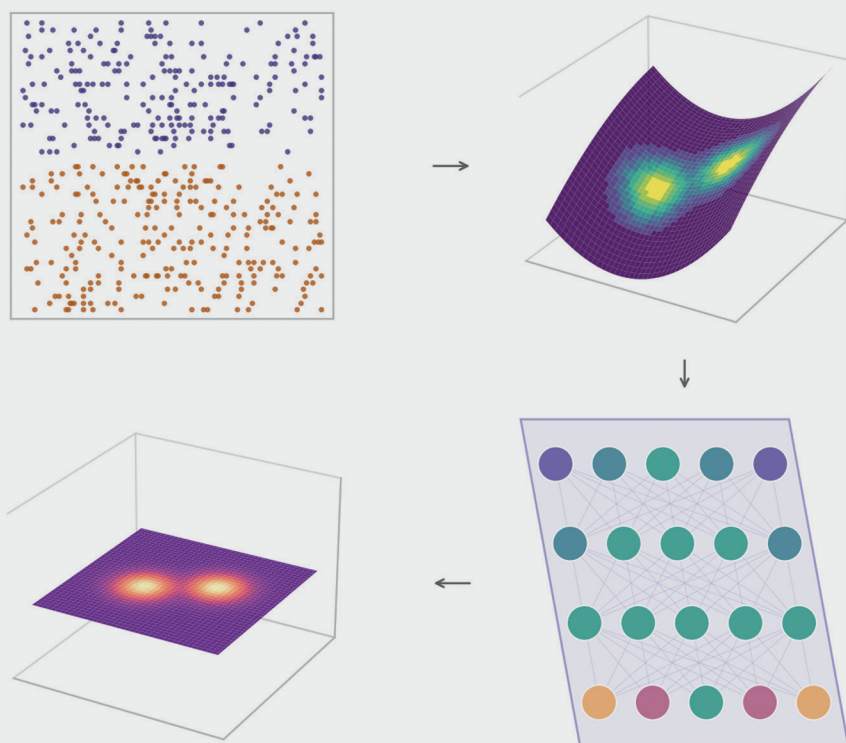




Analysis of Electrophysiological Data with Invertible Neural Networks

Peter Bouss

Information

Band / Volume 130

ISBN 978-3-95806-945-9

Forschungszentrum Jülich GmbH
Institute for Advanced Simulation (IAS)
Computational and Systems Neuroscience (IAS-6)

Analysis of Electrophysiological Data with Invertible Neural Networks

Peter Bouss

Schriften des Forschungszentrums Jülich
Reihe Information / Information

Band / Volume 130

ISSN 1866-1777

ISBN 978-3-95806-945-9

Bibliografische Information der Deutschen Nationalbibliothek.
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der
Deutschen Nationalbibliografie; detaillierte Bibliografische Daten
sind im Internet über <http://dnb.d-nb.de> abrufbar.

Herausgeber
und Vertrieb: Forschungszentrum Jülich GmbH
Zentralbibliothek, Verlag
52425 Jülich
Tel.: +49 2461 61-5368
Fax: +49 2461 61-6103
zb-publikation@fz-juelich.de
www.fz-juelich.de/zb

Umschlaggestaltung: Grafische Medien, Forschungszentrum Jülich GmbH

Druck: Grafische Medien, Forschungszentrum Jülich GmbH

Copyright: Forschungszentrum Jülich 2026

Schriften des Forschungszentrums Jülich
Reihe Information / Information, Band / Volume 130

D 82 (Diss. RWTH Aachen University, 2026)

ISSN 1866-1777
ISBN 978-3-95806-944-2 (Print)
ISBN 978-3-95806-945-9 (E-Book)

Vollständig frei verfügbar über das Publikationsportal des Forschungszentrums Jülich (JuSER)
unter www.fz-juelich.de/zb/openaccess.



This is an Open Access publication distributed under the terms of the [Creative Commons Attribution License 4.0](https://creativecommons.org/licenses/by/4.0/),
which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Declaration of authorship

I, Peter Bouss,
declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

I do solemnly swear that / Hiermit erkläre ich an Eides statt:

1. This work was done wholly or mainly while in candidature for the doctoral degree at this faculty and university;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this university or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others or myself, this is always clearly attributed;
4. Where I have quoted from the work of others or myself, the source is always given. This thesis is entirely my own work, with the exception of such quotations;
5. I have acknowledged all major sources of assistance;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published before as:
Stella*, A., Bouss*, P., Palm, G., and Grün, S. (2022). Comparing surrogates to evaluate precisely timed higher-order spike correlations. *eNeuro*, 9(3):ENEURO.0505–21.2022.
Schutzeichel, L., Bauer, J., Bouss, P., Musall, S., Dahmen, D., and Helias, M. (2025). Transient recurrent dynamics shape representations in mice. *BioRxiv*, 2025.
Merger, C., René, A., Fischer, K., Bouss, P., Nestler, S., Dahmen, D., Honerkamp, C., and Helias, M. (2023). Learning interacting theories from data. *Physical Review X*, 13:041033.
Bouss*, P., Nestler*, S., Fischer, K., Merger, C., René, A., and Helias, M. (2025). Characterizing neural manifolds' properties and curvatures using normalizing flows. *arXiv preprint arXiv:2506.12187*.

München, Mai 2026

Zusammenfassung

Die Frage, wie neuronale Populationen Informationen darstellen, ist eine zentrale Herausforderung in den Neurowissenschaften. Angesichts der parallelen neuronalen Aufzeichnungen vieler Neuronen stellt sich eine Kernfrage: Wie lassen sich aus komplexen Populationsdynamiken aufschlussreiche Strukturen herauslesen? Diese Arbeit entwickelt neue Ansätze zur Analyse neuronaler Daten. Der Fokus richtet sich dabei auf die zeitliche Präzision von Spike-Timing-Korrelationen und Korrelationen höherer Ordnung in kontinuierlichen Populationsaktivitäten.

Gruppen von Neuronen zeigen hochstrukturierte, zeitlich präzise Korrelationen, obwohl ihre Reaktionen individuell variabel erscheinen. Die Kernfrage lautet: Sind die beobachteten Korrelationen lediglich Nebenprodukte von Feuerratenmodulationen oder spiegeln sie eine wirkliche zeitliche Koordination mit zusätzlicher Rechenkraft wider?

Die neuronale Populationsaktivität findet typischerweise auf niedrigdimensionalen Mannigfaltigkeiten innerhalb hochdimensionaler Zustandsräume statt. Herkömmliche Methoden wie die Hauptkomponentenanalyse sind durch Linearitäts- und Gaußsche Annahmen eingeschränkt und können gekrümmte Mannigfaltigkeiten und nicht-lineare Abhängigkeiten, die für das Verständnis neuronaler Berechnungen entscheidend sind, nicht erkennen.

Diese Arbeit befasst sich mit beiden Herausforderungen in vier miteinander verbundenen Studien, die diskrete und kontinuierliche neuronale Signale umfassen. Das erste Kapitel befasst sich mit der Analyse zeitlich präziser Korrelationen durch eine Bewertung verschiedener Ansätze zur Erzeugung von Surrogaten und gibt Empfehlungen für die Auswahl von Methoden für bestimmte experimentelle Ziele. Das zweite Kapitel untersucht, wie rekurrente Schaltkreisdynamiken neuronale Repräsentationen transformieren. Dazu führt es eine Mean-Field-Theorie ein, die komplexes Netzwerkverhalten auf drei kritische Variablen reduziert: populationsgemittelte Aktivität und Korrelationsmaße, die die Reaktionsvariabilität innerhalb und zwischen Reizkategorien quantifizieren. Das dritte Kapitel entwickelt Methoden zur Extraktion verborgener Abhängigkeiten mithilfe invertierbarer neuronaler Netzwerke und demonstriert Techniken zur Identifizierung mikroskopischer Interaktionen direkt aus beobachteten Daten ohne strenge Vorannahmen. Das vierte Kapitel stellt ein umfassendes System zur neuronalen Mannigfaltigkeitsanalyse vor, das invertierbare Netzwerke nutzt. Wir setzen neue Erweiterungen ein, darunter eine spezielle Verlustfunktion, einen multimodalen latenten Raum und eine analytische Darstellung, mit der Statistiken höherer Ordnung und Krümmungsmaße direkt berechnet werden können. Die Analyse von Aufzeichnungen aus dem visuellen Kortex von Makaken zeigt, dass unterschiedliche Verhaltenskontexte Mannigfaltigkeiten mit charakteristischen geometrischen Eigenschaften erzeugen.

Diese Fortschritte bieten neue Möglichkeiten zur Erforschung der Informationsverarbeitung in neuronalen Systemen und verbessern unser Verständnis der Rechenprinzipien, die der Gehirnfunktion zugrunde liegen.

Summary

Understanding how neural populations represent information remains a central challenge in neuroscience. With highly parallel neural recordings from many neurons, a key question arises: how can we extract meaningful structure from complex population dynamics? This thesis develops new approaches for analyzing neural data, focusing on temporal precision of spike-timing correlations and higher-order correlation structure in continuous population activity.

Groups of neurons exhibit highly structured, precisely timed correlations despite individual responses appearing variable. The fundamental question becomes: are observed correlations merely byproducts of firing rate modulations, or do they reflect genuine temporal coordination that provides additional computational power?

Population-level neural activity typically resides on low-dimensional manifolds within high-dimensional state spaces. Traditional methods like Principal Component Analysis are constrained by linearity and Gaussianity assumptions, unable to reveal curved manifolds and nonlinear dependencies crucial for understanding neural computation.

This thesis tackles both challenges through four interconnected studies spanning discrete to continuous neural signals. The first chapter advances analyzing precisely timed correlations through exhaustive assessment of surrogate generation approaches, offering recommendations for method selection tailored to specific experimental objectives. The second chapter examines how recurrent circuit dynamics transform neural representations, introducing a mean-field theory that distills complex network behavior into three critical variables: population-averaged activity and correlation measures quantifying response variability within and between stimulus categories. The third chapter develops methods for extracting hidden interdependencies using invertible neural networks, demonstrating techniques for identifying microscopic interactions directly from observed data without strong prior assumptions. The fourth chapter presents a comprehensive framework for neural manifold analysis using invertible networks. We implement novel extensions including a specialized loss function, a multimodal latent space, and an analytical representations enabling direct computation of higher-order statistics and curvature measures. Analysis of macaque visual cortex recordings reveals that distinct behavioral contexts produce manifolds with characteristic geometric properties.

These innovations provide new capabilities for deciphering how neural systems encode and process information, advancing our understanding of computational principles governing brain function.

Acknowledgments

This thesis would not have been possible without the support, guidance, and friendship of many remarkable people who have enriched both my scientific journey and personal growth over these years.

My deepest gratitude goes to my supervisors, Sonja and Moritz, whose complementary guidance shaped not only this work but also my development as a scientist. Sonja, thank you for your thoughtful advice, your unwavering support during challenging times, and for transmitting your genuine passion for science. Your dedication to fostering a vibrant lab culture and your care for the well-being of everyone in the lab created an environment where both science and personal growth could flourish. Moritz, I am grateful for your creative insights and ideas, our enjoyable meetings filled with shared humor, and your consistently positive and optimistic outlook. Your ability to approach complex problems with enthusiasm and to maintain such a supportive atmosphere made even the most challenging aspects of this research feel manageable and rewarding.

I extend my sincere appreciation to the members of my thesis advisory committee, Björn and Michael, whose valuable feedback and constructive discussions helped refine the direction and quality of this research.

To my fellow PhD and master students, both those who completed their journeys before me and those who continued alongside me, who transformed this experience into something truly memorable: Alessandra, thank you for your warm supervision, genuine friendship, and for fighting alongside me through the complexities of the surrogates paper. Sandra, I deeply appreciate your patience as we worked through the intricacies of the INN project, your friendship, and your commitment to staying closely connected. Anno, your relentless pursuit of truth, warm friendship, and our countless discussions enriched my understanding in ways I never anticipated. Kirsten, your unwavering support, our joint adventures, and the wealth of advice you generously shared have been invaluable. Jonas, thank you for your companionship and for always lending an open ear when needed. Claudia, your infectious passion for science and positive spirit brightened countless days. Lorenzo, our intense friendship and the amazing food you shared created some of my most cherished memories. Lars, your consistency in our joint projects and simple acts of kindness never went unnoticed. Simon, thank you for the warmhearted welcome you provided when I first arrived at the institute. Monica, our regular discussions about pattern quality always sparked new insights. Aitor, your creative ideas for the manifold project opened up new avenues of exploration. Lies, your curious, open, and joyful attitude was a constant source of inspiration. Julio, thank you for your enlightening insights into the world of coffee that made our breaks so much richer. Melissa, your artistic spirit brought a unique perspective to our scientific environment. Jakob, your ability to spread joy and make everyday life feel like a celebration was truly remarkable.

Javed, you have been a wellspring of creative ideas and scientific passion that continually motivated me. Christian, sharing insights during our time in the same office deepened my understanding in numerous ways. Matthias, thank you for the many meaningful moments we shared in our workspace. Gorka, your ability to effortlessly transition into deep, meaningful conversations enriched my perspective on science and life.

To the postdocs who provided valuable mentorship: Alexandre, thank you for the normalizing flows journal club and your help in navigating the literature. David, your positive spirit and willingness to share ideas created a collaborative atmosphere that benefited us all.

My appreciation extends to our dedicated secretaries: Saskia and Janine, thank you for your help with all bureaucratic matters, the happiness and friendship you shared, and your excellent organization of our retreats. Petra, your assistance with trip arrangements was always handled with such care and efficiency.

I am grateful to Günther for his genuine interest in all our projects and his thoughtful engagement with our work. Special thanks to Markus, Abigail, and Sacha for sharing countless insights and ideas that shaped my thinking over the years.

I extend my sincere gratitude to the research teams who generously provided the datasets that were essential to this work: Thomas Brochier, Alexa Riehle and colleagues for the macaque motor cortex data; Pieter Roelfsema, Xing Chen, Aitor Morales-Gregorio and colleagues for the macaque visual cortex data; and Simon Musall and colleagues for the superior colliculus data from mice. Their willingness to share these valuable datasets made much of this research possible.

To our scientific coordinators, Martina and Anne, thank you for your support and guidance in navigating various situations throughout this journey.

I am deeply grateful to the Research Training Group (RTG) and particularly to Marc, Jenny, and Oke for organizing the wonderful retreats over these years. To all RTG members, thank you for the engaging retreats and the wealth of interdisciplinary insights that broadened my scientific horizons in unexpected ways.

Thanks to our IT administrators, Sebastian and Justin, for your patient assistance with all the technical challenges that inevitably arose with the cluster systems.

Finally, my heartfelt appreciation goes to the entire institute for fostering such a warm, collaborative lab culture. The friendships formed and nurtured over these past five years have made this not just a place of work, but a true scientific home where curiosity thrives and human connections flourish. The variety of skills I have learned and experiences I have gained throughout this journey could not have been imagined beforehand and represent a tremendous enrichment to my life, particularly in developing a thorough and critical approach to thinking across diverse topics.

Special thanks to my dedicated proofreaders – Anno, Julio, Jonas, Lars, Claudia, Sandra, Kirsten, and Alessandra – whose careful reading and thoughtful suggestions greatly improved the clarity and quality of this thesis.

Funding

This work was supported by funding that included both direct support for the doctoral candidate and broader project funding. Direct doctoral support was provided by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under grant 368482240/GRK2416. Additional project funding supporting this research was provided by the DFG under grant 491111487; the German Federal Ministry for Education and Research (BMBF) under grants 01IS19077A and 01IS19077B; the European Union's Horizon 2020 Framework Programme for Research and Innovation under Specific Grant Agreements No. 785907 (Human Brain Project SGA2) and No. 945539 (Human Brain Project SGA3); the Helmholtz Association Initiative and Networking Fund under Project Number ZT-I-0003; and the Joint-Lab "Supercomputing and Modeling for the Human Brain".

x

Publications and Scientific Contributions

This chapter summarizes the research contributions made during the course of this doctoral thesis, including peer-reviewed publications, submitted manuscripts, and presentations at international conferences. The work presented here spans two main research areas: the development of statistical methods for higher-order spike correlation analysis and the characterization of neural manifolds using invertible neural networks.

Peer-Reviewed Publications

- Stella*, A., Bouss*, P., Palm, G., & Grün, S. (2022). Comparing surrogates to evaluate precisely timed higher-order spike correlations. *eNeuro*, 9(3). *Equal contribution [**First Main Contribution of this Thesis**]
- Merger, C., René, A., Fischer, K., Bouss, P., Nestler, S., Dahmen, D., Honerkamp, C. & Helias, M. (2023). Learning Interacting Theories from Data. *Physical Review X*, 13(4), 041033.
- Bouss*, P., Nestler*, S., Fischer, K., Merger, C., René, A., & Helias, M. (2025). Characterizing Neural Manifolds' Properties and Curvatures using Normalizing Flows. *Physical Review X Life*, 4(2), 023009. *Equal contribution (under review at Physical Review X Life) [**Second Main Contribution of this Thesis**]

Manuscripts Under Review

- Schutzeichel, L., Bauer, J., Bouss, P., Musall, S., Dahmen, D., & Helias, M. (2025). Transient recurrent dynamics shape representations in mice. *bioRxiv*, 2025.

Conference Presentations and Talks

2025

- Bouss, P., Nestler, S., Fischer, K., Merger, C., René, A., & Helias, M. (2025). Employing normalizing flows to examine neural manifold characteristics and curvatures. DPG Spring Meeting Regensburg. [Oral Presentation]

2024

- Bouss, P., Nestler, S., Fischer, K., Merger C., René, A., & Helias, M. (2024). Assessing Neural Manifold Properties With Adapted Normalizing Flows. *Bernstein Conference 2024*. [Poster]
- Bouss, P., Nestler, S., Fischer, K., Merger C., René, A., & Helias, M. (2024). Exploring Neural Manifold Characteristics Using Modified Normalizing Flows. *ROccella Conference on INFerence and AI, ROCKIN' AI 2024*. [Poster]
- Bouss, P., Nestler, S., Fischer, K., Merger C., René, A., & Helias, M. (2024). Exploring Neural Manifold Characteristics Using Adapted Normalizing Flows. *International Conference on Neuromorphic Computing and Engineering, ICNCE 2024*. [Poster]

2023

- Bouss, P., Nestler, S., Fischer, K., Merger C., René, A., & Helias, M. (2023). Normalizing flows for nonlinear dimensionality reduction of electrophysiological recordings. *Computational Neuroscience Academy 2023*. [Poster]
- Bouss, P., Nestler, S., Fischer, K., Merger C., René, A., & Helias, M. (2023). Non-linear dimensionality reduction with normalizing flows for analysis of electrophysiological recordings. *32nd Annual Computational Neuroscience Meeting*. [Poster]
- Bouss, P., Nestler, S., Fischer, K., Merger C., René, A., & Helias, M. (2023). Dimensionality reduction with invertible neural networks. *PhD and PostDoc Seminar, Bernstein Center Freiburg*. [Oral Presentation]

2022

- Bouss, P., Nestler, S., René, A., & Helias, M. (2022). Dimensionality reduction with normalizing flows. *Bernstein Conference 2022*. [Poster]
- Bouss, P., Stella, A., Palm, G., & Grün, S. (2022). Surrogate techniques to evaluate significance of spike patterns. *6th HBP Student Conference*. [Poster]

2021

- Bouss, P., Stella, A., Palm, G., & Grün, S. (2021). Surrogate methods for spike pattern detection in non-Poisson data. *30th Annual Computational Neuroscience Meeting*. [Poster]
- Bouss, P., Stella, A., Palm, G., & Grün, S. (2021). Surrogate methods for robust significance evaluation of spike patterns in non-Poisson data. *Neuroscience 2021 - 50th Annual Meeting*. [Poster]

Contents

Declaration of authorship	i
Zusammenfassung	iii
Summary	v
Acknowledgments	vii
Funding	ix
Publications and Scientific Contributions	xi
0 Introduction	1
0.1 Neural Computation: From Spike Timing to Manifold Geometry	1
0.1.0 Overview of Contributions and Logical Progression	1
0.1.1 Chapter 1: Surrogate Methods for Higher-Order Spike Correla- tions	2
0.1.2 Chapter 2: Dynamic Recurrent Circuits and the Temporal Evo- lution of Neural Representations	3
0.1.3 Chapter 3: Data-Driven Discovery of Statistical Interaction Prin- ciples	4
0.1.4 Chapter 4: Characterizing Neural Manifolds Using Invertible Neural Networks	6
0.2 Connecting the Four Contributions: A Methodological Framework . .	7
0.3 Conclusion and Outlook	8
0.4 Extended Introductions to the Primary Methodological Contributions	8
0.4.1 Surrogate Methods for Higher-Order Spike Correlations	9
0.4.2 Characterizing Neural Manifolds Using Invertible Neural Net- works	13

1	Surrogate Methods for Higher-Order Spike Correlations	21
1.1	Background	22
1.1.1	Neural Coding Fundamentals	22
1.1.2	Surrogate Methods: Theoretical Framework	24
1.1.3	Methodological Challenges in Spike Train Analysis	25
1.1.4	Statistical Analysis Framework	27
1.1.5	Biological Constraints and Surrogate Design	28
1.1.6	Surrogate Method Classification	29
1.1.7	Experimental Context and Advanced Pattern Detection	30
1.1.8	Chapter Summary	31
1.2	Materials and Methods	31
1.2.1	Surrogate Methods	32
1.2.2	SPADE Algorithm Application	36
1.2.3	Experimental Data and Preprocessing	40
1.2.4	Artificial Non-Stationary Data Generation	41
1.2.5	Code Accessibility and Computational Resources	42
1.3	Results	42
1.3.1	Statistical Evaluation of Surrogate Techniques	42
1.3.2	Binarization Effects on Spike Counts	51
1.3.3	Surrogate Technique Performance on Non-Stationary Data	58
1.3.4	Experimental Data Applications	68
1.4	Discussion and Implications	76
1.4.1	Systematic Biases in Uniform Dithering	76
1.4.2	Performance of Conservative Surrogate Methods	77
1.4.3	Implications for Neural Coordination	78
1.4.4	Methodological Implications	79
1.4.5	Future Directions	79

2	Dynamic Recurrent Circuits and the Temporal Evolution of Neural Representations	81
2.1	The Challenge of Similarity Processing in Recurrent Circuits	82
2.2	A Mean-Field Theory of Neural Representation Dynamics	82
2.2.1	Reducing High-Dimensional Dynamics to Essential Quantities	82
2.2.2	Experimental Validation in Mouse Superior Colliculus	85
2.2.3	Linear Readout and Separability Assessment	86
2.3	Limitations of Traditional Correlation Analysis	86
2.3.1	Extending Beyond Pairwise Correlations	86
2.4	Connecting to Interaction Learning and Manifold Geometry	87
2.4.1	Bridging Discrete and Continuous Perspectives	87
2.4.2	From Rate-Based to Geometric Coding	88
2.5	Implications for Neural Computation	88
2.5.1	Temporal Dynamics and Spatial Organization	88
2.5.2	A Framework for Multi-Scale Analysis	88
2.6	Looking Forward: From Network Dynamics to Statistical Learning . .	89
2.7	Conclusion	89
3	Data-Driven Discovery of Statistical Interaction Principles	91
3.1	From Mean-Field Foundations to Data-Driven Statistical Learning . . .	92
3.2	The Challenge of Inverse Statistical Modeling	92
3.3	Invertible Neural Networks for Statistical Learning	92
3.3.1	A Novel Approach Using Invertible Neural Networks	92
3.3.2	Concrete Example: How Higher-Order Interactions Emerge . .	94
3.3.3	Hierarchical Construction of Interactions	95
3.3.4	Advantages Over Traditional Methods	96
3.3.5	Bridging to Geometric Analysis	96
3.4	Experimental Validation and Performance	97
3.4.1	Testing on Simulated and Real Data	97
3.4.2	Insights into Training Dynamics	97
3.5	Extensions and Future Directions	97
3.5.1	Multimodal Extensions	97
3.5.2	Higher-Order Nonlinearities	98
3.6	Implications for Understanding Neural Computation	98
3.6.1	Statistical Learning as a Unifying Framework	98
3.6.2	Interpretability in Deep Learning for Neuroscience	98
3.7	Conclusion	99

4	Characterizing Neural Manifolds Using Invertible Neural Networks	101
4.1	Motivation	102
4.2	Mathematical Framework	103
4.2.1	Neural Population Coding and Manifold Structure	103
4.2.2	Statistical and Geometric Perspectives	104
4.2.3	Limitations of Current Approaches	104
4.2.4	Normalizing Flows: A Unified Approach	105
4.3	Neural Data and Initial Analysis	106
4.3.1	Dataset and Preprocessing	106
4.3.2	Evidence for Non-Gaussian Structure	107
4.3.3	Biological Context and Multimodal Structure	110
4.4	Normalizing Flow Architecture and Training	111
4.4.1	Volume-Preserving Architecture Design	111
4.4.2	Training Procedure	111
4.4.3	Architecture Choices for Biological Data	112
4.5	Interpretable Latent Representations	112
4.5.1	The Interpretability Challenge in Neural Data Analysis	112
4.5.2	Reconstruction Error Loss for Hierarchical Organization	113
4.5.3	Validation on a Synthetic Dataset	114
4.6	Behavioral State Discovery	116
4.6.1	Multimodal Structure and Behavioral Alignment Strategy	116
4.6.2	Model Selection and Component Determination	117
4.6.3	Discovery of Behavioral State Correspondence	118
4.6.4	Validation of Behavioral Alignment	119
4.6.5	Biological Interpretation and Implications	119
4.7	Statistical Characterization	121

4.7.1	Quadratic Network Approximation for Analytical Tractability .	121
4.7.2	Validation and Error Analysis of Quadratic Approximation . .	122
4.7.3	Higher-Order Statistical Analysis Through Cumulants	123
4.7.4	Analytical Computation via Moment-Generating Functions . .	125
4.7.5	Biological Interpretation of Statistical Complexity	126
4.8	Geometric Analysis	126
4.8.1	Geometric Perspective on Neural Computation	126
4.8.2	From Normalizing Flows to Manifold Curvature	127
4.8.3	Local Geometry Construction and Curvature Measures	128
4.8.4	Geometric Signatures Across Brain States	129
4.8.5	Biological Interpretation of Manifold Curvature	130
4.8.6	Implications for Neural Coding and Computation	131
4.9	Discussion and Implications	132
4.9.1	Behavioral State Discovery Through Geometric Structure	132
4.9.2	Methodological Innovations and Technical Framework	133
4.9.3	Computational Implications of Geometric Structure	134
4.9.4	Multi-Scale Integration and Cross-Chapter Connections	134
4.9.5	Challenging Traditional Approaches to Neural Coding	135
4.9.6	Practical Applications and Future Research Directions	136
4.9.7	A Unified Framework for Neural Computation	137

5	Discussion	139
5.1	Connecting Temporal Precision and Geometric Structure in Neural Computation	139
5.2	Methodological Insights Across Neural Organization Scales	139
5.3	Methodological Advances in Temporal Precision Detection	140
5.4	Network Dynamics and Stimulus Representation	140
5.5	Artificial Data Generation as a Unifying Methodological Theme	141
5.6	Statistical Structure Beyond Pairwise Correlations	142
5.7	Geometric Structure and Statistical Properties in Neural Population Activity	142
5.8	Potential Connections Across Scales	143
5.9	Methodological Contributions and Their Scope	143
5.10	Contributions to Neural Coding Theories	144
5.11	Technical Advances and Scalability	144
5.12	Future Directions Enabled by Our Framework	144
5.13	Limitations and Open Questions	145
5.14	A Methodological Framework for Multi-Scale Analysis	146

	Appendix	147
A	Mathematical Details of Quadratic Representation	149
B	Cumulant Computation via Moment-Generating Function	151
B.1	Analytical Computation of Cumulants	151
B.2	Mathematical Setup and Notation	151
B.3	Moment-Generating Function Evaluation	152
B.4	Cumulant-Generating Function and Series Expansion	153
B.5	Systematic Cumulant Extraction	153

C	Differential Geometry and Curvature Calculation	155
C.1	Geometric Foundation and Assumptions	155
C.2	Manifold Definition and Coordinate System	155
C.3	Mathematical Conventions and Expansion Strategy	156
C.4	Construction of Fundamental Geometric Objects	156
C.4.1	Tangent Vectors and Local Coordinates	156
C.4.2	Metric Tensor Construction	157
C.4.3	Normal Space Construction and Second Fundamental Form . .	157
C.5	Christoffel Symbols and Connection	158
C.6	Curvature Computation	158
C.6.1	Riemannian Curvature Tensor	158
C.6.2	Sectional Curvature	158
C.6.3	Scalar Curvature	159
	References	161
	List of Figures	186
	List of Tables	187

Introduction

0.1 Neural Computation: From Spike Timing to Manifold Geometry

Understanding how neural populations represent and process information remains one of the fundamental challenges in neuroscience (Churchland et al., 2012; Shenoy et al., 2013; Vyas et al., 2020). The neural code – the relationship between patterns of neural activity and the information they encode – operates across multiple spatial and temporal scales, from millisecond-precise spike timing to the geometric structure of population-level activity patterns (Rieke et al., 1997; Pouget et al., 2000; Buzsáki, 2010; Gallego et al., 2017). This geometric perspective views neural activity as trajectories in high-dimensional space that typically lie on lower-dimensional curved surfaces called manifolds (Churchland et al., 2012; Gallego et al., 2017). With the advent of high-density recording technologies that can simultaneously measure electrical activity from hundreds or thousands of neurons, we are now confronted with an unprecedented opportunity to decode neural population dynamics (Cunningham and Yu, 2014).

This thesis develops and applies new methodological approaches to address this challenge, focusing on two complementary aspects of neural computation: the temporal precision of spike-timing correlations in discrete neural signals, and the geometric structure of neural manifolds in continuous population activity. Through four interconnected studies, we progress systematically from discrete to continuous representations, from simple correlations to manifold structures, building a comprehensive framework for analyzing electrophysiological data using invertible neural networks and statistical methods.

0.1.0 Overview of Contributions and Logical Progression

Before introducing each chapter in detail, I first outline how these contributions connect to form a unified analytical framework. Each chapter addresses specific methodological limitations while progressively building toward comprehensive tools to study neural information processing:

- **Chapter 1** establishes methods for evaluating precisely timed correlations in discrete spiking signals (shared first author contribution)
- **Chapter 2** contextualizes correlations within network dynamics, showing how they emerge from recurrent circuit interactions (collaborative contribution)
- **Chapter 3** moves beyond pairwise correlations to learn interaction structures from data using invertible neural networks (collaborative contribution)
- **Chapter 4** extends this framework to continuous signals, revealing how dependencies manifest as statistical and geometric properties of neural manifolds (shared first author contribution)

This progression addresses a fundamental challenge: whereas traditional neuroscience focuses on either temporal precision or population geometry, real neural computation involves both. Our framework develops complementary approaches for analyzing neural data across both temporal and geometric aspects, providing quantitative tools that can be applied across different data types and recording modalities.

0.1.1 Chapter 1: Surrogate Methods for Higher-Order Spike Correlations

0.1.1.1 The Challenge of Temporal Precision in Neural Correlations

Neural computation critically depends on the precise timing of spike correlations (Grün, 2009). Even when individual neurons appear to respond stochastically to repeated stimuli, groups of neurons can exhibit highly structured, precisely timed correlations that operate at millisecond timescales (Stark and Abeles, 2007). These correlations create detectable patterns in collective spiking activity (Grün et al., 2002a). The central question becomes *whether observed correlations arise solely from firing rate modulations, or whether they reflect additional temporal coordination mechanisms that operate at millisecond timescales and contribute specific computational advantages beyond rate-based coding*.

Detecting such patterns in discrete spiking signals requires statistical methods that can distinguish genuine spatio-temporal patterns from chance occurrences (Louis et al., 2010; Quaglio et al., 2018). Traditional approaches often cannot account for the statistical properties of spike trains, including non-stationary firing rates, regularities, inter-spike interval distributions, and autocorrelations (Pazienti et al., 2008; Elsayed and Cunningham, 2017). This situation creates a fundamental methodological challenge: how can we develop statistical frameworks that separate the effects of temporal precision from rate-based coding?

0.1.1.2 **Contribution: Evaluating Surrogate Methods for Spike-Timing Analysis**

As the first major contribution of this thesis (shared first author work), we address this challenge through a comprehensive evaluation of surrogate data methods used in the statistical assessment of temporal precision in neural correlations. We systematically compare different surrogate generation techniques – including uniform dithering, trial shifting, and various other dithering methods – using artificially generated spike trains with known ground truth correlations.

Our analysis reveals that common surrogate methods can significantly distort essential statistical features of the data, leading to over- or under-estimation of neural correlations. We develop analytical tools for assessing how surrogate methods alter spike statistics and provide practical guidelines for choosing appropriate surrogate strategies based on experimental goals and data characteristics. This work establishes a statistical foundation for detecting precisely timed correlations in discrete neural signals, providing analytical tools to assess the statistical significance of temporal correlations while accounting for the statistical properties of spike trains.

0.1.2 **Chapter 2: Dynamic Recurrent Circuits and the Temporal Evolution of Neural Representations**

0.1.2.1 **Understanding How Recurrent Dynamics Shape Neural Representations**

Building on the correlation detection methods established in Chapter 1, this chapter addresses a complementary aspect: how do population-level correlations emerge and evolve through network dynamics? We examine a fundamental computational challenge that every brain must solve: distinguishing different stimuli (like a blue ball from a red ball) while preserving similarities within stimulus categories (recognizing that two red balls are similar despite different contexts) (DiCarlo and Cox, 2007). This similarity processing challenge becomes critical in higher-order areas tightly linked to cognition, but the computational principles emerge earlier in processing areas like the superior colliculus (SC) (Meredith and Stein, 1986).

We shift from analyzing precisely-timed correlations to understanding network-level computation, examining how recurrent circuit dynamics reshape stimulus representations over time (Churchland et al., 2010; Mastrogiuseppe and Ostojic, 2018; Kietzmann et al., 2019). A fundamental property of neural circuits is that identical stimuli never produce identical responses (Shadlen and Newsome, 1998; Faisal et al., 2008). Neural responses exhibit trial-to-trial variability that reflects the ongoing dynamics of recurrent networks (Arieli et al., 1996). This variability follows systematic patterns that change predictably when stimuli are presented (Churchland et al., 2010). Response variability remains elevated during baseline periods but diminishes when

stimuli are presented, suggesting that stimuli interact with ongoing network dynamics to shape representations (Churchland et al., 2010; Litwin-Kumar and Doiron, 2012).

The central question connecting this work to our broader framework is: *How do local recurrent circuits dynamically reshape neural representations over time, and what determines the trade-off between preserving stimulus information and achieving stimulus separability?*

0.1.2.2 Contribution: Dynamic Reshaping of Neural Representations

In this collaborative work (co-author contribution), we investigate how neural representations are dynamically reshaped by recurrent circuits using both experimental data from awake behaving mice and a recurrent neural network model. Our focus is the mouse superior colliculus (SC), a structure important for processing visual information and highly multimodal in nature, making it ideal for studying how networks handle diverse stimulus types (Meredith and Stein, 1986; Ito and Feldheim, 2018; Wang et al., 2020).

The key theoretical contribution is a mean-field theory that reduces high-dimensional network dynamics to three relevant quantities: mean population activity and two covariance measures reflecting response variability within and across stimulus classes (Mazzucato et al., 2016; Benavidez et al., 2021). Our framework reveals how these three quantities shape stimulus representation separability through their dynamic interplay, demonstrating a fundamental trade-off between preserving stimulus-specific information and achieving separability between different stimulus classes.

This correlation-based analysis – analogous to examining ellipsoids defined by covariance matrices – provides only a simplified view of more high-dimensional statistical structures. The limitation of pairwise correlations becomes apparent when considering the dependencies that characterize real neural data. Chapter 3 develops methods for learning statistical structures that extend beyond simple pairwise relationships.

0.1.3 Chapter 3: Data-Driven Discovery of Statistical Interaction Principles

0.1.3.1 Beyond Simple Correlations: Learning Complex Statistical Structures

The limitations of considering only pairwise correlations lead to a fundamental methodological question: *How can we characterize the statistical structures underlying complex data when moving beyond simple pairwise correlations?* This challenge proves

crucial for understanding neural computation because real neural interactions involve higher-order dependencies that cannot be captured by traditional pairwise correlation analysis (Schneidman et al., 2006; Ganmor et al., 2011).

These challenges are not unique to neuroscience. Across scientific domains, understanding systems often requires inferring microscopic interactions from macroscopic measurements. In physics, scientists seek to understand how collective behaviors emerge from underlying interaction rules (Wilson, 1971; Anderson, 1972). When predictive theoretical frameworks are unavailable, inverse inference becomes necessary – reconstructing microscopic models from observations of collective states (Kaipio and Somersalo, 2005; Rechtsman et al., 2006). Solving these inverse problems typically requires computational approaches (Tarantola, 2005).

Recent advances in machine learning offer promising solutions through generative neural networks (Kingma and Welling, 2013; Dinh et al., 2015). These models learn the structures underlying training data and must therefore possess internal representations encoded in their parameters (Bengio et al., 2013). The key insight is that we can extract and interpret these learned models to understand underlying patterns that govern neural population dynamics (Sussillo and Barak, 2013).

To address our methodological question systematically, we examine a controlled setting where we can validate our approach: artificial neural networks where we control the underlying generative process and know the ground truth. This validation strategy allows testing of data-driven methods before applying them to experimental data where true structure remains unknown.

0.1.3.2 **Contribution: Interaction Learning**

In this collaborative work (co-author contribution), we develop a framework for learning hierarchical interaction structures from neural network data using invertible neural networks (INNs), also called normalizing flows – generative models that learn invertible mappings between data distributions and simple latent representations (Dinh et al., 2015, 2016; Kingma and Dhariwal, 2018; Ardizzone et al., 2019). Unlike traditional approaches that compress information, these networks preserve all information through their invertible architecture, making them well-suited for extracting the complete statistical structure (Papamakarios et al., 2021).

Our approach demonstrates how to extract an action (Parisi, 1988) – the negative log-likelihood of the probability density – from observed population activity patterns, revealing underlying dependencies governing the system. The method uncovers interaction patterns at multiple scales through the layered architecture of the invertible network, bridging insights from statistical physics about how neural networks learn to represent data statistics. Our method moves beyond the constraints of pairwise

correlation analysis by learning the complete structure of neural population activity, including higher-order interactions that emerge from the full joint probability distribution.

This work establishes the foundation for applying invertible neural networks to neural data analysis. While Chapter 3 demonstrates how to learn complex structures from artificial data, the next step involves applying similar approaches to neuronal data. Many modern high-density neural recordings capture continuous population-level signals, creating an opportunity to examine neural computation from a statistical and geometric perspective. Chapter 4 develops methods to quantify how dependencies in neural activity manifest as geometric properties of neural manifolds.

0.1.4 Chapter 4: Characterizing Neural Manifolds Using Invertible Neural Networks

0.1.4.1 Geometric Analysis of Continuous Neural Signals

The final chapter represents the culmination of our methodological development, extending the invertible neural network framework from Chapter 3 to continuous population activity. This transition allows us to examine neural computations from a geometric perspective, asking: How do higher-order correlations relate to the geometric structures of the neural manifold, and what do these structures reveal about neural computation?

Neural activity in large populations often lies on low-dimensional manifolds embedded within high-dimensional state spaces (Stopfer et al., 2003; Churchland et al., 2012; Gallego et al., 2017). These manifolds encode relationships between neural activity and latent variables (Seung and Lee, 2000). Traditional linear methods like PCA fail to capture the curved, nonlinear structure of real neural data and the higher-order dependencies that are critical for understanding information representation and transformation (Mitchell-Heggs et al., 2023; Langdon et al., 2023).

This chapter addresses the fundamental question that emerges from our entire methodological progression: *How can we characterize the full geometric and statistical properties of neural manifolds, moving beyond the limitations of both discrete correlation analysis and pairwise methods?*

0.1.4.2 Contribution: Characterizing Neural Manifolds with Invertible Neural Networks

As the second major contribution of this thesis (shared first author work), we develop a framework for characterizing neural manifolds and their geometric properties using normalizing flows (Dinh et al., 2015, 2016; Rezende and Mohamed, 2015; Papamakarios et al., 2021) – the same class of invertible neural networks introduced in

Chapter 3, but now adapted for continuous neural signals. This framework enables interpretable geometric analysis of neural population activity.

We combine several key extensions to standard normalizing flows: a modified loss function that promotes interpretable low-dimensional structure in the latent space (Bekasov and Murray, 2020), a mixture-of-Gaussian latent space to accommodate neural activity patterns with multiple distinct modes (Izmailov et al., 2020), and a quadratic parameterization of the network architecture that enables analytical computation of higher-order statistics and curvature properties. These geometric and statistical features – including local curvature and third- and fourth-order correlations – serve as characteristic signatures of neural activity states and reflect the underlying dependencies between neurons (Shimazaki et al., 2012; Chung and Abbott, 2021).

Through analysis of recordings from macaque visual cortex (Chen et al., 2022; ?), we demonstrate that different behavioral states correspond to different manifold geometries, ranging from nearly flat to curved structures with saddle-point-like features. These findings suggest that manifold geometry provides a compact representation of neural circuit computations, with different geometric structures corresponding to different behavioral states and computational demands (Chung et al., 2018; Cohen et al., 2020; Chou et al., 2025).

0.2 Connecting the Four Contributions: A Methodological Framework

This thesis presents a progression from discrete temporal analysis to continuous geometric characterization of neural computation, unified by the common theme of moving beyond simple correlations to understand higher-order statistical dependencies in neural data. The four chapters are interconnected through their shared methodological development and complementary perspectives on neural information processing:

Chapter 1 establishes the statistical foundation by providing methods for detecting precisely timed correlations in discrete spiking signals, addressing the fundamental challenge of separating temporal precision from rate-based effects.

Chapter 2 shifts focus to network-level dynamics, examining how recurrent circuits shape neural representations through the interplay of mean population activity and correlations within and between stimulus classes.

Chapter 3 transcends the limitations of both approaches developing data-driven methods for learning interaction structures using invertible neural networks, bridging discrete and continuous perspectives.

Chapter 4 extends this framework to continuous neural signals, demonstrating how higher-order correlations manifest as geometric properties of neural manifolds and reveal computational principles.

Together, these contributions offer complementary perspectives – temporal precision, network dynamics, higher-order interactions, and geometric structure – that provide a methodological framework for analyzing neural circuits across multiple scales and data types. The two primary methodological innovations (Chapters 1 and 4) provide principled, interpretable tools that have been successfully applied to experimental data, while the contextualizing chapters (Chapters 2 and 3) illustrate how these methods connect to network function and statistical learning theory.

0.3 Conclusion and Outlook

This thesis contributes to computational neuroscience by developing analytical approaches that address both millisecond-precise spike timing and the geometric structure of neural manifolds. These methods progress from discrete to continuous signals and from simple correlations to manifold geometry, providing tools for analyzing neural activity across multiple scales.

The methodological innovations presented here – statistical methods for temporal analysis, mean-field theory for dynamic reshaping, data-driven statistical learning for interaction structures, and invertible neural networks for manifold characterization – together enable a deeper understanding of how neural systems encode and transform information. These approaches for analyzing electrophysiological data promise to advance our ability to decode the neural code and understand the computational principles underlying brain function by providing quantitative methods to characterize both temporal and geometric aspects of neural population activity.

0.4 Extended Introductions to the Primary Methodological Contributions

The sections that follow provide in-depth introductions to the two main methodological pillars of this thesis. First, we explore in detail the challenges of detecting precisely timed higher-order spike correlations and the critical importance of statistical methods for separating temporal precision from rate-based effects. This extended introduction sets the stage for the comprehensive evaluation and practical guidelines presented in Chapter 1.

Second, we delve into the geometric analysis of neural manifolds using invertible neural networks, examining how higher-order statistical dependencies manifest as curvature properties and how these geometric features reveal computational principles. This detailed motivation provides the theoretical foundation for the manifold characterization framework developed in Chapter 4.

0.4.1 Surrogate Methods for Higher-Order Spike Correlations

The analysis of millisecond-precise spike correlations presents a fundamental methodological challenge. While neural computation may depend critically on precise timing relationships (Grün et al., 2002a), distinguishing genuine temporal coordination from statistical artifacts remains difficult with current approaches.

The problem of temporal precision in neural correlations represents a fundamental intersection between the discrete nature of spike signals and the continuous computational processes they encode (Rieke et al., 1997). While high-density recording technologies now enable simultaneous monitoring of hundreds or thousands of neurons, such technological advancement has exposed the inadequacy of traditional statistical methods for analyzing complex, high-dimensional spike data (Stevenson and Kording, 2011). The central methodological question we address here is: *how can we develop a rigorous statistical framework that separates the effects of temporal precision from rate-based coding mechanisms?*

The Foundation of Surrogate Data Generation

Surrogate data techniques offer a practical alternative when traditional hypothesis testing fails due to complex spike train statistics. Grün's pioneering work (2002a; 2002b; 2009) established surrogate methods as a viable alternative to parametric approaches, while advances by Louis et al. (2010) have highlighted remaining limitations. Classical statistical approaches depend on establishing predefined null-hypotheses that quantify the probability of observed effects emerging purely by chance, yet such methodology needs access to analytically tractable distributions where significance probabilities can be mathematically computed (Efron and Tibshirani, 1993; Brown et al., 2004).

Spike train statistics often violate the independence assumptions required for analytical approaches (Shadlen and Newsome, 1998; Brown et al., 2004). For instance, regularity and dead times create temporal dependencies that standard Poisson models cannot capture, creating the need for simulation-based statistical methods. In these cases, computational statistics has enabled the estimation of significance probabilities through extensive sampling from null-hypothesis distributions, effectively generating artificial datasets that replicate certain statistical properties of the original observations (Efron and Tibshirani, 1993). These methods transform the challenge from mathematical derivation to computational simulation, thereby expanding possibilities for rigorous statistical analysis of complex neural phenomena (Louis et al., 2010). Modern approaches have extended these principles to include point process models that better capture the statistical structure of spike trains (Grün and Rotter, 2010), as well as variational methods for recovering population dynamics from neural data (Buesing et al., 2011).

Surrogate Methods Within the Bootstrap Framework

Surrogate data generation provides a methodological framework that uses the dataset itself as the foundation for statistical inference (Theiler et al., 1992). Rather than imposing predetermined distributional assumptions like the Poisson distribution, this technique constructs modified versions of the observed data that retain essential characteristics while systematically manipulating specific features of interest (Louis et al., 2010).

Surrogate methods extend bootstrap principles (Efron, 1982; Efron and Tibshirani, 1993) by selectively disrupting specific data features while preserving others. This targeted manipulation allows researchers to test whether particular aspects of the data – such as precise spike timing – contribute to observed patterns. The approach strategically eliminates hypothesized critical features while maintaining other statistical properties (Amarasingham et al., 2012). When the modified data generates the observed effect with minimal probability, we obtain evidence supporting the significance of the eliminated features (Kass et al., 2005; Ventura, 2010).

Application to Large-Scale Neural Recordings

Large-scale neural recordings using multi-electrode arrays have revolutionized neuroscience by enabling simultaneous monitoring of hundreds to thousands of neurons across diverse brain areas and species (Jun et al., 2017; Steinmetz et al., 2021). Despite the apparent stochasticity characterizing neural responses to repeated stimuli (Nawrot et al., 2008), precisely timed neural interactions have emerged as a central research focus.

In this work, we focus on recordings from macaque pre- and motor cortex using Utah arrays that simultaneously monitor around 100 neurons (Riehle et al., 2013; Brochier et al., 2018). Research across multiple experimental contexts has evolved from early investigations of spatiotemporal firing patterns among simultaneously recorded neurons (Abeles and Gerstein, 1988; Gerstein et al., 1989) to comprehensive analyses of both pairwise correlations (Grün et al., 2002a,b; Pipa and Grün, 2003; Pipa et al., 2003, 2007; Grün, 2009; Pipa et al., 2013) and higher-order correlation structures in parallel spike train recordings (Villa and Abeles, 1990; Martignon et al., 1995; Riehle et al., 1997; Prut et al., 1998; Villa et al., 1999; Kilavik et al., 2009; Staude et al., 2010a; Shimazaki et al., 2012). The complexity of these temporal dependencies has motivated the development of more sophisticated statistical frameworks, including cumulant-based inference methods for detecting higher-order correlations (Staude et al., 2010b). Modern approaches include methods such as SPADE (SPAtio-temporal patterns Detection and Evaluation) for identifying complex spike patterns (Torre et al., 2013, 2016; Quaglio et al., 2017; Stella et al., 2019).

Temporal Precision and Dithering Approaches

Establishing high temporal precision in multi-neuron interactions via statistical methods requires sophisticated surrogate techniques capable of selectively disrupting correlations at fine temporal scales while maintaining coarser temporal structure (Louis et al., 2010). These approaches must simultaneously preserve the statistical integrity of individual spike trains to the greatest extent possible (Amarasingham et al., 2012). The core strategy entails introducing minor perturbations to spike timing, thereby generating a family of techniques termed dithering, jittering, or teetering (Date et al., 1998; Hatsopoulos et al., 2003; Pazienti et al., 2008; Stark and Abeles, 2009; Louis et al., 2010).

Limitations of Uniform Dithering

Among these approaches, uniform dithering has achieved widespread adoption owing to its conceptual simplicity – each spike is shifted by a small amount sampled from a uniform distribution (Date et al., 1998; Hatsopoulos et al., 1998). But uniform dithering distorts interspike intervals (Louis et al., 2010). This creates serious problems for pattern detection. As we find out these distortions can produce substantial reductions in spike counts, when combined with binning and binarization procedures that represent prerequisites for numerous statistical analysis methods, which subsequently generate systematic underestimation of pattern significance probabilities. The implications of these methodological choices extend beyond technical considerations.

