First LDS analysis of Emmett' data

Joaquín Rapela*

December 18, 2025

1 Introduction

Here we use a Linear Dynamical System model (LDS)[1] to explore the low dimensional structure of Emmett's switching-task recordings (Thompson et al., 2024).

2 Methods

2.1 Data

We analyzed two minutes of continuous electrophysiological recordings starting at (electrophysiological) time 5512 from `EJT178_implant1/recording1_15-032022`, and neurons indices 124 to 223 in the striatum. Figure 1 plots the spike times used in this analysis.

[Figure 1 about here.]

2.2 LDS model

The LDS model estimates K latent variables, $x_0[n], \ldots, x_{K-1}[n]$, $n \in \{0, \ldots, N-1\}$. We concatenated the latent variables into a latent vector $\mathbf{x}[n]$:

*j.rapela@ucl.ac.uk

[1] `https://github.com/joacorapela/ssm`

$$\mathbf{x}[n] = \begin{bmatrix} x_0[n] \\ \vdots \\ x_{K-1}[n] \end{bmatrix} \tag{1}$$

The latents evolve in time following a first-order Markov process with additive Gaussian noise:

$$\mathbf{x}[n] = A\mathbf{x}[n-1] + \mathbf{v}[n] \quad \text{with } \mathbf{v}[n] \sim \mathcal{N}(0, Q)$$

starting from $\mathbf{x}_0 \sim \mathcal{N}(\mathbf{x}_0|\mathbf{m}_0, V_0)$.

The latent variables are combined linearly to approximate the firing rate of the recorded neurons (bin size 20 ms), $\mathbf{y}[n]$, with a constant offset, $\mathbf{a}$, and an additive noise, $\mathbf{w}[n]$.

$$\mathbf{y}[n] = \mathbf{a} + C\mathbf{x}[n] + \mathbf{w}[n] \quad \text{with } \mathbf{w}[n] \sim \mathcal{N}(0, R)$$

We estimated an LDS models with $K = 5$ latent variables (Eq. 1; the selection of the number of latent variables was arbitrary, and we should perform model selection in later iterations).

We estimate the model parameter $\{\mathbf{m}_0, V_0, \mathbf{u}, A, Q, \mathbf{a}, Z, R\}$ using the expectation maximisation algorithm (doEstimateEM.py). We then used the Kalman Filter algorithm, with these parameters, to infer the latents variables, $\mathbf{x}[n]$ in Equation 1 (doKalmanFilteringWithEstimatedParams.py). We also forecasted firing rates of individual neurons (doPlotObservationsForecasts.py).

3 Results

3.1 Simultaneous spikes across the striatum

The inferred latent variables are shown in Figure 2. Latent 0, that captures most of variance in the spike rates, shows large negative deflections every 1 to 3 seconds. Figures 3 and 4 show zoomed-in views of the latents, and Figures 5 and 6 show corresponding zoomed-in views of the spikes times. We note that **at times of large negative deflections of latent 0 there is**

a large number of simultaneous spikes across many neurons in the striatum.

[Figure 2 about here.]

[Figure 3 about here.]

[Figure 4 about here.]

[Figure 5 about here.]

[Figure 6 about here.]

3.2 Forecasting spikes rates

Figures 7-10 show model forecastings of spike rates with horizon ranging from 20 ms to 1.0 secs into the future. In these figures the forecasted time series is plotted at forecasting times. This means that if the forecasting horizon was h seconds, the forecasted trace at time t represents the forecasted spike rate at time t, that was generated using data up to time $t - h$.

[Figure 7 about here.]

[Figure 8 about here.]

[Figure 9 about here.]

[Figure 10 about here.]

The forecasting model is learning two interesting global trends in the recordings. First, when repeated firing is detected, the model learns that the firing rate should first increase to a peak value, and then decrease to baseline. Second, the model learns that increases in firing rate are regular, and sometimes forecasts increases in firing rate, even without previous spiking. This features are more prominent for horizons $h = 0.02, 0.1$ sec. Larger horizons of $h = 0.5 and 1.0$ sec show delays. However, we will perform a proper evaluation of forecasting in the near future.

