iLQR-VAE : control-based learning of input-driven dynamics with applications to neural data

Marine Schimel[@1], Ta-Chu Kao[1], Kristopher T. Jensen[1], and Guillaume Hennequin[1]

[1]Computational and Biological Learning Lab, Department of Engineering, University of Cambridge, Cambridge, U.K.

[@] Corresponding author (mmcs3@cam.ac.uk)

Abstract

Understanding how neural dynamics give rise to behaviour is one of the most fundamental questions in systems neuroscience. To achieve this, a common approach is to record neural populations in behaving animals, and model these data as emanating from a latent dynamical system whose state trajectories can then be related back to behavioural observations via some form of decoding. As recordings are typically performed in localized circuits that form only a part of the wider implicated network, it is important to simultaneously learn the local dynamics and infer any unobserved external input that might drive them. Here, we introduce iLQR-VAE, a novel control-based approach to variational inference in nonlinear dynamical systems, capable of learning both latent dynamics, initial conditions, and ongoing external inputs. As in recent deep learning approaches, our method is based on an input-driven sequential variational autoencoder (VAE). The main novelty lies in the use of the powerful iterative linear quadratic regulator algorithm (iLQR) in the recognition model. Optimization of the standard evidence lower-bound requires differentiating through iLQR solutions, which is made possible by recent advances in differentiable control. Importantly, having the recognition model implicitly defined by the generative model greatly reduces the number of free parameters and allows for flexible, high-quality inference. This makes it possible for instance to evaluate the model on a single long trial after training on smaller chunks. We demonstrate the effectiveness of iLQR-VAE on a range of synthetic systems, with autonomous as well as input-driven dynamics. We further show state-of-the-art performance on neural and behavioural recordings in non-human primates during two different reaching tasks.

1 Introduction

The mammalian brain is a complex, high-dimensional system, containing billions of neurons whose coordinated dynamics ultimately drives behaviour. Identifying and interpreting these dynamics is the focus of a large body of neuroscience research, which is being facilitated by the advent of new experimental techniques that allow large-scale recordings of neural populations (Jun et al., 2017; Stosiek et al., 2003). A range of methods have been developed that identify dynamics from data by learning a generative model, typically composed of a flexible latent dynamical system driven by process noise combined with an appropriate observation model (Archer et al., 2015; Buesing et al., 2012; Duncker et al., 2019; Gao et al., 2016; Hernandez et al., 2018; Kim et al., 2021; Nguyen et al., 2020; She and Wu, 2020).

However, neural recordings are typically only made in a small selection of brain regions, leaving many areas unobserved which might provide relevant task-related input to the recorded one(s). Yet, the aforementioned methods perform Bayesian inference of state trajectories directly, and therefore do not support inference of external input (which they effectively treat as process noise and marginalize out). Indeed, simultaneous learning of latent dynamics *and* inference of unobserved control inputs is a challenging problem that involves teasing apart momentary variations in the data that can be attributed to the system's internal transition function, and those that need to be explained by unobserved inputs. Previously, this distinction has been achieved e.g. by introducing external control in the form of abrupt changes in the latent state transition function, and inferring these switching events (Ghahramani and Hinton, 2000; Linderman et al., 2017). Alternatively, the data can be modelled using an input-driven sequential VAE such as LFADS (Pandarinath

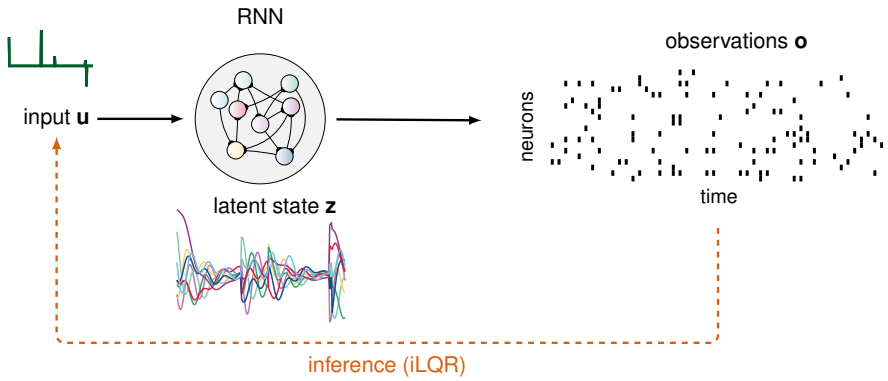

Figure 1: **Illustration of the model.** iLQR-VAE is trained to model the joint density of noisy observations (in this case, a population of neural spike trains), assuming that they are generated by an underlying input-driven dynamical system. At every training iteration, iLQR is used to infer the input $\boldsymbol{u}$ (green; here, the ground truth is a series of sporadic pulses) given the observations $\boldsymbol{o}$ and the current parameters of the generator. Intuitively, the inputs inferred by iLQR are ones that produce latent trajectories in the RNN that are most compatible with the observed data, without overfitting.

et al., 2018) where (amortized) inference is performed at the level of external inputs as well as initial latent states. Encouragingly, LFADS has been able to inferred inputs that are congruent with task-induced perturbations in various reaching tasks in primates (Keshtkaran et al., 2021; Pandarinath et al., 2018).

Here, we introduce iLQR-VAE, a new method for learning input-driven latent dynamics from data. As in LFADS, we use an input-driven sequential VAE to encode observations into a set of initial conditions and external inputs driving an RNN generator. However, while LFADS uses a separate, bidirectional RNN as the encoder, here we substitute the inference network with an *optimization-based recognition model* that relies on the powerful iterative linear quadratic regulator algorithm (iLQR, Li and Todorov, 2004). iLQR solves an optimization problem that finds a mode of the exact posterior over inputs for the current setting of generative parameters. This ensures that the encoder mean remains optimal for every update of the decoder, thus reducing the amortization gap (Cremer et al., 2018). Moreover, having the recognition model implicitly defined by the generative model stabilizes training, circumvents the need for tricks such as KL warmup, and greatly reduces the number of (hyper-)parameters.

