# Building population models for large-scale neural recordings: opportunities and pitfalls

# Building population models for large-scale neural recordings: opportunities and pitfalls

Cole Hurwitz*, Nina Kudryashova*, Arno Onken and Matthias H. Hennig[†]

University of Edinburgh, Institute for Adaptive and Neural Computation

Edinburgh, EH8 9AB, United Kingdom

† Correspondence: `m.hennig@ed.ac.uk`

* equal contribution

April, 2021

## Abstract

Modern recording technologies now enable simultaneous recording from large numbers of neurons. This has driven the development of new statistical models for analyzing and interpreting neural population activity. Here we provide a broad overview of recent developments in this area. We compare and contrast different approaches, highlight strengths and limitations, and discuss biological and mechanistic insights that these methods provide.

## Introduction

Large scale recordings from neurons with dense, high channel count probes now yield sufficient data to understand the precise covariation in neural populations [35, 65, 79]. This departure from analyzing the tuning properties of individual neuron, often pooled across experiments, requires model-based approaches that capture the complexity of high-dimensional population activity while remaining computationally tractable.

Statistical models have emerged as an essential tool for overcoming this challenge. Informally, these models can be grouped into two categories: fully observed models and latent variable models. Fully observed models try to explicitly capture the interactions between neurons by directly modeling the joint activity of the population. Since modeling the full space of activity patterns is intractable, a key ingredient for tractable fully observed models is an efficient description of the interactions between neurons that retains essential characteristics. Latent variable models, on the other hand, assume that population activity can be summarized by a small set of variables called latent factors. Rather than directly modeling interactions, latent variable models capture neuronal interactions through activation of these factors.

This conceptual difference means that the specific hypotheses each of these model classes can address are also different and that conclusions drawn from such analysis have to be considered in the light of the main assumptions the model is based upon. In this review, we first briefly discuss how single neuron activity is extracted from high-density recordings. Next, we provide an overview of current developments and show which questions can and cannot be addressed with each model class. We close with a discussion of limitations and future directions.
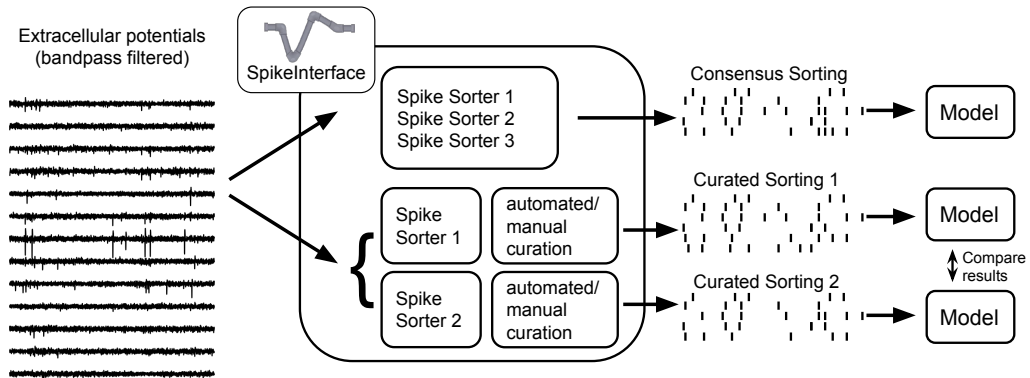
**Figure 1:** The process of extracting the activity of single neurons with a spike sorter is error-prone with different sorters having different failure modes. It is, therefore, recommended to make use of different sorters before performing model-based analysis. A first method is to perform the same analysis on the output of different sorters to see if functional insights are independent of the chosen spike sorter (bottom). Alternatively, one could obtain a consensus sorting (top) where false positive units are largely removed by taking the agreement between different sorters. Any analyses could then be performed on this consensus sorting. These pipelines, together with various pre- and postprocessing steps as well as automated curation, can be easily implemented using the SpikeInterface framework [9] which supports all commonly used spike sorters.

## Extracting modeling data

Large scale recordings can be obtained with high density extracellular probes or using calcium imaging. From extracellular recordings, spike trains of individual neurons are extracted through a process called *spike sorting*. There are several algorithmic approaches for spike sorting of data from dense probes which are reviewed elsewhere [68, 30, 12, 41]. Accurately extracting spike trains is essential since mistakes, including incorrectly merged or split units, unresolved overlapping spikes, or incorrect clustering due to probe drift, have been shown to bias subsequent analysis [4, 91, 92]. For example, in regions with high-firing rates and synchronicity, spike sorting mistakes can lead to artificial correlations between independently firing neurons [4, 17].

It is important to acknowledge that none of the existing spike sorting algorithms are error-free, and that the scale of these recordings now makes manual curation difficult. To aid selection of an appropriate spike sorter, a large collection of ground truth comparisons is available on the SpikeForest website [45]. Moreover, the comparison of spike sorter outputs has shown surprisingly low agreement in part due to a large number of false positive units [9]. A viable approach to mitigate against spurious findings is, therefore, to perform the same analysis on multiple sorter outputs or on an ensemble agreement among multiple sorters; this can be done using a software framework such as SpikeInterface (Figure 1) [9].

While calcium imaging allows unequivocal identification of single neurons, the much lower sampling rate and the binding kinetics of calcium indicators mean the recorded signal may be an incomplete record of the spiking activity of these neurons [32]. A first approach is to recover neural spike trains using deconvolution methods. Similar to spike sorting however, different algorithms have been shown to produce conflicting results when analyzing correlations between neurons [19]. It is therefore prudent to avoid analysis based only on a single convolution algorithm. Second, recent work aims to to include a model of the dynamics giving rise to the calcium signal directly into statistical models [63]. This approach is promising in particular for latent variable models if the different sources of variation (neural and calcium dynamics) can be effectively disentangled.
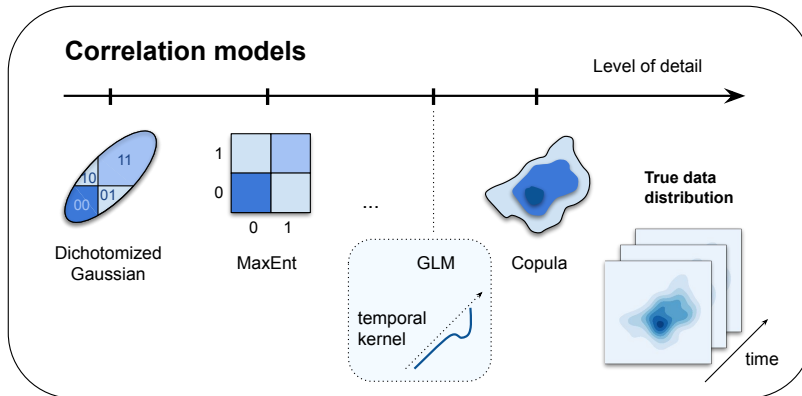
**Figure 2:** Fully observed models construct a joint distribution of population activity with various levels of detail. The horizontal arrow orders the methods according to their expressivity (i.e. how much dependence structure the method can account for).