4 Future work

1. Analyse other time periods from `EJT178_implant1/recording1_1503-2022`.
2. Analyze other animals.
3. Improve the evaluation of forecasters.
4. Perform real-time latent visualisation with unsorted spikes.

References

Thompson, E., Rollik, L., Waked, B., Mills, G., Kaur, J., Geva, B., Carrasco-Davis, R., George, T., Domine, C., Dorrell, W., et al. (2024). Replay of procedural experience is independent of the hippocampus. *bioRxiv*, pages 2024–06.

List of Figures

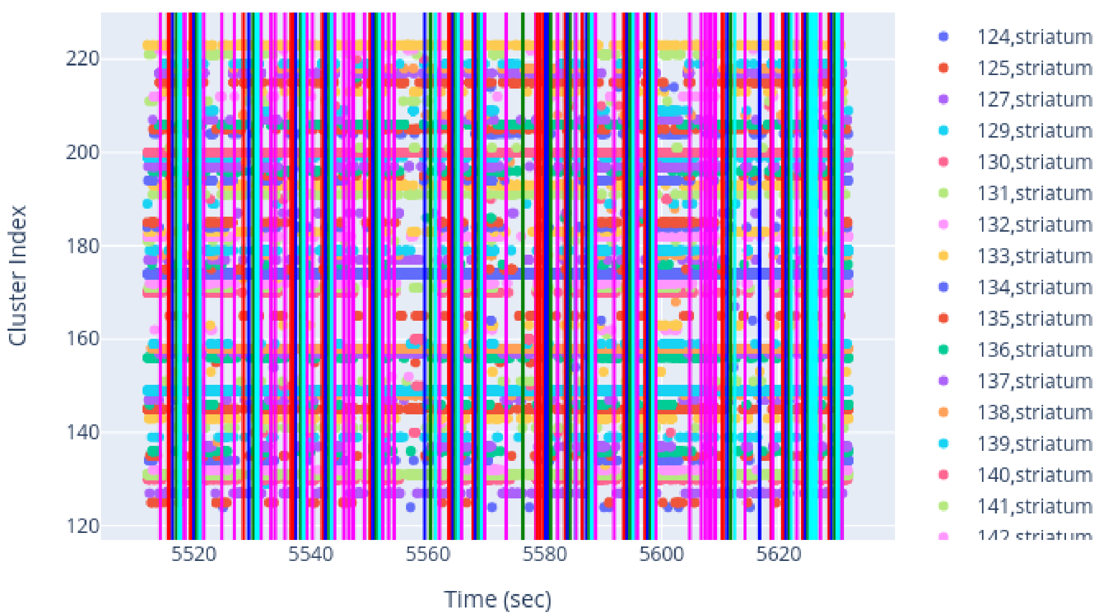

Figure 1: Spikes times used in this analysis: we analyzed two minutes of continuous electrophysiological recordings starting at (electrophysiological) time 5,512 from `EJT178_implant1/recording1_15032022`, and neurons indices 124 to 223 in the striatum. Click on the image to get its interactive version.