Our Methodological Contribution and Validation Framework

To overcome the mentioned limitations, we systematically compare alternative surrogate approaches to find better solutions. We provide practical guidelines for selecting appropriate surrogate methods based on experimental objectives and data characteristics, contributing to research on precise spiking motifs and higher-order spatiotemporal correlations in neurobiological and neuromorphic systems (Grimaldi et al., 2023; Deshpande et al., 2023). Chapter 1 presents our comprehensive investigation assessing these methods relative to uniform dithering across multiple dimensions, with particular emphasis on their impact on statistical evaluation of repeating spatio-temporal spike patterns using the SPADE algorithm (Torre et al., 2013; Quaglio et al., 2017; Stella et al., 2019).

To establish rigorous benchmarks for our comparative analysis, we utilize artificially generated datasets of progressively increasing complexity where any observed spatio-temporal pattern can emerge only by chance. Our controlled experimental

framework enables us to quantify the false-positive rate associated with each surrogate method, thereby providing an essential foundation for method validation. Given the multiple testing challenges inherent in pattern detection across large neural populations, we employ appropriate statistical corrections to control the false discovery rate (Benjamini and Hochberg, 1995). The investigation culminates with application to experimental data, where we assess the detection of significant patterns.

Transition to Geometric Analysis of Neural Manifolds

While the surrogate methods framework provides essential tools for analyzing discrete spike timing with millisecond precision, understanding neural computation requires complementary approaches for characterizing the geometric structure of continuous population activity. When we view neural population activity as points in a high-dimensional space defined by the recorded neurons, this activity typically resides on lower-dimensional curved surfaces called manifolds (Gallego et al., 2017). The temporal precision revealed through surrogate analysis represents one dimension of neural organization, yet the activity of neural populations also exhibits complex geometric relationships that manifest as these curved manifolds in high-dimensional activity space (Stringer et al., 2019a).

The transition from discrete temporal analysis to continuous geometric characterization represents a natural progression in our comprehensive framework for understanding neural computation. Where surrogate methods enable us to evaluate precisely timed correlations in spike trains, geometric analysis of neural manifolds allows us to characterize how higher-order statistical dependencies manifest in the curvature properties of the population activity. This complementary perspective, developed through invertible neural networks, forms the foundation for our second major methodological contribution.

0.4.2 Characterizing Neural Manifolds Using Invertible Neural Networks

From Temporal Precision to Geometric Structure

Spike timing analysis reveals one dimension of neural organization, but population activity exhibits additional structure that becomes apparent only when viewed geometrically. Neural populations operate in high-dimensional spaces, yet their activity typically concentrates on curved, low-dimensional surfaces – neural manifolds. Understanding these manifold structures requires moving beyond linear methods like PCA, which assume flat geometries and Gaussian statistics – assumptions that neural data frequently violate (Cunningham and Yu, 2014; Gallego et al., 2017).

Our second major contribution focuses on characterizing neural population dynamics by examining how higher-order statistical dependencies manifest as geometric properties in neural manifolds (Figure 1). This approach builds upon our theoretical frameworks for recurrent dynamics (Chapter 2) and invertible architectures (Chapter 3), extending these methods to characterize the intrinsic statistical and geometric structure of neural population activity.

The Dual Challenge of Neural Information Processing

Neural information processing presents two interconnected challenges (Knill and Pouget, 2004). First, we must characterize relationships among multiple cellular activities (Georgopoulos et al., 1986; Seung and Sompolinsky, 1993). These relationships can be described through statistical frameworks – like correlation patterns across neuronal ensembles – or geometric approaches that examine subspaces occupied by collective population activity, commonly termed neural manifolds (Saxena and Cunningham, 2019; Vyas et al., 2020).

Second, we must understand how neuronal activity relates to meaningful lower-dimensional latent variables – the unobserved factors that drive the patterns we observe in neural responses (Cunningham and Yu, 2014; Jazayeri and Ostojic, 2021). These latent variables represent the underlying computations or states that cannot be directly measured but generate the observed neural activity patterns (Recanatesi et al., 2021). Motor control signals (Churchland et al., 2012; Kaufman et al., 2013; Sussillo et al., 2015; Kao et al., 2021), sensory processing components (Olshausen and Field, 1996; DiCarlo et al., 2012; Mante et al., 2013; Stringer et al., 2019a), cognitive decision variables (Kim et al., 2021; Steinemann et al., 2024; Langdon and Engel, 2025), and other continuous internal brain states (Pandarinath et al., 2018; Whiteway and Butts, 2019) all represent potential underlying latent variables.

Addressing these dual challenges – characterizing population relationships and identifying meaningful latent variables – requires extracting interpretable low-dimensional

structure from high-dimensional electrophysiological recordings (Cunningham and Yu, 2014; Gallego et al., 2017). This extraction process must simultaneously characterize statistical relationships within neural populations (Schneidman et al., 2006; Pillow et al., 2008) and establish how behaviorally relevant latent variables drive the observed activity patterns (Pandarinath et al., 2018).

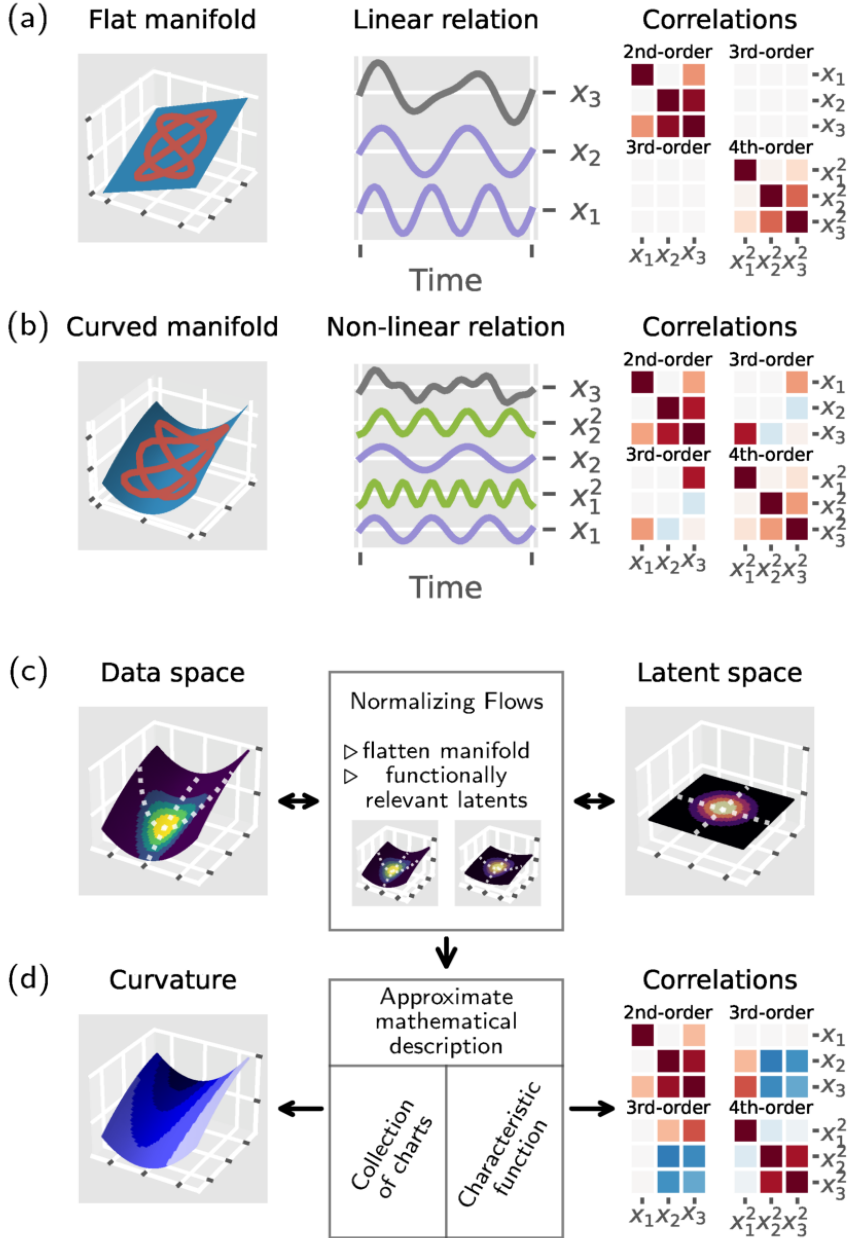


Figure 1: Analysis pipeline for curved neural manifolds using meaningful latent variables.

Figure 1: **Analysis pipeline for curved neural manifolds using meaningful latent variables.** **(a)** A signal trajectory x (red curve) moves through three-dimensional space (left panel) staying on a flat manifold (blue surface), where x_i represents its component in dimension i . The middle panel shows how each signal changes over time, where x_1 and x_2 (purple curves) are independent oscillations, while x_3 (gray curve) is a linear combination of x_1 and x_2 . This creates a linear relationship between the signals. The right panel shows correlations of different orders for combinations of x_i , x_i^2 , x_j , and x_j^2 . Third-order correlations are zero, while fourth-order correlations originate only from the second-order terms. **(b)** Similar setup shows a signal trajectory x (red curve) on a curved manifold (blue surface). The middle panel reveals that x_3 (gray curve) depends on both x_1 and x_2 (purple curves) and their squared terms x_1^2 and x_2^2 (green curves). The manifold curves because of these nonlinear relationships. Unlike panel (a), the right panel shows non-zero third-order correlations that are a hint for the manifold curvature. **(c)** A probability distribution lies on a curved manifold (left panel). The Normalizing Flow (middle panel) learns to map this distribution to a simpler form on a flat manifold (right panel). Each layer in this neural network performs invertible transformations. The flat latent space uses a Gaussian distribution with main axes (white dashed lines) that correspond to meaningful curved directions in the original data space (left panel), identifying the underlying latent variables. **(d)** The trained Normalizing Flow is approximated using a quadratic function, which creates coordinate charts and a characteristic function (middle panel). With these charts the manifold's curvature is calculated (left panel), while the characteristic function produces correlations of different orders (right panel).

Manifold Structure in Population Dynamics

Recent investigations reveal that neural activity trajectories occupy low-dimensional manifolds with intrinsic dimensionality substantially lower than their embedding space - the high-dimensional space defined by all simultaneously recorded neurons (Cunningham and Yu, 2014; Gallego et al., 2017; Vyas et al., 2020; Langdon et al., 2023). This principle appears consistently across brain regions and behavioral contexts (Saxena and Cunningham, 2019; Altan et al., 2021; Mitchell-Heggs et al., 2023). Gallego et al. (2017, 2020) demonstrated stable low-dimensional control manifolds in motor cortex, while Stringer et al. (2019a) revealed similar principles in visual processing and Low et al. (2018) in hippocampal navigation. Prefrontal cortex operates through manifold-based computations during cognitive tasks (Mante et al., 2013; Langdon et al., 2023). Studies spanning multiple brain areas confirm this widespread organizational principle (Safaie et al., 2023; Fortunato et al., 2024).

This manifold organization reflects how neurons function through complex interdependencies rather than independently (Schneidman et al., 2006; Tkačik et al., 2014). Representational drift and intrinsic network structure influence these dependencies,

while correlated variability shapes population coding accuracy (Elsayed and Cunningham, 2017; Rule et al., 2019). Characterizing these geometric properties becomes crucial for understanding how the brain represents and transforms information across different computational demands (Chung and Abbott, 2021; Jazayeri and Ostojic, 2021).

Neural manifolds show both universal principles and region-specific adaptations (Saxena and Cunningham, 2019). Motor areas employ manifolds for learning and adaptive control of precise motor commands (Churchland et al., 2012; Dhawale et al., 2017). Prefrontal regions use manifold geometry for flexible cognitive computations (Mante et al., 2013; Remington et al., 2018). Hippocampal circuits organize spatial and episodic memory through manifold representations (Low et al., 2018). Visual cortex demonstrates multiple aspects of manifold organization: population dynamics during spontaneous activity (Stringer et al., 2019b), hierarchical processing for object recognition (DiCarlo et al., 2012), and distinct manifolds with top-down modulation between areas (?). While the low-dimensional principle appears universal, manifold properties vary across brain regions in ways that may reflect different computational demands (Gao and Ganguli, 2015; Chung and Abbott, 2021).

Limitations of Linear Methods

Principal Component Analysis (PCA) and related linear techniques have dominated neural manifold studies (Pearson, 1901; Hotelling, 1933; Bishop, 1998; Zou et al., 2006; Kobak et al., 2016; Mitchell-Heggs et al., 2023), alongside tensor-based extensions (Williams et al., 2018; Pellegrino et al., 2024). However, PCA's linear transformation constraints create substantial limitations for characterizing curved neural manifolds or complex neural dependencies (Sussillo and Barak, 2013; Cunningham and Yu, 2014; Pandarinath et al., 2018). These constraints become particularly apparent when analyzing brain-wide interactions and population-level dynamics that exhibit rich temporal structure (Perich et al., 2024).

Linear descriptions work adequately only when data reside on flat neural manifolds and follow multivariate Gaussian distributions – for instance, when one variable represents a linear combination of others and pairwise correlations capture all statistical dependencies (Bishop, 2006; Dayan and Abbott, 2001). Figure 1a illustrates this scenario, where signal x_3 is a weighted sum of independent oscillatory signals x_1 and x_2 . Such linear relationships produce vanishing third-order correlations and fourth-order correlations that can be predicted from second-order statistics alone (Schneidman et al., 2006).

The Need for Nonlinear Frameworks

Neural recordings from behaving animals rarely conform to these restrictive assumptions of flat manifolds (Duncker and Sahani, 2021; Abbaspourazad et al., 2024). Mod-

ern electrophysiological recordings frequently exhibit non-Gaussian statistical properties and manifold curvature that necessitate nonlinear analytical tools (Altan et al., 2021; Acosta et al., 2023; Fortunato et al., 2024). Neural dynamics are often dominated by richly varied movements and behavioral states, with single-trial patterns showing complex structure that reflects underlying manifold organization (Stringer et al., 2019a; Musall et al., 2019; Zatka-Haas et al., 2021).

This complexity arises because curved neural manifolds emerge when variables exhibit quadratic or higher-order dependencies, leading to state space curvature. Such nonlinear relationships can generate non-trivial higher-order correlations that linear methods cannot capture. Figure 1b demonstrates how variables depending not only on linear combinations but also on quadratic terms (x_1^2 and x_2^2) produce manifold curvature and yield non-zero third-order correlations – providing statistical signatures that reflect the underlying geometric structure.

Multiple nonlinear methods have emerged to address these curved manifold structures. Classical manifold learning techniques offer one class of solutions (Tenenbaum et al., 2000; Roweis and Saul, 2000; Belkin and Niyogi, 2003; van der Maaten and Hinton, 2008; McInnes et al., 2018). Gaussian process factor analysis has proven particularly influential for single-trial neural population dynamics (Yu et al., 2008). Deep generative models represent another major direction (Kingma and Welling, 2013; Pandarinath et al., 2018; Brehmer and Cranmer, 2020; Kim et al., 2020; Horvat and Pfister, 2022; Flouris and Konukoglu, 2024).

Despite some success in capturing certain curved manifold aspects, these methods often suffer from limited interpretability regarding behavioral description through latent variables and tend to lack explicit characterization of underlying manifold geometry. This gap motivates the need for analytical frameworks that can simultaneously capture geometric structure and statistical dependencies while producing interpretable low-dimensional representations.

Our Methodological Framework

Our framework addresses three key challenges: learning interpretable low-dimensional representations, characterizing manifold geometry, and connecting geometric properties to statistical dependencies. Existing methods tackle these objectives in isolation; we integrate them within a unified approach.

Extending Normalizing Flows for Neural Manifold Analysis

To address the limitations of linear methods and the need for interpretable nonlinear frameworks, we turn to invertible neural networks as our analytical foundation.

Specifically, we employ Normalizing Flows (NFs) (Dinh et al., 2015, 2016; Kingma and Dhariwal, 2018) – neural networks that learn invertible transformations between data and latent spaces. These networks map complex neural activity patterns to simpler latent representations while preserving all information through their invertible architecture.

Figure 1c illustrates this process: the Normalizing Flow transforms both the probability distribution and its underlying curved manifold into a latent distribution on a flat manifold. Each network layer implements an invertible transformation, ensuring that complex curved relationships in neural space become linear relationships in latent space, making them easier to interpret and analyze.

This approach leverages existing work in flow-based manifold learning (Brehmer and Cranmer, 2020; Kim et al., 2020; Horvat and Pfister, 2021, 2022, 2023; Loaiza-Ganem et al., 2022; Flouris and Konukoglu, 2024) but our framework uniquely addresses all three core objectives simultaneously.

Our framework incorporates three extensions to standard NFs. First, we adapt a modified loss function that encourages low-dimensional latent structure (Bekasov and Murray, 2020), with the goal that the most informative latent directions may correspond to meaningful neural population dynamics. Backward mapping from latent to data space can yield curved, low-dimensional neural manifolds that potentially capture meaningful features of the data. The latent space axes (shown as white dashed lines in Figure 1c) aim to correspond to functionally relevant curved axes in the original data distribution.

Second, we replace standard Gaussian latent distributions with mixture-of-Gaussians representations in the latent space, enabling the network to model multimodal neural data essential for capturing state-dependent activity (Izmailov et al., 2020).

Third, we approximate the learned inverse network mappings using quadratic functions, as illustrated in Figure 1d. This enables extraction of correlations between neuronal activations through approximate characteristic functions. The approach obtains local coordinates that yield Riemannian curvature tensors for neural manifolds. This integrated approach offers both geometric and statistical characterization of the underlying structure.

While Riemannian geometry has been applied to neural decoding in other domains such as EEG analysis (Wilson et al., 2025), our approach appears to be among the first applications of Riemannian curvature analysis to characterize neural manifold geometry in population activity data. Our framework draws from recent developments in diffusion geometry and curvature estimation methods (Shao et al., 2018; Acosta et al., 2023; Jones, 2024). However, we distinctively integrate curvature-based geometric insights with statistical modeling of correlations. This enables analysis of large-scale electrophysiological recordings from hundreds of simultaneously recorded neurons. The approach produces interpretable descriptions of how neurons vary together, capturing complex statistical relationships that go beyond simple pairwise correlations.

Validation Through Macaque Visual Cortex Analysis

We demonstrate these capabilities through application to electrophysiological recordings from macaque visual cortex (Chen et al., 2022; ?). The analysis reveals distinct components where some clearly align with behavioral states while others exhibit pronounced non-Gaussian higher-order correlations. Components strongly aligned with behavioral states show reduced deviations from Gaussianity and reside on less curved neural manifolds. Other components display saddle-point-like manifold geometries, offering insights into how geometric structure relates to functional organization. A detailed analysis of these results is presented in Chapter 4.

This methodological framework establishes our second major contribution, complementing discrete temporal precision analysis with continuous geometric characterization. Together, these approaches form a comprehensive framework for understanding neural computation across multiple scales and representations.

Surrogate Methods for Higher-Order Spike Correlations

This chapter is based on

Stella*, A., Bouss*, P., Palm, G., and Grün, S. (2022). Comparing surrogates to evaluate precisely timed higher-order spike correlations. *eNeuro*, 9(3):ENEURO.0505–21.2022.

Author contributions

As joint first author, I designed the surrogate techniques and their implementation, performed the evaluation of the statistical properties of the surrogate techniques, and jointly led the writing of the manuscript. I contributed to the design of artificial data and participated in the analysis of both artificial and experimental data. I refined and updated the analytical code after the original publication to provide enhanced accuracy in pattern detection, resulting in revised quantitative estimates compared to the published version. An earlier version of this material was previously presented in the doctoral thesis of Alessandra Stella (2021). The work was conducted under the supervision of Günther Palm and Sonja Grün. Figures in this chapter were reproduced from Stella, Bouss, et al. (2022). Sebastian Lehmann helped in the design and development of the Figures 1.1 & 1.2.

1.1 Background

This chapter presents a comprehensive evaluation of surrogate techniques and their impact on pattern detection accuracy. The central goal is to identify patterns of neural activity which reflect true instances of temporal coding. These patterns represent precise spike timing relationships that cannot be explained by rate coding alone. While theoretical foundations provide the conceptual framework, practical implementation requires careful consideration of how different methods preserve or distort statistical properties of neural spike trains.

1.1.1 Neural Coding Fundamentals

Spike Trains as Point Processes

Neural communication occurs through electrical impulses called action potentials or spikes, which serve as basic units of information processing in biological neural networks (Rieke et al., 1997; Dayan and Abbott, 2001). These action potentials are elicited by incoming stimuli that can be either electrical or chemical in nature (Kandel et al., 2013; Purves et al., 2018). Electrical stimuli are transmitted directly through gap junctions between neurons, while chemical stimuli involve synaptic transmission through the release and binding of neurotransmitters with their respective receptors (Bennett and Zukin, 2004; Südhof, 2013).

From a mathematical perspective, the resulting spike trains can be represented as point processes where individual spike occurrences are characterized by their precise timing (Brillinger, 1992; Brown et al., 2001). For a single neuron, a spike train in the time interval $[0, T]$ can be described as a sequence of spike times: $t_1, t_2, \dots, t_n$ where $0 < t_1 < t_2 < \dots < t_n < T$. Such temporal precision enables investigation of neural coding mechanisms operating across multiple timescales, from millisecond-precise spike timing to slower firing rate modulations occurring over hundreds of milliseconds (Gerstein and Perkel, 1969; Abeles, 1991).

Coding Paradigms and Their Interactions

Temporal patterns of spikes contain information about neural computation, sensory processing, and motor control (Reich et al., 2001; Grün et al., 2002a). Previous research on information encoding and communication has traditionally aimed to distinguish between rate coding and temporal coding schemes (Shadlen and Newsome, 1998; Victor, 2005). However, this traditional dichotomous framing has been problematic because it approaches neural coding from an external observer's perspective, focusing on distinguishing between coding schemes rather than addressing the more

fundamental question of which mechanisms play causal roles in neural computation (Brette, 2015).

Rate coding operates on the principle that information is encoded in average firing rates over integration time windows that can range from tens of milliseconds to several seconds, depending on the context (Adrian, 1928; Georgopoulos et al., 1986). This scheme assumes that the time of spikes within the integration windows is less important than the overall firing frequency. Therefore, this view neglects spike timing over the coarser rate signal, treating the precise time of each spike as random noise. However, this rate-based view has been challenged as an ad hoc methodological postulate with limited empirical or theoretical support (Brette, 2015).

Temporal coding, in contrast, proposes that information is encoded in precise spike timing and relative temporal relationships between spikes (Carr and Konishi, 1990; Singer, 1999). This encompasses several types of temporal relationships: synchronous firing between neurons (Gray et al., 1989; ?), spike timing relative to oscillatory brain rhythms (Buzsáki and Draguhn, 2004; Fries, 2005), and temporal sequences of neural activity (Abeles, 1991; Diesmann et al., 1999). Modern formulations recognize that temporal coding can operate across multiple timescales and may involve complex patterns of neural coordination (Kiselev, 2016).

Recent empirical and theoretical work has increasingly challenged this traditional dichotomy, demonstrating that rate and temporal coding mechanisms often coexist and interact rather than represent mutually exclusive alternatives (Kayser et al., 2009; Panzeri et al., 2010; Zuo et al., 2015). Neural circuits may employ hybrid coding strategies. Rate modulations provide coarse information while temporal patterns convey fine-grained details. Indeed, it has been suggested that both spike-rate and spike-timing information coding schemes contribute complementarily to perceptual decisions (Zuo et al., 2015).

More fundamentally, some argue that the neural coding framework itself may be inadequate to understand brain function. It attempts to fit the dynamic, circular, and distributed causal structure of the brain into a linear chain of transformations between observables (Brette, 2019). This linear perspective assumes that neural activity can be decomposed into discrete coding steps, from stimulus encoding to motor output. However, the brain operates through complex feedback loops and distributed processing that resist such compartmentalization. The coding framework may therefore impose artificial constraints on our understanding of how neural circuits actually compute and process information.

Despite these theoretical critiques and evidence for hybrid mechanisms, methodological approaches are still needed to test specific hypotheses about temporal coordination in neural data (Stark and Abeles, 2009). Surrogate methods provide a statistical framework for testing whether spike patterns reflect temporal coordination beyond what would be expected from rate-based mechanisms alone. This approach allows us to evaluate the evidence for precise temporal relationships while acknowledging the complexity of neural coding strategies.

1.1.2 Surrogate Methods: Theoretical Framework

Implementing Null Hypotheses for Rate versus Temporal Coding

Since the aim of this work is to evaluate how meaningful precisely timed spike correlations identified by pattern detection algorithms are, surrogate methods are of special interest. Surrogate methods can specifically test whether the observed correlations can be fully explained by rate coding mechanisms rather than requiring precise temporal coordination between neurons. This approach involves testing the null hypothesis that rate-based processes alone can account for the detected patterns (Grün, 2009; Louis et al., 2010). Surrogate data methods implement specific null hypotheses when analytical solutions are unavailable or intractable, providing a computational approach to statistical testing (Date et al., 1999; Grün, 2009; Louis et al., 2010).

The fundamental concept involves generating multiple surrogate datasets that preserve the rate-based structure of the original data while destroying precise temporal relationships. This allows evaluation of whether identified correlations exceed those predicted by rate coding alone. The methodological approach involves computing test statistics such as pattern counts or correlation strength on both original and surrogate data, enabling calculation of significance levels based on the original statistic's rank within surrogate distributions (Efron and Tibshirani, 1993).

This nonparametric approach provides flexibility in handling complex data structures and arbitrary test statistics that may not have known analytical distributions. However, this framework assumes that rate and temporal coding can be cleanly separated, an assumption that may not hold in biological systems where these mechanisms interact dynamically (Kayser et al., 2009; Panzeri et al., 2010).

Design Principles and Statistical Foundations

Effective surrogate methods must balance competing objectives that often create tension in practical implementation (Grün, 2009; Louis et al., 2010). The primary requirement is to preserve all aspects of the data that could contribute to rate-based correlations while specifically destroying temporal correlations that operate beyond rate coding mechanisms. This represents a significant methodological challenge, as distinguishing between these two types of correlations in practice can be complex and may not always be cleanly achievable. This typically involves maintaining firing rate profiles and single-neuron spiking statistics while eliminating precise timing relationships between neurons (Pazienti et al., 2008).

1.1.3 Methodological Challenges in Spike Train Analysis

Data Quality and Modern Recording Technologies

Spike train analyses begin with spike sorting – the process of reconstructing single unit spike times from extracellular multi-unit recordings (Lewicki, 1998; Rey et al., 2015). Modern high-density silicon probes can record from hundreds to thousands of electrodes simultaneously (Jun et al., 2017; Steinmetz et al., 2021), creating unprecedented opportunities for population-level analysis but also amplifying the impact of sorting errors. Template matching algorithms used in contemporary spike sorting can achieve high accuracy but may introduce systematic biases in correlation detection, particularly for neurons with similar waveforms or overlapping spatial footprints (Pachitariu et al., 2016; Yger et al., 2018).

Spike sorting inevitably introduces some level of error, impacting correlation analysis in general (Wood et al., 2004; Pazienti and Grün, 2006). Even if this is the case, for the purposes of this study we assume spike sorting represents the true activity of individual neurons. This assumption allows us to focus on the comparative evaluation of surrogate methods applied to well-sorted spike train data.

Multi-scale Temporal Structure and Network Dynamics

Neural activity exhibits structure across multiple temporal scales, necessitating analysis methods that can separate fast correlations occurring on milliseconds to tens of milliseconds from slower rate modulations extending over hundreds of milliseconds to seconds (Vaadia et al., 1995; Riehle et al., 1997). This multi-scale nature reflects the hierarchical organization of neural circuits, where local microcircuit dynamics interact with global network states to produce complex correlation patterns (Honey et al., 2007; Deco et al., 2011). Correlation structures evaluated by surrogate methods represent local statistical dependencies that may emerge from global network properties such as small-world topology or modular organization (Bullmore and Sporns, 2009).

The relationship between local correlations and global network structure is particularly important when interpreting surrogate testing results, as correlations that appear significant locally may reflect network-wide phenomena rather than direct neuronal interactions (Gerstner et al., 2014). The temporal scales of correlation detection interact with network dynamics in complex ways, creating multiple layers of interpretation challenges.

Fast correlations may reflect direct synaptic connections or shared local inputs, while slower correlations may emerge from network-wide oscillations or global state changes (Buzsáki and Draguhn, 2004; Singer, 2013). This multi-scale structure creates a

complex scenario where the timescales of individual neuron dynamics may differ substantially from the timescales of correlations between neurons or across brain areas (Deco et al., 2013; Breakspear, 2017).

These considerations highlight important limitations in current surrogate method design. Ideally, surrogate methods should consider both the timescales of correlations being tested and their relationship to underlying network architecture, including potential mismatches between local dynamics and inter-neuronal correlation timescales. However, current approaches typically operate without knowledge of the underlying connectivity structure, limiting their ability to account for these network-level effects.

High-Dimensional Data and Statistical Challenges

Modern recording techniques create high-dimensional datasets requiring careful statistical treatment to avoid spurious findings (Stevenson and Kording, 2011). When analyzing n neurons for pairwise correlations, the number of tests equals $n(n-1)/2$ pairs. For 100 neurons, if one tests all pairs, this yields 4,950 tests, making it highly likely that some correlations will appear significant purely by chance even when no true correlations exist.

The high-dimensional nature of modern neural datasets also raises questions about the relationship between observed correlations and underlying network structure. Computational models provide a framework for understanding how local pairwise correlations relate to global population dynamics (Schneidman et al., 2006; Shlens et al., 2006). These models suggest that in many cases, pairwise correlations can capture most of the statistical structure in neural population activity, though higher-order interactions become increasingly important in larger populations (Ganmor et al., 2011; Macke et al., 2011).

Non-Stationarity and Temporal Dynamics

Neural activity often exhibits time-varying properties due to behavioral context, learning, or experimental conditions, introducing non-stationarity that complicates statistical analysis (Brown et al., 2001; Shimokawa and Shinomoto, 2009). This non-stationarity can be characterized as time-dependent changes in the statistical properties of spike trains, including firing rates and correlations between neurons (Kilavik et al., 2009; Churchland et al., 2010).

The implications of dealing with non-stationary data for surrogate methods are particularly important (Louis et al., 2010). Different surrogate techniques make different assumptions about temporal stability, and these assumptions may be violated when neural dynamics change over time. For example, trial shifting assumes that correlation structures remain constant across trials, making it inappropriate for detecting correlations that emerge during learning. In contrast, dithering methods that operate within individual trials may be more robust to slow changes across trials but could miss genuine correlations that develop consistently over time.

Subsampling and Network Inference Limitations

Despite advances in recording technology, available data represents only a tiny fraction of the millions or billions of neurons in complete neural circuits (Levina et al., 2022). This spatial subsampling creates substantial systematic biases when inferring collective properties of neural networks from observed subsets of neurons. When recording from a sparse sample of neurons, the observed correlation structures may fundamentally differ from the true population dynamics.

Subsampling can dramatically alter apparent network properties, with important implications for correlation detection and interpretation (Decker et al., 2009). These subsampling effects must be considered when interpreting surrogate-based analyses and may require different analytical approaches depending on the density and spatial organization of the recorded neuronal population.

1.1.4 Statistical Analysis Framework

Cross-Correlation Analysis

Cross-correlation analysis quantifies temporal relationships between spike trains by measuring the probability of observing spikes from one neuron at time τ relative to spikes from another neuron (Perkel et al., 1967a,b; Aertsen et al., 1989). The cross-correlation function is mathematically defined as $C_{AB}(\tau) = \frac{1}{T} \sum_i \sum_j \delta(t_i^A - t_j^B - \tau)$, where t_i^A and t_j^B represent spike times of neurons A and B respectively, δ represents the Dirac delta function, and T is the total observation time. Significant peaks or troughs in cross-correlograms indicate exceptional temporal correlations (Aertsen and Gerstein, 1985; Gray et al., 1989).

Higher-Order Patterns and Population Analysis

While pairwise correlations provide important insights into neural interactions, neural computation may depend on higher-order correlations involving multiple neurons simultaneously (Schneidman et al., 2006; Shimazaki et al., 2012). These higher-order interactions cannot be captured by simple pairwise analysis and may represent fundamental computational units in neural circuits (Yu et al., 2009; Ohiorhenuan et al., 2010).

Computational models provide a framework for understanding the relationship between pairwise and higher-order interactions in neural populations. These models can help determine whether pairwise correlations are sufficient to explain population activity or whether higher-order interactions are necessary (Shlens et al.,

2006). Pattern detection methods search for statistically significant patterns of near-synchronous activity across multiple neurons, extending beyond traditional analysis to capture the complexity of neural population dynamics (Grün et al., 2002a).

The challenge lies in distinguishing genuine higher-order correlations from those that can be explained by pairwise interactions, a problem that becomes increasingly complex as population size increases (Martignon et al., 2000; Ganmor et al., 2011).

Multiple Comparisons and Correction Strategies

When searching for neural patterns or correlations, researchers examine many potential patterns simultaneously, creating a multiple comparisons problem where the probability of finding significant results by chance increases dramatically with test numbers (Holm, 1979; Miller, 1981). Without appropriate correction procedures, the false positive rate can become unacceptably high, leading to spurious conclusions about neural coordination and connectivity.

Two primary approaches address this challenge with different trade-offs between statistical rigor and discovery power. Bonferroni correction represents the most conservative approach, adjusting the significance threshold α by dividing by the number of tests performed (Bonferroni, 1936; Dunn, 1961). False Discovery Rate (FDR) control offers a less conservative alternative, controlling the expected proportion of false discoveries among rejected hypotheses (Benjamini and Hochberg, 1995). This approach provides better balance between discovery power and statistical rigor, making it well-suited for high-dimensional neural data analysis (Genovese et al., 2002; Storey, 2002).

1.1.5 Biological Constraints and Surrogate Design

Neurobiological Realism in Surrogate Generation

Multiple factors complicate appropriate surrogate method design for spike trains (Pazienti et al., 2008; Nawrot et al., 2008). Complex interspike interval distributions represent a fundamental challenge, as neural spike trains often exhibit ISI distributions differing markedly from simple Poisson processes (Softky and Koch, 1993; Shadlen and Newsome, 1994). These non-Poissonian characteristics include bursting behavior, regular spiking, and complex multimodal distributions that must be preserved in surrogate data to maintain biological realism.

Refractory periods create additional constraints that must be respected in surrogate generation (Berry and Meister, 1998). Both spike sorting algorithms and biological neurons impose minimum intervals between detected spikes, typically ranging from 1-2 milliseconds for technical limitations to several milliseconds for biological absolute refractory periods. The biological absolute refractory period reflects the time

required for voltage-gated sodium channels to recover from inactivation, while the relative refractory period extends this constraint with gradually decreasing effect (Hodgkin and Huxley, 1952). Surrogate methods that violate these constraints may introduce artificial correlations or destroy genuine temporal structure.

Network-Level Constraints and Circuit Properties

The relationship between ISI distributions and network structure adds another layer of complexity to surrogate design. Different neuron types exhibit characteristic firing patterns that reflect their roles in circuit computation (McCormick et al., 1985; Connors and Gutnick, 1990), and ideally surrogate methods should preserve these patterns while destroying the specific correlations being evaluated. However, the heterogeneity of firing patterns contributes to the overall dynamics of neural circuits in ways that are difficult to capture in practical surrogate implementations, potentially leading to unintended alterations of circuit dynamics that affect correlation detection.

The balance between excitation and inhibition in neural circuits represents another challenging constraint for surrogate design (van Vreeswijk and Sompolinsky, 1996; Brunel, 2000). Correlations in neural networks often emerge from the interplay between excitatory and inhibitory influences, but surrogate methods typically lack the information necessary to account for how their manipulations affect this balance. Network models suggest that correlation structures may be particularly sensitive to changes in excitation-inhibition balance (Renart et al., 2010), yet practical surrogate implementations rarely have access to the cell-type information needed to preserve these relationships.

1.1.6 Surrogate Method Classification

Overview of Surrogate Approaches

Understanding the specific mechanisms and appropriate applications of different surrogate methods is crucial for evaluating whether identified correlations reflect rate-based or temporal coding mechanisms. Main methods for surrogate data generation include uniform dithering, which displaces spikes by small random amounts, preserving approximate firing rates while disrupting precise timing relationships, and trial shifting, which shifts entire trials independently for each neuron, thus preserving within-trial structure. Additional methods including ISI-based approaches and their variants provide alternative strategies for implementing rate-coding null hypotheses. A detailed examination of these methods and their comparative performance forms the core of this study, building on the theoretical framework established in this background section. The specific implementation details, parameter selection criteria, and performance characteristics of these methods are presented in the subsequent Materials and Methods section.

1.1.7 Experimental Context and Advanced Pattern Detection

Motor Cortex as a Model System

Motor cortex recordings during precision grip tasks provide an ideal testing ground for surrogate method evaluation due to their well-characterized neural dynamics and clear behavioral correlates (Kalaska and Crammond, 1992; Scott, 2004). Motor cortical networks exhibit distinct organizational principles that may affect correlation detection and surrogate method performance. The columnar organization of motor cortex creates structured patterns of connectivity that may generate specific correlation signatures (Rosen and Asanuma, 1972; Huntley and Jones, 1991). These columns represent vertical arrangements of interconnected neurons that share similar functional properties and preferred movement directions, creating local correlation structures that reflect the underlying anatomical organization (Strick and Preston, 1982; Donoghue and Wise, 1982). Despite this organizational structure, the functional significance of columns in motor cortex remains more debated than in sensory areas, where columnar organization is more clearly defined (Georgopoulos et al., 1986; Ashe and Georgopoulos, 1994; Scott, 2004).

The relationship between motor cortical activity and peripheral motor output introduces additional complexity for correlation analysis. Correlations between motor cortical neurons may reflect shared projections to spinal circuits rather than intrinsic cortical computation, raising questions about the interpretation of temporal coding in this system (Lemon, 2008; Rathelot and Strick, 2009). Recent work suggests that motor cortical activity may be constrained to lower-dimensional manifolds that reflect the true computational space of internal processing (Churchland et al., 2012; Gallego et al., 2017). This manifold structure could reduce the apparent correlation complexity to its essential computational components, potentially revealing genuine temporal coding mechanisms that are obscured by the high-dimensional correlation patterns observed in the full neural population (Shenoy et al., 2013; Vyas et al., 2020).

Within this theoretical framework, precision grip tasks provide controlled experimental paradigms where subjects grasp objects between thumb and fingers, enabling systematic control of force requirements and movement parameters while examining neural activity during different movement phases (Castiello, 2005; Brochier et al., 2018).

SPADE Algorithm and Pattern Detection

SPADE (Spike Pattern Detection and Evaluation) represents an advanced approach to identifying statistically significant spatiotemporal patterns in parallel spike trains (Torre et al., 2013; Quaglio et al., 2017; Stella et al., 2019). The algorithm extends beyond simple pairwise correlation analysis to detect higher-order spatiotemporal

patterns that may reflect genuine network-level computation. The algorithm operates through different steps including data discretization, frequent itemset mining, statistical evaluation using surrogate data methods, and multiple testing correction (Torre et al., 2013; Quaglio et al., 2017; Stella et al., 2019).

SPADE's dependence on surrogate methods means that all theoretical limitations of surrogate approaches apply directly to pattern detection results. The assumption that rate and temporal coding can be clearly separated becomes even more problematic for higher-order patterns, where interactions between coding mechanisms may be more complex. Furthermore, the multiple testing burden increases dramatically with pattern complexity, requiring sophisticated correction procedures to maintain statistical validity while preserving discovery power.

1.1.8 Chapter Summary

This background chapter has established the theoretical foundation for evaluating surrogate methods in the context of detecting precisely timed spike correlations. The framework acknowledges both the utility and limitations of traditional approaches to distinguishing rate-based from temporal coding mechanisms. Key challenges include the potential interaction between coding mechanisms, the relationship between local correlations and global network structure, the impact of experimental design on correlation detection, and the fundamental problem of establishing ground truth for temporal coding.

The systematic evaluation presented in subsequent sections addresses these challenges through complementary approaches using both artificial data with known ground truth and experimental motor cortex data with behavioral validation. The analysis framework recognizes that no single surrogate method may be optimal for all circumstances, and that method selection must consider the specific experimental context, network structure, and coding hypotheses being tested.

The motor cortex precision grip paradigm provides a well-characterized experimental system for surrogate method evaluation, though limitations in generalizability require careful consideration. The Brochier et al. (2018) dataset offers detailed behavioral monitoring and high-quality neural recordings that enable systematic investigation of surrogate method performance under realistic experimental conditions. Through this comprehensive evaluation framework, we aim to advance methodological foundations for detecting temporal coding mechanisms while acknowledging the theoretical complexities and practical limitations that constrain current approaches.

1.2 Materials and Methods

This chapter presents a systematic comparison of six surrogate techniques and their effects on neural pattern detection accuracy. Our empirical investigation bridges

the gap between surrogate methodology theory and real-world application, offering concrete guidance for method selection across diverse experimental scenarios.

1.2.1 Surrogate Methods

1.2.1.1 Theoretical Foundation

When analytical or generative probabilistic models prove unavailable for statistical analysis, computational null hypothesis testing through surrogate data becomes essential (Louis et al., 2010). Dithering techniques address a fundamental challenge in multi-spike-train analysis by separating correlations that arise from two distinct temporal scales. These correlations originate from firing rate modulations operating on timescales of tens of milliseconds (related to rate coding mechanisms) versus precise temporal spike coordination occurring on millisecond scales (related to temporal coding mechanisms) (Abeles, 1991; Grün et al., 2002a,b; Grün, 2009).

Our approach employs selective temporal disruption (Grün, 2009; Louis et al., 2010). By independently shifting spikes by intervals under 25 milliseconds, we eliminate all millisecond-scale dependencies while preserving dependencies at broader temporal scales used for firing rate estimation (Grün et al., 2002a,b). This selective process maintains correlation structures that link different neurons and connect neurons with external stimuli or experimental conditions at temporal resolutions exceeding 25 milliseconds.

1.2.1.2 Method Selection Criteria

Our comparative framework emphasizes surrogate techniques that maintain original data firing rate modulations alongside other critical spike train properties, particularly interspike intervals (ISIs) and associated measurements. This preservation strategy addresses uncertainties about how ISI modifications might influence pattern significance evaluation (Brown et al., 2004; Pipa et al., 2008). While we cannot perfectly preserve all spike train features, our investigation determines which surrogate strategies produce minimal false positive outcomes (Grün, 2009; Ventura, 2010).

We systematically evaluate six surrogate techniques, comprising four well-established methods from published research and two innovative strategies we developed specifically for this investigation. Our method selection deliberately excludes surrogates that compromise firing rate profiles. Instead, we focus on strategies engineered to maximize preservation of spike train properties, including ISI distributions and related parameters like the coefficient of variation (CV), which measures the ratio of the standard deviation to the mean of interspike intervals.

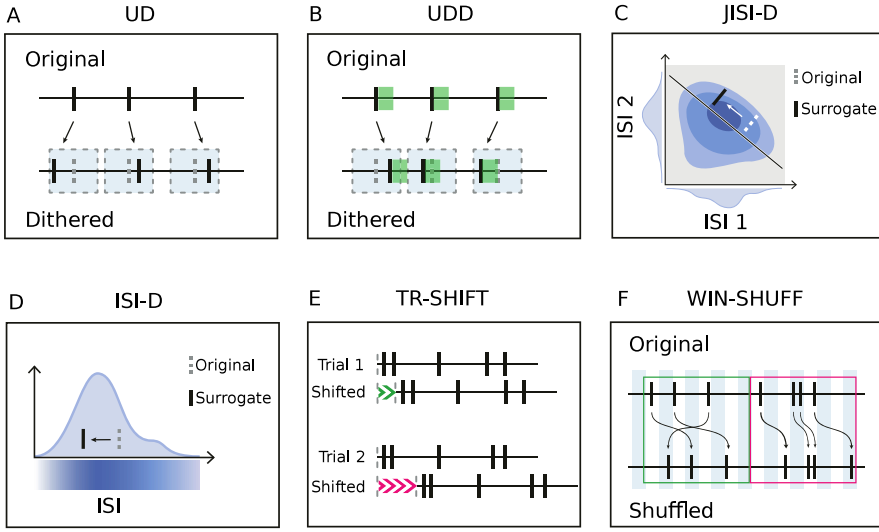


Figure 1.1: Six surrogate generation methods for spike train analysis. Schematic illustrations of different approaches to generate surrogate data while preserving specific statistical properties of the original spike trains. **(A) Uniform Dithering (UD).** Individual spikes are randomly displaced within fixed temporal windows (gray dotted rectangles) using uniform probability distributions. This method eliminates precise spike timing while preserving average firing rates. **(B) Uniform Dithering with Dead-Time (UDD).** Similar to UD but enforces minimum separation distances between spikes. Green shaded regions show exclusion zones after each spike due to refractory period constraints, preventing biologically unrealistic spike clustering. **(C) Joint ISI Dithering (JISI-D).** Spike repositioning preserves the joint distribution of consecutive interspike intervals from the original data. The blue heat map shows the bivariate probability distribution of ISI pairs, with marginal distributions displayed on the axes representing first and second ISI durations in spike triplets. **(D) ISI Dithering (ISI-D).** Spike displacement follows the neuron’s empirical ISI distribution (blue histogram) while ignoring correlations between consecutive intervals. This preserves ISI statistics but not their sequential relationships. **(E) Trial Shifting (TR-SHIFT).** Entire trials are shifted by random temporal offsets (green and pink arrows show different shifts per trial), preserving all within-trial spike relationships while eliminating cross-trial correlations. **(F) Window Shuffling (WIN-SHUFF).** Spike counts within temporal bins are randomly permuted within non-overlapping windows (green and pink regions), preserving local spike counts while disrupting temporal order.

1.2.1.3 Uniform Dithering

Uniform dithering shifts spikes from each neuron using small, uniformly distributed random displacements $\sim U[-\Delta, +\Delta]$ relative to their original temporal locations (Figure 1.1A). Also known as jittering or teetering, this technique has found extensive application across experimental investigations and represents a standard approach in several major analysis frameworks (Abeles and Gat, 2001; Hatsopoulos et al., 2003; Gerstein, 2004; Shmiel et al., 2006; Maldonado et al., 2008; Torre et al., 2016).

The method’s computational efficiency makes it suitable for pairwise synchrony detection via cross-correlogram significance testing, higher-order synchrony evaluation, and complex pattern identification tasks (Grün, 2009; Torre et al., 2013; Quaglio et al., 2017; Stella et al., 2019). However, Louis et al. (2010) demonstrated that uniform dithering generates false positives when analyzing neurons with regular firing characteristics ($CV < 1$), motivating systematic comparison of alternative methods.

The displacement parameter $\Delta > 0$ controls maximum spike movement and requires careful calibration. Typical ranges span 15-25 milliseconds, with small Δ values causing insufficient displacement and large values creating excessive firing rate smoothing hypotheses (Pazienti et al., 2008; Torre et al., 2016).

1.2.1.4 Uniform Dithering with Dead-Time

We develop a novel variant that preserves post-spike dead-time periods, typically extending to 2 milliseconds in experimental datasets (Berry and Meister, 1998; Brochier et al., 2018) (Figure 1.1B). This enhancement addresses a biological constraint that standard uniform dithering ignores, potentially compromising null hypothesis validity.

Our modification extracts the dead-time d for each neuron from experimental observations and restricts dithering windows based on neighboring spike intervals reduced by the estimated dead-time. This prevents modified spikes from violating the dead-time constraint, preserving biological realism for valid statistical inference.

We calculate dead-time using minimum interspike interval measurements across complete experimental trials. However, low firing rates can produce minimum ISI values reaching hundreds of milliseconds, creating challenges for meaningful dead-time determination. Therefore, we establish a maximum dead-time threshold $d_{\max}$ (4 milliseconds in this work) to maintain biological plausibility while ensuring adequate temporal displacement for correlation elimination.

1.2.1.5 Joint-ISI Dithering

Joint-ISI dithering, developed by Gerstein (2004), relocates spikes based on joint distributions of neighboring ISI intervals (Figure 1.1C). This technique eliminates correlations of interest while preserving the statistical structure of interspike intervals (Softky and Koch, 1993; Shinomoto et al., 2003), addressing a fundamental limitation of uniform dithering.

The method preserves distributions of both preceding and succeeding interspike intervals relative to each spike. We derive probability functions from original datasets by calculating joint-ISI histograms using 1-millisecond bins, providing sufficient temporal detail while maintaining computational tractability.

Spike relocation follows the two-dimensional histogram structure, corresponding geometrically to displacement along joint-ISI distribution anti-diagonals (Gerstein, 2004; Louis et al., 2010). This ensures temporal structure remains statistically consistent with original data, preserving relationships between consecutive interspike intervals essential for biological realism.

Limited recording duration can restrict spike samples available for robust probability distribution calculation. We address this through two-dimensional Gaussian smoothing of joint-ISI histograms with a σ parameter of a few milliseconds (Louis et al., 2010). This improves distribution estimation reliability while preserving essential temporal relationships.

1.2.1.6 ISI Dithering

ISI dithering (ISI-D) represents a simplified version of joint-ISI dithering that repositions spikes based on marginal ISI probability distributions while ignoring sequential connections between consecutive intervals (Figure 1.1D). This approach offers computational advantages when data limitations prevent reliable joint-ISI estimation, particularly for datasets with limited recording duration or low spike counts.

ISI-D assumes consecutive ISI independence, reducing the joint-ISI histogram to the product of marginal ISI histograms ($p_{\text{joint-ISI}}(\tau, \tau') = p_{\text{ISI}}(\tau) \cdot p_{\text{ISI}}(\tau')$). This independence assumption simplifies computation while maintaining many desirable statistical properties, though it sacrifices preservation of sequential ISI relationships.

The technique maintains marginal ISI distributions independently of temporal sequence, representing a trade-off between computational tractability and biological realism. Despite this limitation, ISI-D provides substantial improvement over uniform dithering in statistical property preservation while remaining computationally efficient for large-scale applications.

1.2.1.7 Trial Shifting

Trial shifting displaces complete spike trains through temporal shifts of entire trials (Figure 1.1E; Pipa et al., 2008). This method suits neuroscience experiments that organize recordings into discrete temporal segments or trials, a common approach in motor cortical and other behavioral studies (Riehle et al., 1997).

The strategy completely maintains firing rate and ISI structure within trials while eliminating correlations among distinct spike trains. We shift all spike times within trials uniformly by random amounts $\sim U[-\Delta, +\Delta]$, applied independently across each neuron-trial combination.