# Fully observed models

When recording from just 20 neurons, there are over one million possible instantaneous ON/OFF patterns of spiking ($2^{20}$) for a small time bin; this number grows to one billion for 30 neurons ($2^{30}$). While the number of spike patterns that are actually observed is much smaller and constrained by the connectivity between neurons, estimating all pattern frequencies reliably from typical recordings is impossible even for small populations. Fully observed statistical models are a tool to achieve precisely this by approximating this full distribution using as few parameters as possible together with suitable constraints [3, 20, 100]. As always, different approaches exist with different trade-offs between complexity, tractability and interpretability (see Fig. 2 and Tab. 1). A main purpose of these models is to ask how much information recorded activity contains about stimuli or behavior, which can be addressed using information-theoretical or decoding approaches.

### Models

*Maximum Entropy models* (MaxEnt) provide a principled way to construct probability distributions for spiking activity by finding the least structured distribution that fulfills specific constraints obtained from the data [78]. Typical constraints for MaxEnt models are the firing rates of and pairwise correlations between the neurons [80, 83, 88] and the population synchrony [88]. While MaxEnt models have been shown to model spiking activity well, the computational cost of fitting them is prohibitive for large neuronal populations. This has led to the development of analytically tractable models constrained by the dependencies between individual neurons and the aggregate population activity [88, 25, 53, 51]. Related and more scalable is the *Dichotomized Gaussian model* [2, 43] which can be used to model spike count variables with arbitrary marginal distributions [43].

MaxEnt models cannot easily take external inputs into account. If these are well specified, *generalized linear models* (GLMs) can successfully describe the activity of large populations [90, 60, 61]. In these models, external inputs, inputs from other neurons, and each neuron's spiking history are weighted, summed up and transformed by a rate function, which drives a stochastic process to model spiking activity. For typical rate functions, fitting GLMs is a convex optimization problem which can be solved efficiently, making GLMs scalable to large-scale neural recordings [90, 61]. Along with modeling the observed neural population, the activity of unobserved neurons can be inferred [5], and knowing the topology of the external covariates can further aid in inferring the activity of the unseen neurons [86]. Despite their flexibility, GLMs

| Model | References | Number of parameters | Closed-form pattern probabilities? | Fit for large $N$? |
|---|---|---|---|---|
| Dichotomized Gaussian | Amari et al. [2] Macke et al. [43] | $\sim N^2$ | No | Yes |
| Pairwise MaxEnt | Schneidman et al. [80] Shlens et al. [83] | $\sim N^2$ | No[1] | Difficult |
| Tractable MaxEnt | Tkačik et al. [88] Gardella et al. [25] O'Donnell et al. [51] | $\sim N^2$ | Yes[1] | Yes |
| GLM | Pillow et al. [61] | $\sim DN^2$ | No | Difficult |
| Vine Copula | Aas et al. [1] Onken and Panzeri [54] | $\sim DN^2$ | Yes | Yes |

**Table 1:** Fully observed models: A table (adapted from O'Donnell et al. [51]) characterizing model properties and limitations. Here, $N$ is the number of neurons and $D$ is the number of coefficients per interaction term, such as filter sizes for GLM or number of parameters for parametric copula families. Sampling is possible from all of these models. 1. For a more detailed comparison of MaxEnt models, see Table 1 in [78].

cannot model networks with neurons that perform non-linear integration of multiple external inputs [6].

A more direct representation of statistical dependencies is provided by *copula models*, which decompose neural population activity into the distributions of individual neuron activities and the dependence between these neurons, represented by a copula [34]. Copula models can represent inter-neuronal dependencies that go beyond linear (Gaussian) correlations [7], as well as non-linear stimulus-response relationships [37]. These highly flexible models can for example be used to jointly model different recording modalities such as local field potentials and individual neurons' spiking activities [1, 54]. They are particularly promising for analyzing recordings of neural data and exploring its relation to behavior and stimuli which typically follow different statistics and tend to have complex dependencies [37, 33].

**Interpretation and biological insights**

Fully observed models can be used in two main areas. First, they provide access to the full joint activity distribution of the population in place of direct, typically biased estimates from the recordings for information theoretic analyses and for decoding [33, 37, 71, 59]. GLMs and copula models that were fit to population activity have been shown to have high decoding performance of external variables [90, 61, 38, 39, 33, 37]. For instance, GLMs with coupling filters were shown to capture 40% more visual information from the retina than optimal linear decoding [61], indicating that GLMs can model additional details in the activity that are relevant for representing the stimulus. In large-scale datasets with complex statistics, GLMs may however not be appropriate. In these cases, copula models can be useful for measuring the amount of information in the population response about external covariates [54, 33, 37]. Additionally, for complex stimuli with spatial or temporal correlations, one must distinguish 'stimulus correlations' that arise from stimulus statistics from 'noise correlations' due to circuit interactions. Including this yields better performance in MaxEnt models [26], and prevents spurious effects in GLMs [46]. More generally, comparisons of these models with linear decoders can quantify the information that emerges from the network interactions, i.e. how the whole network is different from the sum of its parts [60, 61, 80, 52].