While iLQR-VAE could find applications in many fields as a general approach to learning stochastic nonlinear dynamical systems, here we focus on neuroscience case studies. We first demonstrate in a series of synthetic examples that iLQR-VAE can recover the true dynamics in both autonomous and input-driven systems. Next, we show state-of-the art results on recordings performed in the primary motor cortex (M1) of monkeys during two types of reaching tasks (Churchland et al., 2010; O'Doherty et al., 2018). In particular, we show that hand kinematics can be accurately decoded from inferred latent state trajectories, and that the inferred inputs are consistent with recently proposed theories of motor preparation.

2 Method

iLQR-VAE models a set of temporal observations, such as behavioural and/or neural recordings, through a shared input-driven nonlinear latent dynamical system (Figure 1). The input encapsulates both process noise (as in traditional latent dynamics models), initial inputs that set the initial condition of the dynamics, and any meaningful task-related control input. In this section, we describe the architecture of the generative model, and the control-based variational inference strategy used for training the model and making predictions.

2.1 Generative model

We consider the following generative model:

$$\text{latent state} \quad \boldsymbol{z}_{t+1} = f_\theta(\boldsymbol{z}_t, \boldsymbol{u}_t, t) \tag{1}$$

$$\text{observations} \quad \boldsymbol{o}_t|\boldsymbol{z}_t \sim p_\theta(\boldsymbol{o}_t|\boldsymbol{z}_t) \tag{2}$$

where $\boldsymbol{u}_t \in \mathbb{R}^m$, $\boldsymbol{z}_t \in \mathbb{R}^n$ and $\boldsymbol{o}_t \in \mathbb{R}^{n_o}$ are the input, latent state and observations at time t, respectively. Here, observations may comprise either neural activity, behavioural variables, or both – the distinction will be made later where relevant. We use the notation θ to denote the set of all parameters of the generative model. We use $\boldsymbol{u}_0$ to set the initial condition $\boldsymbol{z}_1 = f_\theta(\boldsymbol{0}, \boldsymbol{u}_0, 0)$ of the network[1]. This way, the latent state trajectory of the network $\boldsymbol{z}(\boldsymbol{u}) = \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_T\}$ is entirely determined by the input sequence $\boldsymbol{u} = \{\boldsymbol{u}_0, \ldots, \boldsymbol{u}_T\}$ and the state transition function $f_\theta(\cdot)$, according to Equation 1. For $f_\theta(\cdot)$, we use either standard linear or GRU-like RNN dynamics (see Appendix A for details). For the likelihoods, we use Gaussian or Poisson distributions with means given by linear or nonlinear readouts of the network state of the form $\bar{\boldsymbol{o}}_t = h(\boldsymbol{C}\boldsymbol{z}_t + \boldsymbol{b})$ (Appendix B).

We place a Gaussian prior over $\boldsymbol{u}_{t\leq 0}$. We then consider two alternative choices for the prior over $\boldsymbol{u}_{t>0}$. The first is a Gaussian prior

$$p_\theta(\boldsymbol{u}_{t>0}) = \mathcal{N}(0, \boldsymbol{S}^2) \tag{3}$$

with $\boldsymbol{S} = \text{diag}(s_1, \ldots, s_m)$. In many settings however, we expect inputs to enter the system in a sparse manner. To explicitely model this, we introduce a second prior over $\boldsymbol{u}$ in the form of a heavy-tailed distribution constructed hierarchically by assuming that the ith input at time $t > 0$ is

$$u_{it} = s_i \epsilon_{it} \sqrt{\nu/\alpha_t} \tag{4}$$

where $s_i > 0$ is a scale factor, $\epsilon_{it} \sim \mathcal{N}(0, 1)$ is independent across i and t, and $\alpha_t \sim \chi^2_\nu$ is a shared scale factor drawn from a chi-squared distribution with ν degrees of freedom. Thus, inputs are spatially and temporally independent a priori, such that any spatio-temporal structure in the observations will have to be explained by the coupled dynamics of the latent states. Moreover, the heavy-tailed nature of this prior allows for strong inputs when they are needed. Finally, the fact that the scale factor is shared across input dimensions means that inputs are either all weak or potentially all strong at the same time for all input channels, expressing the prior belief that inputs come as shared events.

This hierarchical construction induces a multivariate Student prior at each time step:

$$p_\theta(\boldsymbol{u}_t) = \frac{\Gamma\left[(\nu+m)/2\right]}{\Gamma\left[\nu/2\right](\nu\pi)^{m/2}|\boldsymbol{S}|}\left[1 + \frac{1}{\nu}\boldsymbol{u}_t^T \boldsymbol{S}^{-2}\boldsymbol{u}_t\right]^{-(\nu+m)/2} \tag{5}$$

where $\boldsymbol{S} = \text{diag}(s_1, \ldots, s_m)$. Note that both $\boldsymbol{S}$ and ν are parameters of the generative model, which we will learn.

2.2 iLQR-VAE: a novel control-based variational inference strategy

To train the model, we optimize θ to maximize the log-likelihood of observing a collection of independent observation sequences $\boldsymbol{\mathcal{O}} = \{\boldsymbol{o}^{(1)}, \ldots, \boldsymbol{o}^{(K)}\}$, or "trials", given by:

$$\log p_\theta(\boldsymbol{\mathcal{O}}) = \sum_{k=1}^{K} \log \int p_\theta(\boldsymbol{o}^{(k)}|\boldsymbol{z}(\boldsymbol{u}))p_\theta(\boldsymbol{u})\, d\boldsymbol{u}. \tag{6}$$

As the integral is in general intractable, we resort to an amortized variational inference strategy by introducing a recognition model $q_\phi(\boldsymbol{u}|\boldsymbol{o}^{(k)})$ to approximate the posterior $p_\theta(\boldsymbol{u}|\boldsymbol{o}^{(k)})$. Following standard practice (Kingma

[1]Note that when $m < n$, $\boldsymbol{u}_0$ can only reach an m-dimensional subspace of initial conditions, which could be limiting. We can circumvent this problem by spreading $\boldsymbol{u}_0$ over multiple surrogate time bins before the start of the trial, i.e. introduce $\{\boldsymbol{u}_{-n/m}, \ldots, \boldsymbol{u}_{-2}, \boldsymbol{u}_{-1}, \boldsymbol{u}_0\}$ together with an appropriate dependence of f_θ on $t \leq 0$ in Equation 1, such that each of these surrogate inputs target a different latent subspace with purely integrating ("sticking") linear dynamics before $t = 1$.