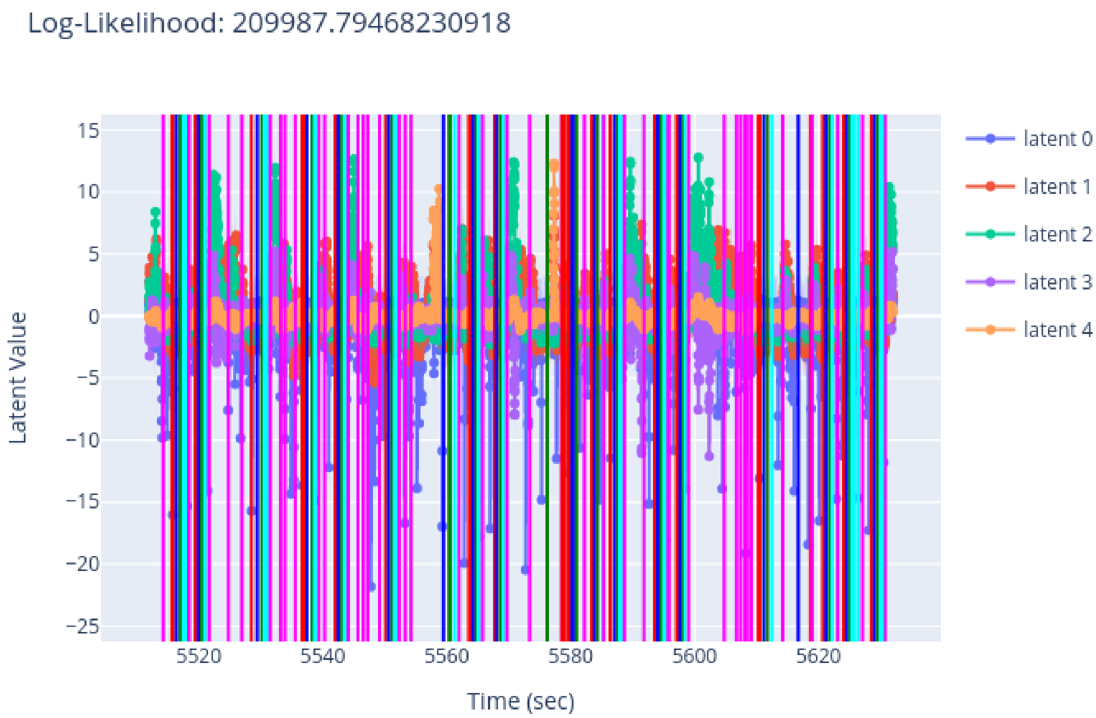

Figure 2: Orthonormalised latent variables ($\mathbf{x}(t)$ in Eq. 1) inferred by the Kalman filter algorithm. The colored vertical lines indicate 'poke_in' times, with blue, red, cyan, yellow, purple, green and magenta corresponding to ports 1 to 7, respectively. The correct sequence is $2 \rightarrow 1 \rightarrow 6 \rightarrow 3 \rightarrow 7$ corresponding to red $\rightarrow$ blue $\rightarrow$ green $\rightarrow$ cyan $\rightarrow$ magenta. Note the sharp downward peaks of latent 0 (the latent variable capturing most of the variance in the response). Zoomed versions of this image appear in Figures 3 and 4. Click on the image to get its interactive version.

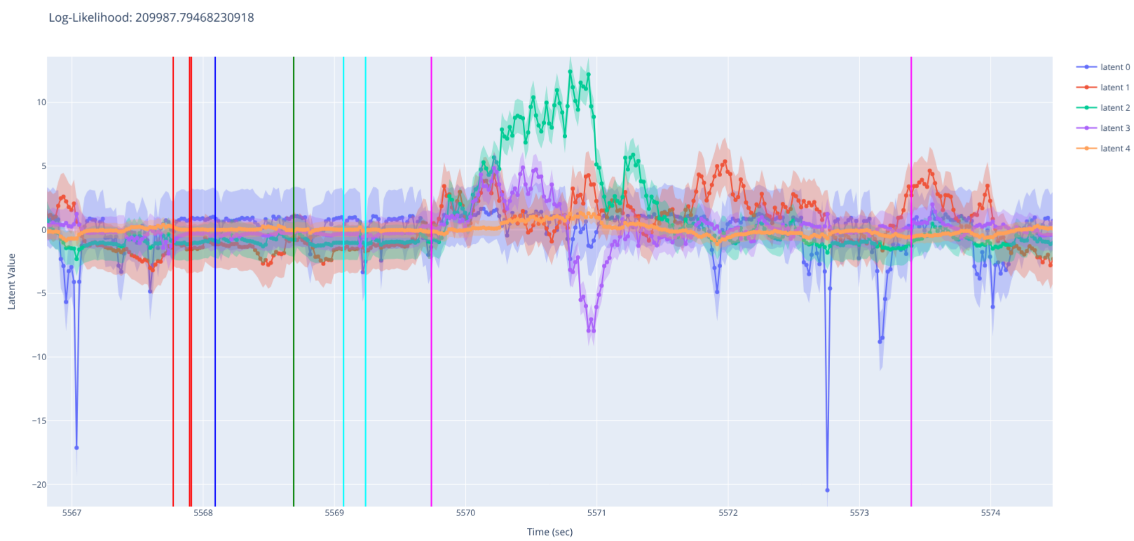

Figure 3: Zoomed-in view one of the latents in Figure 2
.