Second, the parameters in many fully observed models can be interpreted as the strength of the interactions between neurons. For example, the coupling parameters in dichotomized
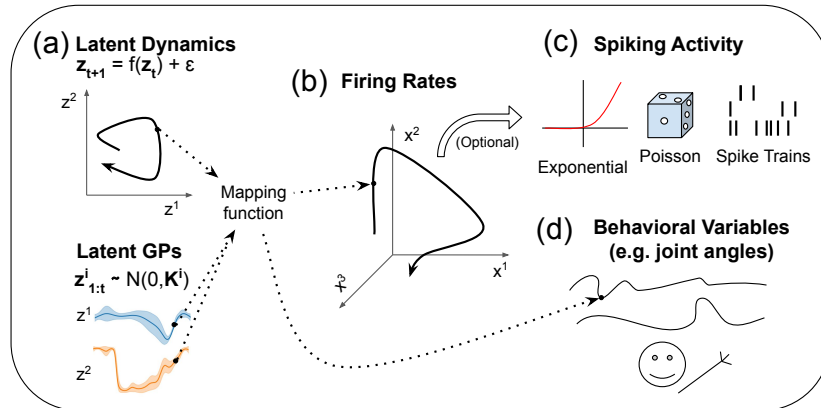
**Figure 3:** Latent variable models assume that neural activity is generated by a set of low-dimensional latent states. (a) When modeling the temporal evolution of the latent states, one can explicitly model the dynamics with a state-space model or can instead model the statistics of the latent trajectories with a Gaussian Process (GP). (b) The latent trajectories can be related back to neural activity (e.g. the firing rates) with a variety of mapping functions including linear functions, neural networks, and GPs. (c) An additional step to further constrain the dynamics is to reconstruct the spiking activity using a point-process distribution such as the Poisson distribution. (d) Recently, it has been shown that explicitly modeling the relationship between the latent factors and the observed behavioral variables can further constrain the factors to be low-dimensional and 'behaviorally relevant' [76]

Gaussian and in pairwise MaxEnt models describe the strength of pairwise interactions between single neurons [43, 80]. As these connections are inferred from the population activity alone, they are usually referred to as functional interactions or connections, to emphasize the difference to synaptic connections. This analysis has been used to uncover the spatial extent of interactions in the retina [83, 52] and to characterize interactions between neurons with different selectivity in the entorhinal cortex [14].

Whether or not inferred functional connectivity is biologically interpretable critically depends on the identifiability of the relevant parameters, and here caution should be exercised. For instance, while the coupling filters in GLMs have unique solutions, filter parameter changes do not necessarily change the GLM responses much [8] suggesting that, in practice, coupling filters may not be properly identifiable. Similarly, in MaxEnt models inferred functional connectivity is not well constrained by the recorded spike trains. As a result, many coupling parameters, even those with large absolute values, can be altered without significantly changing the network activity predicted by the model [55, 62].

Limited data can be a main contributor to poor parameter identifiability. To address this, the uncertainty of the model parameters can be quantified with approximate Bayesian inference methods [99]. A recently developed LR-GLM utilized low-rank data approximations to scale approximate Bayesian inference methods to high-dimensional real datasets making these methods more applicable to large-scale neural recordings [89]. However, closer inspection of MaxEnt models suggests weakly constrained couplings could also be an intrinsic property of neural circuits. Simulations have shown that re-wiring even a large number of synapses may only have minor functional consequences [48], hence the corresponding inferred connections are also weakly specified. Analysis of long term recordings has shown that these less important connections are also subject to more intensive re-modeling over time [55, 62]. Therefore, taking into account the reliability of the inferred couplings in functional connectivity models allows for the identification of the neurons, connections and circuit motifs that most determine the population activity statistics [31, 55, 62].

# Latent variable models

Despite the apparent high-dimensionality of population activity, recent studies suggest that the activity patterns underlying neural function are confined to low-dimensional manifolds [47, 13, 74, 29, 22, 18]. In other words, neural computation results from the activation of specific population-wide activity patterns rather than the independent modulation of single neurons [21]. In this regime, latent variable models are a natural choice as they aim to capture arbitrarily complicated response structure in the neural activity with only a few latent variables.

## Models

Latent variable models aim to capture the population activity structure through the temporal evolution of their learned latent variables. There are two main approaches for modeling the latent trajectories (shown in Figure 3 and summarized in Table 2): *State-space representations*, where the dynamics of the latent trajectories are modeled explicitly with an evolution function that relates past latent states to future latent states, and *moment representations* where, rather than having an explicit model of the dynamics, the statistics of the latent trajectories over time are modeled. With both approaches, there is an *observation model* which relates the neural trajectories back to the observed activity.

Traditional dimensionality reduction techniques such as principal component analysis (PCA) and factor analysis (FA) can be re-interpreted as *static* state-space models without temporal dependence between states [69]. Extending these, *linear* dynamical state-space models such as linear dynamical systems (LDS) [85, 38, 10, 58] and jPCA [13, 82, 50] can capture simple temporal dependencies between latent states. While efficient, these methods cannot model non-linear dynamics which are thought to underlie rhythmic motor patterns [72, 28], decision making [64], and pathologies such as epilepsy [27]. To capture these dynamics, *non-linear* dynamical state-space models have been introduced including approaches that model dynamics with recurrent neural networks (RNNs) [56, 81], piecewise-linear RNNs [15], switching linear dynamics [57], and recurrent switching linear dynamics [42, 98]. A particularly efficient RNN-based model is the recently introduced latent factor analysis via dynamical systems (LFADS) which utilizes neural networks to reduce the computational cost of inferring dynamics, making it more practical for large-scale neural recordings [56]. Despite these recent developments, non-linear state-space models are typically computationally expensive and parameter inference is difficult [73].

An alternative to explicitly learning a dynamical model for the neural population is to model a statistical representation of the trajectories, typically as a Gaussian process (GP). GPs define a distribution over functions and allow to constrain certain features of the function such as its smoothness. Unlike state-space models, GP-based models provide both uncertainty quantification and principled model selection. The most common GP-based latent variable model is the Gaussian-process factor analysis (GPFA) [11]. In this model, each trajectory of the latent state is sampled from an independent GP. This independence constraint, however, limits the ability of GPFA to model joint temporal dependencies making them less useful for uncovering temporal structure in neural data; for instance, GPFA cannot disentangle motor cortex trajectories during hand reach [73]. Gaussian Process Factor Analysis with Dynamical Structure (GPFADS) seeks to remedy this by constraining the learned trajectories to have lower probabilities of occurring in reverse [73]. As a result, GPFADS can disentangle the trajectories generated by non-linear non-reversible dynamical systems such as primary motor cortex arm reach data. Overall, while GP-based latent variable models provide useful analytical properties and are generally quite efficient, they still need to demonstrate that they can capture the complex temporal dynamics that underlie neural activity as well as state-space models can.