and Welling, 2013; Rezende et al., 2014), we thus train the model by maximizing the evidence lower-bound (ELBO):

$$\mathcal{L}(\boldsymbol{\mathcal{O}}, \theta, \phi) = \sum_k \mathbb{E}_{q_\phi(\boldsymbol{u}|\boldsymbol{o}^{(k)})} \left[\log p_\theta(\boldsymbol{o}^{(k)}|\boldsymbol{u}) + \log p_\theta(\boldsymbol{u}) - \log q_\phi(\boldsymbol{u}|\boldsymbol{o}^{(k)}) \right] \tag{7}$$

$$= \sum_k \mathbb{E}_{q_\phi(\boldsymbol{u}|\boldsymbol{o}^{(k)})} \left[\sum_{t=1}^{T} \log p_\theta(\boldsymbol{o}_t^{(k)}|\boldsymbol{z}_t) + \log p_\theta(\boldsymbol{u}_t) - \log q_\phi(\boldsymbol{u}_t|\boldsymbol{o}^{(k)}) \right] \tag{8}$$

$$\leq \log p_\theta(\boldsymbol{\mathcal{O}}). \tag{9}$$

with respect to both θ and ϕ.

Here, the main novelty is the use of an optimization-based recognition model. We reason that maximizing the exact log posterior, i.e. computing

$$\boldsymbol{u}^\star(\boldsymbol{o}^{(k)}) = \underset{\boldsymbol{u}}{\operatorname{argmax}} \ \log p_\theta(\boldsymbol{u}|\boldsymbol{o}^{(k)}) \tag{10}$$

$$= \underset{\boldsymbol{u}}{\operatorname{argmax}} \left[\sum_{t=1}^{T} \log p_\theta(\boldsymbol{o}_t^{(k)}|\boldsymbol{u}) + \log p_\theta(\boldsymbol{u}_t) \right] \tag{11}$$

subject to the generative dynamics of Equations 1 and 2, is a standard nonlinear control problem: $\log p_\theta(\boldsymbol{o}_t^{(k)}|\boldsymbol{u})$ acts as a running cost penalizing momentary deviations between desired outputs $\boldsymbol{o}_t$ and the actual outputs caused by a set of controls $\boldsymbol{u}$, and $\log p_\theta(\boldsymbol{u}_t)$ acts as an energetic cost on those controls. Importantly, there exists a general purpose, efficient algorithm to solve such nonlinear control problems: iLQR (Li and Todorov, 2004; Appendix C). We thus propose to use a black-box iLQR solver to parameterize the mean of the recognition density $q_\phi(\boldsymbol{u}|\boldsymbol{o})$ for any $\boldsymbol{o}$, and to model uncertainty separately using a multivariate Gaussian density common to all trials. Therefore, we parametrize the recognition model as follows:

$$q_\phi(\boldsymbol{u}|\boldsymbol{o}) = \mathcal{N}(\boldsymbol{u}; \boldsymbol{u}^\star(\boldsymbol{o}), \boldsymbol{\Sigma}_{\text{s}} \otimes \boldsymbol{\Sigma}_{\text{t}}) \tag{12}$$

$$\text{with } \boldsymbol{u}^\star(\boldsymbol{o}) = \text{iLQRsolve}(\boldsymbol{o}, \theta). \tag{13}$$

where we use a separable posterior covariance (the Kronecker product of a spatial factor $\boldsymbol{\Sigma}_{\text{s}}$ and a temporal factor $\boldsymbol{\Sigma}_{\text{t}}$).

To optimize the ELBO, we estimate the expectation in Equation 8 by drawing samples from $q_\phi(\boldsymbol{u}|\boldsymbol{o}^{(k)})$ and using the reparameterization trick (Kingma et al., 2015) to obtain gradients. A major complication that would normally preclude the use of optimization-based recognition models is the need to differentiate through the mean of the posterior. In this case, this involves differentiating through an entire optimization process. Using automatic differentiation within the iLQR solver is in general impractically expensive memory-wise. However, recent advances in differentiable model predictive control enable implicit differentiation through iLQRsolve with a memory cost that does not depend on the number of iterations (Amos et al., 2018; Blondel et al., 2021; Appendix D).

2.3 Complexity and implementation

We optimize the ELBO using Adam (Kingma and Ba, 2014) with a decaying learning rate $\propto 1/\sqrt{i}$ where i is the iteration number. Averaging over data samples can be easily parallelized; we do this here using the MPI library and a local CPU cluster. In each iteration and for each data sample, obtaining the approximate posterior mean through iLQR is the main computational bottleneck, with a complexity of $\mathcal{O}(T(n^3 + n^2 n_o))$. To help mitigate this cost, we find it useful to re-use the previously inferred control inputs to initialize each iLQRsolve.