Independent temporal shifts across neurons eliminate correlations between spike trains while preserving individual firing patterns within each trial. When experimental designs lack clear trial boundaries, spike trains can be divided into extended temporal segments or sequences separated by substantial interspike intervals Harrison and Geman (2009) suggested. Trial shifting preserves complete spike train structural integrity within these segments.

1.2.1.8 Window Shuffling

Window shuffling (WIN-SHUFF) represents our second novel design that preserves spike counts from discretized original spike trains throughout all surrogate realizations (Figure 1.1F). We partition spike trains into consecutive, non-overlapping temporal windows of duration Δ_{WS} , then subdivide windows into bins of length b (requiring Δ_{WS} to be a multiple of b). We shuffle bins within windows before randomizing spike times within individual bins.

This local spike rearrangement disrupts temporal correlations while creating stationarity within Δ_{WS} duration windows. For comparison consistency with other techniques, we maintain $\Delta_{WS} = 2\Delta$ throughout the analysis.

1.2.2 SPADE Algorithm Application

1.2.2.1 Framework Overview

We assess pattern detection using SPADE (Spike Pattern Detection and Evaluation), a software framework for evaluating spatiotemporal spike patterns in multi-unit recordings (Torre et al., 2013; Quaglio et al., 2017; Stella et al., 2019). SPADE integrates surrogate techniques with frequent itemset mining (FIM) algorithms to address computational challenges from vast numbers of recurring patterns in original and surrogate datasets (Borgelt, 2012). Figure 1.2 illustrates the complete workflow.

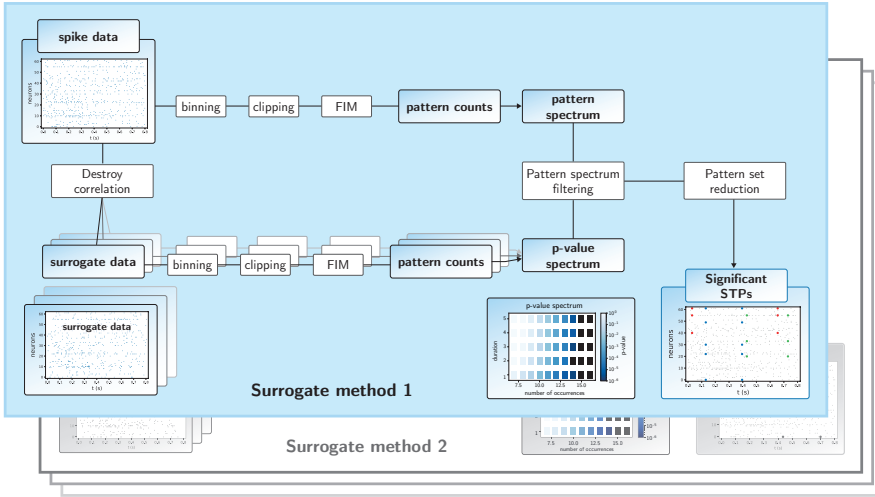


Figure 1.2: **SPADE analysis workflow for pattern detection and significance testing.** Schematic overview of the complete SPADE pipeline showing parallel processing of original and surrogate data. **Upper pathway:** Original experimental data undergoes preprocessing (binning and clipping), followed by frequent itemset mining to identify candidate patterns and generate a pattern occurrence spectrum. **Lower pathway:** Surrogate data generation using dithering methods (multiple techniques can be applied, as indicated by "Surrogate method 2"), followed by identical preprocessing and mining steps. The overlapping panels represent multiple surrogate iterations (typically 5000) that collectively build the null distribution for statistical testing. **Statistical evaluation:** Pattern spectrum filtering compares observed pattern counts against the null distribution to compute p-values, followed by pattern set reduction to eliminate spurious patterns caused by chance overlaps. **Output:** Statistically significant spatio-temporal patterns (STPs) that cannot be explained by firing rate fluctuations alone. This framework allows systematic comparison of different surrogate methods while maintaining consistent analytical procedures throughout the pipeline.

1.2.2.2 Data Preprocessing and Pattern Mining

Spike trains are discretized into non-overlapping temporal bins, establishing acceptable temporal imprecision thresholds. We count spikes within bins (binning) then reduce multiple spikes per bin to 1 (clipping), constituting binarization (Grün et al., 2002a; Torre et al., 2013).

We identify candidate patterns using the Frequent Itemset Mining (FIM) algorithm – a computational method that systematically searches through binarized data to find sets of neurons that spike together repeatedly within specified time windows (Zaki, 2004; Borgelt, 2012; Picado-Muiño et al., 2013). The FIM algorithm efficiently enumerates all possible neuron combinations and count how often each pattern occurs across the dataset. The output includes occurrence frequencies and precise temporal locations for patterns meeting minimum frequency thresholds (non-trivial patterns that cannot be explained as subsets of larger, equally frequent patterns).

1.2.2.3 Pattern Spectrum and Significance Assessment

We organize pattern data into three-dimensional histograms (pattern spectra) defined by: number of neurons (z), occurrence count (c), and pattern duration (d). Patterns sharing identical signatures are aggregated, reducing statistical tests required for multiple testing correction.

SPADE tests whether patterns result from precise temporal coordination or random spike generation. The null hypothesis assumes spike train independence given firing rate modulations. We generate 5000 surrogates undergoing identical processing, then compute three-dimensional p-value spectra (Torre et al., 2013).

1.2.2.4 Statistical Filtering and Optimization

We apply Benjamini-Hochberg correction (Benjamini and Hochberg, 1995) for multiple testing. and filter patterns using a significance threshold ($\alpha = 0.05$). Pattern set reduction eliminates spurious false positives from overlapping patterns (Torre et al., 2013).

For computational efficiency, we use dynamic minimum occurrence counts that vary by pattern size. We first determine the expected occurrence range for each pattern size, then set `min_occ` values accordingly when computationally feasible, with an absolute minimum of $\lceil \# \text{trials} / 3 \rceil$. This adaptive approach reduces the search space during FIM by filtering out patterns with very low occurrence counts that are unlikely to reach statistical significance, thereby reducing computational burden without affecting the detection of meaningful patterns (Stella et al., 2019).

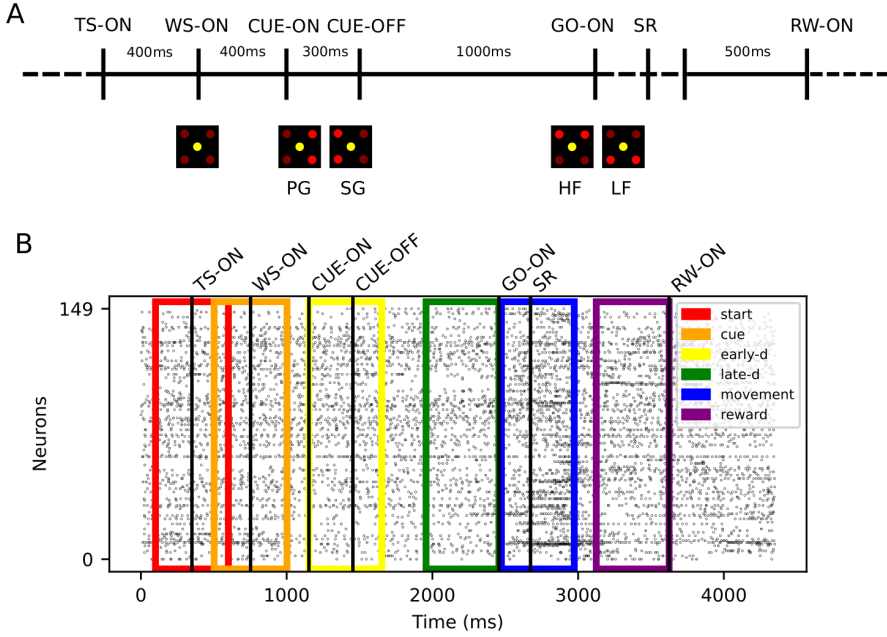


Figure 1.3: Experimental paradigm and neural data structure. (A) Behavioral task timeline. Schematic of the reach-to-grasp task sequence. Monkeys initiate trials voluntarily (TS: trial start), then receive a waiting signal (WS-ON) followed by grip type instruction (CUE-ON: precision grip PG or side grip SG). After a 1000 ms delay, a GO signal appears indicating force requirement (high HF or low LF) and cueing movement initiation. Key behavioral events include switch release (SR: movement onset), object touch, and reward delivery (RW-ON) after maintaining the grip for 500 ms. Timing variability in SR and RW reflects natural differences in reaction time and movement execution. **(B) Neural population activity example.** Raster plot showing simultaneous spike activity from all recorded neurons (y-axis) over time (x-axis) for a representative PGHF trial from session i140703-001. Each dot represents a single spike, aligned to trial start (TS-ON). Colored background regions indicate the six analysis epochs used for pattern detection: start (red), cue (orange), early-delay (yellow), late-delay (green), movement (blue), and hold (purple). This temporal segmentation allows examination of neural coordination patterns across different phases of the behavioral task.

1.2.3 Experimental Data and Preprocessing

1.2.3.1 Experimental Design

Our analytical strategy combines two complementary approaches using experimental data from Brochier et al. (2018). We generate synthetic spike trains replicating experimental statistical properties to evaluate false positive rates, and directly analyze experimental data using multiple surrogate techniques to assess performance under realistic conditions.

1.2.3.2 Data Collection Protocol

Brochier et al. (2018) collected data during reaching-and-grasping experiments from pre-motor and motor cortex recordings in two macaque monkeys using chronic 100-electrode Utah arrays (Blackrock Microsystems). The experimental design is detailed in previous publications (Riehle et al., 2013; Brochier et al., 2018; Riehle et al., 2018).

Figure 1.3A illustrates the experimental protocol. Monkeys initiated trials by pressing a start button (TS), followed by a 400ms interval and yellow LED waiting signal (WS). After another 400ms, visual cues appeared for 300ms (CUE-ON to CUE-OFF) specifying grip type: full-hand side grip (SG) or precision grip (PG). Following a 1000ms interval, the GO-signal specified grip force (high/low) through LEDs. Movement onset was detected through switch release (SR), followed by object touch (OT), holding period (HS), and reward delivery (RW) after 500ms retention. Behavioral conditions were randomized to prevent anticipatory effects.

1.2.3.3 Data Processing and Quality Control

The datasets comprise two sessions (i140703-001 and l101210-001) with 15-minute recordings containing ~35 trials per condition. Spike sorting used Plexon Offline Spike Sorter (v3.3), with data available at https://gin.g-node.org/INT/multielectrode_grasp.

Neuron inclusion required spike shape signal-to-noise ratios >2.5 and firing rates <70 Hz. Hyper-synchronous artifact spikes from electrical cross-talk were automatically identified and removed following established protocols (Torre et al., 2016; Oberste-Frielinghaus et al., 2025). Analysis included only successful behavioral trials.

1.2.3.4 Temporal Segmentation

We divided trials into six partially overlapping 500ms epochs spanning complete trials, facilitating behavioral context association. Matching epoch segments within trial types were concatenated to produce 24 datasets per session (4 trial types $\times$ 6 epochs) for independent analysis (Torre et al., 2016). Figure 1.3B illustrates temporal organization and epoch positioning relative to behavioral events.

1.2.4 Artificial Non-Stationary Data Generation

1.2.4.1 Modeling Framework

To systematically evaluate surrogate performance under controlled conditions, we generate artificial non-stationary spike trains that preserve only specific statistical properties from experimental data while eliminating all temporal correlations. We use two point process models: Poisson processes with dead-time (PPD; Deger et al., 2012) and Gamma processes. These models replicate experimental firing rates and either dead-time characteristics (PPD) or regularity patterns (Gamma), but contain no genuine spike correlations, providing controlled null datasets for surrogate technique validation.

1.2.4.2 Firing Rate Profile Estimation

We create artificial spike trains using original firing rate profiles calculated via optimized kernel density estimation on trial bases (Shinomoto, 2010; Shimazaki and Shinomoto, 2010). Time-varying firing rates constitute essential experimental characteristics requiring incorporation into null hypothesis frameworks.

1.2.4.3 Poisson Process with Dead-Time

PPD modifies classical Poisson processes by preventing spike generation within dead-time intervals d following preceding spikes. We calculate the dead-time for each neuron and epoch/trial-type combination using minimum ISI measurements. The implementation adjusts rate signals to compensate for dead-times, generates non-stationary Poisson spike trains, then removes spikes violating the dead-time constraint through post-processing.

1.2.4.4 Gamma Process Creation

We create Gamma processes using shape factors calculated from coefficient of variation (CV) measurements in operational time, converting CV values to shape factors ($\gamma = 1/\text{CV}^2$; Nawrot et al. 2008; van Vreeswijk 2010). Process creation occurs in operational time with subsequent real-time transformation. We assess regularity using CV2 measurements compensating for non-stationary firing rates (Holt et al., 1996).

Final CV2 distributions closely approximate experimental data. Although Gamma processes lack absolute dead-times, $\gamma > 1$ conditions produce low probabilities for small ISIs, effectively incorporating relative dead-times (Nawrot et al., 2008).

1.2.4.5 Dataset Generation

We model each of the 24 experimental datasets using both PPD and Gamma strategies, producing 96 total artificial datasets (2 sessions $\times$ 24 conditions $\times$ 2 models) for comprehensive surrogate technique assessment.

1.2.5 Code Accessibility and Computational Resources

Complete analytical code and figure reproduction scripts are available at https://github.com/INM-6/SPADE_surrogates. Experimental data can be obtained at https://gin.g-node.org/INT/multielectrode_grasp, with artificial data generation included in the SPADE_surrogates repository. SPADE methodology and surrogate techniques are implemented in the Elephant Python package (<https://elephant.readthedocs.io>).

Large-scale analyses may require several hours of cluster computing time depending on dataset size and surrogate count, despite computational performance enhancements.

1.3 Results

1.3.1 Statistical Evaluation of Surrogate Techniques

1.3.1.1 Experimental Design for Statistical Characterization

To develop a comprehensive understanding of how surrogate techniques affect spike train statistical characteristics, we assess their performance using stationary synthetic data with well-characterized properties. This controlled approach provides the foundation for interpreting experimental results by establishing baseline performance metrics under idealized conditions.

Our controlled approach utilizes three point process models with precisely defined attributes. The baseline Poisson process provides a reference for random spike generation, establishing expected behaviors under null conditions. The Poisson process with dead-time constraints (PPD) replicates refractory period effects commonly observed in biological neurons, addressing temporal constraints that prevent immediate spike generation following previous spikes. The Gamma process models regularity variations observed in cortical neurons, capturing the spectrum of firing patterns from highly regular to highly irregular that characterizes real neural populations.

These models specifically replicate ISI distribution features characteristic of experimental observations, ensuring that our synthetic data maintains biological relevance

while providing controlled conditions for evaluation. The use of stationary processes deliberately removes confounding statistical variables, enabling focused study of surrogate technique effects on core spike train attributes without complexity from time-varying firing rates.

We calibrate model parameters to match experimental data characteristics from macaque motor cortex recordings (Riehle et al., 2013; Torre et al., 2016; Brochier et al., 2018), ensuring meaningful result interpretation for subsequent experimental dataset applications. Figure 1.4 summarizes our comprehensive findings from this stationary, independent data study, providing the foundation for understanding surrogate behavior under controlled conditions.

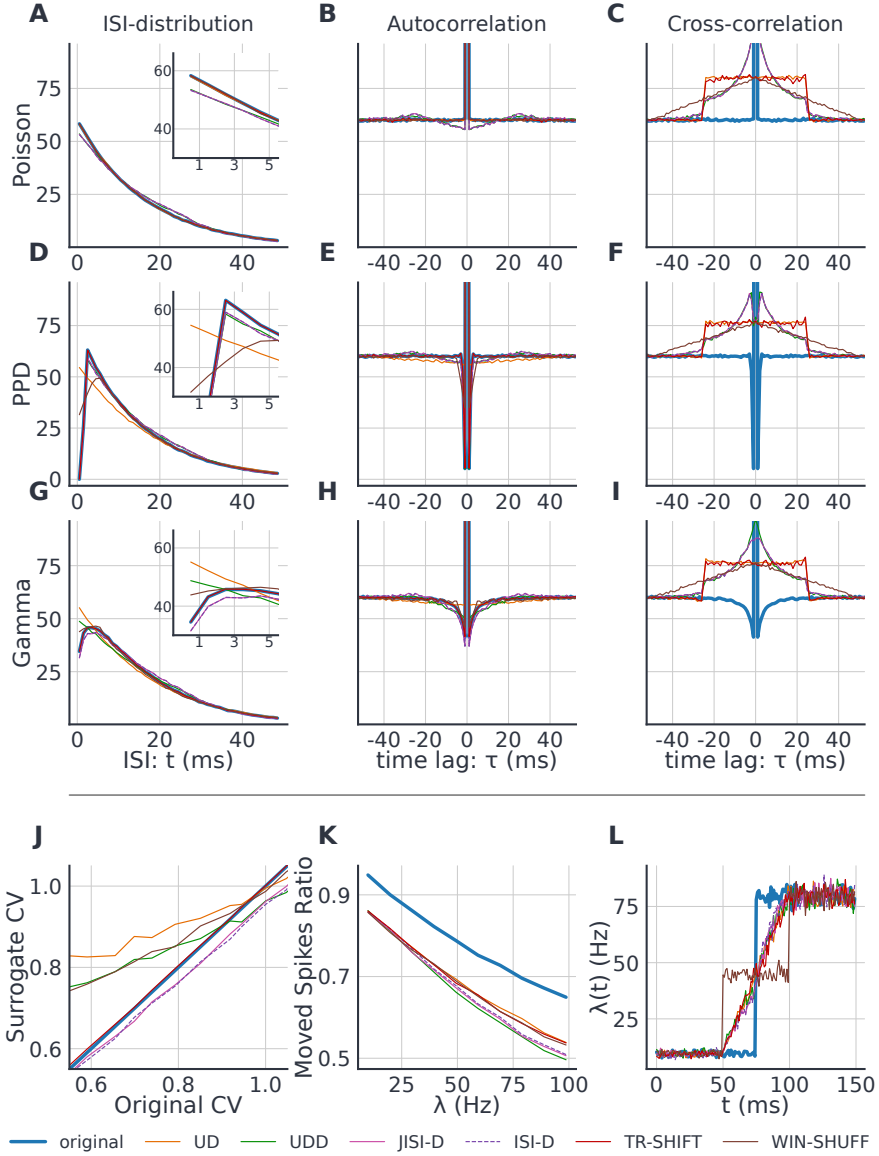


Figure 1.4: Statistical property preservation across surrogate methods.

Figure 1.4: Statistical property preservation across surrogate methods. Comprehensive evaluation of how six surrogate techniques (UD: orange, UDD: green, JISI-D: pink, ISI-D: violet, TR-SHIFT: red, WIN-SHUFF: brown) affect spike train statistics compared to original data (blue). All analyses use synthetic spike trains with ~500,000 spikes at 60 Hz unless noted otherwise. **(A,D,G) Interspike interval distributions (1/s).** ISI probability distributions for three process types (Poisson, PPD, Gamma; top to bottom) showing preservation of temporal structure. Inset: magnified view of ISI ≤ 5 ms range. **(B,E,H) Autocorrelation analysis (1/s).** Temporal structure within spike trains for original (blue) versus surrogate data (colors as in A-C) over the time lag τ (ms). **(C,F,I) Cross-correlation analysis (1/s).** Similarity between original and surrogate spike trains across process types. shown over the time lag τ (ms). **(J) Coefficient of variation preservation.** Original CV versus surrogate CV for Gamma processes with CV ranging 0.4-1.2 in 0.05 steps. Perfect preservation would follow the diagonal. **(K) Spike repositioning efficacy.** Fraction of spikes moved from original positions across firing rates λ (Hz) 10-100 Hz (Gamma processes, $\gamma = 1.23$). Higher values indicate better correlation elimination (dimensionless ratio). **(L) Firing rate profile fidelity (Hz).** Preservation of temporal rate modulations $\lambda(t)$ during abrupt transitions (10 $\rightarrow$ 80 Hz step function) in Gamma spike trains. Time t (ms). Computed from 10,000 trials, 150 ms each, using 1 ms bins. Systematic deviations indicate rate profile distortion by surrogate methods.

1.3.1.2 ISI Distribution Characterization

The characterization of ISI distributions (Figure 1.4A,D,G) reveals distinct patterns across surrogate methods and process categories, providing fundamental insights into how different techniques modify the temporal structure of spike trains. Our analysis uncovers critical patterns that distinguish the methods in their treatment of brief interspike intervals, which prove crucial for understanding their relative strengths and limitations.

Across all surrogate datasets, ISI distributions consistently manifest exponential decay properties, though with marked differences in their treatment of brief ISIs. The methods JISI-D, ISI-D, and UDD show markedly reduced probabilities for brief ISIs compared to their unconstrained counterparts, reflecting their inherent design constraints that limit repositioning based on neighboring spike locations. Our analysis of Poisson data, especially at increased rates, confirms frequent brief ISI occurrences that match underlying process expectations.

The analysis of PPD processes reveals that uniform dithering commonly relocates spikes into time windows extending to 5 milliseconds, precisely matching the bin width specifications used in our analysis. This relocation behavior demonstrates how UD can violate the biological constraints imposed by refractory periods (Berry and Meister, 1998), potentially creating unrealistic spike patterns that compromise the validity of the null hypothesis. In contrast, other surrogate methods that preserve dead-time limitations adhere more closely to original ISI distribution patterns, maintaining the biological realism essential for accurate significance testing.

For Gamma processes utilizing $\gamma = 1.23$, which falls within the range of shape factors observed in cortical neurons (Shinomoto et al., 2003), ISI distributions undergo maximum modification under UD and UDD methods. These modifications fundamentally alter the regularity characteristics of the original spike trains, potentially leading to false positive results when the null hypothesis assumes preservation of these properties. The importance of maintaining these regularity characteristics reflects the functional significance of irregular firing patterns in cortical neurons, where deviations from regular firing carry computational information (Softky and Koch, 1993). Other techniques demonstrate superior preservation of brief ISI properties, with distributions that nearly match original characteristics while still providing adequate temporal disruption for correlation elimination.

1.3.1.3 Spike Repositioning and Cross-Correlation Analysis

Cross-correlation analysis between synthetic datasets and their matching surrogate versions (Figure 1.4A,E,H) reveals systematic differences in spike repositioning patterns that fundamentally distinguish the approaches and their underlying assumptions. This analysis builds on established methods for detecting spike correlations and assessing their statistical significance (Perkel et al., 1967b; Aertsen et al., 1989). The distinct repositioning patterns provide crucial insights into how each method balances the competing requirements of correlation elimination and statistical property preservation.

Uniform methods including UD and TR-SHIFT produce characteristically uniform distributions within their respective dither boundaries, reflecting their unconstrained approach to spike repositioning. This uniform distribution indicates that spikes can be moved with equal probability to any location within the allowed temporal window, providing maximum disruption of temporal correlations. Window shuffling creates distinctive triangular configurations with high zero-lag probability, resulting from its binned approach to spike redistribution that favors local rearrangements over distant repositioning.

In marked contrast, ISI-based methods including JISI-D, ISI-D, and UDD demonstrate exponential-type decay patterns reaching the 25-millisecond dither boundary. This characteristic decay reflects their constrained repositioning capabilities, where the probability of spike movement depends on the local ISI structure surrounding each spike. The exponential nature of this decay indicates that shorter repositioning distances occur more frequently than longer ones, preserving more of the original temporal structure while still providing some degree of correlation disruption.

The design constraints inherent in JISI-D, ISI-D, and UDD methods fundamentally limit dithering to time windows between adjacent spikes, creating a dependency between repositioning capability and local spike density. For Poisson data, particularly at high rates where spikes occur in close temporal proximity, these available dither

boundaries may remain partially exploited due to neighboring spike constraints. This limitation results in systematically reduced repositioning on average, thereby preserving closer proximity to original temporal locations compared to unconstrained methods..

1.3.1.4 Auto-correlation Function Analysis

Auto-correlation studies of synthetic processes and their surrogate-altered versions (Figure 1.4C,F,I) provide fundamental insights into how ISI properties undergo modification during surrogate generation, following established approaches for analyzing temporal dependencies in neural spike trains (Perkel et al., 1967a). Our analysis reveals the deeper mechanisms that distinguish different approaches to temporal disruption and illuminates the trade-offs between correlation elimination and statistical property preservation that define surrogate method performance.

The analysis of Poisson processes demonstrates that auto-correlations exhibit characteristically flat profiles for UD, TR-SHIFT, and WIN-SHUFF methods when central peaks are excluded from consideration. This flatness indicates successful elimination of temporal correlations while maintaining the essential randomness of the underlying Poisson process. However, JISI-D, ISI-D, and UDD display markedly different behavior, showing reduced probabilities for extremely brief ISIs succeeded by characteristic elevations reaching maximum dither width.

This behavioral distinction originates from the spike sequence preservation constraints inherent in ISI-based methods, creating asymmetries in repositioning probabilities that generate observable signatures in the auto-correlation functions. When reference spikes are positioned closer to subsequent spikes than to preceding ones, the algorithms favor reverse temporal repositioning over forward repositioning to maintain ISI distribution properties.

Our examination of PPD processes reveals that auto-correlations for JISI-D, ISI-D, and UDD exhibit reduced brief ISIs relative to matching Poisson surrogate methods, though these reductions occur to lesser magnitudes than observed in pure Poisson cases. This attenuation reflects the interaction between the dead-time constraints of the PPD process and the ISI preservation mechanisms of the surrogate methods. The analysis of Gamma processes confirms excellent preservation of original process properties across most methods, with TR-SHIFT maintaining virtually identical characteristics to original processes.

1.3.1.5 Coefficient of Variation Preservation

Examining coefficient of variation (CV) preservation represents a critical component of our analysis, since surrogate methods that alter ISI distributions directly impact

this crucial regularity measure. CV provides a normalized measure of firing regularity that remains interpretable across different firing rates and experimental conditions, making its preservation essential for maintaining the biological relevance of surrogate-generated null hypotheses.

Our investigation focuses on CV modifications in surrogate data relative to original Gamma processes operating at 60 Hz, with CV values spanning from 0.4 to 1.2 in 0.05 increments (Figure 1.4J). This systematic parameter sweep enables comprehensive evaluation of method performance across the full spectrum of neuronal regularity, from highly regular firing patterns through highly irregular patterns. The analysis reveals that CV non-preservation in surrogate data constitutes a significant potential source of false positives, particularly for neurons with extreme regularity or highly irregular firing patterns.

Uniform dithering demonstrates the most severe CV modifications among all tested methods, dramatically converting original 0.4 values to 0.75, indicating substantial regularity decrease that fundamentally alters neuronal firing characteristics. This transformation represents a qualitative change in firing pattern classification, converting highly regular neurons into moderately irregular ones and potentially invalidating statistical comparisons based on regularity assumptions. For original CV values of 1.2, the method shows increasing deviation with gradual slopes reaching maxima marginally above 1.0, reducing burstiness properties in highly irregular neurons.

Window shuffling and UDD exhibit comparable behaviors to UD in their general trends but demonstrate superior preservation of lower CV values, particularly notable for highly regular neurons with original CV values of 0.4. The ISI-based methods including JISI-D and ISI-D begin with CV values matching those of the original Gamma processes, reflecting their explicit design to preserve ISI distribution properties. However, they show systematically reduced slopes relative to the source processes and attain elevated final values approximately 0.05 below maximum CV levels at 1.2.

In stark contrast, trial shifting sustains virtually perfect CV preservation across all examined values, establishing it as the gold standard for regularity maintenance in surrogate generation. This exceptional performance reflects its fundamental difference in approach: complete spike trains are temporally shifted rather than individual spikes being repositioned, preserving all within-trial temporal relationships while eliminating between-trial correlations.

1.3.1.6 Spike Repositioning Efficacy

The assessment of spike repositioning efficacy represents a fundamental requirement for effective surrogate performance, as adequate displacement of spikes from their

original bin locations becomes essential for eliminating intended correlations while preventing false negative outcomes. We quantify this repositioning efficacy by measuring spike relocation ratios – the proportion of spikes moved from their original bin locations relative to total spike counts – across systematically varied firing rates ranging from 10 to 100 Hz in 10 Hz increments (Figure 1.4K).

Our comprehensive study incorporates critical reference comparisons using two independent Gamma spike train realizations with $\gamma = 1.23$ to establish optimal repositioning targets. These independent realizations provide a benchmark representing complete decorrelation while maintaining the same underlying statistical properties, thereby defining the performance ceiling for surrogate methods. Our analysis reveals that no tested method achieves this optimum, with all techniques consistently exhibiting approximately 10% fewer repositioned spikes relative to the independent implementation benchmarks.

The systematic evaluation reveals distinct performance tiers among the surrogate methods. High efficacy methods including WIN-SHUFF, UD, and TR-SHIFT demonstrate consistently superior repositioning capabilities across all examined firing rates, maintaining relatively stable performance even as firing rates increase and bin occupancy becomes more dense. In contrast, moderate efficacy methods including UDD, ISI-D, and JISI-D show progressively deteriorating repositioning performance, particularly at elevated firing rates where their design constraints become increasingly limiting. At maximum firing rates, approximately 50% of spikes processed by JISI-D, ISI-D, and UDD methods remain in their original bins, demonstrating the fundamental trade-off between statistical property preservation and repositioning efficacy.

1.3.1.7 Firing Rate Profile Alterations

The examination of firing rate profile modifications in surrogate data addresses a crucial concern in surrogate methodology, as such alterations represent potentially significant sources of false positives in pattern detection algorithms (Grün, 2009). The preservation of firing rate profiles constitutes a fundamental requirement for maintaining the validity of the null hypothesis, which assumes that observed patterns result from chance coincidences given the underlying firing rate structure rather than from genuine temporal coordination.

We assess this critical criterion using deliberately extreme test scenarios incorporating abrupt rate transitions from 10 Hz to 80 Hz, following established analytical methodologies (Louis et al., 2010) as illustrated in Figure 1.4L. Our systematic analysis reveals consistent smoothing behaviors across the majority of tested methods, with most techniques transforming sharp firing rate transitions into characteristic linear ramps that span the temporal extent of the dither window. These ramps invariably begin 25 ms before the original transition point and conclude 25 ms afterward,

creating a smoothing window that directly corresponds to the dither specification parameter Δ .

Window shuffling demonstrates notably different behavior by introducing secondary transitions through its mechanism of generating locally uniform rates within the 50 ms shuffling boundaries. These secondary transitions create additional rate modulations that are absent in the original data, potentially generating artificial structure that could be misinterpreted as genuine neural coordination patterns. The observed smoothing behavior precisely matches the theoretical prediction of convolving original firing rate profiles with boxcar functions, a relationship that previous studies confirmed analytically (Louis et al., 2010).

1.3.1.8 Comprehensive Statistical Impact Assessment

Feature / Method	UD	UDD	ISI-D	JISI-D	TR-SHIFT	WIN-SHUFF
ISI	no	no	approx.	approx.	yes	approx.
Dead-time	no	yes	yes	yes	yes	no
Auto-correlation	no	no	no	no	yes	approx.
Firing rate modulation	approx.	approx.	approx.	approx.	approx.	approx.
Spike train regularity (CV < 1; regular)	no	no	approx.	approx.	yes	no
Spike train regularity (CV > 1; bursty)	no	no	approx.	approx.	yes	approx.

Table 1.1: **Statistical property preservation across surrogate methods.** Systematic evaluation of how well each surrogate technique preserves key statistical characteristics of the original spike trains. Properties are rated as: **yes** (well preserved), **approx.** (approximately preserved), or **no** (substantially altered). Dead-time preservation is assessed only for PPD processes (which contain explicit dead-time constraints), while other properties are evaluated across all three process types (Poisson, PPD, and Gamma). This assessment guides surrogate method selection based on which statistical properties are most critical for specific experimental analyses.

Our comprehensive investigation demonstrates that surrogate data systematically deviate from original data characteristics to varying degrees across all tested techniques, with these impacts maintaining remarkable consistency across the three data models studied. The systematic evaluation reveals that each surrogate method introduces characteristic distortions that reflect its underlying algorithmic approach

and design constraints, representing trade-offs between different aspects of statistical property preservation.

Table 1.1 consolidates preservation properties across surrogate methods using a categorical framework designating characteristics as maintained ('yes'), approximately maintained ('approx.'), or not maintained ('no'). This systematic assessment provides essential guidance for appropriate surrogate selection based on specific experimental requirements and acceptable compromises between different statistical characteristics.

The evaluation reveals that trial shifting emerges as the optimal choice for preserving most statistical properties, demonstrating particular excellence in maintaining CV values critical for regularity-dependent analyses. This superior performance reflects its fundamental approach of shifting complete spike trains rather than individual spikes, thereby preserving all within-trial temporal relationships while eliminating between-trial correlations. However, this method requires well-defined trial structures that may not be available in all experimental paradigms.

Uniform dithering retains computational efficiency advantages that make it attractive for large-scale analyses while simultaneously compromising important regularity measures, creating a fundamental trade-off between processing speed and statistical fidelity. ISI-based methods demonstrate intermediate performance characteristics that balance property preservation with correlation elimination, though they suffer from reduced repositioning efficacy that may limit their sensitivity in detecting genuine patterns. Window shuffling provides the unique advantage of maintaining spike counts by design, eliminating one potential source of false positives, while uniform dithering with dead-time addresses important biological constraints related to refractory periods.

1.3.2 Binarization Effects on Spike Counts

1.3.2.1 Methodological Context and Importance

Experimental spike data analysis requires temporal binning that converts precise spike times into discrete count sequences within millisecond-scale windows, enabling pattern detection at behaviorally relevant timescales. This discretization process represents a fundamental preprocessing step that bridges the gap between individual spike timing and the temporal resolution required for pattern detection algorithms, following established approaches for temporal discretization of neural data (Perkel et al., 1967b; Grün et al., 2002a). The binning process must balance temporal precision with computational tractability, requiring careful consideration of bin sizes that capture relevant neural dynamics while maintaining sufficient data density for statistical analysis.

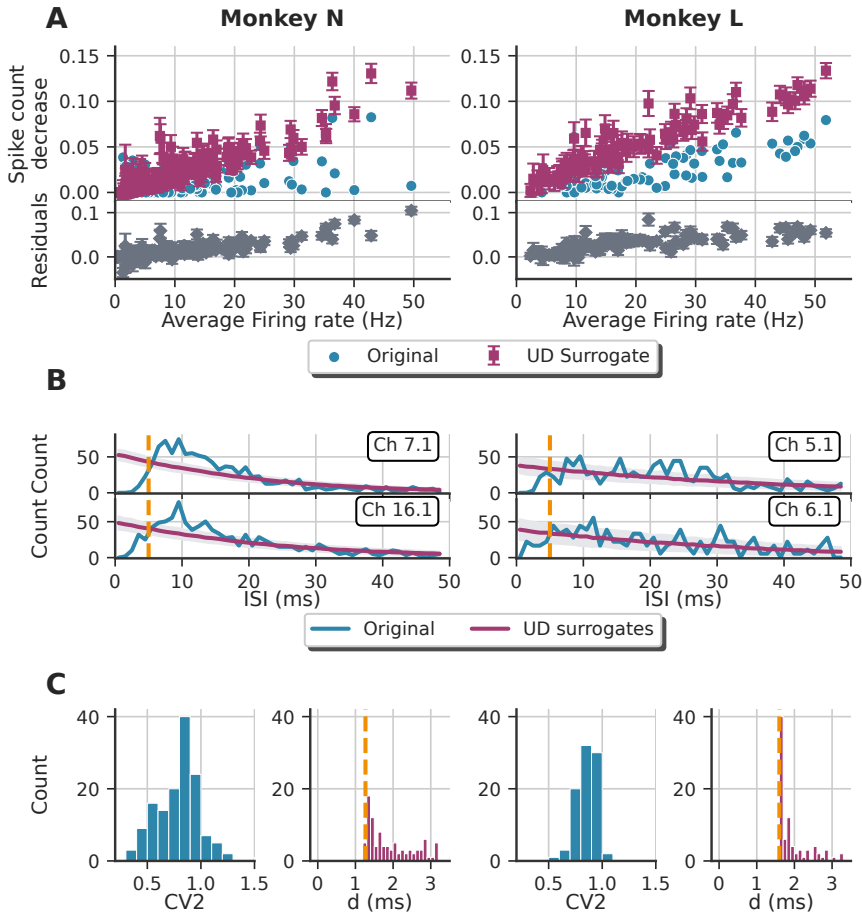


Figure 1.5: Uniform dithering effects on spike counts and temporal structure in experimental data. Analysis of spike count reduction and statistical property changes during uniform dithering and binarization using experimental datasets from monkeys N (left) and L (right) during movement epochs of PGHF trials. **(A) Spike count analysis.** Top: spike count reduction versus mean firing rate for individual neurons. Blue circles show losses from binarization alone; magenta squares show combined losses from uniform dithering plus binarization (mean $\pm$ SD from 100 surrogate realizations). Bottom: difference between original and surrogate spike counts (gray diamonds) using the same normalization. Higher firing rate neurons show greater spike count reduction, particularly with uniform dithering. **(B) Interspike interval distributions.** ISI probability distributions for representative neurons from each monkey (monkey N: channels 7.1 and 16.1; monkey L: channels 5.1 and 6.1). Blue lines show original data; gray traces show uniform dithering surrogates (dark gray: mean; light gray: variability across 500 iterations). Vertical dashed orange lines mark 5 ms bin size. Original data show structured ISI distributions with characteristic peaks, while surrogates exhibit more exponential decay patterns typical of Poisson processes. **(C) Population-level statistics.** Left: CV2 distributions across all recorded neurons, showing original data regularity. Right: minimum ISI (dead time) values for individual neurons. Vertical dashed orange lines indicate spike sorting dead-time thresholds (1.27 ms for monkey N, 1.6 ms for monkey L). Uniform dithering creates ISI values below these biological constraints, violating the dead-time properties of the original data.

Subsequent analysis frequently requires further reduction to binary sequences as outlined for SPADE methodology, where bin contents become 1 for bins containing one or more spikes through a process termed "clipping" and 0 for empty bins. This binarization step transforms continuous spike count data into discrete binary representations that simplify pattern detection algorithms while potentially introducing systematic biases that could affect statistical outcomes. The combination of binning and clipping, collectively referred to as binarization, represents a crucial preprocessing stage that can significantly impact the validity of subsequent analyses.

The importance of understanding binarization effects extends beyond methodological considerations to fundamental questions about pattern detection reliability. Systematic differences in spike counts between original and surrogate data can create artificial significance patterns that masquerade as genuine neural coordination, leading to false positive results that compromise scientific conclusions, a critical concern in neural pattern analysis (Grün, 2009).

1.3.2.2 Systematic Spike Count Reduction

Our systematic examination reveals that specific neurons display substantially reduced total spike counts in uniform dithering surrogate data relative to original recordings, creating systematic disparities that fundamentally compromise the null hypothesis framework. This reduction occurs despite uniform dithering not directly removing spikes from the data, indicating that the interaction between spike repositioning and binarization creates emergent effects that exceed the sum of their individual contributions.

Spike count reduction correlates positively with neuron firing rates, generating systematic disparities between original and surrogate data that become increasingly pronounced as firing rates increase. This rate-dependent relationship suggests that the mechanisms underlying spike count reduction involve the statistical properties of spike timing rather than simple algorithmic artifacts, with higher firing rates creating conditions that exacerbate the effects of temporal repositioning on subsequent binarization outcomes.

Our comprehensive analysis of datasets from two distinct monkeys reveals spike count disparities reaching 10% between uniform dithering surrogates and original data, causing systematic pattern significance overestimation as illustrated in Figure 1.5A. These substantial count differences create fundamental problems for statistical inference because they reduce expected pattern counts in surrogate data relative to original recordings, thereby lowering the statistical threshold for significance and increasing the likelihood of false positive results. The magnitude of these disparities proves particularly concerning given their systematic nature, which creates predictable biases rather than random variations that might be averaged out across analyses.

1.3.2.3 Mechanisms Driving Spike Count Reduction

The investigation of mechanisms driving spike count reduction requires careful examination of how uniform dithering alterations influence interspike interval (ISI) statistics relative to original spike trains. The key insight is that uniform dithering procedures do not directly remove spikes from the data, making spike elimination an emergent property of the interaction between repositioning and binarization rather than a direct algorithmic consequence.

Understanding this phenomenon requires examining the statistical transformation of ISI distributions under uniform dithering and its subsequent impact on the probability of multiple spikes occupying the same temporal bin. The fundamental mechanism involves the conversion of regular ISI patterns into more random distributions that increase the likelihood of brief intervals, thereby creating conditions where multiple spikes are more likely to fall within the same bin and be reduced to a single spike through the clipping process.

Our experimental data analysis reveals characteristic patterns that illuminate this mechanism. Original data ISI distributions consistently display peaks at specific time windows, indicating enhanced regularity that reflects the underlying neural dynamics and refractory properties (Berry and Meister, 1998). These peaks represent preferred ISI values that arise from the interaction between intrinsic neural properties and network dynamics (Softky and Koch, 1993), creating temporal structure that extends beyond simple refractory constraints.

In contrast, surrogate data ISI distributions decay exponentially and populate brief ISI regions more densely than their original counterparts, reflecting the transformation of structured temporal patterns into more random arrangements. This transformation increases the probability of brief intervals because uniform repositioning eliminates the natural temporal spacing constraints that normally prevent spikes from clustering together. When spikes are randomly repositioned without regard to their original ISI structure, the protective mechanisms that maintain appropriate inter-spike distances are disrupted, leading to increased clustering and subsequent spike loss during binning. The result is a redistribution of ISI probabilities that favors shorter intervals and increases the likelihood of multiple spikes occurring within individual bins.

The regularity characteristics of experimental data receive confirmation through CV2 measurements, which quantify local firing regularity by comparing consecutive interspike interval variability while accounting for non-stationary firing rates (Holt et al., 1996). These measurements show distributions concentrating beneath 1.0, indicating firing patterns that are more regular than expected from Poisson processes. Minimum ISI distributions for individual neurons reveal dead-times of 1.27 ms for monkey N and 1.6 ms for monkey L, values that correspond to constraints imposed by

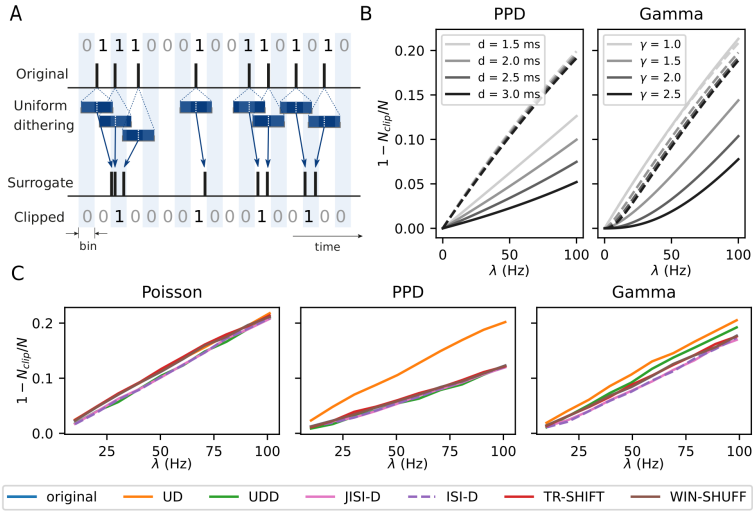


Figure 1.6: Spike count reduction mechanisms during binarization. (A) Illustration of binarization effects. Schematic showing how uniform dithering affects spike counts during temporal discretization. Top: original regularly-spaced spike train maintains consistent interspike intervals. Middle: uniform dithering redistributes spike times, causing multiple spikes to fall within the same temporal bins (5 ms). Bottom: binarization (clipping) reduces multiple spikes per bin to single events, resulting in spike count loss. The regular firing pattern (top) experiences less spike loss than the dithered version (bottom) due to its temporal structure. **(B) Theoretical predictions of spike count reduction.** Spike count reduction ($1 - N_{\text{clip}}/N$, where N_{clip} = clipped counts, N = original counts) following 5 ms binning for two process types. Left panel: PPD processes with varying dead-times ($d = 1.5, 2.0, 2.5, 3.0$ ms, grayscale). Right panel: Gamma processes with different shape factors ($\gamma = 1, 1.5, 2, 2.5$, grayscale). Solid curves show original processes; dashed curves show uniform dithering surrogates. Higher firing rates increase spike loss, with UD surrogates showing greater attenuation than original processes. **(C) Empirical validation using synthetic data.** Spike count reduction measured from synthetic spike trains across six surrogate methods (color-coded) for firing rates from 10-100 Hz (10 Hz increments). Each spike train contained 10,000 expected spikes regardless of rate. Results confirm theoretical predictions, with uniform dithering showing the highest spike count reduction, particularly at elevated firing rates.

spike sorting algorithms for resolving overlapping spikes in the experimental setup (Brochier et al., 2018).

Importantly, surrogate data ISI distributions exhibit time windows briefer than these experimental minima, confirming that dithering procedures produce briefer ISIs in uniform dithering surrogates than in original data as shown in Figure 1.5B,C. This violation of biological constraints represents a fundamental flaw in the null hypothesis generated by uniform dithering, as it creates spike patterns that could not occur in the original data due to physical limitations of the recording process.

1.3.2.4 Model-Specific Spike Count Behaviors

Our comprehensive spike count reduction analysis across all surrogate types applied to three distinct spike train models reveals that elevated firing rates consistently generate enhanced spike elimination, though the extent of reduction varies significantly among data models and surrogate methods. This systematic variation provides insights into the mechanisms underlying spike count reduction and their dependence on the statistical properties of the underlying spike generation processes.

The analysis of Poisson processes exhibits spike count elimination that increases with firing rates across all surrogate methods, reaching elimination rates of 20% at maximum firing rates of 100 Hz. This relationship reflects the fundamental statistical properties of Poisson processes, where the probability of multiple spikes occurring within a given time window increases with firing rate. Notably, only JISI-D, ISI-D, and UDD exhibit marginally reduced overall elimination relative to other techniques, reflecting their design constraints that limit repositioning to intervals between adjacent spikes.

The analysis of PPD processes reveals considerably reduced spike elimination relative to alternative models, with the notable exception of uniform dithering which displays enhanced elimination despite the dead-time constraints. This exception reflects the ISI distribution observations where uniform dithering relocates many spikes into time windows up to 5 ms in duration, precisely matching the bin widths used in our analysis and creating artificial clustering that violates the dead-time constraints of the original process.

The reduced elimination in PPD processes occurs because strict dead-time constraints prevent the formation of brief ISI intervals that would lead to multiple spikes within single bins. When dead-time duration approaches bin sizes, the occurrence of two spikes dithering into identical bins becomes increasingly improbable, thereby reducing spike elimination during the binarization process. This protective effect demonstrates the importance of preserving biological constraints during surrogate generation.

Gamma process analysis reveals spike elimination levels that surpass PPD levels but remain substantially below Poisson levels, reflecting the intermediate regularity characteristics of these processes. Uniform dithering exhibits maximum elimination whereas JISI-D and ISI-D show minimal elimination, consistent with their design emphasis on preserving ISI structure. Window shuffling maintains original spike counts by design, eliminating this source of bias entirely, whereas trial shifting displays negligible count differences as illustrated in Figure 1.6C.

1.3.2.5 Significance Evaluation Consequences

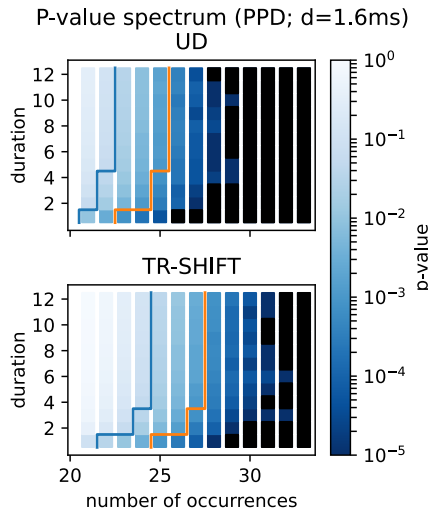


Figure 1.7: **Binarization-induced bias in pattern significance testing.** Comparison of p-value distributions for triplet spike patterns using uniform dithering (UD, upper panel) versus trial shifting (TR-SHIFT, lower panel) as surrogate methods. Heat maps show significance levels across pattern occurrence counts (x-axis) and temporal durations in bins (y-axis), with 5 ms bin size. Color scale ranges from dark blue (highly significant, low p-values) to light blue (not significant, high p-values). Data represent 20 independent analyses of stationary PPD spike trains ($n = 20$ parallel trains, 2-second duration, 60 Hz firing rate, 1.6 ms dead-time). Surrogate distributions were generated using 5000 iterations with $\Delta = 25$ ms dither width; TR-SHIFT additionally used 100 ms trial segments. Blue horizontal lines mark the $\alpha = 0.05$ significance threshold; orange lines show multiple testing corrected thresholds. Patterns to the right of these lines are considered significant. **Key finding:** UD requires fewer pattern occurrences for significance compared to TR-SHIFT (note leftward shift of significance boundaries), demonstrating systematic bias due to spike count reduction during binarization. This bias leads to inflated significance levels and increased false positive rates when using UD with regular firing patterns.

Having established the occurrence and mechanisms of spike count elimination through uniform dithering binarization, we now demonstrate how this elimination system-

atically reduces expected pattern counts in surrogate data, thereby producing false positive results that compromise the statistical validity of pattern detection analyses.

Our analysis employs 20 repetitions of p-value spectrum generation using datasets with 20 independent, stationary PPD spike trains, averaging results across uniform dithering and trial shifting comparisons to eliminate random variation and focus on systematic effects. The p-value spectrum analysis reveals characteristic transitions from elevated p-values at low occurrence counts to reduced p-values at higher occurrences, reflecting the fundamental relationship between pattern frequency and statistical significance in the SPADE framework.

The critical finding is that trial shifting requires approximately two additional pattern occurrences relative to uniform dithering for equivalent significance classification, representing a substantial shift in the statistical threshold that affects the interpretation of experimental results. For patterns with 4-bin durations, uniform dithering requires only 25 occurrences for significance classification versus 28 for trial shifting, a difference that translates into increased false positive rates for experimental analyses employing uniform dithering.