An important aspect of all latent variable models is the observation model which relates the trajectories back to the neural activity. Common approaches include modeling neural activity with a conditional Gaussian model or the more plausible conditional Poisson model [44]. The mapping between the latent trajectories and neural activity distribution can take on a number

of functional forms including linear maps [85, 11, 56, 96], GPs [40, 94, 81], or neural networks [24]. Most models, however, rely on a simple linear mapping function as it forces the latent variables to capture more information about the joint neural activity. Recently, observation models that take into account external behavioral variables have been introduced for capturing 'behaviorally relevant' latent trajectories. Sani et al. [76] introduced a novel linear Gaussian state-space model (PSID) that jointly models the neural activity and the recorded behavioral variables. Intriguingly, PSID found that behaviorally relevant latent trajectories are significantly lower-dimensional than previously thought.

**Interpretation and biological insights**

Unlike fully observed models, latent variable models provide a very succinct description of joint population activity in the form of latent trajectories. A simple yet effective approach to interpreting these latent trajectories is visualization; reducing the number of latent variables down to just two or three allows for insightful visualizations. For example, Santhanam et al. [77] used factor analysis to reduce preparatory neural activity in the premotor cortex to just three dimensions that, when visualized, provided evidence that the population activity contained information that could discriminate between goal-related target conditions in a delayed center-out reach task.

It may however not be appropriate to reduce the number of latent variables to two or three for visualization without sacrificing model fit[1]. A more principled and flexible way to interpret the learned latent trajectories is to relate them back to a measured external variable [93]. For example, one can measure how informative a latent variable is about an external variable by decoding the external variable at each time step using a simple method such as a least squares regression. While this approach has been used to interpret population response structure in the olfactory and visual system [75, 16, 84, 95, 87], its value is best demonstrated in studies that correlate latent trajectories in the premotor and motor cortex to arm reaching tasks. Churchland et al. [13] found that low-dimensional rotational trajectories in the primary motor cortex were correlated with arm reaches. Gallego et al. [21] extended this finding, demonstrating that these low-dimensional trajectories are a stable neural correlate for consistent execution of arm reaches over the course of many years. Moreover, Ramanathan et al. [66] showed that following a stroke, diminished reaching function in the contralesional arm was correlated with a loss of motor cortical neural trajectories. The trajectories only reemerged after motor recovery and proved to be a useful neuromodulatory target for therapeutic electrical stimulation.

Despite the success of latent variable models, their usefulness can be dependent on the properties of the neuronal population and the specific experimental setting. Firstly, latent variable models are most useful when the neural population activity is, in fact, low-dimensional. While a number of studies have independently found this to be the case for their datasets, recent work by Stringer et al. [87] suggests that the manifold dimensionality of stimulus-evoked activity in the visual cortex is actually as high as it can be without becoming non-differentiable. Moreover, Gao et al. [23] suggests that population activity is as high-dimensional as possible given the simplicity of the given stimuli or tasks. While latent variable models can still be used to model this data (by increasing the number of latent variables), they may no longer be as interpretable. Another assumption in some models is that the latent trajectories are autonomous[2], i.e. they only depend on an initial condition (usually trial-specific) and an evolution function. This limits the ability of these models to capture unmeasured inputs to the neural population from other brain regions. Finally, and most importantly, latent variable models, by design, will mix all sources of neural variability in the latent space [93]. This can make interpreting the latent variables challenging in particular for complex behaviors or diverse stimulus ensembles. These

---

[1]Model fit can be measured in different ways including the amount of variance explained or the cross-validated likelihood of observed activity given the model. GP-based latent variable models can also provide uncertainty estimates for the latent trajectories.

[2]Pandarinath et al. [56] showed that LFADS could capture task-related inputs to a neural population by incorporating them into the generative process. It is still an open question how generalizable this approach is to other types of inputs.

| Model | Trajectories | Mapping Function | Observation | Single-trial[1] |
|---|---|---|---|---|
| PCA/FA [69] | Static | Linear | Gaussian | No |
| dPCA [36] | Static | Linear | Gaussian[2] | No |
| jPCA [13] | Linear[3] | Linear | Gaussian | No |
| LDS [85] | Linear | Linear | Gaussian | Yes |
| PLDS [44] | Linear | Linear | Poisson | Yes |
| PSID [76] | Linear | Linear | Gaussian[4] | Yes |
| PfLDS [24] | Linear | Neural Network | Poisson | Yes |
| SLDS [57] | Switching Linear | Linear | Gaussian | Yes |
| RSLDS [42, 98] | Recurrent Switching Linear[5] | Linear | Gaussian | Yes |
| PLRNN-SSM [15] | Piecewise-linear RNN | Linear | Gaussian | Yes |
| LFADS [56] | RNN | Linear | Poisson | Yes |
| GP-RNN [81] | RNN | GP | Poisson/Gaussian | Yes |
| GPFA [11] | GP | Linear | Gaussian | Yes |
| GPFADS [73] | GP[6] | Linear | Gaussian | Yes |
| vLGP [96] | GP | Linear | Poisson[7] | Yes |
| P-GPLVM [94] | GP | GP | Poisson | Yes |

**Table 2:** Latent Variable Models: A table (adapted from [81]) characterizing the latent variable models referenced in this review. 1. The ability to extract dynamics from single-trials is important for capturing important variability that can be obscured with trial-averaging [49]. 2. dPCA models trial-averaged and task averaged neural activity to demix different sources of variability in the data. 3. jPCA is equivalent to a linear dynamical model with specific constraints for modeling rotational dynamics [50]. 4. PSID models both the observed neural activity and behaviour with separate Gaussian observation models. 5. RSLDS can model multiple neural populations and their time-varying interactions using switching linear dynamics that is governed by a discrete latent state. 6. GPFADS introduces a new GP kernel which allows for modeling the dynamical behaviour of temporal-irreversibility. 7. The activity of each neuron in vLGP is linearly dependent on the latent variable and the self-history of the neuron.

constraints may be the reason why many of the conceptual breakthroughs provided by latent variable models are for motor cortical activity where the dimensionality appears low, the trajectories are largely autonomous given the trial conditions, and the variation is largely explained by the measured behavior. A possible way forward to improve the interpretability of these models in more complex experimental regimes is to explicitly model external variables either in the observation model [36, 76] or in the latent space [97].