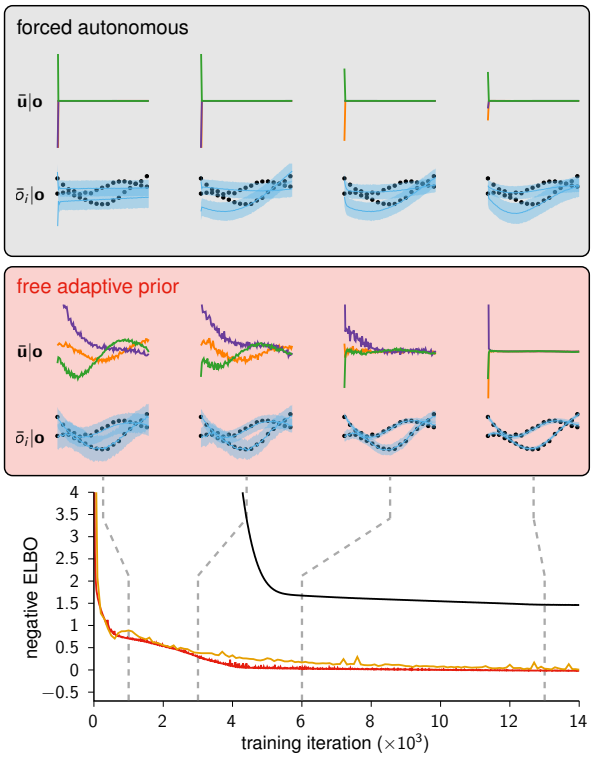

Figure 2: **Fast and robust learning in iLQR-VAE.** Example run of iLQR-VAE on a synthetic dataset generated by an autonomous linear dynamical system seeded with random initial conditions. iLQR-VAE with an adaptive prior over control inputs (red, 'free adaptive prior') uses large inputs to fit the observations initially, but gradually pushes those back into initial conditions as it acquires the ground truth autonomous dynamics. In contrast, iLQR-VAE with a rigid input prior imposing autonomous dynamics throughout learning struggles to learn (black, 'forced autonomous'; see text for details). For each setting, insets show the three inferred inputs for a given test trial (top; posterior mean $\bar{\boldsymbol{u}}|\boldsymbol{o}$), and the posterior predictions for the first two corresponding outputs ($\bar{o}_i|\boldsymbol{o}$; black dots: ground truth; blue: posterior mean with 95% c.i.). We also compare learning curves with LFADS (yellow curve). We used the same generator architecture in all scenarios, and the learning rate was hand-tuned for each method.

3 Experiments and results

3.1 iLQR-VAE enables fast learning of dynamics

Before demonstrating the method on a number of synthetic and real datasets involving ongoing external inputs, we begin with a simpler example meant to illustrate some of the main advantages of iLQR-VAE (Figure 2). We generated data from an autonomous (i.e. non-input-driven) linear dynamical system ($n = 8$, $m = 3$ where n is the number of latents and m the number of input channels) seeded with a random initial condition in each of 56 trials. The state $\boldsymbol{z}_t$ was linearly decoded with added Gaussian noise to produce observation data, which we used to train a model in the same class.

At the beginning of learning, iLQR-VAE originally relies on large ongoing inputs that control the generator into producing outputs very similar to the observations in the data (Figure 2, red box, left), resulting in a rapidly decreasing loss. Subsequently, the amount of input required to fit the observations gradually decreases as the system learns the internal dynamics of the ground truth system. Eventually, the inferred control inputs become confined to the first time bin, i.e. they act as initial conditions for the now autonomous dynamics of the generative model. Thus, iLQR-VAE operates in a regime where the output of the generator explains the data well at all times, and learning consists in make the inputs more parsimonious. We note that this regime is facilitated here by our choice of generator dynamics, which we initialised to be very weak initially and therefore easily controllable.

We contrast this with learning in a modified version of iLQR-VAE where we allowed $\boldsymbol{u}_0$ to vary freely (with a Gaussian prior of adjustable variance) but effectively fixed $\boldsymbol{u}_{t>0}$ to be 0. In other words, we constrained the dynamics of the generator to remain (near-)autonomous throughout learning (Figure 2, grey box, top). Although this incorporates important information about the ground truth generator (which is itself autonomous), counter-intuitively we found that it makes learning worse. Indeed, iLQR is unable to find initial conditions that would explain the data well, resulting in a much higher initial loss. The loss does not decrease as much as in normal iLQR-VAE, indicating that the ground truth dynamics has not been acquired.

On the same toy dataset, we also compare iLQR-VAE to LFADS (Pandarinath et al., 2018), keeping the generative model in the same model class (see Appendix F for details). We find that LFADS learns in a

similar manner to iLQR-VAE (Figure 2), also progressively doing away with inputs.

3.2 iLQR-VAE for nonlinear system identification

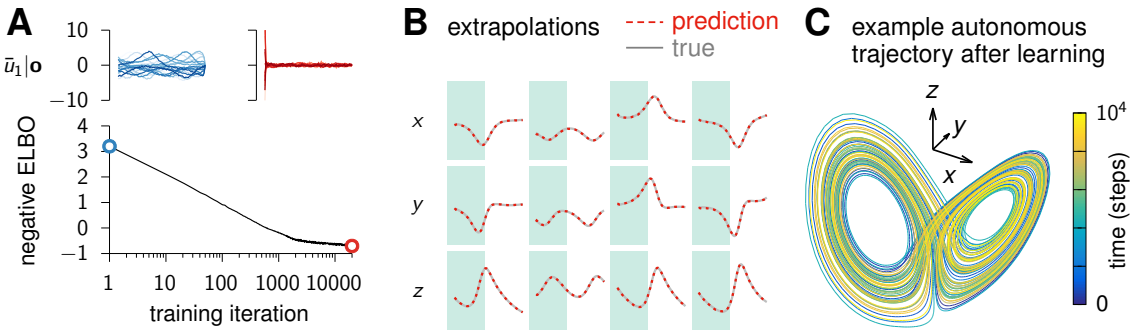

Figure 3: **Learning nonlinear autonomous dynamics.** **(A)** Evolution of the negative ELBO during training of an iLQR-VAE with a nonlinear RNN and a Gaussian likelihood ($n = 20$, $m = 5$, $n_o = 3$). The inputs inferred by iLQR are initially strong (blue, before learning), but are progressively pushed to initial conditions (red, after learning) as the autonomous dynamics of the Lorenz attractor are approximated with increasing accuracy. **(B)** Four example test trajectories (columns), conditioning on the noisy data within the first half (green shading), and predicting in the second half. **(C)** Single long autonomous trajectory after training (setting $\boldsymbol{u}_{t>0} = 0$), starting from a random initial condition and running the dynamics for 10000 steps. The model displays the butterfly topology characteristic of the Lorenz system, and completes multiple cycles without deviating from the attractor, suggesting that the ground-truth dynamics have been learned.

Next, we illustrate the method on an autonomous nonlinear dynamical system, the chaotic Lorenz attractor (Lorenz, 1963; Appendix H). This is a standard benchmark to evaluate system identification methods on nonlinear dynamics (Champion et al., 2019; Hernandez et al., 2018; Nguyen et al., 2020), and one typically considers the dynamics to be learned if the trained model can recreate the whole attractor structure.