This systematic difference arises because surrogate-induced spike elimination reduces expected pattern counts in surrogate data relative to original recordings, thereby lowering the statistical threshold for significance and increasing the likelihood that chance patterns will be classified as significant. Our analysis of the remaining four surrogate methods demonstrates behavior that closely resembles trial shifting, suggesting that uniform dithering represents a unique case of systematic bias rather than a general property of surrogate methods.

The implications extend beyond methodological considerations to fundamental questions about the reliability of published results that employed uniform dithering for statistical validation. These findings lead to the conclusion that uniform dithering inadequately serves spike data featuring hard dead-times or regular spiking behaviors characteristic of motor cortex recordings (Mochizuki et al., 2016). The systematic bias introduced by spike count reduction renders uniform dithering inappropriate for many experimental applications, necessitating the adoption of alternative methods that preserve spike counts or the development of correction procedures that account for systematic biases.

1.3.3 Surrogate Technique Performance on Non-Stationary Data

1.3.3.1 Experimental Framework and Motivation

Building upon our comprehensive analysis of stationary synthetic data, we now advance to examine the surrogate method performance under the more realistic conditions of non-stationary neural activity. This extension addresses a fundamental gap

between controlled theoretical evaluation and the complex temporal dynamics characteristic of actual experimental recordings, where firing rates fluctuate dramatically across behavioral epochs and experimental conditions.

The experimental datasets underlying our synthetic data generation comprise two comprehensive sets of recordings from approximately 100 parallel spike trains obtained from macaque monkey motor and pre-motor cortex during performance of a reach-to-grasp task (Riehle et al., 2013). These recordings capture the rich temporal dynamics of cortical activity during skilled motor behavior, including periods of low baseline activity during start and cue epochs, rapid rate increases during movement preparation in delay periods, and sustained elevated activity during movement execution. The behavioral paradigm provides a naturalistic context for evaluating surrogate performance across diverse neural states and temporal dynamics.

We apply our six surrogate methods to synthetic datasets produced from these two experimental sessions to systematically study surrogate method impacts on false positive occurrence rates under non-stationary conditions. This approach allows us to maintain controlled conditions for definitive false positive assessment while incorporating the temporal complexity that defines real neural data. The analysis presented here employs updated analytical code that offers more accurate pattern detection compared to earlier implementations, resulting in revised false positive estimates relative to those reported in Stella et al. (2022).

We employ two distinct non-stationary point processes to model the diverse characteristics observed in experimental data. PPD models (Section 1.2.4.3) accurately replicate dead-time effects that arise from spike sorting artifacts and genuine refractory periods, while Gamma processes enable precise modeling of CV2 characteristics that capture the distribution of firing regularity observed across the recorded population. This dual approach ensures comprehensive coverage of the statistical properties that influence surrogate method performance under realistic conditions.

Using these modeling frameworks, our approach incorporates firing rate modifications by producing synthetic spike trains that precisely match the firing rate profiles of experimental data while maintaining controlled null hypothesis conditions. This methodology enables systematic evaluation of surrogate performance under realistic temporal dynamics while preserving the ability to definitively identify false positive results. The generated data successfully replicate both individual neuron and population response properties from experimental recordings, ensuring that our synthetic evaluation reflects the statistical challenges encountered in real experimental analyses.

1.3.3.2 Synthetic Data Validation and Characteristics

Before presenting our analysis of surrogate performance on non-stationary data, we conduct thorough validation to confirm that our synthetic datasets accurately capture

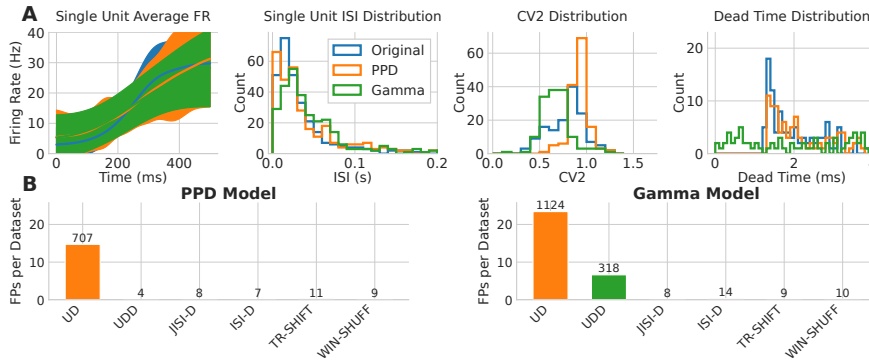


Figure 1.8: Synthetic data validation and false positive assessment for surrogate methods. (A) Synthetic data validation. Comparison of synthetic data models with experimental recordings to verify model fidelity. Data are shown for experimental data (blue), PPD synthetic data (orange), and Gamma process synthetic data (green). From left to right: firing rate profile for a representative neuron (session i140703-001, movement epoch, PGHF condition, electrode 21, unit 1) showing smoothed average firing rate computed across 500 ms epochs and multiple trials; interspike interval (ISI) distribution for the same neuron; coefficient of variation (CV2) values averaged across all neurons in the recording session, with lower values indicating more regular firing; minimum ISI measurements representing dead-time estimates for all neurons in the dataset. **(B) False positive rates across surrogate methods.** Number of false positive patterns detected by each surrogate method (color-coded along x-axis) averaged across 48 datasets (2 sessions $\times$ 6 epochs $\times$ 4 conditions). Left panel shows results for PPD synthetic data; right panel shows Gamma synthetic data results. Numbers above bars indicate total false positives detected across all 48 datasets. Most methods show approximately one false positive per dataset, while uniform dithering (UD) shows substantially higher false positive rates, confirming its systematic bias identified in controlled analyses.

the essential characteristics of experimental recordings. This validation proves crucial for ensuring that our conclusions regarding surrogate performance will generalize to real experimental applications, where similar temporal dynamics and statistical properties determine method effectiveness.

We focus our validation on movement epoch datasets, where neurons consistently exhibit high firing rates across all trial types. These epochs provide the most stringent test of synthetic data fidelity due to their maximum firing rate modifications. Figure 1.8A presents comprehensive comparisons of characteristics across experimental (blue), PPD (orange), and Gamma process (green) data, revealing both the successes and limitations of our synthetic modeling approach.

Our synthetic data demonstrate excellent preservation of firing rate dynamics, closely matching the temporal patterns of experimental recordings. This successful replication occurs precisely as expected given our construction methodology, which explicitly targets rate profile matching through thinning algorithms and continuous rate variation simulation. The preserved rate dynamics ensure that our synthetic evaluation captures the essential temporal structure influencing surrogate performance, particularly the rapid rate transitions that challenge surrogate methods most severely.

Individual unit ISI properties show strong correspondence between synthetic and experimental data, with systematic differences that align with our modeling design choices. Both PPD and Gamma processes incorporate brief ISIs absent in experimental recordings due to dead-time constraints and spike sorting limitations. This difference reflects our deliberate modeling approach: while PPD models explicitly incorporate dead-time effects, Gamma processes prioritize CV2 matching over absolute refractory period enforcement. The overall ISI structure preservation validates our modeling framework while highlighting the specific trade-offs inherent in each approach.

CV2 distributions demonstrate that our synthetic data successfully approximate experimental firing regularity patterns, though with expected compression of the extreme ranges. The synthetic data approach experimental means despite broader experimental distributions that include neurons with particularly low CV2 values. This compression occurs as anticipated given our modeling constraints, which cannot capture the full diversity of firing patterns in the experimental population. Nevertheless, our synthetic data span the CV2 range most critical for surrogate evaluation, providing comprehensive coverage for analytical purposes. The use of CV2 proves particularly appropriate here, as it provides local firing regularity measures that remain valid under non-stationary conditions, unlike traditional coefficients of variation that assume rate stationarity (Holt et al., 1996).

Dead-time characteristics exhibit the expected high fidelity between experimental and PPD data, with both showing characteristic 1.6 ms dead-times reflecting spike sorting constraints. This precise matching validates our PPD implementation and

confirms successful replication of refractory period effects. In contrast, Gamma processes permit near-zero inter-spike intervals by design, highlighting the fundamental difference between modeling approaches and their respective optimization targets.

These validation results confirm that our synthetic data successfully capture the essential experimental characteristics while maintaining the controlled conditions necessary for definitive false positive assessment. The observed matches and systematic differences align precisely with our construction methodology: rate profiles are preserved by design, ISI and CV2 properties reflect our modeling priorities, and dead-time effects demonstrate the specific strengths of each approach. This systematic validation establishes that our synthetic evaluation framework provides meaningful insights into surrogate performance under realistic experimental conditions.

1.3.3.3 False Positive Analysis Using Non-Stationary Conditions

Our systematic analysis of false positive occurrence under non-stationary conditions employs synthetic data generated from independent processes by design, ensuring that all observed spike patterns represent chance occurrences qualifying as definitive false positives. This controlled framework offers unambiguous criteria for evaluating surrogate method performance, eliminating the interpretive ambiguities that complicate experimental data analysis.

We conduct comprehensive SPADE analysis using all datasets separately, with each surrogate method applied using standardized parameters that ensure fair comparison across methods. These parameters include 5-ms bin sizes that balance temporal precision with adequate spike counts, 60-ms maximum pattern durations that span the range of biologically relevant coordination timescales, $\alpha = 0.05$ significance levels that match conventional statistical criteria, 5000 surrogate iterations that provide robust statistical estimation, and $\Delta = 25$ ms dither specifications that represent typical values used in experimental applications.

Overall False Positive Rates The results reveal dramatic differentiation between uniform dithering and all other surrogate methods, with uniform dithering producing substantially more false positives than any alternative approach. Specifically, uniform dithering generates 707 false positives for PPD processes and 1124 for Gamma processes across 96 datasets, representing error rates that far exceed acceptable levels for scientific analysis. These elevated false positive rates confirm our predictions based on spike count reduction analysis, revealing that the systematic biases observed in controlled conditions translate into severe practical problems under realistic experimental conditions.

In stark contrast, alternative surrogate methods generate substantially fewer false positives. UDD produces 4 false positives for PPD processes and 318 for Gamma

processes, representing a dramatic improvement over uniform dithering while still showing elevated rates for Gamma data. The most conservative methods exhibit excellent performance, with JISI-D producing 8 false positives for both PPD and Gamma processes, ISI-D generating 7 false positives for PPD and 14 for Gamma processes, and TR-SHIFT producing 11 false positives for PPD and 9 for Gamma processes. WIN-SHUFF maintains consistent performance with 9 false positives for PPD and 10 for Gamma processes.

Family-Wise Error Rate Analysis in the Context of SPADE's Built-in FDR Control

The analysis of family-wise error rates (FWER) provides crucial insights into the practical implications of different surrogate methods for experimental applications, with particular significance given that SPADE already incorporates Benjamini-Hochberg FDR correction as an integral component of its statistical testing framework. It is essential to recognize that all false positive rates reported here represent patterns that have already undergone FDR correction within the SPADE algorithm, making the observed systematic biases in uniform dithering particularly concerning from a methodological perspective.

SPADE's statistical testing procedure applies the Benjamini-Hochberg method to control false discovery rates across all pattern signatures $\{z, c, d\}$ identified in each dataset, with the expectation that this correction will maintain false positive rates at acceptable levels regardless of the number of patterns tested. The fact that uniform dithering produces false positives at rates approaching or exceeding 90% of datasets (FWER: 0.917 for PPD, 0.958 for Gamma) despite this built-in FDR control reveals that the systematic bias introduced by uniform dithering is so severe that it overwhelms the multiple comparison correction mechanism. This extraordinarily high family-wise error rate confirms that uniform dithering should be completely avoided for realistic neural data analysis, as it fails to provide valid statistical control even when subjected to appropriate multiple comparison correction procedures.

The fundamental problem with uniform dithering lies in its violation of the basic assumptions underlying the Benjamini-Hochberg procedure. FDR control methods assume that the null hypothesis testing framework provides valid p-values as input, with the expectation that under true null conditions, p-values will be uniformly distributed. The systematic bias in uniform dithering distorts this p-value distribution, leading to an excess of small p-values that are incorrectly interpreted as significant patterns even after FDR correction. This represents a fundamental failure of the surrogate method rather than a limitation of the multiple comparison correction approach.

The conservative methods – including JISI-D, ISI-D, TR-SHIFT, and WIN-SHUFF – exhibit consistently reliable performance across both process types, establishing the valid statistical foundation necessary for SPADE's built-in FDR control to function effectively. JISI-D shows identical performance across process types (8 false positives

each, FWER 0.125), ensuring robust statistical control that enables the Benjamini-Hochberg procedure to function as intended. ISI-D shows slightly elevated false positive rates for Gamma processes (7 vs. 14 false positives, FWER 0.125 vs. 0.188) but maintains acceptable error control within bounds that preserve FDR correction validity. TR-SHIFT achieves excellent and consistent performance (11 vs. 9 false positives, FWER 0.167 vs. 0.104), while WIN-SHUFF shows stable performance across process types (9 vs. 10 false positives, FWER 0.167 vs. 0.146).

The total false positive counts for conservative methods range from 7-14 patterns across 96 datasets, representing error rates consistent with chance expectations given our significance criteria ($\alpha = 0.05$). This confirms that SPADE's built-in Benjamini-Hochberg correction functions appropriately when provided with valid surrogate data. These methods demonstrate that proper surrogate design enables SPADE's statistical testing framework to achieve reliable false discovery rate control regardless of the underlying regularity characteristics of neural data. The observed FWER values for conservative methods (0.104-0.188) confirm that these approaches provide the statistical validity necessary for SPADE's multiple comparison correction procedures to function effectively.

1.3.3.4 Method-Specific Performance Analysis and Underlying Mechanisms

Building upon the overall false positive assessment, we now examine the specific performance characteristics of individual surrogate methods to understand the mechanistic basis for their differential effectiveness. This analysis reveals fundamental insights about surrogate design principles and their relationship to different types of neural data characteristics.

Uniform Dithering: Systematic Failure Across Conditions Uniform dithering represents a uniquely problematic approach that systematically violates statistical assumptions, producing false positives in over 90% of analyzed datasets. The consistent failure across both PPD and Gamma data types reveals that the fundamental flaws in uniform dithering transcend specific data characteristics and reflect core algorithmic limitations that cannot be addressed through parameter adjustment or selective application.

UDD: Process-Dependent Performance and Design Limitations UDD exhibits process-dependent performance that reveals important insights about surrogate method design. For PPD data, UDD achieves excellent false positive control with only 4 false positives across 96 datasets and an FWER of 0.083, suggesting that preserving dead-time constraints offers effective bias control when analyzing data with explicit refractory periods. However, for Gamma data, UDD shows substantially elevated false

positive rates (318 false positives, FWER 0.583), indicating systematic bias when applied to neurons with intermediate regularity characteristics that lack explicit dead-time constraints but possess other forms of temporal structure.

This dramatic performance difference reveals a fundamental limitation of the UDD approach: it provides effective bias control only when applied to spike trains that genuinely possess hard dead-time constraints, as in the PPD model. For Gamma processes, which achieve regularity through shape parameter modulation rather than absolute refractory periods, the dead-time preservation mechanism of UDD becomes irrelevant and fails to address the underlying statistical biases that affect uniform dithering. The method's inability to handle non-dead-time-based regularity explains why UDD performance on Gamma data remains much closer to standard uniform dithering than to truly conservative methods.

Conservative Methods: Robust Performance Across Data Types The conservative methods establish that proper surrogate design can yield valid statistical frameworks for pattern detection, with false positive rates that remain within expected bounds given standard significance criteria. The modest variations in performance across process types for some methods (ISI-D, TR-SHIFT) remain within acceptable limits while highlighting the importance of method selection based on dataset characteristics. The consistency of performance across diverse data types suggests that these methods achieve effective bias control through comprehensive preservation of statistical properties rather than reliance on specific features that may not be universally relevant.

1.3.3.5 Neuronal Characteristics Associated with False Positive Susceptibility

Having established the differential performance of surrogate methods across synthetic data types, we now examine the specific neuronal characteristics that predispose individual units to false positive pattern detection. This analysis provides practical guidance for identifying neurons and experimental conditions that require particular attention to surrogate method selection.

Firing Regularity Characteristics Analysis of the neuronal properties contributing to false positive patterns yields critical insights into conditions that predispose neurons to systematic bias in pattern detection algorithms. The examination of firing regularity characteristics shows that neurons involved in false positive patterns for PPD data exhibit CV2 values mostly in the range of 0.8 to 1.0, while those in Gamma data show CV2 values primarily in the range of 0.5 to 0.8. The use of CV2 rather than traditional coefficient of variation is essential for non-stationary data analysis, as CV2 offers a local measure of spike train regularity that remains valid under time-varying

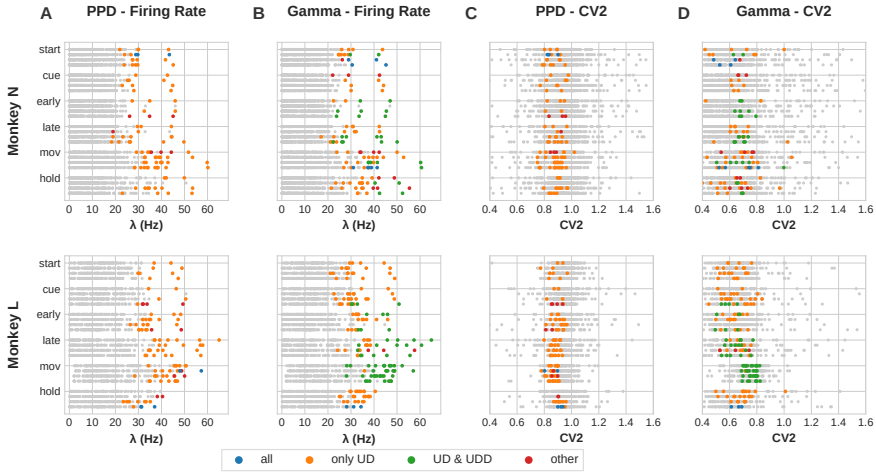


Figure 1.9: Neuronal characteristics associated with false positive pattern detection. Analysis of firing properties for neurons contributing to false positive patterns in synthetic data across different surrogate methods. **(A) PPD firing rate analysis.** Distribution of mean firing rates for neurons involved in false positive patterns compared to the complete neuronal population in PPD synthetic data. **(B) Gamma firing rate analysis.** Distribution of mean firing rates for neurons in Gamma synthetic data. **(C) PPD regularity analysis.** Coefficient of variation (CV2) distributions for neurons in PPD synthetic data. **(D) Gamma regularity analysis.** CV2 distributions for neurons in Gamma synthetic data. Data are organized by monkey (N: upper row, L: lower row) and behavioral epoch (vertical arrangement, inverted axis). Within each epoch, experimental conditions are arranged horizontally as PGHF, PGLF, SGHF, SGLF. Neurons are color-coded by their pattern of false positive involvement: blue circles indicate neurons in false positives detected by all surrogate methods (highest visibility); red circles indicate neurons contributing to false positives in other surrogate combinations; green circles represent neurons involved in both UD and UDD false positives only; orange circles show neurons specific to uniform dithering (UD) false positives. Gray circles show firing rates for neurons not involved in any false positive patterns. Lower CV2 values indicate more regular firing patterns, while higher values indicate more irregular firing.

rate conditions (Holt et al., 1996). This local measurement approach enables accurate characterization of firing regularity even when neurons exhibit dramatic rate changes across behavioral epochs, as commonly observed in motor cortical recordings.

Importantly, these CV2 ranges fall within normal physiological parameters observed in motor cortical neurons, indicating that false positive susceptibility does not depend on aberrant firing regularity but rather reflects systematic algorithmic bias affecting neurons with typical firing characteristics. The concentration of false positive patterns within these CV2 ranges suggests that the biases introduced by inappropriate surrogate methods are most pronounced for neurons with intermediate regularity characteristics, which represent a substantial proportion of the motor cortical population.

Firing Rate Dependencies and Motor Cortical Context More striking than regularity characteristics is the pronounced concentration of false positive patterns in neurons with elevated firing rates, typically ranging from 30 to 70 Hz. This firing rate range is particularly significant in the context of reach-to-grasp behaviors in macaque motor and pre-motor cortex, where such elevated rates commonly occur during active movement execution and represent peak activity levels associated with motor command generation and control (Churchland et al., 2012). These high firing rates reflect the intense neural activity required for precise motor coordination during skilled grasping movements, making them prevalent in experimental recordings from motor cortical areas.

The concentration of false positive patterns specifically in neurons operating at these peak physiological firing rates has important implications for experimental interpretation. Since elevated firing rates represent the conditions most relevant for understanding motor cortical function during active behavior, the systematic bias affecting high-rate neurons could lead to substantial overestimation of coordination patterns precisely when they are most behaviorally relevant. This finding underscores the critical importance of appropriate surrogate method selection for studies of motor cortical coordination during active movement phases.

These results align closely with our expectations based on spike count reduction analysis, confirming that the systematic biases observed in controlled stationary conditions translate predictably into false positive behavior under more realistic non-stationary conditions. The dramatic differences in false positive rates between uniform dithering and other methods reveal that the problems with uniform dithering represent specific methodological flaws rather than inherent limitations of surrogate approaches. The consistency of conservative performance across JISI-D, ISI-D, TR-SHIFT, and WIN-SHUFF establishes confidence in properly implemented surrogate techniques when appropriate statistical controls are maintained. Critically, these methods enable SPADE's built-in Benjamini-Hochberg FDR correction to function as intended, establishing the valid statistical foundation necessary for reliable pattern detection in high-dimensional neural data analysis.

1.3.4 Experimental Data Applications

1.3.4.1 Real-World Performance Assessment Framework

Our final and most crucial analysis applies SPADE with all six surrogate methods to experimental data sessions, examining significance testing differences that result from surrogate method applications in the absence of known ground truth regarding pattern presence and occurrence counts. This analysis represents the ultimate test of surrogate method performance in realistic neuroscience applications, where the complex interactions between neural dynamics, recording procedures, and analytical methods determine the validity of scientific conclusions.

The experimental data analysis differs fundamentally from our synthetic data studies because it lacks definitive criteria for distinguishing genuine neural coordination from chance coincidences. This absence of ground truth creates interpretive challenges that require careful consideration of converging evidence from multiple analytical approaches. However, it also provides the most realistic assessment of method performance under the conditions actually encountered in experimental neuroscience, where researchers must make inferences about neural coordination based on statistical criteria rather than known truth.

We present results displaying significant pattern counts for each behavioral epoch and trial-type, with data presented separately for each monkey due to substantial differences in neural characteristics that may influence surrogate method performance. This separation allows examination of how individual differences in neural characteristics might influence pattern detection across methods, providing insights into the extent to which conclusions about neural coordination may depend on both the analytical approach and the specific characteristics of the recorded population.

1.3.4.2 Overall Pattern Detection Results and Method Comparisons

The comprehensive analysis of pattern counts across all experimental conditions reveals dramatic method-dependent differences that confirm our predictions based on synthetic data analysis while providing new insights into the magnitude of these effects under realistic experimental conditions (Figure 2.10). The results demonstrate that surrogate method choice profoundly influences conclusions about neural coordination, with implications that extend far beyond methodological considerations to fundamental questions about the nature of cortical processing.

Quantitative Summary of Pattern Detection For monkey N (session i140703-001), uniform dithering detected 539 total patterns with dramatic elevation across all

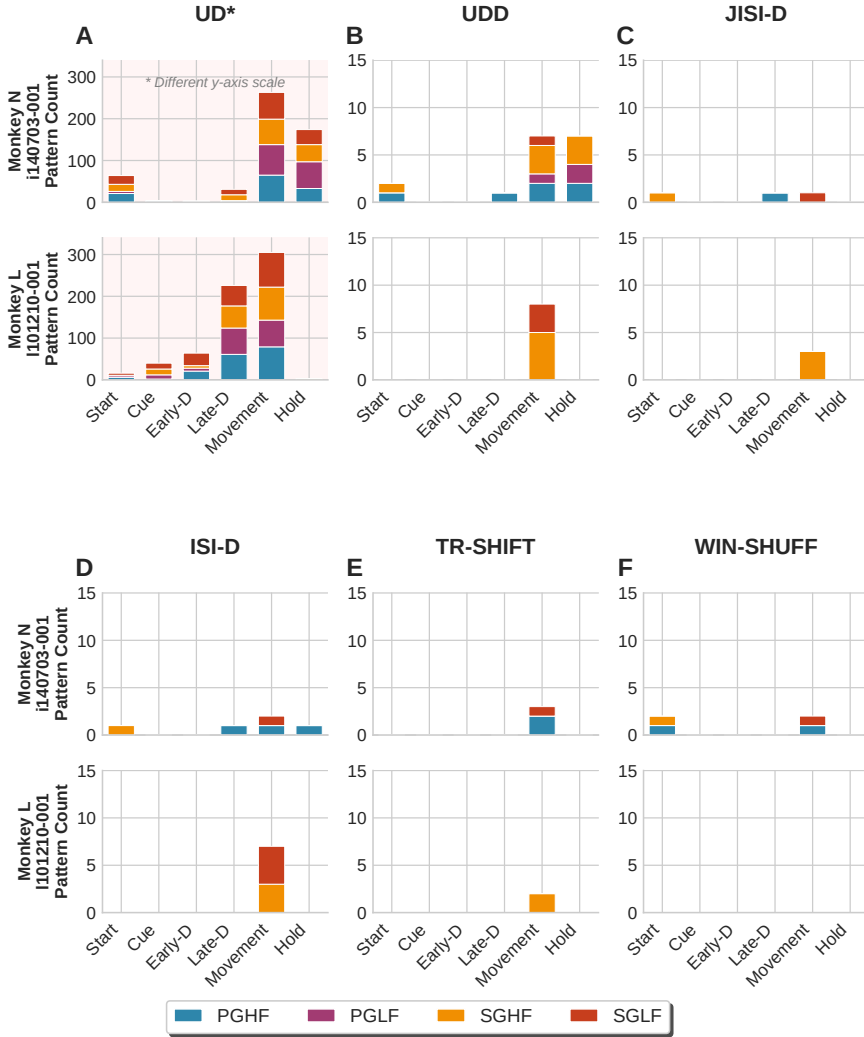


Figure 1.10: Pattern detection results from experimental data analysis. SPADE analysis results for two experimental sessions: monkey N (session i140703-001, top row) and monkey L (session i101210-001, bottom row). Bar plots show the number of statistically significant patterns detected in each behavioral epoch (start, cue, early-delay, late-delay, movement, and hold) across different experimental conditions. Colors indicate grip type and force combinations: precision grip with low force (PGLF), precision grip with high force (PGHF), side grip with low force (SGLF), and side grip with high force (SGHF). Each column represents results from a different surrogate method. Note that uniform dithering (UD, first column) uses a different y-axis scale due to substantially higher pattern counts compared to other methods.

epochs, while the conservative methods showed substantially lower counts. UDD detected 17 total patterns concentrated in start, late-delay, movement, and hold epochs. The highly conservative JISI-D method detected only 3 patterns across start, late-delay, and movement epochs, while ISI-D found 5 patterns distributed across start, late-delay, movement, and hold epochs. TR-SHIFT proved most conservative with only 3 patterns detected exclusively during movement epochs, and WIN-SHUFF identified 4 patterns across start and movement epochs.

For monkey L (session I101210-001), uniform dithering again showed dramatically elevated detection with 654 total patterns across most epochs. UDD detected 8 patterns concentrated solely in movement epochs, while JISI-D found 3 patterns highly localized to movement epochs only. ISI-D detected 7 patterns exclusively during movement epochs, TR-SHIFT maintained its conservative profile with only 2 patterns during movement, and WIN-SHUFF detected no significant patterns whatsoever in this animal.

False Discovery Rate Considerations The dramatic elevation of UD pattern counts relative to other methods raises critical questions about false discovery rates and the appropriateness of different null hypotheses. With significance thresholds of $\alpha = 0.05$, we would expect approximately 5% of tested pattern signatures to achieve significance by chance alone. The fact that UD detects 180-327 times more patterns than the most conservative methods (TR-SHIFT and JISI-D) strongly suggests systematic bias rather than enhanced sensitivity to genuine patterns.

Our synthetic data analysis demonstrated that UD produces false positive rates of 707-1124 patterns across 96 datasets (approximately 7.4-11.7 patterns per dataset), while other conservative methods produced 7-14 false positives total (approximately 0.07-0.15 patterns per dataset). The experimental data show similar relative magnitudes, with UD detecting approximately 22-27 patterns per dataset ($539/24 \approx 22$ for monkey N, $654/24 \approx 27$ for monkey L) while conservative methods detect 0.08-0.29 patterns per dataset. This concordance between synthetic and experimental results provides strong evidence that the vast majority of UD-detected patterns represent false positives.

1.3.4.3 Temporal and Behavioral Specificity of Pattern Detection

Epoch-Specific Analysis for Conservative Methods Excluding UD due to its demonstrated systematic bias, the analysis of conservative surrogate methods reveals extremely sparse pattern detection across experimental conditions. The small number of patterns detected by each method limits the scope for detailed interpretation, but several basic observations can be made regarding their temporal and behavioral distribution.

Monkey N Coordination Patterns: Conservative methods detected between 3 and 5 patterns total, with the majority occurring during movement epochs. JISI-D identified 3 patterns distributed across start (SGHF), late-delay (PGHF), and movement (SGLF) epochs, while ISI-D detected 5 patterns spanning start (SGHF), late-delay (PGHF), movement (PGHF, SGLF), and hold (PGHF) epochs. TR-SHIFT detected 3 patterns exclusively during movement, specifically 2 PGHF and 1 SGLF patterns. UDD showed broader temporal distribution with patterns occurring across start, late-delay, movement, and hold epochs, though the total count remained relatively low at 17 patterns. WIN-SHUFF detected 4 patterns distributed across start and movement epochs. While the concentration during movement epochs appears notable, the small sample sizes preclude drawing strong conclusions about condition-specific coordination requirements.

Monkey L Coordination Patterns: Monkey L demonstrated even sparser pattern detection, with conservative methods detecting between 0 and 8 patterns total, concentrated exclusively in movement epochs. JISI-D detected 3 patterns during movement, all occurring in SGHF conditions, while ISI-D detected 7 patterns during movement comprising 3 SGHF and 4 SGLF patterns. TR-SHIFT detected 2 patterns during movement, both occurring in SGHF conditions. UDD showed patterns only in movement epochs, specifically 5 SGHF and 3 SGLF patterns. WIN-SHUFF detected no patterns whatsoever in monkey L, creating a striking contrast with the few patterns detected in monkey N.

Behavioral Context Dependencies Given the extremely small number of patterns detected by conservative methods (2-8 per method per monkey), detailed analysis of behavioral dependencies would be inappropriate and potentially misleading. The apparent concentration of patterns during movement epochs appears consistent across both animals, suggesting this may represent the most detectable coordination activity, though the sparse detection overall indicates either highly selective coordination mechanisms or limitations in current detection sensitivity.

1.3.4.4 Critical Evaluation of Null Hypothesis Assumptions

Strictness of the Independence Assumption The low pattern counts detected by conservative surrogate methods raise important questions about whether the null hypothesis of spike train independence given firing rate modulations may be overly restrictive for detecting biologically relevant coordination patterns. Our analysis reveals that significant patterns occur almost exclusively during high firing rate epochs (movement) and primarily in neurons with elevated baseline firing rates, suggesting that current methods may have limited sensitivity for detecting coordination at moderate firing rates.

This limitation has several potential explanations. First, the temporal precision required for pattern detection (5 ms bins) may be too strict for coordination mechanisms operating at slightly broader timescales. Second, the minimum occurrence thresholds (10 occurrences across trials) may be too stringent for detecting coordination patterns that occur less frequently but remain biologically meaningful. Third, the null hypothesis assumption of complete independence may not adequately account for loose temporal relationships that preserve some coordination while disrupting precise synchrony.

Firing Rate Dependencies and Detection Sensitivity Our experimental results confirm the synthetic data prediction that pattern detection shows strong dependence on firing rate characteristics. The concentration of patterns in movement epochs, which exhibit the highest firing rates, and the rarity of patterns in low-firing-rate epochs like cue and early-delay, suggests that current methods may be biased toward detecting coordination only under high-activity conditions.

This bias has important implications for understanding cortical processing, as it may lead to underestimation of coordination during cognitive phases of behavior that involve more moderate firing rate changes. The development of methods with enhanced sensitivity to moderate-rate coordination represents an important direction for future methodological development.

1.3.4.5 Comparison with Literature and Methodological Implications

Consistency with Previous Findings Our results align with previous literature demonstrating that surrogate method choice critically influences pattern detection outcomes. Studies by Louis et al. (2010) reported concerns about uniform dithering producing false positives in regular firing neurons, consistent with our finding that UD dramatically overestimates pattern significance with family-wise error rates exceeding 90% in synthetic data. The concentration of patterns in high-firing-rate neurons and movement epochs also aligns with previous reports of motor cortical coordination during active movement execution.

Our findings raise important questions about the interpretation of previous studies that employed uniform dithering, such as Torre et al. (2016), though their use of different pattern detection algorithms (CoCoNAD rather than FIM-based approaches) may have provided some protection against the systematic biases we identify. The substantial differences in false positive rates between uniform dithering and conservative methods suggest that some coordination patterns reported in earlier studies may warrant reexamination using more rigorous surrogate approaches.

However, our finding of very low pattern counts with conservative methods contrasts with some previous studies that reported more extensive coordination patterns. This

discrepancy may reflect differences in analytical parameters, recording techniques, or the specific motor tasks employed. The systematic comparison across multiple surrogate methods, validated against synthetic data showing that conservative methods maintain acceptable false positive rates (7-14 patterns across 96 datasets), provides a more rigorous assessment than previous studies that relied primarily on uniform dithering.

Methodological Recommendations Based on our comprehensive evaluation, we recommend specific guidelines for surrogate method selection in experimental applications. For exploratory analyses, researchers should employ TR-SHIFT or JISI-D methods to minimize false positive rates while maintaining reasonable sensitivity. These methods provide the most conservative estimates of coordination prevalence and should be preferred when false positive control represents the primary concern.

For confirmatory analyses, investigators should employ multiple surrogate methods and require cross-method validation for pattern significance. Patterns detected consistently across TR-SHIFT, ISI-D, and JISI-D methods are most likely to represent genuine coordination, as the convergence of evidence from independent approaches strengthens confidence in the biological relevance of detected patterns.

Our results provide compelling evidence that uniform dithering should be avoided entirely for experimental data containing regular firing patterns or elevated firing rates, conditions commonly encountered in motor cortical recordings. The systematic bias demonstrated across both synthetic and experimental data makes UD unsuitable for reliable pattern detection in realistic neuroscience applications.

Researchers should remain aware that current conservative methods may have limited sensitivity for detecting coordination at moderate firing rates, though this limitation appears preferable to the systematic overestimation produced by less rigorous approaches. Future methodological development should focus on enhancing sensitivity while maintaining the stringent false positive control demonstrated by conservative methods.

Implications for Understanding Cortical Coordination Our findings suggest that genuine neural coordination in motor cortical areas may be sparser than previously estimated, occurring primarily during active movement execution and involving neurons with elevated firing rates. This pattern is consistent with theories of motor cortical function that emphasize coordination during movement generation rather than sustained background coordination.

However, the possibility that current methods have limited sensitivity for moderate-rate coordination suggests that cortical coordination may be more extensive than our analysis reveals. The development of methods optimized for different firing rate

regimes and temporal scales represents an important direction for advancing our understanding of cortical network dynamics.

The individual differences observed between monkeys, with distinct patterns of grip-type specificity and temporal organization, highlight the importance of considering individual variations in cortical organization when interpreting coordination analyses. These differences may reflect genuine biological variation in motor control strategies rather than methodological limitations.

1.3.4.6 Synthesis and Future Directions

Our experimental data applications demonstrate that surrogate method choice profoundly influences conclusions about neural coordination, with implications that extend from methodological considerations to fundamental questions about cortical processing. The systematic bias demonstrated by uniform dithering necessitates revision of previous studies that relied primarily on this method, while the low pattern counts detected by conservative methods highlight the need for enhanced sensitivity in coordination detection.

Methodological Limitations and Extensions The sparse pattern detection observed with conservative methods raises important questions about whether SPADE's item-set mining approach, while computationally efficient, may be optimally suited for detecting the subtle coordination patterns that characterize cortical networks during complex behaviors. Future investigations should extend this comparative framework to alternative pattern detection methodologies, including continuous time approaches, sliding window techniques, and methods that capture coordination across broader temporal scales. Such extensions would provide valuable insights into whether our findings reflect fundamental properties of cortical coordination or limitations specific to the SPADE framework.

Additionally, our current analysis employs a conservative approach by restricting pattern detection to features within individual trials, which may overlook coordination patterns that manifest across trial boundaries or require longer temporal integration windows. Future work should explore the incorporation of cross-trial features and trial shuffling approaches that could reveal coordination patterns operating at longer timescales while maintaining appropriate statistical controls. This extension would be particularly valuable for understanding coordination during cognitive aspects of motor control that may not be confined to individual trial boundaries.

Refinement of Previous Findings Our results provide important updates to previous findings in this field, including refinements to conclusions from Stella et al. (2022), where subsequent code verification and methodological improvements have

led to more conservative pattern count estimates. These updates reflect the ongoing evolution of analytical methods in computational neuroscience and highlight the importance of rigorous code validation and methodological transparency in pattern detection studies. The lower pattern counts observed in our analysis, while initially surprising, may provide a more accurate representation of genuine coordination patterns when appropriate statistical controls are applied.

Theoretical Implications for Neural Coding The complexity of our findings also challenges the traditional dualistic framework that sharply distinguishes temporal coding from rate coding mechanisms. The observation that pattern detection sensitivity depends strongly on firing rate characteristics suggests that these coding schemes may operate along a continuum rather than representing discrete alternatives. Neural coordination patterns may simultaneously involve rate modulations and temporal precision, with their detectability depending on the dynamic range and temporal characteristics of the underlying neural population. Future theoretical frameworks should consider more nuanced models that integrate rate and temporal coding mechanisms rather than treating them as mutually exclusive hypotheses.

Convergence and Validation Framework The convergence of synthetic and experimental results provides confidence in our methodological conclusions while raising important questions about the nature and extent of cortical coordination. Future research should focus on developing methods that balance false positive control with enhanced sensitivity for coordination at diverse firing rates and temporal scales, while considering individual variations in cortical organization and motor control strategies.

The framework established by this study, combining synthetic data validation with systematic experimental application, provides a model for rigorous evaluation of analytical methods in computational neuroscience. This approach enables principled method selection based on specific experimental requirements while advancing our understanding of the neural mechanisms underlying skilled motor behavior. The integration of multiple surrogate methods and systematic bias assessment represents a significant advancement in the reliability of coordination detection studies, though continued methodological development remains essential for fully characterizing the rich dynamics of cortical network activity.

1.4 Discussion and Implications

1.4.1 Systematic Biases in Uniform Dithering

Our evaluation demonstrates that uniform dithering introduces biases severe enough to compromise temporal precision analyses, despite widespread use across computational neuroscience studies (Abeles and Gat, 2001; Hatsopoulos et al., 2003; Shmiel et al., 2006; Maldonado et al., 2008; Torre et al., 2016).

Nature and Magnitude of Statistical Issues

Multiple lines of evidence reveal significant problems with this approach. When applied to realistic synthetic data matching experimental conditions, uniform dithering generated false positives in over 90% of datasets – a family-wise error rate that undermines statistical reliability. Specifically, uniform dithering produced 707-1124 false positive patterns across 96 synthetic datasets, while conservative methods generated only 7-14 false positives total. These findings align with earlier methodological concerns (Pazienti et al., 2008; Louis et al., 2010) and highlight ongoing challenges in spike train analysis first recognized in foundational work on neural point processes (Rieke et al., 1997).

The Critical Role of Binning and Clipping

The problems become particularly acute when uniform dithering combines with binning and clipping procedures common in pattern detection algorithms. This combination creates a cascade of distortions: uniform dithering first violates biological constraints by generating unrealistically brief interspike intervals, then binning and clipping reduce spike counts by up to 20% while transforming regular firing patterns into Poisson-like distributions. Earlier studies documented individual aspects of these effects (Harrison et al., 2007; Pazienti et al., 2008), while subsequent work identified interspike interval alterations as a fundamental challenge in surrogate generation (Harrison and Geman, 2009; Louis et al., 2010). While each distortion alone compromises statistical validity, their combination with temporal discretization creates compounding biases that demand caution in applications requiring binning and clipping.

Implications for Previous Studies

The widespread adoption of uniform dithering across temporal precision studies raises questions about result reliability, particularly for analyses combining dithering

with binning and clipping. Our findings reveal a stark contrast: uniform dithering identified 539-654 patterns in experimental data while conservative methods detected 2-17 patterns per session across identical datasets, suggesting some reported patterns may reflect statistical artifacts rather than genuine neural coordination (Baker and Lemon, 2000; Louis et al., 2010).

The severity of these issues depends critically on analytical pipelines. Studies employing uniform dithering without subsequent binarization face reduced risk from the biases we identify, while spatiotemporal pattern detection methods vary in their susceptibility to these effects (Grün et al., 2002a). For instance, Torre et al. (2016) used CoCoNAD rather than FIM-based methods, potentially providing some protection against these biases, though this relationship requires further investigation. However, determining the extent of impact on published findings remains challenging given diverse analytical approaches across the field.

1.4.2 Performance of Conservative Surrogate Methods

Having identified uniform dithering’s limitations, our assessment of alternative approaches shows that rigorous temporal precision detection remains achievable through principled surrogate methods.

Trial Shifting as a Reference Method

Trial shifting demonstrated superior performance through an elegant solution to correlation elimination. By applying independent temporal displacements to complete spike trains on a trial-by-trial basis, this method preserves all within-neuron statistical properties while eliminating cross-neuron temporal relationships. This approach represents a highly conservative test: patterns surviving such disruption likely reflect coordination beyond rate modulations alone.

Performance remained consistent across validation tests, with family-wise error rates of 10.4-16.7% (meaning 10.4-16.7% of datasets contained any false positives) while preserving genuine temporal coordination detection. Beyond empirical success, trial shifting provides clear theoretical grounding – it directly tests whether temporal precision extends beyond rate-based effects.

Alternative Conservative Methods

ISI-based methods (JISI-D, ISI-D) and window shuffling offered acceptable performance with family-wise error rates of 12.5-18.8%. These approaches generated 7-14

false positives across 96 synthetic datasets, representing a substantial improvement over uniform dithering while maintaining reasonable statistical control.

Uniform dithering with dead-time (UDD) showed process-dependent performance, demonstrating that preserving biological constraints can dramatically improve statistical validity. When applied to PPD synthetic data with explicit refractory periods, UDD achieved excellent control with only 4 false positives across 96 datasets (FWER 8.3%). However, performance deteriorated substantially for Gamma process data (318 false positives, FWER 58.3%), indicating effectiveness depends critically on data characteristics.

Consistency Across Methods

The general agreement across conservative approaches when applied to experimental data supports their reliability. Multiple independent methods yielding similar conclusions increases confidence in findings. The sharp contrast with uniform dithering's performance further validates conservative approaches.

1.4.3 Implications for Neural Coordination

Our analysis offers limited insights into temporal precision prevalence in neural systems, with interpretations remaining cautious given the small number of patterns detected by conservative methods.

Apparent Sparsity of Detected Coordination

Conservative methods revealed substantially fewer coordination patterns than previously reported, though this observation permits only tentative inferences about temporal precision's true prevalence. The small pattern count (2-17 per method per session) constrains our ability to draw strong conclusions about underlying neural mechanisms. This apparent sparsity might align with theoretical considerations regarding potential metabolic constraints on precise coordination and computational trade-offs in neural coding (Softky and Koch, 1993; Lennie, 2003), though alternative explanations remain plausible.

In our experimental data, fine-temporal patterns predominantly involved neurons with high firing rates, though this observation may reflect methodological artifacts of the SPADE approach rather than fundamental neural properties. Pattern detection algorithms can exhibit biases toward high-firing neurons due to increased statistical power for detecting coincidences, making it difficult to determine whether this represents genuine selective coordination during periods of enhanced activity (Riehle et al., 1997; Fries, 2005) or detection limitations. Limited sample size prevents definitive conclusions about the relationship between firing rates and temporal coordination.

1.4.4 Methodological Implications

Our findings carry important implications for interpreting temporal precision claims and designing future experiments.

Interpreting Temporal Precision Findings

Conservative surrogate methods for evaluating the statistical significance of detected patterns reveal substantially fewer significant patterns than uniform dithering, suggesting temporal precision, while genuine, may be less prevalent than some studies indicate. This need not diminish temporal coding mechanisms' significance but rather suggests they operate more selectively, connecting to fundamental questions about neural population information encoding and the complementary relationship between rate and temporal coding (Theunissen and Miller, 1995; Pouget et al., 2000; Panzeri et al., 2010).

This selectivity could reflect computational principles. If neural systems employ temporal precision primarily during demanding computational periods, conservative surrogate testing may illuminate when and why precise coordination emerges. Patterns surviving conservative statistical evaluation likely represent coordination mechanisms warranting further investigation.

Methodological Recommendations

Our analysis suggests guidelines for future temporal precision studies. Trial shifting appears well-suited for most applications when evaluating pattern significance, with other conservative methods employed when specific experimental constraints require alternatives. Uniform dithering should be used cautiously for significance testing with realistic neural data, particularly when analyzing neurons with regular firing patterns or high firing rates.

This work underscores the importance of rigorous surrogate validation in computational neuroscience, building on established principles for neural data statistical testing (Brown et al., 2004; Kass et al., 2005; Ventura, 2010). Our methodological framework aligns with broader efforts to improve neuroscience research reproducibility and statistical rigor.

1.4.5 Future Directions

Integration with Emerging Technologies

As recording technologies enable simultaneous monitoring of thousands of neurons (Jun et al., 2017; Steinmetz et al., 2021), the statistical challenges addressed here will

gain increasing importance. Principles demonstrated by our conservative methods – preserving essential statistical properties while selectively destroying hypothesized relationships – offer guidance for developing surrogate techniques appropriate for large-scale neural data, where both opportunities for detecting genuine coordination and risks of spurious findings are amplified (Brown et al., 2004; Stevenson and Kording, 2011).

Theoretical Considerations

Temporal precision validated through conservative methods connects to theoretical frameworks about neural computation, including synfire chain models and precisely timed interactions (Abeles, 1991; Diesmann et al., 1999; Izhikevich, 2006). However, observed sparsity suggests these mechanisms may operate more selectively than some theoretical frameworks assume. Recent work on neural population dynamics may provide complementary perspectives on how temporal coordination emerges within broader population-level computation contexts (Cunningham and Yu, 2014; Saxena and Cunningham, 2019; Gallego et al., 2020).

Our work demonstrates that rigorous temporal precision detection remains achievable through conservative methods that preserve statistical integrity. This methodological framework provides computational neuroscience with improved tools for investigating temporal coordination, demonstrating that methodological precision proves equally essential in the pursuit of neural precision.

Dynamic Recurrent Circuits and the Temporal Evolution of Neural Representations

This chapter summarizes the following publication:

Schutzeichel, L., Bauer, J., Bouss, P., Musall, S., Dahmen, D., and Helias, M. (2025). Transient recurrent dynamics shape representations in mice. *BioRxiv*, 2025–06.

Author contributions

Lars Schutzeichel conducted the main research and analysis under the supervision of David Dahmen and Moritz Helias. I contributed to discussions and manuscript writing.

2.1 The Challenge of Similarity Processing in Recurrent Circuits

Neural circuits face a basic problem: they must distinguish between different stimuli while recognizing that the same stimulus can appear in different contexts. A red ball looks different in bright versus dim light, yet we instantly recognize both as the same object. How do recurrent networks solve this similarity processing challenge? This dual requirement – to separate different stimuli while preserving similarities within stimulus categories – lies at the heart of how recurrent neural circuits process sensory information.

The brain's solution to this challenge becomes particularly critical in higher-order areas that are tightly linked to cognition, where final decisions about stimulus identity must be made. Yet the computational principles underlying this similarity processing begin to emerge much earlier in the processing hierarchy. By examining how recurrent circuits in early processing areas like the superior colliculus (SC) handle the separation and grouping of sensory representations, we can uncover the fundamental mechanisms that ultimately enable sophisticated cognitive discriminations.

Building on the correlation detection methods from Chapter 1 we now ask: *how do recurrent circuits dynamically reshape representations to balance similarity preservation with discrimination?*

The rigorous surrogate methods developed in Chapter 1 provide the statistical foundation for detecting precisely timed correlations in discrete neural signals. However, understanding when and why these correlations emerge requires examining the network-level dynamics that generate them (Sompolinsky et al., 1988; Renart et al., 2010). This chapter therefore shifts focus from detecting temporal precision to understanding how recurrent circuit dynamics fundamentally shape neural representations over time, setting the stage for the more sophisticated analytical approaches developed in subsequent chapters.