## Outlook

While both fully observed and latent variable models can be used to overcome the curse of dimensionality in large-scale recordings, the two classes yield distinct yet complementary insights into circuit function. In addition, they each have important limitations that are not easily resolved. Fully observed models can suffer from poor parameter identifiability because data sets have a limited size and potentially because many parameters are actually not well constrained in real neural circuits. Possible future direction are to either simplify the description of the interactions to reduce the number of parameters [67] or to augment a model with biologically informed latent dynamics [70].

Latent variable models assume that circuit dynamics are inherently low-dimensional and place specific constraints on how these dynamics unfold. It is perhaps for this reason that we have, so far, few insights into neural dynamics during complex naturalistic behavior; experimental design still favors simple, repeated stimuli and behaviors where sources of neural vari-

ability are easy to de-mix. In more complex experimental regimes, it may no longer be sufficient to generically model the response structure of a neural population and perform ex post facto analyses to relate it back to the variables of interest. Instead, there is increasing evidence that approaches that explicitly model the measured external variables have improved interpretability and can reveal novel insights about neural function [76, 97].

## Acknowledgments

## References

[1] Aas, K., Czado, C., Frigessi, A., and Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and economics*, 44(2):182–198.

[2] Amari, S. I., Nakahara, H., Wu, S., and Sakai, Y. (2003). Synchronous firing and higher-order interactions in neuron pool. *Neural Comput*, 15.

[3] Averbeck, B. B., Latham, P. E., and Pouget, A. (2006). Neural correlations, population coding and computation. *Nat Rev Neurosci*, 7(5):358–366.

[4] Bar-Gad, I., Ritov, Y., Vaadia, E., and Bergman, H. (2001). Failure in identification of overlapping spikes from multiple neuron activity causes artificial correlations. *J Neurosci Methods*, 107(1-2):1–13.

[5] Battistin, C., Hertz, J., Tyrcha, J., and Roudi, Y. (2015). Belief propagation and replicas for inference and learning in a kinetic ising model with hidden spins. *J Stat Mech Theory Exp*, 2015(5):P05021.

[6] Benjamin, A. S., Fernandes, H. L., Tomlinson, T., Ramkumar, P., VerSteeg, C., Chowdhury, R. H., Miller, L. E., and Kording, K. P. (2018). Modern machine learning as a benchmark for fitting neural responses. *Front Comput Neurosc*, 12:56.

[7] Berkes, P., Wood, F., and Pillow, J. (2008). Characterizing neural dependencies with copula models. In *Advances in Neural Information Processing Systems*, volume 21, pages 129–136.

[8] Brinkman, B. A., Rieke, F., Shea-Brown, E., and Buice, M. A. (2018). Predicting how and when hidden neurons skew measured synaptic interactions. *PLoS Comput Biol*, 14(10):e1006490.

[9] Buccino, A. P., Hurwitz, C. L., Garcia, S., Magland, J., Siegle, J. H., Hurwitz, R., and Hennig, M. H. (2020). SpikeInterface, a unified framework for spike sorting. *eLife*, 9:e61834.

* This paper describes an open source framework to create fully reproducible analysis pipelines for extracellular recordings. It allows runnning all major current spike sorting algorithms on the same data sets, includes common methods for pre- and postprocessing and automated curation, and a systematic comparison between sorter outputs. Use of this toolkit can both help to avoid failure modes of common spike sorters, and aid more reproducible analysis.

[10] Buesing, L., Macke, J. H., and Sahani, M. (2013). Spectral learning of linear dynamics from generalised-linear observations with application to neural population data. In *Advances in Neural Information Processing Systems 25: 26th Conference on Neural Information Processing Systems (NIPS 2012)*, pages 1691–1699.

[11] Byron, M. Y., Cunningham, J. P., Santhanam, G., Ryu, S. I., Shenoy, K. V., and Sahani, M. (2009). Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. In *Advances in Neural Information Processing Systems*, pages 1881–1888.

[12] Carlson, D. and Carin, L. (2019). Continuing progress of spike sorting in the era of big data. *Curr Opin Neurobiol*, 55:90–96.

[13] 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):51–56.

[14] Dunn, B., Mørreaunet, M., and Roudi, Y. (2015). Correlations and Functional Connections in a Population of Grid Cells. *PLoS Comput Biol*, 11(2):e1004052.

[15] Durstewitz, D. (2017). A state space approach for piecewise-linear recurrent neural networks for identifying computational dynamics from neural measurements. *PLoS computational biology*, 13(6):e1005542.

[16] Ecker, A. S., Berens, P., Cotton, R. J., Subramaniyan, M., Denfield, G. H., Cadwell, C. R., Smirnakis, S. M., Bethge, M., and Tolias, A. S. (2014). State dependence of noise correlations in macaque primary visual cortex. *Neuron*, 82(1):235–248.

[17] Ecker, A. S., Berens, P., Keliris, G. A., Bethge, M., Logothetis, N. K., and Tolias, A. S. (2010). Decorrelated neuronal firing in cortical microcircuits. *science*, 327(5965):584–587.

[18] Elsayed, G. F. and Cunningham, J. P. (2017). Structure in neural population recordings: an expected byproduct of simpler phenomena? *Nat Neurosci*, 20(9):1310.

[19] Evans, M. H., Petersen, R. S., and Humphries, M. D. (2019). On the use of calcium deconvolution algorithms in practical contexts. *bioRxiv*, page 871137.

[20] Franke, F., Fiscella, M., Sevelev, M., Roska, B., Hierlemann, A., and da Silveira, R. A. (2016). Structures of neural correlation and how they favor coding. *Neuron*, 89.

[21] 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. *Nat Neurosci*, 23(2):260–270.

** The paper demonstrates that latent dynamics are a stable neural correlate for the execution of learned behaviors. They show that even across years of recordings, latent dynamics in the primary motor cortex during reach remain surprisingly consistent despite significant changes in single neuron activities.