Here, we show that iLQR-VAE can learn these complex nonlinear dynamics. Before training, the inferred inputs were large and explained the output observations by forcing the internal state of the generator into appropriate trajectories (Figure 3A, top). At the end of learning, the inputs were confined to initial conditions, with only weak ongoing corrections to the now stronger, near-autonomous dynamics of the generator. In Figure 3B we show that, conditioned on an initial bout of test data, the model perfectly predicts the rest of the trajectory. Moreover, starting from a random initial condition, the model can recreate the whole attractor structure (Figure 3C).

To quantitatively assess how well the dynamics have been learned, we computed the k-step coefficient of determination, R_k^2 (Appendix H), as in Hernandez et al. (2018). This metric evaluates how well the model can predict the true state k steps into the future, starting from any state inferred along a test trajectory. Hernandez et al. reported $R_{30}^2 \approx 1$ but did not show results for larger k. For iLQR-VAE, $R_{30}^2 = 0.998$ and the forward interpolation was still very high at 50 time steps, with $R_{50}^2 = 0.996$.

3.3 Inferring sparse inputs

To demonstrate iLQR-VAE's ability to infer unobserved inputs and learn the ground truth dynamics of an input-driven system, we generated synthetic data from a system with $n = 3, m = 3$ and $n_o = 10$, which evolves with linear dynamics for $T = 1000$ time steps (Figure 4A). The system was driven by sparse inputs, and the output corrupted with Gaussian noise. Input events were drawn in each time step from a Bernoulli distribution with mean $p = 0.03$. Whenever an input event occurred, the magnitude of inputs in each channel was drawn from a standard mutivariate Gaussian distribution.

iLQR-VAE captured most of the variance in these sparse inputs ($R^2 = 0.94 \pm 0.02$; 5 random seeds), and recovered the eigenvalue spectrum of the transition matrix almost perfectly (Figure 4B). We also evaluated

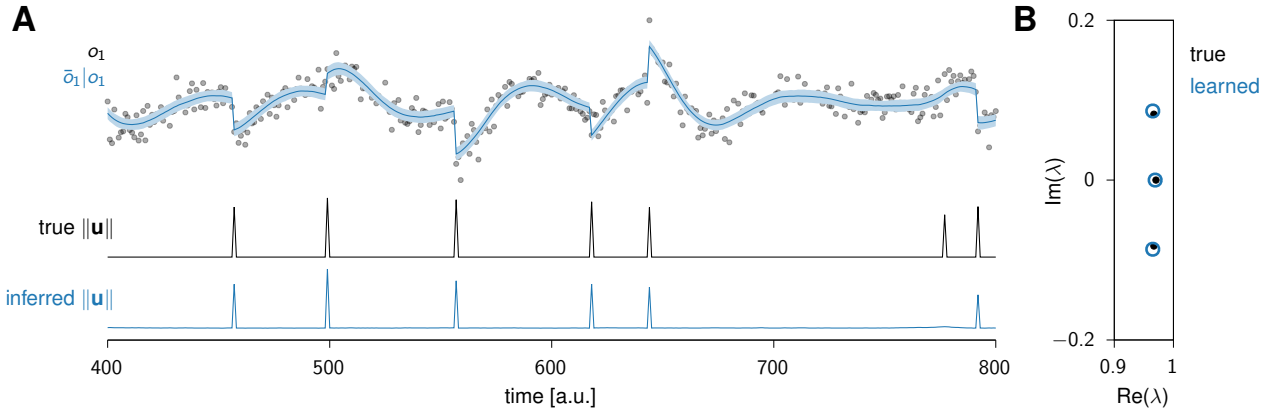

Figure 4: **Inferring sparse inputs to a linear system. (A)** Top: example observations (black dots) and inferred posterior mean (blue line). Bottom: true and inferred inputs. iLQR-VAE is able to infer the timing and magnitude of the inputs almost perfectly, despite being trained on only a single timeseries of 1000 time steps. Note that iLQR-VAE fails to infer one of the smallest inputs, whose effect on the observations is largely masked by observation noise. **(B)** Comparison of the true (black) and learned (blue) eigenvalue spectra. This shows that iLQR-VAE recovers the ground-truth dynamical system up to a similarity transformation.

LFADS on the same data, using a generator with $n = m = 3$ and linear dynamics, and manually optimized hyperparameters. LFADS performed poorly in this case ($R^2 = 0.05 \pm 0.02$ for input reconstruction; 3 random seeds), as well as in several other similar comparisons on datasets of different sizes and trial numbers (Appendix G). This is unsurprising, as LFADS assumes a dense (auto-regressive Gaussian) prior over the inputs, which is not overridden by the relatively small amount of data used here. Nonetheless, when applied to a set of 56 trials of 100 time steps driven by Gaussian autoregressive inputs, iLQR-VAE still captured the structure in the inputs more accurately than LFADS did ($R^2 = 0.81 \pm 0.01$ vs. 0.29 ± 0.06).

3.4 Predicting hand kinematics from primate M1 recordings

3.4.1 Trial-structured maze task

To highlight the utility of iLQR-VAE for analyzing experimental neuroscience data, we next apply it to recordings of monkey motor (M1) and dorsal premotor (PMd) cortices during a delayed reaching task ('Maze' dataset of Kaufman et al., 2016; DANDI 000128). This dataset contains 108 different reach configurations over nearly 3000 trials, and has recently been proposed as a 'neuroscience MNIST' to benchmark neural data analysis methods (Pei et al., 2021). Several latent variable models were evaluated on this dataset in (Pei et al., 2021), allowing us to directly compare the performance of iLQR-VAE to those reported results.

For this experiment, we fitted iLQR-VAE to the neural activity, using a model with GRU dynamics ($n = 60$), a Student prior over inputs ($m = 15$), and a Poisson likelihood ($n_o = 182$ neurons). To allow for direct comparison with baselines reported in Pei et al. (2021), we trained models on 1340 trials spanning all reach conditions and restricting data to a time window starting 250 ms before and ending 450 ms after movement onset, binned at 5ms.

iLQR-VAE inferred inputs that were confined to initial conditions, from which smooth single-trial dynamics evolved near-autonomously (Figure 5A). As a first measure of performance, we evaluated the model on a "co-smoothing" metric defined as the normalized log-likelihood score of held-out neurons' activity (see Appendix I) conditioned on held-in neurons. Conditioning of 137 neurons (i.e using 45 held-out neurons), we obtained a co-smoothing of 0.331 ± 0.001 (over 5 random seeds). For comparison, Pei et al. (2021) reports 0.187 for GPFA (Yu et al., 2009), 0.225 for SLDS (Linderman et al., 2017), 0.323 for Neural Data Transformers (Ye and Pandarinath, 2021) and $R^2 = 0.336$ for AutoLFADS (LFADS with large scale hyperparameter optimization Keshtkaran et al., 2021) on the same dataset.

