LEARNING OSCILLATORY BRAIN DYNAMICS USING NONLINEAR KALMAN SMOOTHING

Germán Abrevaya
University of Buenos Aires
germanabrevaya@gmail.com

Aleksandr Aravkin
University of Washington saravkin@uw.edu

Guillermo A. Cecchi, Pablo Polosecki, & Irina Rish
IBM Thomas J. Watson Research Center
{gcecchi, pipolose, rish}@us.ibm.com

ABSTRACT

Motivated by the challenging problem of modeling nonlinear dynamics of brain activations in calcium imaging, we propose a novel approach for learning a nonlinear differential equation model: a variable-projection optimization approach for estimating the parameters of the multivariate (coupled) van der Pol oscillator. To the best of our knowledge, we are the first to propose a successful approach to learning such nonlinear dynamical model, and to demonstrate that it can accurately capture nonlinear dynamics of the brain data. Furthermore, in order to further improve the predictive accuracy when forecasting future brain activity, we use the learned analytical van der Pol model to generate large amounts of simulated data for LSTM pretraining, since brain imaging datasets are often limited in size, and since the generic LSTM model may benefit from the oscillator prior imposed via such pretraining. Indeed, the proposed combination of the analytical approach with the general-purpose statistical LSTM model improves the performance of both methods.

1 INTRODUCTION

Neuroscience is a source of exceedingly interesting and challenging data for analysis and modeling, generated by the underlying multivariate coupled nonlinear dynamical system, a.k.a. the brain. Brain activity clearly exhibits a highly nonlinear, oscillatory and periodic (or nearly periodic) behavior, with sharp phase transitions between different states; thus, modeling such time series can be nontrivial. Also, state-of-art temporal models such as LSTM networks may require large amounts of training data, not always readily available in biological applications, in order to capture complex nonlinear dynamical processes; moreover, such models generally lack interpretability.

In this work, we consider brain-wide calcium imaging larval zebrafish data Ahrens & et al (2013) and propose a novel approach for estimating the parameters of van der Pol oscillator – an analytical model governed by nonlinear ordinary differential equations, which describe calcium dynamics in the brain using both observed voltage-like variables, reflecting voxel activity, and hidden recovery-like variables, reflecting excitability. To the best of our knowledge, we are the first ones to propose an approach to this problem, and demonstrate its successful application on real data. Note that fitting parameters of *linear* state space systems is a well-studied problem. However, inferring parameters from nonlinear ODEs requires more care, and we develop a specialized algorithm that is different from the classic EM approach typically used for linear systems. Namely, we propose a variable-projection optimization approach to estimate the parameters of the multivariate (coupled) van der Pol oscillator, and demonstrate that the proposed model can accurately capture the nonlinear dynamics of the brain data, and provide interpretable results as it learns a coupling matrix indicating interactions across brain areas.

Furthermore, we improve the accuracy of predicting future brain activity using a combination of both analytical van der Pol model and general-purpose statistical model such as LSTM, which outperforms both techniques considered separately. Namely, the van der Pol model learned on relatively small

training dataset is used as a simulator of unlimited pre-training data which effectively impose an oscillatory "prior" on LSTM, i.e. a kind of regularizer which directs LSTM model towards specific oscillatory patterns we assume are present in the data.

2 OUR APPROACH

Whole-Brain Calcium Imaging Data. In Ahrens & et al (2013), light-sheet microscopy was used to record the neural activity of a whole brain of the larval zebrafish, reported by a genetically-encoded calcium marker, in vivo and at 0.8 Hz sampling rate. From the publicly available data Channel (2013) it is possible to obtain a movie of 500 frames with a 2D collapsed view of 80% of the approximately 40,000–100,000 neurons in the brain, with a resolution of 400 by 250 pixels (approximately 600 by 270 microns). We performed SVD to reduce the dimensionality of the data and we used the top 6 components explaining most of the variance (spatial components corresponding to brain areas; we learn a coupled oscillator on time components).

Van der Pol Model. Because neuronal calcium dynamics are largely driven by transmembrane voltage and voltage-dependent calcium channels, we propose to describe the calcium dynamics of a neuron, or small clusters of them, as a 2D differential equation with a voltage-like variable (activity), and a recovery-like variable (excitability), following similar approaches in the literature M. (2007). In this paper, we propose the following nonlinear oscillator model for each scalar component:

$$\begin{aligned} \dot{x}_{1i}(t) &= \alpha_1 x_{1i}(t)(1 - {x_{1i}}^2(t)) + \alpha_2 x_{2i}(t) + \sum_{j=1}^{m} W_{ij} x_{1j}(t) \\ \dot{x}_{2i}(t) &= -\alpha_3 x_{1i}(t) \end{aligned} \tag{1}$$

where $x_{1i}(t)$ and $x_{2i}(t)$ represent the (observed) activity- and the (hidden) recovery-like variables of the i-th neural unit, respectively, and where the W matrix represents the coupling strength between the observed variables, or neural units. The parameters α_k determine the bifurcation diagram of the system, allowing for a rich availability of possible dynamical states including oscillations and spike-like responses S. (2003); M. (2007).