2.2 A Mean-Field Theory of Neural Representation Dynamics

2.2.1 Reducing High-Dimensional Dynamics to Essential Quantities

To understand how recurrent circuits solve the similarity processing challenge, we need a theoretical framework that captures the essential dynamics without getting lost in the complexity of individual synaptic interactions. Mean-field theory reduces network complexity to three key variables: population firing rate (R), within-class

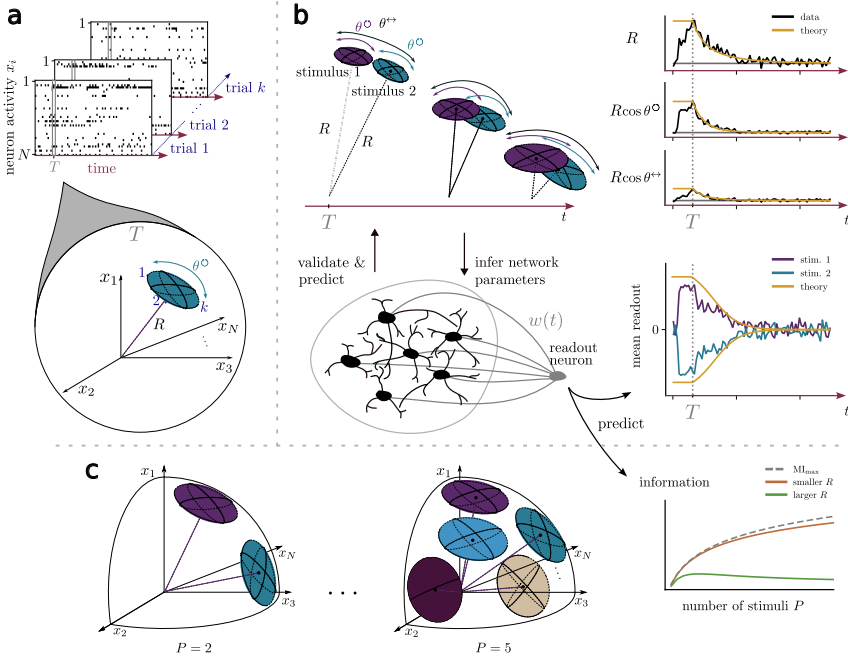


Figure 2.1: Mean-field theory captures recurrent network dynamics through geometric representation analysis. (a) Population spike trains recorded from mouse superior colliculus during multimodal sensory stimulation (exemplified by whisker air puff stimulation) reveal structured trial-to-trial variability. Each neural population state at time T across trials maps to coordinates in high-dimensional neuronal space, with individual trials appearing as discrete points. Distinct sensory modalities (visual, tactile) define separate "stimulus classes" that cluster into characteristic point clouds when visualized in this neural state space. The geometric properties of these clouds – specifically their radial distance R from the origin and their angular extent θ° – can be mathematically approximated as ellipsoidal structures. (b) Dynamic trajectories of two stimulus classes (tactile in blue, visual in purple) demonstrate how neural representations evolve temporally following stimulus presentation at time T_0 , with experimental measurements closely matching theoretical predictions from mean-field analysis. The relative angular separation $\theta^{\leftrightarrow}$ between stimulus classes determines their discriminability over time. A linear classifier $w(t)$ trained to optimally discriminate between stimulus classes provides a time-resolved measure of representational separability, directly determined by the temporal dynamics of population rate $R(t)$ and the geometric overlaps $Q^\circ(t) = R(t)\cos\theta^\circ(t)$ and $Q^{\leftrightarrow}(t) = R(t)\cos\theta^{\leftrightarrow}(t)$, which quantify within-class and cross-class correlations respectively. (c) Information-theoretic analysis reveals how separability metrics scale to multi-stimulus scenarios ($P \geq 2$ classes), where increased stimulus diversity enhances information capacity but neural space constraints create competitive interference between representations. Population firing rate R governs a critical phase transition: low rates (brown region) support expanding information transmission, while high rates (green region) lead to representational collapse and information loss.

correlations ($Q^\odot$), and across-class correlations ($Q^{\leftrightarrow}$). Together, these determine how well circuits separate stimuli while preserving category structure.

The geometric picture is intuitive. Present a red ball repeatedly – each trial yields slightly different neural responses, forming a ‘cloud’ in high-dimensional space. Blue ball responses form their own cloud. Discrimination depends on cloud separation; categorization depends on cloud compactness.

Mean-field theory captures recurrent network dynamics through geometric representation analysis (Figure 2.1). Population spike trains recorded from neural circuits during sensory stimulation reveal structured trial-to-trial variability. Each neural population state at time T across trials maps to coordinates in high-dimensional neuronal space, with individual trials appearing as discrete points. Distinct sensory modalities or stimulus categories define separate “stimulus classes” that cluster into characteristic point clouds when visualized in this neural state space.

This chapter presents a concise summary of published work in which I contributed as co-author (Schutzeichel et al., 2025), investigating how neural representations are dynamically reshaped by recurrent circuits using both experimental data from awake behaving mice and recurrent neural network models. Complete experimental methods and detailed results can be found in the original publication.

The central theoretical contribution is a mean-field framework that reduces the complex dynamics of high-dimensional recurrent networks to three observable quantities (Figure 2.1b). Understanding the geometrical structure underlying neuronal responses has led to insights into representations in both artificial (Chung et al., 2016, 2018; Cohen et al., 2020) and biological networks (Bernardi et al., 2020; Fascianelli et al., 2024). The geometry of neuronal responses in high-level areas lies on low-dimensional manifolds (Gallego et al., 2017), whose size, shape, and dimensionality determine pattern classification accuracy and capacity.

Animals perceive external stimuli through intricate networks of neurons, with information encoded in the activity of neuronal populations – a concept first established in motor cortex studies (Georgopoulos et al., 1986) and now recognized as fundamental to neural computation across brain areas (Saxena and Cunningham, 2019; Vyas et al., 2020).

However, population activity alone is insufficient to understand how recurrent circuits solve the similarity processing challenge. The key insight is that the *relationships* between population responses – how similar responses are within a stimulus category versus across different categories – determine the circuit’s computational capacity. Our mean-field approach captures these essential relationships through three dynamical quantities:

- R : The average firing rate across neurons (population mean activity)

- $Q^\odot$: Overlap of responses to the same stimulus (within-class correlations)
- $Q^{\leftrightarrow}$: Overlap of responses to different stimuli (across-class correlations)

These three quantities jointly determine the separability of neural representations – a critical factor for downstream information processing. To identify and classify stimuli, neural representations must preserve similarities between repetitions of the same stimulus while remaining distinguishable between different stimulus classes.

In the brain, recurrence plays a crucial role in understanding how neural representations deform over time (Kietzmann et al., 2019). Using statistical physics techniques and tools from disordered systems (Helias and Dahmen, 2020), we derive closed-form equations describing how these variables evolve over time in large networks with chaotic dynamics (Sompolinsky et al., 1988; van Vreeswijk and Sompolinsky, 1996; Keup et al., 2021).

2.2.2 Experimental Validation in Mouse Superior Colliculus

The theory is validated using Neuropixels recordings from mouse superior colliculus (SC), a midbrain region that integrates multisensory information and is crucial for processing visual information in mice (Wang et al., 2020; Ito and Feldheim, 2018). The superior colliculus is not only important for visual processing but is also highly multimodal, making it suitable for analyzing different stimulus modalities including visual, auditory, and somatosensory inputs (Meredith and Stein, 1986; White and Munoz, 2011; Benavidez et al., 2021).

The experimental data reveal fundamental properties of neural variability: when identical stimuli are presented, evoked responses share characteristics but also show trial-to-trial variability across repeated presentations (Arieli et al., 1996). This neural variability is present even in the absence of stimuli (spontaneous activity) and typically decreases upon stimulus application – a phenomenon known as quenching of neural variability (Churchland et al., 2010; Litwin-Kumar and Doiron, 2012; Mazzucato et al., 2016).

In the high-dimensional neural state space, stimulus representations observed in our data can be approximated as ellipsoidal structures. While the actual shape of response clouds may deviate from perfect ellipsoids, we can characterize each stimulus class by finding an ellipsoid that encompasses the point cloud, providing a first-order geometric approximation. In this framework, the distance of the ellipsoid’s center from the origin represents the population mean firing rate R , and the ellipsoid’s extent reflects the trial-to-trial variability (Figure 2.1a). The angular separation between ellipsoids from different stimulus classes quantifies their overlap. Upon stimulus onset, neural trajectories briefly align and then diverge, with this temporal evolution governed by the dynamics of the three observables: R (population mean), $Q^\odot$ (within-class overlap), and $Q^{\leftrightarrow}$ (across-class overlap) (Figure 2.1b).

Remarkably, this minimal mean-field model – requiring no fine-tuning of individual synaptic weights, only appropriate statistical parameters such as connectivity strength and excitation-inhibition balance – quantitatively reproduces key features observed in experimental recordings (Figure 2.1b). The model assumes random Gaussian connectivity and operates in the chaotic regime, where the dynamics promote dissimilarities across stimulus responses (Renart et al., 2010; Tetzlaff et al., 2012). Negative feedback from inhibitory neurons stabilizes the dynamics and decorrelates neural activity, while the intrinsic chaos amplifies small differences between initial conditions. This suggests that many computational properties are universal features of such networks, independent of detailed circuit structure.

2.2.3 Linear Readout and Separability Assessment

To quantify stimulus separability, we employ a linear readout neuron that is optimally trained using a Bayesian framework to distinguish between different stimulus classes (Figure 2.1b). The readout performance directly depends on the three mean-field observables R , $Q^\odot$, and $Q^{\leftrightarrow}$, establishing a quantitative link between network dynamics and information processing capacity.

For a network with N neurons, the separability between two stimulus classes can be expressed analytically in terms of these observables. Separability μ depends on the ratio of total variability to signal: it increases when the total variability term ($\kappa + R - Q^\odot$) decreases relative to the signal term ($Q^\odot - Q^{\leftrightarrow}$), where κ represents the variance (strength) of the noise in the readout process.

This framework extends naturally to multi-class scenarios ($P \geq 2$), where encoding more stimuli generally increases information content, but neural space constraints create competitive interference between representations (Figure 2.1c). The average population rate R governs a critical phase transition: low rates (sparse coding) support expanding information transmission, while high rates lead to representational collapse.

The separability formula reveals why sparse coding emerges as a computational necessity for efficient information processing (Olshausen and Field, 1996; Vinje and Gallant, 2000). When neural responses become sparse, reducing the average population firing rate R , the total variability term ($\kappa + R - Q^\odot$) diminishes more rapidly than the signal term ($Q^\odot - Q^{\leftrightarrow}$), resulting in improved signal-to-noise characteristics that enhance representational separability and support the brain's preference for sparse neural codes.

2.3 Limitations of Traditional Correlation Analysis

2.3.1 Extending Beyond Pairwise Correlations

While this mean-field approach successfully captures essential network dynamics and provides quantitative predictions that match experimental data, it also reveals

compelling opportunities for extending the analysis to higher-order statistical structures. The mean-field reduction to three scalar variables (R , $Q^\odot$, $Q^\leftrightarrow$) effectively characterizes the fundamental dynamics of neural representations, yet the complexity of neural population activity suggests that additional insights may be obtained through more comprehensive statistical analysis.

The success of this pairwise correlation framework demonstrates that examining correlations provides a principled approach to understanding the complex, high-dimensional statistical structures present in neural population activity. Real neural networks exhibit rich patterns of higher-order interactions, non-Gaussian statistics, and structured correlations that extend beyond what the mean-field approximation captures (Schneidman et al., 2006; Macke et al., 2011). The ellipsoid representation of stimulus clouds in neural space (Figure 2.1a), while mathematically tractable and empirically validated, represents an initial characterization of what may be more complex geometric structures underlying neural representations.

This recognition that neural dynamics exhibit more intricate statistical structures than those captured by pairwise correlations motivates the development of more sophisticated analytical approaches that can characterize the full complexity of neural representations while building upon the theoretical foundation established by the mean-field framework.

2.4 Connecting to Interaction Learning and Manifold Geometry

2.4.1 Bridging Discrete and Continuous Perspectives

This limitation directly sets the stage for Chapter 3, which addresses how to characterize complex statistical structures that emerge in neural networks when moving beyond simple pairwise correlations. The data-driven statistical learning methods developed using invertible neural networks (Dinh et al., 2015; Rezende and Mohamed, 2015) provide a principled way to extract hierarchical interaction structures from population activity patterns.

Furthermore, the finding that neural representations lie on low-dimensional manifolds whose geometry affects classification accuracy connects directly to Chapter 4's geometric analysis. Models have shown that such low-dimensional dynamics emerge from low-rank connectivity in recurrent networks (Mastrogiuseppe and Ostojic, 2018). The mean-field theory reveals *what* happens to neural representations over time, but the geometric perspective reveals *how* these transformations are embedded in the structure of neural state space, bridging state-dependent computations and representational geometry (Buonomano and Maass, 2009; Kriegeskorte and Kievit, 2013).

2.4.2 From Rate-Based to Geometric Coding

The discovery that firing rate alone is insufficient for determining information content – that different firing rate dynamics can yield identical separability – suggests that the geometric organization of neural activity may be more fundamental than previously appreciated (Churchland et al., 2012). This insight bridges the temporal precision analysis of Chapter 1 with the manifold geometry analysis of Chapter 4: precisely timed correlations may reflect the brain’s strategy for maintaining separable representations within the constraints of curved, high-dimensional neural manifolds, connecting computation through neural population dynamics to representational geometry (Vyas et al., 2020).

2.5 Implications for Neural Computation

2.5.1 Temporal Dynamics and Spatial Organization

The transient nature of representation separability identified in this work (Figure 2.1b) suggests potential connections between temporal precision (Chapter 1) and representational geometry (Chapter 4). The brief windows during which neural trajectories align and diverge indicate that specific temporal epochs may be associated with structured changes in neural manifold geometry, potentially enabling representational transformations.

This temporal-geometric relationship raises the possibility that the precisely timed correlations assessed for temporal versus rate coding using the surrogate methods of Chapter 1 could contribute to coordinating population activity during periods of representational change. The geometric analysis of Chapter 4 may help reveal whether and how these temporal coordination events relate to changes in manifold curvature and higher-order statistical structure, potentially linking the dynamic evolution of neural circuits to their representational geometry (Stopfer et al., 2003; Monteforte and Wolf, 2012).

2.5.2 A Framework for Multi-Scale Analysis

Together, this mean-field approach and the methodological innovations of Chapters 1 and 4 provide complementary perspectives on neural computation (Ganguli and Sompolinsky, 2012). The temporal precision analysis provides insights into when coordination may occur, the mean-field theory offers explanations for why such patterns might emerge from network dynamics, and the geometric analysis characterizes how coordination could manifest in population-level structure. This framework suggests potential bridges between the neural population doctrine and representational geometry, indicating how computation might emerge through the interplay of temporal precision and representational geometry (Kriegeskorte and Kievit, 2013; Saxena and Cunningham, 2019).

2.6 Looking Forward: From Network Dynamics to Statistical Learning

The insights gained from this chapter – that traditional correlation-based approaches provide a foundational yet simplified characterization of complex, high-dimensional statistical structures – directly motivate the data-driven approaches of the subsequent chapters. The recognition that real neural dynamics exhibit more intricate structures than captured by pairwise correlations establishes the foundation for the advanced statistical learning methods that follow.

Chapter 3 addresses how to learn these complex statistical structures directly from data using invertible neural networks, while Chapter 4 extends this framework to reveal how higher-order correlations relate to geometric structures in neural manifolds. The progression from this chapter’s mean-field reduction to the full characterization of manifold geometry represents a fundamental advance in our ability to understand neural computation across multiple scales.

2.7 Conclusion

This chapter demonstrates how recurrent circuit dynamics fundamentally shape neural representations through a minimal yet powerful theoretical framework. By reducing complex network dynamics to three essential quantities, we reveal the computational principles underlying stimulus separability and information transmission. The identification of sparse coding as necessary for efficient representation in large networks provides a crucial constraint for understanding neural computation.

However, the scope of this approach – particularly its reduction of rich, high-dimensional statistical structures to tractable correlation measures – presents opportunities for more sophisticated analytical tools. The subsequent chapters build upon this foundation by developing data-driven methods for characterizing complex statistical dependencies and their geometric manifestations, ultimately providing a comprehensive framework for understanding neural computation from temporal precision to manifold geometry.

Data-Driven Discovery of Statistical Interaction Principles

This chapter summarizes the following publication:

Merger, C., René, A., Fischer, K., Bouss, P., Nestler, S., Dahmen, D., Honerkamp, C., and Helias, M. (2023). Learning interacting theories from data. *Physical Review X*, 13:041033.

Author contributions

Claudia Merger conducted the main research and analysis under the supervision of David Dahmen, Carsten Honerkamp, and Moritz Helias. I contributed to discussions, the code base, and manuscript writing. An extended version of this material was previously presented in the doctoral thesis of Claudia Merger (2023).

3.1 From Mean-Field Foundations to Data-Driven Statistical Learning

The mean-field analysis in Chapter 2 successfully captured recurrent network dynamics with three essential variables, providing quantitative predictions that matched experimental data. But this success came with a trade-off: reducing neural complexity to three observables, while mathematically elegant, necessarily discarded the rich statistical dependencies present in neural population activity.

We developed an approach to learn interaction structures directly from data using invertible neural networks – models that can extract the statistical ‘rules’ governing a system without prior assumptions.

While Chapter 2 showed what happens to neural representations over time, this chapter asks *how data structures emerge from underlying interaction principles that we can learn directly from observed data.*

3.2 The Challenge of Inverse Statistical Modeling

This chapter presents a concise summary of published work in which I contributed as co-author (Merger et al., 2023), addressing the inverse problem in physics: inferring microscopic interactions from macroscopic observations (Figure 3.1). Complete methods and detailed results can be found in the original publication. In physics, interactions typically form the building blocks of theories and are described by polynomial terms in an action. The conventional approach derives these terms from elementary processes and uses the resulting model to make predictions about the entire system.

We explore the reverse approach: learning the microscopic action by observing the complete system, even when the underlying processes are unknown (Figure 3.1). This approach complements the analysis in Chapter 2, where mean-field reduction successfully characterized network dynamics through essential observables, but necessarily operated at a level that did not examine the underlying interaction mechanisms. The central question becomes whether we can reverse-engineer the fundamental interaction principles that give rise to neural population dynamics directly from observed data, going beyond the constraints of maximum entropy approaches (Jaynes, 1957; Schneidman et al., 2006).

3.3 Invertible Neural Networks for Statistical Learning

3.3.1 A Novel Approach Using Invertible Neural Networks

To tackle this inverse problem, we employ invertible neural networks (INNs), also known as normalizing flows, as generative models (Dinh et al., 2015, 2016; Rezende

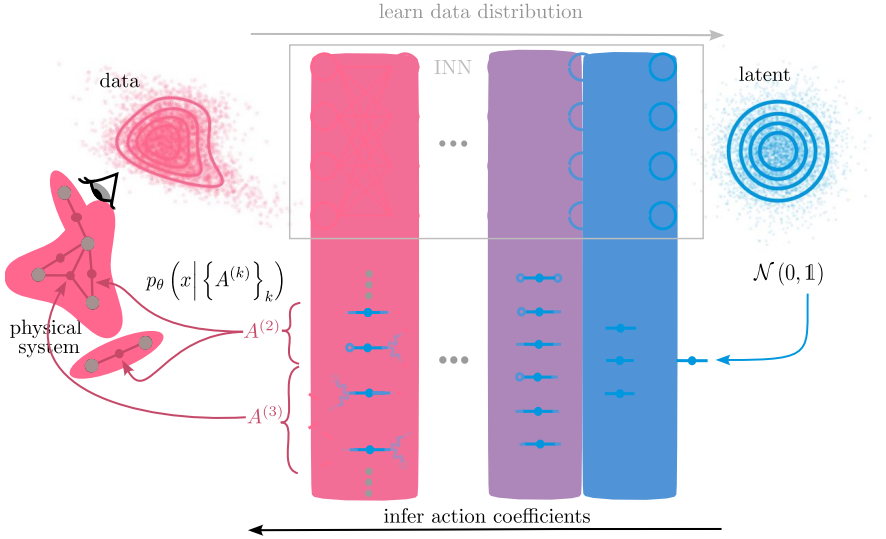


Figure 3.1: **Invertible neural network framework for extracting microscopic interaction theories from macroscopic system observations.** The methodology begins with measurements from a complex physical system containing multiple interacting components (gray dots) where the underlying interaction mechanisms remain unknown (shaded regions indicating hidden coupling structures). An invertible neural network architecture is trained in an unsupervised manner to capture the complete probability distribution underlying the observed measurements. Through a systematic diagrammatic expansion protocol, the trained network parameters are decoded layer by layer to reconstruct the microscopic action governing system behavior. The resulting action coefficients $A^{(k)}$ (pink nodes) provide quantitative estimates of the learned interaction strengths that generate the observed macroscopic dynamics.

and Mohamed, 2015; Kobyzev et al., 2020). Invertible neural networks learn by mapping complex data to a simple Gaussian distribution (Figure 3.1). The key insight: if we can learn this mapping, we can extract the statistical rules that created the original complexity. This approach leverages the change of variables formula central to normalizing flows: for a bijective mapping $f : X \rightarrow Z$, the probability density transforms as

$$p_X(x) = p_Z(f(x)) \left| \det \left(\frac{\partial f}{\partial x} \right) \right|,$$

where the Jacobian determinant accounts for volume changes under the transformation.

The key innovation lies in designing the INN architecture with specific properties: we use quadratic nonlinearities as activation functions, which enables extraction of an explicit microscopic action from the trained network parameters. This differs from typical machine learning applications where networks function as black boxes, as our approach allows direct interpretation of the learned representations through principles from statistical mechanics (Mézard and Montanari, 2009).

3.3.2 Concrete Example: How Higher-Order Interactions Emerge

To illustrate how higher-order interactions emerge from our approach, consider a simple 2D example that demonstrates the core mechanism. Suppose we have two variables x_1 and x_2 that follow standard Gaussian distributions in the latent space. The invertible network implements a transformation that maps these independent variables to correlated observations in the data space through

$$z_1 = x_1 \text{ and } z_2 = x_2 + \alpha x_1^2,$$

where α is a learnable parameter that characterizes the interaction strength between the components.

This transformation reveals how statistical dependencies arise from simple nonlinear couplings. While the latent variables (x_1, x_2) are independent with the standard Gaussian action

$$S_x = (x_1^2 + x_2^2)/2,$$

the transformed variables (z_1, z_2) in the data space exhibit complex correlations. To understand these correlations, we derive the action governing the z variables by applying the change of variables formula with the inverse transformation

$$x_1 = z_1 \text{ and } x_2 = z_2 - \alpha z_1^2.$$

Since the Jacobian determinant of this transformation equals unity, the action in the data space becomes

$$S_z = (z_1^2 + z_2^2)/2 - \alpha z_2 z_1^2 + \alpha^2 z_1^4/2.$$

The quadratic coupling αx_1^2 in the transformation has produced both cubic interaction terms ($z_2 z_1^2$) and quartic self-interaction terms (z_1^4) in the resulting action, demonstrating how higher-order statistical dependencies emerge from the network's learned transformations.

The example illustrates three fundamental principles of our approach. First, higher-order interactions arise automatically from simple transformations, showing how complex statistical structures can emerge from basic nonlinear couplings. Second, the network creates a hierarchical structure where interactions in one layer generate increasingly complex dependencies in subsequent representations, explaining why network depth corresponds to the complexity of learnable interactions. Third, the learned parameter α directly characterizes the strength of the underlying interaction, making the network's internal representations physically interpretable rather than abstract feature encodings.

This concrete example demonstrates how our invertible neural network framework can extract microscopic interaction principles from observed correlations, moving beyond simple pairwise statistics to capture the full hierarchical structure of statistical dependencies that characterize complex systems.

3.3.3 Hierarchical Construction of Interactions

Building on the 2D example, our diagrammatic formalism generalizes this process to arbitrary dimensions and network depths. Each layer of the network transforms and composes pairwise interactions hierarchically into increasingly complex, higher-order terms (Figure 3.1). Just as the simple transformation $z_2 = x_2 + \alpha x_1^2$ generated cubic and quartic terms, deeper architectures create progressively more complex interaction hierarchies.

This hierarchical construction explains why deep architectures may be necessary for learning rich, non-Gaussian statistics that cannot be captured by low-order approximations. The depth-complexity relationship has implications for understanding neural computation more broadly: complex population dynamics may require hierarchical interaction structures to generate the rich statistical patterns observed in neural data.

3.3.4 Advantages Over Traditional Methods

Our approach offers several advantages over existing inverse modeling methods. Traditional approaches face certain constraints: statistical physics methods like Boltzmann machines (Ackley et al., 1985; Hinton and Sejnowski, 1986) are typically restricted to discrete variables and specific interaction forms (often pairwise), while graphical model approaches (Koller and Friedman, 2009) generally require explicit specification of network structure and interaction types.

Our invertible neural network framework learns statistical structures directly from data through end-to-end training, supporting continuous variables and learning interaction hierarchies without predetermined assumptions about their form or order. This provides methodological flexibility that can adapt to different systems without requiring domain-specific modifications.

Additionally, our approach maintains interpretability while providing this flexibility. While modern generative models such as VAEs (Kingma and Welling, 2013) and GANs (Goodfellow et al., 2014) achieve high expressiveness through complex architectures, our method makes a different trade-off: we use quadratic nonlinearities to enable extraction of physically meaningful action coefficients that can be interpreted as interaction strengths (Figure 3.1). This design choice prioritizes interpretability over maximum expressiveness, which can be valuable for scientific applications where understanding the underlying mechanisms is important.

3.3.5 Bridging to Geometric Analysis

The learned transformations in our invertible networks map complex, high-dimensional patterns to structured latent spaces. This mapping provides insight into how the curved, nonlinear manifold structures analyzed in Chapter 4 might emerge from underlying statistical principles rather than being imposed by specific architectural constraints.

The network effectively implements a transformation that maps interacting configurations in data space to independent Gaussian samples in latent space. By learning this statistical decoupling, the method may reveal how complex geometric features could arise naturally from interaction principles. This perspective suggests that the manifold structures identified through geometric analysis might reflect the statistical organization learned by the invertible transformation, providing a bridge between statistical mechanics and geometric approaches to understanding neural computation.

3.4 Experimental Validation and Performance

3.4.1 Testing on Simulated and Real Data

We test our framework on both simulated datasets from known interacting theories and real-world data. The network can reproduce broad classes of unimodal distributions; for distributions outside this class, it finds effective theories that approximate the data statistics up to the third cumulant. This performance is evaluated against ground truth for synthetic data.

Our analysis explicitly demonstrates how network depth and data quantity jointly improve agreement between learned and true models. The model shows increasing accuracy as both network depth and data volume grow, supporting the expectation that more complex theories require more expressive models and larger datasets for effective learning.

3.4.2 Insights into Training Dynamics

The formalism provides insights into training dynamics that connect to statistical physics principles. Limited data introduce bias, particularly for higher-order statistics, while the number of parameters grows rapidly with network depth. This necessitates careful balance between model expressiveness and computational tractability – echoing the information-capacity trade-offs observed in Chapter 2’s neural circuit analysis.

These findings reflect common challenges in deep learning: balancing model complexity with available data and managing the rapid parameter growth with network depth. This necessitates the same expressiveness-tractability trade-offs found throughout deep network theory.

The trained INN can be viewed as implementing a transformation that maps complex, high-dimensional dynamics to a structured latent space where interactions may become more tractable. This provides one perspective on how neural circuits might implement computation: rather than relying solely on specific architectural constraints, they may exploit statistical principles that emerge from optimization processes

3.5 Extensions and Future Directions

3.5.1 Multimodal Extensions

The current framework is limited to unimodal distributions due to our use of Gaussian latent space and quadratic nonlinearities. Future extensions could address multimodal cases using mixture models in the latent space (Bishop, 2006), where each

mixture component would correspond to an effective theory for different system phases. This approach could incorporate advances in normalizing flows for handling multimodal distributions (Izmailov et al., 2020). Chapter 4 explores a related approach using volume-preserving invertible networks with multimodal latent spaces and more complex nonlinearities, while maintaining interpretability through different methods.

3.5.2 Higher-Order Nonlinearities

Our approach could be extended to incorporate higher-order nonlinearities in the analytical derivations, such as cubic or quartic terms in addition to the current quadratic framework. If such extensions prove analytically tractable, they could enable representation of more complex interaction hierarchies while maintaining the interpretability of action coefficient extraction. While the path from cubic or quartic terms to the ELU nonlinearities used in Chapter 4 remains substantial, integrating these simpler higher-order nonlinearities into Chapter 4’s framework could reveal how different polynomial orders affect the higher-order correlations and manifold curvature properties extracted there. This could provide insights into the relationship between interaction complexity and geometric structure in neural manifolds.

3.6 Implications for Understanding Neural Computation

3.6.1 Statistical Learning as a Unifying Framework

This work positions statistical learning as a unifying framework that connects the discrete temporal analysis of Chapter 1, the continuous dynamics of Chapter 2, and the geometric analysis of Chapter 4. The invertible networks provide a principled way to move between these different perspectives while preserving the essential statistical structure of neural population activity.

The framework suggests that neural computation might be understood as a process of statistical inference, where neural circuits learn to extract and manipulate complex statistical dependencies in sensory and motor data. This perspective offers potential insights into how the brain might implement sophisticated computations through relatively simple, biologically plausible mechanisms.

3.6.2 Interpretability in Deep Learning for Neuroscience

Unlike traditional deep learning applications that treat neural networks as black boxes, this approach provides interpretability often missing in modern machine

learning. By explicitly extracting the action and its coefficients, the learned model becomes not just a statistical tool but a physically meaningful representation of neural computation principles (Figure 3.1).

This interpretability is important for neuroscience applications, where understanding the underlying mechanisms is as important as achieving good predictive performance. The framework provides a potential path toward developing AI systems that are both powerful and interpretable, combining the flexibility of deep learning with the rigor of statistical physics.

3.7 Conclusion

This chapter demonstrates how invertible neural networks can address limitations of traditional correlation-based approaches by learning complex statistical structures directly from data. The framework provides a method for extracting microscopic actions from macroscopic observations, offering a data-driven approach to understanding interaction principles (Figure 3.1).

The work establishes theoretical connections between discrete temporal analysis, continuous dynamics, and geometric characterization of neural computation. By moving beyond simple pairwise correlations to capture hierarchical statistical structures, the approach provides tools for analyzing neural data and insights into computational principles.

The framework represents a step toward understanding neural computation as statistical inference, suggesting that computational properties may emerge from optimization processes rather than requiring specific architectural constraints. This perspective contributes to advancing understanding of both biological and artificial intelligence systems by preserving interpretability within powerful machine learning frameworks.

Characterizing Neural Manifolds Using Invertible Neural Networks

This chapter and the appendices A, B, and C are based on the following publication

Bouss*, P., Nestler*, S., Fischer, K., Merger, C., René, A., and Helias, M. (2025). Characterizing neural manifolds' properties and curvatures using normalizing flows. *arXiv* preprint arXiv:2506.12187.

Author contributions

I share first authorship of this publication with Sandra Nestler. Under the supervision of Moritz Helias, Sandra Nestler and I jointly conceived the approach. A preliminary version of this project with initial ideas, though none of the current results, formed part of Sandra Nestler's doctoral thesis (Nestler, 2023). With assistance from Sandra Nestler, I developed the computational framework, created the training routines, and developed the analytical components for the characteristic function and curvature analysis. I executed the empirical analyses, applying the method to macaque visual cortex recordings. Figures in this chapter were reproduced from Bouss, Nestler, et al. (2025). All co-authors participated in discussions throughout the research process and contributed to the manuscript preparation.

4.1 Motivation

Understanding how neural activity patterns encode and process information represents a central challenge in neuroscience (Churchland et al., 2012; Cunningham and Yu, 2014). Cortical neural populations exhibit complex dynamics that encode sensory information, motor commands, and cognitive states. These populations coordinate their activity across hundreds to thousands of neurons, creating high-dimensional patterns that reflect both underlying computations and behavioral states.

Traditional neural data analysis often assumes neural population activity follows Gaussian distributions (Paninski et al., 2007; Pillow et al., 2008). This assumption enables computationally efficient linear methods such as linear decoders (Georgopoulos et al., 1986), optimal linear estimators (Salinas and Abbott, 1994), and dimensionality reduction techniques like Principal Component Analysis (Cunningham and Yu, 2014). However, neural population activity often exhibits non-Gaussian structure that second-order statistics cannot capture (Macke et al., 2011; Stevenson and Kording, 2011). Assuming Gaussianity implies that pairwise correlations fully characterize all statistical dependencies between neurons. This overlooks higher-order interactions that may be crucial for neural computation, though analyzing such interactions in high-dimensional data remains computationally challenging (Schneidman et al., 2006; Macke et al., 2011).

The precisely timed correlations from Chapter 1 and dynamic representational changes from Chapter 2 suggest neural computation operates across multiple organizational scales. Our geometric framework aims to characterize how neural population activity might be organized on curved manifolds, potentially revealing connections between statistical complexity and geometric structure. We hypothesize that millisecond-precise temporal correlations may reflect the brain’s strategy for navigating curved regions of neural state space, while transient increases in stimulus separability might correspond to trajectories exploiting specific geometric properties of neural manifolds.

We developed a framework combining statistical density estimation with geometric manifold analysis using invertible neural networks. Our hypothesis is that statistical and geometric aspects both contribute to neural population coding, and that a unified framework can provide insights that methods focusing on either aspect alone cannot achieve. Using normalizing flows with interpretable latent representations, we pursue three related goals: extract coordinated activity patterns that might be describable both statistically as correlations and geometrically as manifolds; identify latent variables that could capture these structures; and provide analytical tools for characterizing statistical properties through characteristic functions while describing manifold geometry through coordinate charts.

Our analysis suggests an interesting pattern that forms a central contribution of this work: *different behavioral states appear to correspond to neural manifolds with distinct geometric properties*. In our dataset, active visual processing during eyes-open periods

seems to occur on more curved manifolds, potentially indicating complex, nonlinear neural coordination. Resting states during eyes-closed periods appear to lie on flatter surfaces, possibly suggesting simpler coordination patterns dominated by pairwise correlations rather than higher-order interactions. While this geometric perspective offers new insights into brain organization, the generality of these findings beyond our specific experimental conditions requires further investigation.

The approach attempts to connect discrete temporal precision from Chapter 1, continuous population dynamics from Chapter 2, and hierarchical statistical structures from Chapter 3 through a unified geometric perspective. Geometric characterization of neural manifolds using normalizing flows represents a natural extension of the statistical learning approach from Chapter 3. Meanwhile, curvature analysis might offer new insights into computational principles underlying the dynamic representations characterized in Chapter 2. This integrated approach could potentially enable more complete characterization of neural population dynamics than methods focusing on statistical or geometric aspects alone.

4.2 Mathematical Framework

4.2.1 Neural Population Coding and Manifold Structure

Neural populations in cortical areas coordinate their activity to encode and process information (Georgopoulos et al., 1986; Shadlen and Newsome, 1998). The collective activity of hundreds to thousands of neurons creates high-dimensional structure. The population signal $x \in \mathbb{R}^N$ reflects both the underlying neural computations and the behavioral states of the organism, where N represents the dimensionality of the neural recording space. This population-level organization connects directly to the mean-field dynamics explored in Chapter 2, where we saw how three macroscopic variables capture essential aspects of population activity.

High-dimensional neural population activity often exhibits low-dimensional structure, suggesting that neural dynamics are constrained to lie on or near lower-dimensional manifolds embedded in the full neural state space $\mathbb{R}^N$ (Stopfer et al., 2003; Mazor and Laurent, 2005). This dimensionality reduction reflects both anatomical constraints from connectivity patterns and functional organization from coordinated computation. The concept of neural manifolds has become central to understanding population dynamics (Churchland et al., 2012; Gallego et al., 2017). Rather than each neuron contributing independently to the population code, neurons tend to covary in systematic ways that create structured patterns in the high-dimensional activity space.

The manifolds created by these patterns can be characterized by their geometric properties, including their intrinsic dimensionality, curvature, and topological features.

This geometric perspective provides a natural connection to the statistical learning framework of Chapter 3, where invertible neural networks learn transformations $f_\theta : \mathbb{R}^N \rightarrow \mathbb{R}^N$ that map complex, high-dimensional data to structured latent spaces $z \in \mathbb{R}^N$, effectively discovering the higher-order statistical dependencies present in the data. The geometric analysis developed in this chapter provides tools for characterizing the curvature and higher-order properties of the underlying manifolds.

4.2.2 Statistical and Geometric Perspectives

Statistical and geometric perspectives on data structure are complementary and mutually informative (Amari and Nagaoka, 2000). While statistical methods focus on probabilistic relationships and dependencies between variables through probability distributions $p_x(x)$, geometric methods characterize the shape and curvature of the data manifold. The fundamental insight driving our approach is that these perspectives might be intimately connected: higher-order statistical correlations could correspond to curved geometric structure, as both potentially reflect deviations from simple linear relationships that would be captured by Gaussian statistics.

Cumulants provide a systematic way to characterize deviations from Gaussianity (Gardiner, 2009; Helias and Dahmen, 2020). While the first two cumulants are the mean and covariance, higher-order cumulants capture additional statistical structure. Third-order cumulants measure asymmetry in the distribution and can indicate directional biases in neural activity patterns, while fourth-order cumulants measure the heaviness of distribution tails and can reveal the presence of rare, high-amplitude events or sparse activity patterns. For any Gaussian distribution, all cumulants beyond the second order are exactly zero, so non-zero higher-order cumulants provide direct evidence for non-Gaussian structure in neural data (Onken et al., 2009).

This statistical characterization connects to the geometric analysis developed in this chapter: higher-order cumulants might correspond to curved geometric structure, as both potentially reflect deviations from Gaussian statistical assumptions. The geometric perspective provides a complementary way to understand these same deviations through manifold curvature properties. Conversely, flat manifolds tend to exhibit approximately Gaussian statistics, providing a potential bridge between the statistical learning framework of Chapter 3 and the geometric analysis developed here.

4.2.3 Limitations of Current Approaches

Current dimensionality reduction techniques such as Principal Component Analysis (PCA), Independent Component Analysis (ICA), and Linear Discriminant Analysis (LDA) assume linear relationships between observed variables and underlying factors (Jolliffe, 2002; Hyvärinen and Oja, 2000; Hastie et al., 2009). These methods

have provided valuable insights into neural population organization, revealing low-dimensional structure and identifying coordinated activity patterns across neural populations. However, by making linear assumptions, these approaches correspond primarily to second-order statistical relationships and assume the data manifold is flat everywhere. While computationally efficient and interpretable, this limits their ability to capture complex, nonlinear structure and higher-order statistical dependencies that may characterize neural population activity. We extend beyond this foundational work by developing methods capable of characterizing both curved manifold geometry and non-Gaussian statistical structure.

Neural data often exhibit multimodal structure, with distinct modes corresponding to different behavioral or physiological states (Churchland et al., 2012; Cunningham and Yu, 2014). For example, neural population activity during awake, alert states may form one cluster in high-dimensional activity space, while activity during quiet rest forms a separate cluster.

Gaussian Mixture Models provide a framework for modeling multimodal distributions as weighted combinations of Gaussian components (McLachlan and Peel, 2000). While GMMs could theoretically approximate complex structure by using many components per behavioral mode, this becomes computationally intractable and loses interpretability. When using a practical number of components to maintain model simplicity, the Gaussian restriction may inadequately capture the rich statistical patterns within each behavioral mode. This limitation motivates the normalizing flow approach developed in this chapter, which extends the invertible neural network framework from Chapter 3 to enable flexible modeling of non-Gaussian structure within each behavioral mode while maintaining interpretability.

4.2.4 Normalizing Flows: A Unified Approach

Normalizing flows represent a class of deep generative models that learn invertible transformations between complex data distributions and simple latent distributions (Dinh et al., 2015; Rezende and Mohamed, 2015). This approach directly extends the invertible neural network framework developed in Chapter 3 to the geometric analysis of neural manifolds. While Chapter 3 focused on quadratic nonlinearities to enable extraction of physically interpretable actions, the geometric analysis in this chapter employs more general nonlinearities that can capture richer manifold structures while maintaining invertibility. Additionally, we encourage low-dimensional interpretable latent representations through an auxiliary training objective, enabling identification of the most important dimensions for characterizing neural population structure.

Normalizing flows enable both density estimation, and sampling and, unlike other generative models, provide exact likelihood computation. A normalizing flow defines a bijective mapping $f_\theta : \mathbb{R}^N \rightarrow \mathbb{R}^N$ parameterized by neural network weights θ .

The transformation maps complex data x to latent variables z that follow a simple distribution $p_z(z)$:

$$\begin{aligned} z &= f_\theta(x) \\ x &= f_\theta^{-1}(z) \end{aligned}$$

The probability density estimator is computed using the change of variables formula:

$$\hat{p}_{x;\theta}(x) = p_z(f_\theta(x)) \left| \det \left(\frac{\partial f_\theta}{\partial x} \right) \right|,$$

where $|\det(\partial f_\theta/\partial x)|$ represents the absolute value of the Jacobian determinant that accounts for the volume change induced by the transformation.

Volume-preserving flows (Dinh et al., 2015) represent a special class of invertible networks that maintain constant Jacobian determinants ($|\det(\partial f_\theta/\partial x)| = \Delta V_\theta$, independent of x), where ΔV_θ represents the volume scaling factor, simplifying this to:

$$\hat{p}_{x;\theta}(x) = p_z(f_\theta(x)) \Delta V_\theta.$$

Volume-preserving flows are particularly well-suited for neural data analysis because they preserve the multimodal structure of the data distribution while enabling complex nonlinear transformations.

The training of normalizing flows proceeds by maximizing the log-likelihood of observed data:

$$L_\Pi(\theta) = -\frac{1}{N} \frac{1}{|X|} \sum_{x \in X} \log(\hat{p}_{x;\theta}(x)), \quad (4.1)$$

where X represents the dataset providing training samples and $|X|$ denotes the number of samples. This objective encourages the model to assign high likelihood to observed data points while maintaining a valid probability distribution (Papamakarios et al., 2021). For the neural manifold analysis developed in this work, we extend this training objective by adding an additional term that enforces geometric constraints, ensuring that the learned latent representations preserve meaningful manifold structure. This modification is crucial for achieving interpretable a geometric characterization of the neural population activity.

4.3 Neural Data and Initial Analysis

4.3.1 Dataset and Preprocessing

We analyze comprehensive electrophysiological recordings from visual cortices V1 and V4 of macaque monkeys, recorded by Chen et al. (2022). The dataset contains

1024 channels from 16 Utah arrays (1.5 mm electrodes, Blackrock Microsystems): 14 arrays in V1 (primary visual cortex) and 2 arrays in V4 (visual area 4), primarily targeting layer 5.

We follow the preprocessing pipeline from ? for multi-unit activity envelope (MUAe) extraction. The MUAe is obtained through several steps: band-pass filter the raw signal between 0.5 and 9 kHz to isolate action potential components; full-wave rectify by taking the absolute value; low-pass filter at 200 Hz using a 4th-order Butterworth filter to extract the smooth temporal envelope of population spiking activity. This processing yields a continuous measure of local neural activity, bridging discrete spike trains analyzed in Chapter 1 with the continuous dynamics framework developed here.

Additional preprocessing includes band-stop filtering at 50, 100, and 150 Hz to remove European electric grid noise and harmonics. We remove electrodes with signal-to-noise ratios below 2 or those contributing to high-frequency crosstalk (according to ?), excluding around 200 electrodes per session. We further downsample MUAe to 100 Hz using finite impulse response filter-based decimation, remove edge samples to eliminate boundary artifacts, and apply z-scoring to each channel independently. We concatenate normalized signals from all arrays into a unified matrix for subsequent analysis.

Chen et al. (2022) tracked the monkey's eye state (open or closed) using camera monitoring during resting state recordings with no behavioral task. The findings of ? demonstrate that these two behavioral states correspond to distinct neural manifolds when analyzed using Gaussian mixture models. Here, we go beyond their work and conduct a more detailed examination of these two manifolds using our geometric framework.

4.3.2 Evidence for Non-Gaussian Structure

Figure 4.1 presents dataset characteristics using data from an exemplary recording session (L_RS_250717). The dataset encompasses 774 signal channels, each representing z-scored multi-unit activity envelope (MUAe) data. MUAe represent aggregate spiking activity from multiple neurons recorded at a single electrode, providing a mesoscopic view of local population dynamics (Buzsáki, 2004; Einevoll et al., 2013).

Figure 4.1b shows temporal dynamics of the ten principal components with highest eigenvalues, along with eye state labels. Several principal components exhibit clearly distinguishable patterns between eye states. The rapid decay of PC eigenvalues demonstrates that few variables account for most variance in the data, providing evidence that the data reside on a low-dimensional manifold.

Figure 4.1c shows marginal distributions for the first ten principal components. The data conform neither to Gaussian nor to simple two-component Gaussian mixture

distributions. Several PCs display near-Gaussian marginal distributions (PCs 5, 7, and 10), while others show marked deviations from Gaussianity, particularly the first four PCs which exhibit heavy-tailed distributions. These deviations from Gaussianity are quantified systematically using fourth-order cumulants.

The two-dimensional probability density marginalization along the first two PCs (Figure 4.1d) exhibits bimodal characteristics, consistent with successful Gaussian mixture model fitting by ?. However, the overall distributional shape differs dramatically from a Gaussian mixture. Rather than exhibiting near-elliptical contours, probability density concentrates within a clearly defined triangular region, indicating substantial deviations from Gaussian statistics.

These visualizations provide initial evidence for non-Gaussian statistics. Higher-order cumulants offer a more rigorous approach to quantifying such deviations, capturing statistical dependencies within high-dimensional spaces. Since cumulants beyond second order equal zero for any Gaussian distribution, non-zero values indicate deviations from Gaussianity. Our analysis reveals substantial third- and fourth-order cumulants in this dataset (Figure 4.1e-f). Higher-order cumulants approach zero for near-Gaussian PCs 5, 7, and 10, but show distinctly negative (blue) or positive (red) values for other PCs, particularly PCs 1 and 2. Pronounced positive values along the fourth-order cumulant diagonal suggest heavy tails in these directions.

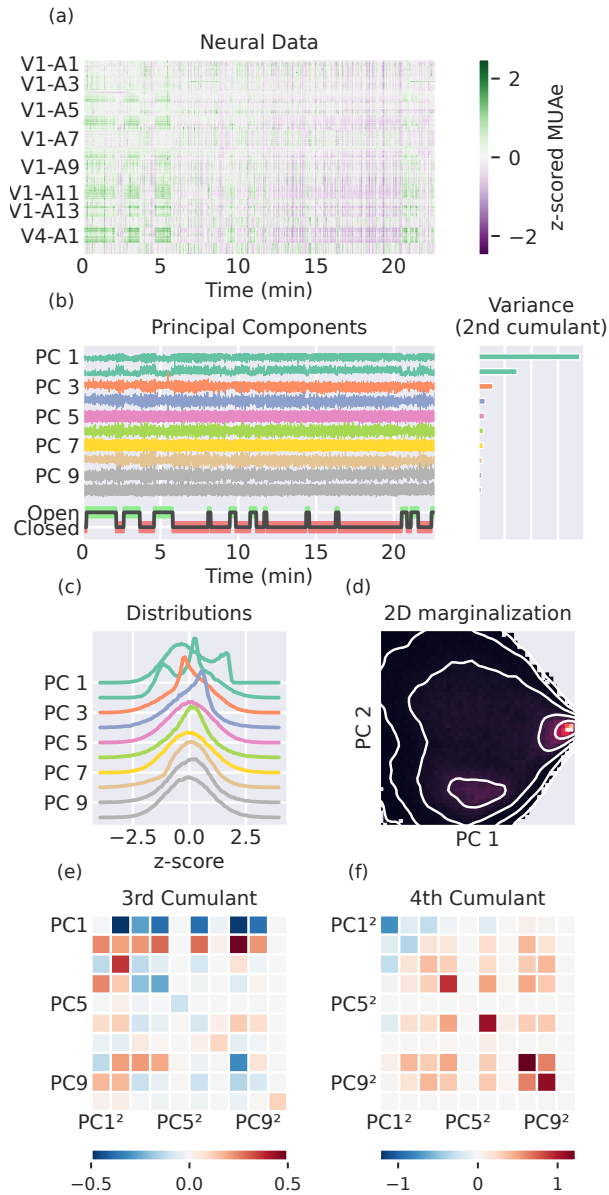


Figure 4.1: Neural population activity demonstrates non-Gaussian statistical structure and higher-order cumulants.

Figure 4.1: Neural population activity demonstrates non-Gaussian statistical structure and higher-order cumulants. (a) Multi-unit activity envelope (MUAe) data from recording session L_RS_250717, acquired at 100 Hz sampling rate over 22.7 minutes of recording time. The data comprise 774 channels distributed across 16 Utah electrode arrays (14 positioned in V1 (primary visual cortex), 2 in V4 (visual area 4)), with each channel normalized independently through z-scoring (zero mean, unit variance) for consistent amplitude scaling. (b) Temporal dynamics of the ten principal components (PCs) with largest variance from the normalized data. Raw principal component (PC) traces are displayed as transparent lines, with temporally smoothed versions (Gaussian kernel, 1 s bandwidth) shown as solid lines. Behavioral state annotations (eyes open/closed) appear beneath the traces. The eigenvalue spectrum σ_i^2 (where σ_i^2 represents the variance explained by the i -th principal component) from the covariance matrix is plotted on the right, demonstrating rapid decay that indicates low-dimensional manifold structure. (c) Univariate probability distributions for the first ten principal components. Several components exhibit near-Gaussian characteristics, while others display pronounced non-Gaussian features. (d) Bivariate probability density projection onto the first two principal component dimensions. (e-f) Third-order and fourth-order cumulant structure visualization. Prior to cumulant computation, PC signals undergo normalization by their respective standard deviations σ_i (where σ_i is the standard deviation of the i -th principal component) to account for amplitude differences. Panel (e) displays normalized third-order cumulants (x_i/σ_i versus x_j^2/σ_j^2 , where x_i represents the i -th principal component value), while panel (f) shows fourth-order cumulants (x_i^2/σ_i^2 versus x_j^2/σ_j^2). Colormap intensity reflects cumulant magnitude values.

4.3.3 Biological Context and Multimodal Structure

These analyses establish that the data contain structured, non-Gaussian statistical properties. As highlighted in Section 0.4.2, such higher-order correlations are intimately connected to curvature properties of underlying manifolds. This relationship underscores the necessity for methodological approaches capable of capturing these structures to achieve a more comprehensive understanding of neural representations.

Neural data often exhibit multimodal structure reflecting different behavioral states during recording (Niell and Stryker, 2010). In our dataset, the macaque alternated between eyes-open and eyes-closed periods, creating two distinct behavioral conditions. The bimodal probability density in the high-dimensional data space suggests these modes correspond to distinct system states. Each mode’s center serves as an archetypal representation of its corresponding behavioral state, provided the mode’s spatial extent remains sufficiently compact. Our geometric framework provides tools for characterizing how these behavioral modes correspond to different manifold geometries, revealing organizational principles invisible to traditional correlation-based approaches.