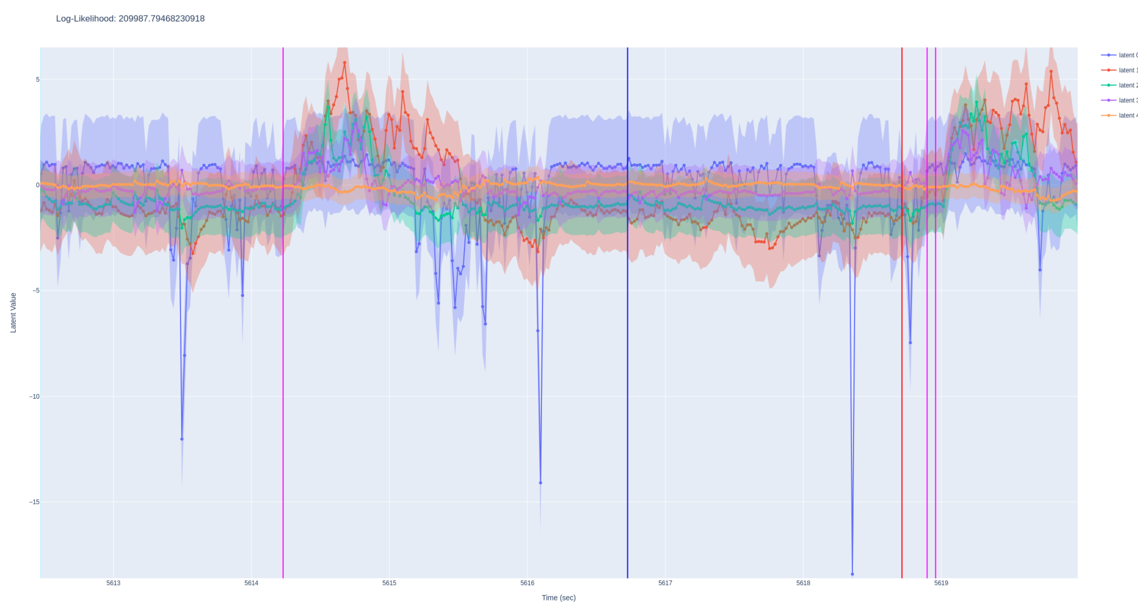

Figure 4: Zoomed-in view two of the latents in Figure 2
.