Our ultimate goal is to learn the dynamical model (1), i.e. to estimate the parameters $\alpha_i, i = 1, ..., 5$ and W_{ij}, from calcium imaging data. We discretize the ODE model in equation (1), and formulate a joint inference problem for the state space x and parameters α, W that is informed by noisy direct observations of some components; and constrained by the discretized dynamics.

Parameter estimation. We consider learning one component firs, and then extend the approach to multivariate model. Let $x^k \in \mathbb{R}^2$ denote a component of the van der Pol model given earlier in the equation (1), where $x^k = (x_1^k, x_2^k)$, i.e. a combination of both the observed and the hidden variables, and k is the time index. The discretized dynamics governing the evolution of the single component $x = (x_1, x_2)$ can be written as $x^{k+1} = g(x^k, \alpha)$, where g is a first-order Euler discretization of the nonlinear ODE (1). The same α inform the evolution of the entire time series $x = \begin{bmatrix} x_1^T & \cdots & x_N^T \end{bmatrix}^T$, which we write as $G(x, \alpha) = \eta^0$, with

$$G(x, \alpha) = \begin{bmatrix} x^1 \\ x^2 - g(x^1, \alpha) \\ \vdots \\ x^N - g(x^{N-1}, \alpha) \end{bmatrix}, \quad \eta^0 = \begin{bmatrix} x^0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{2}$$

Given noisy observations $z^k = H_k x^k + \omega_k$, we obtain consider ODE-constrained optimization problem

$$\min_{x,\alpha} \quad \frac{1}{2}\|z - Hx\|^2 \quad \text{s.t.} \quad G(x, \alpha) = \eta^0, \tag{3}$$

To solve this problem, we use a recent technique from PDE-constrained optimization, called the *penalty method* van Leeuwen & Herrmann (2015). The technique is a new take on the quadratic penalty method. Rewriting (3) with a quadratic penalty, we obtain the relaxed problem

$$\min_{x,\alpha} f_\lambda(x, \alpha) := \frac{1}{2}\|z - Hx\|^2 + \frac{\lambda}{2}\|G(x, \alpha) - \eta^0\|^2. \tag{4}$$

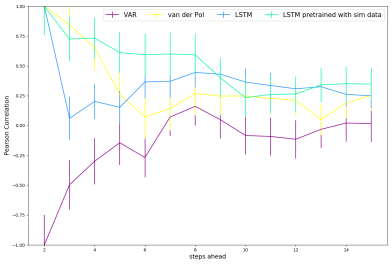

Figure 1: Predictive performance: correlation between the actual and predicted time series.

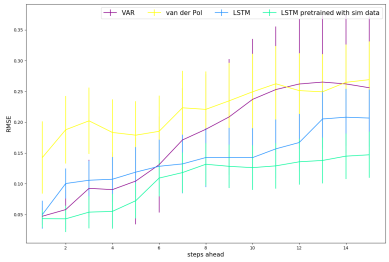

Figure 2: Predictive performance: root mean square error (RMSE) between the actual and predicted time series.

The key idea is to then use partial minimization with respect to x at each iteration of α (see appendix for more detail).

Extending the model to m components, and incorporating W parameters besdies α, we obtain the coupled model shown previously in (1), (more details on the optimization are provided in the Appendix).

3 Experiments

First, we evaluated multiple runs of van der Pol estimation procedure described above, and obtained high correlations between the actual data and the model's prediction, ranging from 0.76 to 0.83 for all six components, demonstrating that the model was able to capture well the underlying dynamics in the data.

We next compared our van der Pol model, LSTM and hybrid O-LSTM approach (which consists in simulating multiple times series from the van der Pol model estimated above and pretraining LSTM before training ot on real data), as well as baseline Vector Auto-Regressive (VAR) model. Figures 1 and 2 show the median correlation and the root-means square error, respectively, for several predictive methods: vector autoregressive (VAR) model (magenta), van der Pol model (yellow), LSTM (blue), and O-LSTM, i.e. LSTM pretrained on the data simulated using the above van der Pol model (green). Here we estimated parameters of the models on 100 consequtive points of training data, and then predicted the next 30 points (x-axis plots the index of the time points being predicted).

Note that the linear VAR model (magenta) performs poorly, unable to capture the nonlinear dynamics; both van der Pol (yellow) and LSTM (blue) have somehwat comparable perfromance, outperforming each other in shorter or in longer run; however the combination of both, O-LSTM model appears to be superior to both, achieving best accuracy for majority of future time points.

References

MB Ahrens and et al. Whole-brain functional imaging at cellular resolution using light-sheet microscopy. *Nature Methods*, 10:413–420, 2013.

Discover Magazine Youtube Channel. Calcium imaging data from ahrens et al, (2013), March 2013. URL `https://www.youtube.com/watch?v=lppAwkek6DI`.

Izhikevich E. M. *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*. MIT Press,, 2007.

Wiggins S. *Introduction to Applied Nonlinear Dynamical Systems and Chaos*. Springer,, 2003.

Tristan van Leeuwen and Felix J Herrmann. A penalty method for pde-constrained optimization in inverse problems. *Inverse Problems*, 32(1):015007, 2015.