[22] 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.

[23] Gao, P., Trautmann, E., Yu, B., Santhanam, G., Ryu, S., Shenoy, K., and Ganguli, S. (2017). A theory of multineuronal dimensionality, dynamics and measurement. *BioRxiv*.

[24] Gao, Y., Archer, E., Paninski, L., and Cunningham, J. P. (2016). Linear dynamical neural population models through nonlinear embeddings. *arXiv preprint arXiv:1605.08454*.

[25] Gardella, C., Marre, O., and Mora, T. (2016). A tractable method for describing complex couplings between neurons and population rate. *eNeuro*, 3.

[26] Granot-Atedgi, E., Tkačik, G., Segev, R., and Schneidman, E. (2013). Stimulus-dependent maximum entropy models of neural population codes. *PLoS Comput Biol*, 9(3):e1002922.

[27] Haghighi, H. S. and Markazi, A. (2017). A new description of epileptic seizures based on dynamic analysis of a thalamocortical model. *Sci Rep*, 7(1):1–10.

[28] Hall, T. M., de Carvalho, F., and Jackson, A. (2014a). A common structure underlies low-frequency cortical dynamics in movement, sleep, and sedation. *Neuron*, 83(5):1185–1199.

[29] Hall, T. M., Nazarpour, K., and Jackson, A. (2014b). Real-time estimation and biofeedback of single-neuron firing rates using local field potentials. *Nat Commun*, 5(1):1–12.

[30] Hennig, M. H., Hurwitz, C., and Sorbaro, M. (2019). Scaling Spike Detection and Sorting for Next-Generation Electrophysiology. In Chiappalone, M., Pasquale, V., and Frega, M., editors, *In Vitro Neuronal Networks: From Culturing Methods to Neuro-Technological Applications*, Advances in Neurobiology, pages 171–184. Springer International Publishing, Cham.

[31] Herzog, R., Escobar, M.-J., Cofre, R., Palacios, A. G., and Cessac, B. (2018). Dimensionality Reduction on Spatio-Temporal Maximum Entropy Models of Spiking Networks. *bioRxiv*, page 278606.

[32] Huang, L., Ledochowitsch, P., Knoblich, U., Lecoq, J., Murphy, G. J., Reid, R. C., de Vries, S. E., Koch, C., Zeng, H., Buice, M. A., et al. (2021). Relationship between simultaneously recorded spiking activity and fluorescence signal in gcamp6 transgenic mice. *Elife*, 10:e51675.

[33] Ince, R. A., Giordano, B. L., Kayser, C., Rousselet, G. A., Gross, J., and Schyns, P. G. (2017). A statistical framework for neuroimaging data analysis based on mutual information estimated via a gaussian copula. *Hum Brain Mapp*, 38(3):1541–1573.

[34] Jenison, R. L. and Reale, R. A. (2004). The shape of neural dependence. *Neural Comput*, 16.

[35] 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.

[36] 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.

[37] Kudryashova, N., Amvrosiadis, T., Dupuy, N., Rochefort, N., and Onken, A. (2020). Parametric Copula-GP model for analyzing multidimensional neuronal and behavioral relationships. *arXiv preprint arXiv:2008.01007*.

* The paper proposes a scalable framework for modeling neuronal population activity and external covariates (e.g. behavior). The advantage of the method is the mutual information estimation: it provides more accurate estimates than non-copula methods, especially for large datasets. The GP parametrization allows for conditioning on continuous variables, e.g. on time or position, making the method suitable for studying dynamic trajectories.

[38] Kulkarni, J. E. and Paninski, L. (2007). Common-input models for multiple neural spike-train data. *Network: Computation in Neural Systems*, 18(4):375–407.

[39] Lawhern, V., Wu, W., Hatsopoulos, N., and Paninski, L. (2010). Population decoding of motor cortical activity using a generalized linear model with hidden states. *Journal of neuroscience methods*, 189(2):267–280.

[40] Lawrence, N. and Hyvärinen, A. (2005). Probabilistic non-linear principal component analysis with gaussian process latent variable models. *J Mach Learn Res*, 6(11).

[41] Lefebvre, B., Yger, P., and Marre, O. (2016). Recent progress in multi-electrode spike sorting methods. *Journal of Physiology-Paris*, 110(4, Part A):327–335.

[42] Linderman, S., Johnson, M., Miller, A., Adams, R., Blei, D., and Paninski, L. (2017). Bayesian learning and inference in recurrent switching linear dynamical systems. In *Artificial Intelligence and Statistics*, pages 914–922. PMLR.

[43] Macke, J. H., Berens, P., Ecker, A. S., Tolias, A. S., and Bethge, M. (2009). Generating spike trains with specified correlation coefficients. *Neural Comput*, 21.

[44] Macke, J. H., Buesing, L., Cunningham, J. P., Yu, B. M., Shenoy, K. V., and Sahani, M. (2012). Empirical models of spiking in neural populations. In *Advances in Neural Information Processing Systems 24: 25th conference on Neural Information Processing Systems (NIPS 2011)*, pages 1350–1358.

[45] Magland, J., Jun, J. J., Lovero, E., Morley, A. J., Hurwitz, C. L., Buccino, A. P., Garcia, S., and Barnett, A. H. (2020). SpikeForest, reproducible web-facing ground-truth validation of automated neural spike sorters. *eLife*, 9:e55167.

[46] Mahuas, G., Isacchini, G., Marre, O., Ferrari, U., and Mora, T. (2020). A new inference approach for training shallow and deep generalized linear models of noisy interacting neurons. *arXiv preprint arXiv:2006.06497*.

** The paper proposes a method for training GLMs with complex correlated stimuli. It disentangles stimulus correlations from couplings, which improves generalization and prevents self-excitation in the network dynamics.

[47] 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.

[48] Mongillo, G., Rumpel, S., and Loewenstein, Y. (2018). Inhibitory connectivity defines the realm of excitatory plasticity. *Nat Neurosci*, 21(10):1463–1470.

[49] Nawrot, M., Aertsen, A., and Rotter, S. (1999). Single-trial estimation of neuronal firing rates: from single-neuron spike trains to population activity. *Journal of neuroscience methods*, 94(1):81–92.