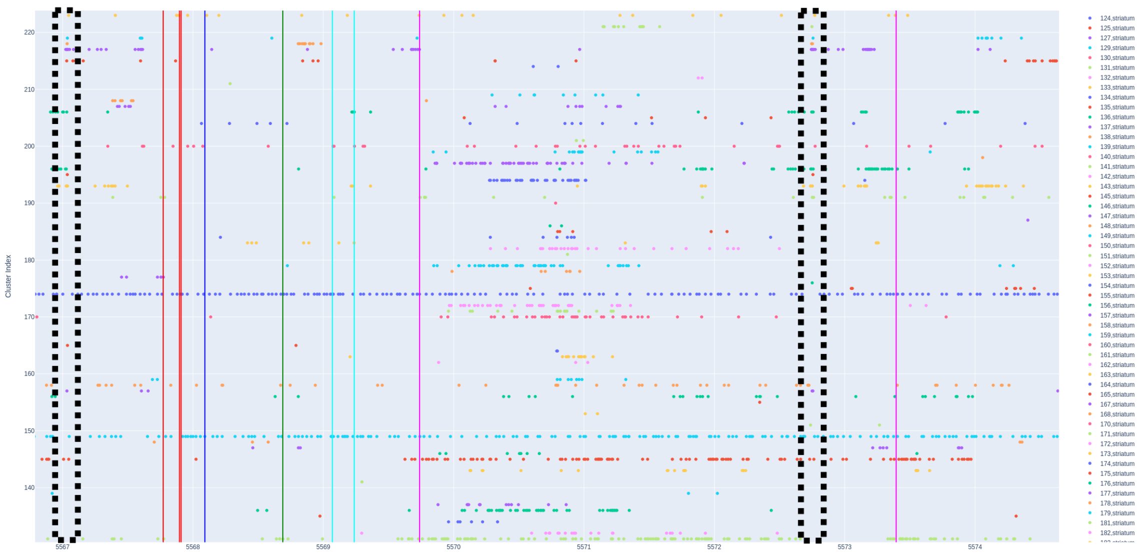

Figure 5: Zoomed-in view of the spikes-times in Figure 1 with bounding boxes around the times of large negative deflections of latent 0 in Figure 3.

Figure 6: Zoomed-in view of the spikes-times in Figure 1 with bounding boxes around the times of large negative deflections of latent 0 in Figure 4.

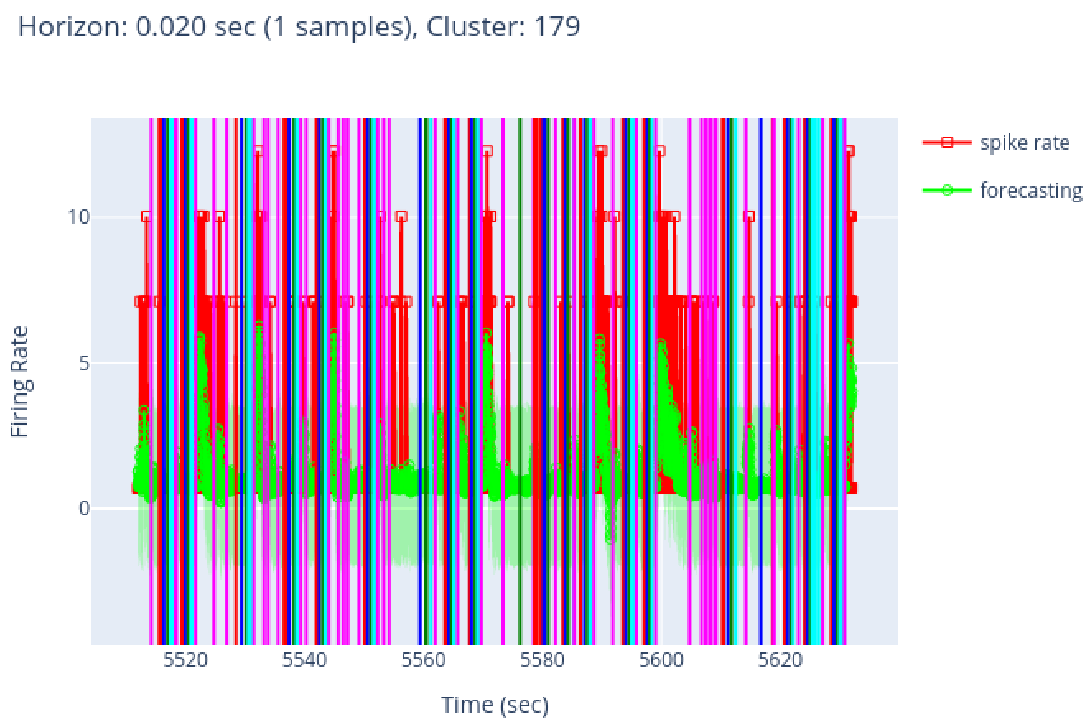

Figure 7: Forecasted spikes rates of example cluster 179 with an horizon $h = 0.02$ seconds. The forecasted time series is plotted at forecasting times. This means that if the forecasting horizon was h seconds, the forecasted trace at time t represents the forecasted spike rate at time t, that was generated using past data up to time $t - h$. Click on the image to get its interactive version.