Next, we assessed how well hand velocity could be decoded from neural activity (another highly relevant

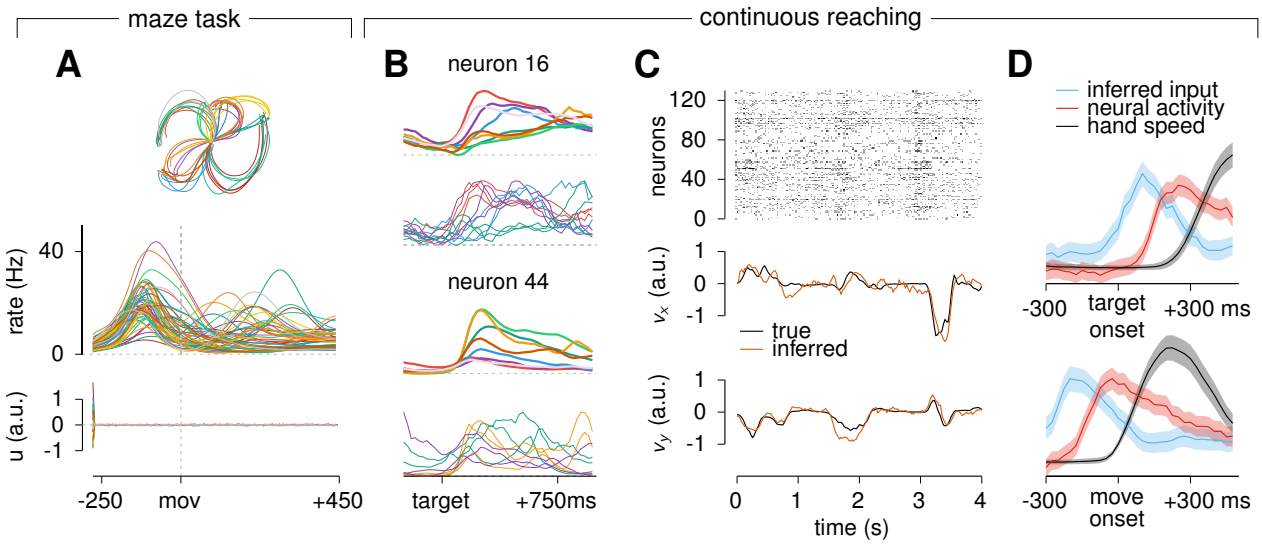

Figure 5: **iLQR-VAE infers smooth firing rates from spiking data, which can be used to accurately decode kinematic variables.** (**A**) 50 example hand trajectories (top) from the monkey reaching 'Maze' dataset, corresponding single-trial firing rate timecourse (middle; one example neuron), and inferred input (bottom; one example input dimension). (**B**) Example spike raster (top) and hand kinematics (bottom) for a 4 second-long chunk of test data in the continuous monkey reaching task. v_x and v_y refer to hand x- and y-velocities respectively. (**C**) Mean (thick) and single-'trial' (thin) firing rates of two example neurons during reaches of various directions (colours), aligned to target onset. Note that three single-'trial' firing rates are shown for only three of the 8 reach directions for which averages are shown. Interestingly, single 'trial' activities evolve tightly around their trial averages, and resembles the firing rate responses shown in A (middle). (**C**) Overall magnitude of the inferred inputs $\|\boldsymbol{u}_t\|$ (blue), average population spiking activity (red), and hand velocity $\|\boldsymbol{y}_t\|$ (black), each z-scored, averaged across movement episodes, and aligned to target onset (top) or movement onset (bottom).

metric for neural data). We applied ridge regression to predict the monkey's hand velocity (with a 100 ms lag) from momentary neuronal firing rates (mean of the posterior predictive distribution) on test data. This reconstruction could be performed with very high accuracy $R^2 = 0.896 \pm 0.002$ (over 5 random seeds), compared to 0.640 for GPFA, 0.795 for SLDS, 0.886 for Neural Data Transformers and 0.910 for AutoLFADS (Pei et al., 2021).

These experiments place iLQR-VAE on par with state-of-the-art methods, without any extensive hyperparameter optimization.

3.4.2 Continuous reaching task

While a large number of neuroscience studies perform neural and behavioural recordings during trial-structured tasks, much can be learned by analyzing the dynamics of more naturalistic, less constrained behaviours. iLQR-VAE's flexible recognition model is well-suited to the analysis of such less structured tasks, as it can easily be trained and tested on trials of heterogeneous lengths. To illustrate this, we applied iLQR-VAE to a self-paced reaching task during which a monkey had to reach to consecutive targets randomly sampled from a 17x8 grid on a screen (Makin et al., 2018; O'Doherty et al., 2018). This dataset consists of both neural recordings from primary motor cortex (M1) together with continuous behavioural recordings in the form of x- and y-velocities of the fingertip.

In this experiment, we used GRU dynamics ($n = 50$), a Gaussian prior over inputs ($m = 10$), and a combination of a Poisson likelihood (130 neurons) and a Gaussian likelihood (2 kinematic variables). Here, we also experimented with fitting the spike trains and hand velocities jointly, as we found that it allowed iLQR-VAE to reach a similar kinematic decoding performance as when fitting neural activity alone, but

using a smaller network. More generally, we reason that a more natural approach to making behavioural predictions from neural data using a probabilistic generator is to fit it to both jointly, and then use the posterior predictive distribution over behavioural variables (conditioning on spike trains only) as a nonlinear decoder. In future work, this could provide more accurate predictions in those motor tasks where linear regression struggles (see e.g. Schroeder et al., 2021).

For our analyses, we used the first $\sim$ 22 minutes of a single recording session (`indy_20160426`), excluded neurons with overall firing rates below 2 Hz, and binned data at 25 ms resolution. Although it is not a formal requirement of our method, we chunked the data into 336 non-overlapping pseudo-trials of 4 s each, in order to enable parallelization of the ELBO computation during training. We only trained the model on a random subset of 168 trials.