[50] Nemati, S., Linderman, S. W., Chen, Z., et al. (2014). A probabilistic modeling approach for uncovering neural population rotational dynamics. In *Computational and Systems Neuroscience (Cosyne)*.

[51] O'Donnell, C., alves, J. T. G., Whiteley, N., Portera-Cailliau, C., and Sejnowski, T. J. (2017). The Population Tracking Model: A Simple, Scalable Statistical Model for Neural Population Data. *Neural Comput*, 29(1):50–93.

[52] 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*, pages 1–6.

[53] Okun, M., Steinmetz, N. A., Cossell, L., Iacaruso, M. F., Ko, H., Barthó, P., Moore, T., Hofer, S. B., Mrsic-Flogel, T. D., Carandini, M., et al. (2015). Diverse coupling of neurons to populations in sensory cortex. *Nature*, 521(7553):511–515.

[54] Onken, A. and Panzeri, S. (2016). Mixed vine copulas as joint models of spike counts and local field potentials. In *Advances in Neural Information Processing Systems*.

[55] Panas, D., Amin, H., Maccione, A., Muthmann, O., van Rossum, M., Berdondini, L., and Hennig, M. H. (2015). Sloppiness in spontaneously active neuronal networks. *J Neurosci*, 35(22):8480–8492.

[56] Pandarinath, 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. *Nat Methods*, 15(10):805–815.

** This paper introduces a novel RNN-based, state-space model termed latent factor analysis via dynamical systems (LFADS). LFADS utilizes a variational autoencoder to amortize the cost of inference over the dynamics making it more practical for large-scale neural recordings.

[57] Petreska, B., Yu, B. M., Cunningham, J. P., Santhanam, G., Ryu, S. I., Shenoy, K. V., and Sahani, M. (2011). Dynamical segmentation of single trials from population neural data. In *Proceedings of the 24th International Conference on Neural Information Processing Systems*, NIPS'11, pages 756–764, Red Hook, NY, USA. Curran Associates Inc.

[58] Pfau, D., Pnevmatikakis, E. A., and Paninski, L. (2013). Robust learning of low-dimensional dynamics from large neural ensembles. In *Advances in Neural Information Processing Systems*.

[59] Pillow, J. W., Ahmadian, Y., and Paninski, L. (2011). Model-based decoding, information estimation, and change-point detection techniques for multineuron spike trains. *Neural Comput*, 23(1):1–45.

[60] Pillow, J. W., Paninski, L., Uzzell, V. J., Simoncelli, E. P., and Chichilnisky, E. J. (2005). Prediction and decoding of retinal ganglion cell responses with a probabilistic spiking model. *J Neurosci*, 25(47):11003–13.

[61] 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(7207):995–999.

[62] Ponce-Alvarez, A., Mochol, G., Hermoso-Mendizabal, A., de la Rocha, J., and Deco, G. (2020). Cortical state transitions and stimulus response evolve along stiff and sloppy parameter dimensions, respectively. *eLife*, 9:e53268.

* Using maximum entropy models, this paper confirms that population activity in cortical networks is constrained only by a small fraction of highly sensitive coupling parameters. In *in vivo* recordings from auditory cortex, the authors show that transitions between synchronised and de-synchronised states reflect changes in sensitive parameters. In contrast, auditory stimuli affect primarily the insensitive couplings, hence stimulation has only a minor effect on the overall statistics of cortical activity.

[63] Prince, L. Y., Bakhtiari, S., Gillon, C. J., and Richards, B. A. (2021). Parallel inference of hierarchical latent dynamics in two-photon calcium imaging of neuronal populations. *bioRxiv*.

[64] Rabinovich, M. I., Huerta, R., Varona, P., and Afraimovich, V. S. (2008). Transient cognitive dynamics, metastability, and decision making. *PLoS Comput Biol*, 4(5):e1000072.

[65] Raducanu, B. C., Yazicioglu, R. F., Lopez, C. M., Ballini, M., Putzeys, J., Wang, S., Andrei, A., Rochus, V., Welkenhuysen, M., Helleputte, N. v., et al. (2017). Time multiplexed active neural probe with 1356 parallel recording sites. *Sens*, 17(10):2388.

[66] Ramanathan, D. S., Guo, L., Gulati, T., Davidson, G., Hishinuma, A. K., Won, S.-J., Knight, R. T., Chang, E. F., Swanson, R. A., and Ganguly, K. (2018). Low-frequency cortical activity is a neuromodulatory target that tracks recovery after stroke. *Nat Med*, 24(8):1257–1267.

[67] René, A., Longtin, A., and Macke, J. H. (2020). Inference of a mesoscopic population model from population spike trains. *Neural Comput*, 32(8):1448–1498.

[68] Rey, H. G., Pedreira, C., and Quian Quiroga, R. (2015). Past, present and future of spike sorting techniques. *Brain Res Bull*, 119:106–117.

[69] Roweis, S. and Ghahramani, Z. (1999). A unifying review of linear gaussian models. *Neural Comput*, 11(2):305–345.

[70] Rule, M. E., Schnoerr, D., Hennig, M. H., and Sanguinetti, G. (2019). Neural field models for latent state inference: Application to large-scale neuronal recordings. *PLoS Comput Biol*, 15(11):e1007442.

[71] Runyan, C. A., Piasini, E., Panzeri, S., and Harvey, C. D. (2017). Distinct timescales of population coding across cortex. *Nature*, 548.

[72] Russo, A. A., Khajeh, R., Bittner, S. R., Perkins, S. M., Cunningham, J. P., Abbott, L. F., and Churchland, M. M. (2020). Neural trajectories in the supplementary motor area and motor cortex exhibit distinct geometries, compatible with different classes of computation. *Neuron*, 107(4):745–758.

* The paper explores a new non-linear pattern of activity in supplementary motor area during multi-cycle rhythmic movement: helical population-trajectory geometry. The separation between single-cycles in these trajectories allows for counting the number of revolutions.

[73] Rutten, V., Bernacchia, A., Sahani, M., and Hennequin, G. (2020). Non-reversible gaussian processes for identifying latent dynamical structure in neural data. In *Advances in Neural Information Processing Systems*, volume 33.