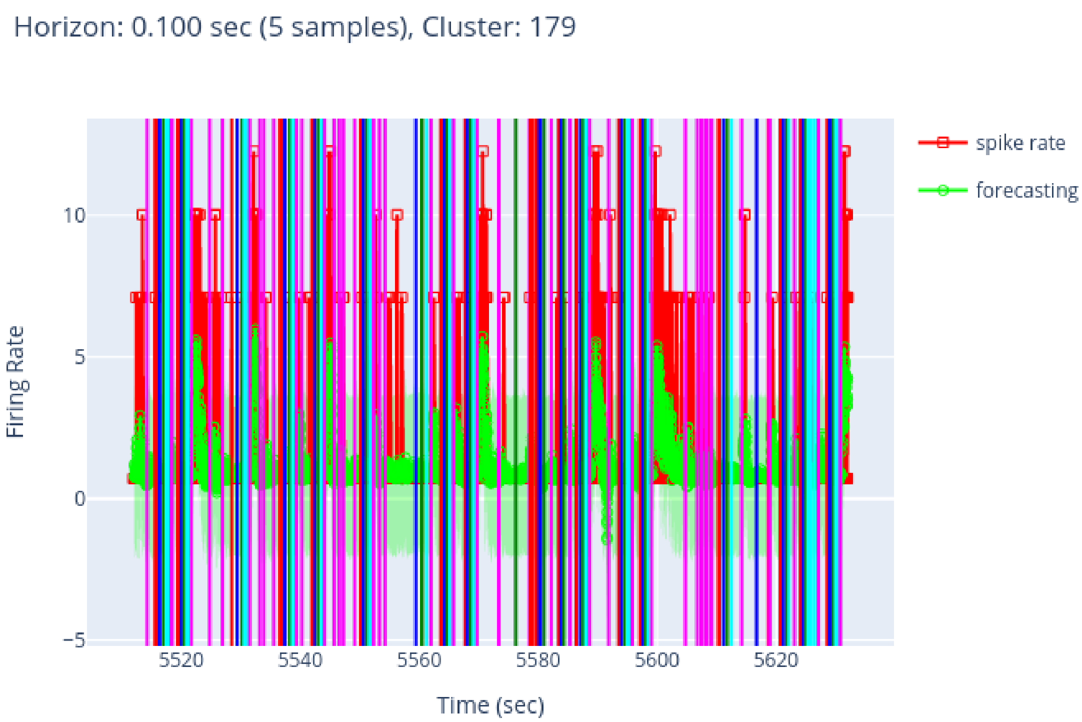

Figure 8: Forecasted spikes rates of example cluster 179 with an horizon $h = 0.10$ seconds. Same format as in Figure 7. Click on the image to get its interactive version.