To highlight the flexibility of iLQR as a recognition model, we then evaluated the model by performing inference on the first 9 minutes of the data, as a single long chunk of observations. Note that this is not generally possible in LFADS or any sequential VAE where an encoder RNN has been trained exclusively on trials of the same fixed length. Despite the lack of trial structure, we found that neurons display a stereotyped firing pattern across multiple instances of each reach. This is revealed by binning the angular space into 8 reach directions, temporally segmenting and grouping the inferred firing rates according to the momentary reach direction, and aligning these segments to the time of target onset (Figure 5B). Moreover, hand kinematics could be linearly decoded from the inferred firing rates with high accuracy (Figure 5C; $R^2 = 0.75 \pm 0.01$ over 5 random seeds), on-par with AutoLFADS ($R^2 = 0.76$; Keshtkaran et al., 2021), and considerably higher than GPFA and related approaches (Jensen et al., 2021).

We next wondered if we could use iLQR-VAE to address an open question in motor neuroscience, namely the extent to which the peri-movement dynamics of the motor cortex rely on external inputs (possibly from other brain areas). Such inputs could arise during movement preparation, execution, neither, or both. We thus examined the relationship between the inputs inferred by iLQR-VAE and the concurrent evolution of the neuronal firing rates and hand kinematics. Overall, neuronal activity tends to rise rapidly starting 150 ms before movement onset (Figure 5D, red), consistent with the literature (Churchland et al., 2012; Shenoy et al., 2013). Interestingly, we found that inputs tend to arise much earlier (around the time of target onset), and start decaying well before the mean neural activity has finished rising (Figure 5D top), about 150 ms before the hand started to move (Figure 5D, bottom). While these results must be interpreted cautiously, as inference was performed using information from the whole duration of the trial (i.e. using iLQR as a smoother), they show that the data is best explained by large inputs prior to movement onset, rather than during movement itself. Interestingly, the timing of these inputs is globally consistent with target-induced visual inputs driving preparatory activity in M1, whose dynamics then evolve in a more autonomous manner to drive subsequent motion.

4 Discussion

Limitations and future work

While we have demonstrated that iLQR-VAE performs well on various toy and real datasets, the method has a number of limitations that could be addressed in future work. Firstly, using iLQR as a way of solving maximum a posteriori inference in state space models comes at a high computational cost, and with the risk that iLQR may converge to a local maximum of the log posterior density. Note that both these issues could potentially be tackled at once if process noise in the generator was modelled separately from control inputs, as the MAP estimation problem could then be solved using some of the highly efficient algorithms available in the framework of linearly solvable stochastic control (Dvijotham and Todorov, 2013; Kappen, 2005; Todorov, 2009). Moreover, for simplicity we modelled posterior input uncertainty using a common covariance across all data samples. This might be limiting, for example when modelling neural populations that exhibit coordinated global firing fluctuations giving rise to data samples with highly variable information content. A better solution would be to amortize the computation of the posterior uncertainty, ideally by reusing some of the computations performed in iLQR. We plan to investigate this in future work.

Conclusion

The rise of new tools and software now makes it possible to record from thousands of neurons while monitoring behaviour in great detail (Jun et al., 2017; Mathis et al., 2018; Musk et al., 2019). These datasets create unique opportunities for understanding the brain dynamics that underlie neural and behavioural observations. However, identifying complex dynamical systems is a hard nonlinear filtering and learning problem that calls for new computational techniques (Kutschireiter et al., 2020). Here, we exploited the duality between control and inference (Kappen and Ruiz, 2016; Levine, 2018; Toussaint, 2009) to bring efficient algorithms for nonlinear control to bear on learning and inference in nonlinear state space models. The method we proposed uses iLQR, a powerful general purpose nonlinear controller, to perform amortized inference over inputs in an RNN-based generative model. iLQR brings important flexibility at test time, enabling predictions on arbitrary, heterogeneous sequences of observations. Finally, we expect the method to lend itself very naturally to applications to control problems, for example in the area of brain-computer interfaces (Makin et al., 2018; Willett et al., 2021).

Acknowledgements

We thank Jasmine Stone and Javier Antorán for helpful comments on the manuscript, and Anil Madhavapeddy for providing computing resources at early stages. M.S. was funded by an EPSRC DTP studentship and K.T.J. was funded by a Gates Cambridge scholarship. This work was performed using resources provided by the Cambridge Tier-2 system operated by the University of Cambridge Research Computing Service (http://www.hpc.cam.ac.uk) funded by EPSRC Tier-2 capital grant EP/P020259/1.

References

Amos, B., Rodriguez, I. D. J., Sacks, J., Boots, B., and Kolter, J. Z. (2018). Differentiable MPC for end-to-end planning and control. *arXiv preprint arXiv:1810.13400*.

Archer, E., Park, I. M., Buesing, L., Cunningham, J., and Paninski, L. (2015). Black box variational inference for state space models. *arXiv preprint arXiv:1511.07367*.

Blondel, M., Berthet, Q., Cuturi, M., Frostig, R., Hoyer, S., Llinares-López, F., Pedregosa, F., and Vert, J.-P. (2021). Efficient and modular implicit differentiation. *arXiv preprint arXiv:2105.15183*.

Buesing, L., Macke, J. H., and Sahani, M. (2012). Learning stable, regularised latent models of neural population dynamics. *Network: Computation in Neural Systems*, 23:24–47.

Champion, K., Lusch, B., Kutz, J. N., and Brunton, S. L. (2019). Data-driven discovery of coordinates and governing equations. *Proceedings of the National Academy of Sciences*, 116(45):22445–22451.

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*, 487:51–56.

Churchland, M. M., Cunningham, J. P., Kaufman, M. T., Ryu, S. I., and Shenoy, K. V. (2010). Cortical preparatory activity: representation of movement or first cog in a dynamical machine? *Neuron*, 68(3):387–400.

Cremer, C., Li, X., and Duvenaud, D. (2018). Inference suboptimality in variational autoencoders. In *International Conference on Machine Learning*, pages 1078–1086. PMLR.