The non-Gaussian and multimodal structure observed in our data provide compelling motivation for the geometric analysis framework developed in the subsequent sections. Substantial deviations from Gaussianity indicate neural populations exhibit complex coordination patterns that traditional second-order methods cannot capture. Meanwhile, multimodal structure suggests different behavioral states may correspond to fundamentally different manifold geometries.

4.4 Normalizing Flow Architecture and Training

4.4.1 Volume-Preserving Architecture Design

Our implementation utilizes volume-preserving NFs (Dinh et al., 2016; Draxler et al., 2024), constructed from network components maintaining constant Jacobian determinant. This architectural choice is beneficial for neural data analysis because volume-preserving flows maintain modal structure between data and latent spaces, ensuring equivalent numbers of modes in both domains. Thus, volume preservation allows us to preserve multimodal structure corresponding to different behavioral states.

The network processes probability distributions of individual time samples across all MUAe recording channels simultaneously. As preprocessing, we apply a principal component analysis (PCA) inspired transformation following Cramer et al. (2022), transforming input data onto the empirical covariance matrix eigenbasis while standardizing each component to unit variance. This preprocessing ensures numerical stability and provides a meaningful coordinate system for subsequent processing.

We process only the $N_n = 70$ most significant principal components through the flow transformation, while remaining dimensions bypass flow processing and transfer directly to latent output space. This approach is sensible because the variance of the remaining components is relatively small compared to the principal components, allowing us to treat them as approximately Gaussian. The network consists of ten volume-preserving computational blocks. Each block combines a linear all-to-all mixing transformation with an additive coupling layer (Dinh et al., 2015), employing exponential linear unit (ELU) activation functions and maintaining hidden layer dimensionality of 128. The complete architecture preserves invertibility across all layers, enabling exact likelihood evaluation and efficient probabilistic sampling procedures.

4.4.2 Training Procedure

Training consists of pre-training and main training phases, with parameter initialization designed to ensure numerical stability and convergence efficiency. Orthogonal matrix initialization governs linear all-to-all and additive coupling layers, while

biases undergo uniform sampling from $[-1/N_n, 1/N_n]$, where N_n represents the dimensionality processed by the core network components. This initialization strategy maintains signal propagation integrity and variance preservation throughout the architecture.

Both training phases employ Adam optimization with learning rate 0.001, beta parameters (0.9, 0.999), epsilon $1e-8$, and zero weight decay. Batch processing utilizes 10,000 time points consistently throughout training. Pre-training involves 100 epochs using standard negative log-likelihood loss to establish basic distributional modeling capabilities. The main training phase incorporates the additional reconstruction error (Eq. 4.2) to ensure an interpretable latent representation.

4.4.3 Architecture Choices for Biological Data

Our architectural design decisions reflect specific biological considerations for neural data analysis. Processing principal components rather than raw channels acknowledges that neural population activity often exhibits correlated structure that can be efficiently captured through dimensionality reduction while preserving essential information. The volume-preserving constraint ensures natural clustering of neural activity patterns corresponding to different behavioral states is maintained throughout the transformation.

Additive coupling layers enable complex nonlinear relationships between neural activity patterns while maintaining invertibility necessary for exact likelihood computation. This is crucial because neural population dynamics involve complex interactions that linear transformations cannot capture. Exponential linear unit activation functions provide smooth nonlinearities that facilitate stable training while enabling the network to capture curved geometric structures underlying neural manifolds.

The batch size and optimization parameters balance computational efficiency with the need to capture temporal structure of neural activity patterns. Pre-training establishes basic distributional modeling capabilities, ensuring the network can accurately represent overall statistical structure of neural population activity before incorporating additional constraints for interpretability.

4.5 Interpretable Latent Representations

4.5.1 The Interpretability Challenge in Neural Data Analysis

While normalizing flows accurately model complex distributions, learned representations may not be interpretable or meaningfully organized (Dinh et al., 2015). Since

the latent distribution is assumed to be Gaussian, which is rotationally symmetric, the model has no inherent preference for any particular orientation in latent space. This leads to latent variables that may exhibit arbitrary rotations or permutations, obscuring their relationship to underlying data structure. For neuroscience applications, interpretable representations are crucial for understanding neural computation mechanisms. We need additional constraints or objectives that encourage a meaningful organization of the latent space, linked to the physically interpretable actions extracted in Chapter 3.

4.5.2 Reconstruction Error Loss for Hierarchical Organization

We address this limitation by incorporating an auxiliary reconstruction loss that imposes hierarchical ordering on latent variables according to their importance for data recovery. This hierarchical organization enables identification of a compact subset of latent dimensions that efficiently characterize underlying data manifold structure. Our approach builds upon reconstruction loss proposed by Bekasov and Murray (2020), but differs by employing uniform weights $p_l = 1/N_r$ for reconstruction errors $R_l(\theta)$ across all l dimensions, rather than their variable weighting scheme.

Our implementation first maps each data point x into the latent representation

$$z = f_{\theta}(x),$$

then selectively zeros all latent dimensions beyond the first l components before reconstructing data through the inverse transformation:

$$\tilde{x}_{\theta;l}(x) = f_{\theta}^{-1}\left(\sum_{i=1}^l f_{\theta,i}(x)\hat{e}_i\right)$$

where $\hat{e}_i$ represents the i -th standard basis vector in latent space, $f_{\theta,i}(x)$ denotes the i -th component of $f_{\theta}(x)$, and l represents the number of latent dimensions retained. For data residing on an l -dimensional manifold, this reconstruction procedure yields perfect recovery, satisfying $\tilde{x}_{\theta;l}(x) = x$ for all data points $x \in X$. The reconstruction loss across the entire dataset for l dimensions follows:

$$R_l(\theta) = \frac{1}{|X|} \sum_{x \in |X|} |x - \tilde{x}_{\theta;l}(x)|^2$$

where $|X|$ denotes the number of samples in dataset X , and $|\dots|^2$ represents the squared L2 norm.

Since intrinsic manifold dimensionality remains unknown *a priori*, we compute the average reconstruction loss across all dimensions l up to a specified maximum N_r :

$$L_{\text{rec}}(\theta) = \gamma_{\text{rec}} \frac{1}{N_r} \sum_{l=1}^{N_r} R_l(\theta) \quad (4.2)$$

where γ_{rec} serves as a hyperparameter controlling the reconstruction term's influence.

The biological motivation for this approach stems from observations that neural population activity often exhibits hierarchical organization. Some coordination patterns are more essential for neural computation than others. By encouraging networks to organize latent variables according to their importance for data reconstruction, we promote representations where the most critical neural coordination patterns are captured by the first few latent dimensions. This hierarchical organization reflects the biological reality that neural systems often employ efficient coding strategies prioritizing the most informative features.

4.5.3 Validation on a Synthetic Dataset

Figure 4.2 demonstrates the impact of this additional loss component using synthetic data with two fundamental latent variables: radial distance and angular phase. The data distribution appears as scattered sample points, while the learned distribution is visualized as grid lines representing the discovered latent coordinate system. Standard normalizing flow training using only negative log-likelihood produces accurate distribution estimates (Figure 4.2a), yet learned latent variables fail to correspond to the task's underlying coordinate system.

Incorporating the reconstruction loss term introduced in the previous section during normalizing flow training (Figure 4.2b) enables the network to recover true underlying variables: radius and phase. This synthetic task further demonstrates dimensionality reduction capabilities: constraining the radial latent variable to zero produces a one-dimensional curve traversing the data center while maintaining reasonable data explanation. Although simple, this two-dimensional example demonstrates our method's three core capabilities: learning arbitrary probability distributions from observed data, aligning latent coordinates with interpretable directions, and identifying reduced-dimensional manifolds through systematic elimination of uninformative dimensions.

The success of this approach on synthetic data provides confidence that the method can discover meaningful coordinate systems for neural manifolds. The ability to recover interpretable coordinates while maintaining accurate distributional modeling suggests that the reconstruction loss effectively balances the competing demands of statistical accuracy and interpretable organization.

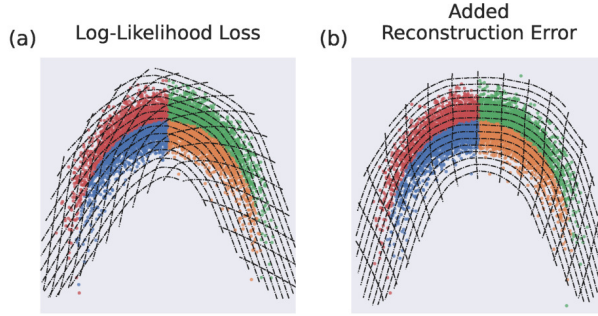


Figure 4.2: Reconstruction loss enables discovery of meaningful latent coordinates in synthetic data. We demonstrate the impact of incorporating reconstruction loss (Eq. 4.2) using a synthetic dataset with inherent two-dimensional structure. The data’s geometric organization naturally suggests representation through two principal latent variables corresponding to orthogonal axes. Color-coding visualizes four quadrants delineated by these axes. We train an invertible normalizing flow network with transformation $x = f_{\theta}^{-1}(z)$, where x represents data points, z denotes latent variables, θ indicates network parameters. Panel (a) shows results from standard training via negative log-likelihood minimization (Eq. 4.1), with latent variables z following $z \sim N(0, \mathbb{I})$ (multivariate normal distribution with zero mean and identity covariance matrix $\mathbb{I}$). Panel (b) shows results when adding a reconstruction error term to the loss function that promotes hierarchical ordering of latent dimensions by reconstruction importance. Both panels visualize the learned coordinate system through grid lines representing $f_{\theta}^{-1}(s \cdot \hat{e}_1 + t \cdot \hat{e}_2)$, where $\hat{e}_1$ and $\hat{e}_2$ denote the first and second latent basis vectors. Grid lines are generated by independently varying each latent variable $s, t \in \mathbb{R}$ while holding the other variable fixed. Adjacent grid lines are separated by half a standard deviation in latent space, with visualization restricted to the region where $t^2 + s^2 \leq 4$. Incorporating reconstruction loss enables accurate identification of the dataset’s intrinsic latent variables: panel (b) shows grid lines aligned with the two fundamental axes, confirming successful recovery of the underlying coordinate system.

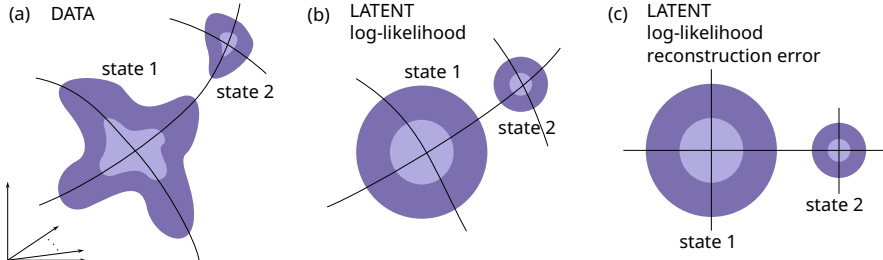


Figure 4.3: **Multimodal distributions in data and latent space.** (a) A bimodal probability density in high-dimensional data space, where the two modes may correspond to distinct system states such as different behavioral conditions. (b) Latent space with bimodal distribution when trained only with negative log-likelihood loss (Eq. 4.1). (c) We observe alignment of sets of latent space axes to the behavioral states due to the added reconstruction error (Eq. 4.2).

4.6 Behavioral State Discovery

4.6.1 Multimodal Structure and Behavioral Alignment Strategy

Neural recordings during different behavioral conditions often produce multimodal data distributions, where each mode potentially corresponds to a distinct behavioral state. Our volume-preserving normalizing flows maintain this modal structure between data and latent spaces, ensuring equivalent numbers of modes in both domains (Dinh et al., 2015; Merger et al., 2023).

We model this multimodal structure using a Gaussian mixture distribution in the latent space (Izmailov et al., 2020). To prevent overfitting while maintaining interpretability, we constrain the mixture components: centers can differ only in the first N_l dimensions, with remaining coordinates fixed at zero, and covariance matrices can deviate from identity only within these same N_l dimensions. The network learns both center coordinates and covariances, while N_l serves as a predetermined hyperparameter.

The training strategy significantly impacts the interpretability of learned representations. Training with negative log-likelihood loss alone produces the configuration shown in Figure 4.3b, where individual latent variables fail to capture behavioral state transitions effectively. However, incorporating reconstruction error as an auxiliary loss term consistently achieves latent variable alignment with behavioral state changes (Figure 4.3c). Although this alignment lacks theoretical guarantees, it demonstrates the model’s tendency toward meaningful latent organization where mode transitions correspond to changes in specific dimensions.

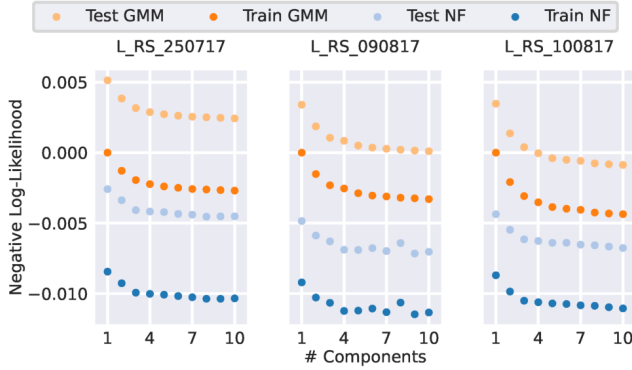


Figure 4.4: **Accuracy of data representation saturates with fewer than 10 latent components.** We evaluate three different recording sessions by measuring negative log-likelihood per dimension for both training and test data as a quality indicator. Normalizing Flows (NFs) are compared against Gaussian Mixture Models (GMMs) to assess the benefits of learning curved representations. NFs demonstrate saturation at 3 to 4 latent components, while GMMs consistently underperform across all component counts. An 80%/20% train-test split remains fixed throughout each recording session analysis.

4.6.2 Model Selection and Component Determination

Given the multimodal framework established above, we now address the important issue of component selection. We apply our normalizing flow framework to the neural activity data presented in Figure 4.1. The most critical hyperparameter is the number of Gaussian mixture components in latent space, which controls the model’s capacity to capture distinct behavioral modes. Each component corresponds to a flat manifold region in latent space that transforms into a curved manifold in data space through the normalizing flow mapping. The component number critically influences model performance: too few components yield suboptimal descriptions that miss important data characteristics, while too many components cause manifold overlap and reduced interpretability.

We evaluate component selection by training networks with varying numbers of mixture components and assessing their log-likelihood performance. Figure 4.4 displays log-likelihood values for up to ten components, with standard Gaussian Mixture Model results included for comparison. Normalizing flows consistently outperform Gaussian Mixture Models, demonstrating the advantages of learning curved representations over linear approaches. The normalizing flow log-likelihood increases monotonically with component number but saturates beyond three or four components, indicating that additional complexity provides diminishing returns.

The superior performance of normalizing flows becomes clear through qualitative evaluation in Figure 4.5, which shows the marginalized probability density along

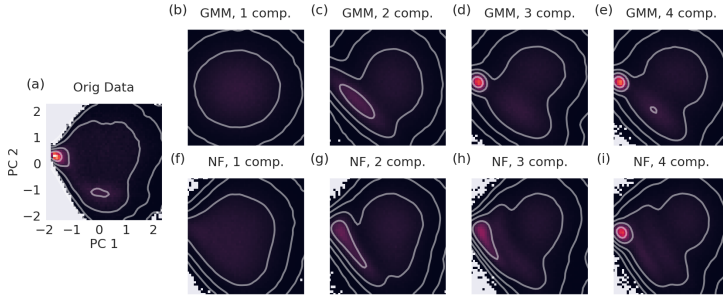


Figure 4.5: Effect of number of latent components on the approximation capabilities: Marginalized distributions for the experimental data and learned models. **(a)** Marginalized distribution of one session of experimental data (L_RS_250717) along the first and second PCA directions. Light-gray isolines represent equal likelihood. **(b–e)** Marginalized probability densities for data sampled from GMMs with increasing numbers of latent components (1 to 4). **(f–i)** Corresponding distributions for NFs, also with 1 to 4 latent components.

the first two Principal Component Analysis variables. The neural data exhibit a stark contrast between small, high-density regions and extensive, low-density areas with triangular probability cutoffs. Gaussian Mixture Models, constrained to elliptical component shapes, cannot accommodate these triangular boundaries, explaining their inferior log-likelihood performance. Normalizing flows excel at modeling such complex boundaries even with few latent components due to their nonlinear transformation capabilities.

Based on these quantitative and qualitative assessments, we select four-component normalizing flows for subsequent analyses, providing optimal balance between modeling accuracy and interpretability.

4.6.3 Discovery of Behavioral State Correspondence

Our dataset contains neural recordings from a macaque alternating between eyes-open and eyes-closed periods, creating two distinct behavioral conditions. Since our four-component model accurately represents the data’s empirical density, we hypothesize that some components should align with these behavioral conditions. Such alignment would demonstrate that different behavioral states generate structurally distinct neural activity patterns, particularly notable since the model was trained without behavioral labels.

We test this hypothesis by computing the posterior probability of each component k given a neural data point x using Bayes' theorem:

$$p(k|x) = \frac{p_k p(x|k)}{\sum_{k'} p_{k'} p(x|k')},$$

where p_k represents component k 's weight and $p(x|k)$ denotes the data point's likelihood under that component. Our normalizing flow approach computes this likelihood using the change-of-variables formula:

$$\hat{p}_{x;\theta}(x|k) = \mathcal{N}(\mu_k, \Sigma_k; f_\theta(x)) \left| \det \left(\frac{\partial f_\theta(x)}{\partial x} \right) \right|, \quad (4.3)$$

where $f_\theta(x)$ maps data x into the latent space with parameters θ . We apply temporal smoothing using a Gaussian kernel (standard deviation 3 seconds) to obtain stable time-varying label probabilities, with dominant component assignment following maximum smoothed probability at each time point.

4.6.4 Validation of Behavioral Alignment

The results reveal correspondence between learned components and behavioral states. Figure 4.6a displays smoothed label probabilities as solid curves, with background shading indicating dominant component assignments. Comparison with true behavioral conditions (Figure 4.6b) shows strong alignment: component C4 corresponds to eyes-open periods, while component C3 aligns with eyes-closed periods. Components C1 and C2 exhibit lower overall probabilities and appear to capture residual variability or transitions between primary states.

Panel c provides quantitative validation through direct comparison of label probability distributions across behavioral conditions. Component C4 shows high probabilities predominantly during eyes-open periods, while C3 peaks during eyes-closed periods, both serving as reliable behavioral indicators. In contrast, C1 and C2 display generally low probabilities, with C1 showing slight preference for eyes-closed conditions. This alignment appears consistently across all recording sessions, confirming our method's ability to extract behaviorally-relevant latent components from neural data.

4.6.5 Biological Interpretation and Implications

The unsupervised discovery of behaviorally-relevant neural population modes provides interesting insights into neural organization. The fact that normalizing flows trained purely on neural activity patterns without behavioral information discovered components that align with behavioral states suggests that different cognitive

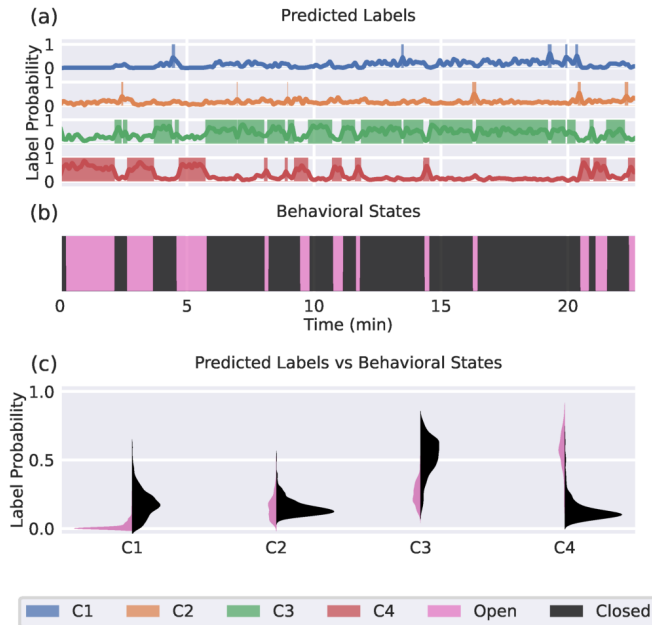


Figure 4.6: Distinct mixture components in the latent space represent behavioral states. (a) Smoothed (Gaussian filter, $\sigma = 3$ s) and normalized likelihoods appear as solid curves for each of the four components (C1–C4). Component assignments result from selecting the component with maximum label probability at each time point, shown through background shading. (b) Ground-truth behavioral states display eyes-open (pink) and eyes-closed (black) conditions for direct comparison with predicted labels in (a). (c) Label probabilities are directly compared against behavioral conditions, with smoothed histograms showing the distribution of label probabilities for each component across the two states (eyes-open: left, eyes-closed: right)

conditions might correspond to distinct neural population dynamics. This observation supports the possibility that behavioral states could be encoded in the statistical structure of population activity patterns, though the generalizability of this finding beyond our specific experimental paradigm requires further investigation.

The four-component structure hints at potential hierarchical organization in neural population dynamics. Components C3 and C4 appear to represent stable states corresponding to eyes-closed and eyes-open conditions, while components C1 and C2 may capture transitional dynamics or mixed states. However, the functional significance of this organization and whether it represents a general principle of neural computation remains to be determined.

This component structure provides the foundation for subsequent geometric analysis by establishing that different behavioral conditions correspond to distinct regions of neural state space. The geometric analysis in following sections will explore whether these behaviorally-distinct components also correspond to manifolds with different curvature properties, potentially revealing connections between behavioral states, statistical structure, and geometric organization in neural population dynamics.

4.7 Statistical Characterization

4.7.1 Quadratic Network Approximation for Analytical Tractability

The trained Normalizing Flow effectively captures complex neural population dynamics, but interpreting its mapping proves challenging due to numerous trained parameters distributed across multiple network layers. We therefore simplify the mapping to facilitate interpretation while preserving essential characteristics. While previous sections examined relationships between high-level behavior and mixture components, we now investigate how latent variables of individual components relate to geometric and statistical properties of the neural activity manifold.

We achieve this by representing the learned mapping for each component using a quadratic function. The quadratic form provides an optimal balance between reduced expressivity and interpretable latent variables, while maintaining latent space structure. This approximation enables analytical computation of statistical and geometric properties, while the full invertible network architecture continues to provide accurate log-likelihood estimation. For each latent component, we represent the inverse transformation $f_\theta^{-1}(z)$ using a quadratic function from latent space z to data space x as:

$$x = q(z) = \sum_{\alpha} \left(\frac{1}{2} \sum_{i,j} A_{ij}^{\alpha} z_i z_j + \sum_i B_i^{\alpha} z_i + c^{\alpha} \right) \hat{e}^{\alpha} \quad (4.4)$$

Upper indices α denote data space dimensions with $\hat{e}^\alpha$ representing the corresponding basis vectors. Lower indices i, j refer to latent space dimensions. Parameters c^α , B_i^α , and A_{ij}^α represent bias, linear, and quadratic terms, respectively. We omit component-specific indices for clarity, noting that $A_{ij}^\alpha = A_{ji}^\alpha$ by symmetry.

This function represents the simplest nonlinear model capable of capturing quadratic interactions among latent variables. Remarkably, we observe that the trained deep network can be accurately represented using this quadratic form. Qualitative validation appears by comparing the upper column (panels a-d) with the middle column (panels e-h) in Figure 4.7, where the inverse transformation and its quadratic representation are respectively applied. Quantitative evaluation of the fitting procedure appears in Appendix A. The close correspondence between these representations demonstrates that despite the network’s capacity to encode vastly more complex function spaces, the quadratic approximation captures the essential structure learned during training.

The success of the quadratic function representation is noteworthy, considering that deep networks can encode vastly more complex function spaces. Whether this representation remains valid depends generally on both dataset characteristics and network architecture. For the neural recordings examined here, it provides a valuable simplification that yields important insights into neural coding, as demonstrated in subsequent sections. The biological interpretation of this finding suggests that neural population interactions, while complex, may be fundamentally characterized through pairwise variable interactions, providing a tractable mathematical framework for understanding coordination patterns.

4.7.2 Validation and Error Analysis of Quadratic Approximation

We describe the procedure for representing each latent component’s inverse transformation with a quadratic function and its validation. This representation aims to simplify interpretation of the complex transformation implemented by a trained Normalizing Flow while maintaining accuracy for subsequent statistical and geometric analysis. The procedure begins with latent components resulting from the training process, where each component is characterized by its mean and covariance parameters. Samples drawn from these components undergo transformation back to data space through the inverse flow.

We perform quadratic fitting for each mixture to represent the inverse transformation from latent space to data space. We construct polynomial features (linear and quadratic terms) for each latent mixture to fit all coefficients in the quadratic mapping equation. Feature construction utilizes polynomial feature expansion from scikit-learn, while linear regression models fit these polynomial features against corresponding data mixtures, capturing both linear and quadratic interactions among

latent variables. The parameters obtained from fitting include bias terms, linear coefficients, and quadratic coefficient matrices, which are crucial for understanding how latent variable changes affect neural population activity.

Empirical analysis shows that cumulants up to order 4 are well-represented by the quadratic function across all components: although deviations occur, particularly for higher-order off-diagonal components, the represented statistics maintain strong correlation with true values. This finding demonstrates that despite reduced expressivity compared to the full network, the quadratic model preserves key statistical properties of neural activity. The close correspondence between original and represented cumulants indicates that important interactions among latent variables are maintained, providing confidence for using this simplified model in further interpretation and analysis.

The validation demonstrates that the quadratic approximation successfully captures the essential statistical structure while dramatically simplifying the mathematical complexity. This simplification enables the analytical computation of moment-generating functions and cumulants that would be intractable for the full network architecture. The preservation of statistical accuracy combined with analytical tractability makes the quadratic approximation particularly valuable for neuroscience applications where interpretability is crucial.

4.7.3 Higher-Order Statistical Analysis Through Cumulants

Neural code understanding requires moving beyond second-order Gaussian descriptions to achieve complete characterization of neural coordination statistics. This becomes evident from the non-Gaussian patterns observed in our data (Figure 4.1). Higher-order interactions can fundamentally shape population dynamics, as demonstrated by previous studies (Montani et al., 2009; Ganmor et al., 2011). The analytically tractable network representation from the quadratic approximation enables computation of means, covariances, and crucially, third- and fourth-order cumulants for each latent component.

These quantities derive from the characteristic function of the approximate network mapping, providing powerful insights into the encoded statistical patterns. These statistics quantify the degree to which neural interactions deviate from Gaussianity within each component. We derive the moment-generating function (Gardiner, 2009; Helias and Dahmen, 2020) from the quadratic representation (Eq. 4.4) of the network mapping for each component as:

$$Z(j) = \frac{e^{j^T c} e^{\frac{1}{2} j^T B (\mathbf{I} - \sum_{\alpha} A^{\alpha} j^{\alpha})^{-1} B^T j}}{\det(\mathbf{I} - \sum_{\alpha} A^{\alpha} j^{\alpha})^{1/2}}$$

where $j = (j^1, \dots, j^N)^T \in \mathbb{R}^N$ is the N -dimensional source vector, c denotes the bias vector, B contains the linear coefficients, and A^{α} are the quadratic coefficient matrices

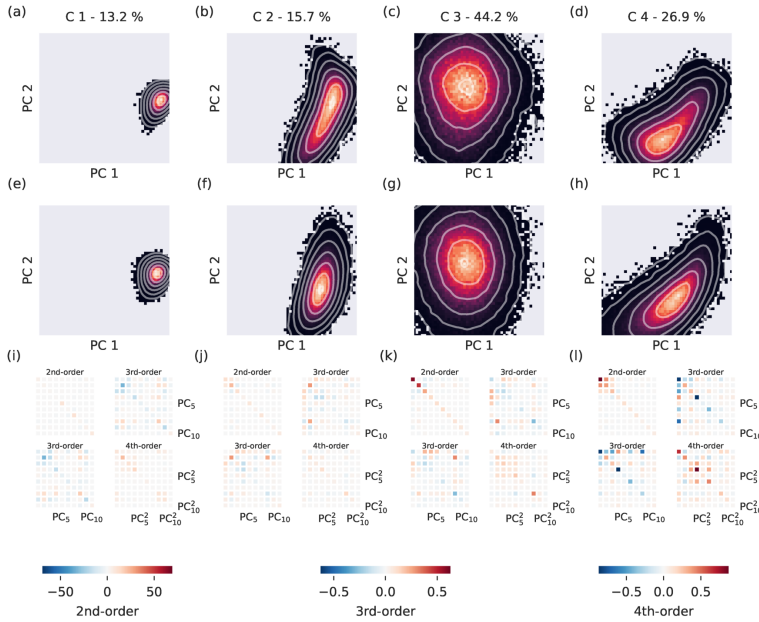


Figure 4.7: Statistical characteristics of the latent components: Approximate marginalized distributions for experimental data and learned models. **(a-d)** Marginalized distributions of each latent component from the experimental session (L_RS_250717), projected onto the first two PCA directions. Light-gray contours indicate isolines of equal likelihood. **(e-h)** Marginalized distribution of each component when the inverse mapping of the network is approximated by a quadratic function (Eq. 4.4). **(i-l)** 2nd to 4th order cumulants computed from this approximation.

for dimension α . This expression captures how the quadratic transformation of latent variables generates moments in the data space.

We also consider the cumulant-generating function $W(j) = \ln Z(j)$, which is provided by the logarithm of the moment-generating function. Detailed derivations for both generating functions appear in Appendix B.

The resulting cumulants serve as effective measures of distributional non-Gaussianity, since all cumulants beyond second order vanish for Gaussian distributions. Figure 4.7e-h displays cumulants for all components, following the visualization scheme from Figure 4.1 by showing complete second-order cumulant values but a subset of third- and fourth-order cumulant values.

Component C4 exhibits pronounced third- and fourth-order cumulants, marking it as non-Gaussian. Components C1 through C3 also demonstrate moderate higher-order patterns.

4.7.4 Analytical Computation via Moment-Generating Functions

We compute cumulants of data variables by representing the trained Normalizing Flow through the second-order quadratic function of latent variables. For each mixture component, latent variables follow a multivariate Gaussian distribution. This structure allows analytical evaluation of the moment-generating function, and subsequently, the cumulant-generating function obtained as the logarithm of the moment-generating function. The moment-generating function and characteristic function are closely related: the characteristic function is obtained by substituting imaginary arguments into the moment-generating function, providing equivalent information about the distribution's statistical properties (Helias and Dahmen, 2020).

Cumulants are extracted as mixed partial derivatives of the cumulant-generating function evaluated at zero. This approach leverages the fundamental property that cumulants measure deviations from Gaussian behavior: for Gaussian distributions, all cumulants beyond second order vanish. Working with the quadratic mapping from latent space to data space, where latent variables are standardized for each mixture component, we derive explicit expressions for cumulants of different orders.

The mathematical framework enables computation of first-order cumulants (means), second-order cumulants (covariances), third-order cumulants measuring asymmetries, and fourth-order cumulants capturing tail behavior. The explicit cumulant formulas involve systematic combinations of the fitted quadratic parameters, providing direct connections between the learned network structure and the statistical properties of neural population activity.

4.7.5 Biological Interpretation of Statistical Complexity

Our results show that the components inferred by the model exhibit distinct signatures, with several showing strong non-Gaussian features across multiple dimensions. This highlights the importance of capturing higher-order structure to completely describe neural population activity across behavioral states. The connection between statistical complexity and behavioral states suggests that different cognitive conditions require different types of neural coordination patterns.

Component C4, which corresponds primarily to the eyes-open behavioral state, exhibits the strongest deviations from Gaussianity through pronounced third- and fourth-order cumulants. This suggests that active visual processing involves complex, hierarchical interactions among neural populations that cannot be captured by traditional correlation-based approaches. In contrast, the components associated with eyes-closed states show more moderate statistical complexity, indicating that resting neural activity involves simpler, more predictable coordination patterns.

The statistical characterization provides crucial groundwork for the geometric analysis developed in the following section. The same quadratic coefficients that generate these higher-order statistical dependencies also determine the curvature properties of neural manifolds, establishing a fundamental connection between statistical complexity and geometric structure. This relationship suggests that understanding neural computation requires tools capable of capturing both non-Gaussian statistics and nonlinear geometry, where our unified framework extracts both aspects from the learned network representation.

4.8 Geometric Analysis

4.8.1 Geometric Perspective on Neural Computation

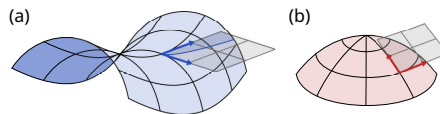


Figure 4.8: **Geometric interpretation of manifold curvature.** Tangent vectors (arrows) define the local tangent space (gray plane) at a specific point on the data manifold. **(a)** Negative sectional curvature: the manifold curves away from the tangent plane in opposite directions, creating a saddle-like geometry where the tangent plane intersects the manifold surface. **(b)** Positive sectional curvature: the manifold curves away from the tangent plane in the same direction, creating a sphere-like geometry where the tangent plane lies entirely outside the manifold surface.

Understanding neural coding principles requires characterizing not only the statistical structure of neural activity but also the geometric shape of population activity manifolds. Neural population activity doesn't just encode information statistically – it also has geometric structure that reveals organizational principles invisible to traditional statistical analysis. Neural states can be conceptualized as points on a curved surface within high-dimensional space: some regions exhibit positive curvature, others display negative curvature or saddle-point geometries (Figure 4.8). This geometric perspective reveals how the brain organizes neural states in ways that reflect underlying computational processes.

The key insight is that curvature – how much the neural manifold bends – reflects the underlying computational structure. Flat regions suggest linear relationships between neural variables, while curved regions indicate complex, nonlinear interactions between neurons. Neural activity patterns typically constrain to complex, nonlinear subspaces within high-dimensional embedding spaces, where curvature serves as a fundamental indicator of such nonlinearity. Having analyzed the various manifolds in our data from a statistical perspective, we now adopt a geometrical viewpoint that provides complementary insights into how the brain organizes neural states.

This geometric perspective proves particularly relevant given the higher-order statistical dependencies identified in the previous section, since curvature can capture underlying nonlinear interactions that second-order statistics overlook. The connection between statistical complexity and geometric structure suggests that regions of the neural manifold exhibiting strong non-Gaussian statistics should also display pronounced curvature properties.

4.8.2 From Normalizing Flows to Manifold Curvature

Our approach establishes direct correspondence between each learned latent component and its associated submanifold in high-dimensional neural activity space. The analyses presented here parameterize each latent component using $N_l = 10$ latent variables, as established in the training procedure. We characterize each mixture component's geometry separately by examining its curvature properties around the point of maximum likelihood.

The quadratic coefficients A_{ij}^α serve as the fundamental building blocks for all geometric analysis in this chapter. These coefficients, obtained from the quadratic approximation of the inverse normalizing flow, directly encode the second derivatives of the embedding mapping and therefore contain all information necessary to compute intrinsic curvature properties of the neural manifold. The quadratic coefficients serve a dual role: they generate statistical complexity as characterized in the previous section and geometric curvature as detailed here. This connection suggests a fundamental relationship between statistical and geometric structure in neural activity.

We construct this analysis using the quadratic approximation of the inverse normalizing flow mapping obtained after training, from which we develop local coordinate charts. These charts establish a smoothly varying atlas that enables computation of intrinsic geometric quantities: scalar curvature, which quantifies overall average curvature, and sectional curvature, which measures manifold bending at specific points relative to two-dimensional tangent planes. Both measures illuminate how neural representations deviate from simple, flat geometries and reveal insights into the complexity of the underlying neural code.

4.8.3 Local Geometry Construction and Curvature Measures

For sectional curvature analysis, we examine pairs of latent variables to determine the curvature of the two-dimensional surface they span on the data manifold. We establish a local coordinate system on the data manifold using tangent vectors, which describe how perturbations in latent variables influence data space coordinates and derive from the quadratic network approximation. These tangent vectors include both linear terms that capture manifold orientation at the reference point and quadratic terms that encode orientation changes away from the reference point, generating curvature through the second derivatives of the embedding mapping.

From these tangent vectors, we construct the metric tensor, which defines the inner product on tangent space and encodes local geometric information including distances and angles between directions (Lee, 2018). This metric tensor proves essential for all subsequent geometric analysis steps and enters calculations both directly and through its inverse. We then introduce the projection operator onto the normal space – the space orthogonal to the tangent plane. This operator ensures curvature calculations respect only deformations perpendicular to the manifold.

The second fundamental form emerges naturally from our quadratic coefficients. For an embedded submanifold, the second fundamental form measures how the manifold curves away from its tangent space and incorporates the quadratic coefficients as the fundamental building blocks. The crucial insight is that our quadratic coefficients directly encode the second derivatives of the embedding mapping, making them the essential components for all curvature calculations. The connection to curvature arises through the Gauss equation, which relates intrinsic curvature to the second fundamental form.

Based on these definitions, we compute sectional curvature $K(e_i, e_j)$ (Lee, 2018 Ch.8), which quantifies manifold bending in the plane spanned by tangent vectors e_i and e_j :

$$K(e_i, e_j) = \frac{\sum_{\alpha, \beta} P^{\alpha\beta} (A_{ii}^\alpha A_{jj}^\beta - A_{ij}^\alpha A_{ij}^\beta)}{g_{ii}g_{jj} - (g_{ij})^2}$$

We obtain scalar curvature R by averaging over all sectional curvatures at a point (Lee, 2018 Ch.7). For our model, this produces the expression:

$$R = \sum_{j,l,c,m=1}^{N_l} \sum_{\alpha,\beta=1}^N g^{jl} g^{cm} P^{\alpha\beta} (A_{jl}^{\alpha} A_{mc}^{\beta} - A_{cj}^{\alpha} A_{ml}^{\beta})$$

The complete derivation of this scalar curvature formula is presented in the detailed mathematical derivations in Appendix C, where we systematically construct it through the Ricci curvature tensor.

4.8.4 Geometric Signatures Across Brain States

	Sectional curvature
Component 1	-0.052
Component 2	-0.011
Component 3	-0.012
Component 4	-0.097

Table 4.1: Scalar curvature for all four components.

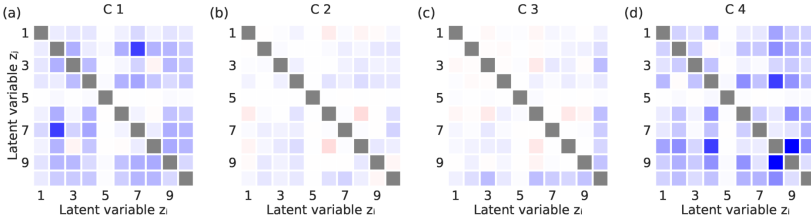


Figure 4.9: **Geometric characteristics of the latent components.** Panels (a-d) show sectional curvature matrices for each of the four latent mixture components, where each matrix element $K(e_i, e_j)$ represents the curvature in the two-dimensional plane spanned by latent basis vectors e_i and e_j . These values are computed from the quadratic approximation of the inverse normalizing flow mapping. Red indicates positive sectional curvature, blue indicates negative sectional curvature. The color scale is consistent across all panels, with the range spanning the scalar curvature values reported in Table 4.1 (scalar curvature represents the average of all sectional curvatures at each point).

Our analysis reveals a striking pattern: the complexity of neural geometry mirrors the complexity of cognitive states. We measure scalar curvature across different latent components as shown in Table 4.1. Component C4 exhibits the most pronounced negative curvature (-0.097), followed by component C1 (-0.052), both indicating saddle-like manifold structures. Conversely, components C2 and C3 display nearly flat geometry with weak negative curvature (-0.011/-0.012), suggesting minimal deviation from flatness.

As demonstrated in Figure 4.6, component C3 aligns predominantly with the eyes-closed state and represents a component with moderate higher-order correlations (Figure 4.7). We now observe that it also corresponds to a component lying on a near-flat manifold. In contrast, component C4, which aligns primarily with the eyes-open state (Figure 4.6), corresponds to the most curved sub-manifold and shows the strongest deviations from Gaussianity (Figure 4.7). The two components that do not clearly align with specific states but rather account for transitions between or periods separating the states (Figure 4.6) exhibit saddle-point-like (C1) and near-flat (C2) geometries.

A more comprehensive picture emerges from the sectional curvature matrices shown in Figure 4.9, where curvature contributions from different pairs of latent variables are visualized. For components C1 and C4, sectional curvature remains predominantly negative (blue-shaded entries), supporting the interpretation of globally saddle-point-like structures. In components C2 and C3, however, curvatures display more balanced distributions with both weakly positive and negative values, aligning with near-zero scalar curvature and suggesting less distinctly curved manifolds.

These results indicate that curvature distributes across the entire latent space rather than confining to exceptional latent variables. The manifold’s curvedness reflects joint interactions among multiple latent variables rather than isolated effects from individual variables. This distributed nature of curvature suggests that neural population coordination involves complex, multidimensional interactions that cannot be understood by examining individual neural variables in isolation.

4.8.5 Biological Interpretation of Manifold Curvature

We observe that the component exhibiting the strongest higher-order statistical correlations resides on the most curved manifold. This alignment between statistical complexity and geometric structure supports the possibility that understanding neural coordination might require tools capable of capturing both non-Gaussian statistics and nonlinear geometry, where our method extracts both aspects from a unified framework.

The connection between our computational approach and classical differential geometry reveals several important relationships for understanding neural manifolds. The same quadratic coefficients that generate higher-order statistical cumulants also determine geometric curvature, suggesting a fundamental relationship between statistical complexity and geometric structure in neural activity patterns. Regions of positive curvature correspond to neural states with sphere-like geometric configurations, while negative curvature regions represent saddle-like geometries. Near-zero curvature indicates neural states lying on nearly flat regions of the manifold.

Our observation that eyes-open states (Component C4) reside on highly curved manifolds while eyes-closed states (Component C3) lie on nearly flat surfaces suggests

that active visual processing might involve more complex, nonlinear neural coordination patterns. This geometric perspective provides one lens for examining how the brain might organize different computational states, though the relationship between neural geometry and cognitive complexity requires further investigation.

The observed relationship between curvature and brain state hints that manifold geometry could play a role in organizing neural population dynamics. Active sensory processing appears to correspond to curved geometric structures, while resting states correspond to simpler flat geometries. Whether this geometric organization reflects computational efficiency principles or represents a more general strategy for managing different computational demands remains an open question.

4.8.6 Implications for Neural Coding and Computation

These findings suggest that geometric analysis might reveal organizational principles of neural computation that are less apparent in purely statistical approaches. The observed relationship between curvature and brain state hints that manifold geometry could serve as an organizing principle for neural population dynamics, though the broader implications for understanding brain function and potential disruptions in neurological conditions require further investigation.

The geometric characterization provides potential insights into how neural populations coordinate their activity to support different cognitive functions. The curved manifolds associated with active visual processing might reflect the complex nonlinear relationships present during sensory processing and perception. In contrast, the flatter manifolds associated with resting states could indicate simpler geometric organization during periods of reduced sensory input.

This geometric perspective also suggests possible directions for understanding neural computation across different brain areas and behavioral conditions. The framework developed here could potentially be applied to characterize manifold geometry in other brain regions, during different behavioral tasks, or in different species. However, whether the connection between statistical complexity and geometric structure represents a general principle of neural organization or is specific to the visual cortical areas and experimental paradigm analyzed in this study remains to be determined.

4.9 Discussion and Implications

We discovered that neural activity exhibits fundamentally different geometric organization depending on behavioral state. During active visual processing, population activity traces curved, saddle-shaped manifolds. During rest, the same populations organize into nearly flat geometric structures. This state-dependent geometry suggests that manifold curvature reflects computational demands.

4.9.1 Behavioral State Discovery Through Geometric Structure

By extending normalizing flows (Dinh et al., 2015; Kobyzev et al., 2020; Papamakarios et al., 2021) to neural data, we gain insights into how neural populations organize information. Through unsupervised training, the learned latent components align with behavioral states observed during experiments, demonstrating the capacity of generative models to discover neural organization.

Remarkably, our method discovered this behavioral alignment without any training labels. The geometric differences between eyes-open and eyes-closed states emerges purely from the statistical structure of neural activity, indicating that behavioral demands shape the geometric organization of neural activity (Niell and Stryker, 2010; Stringer et al., 2019b).

Distinct Geometric Signatures Across Behavioral States

The geometric differences between behavioral states prove quantitatively distinct. Neural activity during eyes-open states consistently occupies curved manifolds with pronounced negative scalar curvatures, exhibiting saddle-point-like geometries that reflect complex, multidimensional coordination patterns. In contrast, eyes-closed states correspond to flatter manifolds with curvatures approaching zero, suggesting simpler geometric organization during periods of reduced computational demands (Renart et al., 2010; Deco et al., 2011).

Linking Statistical and Geometric Complexity

Statistical complexity directly corresponds to geometric curvature. Components with strong non-Gaussian statistics – high third- and fourth-order cumulants – exhibit the most curved manifold structures. This suggests a fundamental relationship between statistical dependencies and neural geometry.

Components with complex curved geometries exhibit rich higher-order correlations, while those with flatter structures display statistical properties closer to Gaussianity. This correspondence supports our integrated analytical framework while suggesting that statistical and geometric perspectives may capture complementary aspects of underlying neural coordination mechanisms.

4.9.2 Methodological Innovations and Technical Framework

Normalizing Flows: Advantages and Extensions

While variational autoencoders (VAEs) represent a natural alternative for this type of analysis (Kingma and Welling, 2013), our normalizing flow approach provides advantages including exact likelihood computation and direct analytical tractability, as demonstrated in Section 4.2.4.

Our framework could extend to VAE architectures since the key innovation – approximating the decoder with a quadratic function and computing characteristic functions, cumulants, and curvature tensors – applies generally to latent-variable models. Advanced VAE variants with disentangled representations (Higgins et al., 2017; Hsu et al., 2017; Kim and Mnih, 2018; Chen et al., 2018) or geometry-aware training (Rey et al., 2019; Koehler et al., 2021) would be particularly promising candidates.

These methods work with any latent variable model, not just normalizing flows. The key extensions could be applied to VAEs, diffusion models, and other generative architectures.

The Power of Quadratic Approximation

The quadratic approximation introduced in Section 4.7.1 proves effective, capturing essential statistical properties despite reducing the complexity of the trained networks. This finding suggests that neural population interactions, while complex, may be well-approximated by a combination of linear and quadratic relationships – a perspective that provides computational tractability.

The approximation’s success enables analytical computation of both statistical cumulants and geometric curvature from the same functional representation, as demonstrated in Sections 4.7 and 4.8.

Multimodal Latent Spaces for State-Dependent Analysis

The multimodal latent space proves essential for capturing how neural geometry varies across behavioral contexts, as demonstrated in Section 4.6. Our extension enables the network to discover distinct geometric modes corresponding to different computational demands without requiring supervised labels, revealing state-dependent organization that emerges naturally from the data structure.

This technical innovation connects directly to the network dynamics insights from Chapter 2, which reveal how neural responses for different stimuli are reshaped. The geometric analysis provides a complementary perspective on how these dynamical reshaping processes manifest as curved manifold structures.

4.9.3 Computational Implications of Geometric Structure

The geometric characterization provides insights into neural computation principles that remain invisible to traditional analytical approaches. We now explore what these geometric properties may reveal about neural information processing.

Saddle-Point Geometries as Computational Motifs

The prevalence of saddle-point geometries (negative curvature) during active processing suggests these structures support flexible computation. Saddle-point manifolds allow neural representations to be dynamically reconfigured while maintaining stability – a key requirement for adaptive behavior (Remington et al., 2018). This interpretation aligns with recent studies demonstrating that nonlinear curvature and manifold structure carry functional relevance across multiple species and brain regions (Gallego et al., 2017, 2020; Altan et al., 2021; Safaie et al., 2023; Fortunato et al., 2024).

The concentration of complex geometric structures during active processing periods aligns with theoretical frameworks that emphasize the computational advantages of curved manifolds for information routing and transformation. During periods of minimal computational demands (eyes-closed states), simpler flat geometries may suffice. When demands increase (eyes-open states), curved structures may emerge to support more sophisticated processing. This relationship connects to broader theoretical work on correlated variability in population codes and the geometry of information coding (Stopfer et al., 2003; Moreno-Bote et al., 2014; Kohn et al., 2016; Azeredo da Silveira and Rieke, 2021).

Adaptive Geometric Organization

The behavioral state-dependent geometric organization demonstrated in Sections 4.6 and 4.8 reveals distinct manifold structures across different computational contexts. This geometric variation provides a population-level complement to the temporal coordination patterns we analyze in Chapter 1, suggesting these represent different organizational principles that may operate across multiple scales of neural activity.

4.9.4 Multi-Scale Integration and Cross-Chapter Connections

Our geometric analysis reveals curved manifolds during active processing states, complementing the temporal coordination patterns analyzed in Chapter 1 and the network dynamics characterized in Chapter 2 .

The invertible networks from Chapter 3 provide the technical foundation for this analysis by learning higher-order statistical dependencies that manifest as geometric structure. This framework demonstrates that neural computation can be analyzed across multiple scales – from temporal patterns to population-level manifold geometry – with statistical dependencies serving as coordination mechanisms that appear differently depending on the analytical perspective employed.

4.9.5 Challenging Traditional Approaches to Neural Coding

Our findings suggest important limitations in current analytical frameworks and point toward new directions for understanding neural computation.

Beyond Linear Assumptions

The differences between curved manifold structure during active processing and flatter geometries during rest that we demonstrate in this chapter suggest potential limitations of linear analytical approaches for understanding neural computation. Traditional methods like PCA, which assume linear relationships and Gaussian statistics (Pearson, 1901; Hotelling, 1936), miss systematically aspects of the geometric complexity that our analysis reveals, despite their continued dominance in neural manifold studies alongside tensor-based extensions (Williams et al., 2018; Mitchell-Heggs et al., 2023; Pellegrino et al., 2024).

The higher-order statistical dependencies, as we characterize in Section 4.7, that generate curved geometric structures may represent information-processing mechanisms that linear methods cannot detect. This finding suggests that some aspects of the computational sophistication of neural circuits remain invisible to traditional analytical approaches, indicating opportunities for advancing understanding through geometric perspectives that go beyond the restrictive assumptions imposed by classical manifold learning techniques (Tenenbaum et al., 2000; Roweis and Saul, 2000; McInnes et al., 2018).