** The first paper to consider time-irreversibility in GPFA models by constructing a new kernel function. This method, termed GPFADS, assigns lower probabilities to the time-reversed variants of the true trajectories. Unlike traditional GPFA, it can disentangle trajectories in the arm reach dataset [13].

[74] Sadtler, P. T., Quick, K. M., Golub, M. D., Chase, S. M., Ryu, S. I., Tyler-Kabara, E. C., Byron, M. Y., and Batista, A. P. (2014). Neural constraints on learning. *Nature*, 512(7515):423–426.

[75] Saha, D., Leong, K., Li, C., Peterson, S., Siegel, G., and Raman, B. (2013). A spatiotemporal coding mechanism for background-invariant odor recognition. *Nat. Neurosci.*, 16(12):1830–1839.

[76] Sani, O. G., Abbaspourazad, H., Wong, Y. T., Pesaran, B., and Shanechi, M. M. (2020). Modeling behaviorally relevant neural dynamics enabled by preferential subspace identification. *Nat Neurosci*, pages 1–10.

** This paper introduces a novel linear state-space model (PSID) that jointly models the neural activity and the recorded behavioural variables as generated by the behaviourally relevant and irrelevant latent dynamics. The authors show that PSID discovers significantly lower-dimensional latent dynamics than other models and is the only one that can capture reverse directional rotational dynamics in the primary motor cortex when the arm is returning versus when the arm is reaching.

[77] Santhanam, G., Yu, B. M., Gilja, V., Ryu, S. I., Afshar, A., Sahani, M., and Shenoy, K. V. (2009). Factor-analysis methods for higher-performance neural prostheses. *J Neurophysiol*, 102(2):1315–1330.

[78] Savin, C. and Tkačik, G. (2017). Maximum entropy models as a tool for building precise neural controls. *Curr Opin Neurobiol*, 46:120–126.

[79] Saxena, S. and Cunningham, J. P. (2019). Towards the neural population doctrine. *Curr Opin Neurobiol*, 55:103–111.

[80] 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.

[81] She, Q. and Wu, A. (2020). Neural dynamics discovery via gaussian process recurrent neural networks. In *Uncertainty in Artificial Intelligence*, pages 454–464. PMLR.

&ast; This paper extends RNN-based dynamical modelling approaches with a Gaussian Process observation model. They show that their more expressive observation model is able to extract insightful low-dimensional dynamics and better fit neural datasets.

[82] Shenoy, K. V., Sahani, M., and Churchland, M. M. (2013). Cortical control of arm movements: a dynamical systems perspective. *Annu Rev Neurosci*, 36.

[83] Shlens, J., Field, G. D., Gauthier, J. L., Grivich, M. I., Petrusca, D., Sher, A., Litke, A. M., and Chichilnisky, E. J. (2006). The structure of multi-neuron firing patterns in primate retina. *J Neurosci*, 26(32):8254–66.

[84] Si, G., Kanwal, J. K., Hu, Y., Tabone, C. J., Baron, J., Berck, M., Vignoud, G., and Samuel, A. D. (2019). Structured odorant response patterns across a complete olfactory receptor neuron population. *Neuron*, 101(5):950–962.

[85] Smith, A. C. and Brown, E. N. (2003). Estimating a state-space model from point process observations. *Neural Comput*, 15(5):965–991.

[86] Spreemann, G., Dunn, B., Botnan, M. B., and Baas, N. A. (2018). Using persistent homology to reveal hidden covariates in systems governed by the kinetic ising model. *Phys Rev E*, 97(3):032313.

[87] Stringer, C., Pachitariu, M., Steinmetz, N., Carandini, M., and Harris, K. D. (2019). High-dimensional geometry of population responses in visual cortex. *Nature*, 571(7765):361–365.

[88] Tkačik, G., Marre, O., Mora, T., Amodei, D., Berry, M. J., and Bialek, W. (2013). The simplest maximum entropy model for collective behavior in a neural network. *J Stat Mech Theory Exp*, 2013.

[89] Trippe, B., Huggins, J., Agrawal, R., and Broderick, T. (2019). Lr-glm: High-dimensional Bayesian inference using low-rank data approximations. In *International Conference on Machine Learning*, pages 6315–6324. PMLR.

[90] Truccolo, W., Eden, U. T., Fellows, M. R., Donoghue, J. P., and Brown, E. N. (2005). A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. *J Neurophysiol*, 93(2):1074–1089.

[91] Ventura, V. (2009). Traditional waveform based spike sorting yields biased rate code estimates. *Proc Natl Acad Sci*, 106(17):6921–6926.

[92] Ventura, V. and Gerkin, R. C. (2012). Accurately estimating neuronal correlation requires a new spike-sorting paradigm. *Proc Natl Acad Sci*, 109(19):7230–7235.

[93] Whiteway, M. R. and Butts, D. A. (2019). The quest for interpretable models of neural population activity. *Curr Opin Neurobiol*, 58:86–93.

[94] Wu, A., Roy, N. A., Keeley, S., and Pillow, J. W. (2017). Gaussian process based nonlinear latent structure discovery in multivariate spike train data. *Adv Neur In*, 30:3496.

[95] Zhao, Y. and Park, I. M. (2016). Interpretable nonlinear dynamic modeling of neural trajectories. *arXiv preprint arXiv:1608.06546.*

[96] Zhao, Y. and Park, I. M. (2017). Variational latent gaussian process for recovering single-trial dynamics from population spike trains. *Neural Comput*, 29(5):1293–1316.

[97] Zhou, D. and Wei, X.-X. (2020). Learning identifiable and interpretable latent models of high-dimensional neural activity using pi-vae. *arXiv preprint arXiv:2011.04798.*

[98] Zoltowski, D., Pillow, J., and Linderman, S. (2020). A general recurrent state space framework for modeling neural dynamics during decision-making. In *International Conference on Machine Learning*, pages 11680–11691. PMLR.

[99] Zoltowski, D. M. and Pillow, J. W. (2018). Scaling the poisson glm to massive neural datasets through polynomial approximations. In *Advances in Neural Information Processing Systems*, volume 31, page 3517. NIH Public Access.

[100] Zylberberg, J., Cafaro, J., Turner, M. H., Shea-Brown, E., and Rieke, F. (2016). Direction-selective circuits shape noise to ensure a precise population code. *Neuron*, 89.