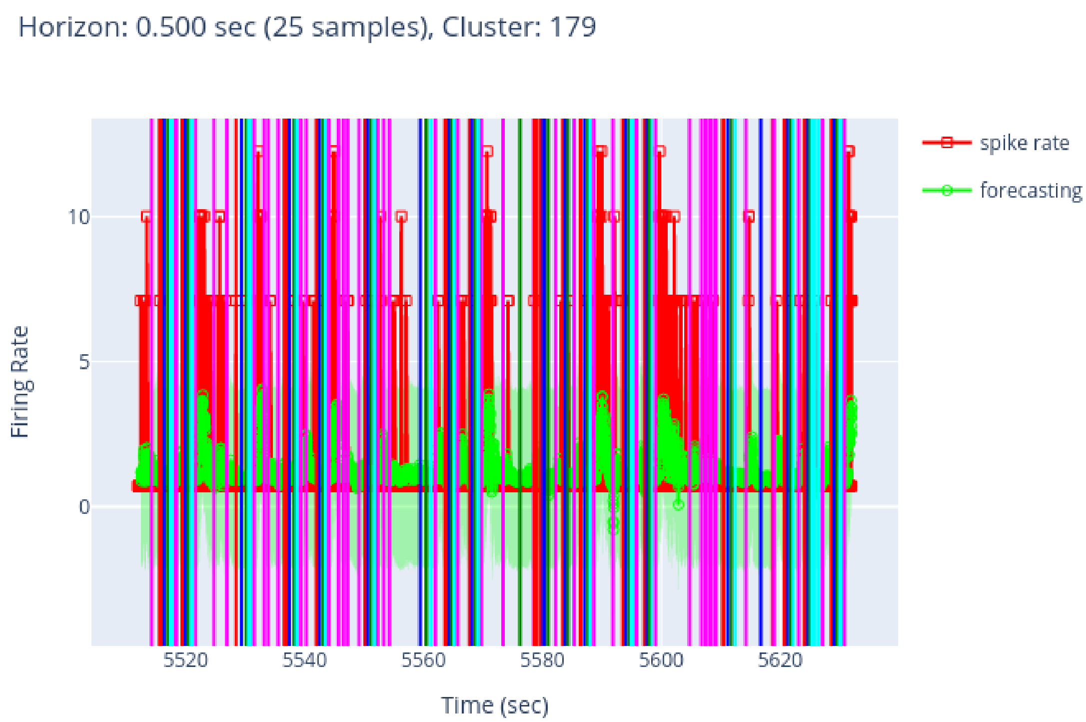

Figure 9: Forecasted spikes rates of example cluster 179 with an horizon $h = 0.50$ seconds. Same format as in Figure 7. Click on the image to get its interactive version.

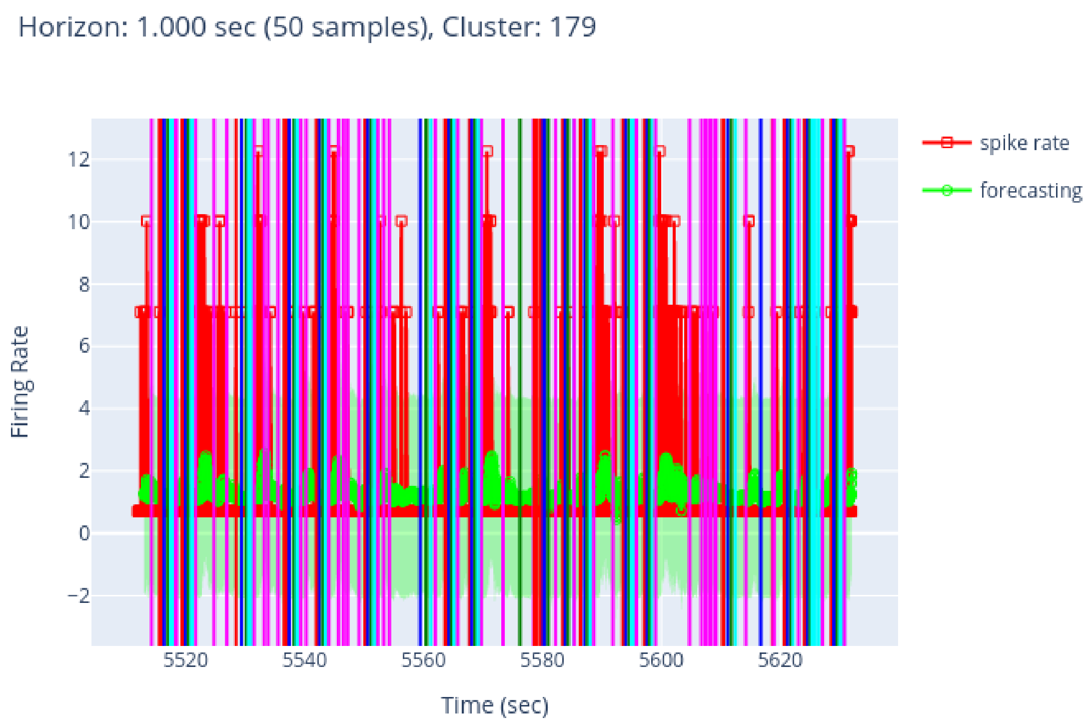

Figure 10: Forecasted spikes rates of example cluster 179 with an horizon $h = 1.00$ seconds. Same format as in Figure 7. Click on the image to get its interactive version.