Toward Geometric Theories of Neural Computation

The systematic relationship between behavioral demands and geometric complexity that our analysis reveals raises questions about the functional role of manifold curvature in neural computation. The consistent emergence of curved geometric structures during demanding processing periods suggests these may represent more than incidental features of neural dynamics, though the functional significance remains to be established.

Our findings contribute to ongoing discussions connecting neural computation to principles from differential geometry and information theory. Our empirical findings provide one data point for these broader theoretical frameworks, though much work remains to establish whether geometric organizational strategies play functional roles in neural information processing.

4.9.6 Practical Applications and Future Research Directions

The methodological innovations we develop in this work open new possibilities for neuroscience research and suggest several promising directions for future investigation.

Scalability for Contemporary Neuroscience

The framework’s demonstrated capability with approximately 800 simultaneously recorded neurons, as we describe in Section 4.3, establishes its relevance for contemporary neuroscience applications. The architectural strategies that enable this scalability – processing only the most informative dimensions – provide foundations for analyzing even larger datasets from emerging recording technologies.

The volume-preserving flow architectures (Dinh et al., 2015) prove essential for training stability while maintaining interpretability within the setting of a multimodal latent space (Izmailov et al., 2020). This training stability will scale to higher-dimensional data, as demonstrated in recent applications of normalizing flows to high-dimensional datasets.

Cross-System Applications

The framework’s generality suggests broad applicability across different brain regions, behavioral contexts, and experimental paradigms. The geometric organizational principles we identify in visual cortex likely operate throughout the nervous system, though specific manifestations may reflect region-specific computational demands, connecting to recent developments in differential geometry applied to complex data (Shao et al., 2018; Jones, 2024).

Our approach learns full probability densities from large-scale recordings, enabling richer descriptions of neural variability compared to other geometric approaches that focus primarily on synthetic data or small-scale applications (Acosta et al., 2023). This builds on established curvature-based methods in neuroscience and statistical analysis on manifolds (Fletcher et al., 2009; Weber et al., 2017; Farooq et al., 2019; Gilbert and O’Neill, 2023). The methodological innovations provide tools for investigating how geometric organization varies across brain regions, potentially revealing general principles while illuminating region-specific adaptations.

Dynamic Geometric Evolution

Understanding how manifold structure evolves during learning, adaptation, and state transitions represents natural extensions of this work. The static geometric analysis we present here could be extended to capture how curvature properties change as neural circuits adapt to new computational demands or acquire new behavioral capabilities.

The temporal precision methods from Chapter 1 could track how precise correlation patterns change during these adaptive processes, while geometric analysis could reveal corresponding changes in manifold structure. This integrated perspective might illuminate fundamental mechanisms of neural plasticity that operate across multiple organizational scales.

Multi-Area Coordination

Different brain regions likely employ distinct geometric organizational strategies depending on their computational roles. The framework provides tools for characterizing these region-specific strategies while investigating how they coordinate to support integrated brain function.

The statistical and geometric analysis could reveal how information flows between brain areas through manifold transformations, providing insights into the computational principles underlying large-scale neural coordination. Such understanding could advance both basic neuroscience and applications in brain-computer interfaces.

4.9.7 A Unified Framework for Neural Computation

This work establishes geometric analysis as a potentially valuable complement to traditional approaches in computational neuroscience. The observation that neural populations organize into curved manifolds with behaviorally-relevant structure provides new perspectives on how neural circuits may implement complex computations.

Synthesis Across Organizational Scales

The geometric characterization provides integration with our comprehensive analytical framework, revealing how the temporal precision we validate in Chapter 1, the network dynamics we characterize in Chapter 2, and the statistical dependencies we learn in Chapter 3 contribute to population-level geometric organization.

This synthesis demonstrates that neural computation may not be fully understood through any single analytical perspective. Instead, insights may emerge when temporal, dynamical, statistical, and geometric approaches are unified to reveal the multi-scale organizational principles that enable neural circuits to implement sophisticated computations.

Methodological Contributions to Computational Neuroscience

The technical innovations – particularly the reconstruction error loss we use in Section 4.5.2, the quadratic approximation enabling analytical tractability we develop in Section 4.7.1, and the mixture-of-Gaussians latent space capturing state-dependent organization we describe in Section 4.6.1 – establish new possibilities for statistical and geometric analysis of neural data. These methods provide computational neuroscience with tools capable of revealing geometric organizational principles while maintaining the interpretability essential for biological understanding.

Toward Understanding Biological Intelligence

The framework’s demonstrated capability to reveal behaviorally-relevant structure through unsupervised learning suggests broad applicability for understanding how neural geometry adapts to computational demands across different brain regions and behavioral contexts. By unifying statistical and geometric perspectives, this approach may advance understanding of neural computation by revealing organizational principles that remain invisible to traditional analytical approaches – principles that may ultimately help explain how the brain’s statistical dependencies and curved geometric structure enable the computational abilities of biological intelligence.

Discussion

5.1 Connecting Temporal Precision and Geometric Structure in Neural Computation

Executive Summary

We developed methods to analyze neural computation across multiple scales – from millisecond-precise spike timing to the curved geometry of population manifolds. We evaluated detecting discrete temporal patterns through conservative surrogate approaches and characterized continuous geometric structures in high-dimensional neural activity. While recurrent network dynamics provide one potential mechanistic framework connecting these scales, the precise relationships between temporal coordination and geometric organization remain an open empirical question.

Neural computation operates across multiple scales simultaneously. Understanding how the brain works requires analytical tools that can capture both discrete temporal patterns and continuous geometric structure. What emerges through our four interconnected studies is not merely a collection of independent methods, but complementary analytical tools for examining discrete temporal coordination and continuous geometric structure as distinct but potentially related aspects of neural organization that underlies how populations represent and process information (Churchland et al., 2012; Shenoy et al., 2013; Vyas et al., 2020).

5.2 Methodological Insights Across Neural Organization Scales

Our progression from discrete temporal analysis to continuous geometric characterization reveals several methodological insights about analyzing neural systems across scales. Most significantly, we develop reliable methods for validating temporal precision through conservative surrogate approaches and demonstrate that

population activity exhibits distinct geometric signatures during different behavioral states. Whether these temporal and geometric features reflect interconnected underlying mechanisms requires further empirical investigation to understand how neural populations encode information across the multiple scales at which the neural code operates (Pouget et al., 2000; Gallego et al., 2017).

5.3 Methodological Advances in Temporal Precision Detection

We discovered that uniform dithering – the most widely used surrogate method – produces false positives in over 90% of datasets. Our conservative alternatives reduce this to 10-15%, suggesting that genuine temporal precision may be much sparser than previously reported. We continued here from the master’s thesis of Bouss (2020).

Our findings indicate that precisely timed correlations, when detected through rigorous statistical methods, appear primarily during high-firing-rate conditions and behaviorally demanding epochs. However, the sparsity of patterns detected through conservative methods suggests that either genuine temporal coordination is rare, or that current detection methods remain insufficient for capturing the full extent of temporal precision in neural activity. The statistical frameworks we develop build upon established surrogate data principles (Louis et al., 2010) while addressing methodological limitations identified in previous studies (Stark and Abeles, 2007).

5.4 Network Dynamics and Stimulus Representation

Chapter 2’s analysis of recurrent network dynamics in mouse superior colliculus demonstrates how stimulus representations are reshaped through three key dynamical quantities: mean population activity and within- and across-stimulus overlaps. This work provides a concrete framework for understanding how network-level properties influence stimulus separability, building on theoretical foundations of recurrent circuit dynamics that have emerged from studies of cortical computation (Churchland et al., 2010; Mastrogiuseppe and Ostojic, 2019; Keup et al., 2021). The superior colliculus represents an ideal system for these investigations due to its multimodal integration properties and well-characterized role in sensory processing (Meredith and Stein, 1986).

The mean-field theory we develop successfully captures essential aspects of how recurrent circuits transform stimulus representations, extending previous theoretical work on population dynamics in neural circuits (Mazzucato et al., 2016; Keup

et al., 2021; Benavidez et al., 2021). However, connecting this population-level framework to the millisecond-precise coordination patterns requires additional investigation to determine whether local temporal coordination systematically contributes to the global covariance relationships we characterize. Our approach addresses fundamental questions about how neural variability patterns change during stimulus presentation, building on seminal observations of stimulus-dependent variability modulation (Arieli et al., 1996; Churchland et al., 2010).

5.5 Artificial Data Generation as a Unifying Methodological Theme

Our methods share a validation strategy: controlled artificial data generation. Surrogate methods create spike trains that selectively destroy temporal patterns while preserving other statistics. Invertible networks learn to generate synthetic neural activity that captures the full statistical structure of real data.

Surrogate Data Generation (Chapter 1): Our surrogate methods generate artificial spike trains that preserve specific statistical properties of the original data while systematically removing hypothesized temporal coordination patterns. This controlled perturbation approach transforms the challenge of statistical significance testing from analytical derivation to computational simulation, enabling rigorous assessment of temporal precision claims in spike train data.

Invertible Neural Networks as Data Generators (Chapters 3 & 4): The invertible neural networks we employ are fundamentally generative models – they learn to produce artificial data that matches the statistical structure of training data. In Chapter 3, we use this generative capability to validate our methods on artificial neural network data where ground truth interaction patterns are known, drawing on advances in flow-based generative modeling (Dinh et al., 2015, 2016). In Chapter 4, we employ a modified architecture of these networks that enables us to generate synthetic neural population activity that captures the geometric and statistical properties of multiunit activity envelope (MUAe) data from Utah array recordings in macaque visual cortex, enabling controlled analysis of manifold structure through approaches that build on recent developments in neural manifold characterization (Cunningham and Yu, 2014).

Both approaches share a key insight: artificial data provides ground truth for rigorous validation. Surrogate methods destroy specific patterns to test statistical significance, while invertible networks learn complete statistical structures for analysis.

Cross-Dataset Applications and Methodological Generalizability: An important strength of our artificial data generation approach is that it enables validation across diverse neural recording types and experimental systems. Chapter 1’s surrogate

methods were developed using sorted spikes from macaque premotor and motor cortex (Riehle et al., 2013; Brochier et al., 2018), Chapter 2’s network dynamics analysis employed sorted spikes from Neuropixels recordings in mouse superior colliculus, and Chapter 4’s geometric analysis utilized MUAe data from macaque visual cortex (Chen et al., 2022; ?). This diversity – spanning different recording technologies (Utah arrays vs. Neuropixels), signal types (sorted spikes vs. multiunit activity envelopes), species (macaque vs. mouse), and brain regions (motor, visual, and midbrain areas) – indicates that our approaches can be applied to different types of neural data.

5.6 Statistical Structure Beyond Pairwise Correlations

Chapter 3’s development of invertible neural networks for learning hierarchical interaction structures addresses fundamental limitations of traditional correlation analysis by extracting interaction coefficients from the learned action (negative log-likelihood) using diagrammatic analysis methods. This approach successfully captures higher-order dependencies in artificial neural network data where ground truth interaction patterns are known, enabling validation of the extracted coefficients against known ground truth.

The invertible neural network framework demonstrates how to extract complete statistical structures from population activity data. While the interaction learning approach from Chapter 3 provides methodological foundations for understanding statistical dependencies, Chapter 4 requires a more sophisticated network architecture to enable the geometric manifold analysis of real neural data.

5.7 Geometric Structure and Statistical Properties in Neural Population Activity

Our manifold analysis using modified normalizing flows reveals that different behavioral states correspond to distinct geometric signatures and statistical properties in neural population activity. Eyes-open states in macaque visual cortex exhibit curved manifolds with pronounced negative curvature, while eyes-closed states show nearly flat geometry. These findings provide new quantitative methods for characterizing neural state differences through both geometric and statistical features, contributing to the growing understanding of how neural activity occupies low-dimensional manifolds embedded in high-dimensional state space (Stopfer et al., 2003; Churchland et al., 2012; Stringer et al., 2019a; Safaie et al., 2023).

The geometric analysis demonstrates that neural population activity occupies curved manifolds that can be characterized through curvature measurements, while simultaneously revealing higher-order statistical dependencies that linear methods cannot

capture. Our quadratic approximation approach enables analytical computation of both higher-order statistics and geometric properties from the same functional representation, providing a unified framework for understanding how statistical structure manifests as manifold geometry. However, the computational significance of these geometric and statistical properties for neural information processing remains an important open question.

Neural manifolds appear to be a fundamental organizing principle across brain regions (Stringer et al., 2019a; Gallego et al., 2020; Langdon et al., 2023; Safaie et al., 2023). Our methods provide new tools for quantifying both manifold structure and the statistical dependencies that create this geometry. The application of normalizing flows to neural data analysis represents an extension of recent advances in deep generative models (Dinh et al., 2015, 2016; Kingma and Dhariwal, 2018) to the specific challenges of neural population dynamics analysis, enabling simultaneous characterization of geometric and statistical properties.

5.8 Potential Connections Across Scales

While our work provides methodological tools for analyzing neural activity at different scales, the connections between temporal precision and geometric structure remain largely speculative. The observation that complex geometric structures appear during active behavioral states, when temporal precision methods also detect coordination patterns, suggests potential relationships. However, establishing direct mechanistic connections requires additional experimental work specifically designed to test these hypotheses.

5.9 Methodological Contributions and Their Scope

Each of our methodological contributions addresses specific analytical challenges while defining clear domains of applicability:

Surrogate Methods (Chapter 1): Provide rigorous statistical foundations for temporal precision detection with conservative error control. While this approach prioritizes specificity over sensitivity, it establishes reliable benchmarks for genuine coordination pattern detection.

Network Dynamics (Chapter 2): Successfully characterize population-level transformations in stimulus representation through mean-field analysis. This correlation-based framework captures essential network properties while providing a foundation for future extensions to higher-order interactions.

Invertible Neural Networks (Chapter 3): Demonstrate effective learning of complex statistical structures from data with validated ground truth. This approach establishes principles for interaction coefficient extraction that can inform applications to experimental data with appropriate validation strategies.

Manifold Characterization (Chapter 4): Provide quantitative geometric and statistical analysis of neural population activity with novel curvature-based characterization methods. These tools enable systematic investigation of how geometric properties relate to neural function across different behavioral states.

5.10 Contributions to Neural Coding Theories

Our findings suggest that neural systems employ multiple coordination strategies – temporal precision, geometric organization, and dynamic network transformations – depending on computational demands (DiCarlo et al., 2012; Vyas et al., 2020). The sparsity of temporal patterns, active reshaping of stimulus representations through recurrent dynamics, and state-dependent geometric complexity together suggest that neural systems employ coordination mechanisms strategically rather than uniformly.

5.11 Technical Advances and Scalability

Our methodological framework scales appropriately with advancing recording technologies (Jun et al., 2017; Steinmetz et al., 2021). The surrogate methods apply to large neural populations, manifold characterization handles datasets with hundreds to thousands of channels, and the invertible neural network approach accommodates high-dimensional data. However, the statistical problem of distinguishing temporal from rate-based coding becomes increasingly complex with larger datasets, making the conservative approaches we develop even more critical for reliable detection of genuine coordination patterns.

The computational efficiency of our approaches ensures their applicability to current and emerging experimental paradigms. The geometric characterization framework, in particular, provides interpretable analysis of large-scale recordings while maintaining analytical tractability through the quadratic approximation approach.

5.12 Future Directions Enabled by Our Framework

The methodological tools we developed enable several specific research directions, particularly given the diverse data types and experimental systems we have validated them on:

Cross-Dataset Method Applications: The diversity of our validation datasets – sorted spikes from Utah arrays in macaque premotor and motor cortex (Chapter 1), sorted spikes from Neuropixels in mouse superior colliculus (Chapter 2), and MUAe from Utah arrays in macaque visual cortex (Chapter 4) – enables cross-application of methods. Temporal precision analysis could be applied to mouse superior colliculus data, while geometric methods could be extended to sorted spike data through appropriate preprocessing. These cross-applications could reveal whether our organizational principles generalize across species and brain regions.

Temporal-Geometric Integration: Simultaneous application of temporal precision and geometric analysis methods to the same datasets could reveal whether coordination patterns systematically relate to manifold properties. This is particularly feasible for Utah array recordings, where both sorted spikes (for temporal precision analysis) and multiunit activity envelopes (for geometric manifold analysis) can be extracted from the same electrode data.

Cross-Species and Cross-Regional Analysis: Our framework provides tools for characterizing region-specific and species-specific coordination strategies. The methodological diversity we achieved – validating our approaches across mouse superior colliculus and macaque motor and visual cortices – enables systematic comparisons that could distinguish universal organizational principles from brain-region-specific or species-specific adaptations.

Recording Technology Integration: The validation across different recording technologies (Utah arrays vs. Neuropixels) and signal types (sorted spikes vs. MUAe) demonstrates that our methods can integrate data from emerging high-density recording systems. This positions the framework to leverage technological advances while maintaining analytical continuity with historical datasets.

Learning and Plasticity: The geometric analysis methods could track how manifold structure evolves during learning, while temporal precision methods could reveal corresponding changes in coordination patterns.

5.13 Limitations and Open Questions

Several important limitations and open questions emerge from our work:

1. **Mechanistic Understanding:** While we provide tools for detecting and characterizing neural coordination patterns, the mechanistic understanding of how these patterns contribute to computation requires additional investigation.
2. **Scale Integration:** The relationship between millisecond-precise temporal patterns and population-level geometric structure remains largely theoretical, requiring targeted experiments designed to test specific hypotheses about their interaction.

3. **Functional Significance:** While our methods can reliably detect temporal coordination patterns and characterize geometric structure in neural manifolds, understanding how these patterns contribute to specific behaviors or computations requires experiments that directly link the detected coordination to task performance or cognitive outcomes.

5.14 A Methodological Framework for Multi-Scale Analysis

This thesis establishes that neural computation can be productively analyzed from both temporal and geometric perspectives using rigorous methodological approaches. Our work provides validated tools for detecting temporal precision, characterizing network dynamics, learning statistical structures, and analyzing geometric properties of neural manifolds.

Rather than claiming to solve how neural computation operates across scales, our work provides methodological foundations that enable more rigorous investigation of these questions. The analytical capabilities we developed – spanning discrete temporal analysis to continuous geometric characterization – offer computational neuroscience complementary tools for examining neural organization. By providing reliable methods for characterizing both temporal coordination and geometric structure, we contribute analytical capabilities that could facilitate more comprehensive investigations of how neural circuits encode and transform information across multiple scales.

Appendix

Mathematical Details of Quadratic Representation

We describe the procedure for representing each latent component’s inverse transformation with a quadratic function and its validation. This representation aims to simplify interpretation of the complex transformation implemented by a trained Normalizing Flow.

The procedure begins with latent components resulting from the training process. Each component k is characterized by its mean μ_k and covariance Σ_k . Samples z drawn from these components undergo transformation back to data space through the inverse flow, $x = f_\theta^{-1}(z)$.

We perform quadratic fitting for each mixture to represent the inverse transformation from latent space z to data space x . We construct features for each latent mixture by incorporating both linear terms (z_i) and quadratic terms ($z_i z_j$) to enable fitting of all coefficients in the quadratic net mapping equation. Feature construction utilizes PolynomialFeatures from scikit-learn. LinearRegression models then fit these polynomial features against corresponding data mixtures, capturing both linear and quadratic interactions among latent variables. Parameters obtained from fitting – bias c^α , linear terms B_i^α , and quadratic terms A_{ij}^α – are crucial for understanding how latent variable changes affect neural population activity.

For each latent component, we represent the inverse transformation $f_\theta^{-1}(z)$ using a quadratic function from latent space z to data space x as:

$$x^\alpha = \frac{1}{2} \sum_{i,j} A_{ij}^\alpha z_i z_j + \sum_i B_i^\alpha z_i + c^\alpha.$$

Upper indices α denote data space dimensions with $\hat{e}^\alpha$ representing the corresponding basis vectors. Lower indices i, j refer to latent space dimensions. We omit component-specific indices for clarity, noting that $A_{ij}^\alpha = A_{ji}^\alpha$ by symmetry.

Empirical analysis shows that cumulants up to order 4 are well-represented by the quadratic function across all components: although deviations occur, particularly for higher-order off-diagonal components, the represented statistics maintain strong correlation with true values (appendix of Bouss et al., 2025). This finding demonstrates that despite reduced expressivity compared to the full network, the quadratic model preserves key statistical properties of neural activity. The close correspondence between original and represented cumulants indicates that important interactions among latent variables are maintained, providing confidence for using this simplified model in further interpretation and analysis.

Cumulant Computation via Moment-Generating Function

B.1 Analytical Computation of Cumulants

We compute cumulants of data variables x^α by representing the trained Normalizing Flow f_θ^{-1} through a second-order (quadratic) function of latent variables z . For each mixture component, latent variables follow a multivariate Gaussian distribution. This structure allows analytical evaluation of the moment-generating function $Z(j)$, and subsequently, the cumulant-generating function $W(j) = \ln Z(j)$. The moment-generating function and characteristic function are closely related: the characteristic function $\varphi(t) = Z(it)$ is obtained by substituting imaginary arguments into the moment-generating function, providing equivalent information about the distribution's statistical properties.

Cumulants $\kappa^{\alpha_1 \dots \alpha_n}$ are extracted as mixed partial derivatives of the cumulant-generating function:

$$\kappa^{\alpha_1 \dots \alpha_n} = \frac{\partial^n W(j)}{\partial j^{\alpha_1} \dots \partial j^{\alpha_n} \big|_{j=0}},$$

where the superscripts $\alpha_1, \dots, \alpha_n$ denote data space indices (each taking values $1, \dots, N$), and the notation $\kappa^{\alpha_1 \dots \alpha_n}$ represents the n -th order cumulant tensor. This approach leverages the fundamental property that cumulants measure deviations from Gaussian behavior: for Gaussian distributions, all cumulants beyond second order vanish.

B.2 Mathematical Setup and Notation

We work with the quadratic mapping from N -dimensional latent space to data space:

$$x^\alpha = \frac{1}{2} \sum_{i,j=1}^N A_{ij}^\alpha z_i z_j + \sum_{i=1}^N B_i^\alpha z_i + c^\alpha, \quad \alpha = 1, \dots, N$$

where $z = (z_1, \dots, z_N)^T \in \mathbb{R}^N$ are standardized latent variables following $\sim \mathcal{N}(0, \mathbb{I})$, A_{ij}^α form symmetric matrices $A^\alpha \in \mathbb{R}^{N \times N}$, B_i^α form vectors $B^\alpha \in \mathbb{R}^N$, and c^α are bias terms.

Component-Specific Standardization: For each mixture component k , latent variables initially follow $z \sim \mathcal{N}(\mu_k, \Sigma_k)$, where $\mu_k \in \mathbb{R}^N$ is the component mean vector and $\Sigma_k \in \mathbb{R}^{N \times N}$ is the positive definite component covariance matrix. We standardize using the affine transformation:

$$z = \Sigma_k^{1/2} \tilde{z} + \mu_k$$

where $\tilde{z} \sim \mathcal{N}(0, \mathbb{I})$ and $\Sigma_k^{1/2} \in \mathbb{R}^{N \times N}$ is the matrix square root satisfying $\Sigma_k = \Sigma_k^{1/2} (\Sigma_k^{1/2})^T$.

This transformation modifies the quadratic mapping parameters according to the change of variables:

$$\begin{aligned} \tilde{c}_k^\alpha &= c^\alpha + (B^\alpha)^T \mu_k + 1/2 \mu_k^T A^\alpha \mu_k, \\ \tilde{B}_k^\alpha &= (\Sigma_k^{1/2})^T B^\alpha + (A^\alpha)^T \mu_k, \\ \tilde{A}_k^\alpha &= (\Sigma_k^{1/2})^T A^\alpha \Sigma_k^{1/2}. \end{aligned}$$

These transformed parameters capture the effect of component-specific mean and covariance on the quadratic mapping. Henceforth, we drop both the tilde notation and component index k , working with the standardized parameters $z \sim \mathcal{N}(0, \mathbb{I})$.

The source vector $j = (j^1, \dots, j^N)^T \in \mathbb{R}^N$ has the same dimensionality as the data space. Since we only require derivatives at $j = 0$ for cumulant extraction, all series convergence conditions are automatically satisfied.

B.3 Moment-Generating Function Evaluation

The moment-generating function is:

$$Z(j) = \langle e^{\sum_{\alpha=1}^N j^\alpha x^\alpha} \rangle_z = \int d^N z (2\pi)^{-N/2} \exp\left(-\frac{1}{2} z^T z\right) \exp\left(\sum_{\alpha=1}^N j^\alpha x^\alpha\right)$$

Substituting the quadratic mapping and completing the square in the exponent, we define effective linear and quadratic terms:

$$\begin{aligned} b(j) &= \sum_{\alpha=1}^N j^\alpha B^\alpha \in \mathbb{R}^N, \\ Q(j) &= \mathbb{I} - \sum_{\alpha=1}^N j^\alpha A^\alpha \in \mathbb{R}^{N \times N}. \end{aligned}$$

The Gaussian integration yields:

$$Z(j) = \exp(j^T c) \exp\left(\frac{1}{2} b(j)^T Q(j)^{-1} b(j)\right) (\det(Q(j)))^{-1/2}.$$

B.4 Cumulant-Generating Function and Series Expansion

Taking the logarithm: $W(j) = \ln Z(j)$, we expand using:

(i) Neumann series:

$$\left(\mathbb{I} - \sum_{\alpha=1}^N j^\alpha A^\alpha\right)^{-1} = \sum_{k=0}^{\infty} \left(\sum_{\alpha=1}^N j^\alpha A^\alpha\right)^k$$

(ii) Matrix logarithm:

$$\ln \det\left(\mathbb{I} - \sum_{\alpha=1}^N j^\alpha A^\alpha\right) = - \sum_{k=1}^{\infty} \frac{1}{k} \text{Tr}\left(\left(\sum_{\alpha=1}^N j^\alpha A^\alpha\right)^k\right)$$

B.5 Systematic Cumulant Extraction

When expanding products like $(\sum_{\alpha=1}^N j^\alpha A^\alpha)^k$, we obtain terms $j^{\alpha_1} j^{\alpha_2} \dots j^{\alpha_k} A^{\alpha_1} A^{\alpha_2} \dots A^{\alpha_k}$. Permutations $P(\alpha_1, \dots, \alpha_n)$ arise naturally because mixed partial derivatives are independent of differentiation order: $\frac{\partial^2 W}{\partial j^\alpha \partial j^\beta} = \frac{\partial^2 W}{\partial j^\beta \partial j^\alpha}$. The explicit cumulant formulas are:

First order (mean):

$$\kappa^\alpha = \langle\langle x^\alpha \rangle\rangle = c^\alpha + \frac{1}{2} \text{Tr}(A^\alpha)$$

Second order (covariance):

$$\kappa^{\alpha\beta} = \langle\langle x^\alpha x^\beta \rangle\rangle = \sum_{i=1}^N B_i^\alpha B_i^\beta + \frac{1}{4} \sum_{(\alpha', \beta') \in P(\alpha, \beta)} \text{Tr}(A^{\alpha'} A^{\beta'})$$

Third order:

$$\begin{aligned} \kappa^{\alpha\beta\gamma} &= \langle\langle x^\alpha x^\beta x^\gamma \rangle\rangle \\ &= \frac{1}{2} \sum_{(\alpha', \beta', \gamma') \in P(\alpha, \beta, \gamma)} \sum_{i,j=1}^N B_i^{\alpha'} A_{ij}^{\beta'} B_j^{\gamma'} + \frac{1}{6} \sum_{(\alpha', \beta', \gamma') \in P(\alpha, \beta, \gamma)} \text{Tr}(A^{\alpha'} A^{\beta'} A^{\gamma'}) \end{aligned}$$

Fourth order:

$$\begin{aligned}
 \kappa^{\alpha\beta\gamma\delta} &= \langle\langle x^\alpha x^\beta x^\gamma x^\delta \rangle\rangle \\
 &= \frac{1}{2} \sum_{(\alpha',\beta',\gamma',\delta') \in P(\alpha,\beta,\gamma,\delta)} \sum_{i,j=1}^N B_i^{\gamma'} \left(\sum_{k=1}^N A_{ik}^{\alpha'} A_{kj}^{\beta'} \right) B_j^{\delta'} \\
 &\quad + \frac{1}{8} \sum_{(\alpha',\beta',\gamma',\delta') \in P(\alpha,\beta,\gamma,\delta)} \text{Tr}(A^{\alpha'} A^{\beta'} A^{\gamma'} A^{\delta'})
 \end{aligned}$$

The $\kappa^{\alpha_1 \dots \alpha_n}$ notation emphasizes tensor structure, while $\langle\langle \dots \rangle\rangle$ emphasizes statistical interpretation as cumulants of specific variable combinations.

Differential Geometry and Curvature Calculation

C.1 Geometric Foundation and Assumptions

The computation of curvature from neural data requires systematic development of differential geometric tools. Beginning with quadratic coefficients A_{ij}^α obtained from the normalizing flow, we establish a mathematical framework that transforms local quadratic information into global geometric understanding. The quadratic representation provides the foundation for all subsequent geometric analysis, where the fitted parameters c^α , B_i^α , and A_{ij}^α encode how latent variable changes affect neural population activity.

To compute scalar curvature, the Ricci tensor must first be obtained. The Ricci tensor requires the Riemannian curvature tensor, which in turn depends on Christoffel symbols derived from the metric tensor and its inverse. This hierarchical structure necessitates systematic construction of each component.

C.2 Manifold Definition and Coordinate System

The quadratically approximated neural manifold is treated as a Riemannian submanifold of the ambient neural activity space. This approach is justified by the validation results which demonstrate that the quadratic representation preserves key statistical properties of neural activity despite reduced expressivity compared to the full network.

The quadratic mapping from latent variables z_j to data space x^α is given by:

$$x^\alpha = \frac{1}{2}A_{ij}^\alpha z_i z_j + B_i^\alpha z_i + c^\alpha.$$

Let $\varphi : U \rightarrow \mathbb{R}^N$ be the smooth map defined by $\varphi(z) = (\varphi^1(z), \dots, \varphi^N(z))$ where $\varphi^\alpha(z)$ represents the α -th component of the mapping above, and $U \subseteq \mathbb{R}^{N_l}$ is an open neighborhood of the origin. The neural manifold M is defined as the image of this map: $M = \varphi(U)$.

For M to be an embedded submanifold of $\mathbb{R}^N$, we require: (1) φ is smooth, (2) the differential $d\varphi_z$ has full rank N_l for all $z \in U$, and (3) φ is a homeomorphism onto its image. These conditions are satisfied locally, where the matrix of first derivatives $\partial\varphi^\alpha/\partial z_i = B_i^\alpha + A_{ij}^\alpha z_j$ has full rank.

C.3 Mathematical Conventions and Expansion Strategy

Einstein summation convention is adopted throughout: repeated indices are summed over their range. Greek letters α, β, γ denote data space indices (range 1 to N), while Latin letters i, j, k, l, m, n denote latent space indices (range 1 to N_l).

All geometric quantities are expanded around $z = 0$, representing the component mean in standardized coordinates. This reference point provides optimal accuracy for the quadratic approximation while enabling generalization to arbitrary points z^* through coordinate translation. The validity of this expansion is supported by the empirical analysis which shows that cumulants up to order 4 are well-represented by the quadratic function across all components (appendix of Bouss et al., 2025).

C.4 Construction of Fundamental Geometric Objects

C.4.1 Tangent Vectors and Local Coordinates

Tangent vectors define the coordinate directions for all subsequent geometric calculations. These vectors represent the rate of change of the embedded manifold with respect to latent coordinates:

$$e_i^\alpha = \frac{\partial x^\alpha}{\partial z_i} = B_i^\alpha + A_{ij}^\alpha z_j.$$

Geometrically, e_i^α describes manifold extension in the α -th data dimension when moving along the i -th latent direction. These vectors span the tangent space $T_p M$ at each point p , forming a local coordinate basis for all tangent directions.

The linear term B_i^α captures manifold orientation at the reference point, while the quadratic term $A_{ij}^\alpha z_j$ encodes orientation changes away from the reference point. This variation generates curvature through the second derivatives of the embedding:

$$\frac{\partial^2 x^\alpha}{\partial z_i \partial z_j} = A_{ij}^\alpha.$$

These second derivatives capture the acceleration of the manifold embedding. When the manifold curves, tangent directions change, and this change is measured by the second derivatives. In classical differential geometry, these second derivatives form the building blocks of the second fundamental form.

C.4.2 Metric Tensor Construction

The metric tensor encodes intrinsic manifold geometry by measuring distances and angles using only information available within the manifold itself:

$$g_{ij} = \langle e_i, e_j \rangle = e_i^\alpha e_j^\alpha.$$

This inner product employs the standard Euclidean metric of the ambient space. Expansion in powers of z yields:

$$g_{ij} = B_i^\alpha B_j^\alpha + (B_i^\alpha A_{jk}^\alpha + B_j^\alpha A_{ik}^\alpha) z_k + A_{ik}^\alpha A_{jl}^\alpha z_k z_l.$$

C.4.3 Normal Space Construction and Second Fundamental Form

Curvature computation requires distinguishing between tangent and normal directions. The tangent space is spanned by $e_1, e_2, \dots, e_{N_l}$, while the normal space contains all orthogonal directions.

The projection operator onto the normal space at $z = 0$ becomes:

$$P^{\alpha\beta}(0) = \delta^{\alpha\beta} - B_n^\alpha \gamma^{np} B_p^\beta.$$

where γ^{ij} is the inverse of the metric tensor at $z = 0$.

The second fundamental form measures manifold curvature away from its tangent plane:

$$\Pi_{ij}^{(a)} = \frac{\partial^2 x^\alpha}{\partial z_i \partial z_j} n_\alpha^{(a)} = A_{ij}^\alpha n_\alpha^{(a)}.$$

The quadratic approximation coefficients A_{ij}^α directly provide the components of this fundamental geometric object. When $A_{ij}^\alpha \neq 0$, the second derivatives possess normal components, indicating curvature away from the tangent space.

C.5 Christoffel Symbols and Connection

Christoffel symbols encode how coordinate systems change during manifold traversal. These connection coefficients enable differentiation of tensor fields on curved manifolds:

$$\Gamma_{kl}^i = \frac{1}{2} g^{im} (g_{mk,l} + g_{ml,k} - g_{kl,m}).$$

Using the symmetry $A_{ij}^\alpha = A_{ji}^\alpha$ inherent in the quadratic fitting procedure, to leading order:

$$\Gamma_{kl}^i = \gamma^{im} B_m^\alpha A_{kl}^\alpha + \mathcal{O}(z).$$

C.6 Curvature Computation

C.6.1 Riemannian Curvature Tensor

The Riemannian curvature tensor provides the fundamental measure of manifold deviation from flat space:

$$R_{klj}^i = \frac{\partial \Gamma_{lj}^i}{\partial z_k} - \frac{\partial \Gamma_{kj}^i}{\partial z_l} + \Gamma_{kp}^i \Gamma_{lj}^p - \Gamma_{lp}^i \Gamma_{kj}^p.$$

After substituting all expressions and carefully collecting terms, the result can be written as:

$$R_{klj}^i = \gamma^{im} P^{\alpha\beta}(0) (A_{jl}^\alpha A_{mk}^\beta - A_{kj}^\alpha A_{ml}^\beta) + \mathcal{O}(z).$$

This result manifests the classical Gauss equation, which relates intrinsic curvature to the second fundamental form.

$$R_{klj}^i = \gamma^{im} (\Pi_{jl}^{(a)} \Pi_{mk}^{(a)} - \Pi_{kh}^{(a)} \Pi_{ml}^{(a)}) + \mathcal{O}(z).$$

C.6.2 Sectional Curvature

For a two-dimensional section spanned by tangent vectors e_i and e_j , sectional curvature is defined as:

$$K(e_i, e_j) = \frac{\langle R(e_i, e_j)e_i, e_j \rangle}{\langle e_i, e_i \rangle \langle e_j, e_j \rangle - \langle e_i, e_j \rangle^2}.$$

Substituting the expressions derived above:

$$K(e_i, e_j) = \frac{P^{\alpha\beta}(0)(A_{ii}^\alpha A_{jj}^\beta - A_{ij}^\alpha A_{ij}^\beta)}{g_{ii}g_{jj} - g_{ij}^2}.$$

C.6.3 Scalar Curvature

Scalar curvature provides a single scalar that summarizes overall geometric character at each manifold point. The Ricci curvature tensor is obtained by contracting the Riemannian curvature tensor:

$$R_{jl} = R_{cjl}^c = \gamma^{cm} P^{\alpha\beta}(0)(A_{jl}^\alpha A_{mc}^\beta - A_{cj}^\alpha A_{ml}^\beta).$$

Scalar curvature is the trace of the Ricci curvature tensor with respect to the metric:

$$R = g^{jl} R_{jl}.$$

Working to leading order where $g^{jl} = \gamma^{jl} + \mathcal{O}(z)$, the final scalar curvature formula becomes:

$$R = g^{jl} g^{cm} P^{\alpha\beta}(0)(A_{jl}^\alpha A_{mc}^\beta - A_{cj}^\alpha A_{ml}^\beta).$$

This expression transforms the original A_{ij}^α coefficients into scalar curvature R , a single number capturing the essential geometric character of each neural state.

References

- Abbaspourazad, H., Erturk, E., Pesaran, B., and Shanechi, M. M. (2024). Dynamical flexible inference of nonlinear latent factors and structures in neural population activity. *Nature Biomedical Engineering*, 8(1):85–108.
- Abeles, M. (1991). *Corticonics: Neural Circuits of the Cerebral Cortex*. Cambridge University Press, Cambridge, 1st edition.
- Abeles, M. and Gat, I. (2001). Detecting precise firing sequences in experimental data. *Journal of Neuroscience Methods*, 107(1–2):141–154.
- Abeles, M. and Gerstein, G. L. (1988). Detecting Spatiotemporal Firing Patterns Among Simultaneously Recorded Single Neurons. *Journal of Neurophysiology*, 60(3):909–924.
- Ackley, D. H., Hinton, G. E., and Sejnowski, T. J. (1985). A learning algorithm for boltzmann machines. *Cognitive Science*, 9(1):147–169.
- Acosta, F., Sanborn, S., Duc, K. D., Madhav, M., and Miolane, N. (2023). Quantifying extrinsic curvature in neural manifolds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 610–619.
- Adrian, E. D. (1928). *The Basis of Sensation*. Norton, New York.
- Aertsen, A. and Gerstein, G. (1985). Evaluation of neuronal connectivity: Sensitivity of crosscorrelation. *Brain Research*, 340:341–354.
- Aertsen, A., Gerstein, G., Habib, M., and Palm, G. (1989). Dynamics of neuronal firing correlation: modulation of "effective connectivity". *Journal of Neurophysiology*, 61(5):900–17.
- Altan, E., Solla, S. A., Miller, L. E., and Perreault, E. J. (2021). Estimating the dimensionality of the manifold underlying multi-electrode neural recordings. *PLOS Computational Biology*, 17(11):e1008591.
- Amarasingham, A., Harrison, M. T., Hatsopoulos, N. G., and Geman, S. (2012). Conditional modeling and the jitter method of spike resampling. *Journal of Neurophysiology*, 107(2):517–531.
- Amari, S.-i. and Nagaoka, H. (2000). *Methods of Information Geometry*, volume 191. American Mathematical Society, Providence, RI.

-
- Anderson, P. W. (1972). More is different. *Science*, 177(4047):393–396.
- Ardizzone, L., Kruse, J., Wirkert, S., Rahner, D., Pellegrini, E. W., Klessen, R. S., Maier-Hein, L., Rother, C., and Köthe, U. (2019). Analyzing inverse problems with invertible neural networks. In *International Conference on Learning Representations*.
- Arieli, A., Sterkin, A., Grinvald, A., and Aertsen, A. (1996). Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. *Science*, 273(5283):1868–1871.
- Ashe, J. and Georgopoulos, A. P. (1994). Movement parameters and neural activity in motor cortex and area 5. *Cerebral Cortex*, 6:590–600.
- Azeredo da Silveira, R. and Rieke, F. (2021). The geometry of information coding in correlated neural populations. *Annual Review of Neuroscience*, 44(1):403–424.
- Baker, S. N. and Lemon, R. (2000). Precise spatiotemporal repeating patterns in monkey primary and supplementary motor areas occur at chance level. *Journal of Neurophysiology*, 84:1770–1780.
- Bekasov, A. and Murray, I. (2020). Ordering dimensions with nested dropout normalizing flows. *arXiv preprint arXiv:2006.08777*.
- Belkin, M. and Niyogi, P. (2003). Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation*, 15(6):1373–1396.
- Benavidez, N. L., Bienkowski, M. S., Zhu, M., Garcia, L. H., Fayzullina, M., Gao, L., Bowman, I., Gou, L., Khanjani, N., Cotter, K., Korobkova, L., Becerra, M., Gao, X., Chun, S., Brash, E. M., Zheng, X., Luo, Q., Hintiryan, H., Tao, H. W., Zhang, L. I., and Dong, H.-W. (2021). Organization of the inputs and outputs of the mouse superior colliculus. *Nature Communications*, 12(1):4004.
- Bengio, Y., Courville, A., and Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8):1798–1828.
- Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1):289–300.
- Bennett, M. V. and Zukin, R. S. (2004). Electrical coupling and neuronal synchronization in the mammalian brain. *Neuron*, 41(4):495–511.
- Bernardi, S., Benna, M. K., Rigotti, M., Munuera, J., Fusi, S., and Salzman, C. D. (2020). The geometry of abstraction in the hippocampus and prefrontal cortex. *Cell*, 183(4):954–967.
- Berry, M. J. and Meister, M. (1998). Refractoriness and neural precision. *Journal of Neuroscience*, 18(6):2200–2211.

-
- Bishop, C. M. (1998). Bayesian PCA. In Kearns, M., Solla, S., and Cohn, D., editors, *Advances in Neural Information Processing Systems*, volume 11. MIT Press.
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer, New York.
- Bonferroni, C. (1936). Teoria statistica delle classi e calcolo delle probabilità. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze*, 8:3–62.
- Borgelt, C. (2012). Frequent item set mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2:437–456.
- Bouss, P. (2020). Statistical evaluation of dithering methods for pattern detection. Master’s thesis, RWTH Aachen University, Aachen, Germany.
- Bouss, P., Nestler, S., Fischer, K., Merger, C., René, A., and Helias, M. (2025). Characterizing neural manifolds’ properties and curvatures using normalizing flows. *arXiv preprint arXiv:2506.12187*.
- Breakspear, M. (2017). Dynamic models of large-scale brain activity. *Nature Neuroscience*, 20(3):340–352.
- Brehmer, J. and Cranmer, K. (2020). Flows for simultaneous manifold learning and density estimation. *Advances in Neural Information Processing Systems*, 33:442–453.
- Brette, R. (2015). Philosophy of the spike: Rate-based vs. spike-based theories of the brain. *Frontiers in Systems Neuroscience*, 9:151.
- Brette, R. (2019). Is coding a relevant metaphor for the brain? *Behavioral and Brain Sciences*, 42:e215.
- Brillinger, D. R. (1992). Nerve cell spike train data analysis: a progression of technique. *Journal of the American Statistical Association*, 87:260–271.
- Brochier, T., Zehl, L., Hao, Y., Duret, M., Sprenger, J., Denker, M., Grün, S., and Riehle, A. (2018). Massively parallel recordings in macaque motor cortex during an instructed delayed reach-to-grasp task. *Scientific Data*, 5:180055.
- Brown, E. N., Barbieri, R., Ventura, V., Kaas, R. E., and Frank, L. M. (2001). The time-rescaling theorem and its application to neural spike train data analysis. *Neural Computation*, 14:325–346.
- Brown, E. N., Kaas, R. E., and Mitra, P. P. (2004). Multiple neural spike train data analysis: state-of-the-art and future challenges. *Nature Neuroscience*, 7(5):456–461.
- Brunel, N. (2000). Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of Computational Neuroscience*, 8(3):183–208.
- Buesing, L., Bill, J., Nessler, B., and Maass, W. (2011). Neural dynamics as sampling: A model for stochastic computation in recurrent networks of spiking neurons. *PLOS Computational Biology*, 7:e1002211.

-
- Bullmore, E. and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3):186–198.
- Buonomano, D. V. and Maass, W. (2009). State-dependent computations: spatiotemporal processing in cortical networks. *Nature Reviews Neuroscience*, 10(2):113–125.
- Buzsáki, G. (2004). Large-scale recording of neuronal ensembles. *Nature Neuroscience*, 7(5):446–451.
- Buzsáki, G. (2010). Neural syntax: Cell assemblies, synapsembles, and readers. *Neuron*, 68(3):362–385.
- Buzsáki, G. and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science*, 304:1926–1929.
- Carr, C. and Konishi, M. (1990). A circuit for detection of interaural time differences in the brain stem of the barn owl. *Journal of Neuroscience*, 10:3227–3246.
- Castiello, U. (2005). The neuroscience of grasping. *Nature Reviews Neuroscience*, 6(9):726–736.
- Chen, R. T., Li, X., Grosse, R. B., and Duvenaud, D. K. (2018). Isolating sources of disentanglement in variational autoencoders. *Advances in Neural Information Processing Systems*, 31.
- Chen, X., Morales-Gregorio, A., Sprenger, J., Kleinjohann, A., Sridhar, S., van Albada, S. J., Grün, S., and Roelfsema, P. R. (2022). 1024-channel electrophysiological recordings in macaque V1 and V4 during resting state. *Scientific Data*, 9(1):77.
- Chou, C.-N., Kim, R., Arend, L. A., Yang, Y.-Y., Mensh, B. D., Shim, W. M., Perich, M. G., and Chung, S. (2025). Geometry linked to untangling efficiency reveals structure and computation in neural populations. *BioRxiv*. v2.
- Chung, S. and Abbott, L. F. (2021). Neural population geometry: An approach for understanding biological and artificial neural networks. *Current Opinion in Neurobiology*, 70:137–144.
- Chung, S., Lee, D. D., and Sompolinsky, H. (2016). Linear readout of object manifolds. *Physical Review E*, 93(6):060301.
- Chung, S., Lee, D. D., and Sompolinsky, H. (2018). Classification and geometry of general perceptual manifolds. *Physical Review X*, 8(3):031003.
- Churchland, M. M., Cunningham, J. P., Kaufman, M. T., Foster, J. D., Nuyujukian, P., Ryu, S. I., and Shenoy, K. V. (2012). Neural population dynamics during reaching. *Nature*, 487(7405):51.

-
- Churchland, M. M., Yu, B. M., Cunningham, J. P., Sugrue, L. P., Cohen, M. R., Corrado, G. S., Newsome, W. T., Clark, A. M., Hosseini, P., Scott, B. B., Bradley, D. C., Smith, M. A., Kohn, A., Movshon, J. A., Armstrong, K. M., Moore, T., Chang, S. W., Snyder, L. H., Lisberger, S. G., Priebe, N. J., Finn, I. M., Ferster, D., Ryu, S. I., Santhanam, G., Sahani, M., and Shenoy, K. V. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nature Neuroscience*, 13(3):369–378.
- Cohen, U., Chung, S., Lee, D. D., and Sompolinsky, H. (2020). Separability and geometry of object manifolds in deep neural networks. *Nature Communications*, 11(1):746.
- Connors, B. W. and Gutnick, M. J. (1990). Intrinsic firing patterns of diverse neocortical neurons. *Trends in Neurosciences*, 13(3):99–104.
- Cramer, E., Mitsos, A., Tempone, R., and Dahmen, M. (2022). Principal component density estimation for scenario generation using normalizing flows. *Data-Centric Engineering*, 3:e7.
- Cunningham, J. P. and Yu, B. M. (2014). Dimensionality reduction for large-scale neural recordings. *Nature Neuroscience*, 17(11):1500–1509.
- Date, A., Bienenstock, E., and Geman, S. (1998). On the temporal resolution of neural activity. In *Society for Neuroscience Abstracts*.
- Date, A., Bienenstock, E., and Geman, S. (1999). A statistical technique for detection of fine temporal structure in multi neuronal spike trains. In *Soc Neurosci Abstr*, volume 25, page 1441.
- Dayan, P. and Abbott, L. F. (2001). *Theoretical Neuroscience*. MIT Press, Cambridge, MA.
- Decker, S., Bollmann, D., Nitzan, M., Barzel, B., and Soriano, J. (2009). Effects of random subsampling on network inference. *Physical Review E*, 79(2):026118.
- Deco, G., Jirsa, V. K., and McIntosh, A. R. (2011). Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nature Reviews Neuroscience*, 12:43–56.
- Deco, G., Ponce-Alvarez, A., Mantini, D., Romani, G. L., Hagmann, P., and Corbetta, M. (2013). Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. *Journal of Neuroscience*, 33(27):11239–11252.
- Deger, M., Helias, M., Bousein, C., and Rotter, S. (2012). Statistical properties of superimposed stationary spike trains. *Journal of Computational Neuroscience*, 32(3):443–463.

-
- Deshpande, S. S., Smith, G. A., and van Drongelen, W. (2023). Third-order motifs are sufficient to fully and uniquely characterize spatiotemporal neural network activity. *Scientific Reports*, 13(1):238.
- Dhawale, A. K., Smith, M. A., and Ölveczky, B. P. (2017). The role of variability in motor learning. *Annual Review of Neuroscience*, 40:479–498.
- DiCarlo, J. J. and Cox, D. D. (2007). Untangling invariant object recognition. *Trends in Cognitive Sciences*, 11(8):333–341.
- DiCarlo, J. J., Zoccolan, D., and Rust, N. C. (2012). How does the brain solve visual object recognition? *Neuron*, 73(3):415–434.
- Diesmann, M., Gewaltig, M.-O., and Aertsen, A. (1999). Stable propagation of synchronous spiking in cortical neural networks. *Nature*, 402(6761):529–533.
- Dinh, L., Krueger, D., and Bengio, Y. (2015). NICE: Non-linear independent components estimation. In *3rd International Conference on Learning Representations (ICLR 2015)*, San Diego, CA.
- Dinh, L., Sohl-Dickstein, J., and Bengio, S. (2016). Density estimation using real nvp. *arXiv preprint arXiv:1605.08803*.
- Donoghue, J. P. and Wise, S. P. (1982). The motor cortex of the rat: cytoarchitecture and microstimulation mapping. *Journal of Comparative Neurology*, 212(1):76–88.
- Draxler, F., Wahl, S., Schnörr, C., and Köthe, U. (2024). On the universality of volume-preserving and coupling-based normalizing flows. *arXiv preprint arXiv:2402.06578*.
- Duncker, L. and Sahani, M. (2021). Dynamics on the manifold: Identifying computational dynamical activity from neural population recordings. *Current Opinion in Neurobiology*, 70:163–170.
- Dunn, O. J. (1961). Multiple comparisons among means. *Journal of the American Statistical Association*, 56(293):52–64.
- Efron, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans*. CBMS-NSF, Society for industrial and applied mathematics Philadelphia, Pennsylvania.
- Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*, volume 57 of *Monographs on Statistics and Applied Probability*. Chapman and Hall/CRC, Boca Raton, London, New York, Washington D.C.
- Einevoll, G. T., Kayser, C., Logothetis, N. K., and Panzeri, S. (2013). Modelling and analysis of local field potentials for studying the function of cortical circuits. *Nature Reviews Neuroscience*, 14(11):770–785.