Duncker, L., Bohner, G., Boussard, J., and Sahani, M. (2019). Learning interpretable continuous-time models of latent stochastic dynamical systems. In *International Conference on Machine Learning*, pages 1726–1734.

Dvijotham, K. and Todorov, E. (2013). Linearly solvable optimal control. In *Reinforcement learning and approximate dynamic programming for feedback control*, volume 17, pages 119–141. Wiley Online Library.

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

Ghahramani, Z. and Hinton, G. E. (2000). Variational learning for switching state-space models. *Neural comput.*, 12:831–864.

Hernandez, D., Moretti, A. K., Wei, Z., Saxena, S., Cunningham, J., and Paninski, L. (2018). Nonlinear evolution via spatially-dependent linear dynamics for electrophysiology and calcium data. *arXiv preprint arXiv:1811.02459*.

Jensen, K. T., Kao, T.-C., Stone, J. T., and Hennequin, G. (2021). Scalable bayesian gpfa with automatic relevance determination and discrete noise models. *bioRxiv*.

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., et al. (2017). Fully integrated silicon probes for high-density recording of neural activity. *Nature*, 551(7679):232–236.

Kappen, H. J. (2005). Linear theory for control of nonlinear stochastic systems. *Phys. Rev. Lett.*, 95:200201.

Kappen, H. J. and Ruiz, H. C. (2016). Adaptive importance sampling for control and inference. *Journal of Statistical Physics*, 162:1244–1266.

Kaufman, M. T., Seely, J. S., Sussillo, D., Ryu, S. I., Shenoy, K. V., and Churchland, M. M. (2016). The largest response component in the motor cortex reflects movement timing but not movement type. *Eneuro*, 3(4).

Keshtkaran, M. R., Sedler, A. R., Chowdhury, R. H., Tandon, R., Basrai, D., Nguyen, S. L., Sohn, H., Jazayeri, M., Miller, L. E., and Pandarinath, C. (2021). A large-scale neural network training framework for generalized estimation of single-trial population dynamics. *bioRxiv*.

Kim, T. D., Luo, T. Z., Pillow, J. W., and Brody, C. (2021). Inferring latent dynamics underlying neural population activity via neural differential equations. In Meila, M. and Zhang, T., editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 5551–5561. PMLR.

Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Kingma, D. P., Salimans, T., and Welling, M. (2015). Variational dropout and the local reparameterization trick. *arXiv preprint arXiv:1506.02557*.

Kingma, D. P. and Welling, M. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.

Kutschireiter, A., Surace, S. C., and Pfister, J.-P. (2020). The hitchhiker's guide to nonlinear filtering. *Journal of Mathematical Psychology*, 94:102307.

Levine, S. (2018). Reinforcement learning and control as probabilistic inference: Tutorial and review. *arXiv preprint arXiv:1805.00909*.

Li, W. and Todorov, E. (2004). Iterative linear quadratic regulator design for nonlinear biological movement systems. In *ICINCO (1)*, pages 222–229. Citeseer.

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.

Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of atmospheric sciences*, 20(2):130–141.

Makin, J. G., O'Doherty, J. E., Cardoso, M. M., and Sabes, P. N. (2018). Superior arm-movement decoding from cortex with a new, unsupervised-learning algorithm. *Journal of neural engineering*, 15(2):026010.

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., and Bethge, M. (2018). Deeplabcut: markerless pose estimation of user-defined body parts with deep learning. *Nature neuroscience*, 21(9):1281–1289.

Musk, E. et al. (2019). An integrated brain-machine interface platform with thousands of channels. *Journal of medical Internet research*, 21:e16194.

Nguyen, D., Ouala, S., Drumetz, L., and Fablet, R. (2020). Variational deep learning for the identification and reconstruction of chaotic and stochastic dynamical systems from noisy and partial observations. *arXiv preprint arXiv:2009.02296*.

O'Doherty, J. E., Cardoso, M. M. B., Makin, J. G., and Sabes, P. N. (2018). Nonhuman Primate Reaching with Multichannel Sensorimotor Cortex Electrophysiology: broadband for indy_20160630_01. This research was supported by the Congressionally Directed Medical Research Program (W81XWH-14-1-0510). JEO was supported by fellowship #2978 from the Paralyzed Veterans of America. JGM was supported by a fellowship from the Swartz Foundation.

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., et al. (2018). Inferring single-trial neural population dynamics using sequential auto-encoders. *Nature methods*, 15(10):805–815.

Pei, F., Ye, J., Zoltowski, D., Wu, A., Chowdhury, R. H., Sohn, H., O'Doherty, J. E., Shenoy, K. V., Kaufman, M. T., Churchland, M., et al. (2021). Neural latents benchmark'21: Evaluating latent variable models of neural population activity. *arXiv preprint arXiv:2109.04463*.

Rezende, D. J., Mohamed, S., and Wierstra, D. (2014). Stochastic backpropagation and approximate inference in deep generative models. In *International conference on machine learning*, pages 1278–1286. PMLR.

Schroeder, K. E., Perkins, S. M., Wang, Q., and Churchland, M. M. (2021). Cortical control of virtual self-motion using task-specific subspaces. *bioRxiv*, pages 2019–12.

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

Shenoy, K. V., Sahani, M., and Churchland, M. M. (2013). Cortical control of arm movements: a dynamical systems perspective. *Annu. Rev. Neurosci.*, 36:337–359.

Stosiek, C., Garaschuk, O., Holthoff, K., and Konnerth, A. (2003). In vivo two-photon calcium imaging of neuronal networks. *Proceedings of the National Academy of Sciences*, 100(12):7319–7324.

Todorov, E. (2009). Efficient computation of optimal actions. *Proc. Natl. Acad. Sci.*, 106:11478–11483.

Toussaint, M. (2009). Robot trajectory optimization using approximate inference. In *Proceedings of the 26th annual international conference on machine learning*, pages 1049–1056.

Willett, F. R., Avansino, D. T., Hochberg, L. R., Henderson, J. M., and Shenoy, K. V. (2021). High-performance brain-to-text communication via handwriting. *Nature*, 593(7858):249–254.

Ye, J. and Pandarinath, C. (2021). Representation learning for neural population activity with neural data transformers. *bioRxiv*.

Yu, Byron, M., 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.