-
- Elsayed, G. F. and Cunningham, J. P. (2017). Structure in neural population recordings: an expected byproduct of simpler phenomena? *Nature Neuroscience*, 20(9):1310–1318.
- Faisal, A. A., Selen, L. P., and Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience*, 9(4):292–303.
- Farooq, H., Chen, Y., Georgiou, T. T., Tannenbaum, A., and Lenglet, C. (2019). Network curvature as a hallmark of brain structural connectivity. *Nature Communications*, 10(1):4937.
- Fascianelli, V., Battista, A., Stefanini, F., Tsujimoto, S., Genovesio, A., and Fusi, S. (2024). Neural representational geometries reflect behavioral differences in monkeys and recurrent neural networks. *Nature Communications*, 15(1):6479.
- Fletcher, P. T., Venkatasubramanian, S., and Joshi, S. (2009). The geometric median on riemannian manifolds with application to robust atlas estimation. *NeuroImage*, 45(1):S143–S152.
- Flouris, K. and Konukoglu, E. (2024). Canonical normalizing flows for manifold learning. *Advances in Neural Information Processing Systems*, 36.
- Fortunato, C., Bennasar-Vázquez, J., Park, J., Chang, J. C., Miller, L. E., Dudman, J. T., Perich, M. G., and Gallego, J. A. (2024). Nonlinear manifolds underlie neural population activity during behaviour. *BioRxiv*. v3.
- Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, 9(10):474–480.
- Gallego, J. A., Perich, M. G., Chowdhury, R. H., Solla, S. A., and Miller, L. E. (2020). Long-term stability of cortical population dynamics underlying consistent behavior. *Nature Neuroscience*, 23(2):260–270.
- Gallego, J. A., Perich, M. G., Miller, L. E., and Solla, S. A. (2017). Neural manifolds for the control of movement. *Neuron*, 94(5):978–984.
- Ganguli, S. and Sompolinsky, H. (2012). Compressed sensing, sparsity, and dimensionality in neuronal information processing and data analysis. *Annual Review of Neuroscience*, 35(1):485–508.
- Ganmor, E., Segev, R., and Schneidman, E. (2011). Sparse low-order interaction network underlies a highly correlated and learnable neural population code. *Proceedings of the National Academy of Sciences of the United States of America*, 108(23):9679–9684.
- Gao, P. and Ganguli, S. (2015). On simplicity and complexity in the brave new world of large-scale neuroscience. *Current Opinion in Neurobiology*, 32:148–155.

- Gardiner, C. (2009). *Stochastic Methods: A Handbook for the Natural and Social Sciences*. Springer, Berlin, Heidelberg, 4th edition.
- Genovese, C. R., Lazar, N. A., and Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, 15(4):870–878.
- Georgopoulos, A., Schwartz, A., and Kettner, R. (1986). Neuronal population coding of movement direction. *Science*, 4771(233):1416–1419.
- Gerstein, G. L. (2004). Searching for significance in spatio-temporal firing patterns. *Acta Neurobiologiae Experimentalis*, 64(2):203–207.
- Gerstein, G. L., Bedenbaugh, P., and Aertsen, A. (1989). Neuronal assemblies. *IEEE Transactions on Biomedical Engineering*, 36:4–14.
- Gerstein, G. L. and Perkel, D. H. (1969). Simultaneously recorded trains of action potentials: analysis and functional interpretation. *Science*, 881(164):828–830.
- Gerstner, W., Kistler, W. M., Naud, R., and Paninski, L. (2014). *Neuronal Dynamics. From Single Neurons to Networks and Models of Cognition*. Cambridge University Press, Cambridge.
- Gilbert, A. C. and O’Neill, K. (2023). Ca-pca: Manifold dimension estimation, adapted for curvature. *arXiv preprint arXiv:2309.13478*.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative adversarial nets. In *Advances in Neural Information Processing Systems*, volume 27, pages 2672–2680.
- Gray, C. M., König, P., Engel, A. K., and Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, 338:334–337.
- Grimaldi, A., Gruel, A., Besnainou, C., Jérémie, J.-N., Martinet, J., and Perrinet, L. U. (2023). Precise spiking motifs in neurobiological and neuromorphic data. *Brain Sciences*, 13(1):68.
- Grün, S. (2009). Data-driven significance estimation of precise spike correlation. *Journal of Neurophysiology*, 101(3):1126–1140.
- Grün, S., Diesmann, M., and Aertsen, A. (2002a). Unitary events in multiple single-neuron spiking activity: I. detection and significance. *Neural Computation*, 14(1):43–80.
- Grün, S., Diesmann, M., and Aertsen, A. (2002b). Unitary events in multiple single-neuron spiking activity: II. nonstationary data. *Neural Computation*, 14(1):81–119.
- Grün, S. and Rotter, S., editors (2010). *Analysis of Parallel Spike Trains*, volume 7 of *Springer Series in Computational Neuroscience*. Springer, Boston, MA.

-
- Harrison, M., Amarasingham, A., and Geman, S. (2007). Jitter methods for investigating spike train dependencies. In *Computational and Systems Neuroscience (COSYNE) 2007*, pages III–17.
- Harrison, M. T. and Geman, S. (2009). A rate and history-preserving resampling algorithm for neural spike trains. *Neural Computation*, 21(5):1244–1258.
- Hastie, T., Tibshirani, R., Friedman, J. H., and Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer.
- Hatsopoulos, N., Geman, S., Amarasingham, A., and Bienenstock, E. (2003). At what time scale does the nervous system operate? *Neurocomputing*, 52:25–29.
- Hatsopoulos, N., Ojakangas, C., Paninski, L., and Donoghue, J. (1998). Information about movement direction obtained from synchronous activity of motor cortical neurons. *Proceedings of the National Academy of Sciences of the United States of America*, 95(26):15706–15711.
- Helias, M. and Dahmen, D. (2020). *Statistical Field Theory for Neural Networks*. Springer International Publishing, Cham.
- Higgins, I., Matthey, L., Pal, A., Burgess, C., Glorot, X., Botvinick, M., Mohamed, S., and Lerchner, A. (2017). beta-vae: Learning basic visual concepts with a constrained variational framework. In *International Conference on Learning Representations*.
- Hinton, G. E. and Sejnowski, T. J. (1986). Learning and relearning in boltzmann machines. In Rumelhart, D. E. and McClelland, J. L., editors, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 1: Foundations*, pages 282–317. MIT Press, Cambridge, MA.
- Hodgkin, A. L. and Huxley, A. F. (1952). Currents carried by sodium and potassium ions through the membrane of the giant axon of *Loligo*. *Journal of Physiology*, 116:449–472.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2):65–70.
- Holt, G. R., Softky, W. R., Koch, C., and Douglas, R. J. (1996). Comparison of discharge variability in vitro and in vivo in cat visual cortex neurons. *Journal of Neurophysiology*, 75(5):1806–1814.
- Honey, C. J., Kötter, R., Breakspear, M., and Sporns, O. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Sciences of the United States of America*, 104(24):10240–10245.
- Horvat, C. and Pfister, J.-P. (2021). Denoising normalizing flow. *Advances in Neural Information Processing Systems*, 34:9099–9111.

-
- Horvat, C. and Pfister, J.-P. (2022). Intrinsic dimensionality estimation using normalizing flows. *Advances in Neural Information Processing Systems*, 35:12225–12236.
- Horvat, C. and Pfister, J.-P. (2023). Density estimation on low-dimensional manifolds: an inflation-deflation approach. *Journal of Machine Learning Research*, 24(61):1–37.
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6):417.
- Hotelling, H. (1936). Relations between two sets of variates. *Biometrika*, 28(3-4):321–r77.
- Hsu, W.-N., Zhang, Y., and Glass, J. (2017). Unsupervised learning of disentangled and interpretable representations from sequential data. *Advances in Neural Information Processing Systems*, 30.
- Huntley, G. W. and Jones, E. G. (1991). Relationship of intrinsic connections to forelimb movement representations in monkey motor cortex: a correlative anatomic and physiological study. *Journal of Neurophysiology*, 66(2):390–413.
- Hyvärinen, A. and Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural Networks*, 13(4-5):411–30.
- Ito, S. and Feldheim, D. A. (2018). The mouse superior colliculus: an emerging model for studying circuit formation and function. *Frontiers in Neural Circuits*, 12:10.
- Izhikevich, E. M. (2006). Polychronization: computation with spikes. *Neural Computation*, 18(2):245–282.
- Izmailov, P., Kirichenko, P., Finzi, M., and Wilson, A. G. (2020). Semi-supervised learning with normalizing flows. In *International Conference on Machine Learning*, pages 4615–4630. PMLR.
- Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical Review*, 106(4):620–630.
- Jazayeri, M. and Ostojic, S. (2021). Interpreting neural computations by examining intrinsic and embedding dimensionality of neural activity. *Current Opinion in Neurobiology*, 70:113–120.
- Jolliffe, I. T. (2002). *Principal Component Analysis*. Springer-Verlag, New York, 2nd edition.
- Jones, I. (2024). Manifold diffusion geometry: Curvature, tangent spaces, and dimension. *arXiv preprint arXiv:2411.04100*.
- Jun, J. J., Steinmetz, N. A., Siegle, J. H., Denman, D. J., Bauza, M., Barbarits, B., Lee, A. K., Anastassiou, C. A., Andrei, A., Aydın, Ç., Barbic, M., Blanche, T. J.,

-
- Bonin, V., Couto, J., Dutta, B., Gratiy, S. L., Gutnisky, D. A., Häusser, M., Karsh, B., Ledochowitsch, P., Lopez, C. M., Mitelut, C., Musa, S., Okun, M., Pachitariu, M., Putzeys, J., Rich, P. D., Rossant, C., Sun, W.-l., Svoboda, K., Carandini, M., Harris, K. D., Koch, C., O’Keefe, J., and Harris, T. D. (2017). Fully integrated silicon probes for high-density recording of neural activity. *Nature*, 551(7679):232–236.
- Kaipio, J. and Somersalo, E. (2005). *Statistical and Computational Inverse Problems*, volume 160 of *Applied Mathematical Sciences*. Springer, New York.
- Kalaska, J. F. and Crammond, D. J. (1992). Cerebral cortical mechanisms of reaching movements. *Science*, 255(5051):1517–1523.
- Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S. A., Hudspeth, A. J., and Mack, S. (2013). *Principles of Neural Science*. McGraw-Hill, New York, 5th edition.
- Kao, T.-C., Sadabadi, M. S., and Hennequin, G. (2021). Optimal anticipatory control as a theory of motor preparation: A thalamo-cortical circuit model. *Neuron*, 109(9):1567–1581.
- Kass, R. E., Ventura, V., and Brown, E. N. (2005). Statistical issues in the analysis of neuronal data. *Journal of Neurophysiology*, 94(1):8–25.
- Kaufman, M. T., Churchland, M. M., and Shenoy, K. V. (2013). The roles of monkey m1 neuron classes in movement preparation and execution. *Journal of Neurophysiology*, 110(4):817–825.
- Kayser, C., Montemurro, M. A., Logothetis, N. K., and Panzeri, S. (2009). Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. *Neuron*, 61(4):597–608.
- Keup, C., Kühn, T., Dahmen, D., and Helias, M. (2021). Transient chaotic dimensionality expansion by recurrent networks. *Physical Review X*, 11(2):021064.
- Kietzmann, T. C., Spoerer, C. J., Sörensen, L. K. A., Cichy, R. M., Hauk, O., and Kriegeskorte, N. (2019). Recurrence is required to capture the representational dynamics of the human visual system. *Proceedings of the National Academy of Sciences of the United States of America*, 116(43):21854–21863.
- Kilavik, B. E., Roux, S., Ponce-Alvarez, A., Confais, J., Grün, S., and Riehle, A. (2009). Long-term modifications in motor cortical dynamics induced by intensive practice. *Journal of Neuroscience*, 29(40):12653–12663.
- Kim, H., Lee, H., Kang, W. H., Lee, J. Y., and Kim, N. S. (2020). Softflow: Probabilistic framework for normalizing flow on manifolds. *Advances in Neural Information Processing Systems*, 33:16388–16397.
- Kim, H. and Mnih, A. (2018). Disentangling by factorising. In *International Conference on Machine Learning*, pages 2649–2658. PMLR.

-
- Kim, T. D., Luo, T. Z., Pillow, J. W., and Brody, C. D. (2021). Inferring latent dynamics underlying neural population activity via neural differential equations. In *International Conference on Machine Learning*, pages 5551–5561. PMLR.
- Kingma, D. P. and Dhariwal, P. (2018). Glow: Generative flow with invertible 1x1 convolutions. *Advances in Neural Information Processing Systems*, 31.
- Kingma, D. P. and Welling, M. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.
- Kiselev, M. (2016). Rate coding vs. temporal coding-is optimum between? In *2016 International Joint Conference on Neural Networks (IJCNN)*, pages 1355–1359. IEEE.
- Knill, D. and Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27(12):712–719.
- Kobak, D., Brendel, W., Constantinidis, C., Feierstein, C. E., Kepecs, A., Mainen, Z. F., Qi, X.-L., Romo, R., Uchida, N., and Machens, C. K. (2016). Demixed principal component analysis of neural population data. *eLife*, 5:e10989.
- Kobyzev, I., Prince, S. J., and Brubaker, M. A. (2020). Normalizing flows: An introduction and review of current methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(11):3964–3979.
- Koehler, F., Mehta, V., Zhou, C., and Risteski, A. (2021). Variational autoencoders in the presence of low-dimensional data: landscape and implicit bias. *arXiv preprint arXiv:2112.06868*.
- Kohn, A., Coen-Cagli, R., Kanitscheider, I., and Pouget, A. (2016). Correlations and neuronal population information. *Annual Review of Neuroscience*, 39(1):237–256.
- Koller, D. and Friedman, N. (2009). *Probabilistic Graphical Models: Principles and Techniques*. Adaptive Computation and Machine Learning. MIT Press, Cambridge, MA.
- Kriegeskorte, N. and Kievit, R. A. (2013). Representational geometry: integrating cognition, computation, and the brain. *Trends in Cognitive Sciences*, 17(8):401–412.
- Langdon, C. and Engel, T. A. (2025). Latent circuit inference from heterogeneous neural responses during cognitive tasks. *Nature Neuroscience*, 28:665–675.
- Langdon, C., Genkin, M., and Engel, T. A. (2023). A unifying perspective on neural manifolds and circuits for cognition. *Nature Reviews Neuroscience*, 24:363–377.
- Lee, J. M. (2018). *Introduction to Riemannian manifolds*, volume 2. Springer.
- Lemon, R. N. (2008). Descending pathways in motor control. *Annual Review of Neuroscience*, 31(1):195–218.
- Lennie, P. (2003). The cost of cortical computation. *Current Biology*, 13(6):493–497.

-
- Levina, A., Priesemann, V., and Zierenberg, J. (2022). Network inference from neural spike trains: a survey. *Neuroscience & Biobehavioral Reviews*, 132:806–828.
- Lewicki, M. S. (1998). A review of methods for spike sorting: the detection and classification of neural action potentials. *Network*, 9(4):R53–78.
- Litwin-Kumar, A. and Doiron, B. (2012). Slow dynamics and high variability in balanced cortical networks with clustered connections. *Nature Neuroscience*, 15(11):1498–1505.
- Loaiza-Ganem, G., Ross, B. L., Cresswell, J. C., and Caterini, A. L. (2022). Diagnosing and fixing manifold overfitting in deep generative models. *arXiv preprint arXiv:2204.07172*.
- Louis, S., Gerstein, G. L., Grün, S., and Diesmann, M. (2010). Surrogate spike train generation through dithering in operational time. *Frontiers in Computational Neuroscience*, 4:127.
- Low, R. J., Lewallen, S., Aronov, D., Nevers, R., and Tank, D. W. (2018). Probing variability in a cognitive map using manifold inference from neural dynamics. *BioRxiv*.
- Macke, J. H., Oppen, M., and Bethge, M. (2011). Common input explains higher-order correlations and entropy in a simple model of neural population activity. *Physical Review Letters*, 106(20):208102.
- Maldonado, P., Babul, C., Singer, W., Rodriguez, E., Berger, D., and Grün, S. (2008). Synchronization of neuronal responses in primary visual cortex of monkeys viewing natural images. *Journal of Neurophysiology*, 100(3):1523–1532.
- Mante, V., Sussillo, D., Shenoy, K. V., and Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature*, 503(7474):78–84.
- Martignon, L., Deco, G., Laskey, K., Diamond, M., Freiwald, W., and Vaadia, E. (2000). Neural coding: Higher-order temporal patterns in the neurostatistics of cell assemblies. *Neural Computation*, 12:2621–2653.
- Martignon, L., von Hasseln, H., Grün, S., Aertsen, A., and Palm, G. (1995). Detecting higher-order interactions among the spiking events in a group of neurons. *Biological Cybernetics*, 73:69–81.
- Mastrogiuseppe, F. and Ostojic, S. (2018). Linking connectivity, dynamics, and computations in low-rank recurrent neural networks. *Neuron*, 99(3):609–623.
- Mastrogiuseppe, F. and Ostojic, S. (2019). A Geometrical Analysis of Global Stability in Trained Feedback Networks. *Neural Computation*, 31(6):1139–1182.

-
- Mazor, O. and Laurent, G. (2005). Transient dynamics versus fixed points in odor representations by locust antennal lobe projection neurons. *Neuron*, 48(4):661–673.
- Mazzucato, L., Fontanini, A., and La Camera, G. (2016). Stimuli reduce the dimensionality of cortical activity. *Frontiers in Systems Neuroscience*, 10:11.
- McCormick, D. A., Connors, B. W., Lighthall, J. W., and Prince, D. A. (1985). Comparative electrophysiology of pyramidal and sparsely spiny neurons of the neocortex. *Journal of Neurophysiology*, 54(4):782–806.
- McInnes, L., Healy, J., and Melville, J. (2018). Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*.
- McLachlan, G. and Peel, D. (2000). *Finite Mixture Models*. John Wiley & Sons, New York.
- Meredith, M. A. and Stein, B. E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology*, 56(3):640–662.
- Merger, C. (2023). *Interactions on structured networks*. PhD thesis, RWTH Aachen University, Aachen, Germany.
- Merger, C., René, A., Fischer, K., Bouss, P., Nestler, S., Dahmen, D., Honerkamp, C., and Helias, M. (2023). Learning interacting theories from data. *Physical Review X*, 13:041033.
- Mézard, M. and Montanari, A. (2009). *Information, Physics and Computation*. Oxford University Press, Oxford.
- Miller, R. G. (1981). *Simultaneous Statistical Inference*. Springer Series in Statistics. Springer-Verlag, New York, 2nd edition.
- Mitchell-Heggs, R., Prado, S., Gava, G. P., Go, M. A., and Schultz, S. R. (2023). Neural manifold analysis of brain circuit dynamics in health and disease. *Journal of Computational Neuroscience*, 51(1):1–21.
- Mochizuki, Y., Onaga, T., Shimazaki, H., Shimokawa, T., Tsubo, Y., Kimura, R., Saiki, A., Sakai, Y., Isomura, Y., Fujisawa, S., Shibata, K.-I., Hirai, D., Furuta, T., Kaneko, T., Takahashi, S., Nakazono, T., Ishino, S., Sakurai, Y., Kitsukawa, T., Lee, J. W., Lee, H., Jung, M. W., Babul, C., Maldonado, P. E., Takahashi, K., Arce-McShane, F. I., Ross, C. F., Sessle, B. J., Hatsopoulos, N. G., Brochier, T., Riehle, A., Chorley, P., Grün, S., Nishijo, H., Ichihara-Takeda, S., Funahashi, S., Shima, K., Mushiake, H., Yamane, Y., Tamura, H., Fujita, I., Inaba, N., Kawano, K., Kurkin, S., Fukushima, K., Kurata, K., Taira, M., Tsutsui, K.-I., Ogawa, T., Komatsu, H., Koida, K., Toyama, K., Richmond, B. J., and Shinomoto, S. (2016). Similarity in neuronal firing regimes across mammalian species. *Journal of Neuroscience*, 36(21):5736–5747.

- Montani, F., Ince, R. A., Senatore, R., Arabzadeh, E., Diamond, M. E., and Panzeri, S. (2009). The impact of high-order interactions on the rate of synchronous discharge and information transmission in somatosensory cortex. *Philosophical Transactions of the Royal Society A*, 367(1901):3297–3310.
- Monteforte, M. and Wolf, F. (2012). Dynamic Flux Tubes Form Reservoirs of Stability in Neuronal Circuits. *Physical Review X*, 2(4):041007.
- Moreno-Bote, R., Beck, J., Kanitscheider, I., Pitkow, X., Latham, P., and Pouget, A. (2014). Information-limiting correlations. *Nature Neuroscience*, 17:1410–1417.
- Musall, S., Kaufman, M. T., Juavinett, A. L., Gluf, S., and Churchland, A. K. (2019). Single-trial neural dynamics are dominated by richly varied movement. *Nature Neuroscience*, 22(10):1677–1686.
- Nawrot, M. P., Boucsein, C., Rodriguez Molina, V., Riehle, A., Aertsen, A., and Rotter, S. (2008). Measurement of variability dynamics in cortical spike trains. *Journal of Neuroscience Methods*, 169:374–390.
- Nestler, S. (2023). *A statistical perspective on learning of time series in neural networks*. PhD thesis, RWTH Aachen University, Aachen, Germany.
- Niell, C. M. and Stryker, M. P. (2010). Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron*, 65(4):472–479.
- Oberste-Frielinghaus, J., Morales-Gregorio, A., Essink, S., Kleinjohann, A., Barthélemy, F., Musall, S., Grün, S., and Ito, J. (2025). Detection and removal of hyper-synchronous artifacts in massively parallel spike recordings. *BioRxiv*. v2.
- Ohiorhenuan, I. E., Mechler, F., Purpura, K. P., Schmid, A. M., Hu, Q., and Victor, J. D. (2010). Sparse coding and high-order correlations in fine-scale cortical networks. *Nature*, 466(7306):617–621.
- Olshausen, B. A. and Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583):607–609.
- Onken, A., Grünewälder, S., Munk, M. H. J., and Obermayer, K. (2009). Analyzing short-term noise dependencies of spike-counts in macaque prefrontal cortex using copulas and the flashlight transformation. *PLOS Computational Biology*, 5(11):1–13.
- Pachitariu, M., Steinmetz, N. A., Kadir, S. N., Carandini, M., and Harris, K. D. (2016). Fast and accurate spike sorting of high-channel count probes with kilosort. In *Advances in Neural Information Processing Systems*, volume 29, pages 4448–4456. Curran Associates, Inc.
- Pandarínath, C., O’Shea, D. J., Collins, J., Jozefowicz, R., Stavisky, S. D., Kao, J. C., Trautmann, E. M., Kaufman, M. T., Ryu, S. I., Hochberg, L. R., Henderson, J. M., Shenoy, K. V., Abbott, L. F., and Sussillo, D. (2018). Inferring single-trial neural population dynamics using sequential auto-encoders. *Nature Methods*, 15(10):805–815.

-
- Paninski, L., Pillow, J., and Lewi, J. (2007). Statistical models for neural encoding, decoding, and optimal stimulus design. *Progress in Brain Research*, 165:493–507.
- Panzeri, S., Brunel, N., Logothetis, N. K., and Kayser, C. (2010). Sensory neural codes using multiplexed temporal scales. *Trends in Neurosciences*, 33(3):111–120.
- Papamakarios, G., Nalisnick, E., Rezende, D. J., Mohamed, S., and Lakshminarayanan, B. (2021). Normalizing flows for probabilistic modeling and inference. *Journal of Machine Learning Research*, 22(57):1–64.
- Parisi, G. (1988). *Statistical field theory*. Addison-Wesley.
- Pazienti, A. and Grün, S. (2006). Robustness of the significance of spike synchrony with respect to sorting errors. *Journal of Computational Neuroscience*, 21(3):329–342.
- Pazienti, A., Maldonado, P. E., Diesmann, M., and Grün, S. (2008). Effectiveness of systematic spike dithering depends on the precision of cortical synchronization. *Brain Research*, 1225:39–46.
- Pearson, K. (1901). LIII. on lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2(11):559–572.
- Pellegrino, A., Stein, H., and Cayco-Gajic, N. A. (2024). Dimensionality reduction beyond neural subspaces with slice tensor component analysis. *Nature Neuroscience*, pages 1–12.
- Perich, M. G., Conti, S., Badi, M., Bogaard, A., Barra, B., Rajan, K., Bloch, J., Courtine, G., Micera, S., Capogrosso, M., and Milekovic, T. (2024). Motor cortical dynamics are shaped by multiple distinct subspaces during naturalistic behavior. *BioRxiv*. v3.
- Perkel, D. H., Gerstein, G. L., and Moore, G. P. (1967a). Neuronal spike trains and stochastic point processes. I. The single spike train. *Biomedical and Pharmacology Journal*, 7(4):391–418.
- Perkel, D. H., Gerstein, G. L., and Moore, G. P. (1967b). Neuronal spike trains and stochastic point processes. II. Simultaneous spike trains. *Biomedical and Pharmacology Journal*, 7(4):419–440.
- Picado-Muiño, D., Borgelt, C., Berger, D., Gerstein, G. L., and Grün, S. (2013). Finding neural assemblies with frequent item set mining. *Frontiers in Neuroinformatics*, 7:9.
- Pillow, J. W., Shlens, J., Paninski, L., Sher, A., Litke, A. M., Chichilnisky, E. J., and Simoncelli, E. P. (2008). Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature*, 454:995–999.
- Pipa, G., Diesmann, M., and Grün, S. (2003). Significance of joint-spike events based on trial-shuffling by efficient combinatorial methods. *Complexity*, 8(4):79–86.

-
- Pipa, G. and Grün, S. (2003). Non-parametric significance estimation of joint-spike events by shuffling and resampling. *Neurocomputing*, 52–54:31–37.
- Pipa, G., Grün, S., and van Vreeswijk, C. (2013). Impact of spike train autostructure on probability distribution of joint spike events. *Neural Computation*, 25(5):1123–1163.
- Pipa, G., Riehle, A., and Grün, S. (2007). Validation of task-related excess of spike coincidences based on neuroxidance. *Neurocomputing*, 70:2064–2068.
- Pipa, G., Wheeler, D. W., Singer, W., and Nikolić, D. (2008). Neuroxidance: reliable and efficient analysis of an excess or deficiency of joint-spike events. *Journal of Computational Neuroscience*, 25(1):64–88.
- Pouget, A., Dayan, P., and Zemel, R. (2000). Information processing with population codes. *Nature Reviews Neuroscience*, 1:125–132.
- Prut, Y., Vaadia, E., Bergman, H., Haalman, I., Slovin, H., and Abeles, M. (1998). Spatiotemporal structure of cortical activity: properties and behavioral relevance. *Journal of Neurophysiology*, 79(6):2857–2874.
- Purves, D., Augustine, G. J., Fitzpatrick, D., Hall, W. C., LaMantia, A.-S., Mooney, R. D., Platt, M. L., and White, L. E. (2018). *Neuroscience*. Sinauer Associates, 6th edition.
- Quaglio, P., Rostami, V., Torre, E., and Grün, S. (2018). Methods for identification of spike patterns in massively parallel spike trains. *Biological Cybernetics*, pages 1–24.
- Quaglio, P., Yegenoglu, A., Torre, E., Endres, D. M., and Grün, S. (2017). Detection and evaluation of spatio-temporal spike patterns in massively parallel spike train data with spade. *Frontiers in Computational Neuroscience*, 11:41.
- Rathelot, J.-A. and Strick, P. L. (2009). Subdivisions of primary motor cortex based on cortico-motoneuronal cells. *Proceedings of the National Academy of Sciences of the United States of America*, 106(3):918–923.
- Recanatesi, S., Farrell, M., Lajoie, G., Deneve, S., Rigotti, M., and Shea-Brown, E. (2021). Predictive learning as a network mechanism for extracting low-dimensional latent space representations. *Nature Communications*, 12(1):1417.
- Rechtsman, M., Stillinger, F. H., and Torquato, S. (2006). Designed interaction potentials via inverse methods for self-assembly. *Physical Review E*, 73(1):011406.
- Reich, D. S., Mechler, F., and Victor, J. D. (2001). Independent and redundant information in nearby cortical neurons. *Science*, 294(5551):2566–2568.
- Remington, E. D., Narain, D., Hosseini, E. A., and Jazayeri, M. (2018). Flexible sensorimotor computations through rapid reconfiguration of cortical dynamics. *Neuron*, 98(5):1005–1019.

-
- Renart, A., De La Rocha, J., Bartho, P., Hollender, L., Parga, N., Reyes, A., and Harris, K. D. (2010). The asynchronous state in cortical circuits. *Science*, 327:587–590.
- Rey, H. G., Pedreira, C., and Quiñero, R. (2015). Past, present and future of spike sorting techniques. *Brain Research Bulletin*, 119:106–117.
- Rey, L. A. P., Menkovski, V., and Portegies, J. W. (2019). Diffusion variational autoencoders. *arXiv preprint arXiv:1901.08991*.
- Rezende, D. and Mohamed, S. (2015). Variational inference with normalizing flows. In *International Conference on Machine Learning*, pages 1530–1538. PMLR.
- Riehle, A., Brochier, T., Nawrot, M., and Grün, S. (2018). Behavioral context determines network state and variability dynamics in monkey motor cortex. *Frontiers in Neural Circuits*, 12:52.
- Riehle, A., Grün, S., Diesmann, M., and Aertsen, A. (1997). Spike synchronization and rate modulation differentially involved in motor cortical function. *Science*, 278(5345):1950–1953.
- Riehle, A., Wirtsohn, S., Grün, S., and Brochier, T. (2013). Mapping the spatio-temporal structure of motor cortical LFP and spiking activities during reach-to-grasp movements. *Frontiers in Neural Circuits*, 7:48.
- Rieke, F., Warland, D., de Ruyter van Steveninck, R., and Bialek, W. (1997). *Spikes: Exploring the Neural Code*. MIT Press, Cambridge, MA.
- Rosen, I. and Asanuma, H. (1972). Peripheral afferent inputs to the forelimb area of the monkey motor cortex: input-output relations. *Experimental Brain Research*, 14(3):257–273.
- Roweis, S. T. and Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326.
- Rule, M. E., O’Leary, T., and Harvey, C. D. (2019). Causes and consequences of representational drift. *Current Opinion in Neurobiology*, 58:141–147.
- Safaie, M., Chang, J. C., Park, J., Miller, L. E., Dudman, J. T., Perich, M. G., and Gallego, J. A. (2023). Preserved neural dynamics across animals performing similar behaviour. *Nature*, 623(7988):765–771.
- Salinas, E. and Abbott, L. F. (1994). Vector reconstruction from firing rates. *Journal of Computational Neuroscience*, 1:89–107.
- Saxena, S. and Cunningham, J. P. (2019). Towards the neural population doctrine. *Current Opinion in Neurobiology*, 55:103–111.

-
- Schneidman, E., Berry, M. J., Segev, R., and Bialek, W. (2006). Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature*, 440:1007–1012.
- Schutzeichel, L., Bauer, J., Bouss, P., Musall, S., Dahmen, D., and Helias, M. (2025). Transient recurrent dynamics shape representations in mice. *BioRxiv*.
- Scott, S. H. (2004). Optimal feedback control and the neural basis of volitional motor control. *Nature Reviews Neuroscience*, 5(7):532–546.
- Seung, H. S. and Lee, D. D. (2000). The manifold ways of perception. *Science*, 290(5500):2268–2269.
- Seung, H. S. and Sompolinsky, H. (1993). Simple models for reading neuronal population codes. *Proceedings of the National Academy of Sciences of the United States of America*, 90(22):10749–10753.
- Shadlen, M. N. and Newsome, W. T. (1994). Noise, neural codes and cortical organization. *Current Opinion in Neurobiology*, 4(4):569–579.
- Shadlen, M. N. and Newsome, W. T. (1998). The variable discharge of cortical neurons: Implications for connectivity, computation, and information coding. *Journal of Neuroscience*, 18(10):3870–3896.
- Shao, H., Kumar, A., and Thomas Fletcher, P. (2018). The riemannian geometry of deep generative models. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 315–323.
- Shenoy, K. V., Sahani, M., and Churchland, M. M. (2013). Cortical control of arm movements: A dynamical systems perspective. *Annual Review of Neuroscience*, 36(1):337–359.
- Shimazaki, H., Amari, S.-i., Brown, E. N., and Grün, S. (2012). State-space analysis of time-varying higher-order spike correlation for multiple neural spike train data. *PLOS Computational Biology*, 8(3):e1002385.
- Shimazaki, H. and Shinomoto, S. (2010). Kernel bandwidth optimization in spike rate estimation. *Journal of Computational Neuroscience*, 29(1):171–182.
- Shimokawa, T. and Shinomoto, S. (2009). Estimating instantaneous irregularity of neuronal firing. *Neural Computation*, 21(7):1931–1951.
- Shinomoto, S. (2010). Estimating the firing rate. In Grün, S. and Rotter, S., editors, *Analysis of Parallel Spike Trains*, volume 7 of *Springer Series in Computational Neuroscience*. Springer, Boston, MA.
- Shinomoto, S., Shima, K., and Tanji, J. (2003). Differences in spiking patterns among cortical neurons. *Neural Computation*, 15(12):2823–2842.

-
- Shlens, J., Field, G. D., Gauthier, J. L., Grivich, M. I., Petrusca, D., Sher, A., Litke, A. M., and Chichilnisky, E. (2006). The structure of multi-neuron firing patterns in primate retina. *Journal of Neuroscience*, 26(32):8254–8266.
- Shmiel, T., Drori, R., Shmiel, O., Ben-Shaul, Y., Nadasdy, Z., Shemesh, M., Teicher, M., and Abeles, M. (2006). Temporally precise cortical firing patterns are associated with distinct action segments. *Journal of Neurophysiology*, 96(5):2645–2652.
- Singer, W. (1999). Neuronal synchrony: a versatile code for the definition of relations? *Neuron*, 24(1):49–65.
- Singer, W. (2013). Cortical dynamics revisited. *Trends in Cognitive Sciences*, 17(12):616–626.
- Softky, W. R. and Koch, C. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *Journal of Neuroscience*, 13(1):334–350.
- Sompolinsky, H., Crisanti, A., and Sommers, H. J. (1988). Chaos in random neural networks. *Physical Review Letters*, 61:259–262.
- Stark, E. and Abeles, M. (2007). Predicting Movement from Multiunit Activity. *Journal of Neuroscience*, 27(31):8387–8394.
- Stark, E. and Abeles, M. (2009). Unbiased estimation of precise temporal correlations between spike trains. *Journal of Neuroscience Methods*, 179(1):90–100.
- Staude, B., Grün, S., and Rotter, S. (2010a). Higher-order correlations in non-stationary parallel spike trains: statistical modeling and inference. *Frontiers in Computational Neuroscience*, 4:16.
- Staude, B., Rotter, S., and Grün, S. (2010b). CuBIC: cumulant based inference of higher-order correlations in massively parallel spike trains. *Journal of Computational Neuroscience*, 29:327–350.
- Steinemann, N., Stine, G. M., Trautmann, E., Zylberberg, A., Wolpert, D. M., and Shadlen, M. N. (2024). Direct observation of the neural computations underlying a single decision. *eLife*, 12:RP90859.
- Steinmetz, N. A., Aydin, C., Lebedeva, A., Okun, M., Pachitariu, M., Bauza, M., Beau, M., Bhagat, J., Böhm, C., Broux, M., Chen, S., Colonell, J., Gardner, R. J., Karsh, B., Kloosterman, F., Kostadinov, D., Mora-Lopez, C., O’Callaghan, J., Park, J., Putzeys, J., Sauerbrei, B., van Daal, R. J. J., Vollan, A. Z., Wang, S., Welkenhuysen, M., Ye, Z., Dudman, J. T., Dutta, B., Hantman, A. W., Harris, K. D., Lee, A. K., Moser, E. I., O’Keefe, J., Renart, A., Svoboda, K., Häusser, M., Haesler, S., Carandini, M., and Harris, T. D. (2021). Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings. *Science*, 372(6539):eabf4588.

-
- Stella, A. (2021). *Higher-order correlation analysis in massively parallel recordings in behaving monkey*. PhD thesis, RWTH Aachen University, Aachen, Germany.
- Stella, A., Bouss, P., Palm, G., and Grün, S. (2022). Comparing surrogates to evaluate precisely timed higher-order spike correlations. *eNeuro*, 9(3):ENEURO.0505–21.2022.
- Stella, A., Quaglio, P., Torre, E., and Grün, S. (2019). 3d-SPADE: Significance Evaluation of Spatio-Temporal Patterns of Various Temporal Extents. *BioSystems*, 185:104022.
- Stevenson, I. H. and Kording, K. P. (2011). How advances in neural recording affect data analysis. *Nature Neuroscience*, 14(2):139–142.
- Stopfer, M., Jayaraman, V., and Laurent, G. (2003). Intensity versus identity coding in an olfactory system. *Neuron*, 39(6):991–1004.
- Storey, J. D. (2002). A direct approach to false discovery rates. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 64(3):479–498.
- Strick, P. L. and Preston, J. B. (1982). Two representations of the hand in area 4 of a primate. i. motor output organization. *Journal of Neurophysiology*, 48(1):139–149.
- Stringer, C., Pachitariu, M., Steinmetz, N., Carandini, M., and Harris, K. D. (2019a). High-dimensional geometry of population responses in visual cortex. *Nature*, 571(7765):361–365.
- Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C. B., Carandini, M., and Harris, K. D. (2019b). Spontaneous behaviors drive multidimensional, brainwide activity. *Science*, 364(6437):eaav7893.
- Südhof, T. C. (2013). Neurotransmitter release: the last millisecond in the life of a synaptic vesicle. *Neuron*, 80(3):675–690.
- Sussillo, D. and Barak, O. (2013). Opening the black box: Low-dimensional dynamics in high-dimensional recurrent neural networks. *Neural Computation*, 25(3):626–649.
- Sussillo, D., Churchland, M. M., Kaufman, M. T., and Shenoy, K. V. (2015). A neural network that finds a naturalistic solution for the production of muscle activity. *Nature Neuroscience*, 18(7):1025–1033.
- Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*. SIAM.
- Tenenbaum, J. B., Silva, V. d., and Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323.

-
- Tetzlaff, T., Helias, M., Einevoll, G. T., and Diesmann, M. (2012). Decorrelation of neural-network activity by inhibitory feedback. *PLOS Computational Biology*, 8(8):e1002596.
- Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., and Farmer, J. D. (1992). Testing for nonlinearity in time series: the method of surrogate data. *Physica D*, 58(1-4):77–94.
- Theunissen, F. and Miller, J. P. (1995). Temporal encoding in nervous systems: A rigorous definition. *Journal of Computational Neuroscience*, 2(6):149–162.
- Tkačik, G., Marre, O., Amodei, D., Schneidman, E., Bialek, W., and Berry II, M. J. (2014). Searching for collective behavior in a large network of sensory neurons. *PLOS Computational Biology*, 10(1):e1003408.
- Torre, E., Picado-Muñoz, D., Denker, M., Borgelt, C., and Grün, S. (2013). Statistical evaluation of synchronous spike patterns extracted by frequent item set mining. *Frontiers in Computational Neuroscience*, 7:132.
- Torre, E., Quaglio, P., Denker, M., Brochier, T., Riehle, A., and Grün, S. (2016). Synchronous spike patterns in macaque motor cortex during an instructed-delay reach-to-grasp task. *Journal of Neuroscience*, 36(32):8329–8340.
- Vaadia, E., Aertsen, A., and Nelken, I. (1995). ‘Dynamics of neuronal interactions’ cannot be explained by ‘neuronal transients’. *Proceedings of the Royal Society - Biological Sciences*, 261(1362):407–410.
- van der Maaten, L. and Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9(86):2579–2605.
- van Vreeswijk, C. (2010). Stochastic models of spike trains. In Grün, S. and Rotter, S., editors, *Analysis of Parallel Spike Trains*, volume 7 of *Springer Series in Computational Neuroscience*. Springer, Boston, MA.
- van Vreeswijk, C. and Sompolinsky, H. (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science*, 274:1724–1726.
- Ventura, V. (2010). Bootstrap tests of hypotheses. In Grün, S. and Rotter, S., editors, *Analysis of Parallel Spike Trains*, volume 7 of *Springer Series in Computational Neuroscience*. Springer, Boston, MA.
- Victor, J. (2005). Spike train metrics. *Current Opinion in Neurobiology*, 15:585–592.
- Villa, A. E. and Abeles, M. (1990). Evidence for spatiotemporal firing patterns within the auditory thalamus of the cat. *Brain Research*, 509(2):325–327.
- Villa, A. E. P., Tetko, I. V., Hyland, B., and Najem, A. (1999). Spatiotemporal activity patterns of rat cortical neurons predict responses in a conditioned task. *Proceedings of the National Academy of Sciences of the United States of America*, 96(3):1106–1111.

-
- Vinje, W. E. and Gallant, J. L. (2000). Sparse coding and decorrelation in primary visual cortex during natural vision. *Science*, 287(5456):1273–1276.
- Vyas, S., Golub, M. D., Sussillo, D., and Shenoy, K. V. (2020). Computation Through Neural Population Dynamics. *Annual Review of Neuroscience*, 43(1):249–275.
- Wang, L., McAlonan, K., Goldstein, S., Gerfen, C. R., and Krauzlis, R. J. (2020). A causal role for mouse superior colliculus in visual perceptual decision-making. *Journal of Neuroscience*, 40(19):3768–3782.
- Weber, M., Stelzer, J., Saucan, E., Naitsat, A., Lohmann, G., and Jost, J. (2017). Curvature-based methods for brain network analysis. *arXiv preprint arXiv:1707.00180*.
- White, B. J. and Munoz, D. P. (2011). The superior colliculus. In *Oxford Handbook of Eye Movements*, pages 195–213. Oxford University Press, Oxford, UK.
- Whiteway, M. R. and Butts, D. A. (2019). The quest for interpretable models of neural population activity. *Current Opinion in Neurobiology*, 58:86–93.
- Williams, A. H., Kim, T. H., Wang, F., Vyas, S., Ryu, S. I., Shenoy, K. V., Schnitzer, M., Kolda, T. G., and Ganguli, S. (2018). Unsupervised discovery of demixed, low-dimensional neural dynamics across multiple timescales through tensor component analysis. *Neuron*, 98(6):1099–1115.
- Wilson, D., Schirrmester, R. T., Gemein, L. A., and Ball, T. (2025). Deep riemannian networks for end-to-end eeg decoding. *Imaging Neuroscience*, 3:imag_a_00511.
- Wilson, K. G. (1971). Renormalization group and critical phenomena. i. renormalization group and the kadanoff scaling picture. *Physical Review B*, 4(9):3174–3183.
- Wood, F., Black, M. J., Vargas-Irwin, C., Fellows, M., and Donoghue, J. P. (2004). On the variability of manual spike sorting. *IEEE Transactions on Biomedical Engineering*, 51(6):912–918.
- Yger, P., Spampinato, G. L., Esposito, E., Lefebvre, B., Deny, S., Gardella, C., Stimberg, M., Jetter, F., Zeck, G., Picaud, S., et al. (2018). A spike sorting toolbox for up to thousands of electrodes validated with ground truth recordings in vitro and in vivo. *eLife*, 7:e34518.
- Yu, B. M., Cunningham, J. P., Santhanam, G., Ryu, S., Shenoy, K. V., and Sahani, M. (2008). Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. *Advances in Neural Information Processing Systems*, 21.
- Yu, S., Huang, D., Singer, W., and Nikolic, D. (2009). Higher-order interactions characterized in cortical activity. *Journal of Neuroscience*, 29(48):15524–15533.
- Zaki, M. J. (2004). Mining non-redundant association rules. *Data Mining and Knowledge Discovery*, 9:223–248.

- Zatka-Haas, P., Steinmetz, N. A., Carandini, M., and Harris, K. D. (2021). Sensory coding and the causal impact of mouse cortex in a visual decision. *eLife*, 10:e63163.
- Zou, H., Hastie, T., and Tibshirani, R. (2006). Sparse principal component analysis. *Journal of Computational and Graphical Statistics*, 15(2):265–286.
- Zuo, Y., Safaai, H., Notaro, G., Mazzoni, A., Panzeri, S., and Diamond, M. E. (2015). Complementary contributions of spike timing and spike rate to perceptual decisions in rat s1 and s2 cortex. *Current Biology*, 25(3):357–363.

List of Figures

1	Analysis pipeline for curved neural manifolds using meaningful latent variables.	15
1.1	Six surrogate generation methods for spike train analysis.	33
1.2	SPADE analysis workflow for pattern detection and significance testing.	37
1.3	Experimental paradigm and neural data structure.	39
1.4	Statistical property preservation across surrogate methods.	44
1.5	Uniform dithering effects on spike counts and temporal structure in experimental data.	52
1.6	Spike count reduction mechanisms during binarization.	55
1.7	Binarization-induced bias in pattern significance testing.	57
1.8	Synthetic data validation and false positive assessment for surrogate methods.	60
1.9	Neuronal characteristics associated with false positive pattern detection.	66
1.10	Pattern detection results from experimental data analysis.	69
2.1	Mean-field theory captures recurrent network dynamics through geometric representation analysis.	83
3.1	Invertible neural network framework for extracting microscopic interaction theories from macroscopic system observations.	93
4.1	Neural population activity demonstrates non-Gaussian statistical structure and higher-order correlations.	109

4.2	Reconstruction loss enables discovery of meaningful latent coordinates in synthetic data.	115
4.3	Multimodal distributions in data and latent space.	116
4.4	Accuracy of data representation saturates with fewer than 10 latent components.	117
4.5	Effect of number of latent components on the approximation capabilities.	118
4.6	Distinct mixture components in the latent space represent behavioral states.	120
4.7	Statistical characteristics of the latent components.	124
4.8	Geometric interpretation of manifold curvature.	126
4.9	Geometric characteristics of the latent components.	129

List of Tables

1.1 Statistical property preservation across surrogate methods. 50

4.1 Scalar curvature for all four components. 129

Band / Volume 117

Electrical anisotropy and shear-resistant topology in the quasi one-dimensional van-der-Waals material α -Bi₄Br₄

J.K. Hofmann (2025) xv, 129 pp

ISBN: 978-3-95806-869-8

Band / Volume 118

Spin-orbital mixing in the topological ladder of the two-dimensional metal PtTe₂

M. Qahosh (2025), ix, 170 pp

ISBN: 978-3-95806-872-8

Band / Volume 119

Functions of SNNs constrained by biology

A. Korcsak-Gorzo (2025), xvi, 145 pp

ISBN: 978-3-95806-876-6

Band / Volume 120

On Scalable Integrated Charge State Tuning for Semiconductor Quantum Dot Devices

F. A. Hader (2026), xx, 145 pp

ISBN: 978-3-95806-884-1

Band / Volume 121

Efficient Massively Space-Time-Parallel Simulations with Adaptive Spectral Deferred Correction

T. Saupe (2026), v, 149 pp

ISBN: 978-3-95806-886-5

Band / Volume 122

Controlled single-molecule manipulation

P. Kopp (2026), viii, 136 pp

ISBN: 978-3-95806-887-2

Band / Volume 123

Growth and Excitation Dynamics of Epitaxial 2D Materials and Molecular Layers Studied by Electron and Momentum Microscopy

M. E. Rath (2026), xvi, 161, XXIX pp

ISBN: 978-3-95806-892-6

Band / Volume 124

Ferroelectric and Resistive Switching in Epitaxial Hf_{0.5}Zr_{0.5}O₂

J. R. Knabe (2026), 145 pp

ISBN: 978-3-95806-895-7

Band / Volume 125

Temperature Influences on Frequency Mixing

Magnetic Detection for mobile Drinking Water Analysis in Crisis Areas

M. Jessing (2026), xx, 113 pp

ISBN: 978-3-95806-902-2

Band / Volume 126

Light-matter interaction from ab-initio Keldysh formalism

M. D. Merte (2026), xvi, 184, A-27 pp

ISBN: 978-3-95806-906-0

Band / Volume 127

Towards high-quality decoupled Graphene heterostructures

M. Hutter (2026), VIII, 149 pp

ISBN: 978-3-95806-912-1 (Print)

ISBN: 978-3-95806-913-8 (E-Book)

Band / Volume 128

**Adaptive-precision interatomic potentials
for large-scale atomistic simulations**

D. Immel (2026), xxiv, 135 pp

ISBN: 978-3-95806-934-3 (Print)

ISBN: 978-3-95806-935-0 (E-Book)

Band / Volume 129

**Competing spin-orbit coupling and exchange
interaction in a two-dimensional ferromagnet**

X. L. Tan (2026), ii, 113 pp

ISBN: 978-3-95806-938-1 (Print)

ISBN: 978-3-95806-939-8 (E-Book)

Band / Volume 130

Analysis of Electrophysiological Data with Invertible Neural Networks

P. Bouss (2026), xxii, 187 pp

ISBN: 978-3-95806-944-2 (Print)

ISBN: 978-3-95806-945-9 (E-Book)

Information
Band / Volume 130
ISBN 978-3-95806-945-9