Exact Continuous Spiking Rate Inference

Merav Stern

October 18, 2024

Abstract

Many cognitive functions involve multiple brain areas simultaneously processing, distributing, and sharing information. The wide-field imaging technique can record such brain-wide activity since it offers a unique combination of simultaneous recordings from a wide field of view at a high rate. This unique combination is achieved by using a single-photon camera, which is set to capture, from a large area, neural activity-driven light emission generated by calcium indicators. Adequately analyzing this captured data requires inferring the underlying neural activity from the recorded fluorescence, a challenging mathematical problem. The challenge arises from the calcium indicator dynamics, which distort the neural dynamics as it transforms the neural activity into light emission and from the presence of noise. The wide-field setting adds a distinctive challenge. Its wide field of view recordings at a high rate constrains the spatial resolution to be limited. As a result, each fluorescence trace captured by each camera's pixel originates from many neuron activities and not a single neuron as typical of other recording techniques. No previous, rigorously studied analytic solution exists for inferring neural activity from recorded fluorescence in the wide-field setting. Here, we phrase the inference problem arising from wide-field recordings as an optimization problem and solve it exactly. To ensure the robustness of our solution and provide a solid foundation for its application, we rigorously verify it using real data. We further suggest a novel approach for the optimization problem parameter-tuning. Beyond recovering the neural dynamics, our inference will allow future studies to retrieve accurate, correlation-based analyses of brain-wide activity.

1 Introduction

Optogenetic tools use calcium indicators to bind spikes initiated by neurons with fluorescence emission. The produced fluorescence can then be captured by a camera, making neural activity visible and, hence, more easily detectable.

In the wide-field setting, a vast amount of fluorescence emitted due to neural activity is captured by a standard (one photon) camera across a vast field of view (Cardin et al. 2020). A typical example of such a wide-field view is a mouse brain's whole dorsal surface area (Silasi et al. 2016). In addition to the large field of view, the wide-field setting also permits fast, simultaneous recordings of the whole surface (Gilad 2024). These enable wide-field recordings to reliably portray the fluorescence signal's time evolution across all recorded areas. However, due to the large field of view, the resolution is limited; many thousands of neurons' fluorescence contribute to each single light trace captured by each pixel in the camera, resulting in many thousands of neurons contributing their activity-generated fluorescence to a single recorded fluorescence trace. In turn,

the vast amount of fluorescence captured in each pixel allows the recordings to take place via the skull (of a mouse for example, (Silasi et al. 2016)), leaving the brain intact and making wide-field minimally invasive. With its minimal invasion, wide-field can be easily utilized for recording awake-behaving animals across multiple sessions (days) with the ability to relate and compare all sessions (Musall et al. 2018, MacDowell & Buschman 2020). The long list of advantages of the wide-field method presented above hence makes the wide-field a popular recording technique despite its limited spatial resolution.

Inferring the total neural activity contributing to each recorded fluorescence trace is a challenging mathematical problem. The limited spatial resolution, a distinctive property of the wide-field method, and the presence of noise, require a distinctive solution that can extract such multiple-neural activity. Many solutions exist for extracting spiking activity from single-neuron-generated fluorescence recordings (Jewell et al. 2019, Friedrich et al. 2017, Pachitariu et al. 2018, Pnevmatikakis et al. 2016, Berens et al. 2018, for example). However, as we show in this work, the mathematical challenge that stems from the need to infer the total spiking activity of many neurons is unique and cannot be borrowed from the single-neuron scenario, hence requiring a new tailored solution.

The relation between the neural activity and the dynamics of the calcium indicator, which is responsible for the fluorescence, is known in calcium imaging (Zhang et al. 2023, Chen et al. 2013, for example). Nevertheless, since, in essence, the calcium integrates the spiking activity of a neuron, recovering the neural spiking activity from the fluorescence is fundamentally a derivative operation. The presence of noise makes utilizing such naive inverse, derivative-based operation impractical since it amplifies the noise to a level that makes inferring the underlying neural activity not plausible. Hence, naive approaches to recover neural activity from fluorescence recordings by inverting the calcium dynamics fail, requiring mathematical solutions.

While the level of fluorescence is a direct indication of the calcium concentration, it conveys only to first-order the volume of recent neural activity. But since neural activity holds much information about brain functions, it is almost always that information we are after. Hence, inferring its accurate amount (beyond first-order) is important. However, even after inference, in the wide-field setting, the information about neural activity remains limited since this setting cannot portray the absolute count of spikes per second (nor the absolute amount of neurons that emitted them). This information simply does not exist since, prior to inference, the fluorescence itself is given in arbitrary units and is meaningless in terms of the absolute volume of brain activity. Only information about relative changes is in the data. This includes information about relative changes in neural activity in time. This information is preserved and becomes available upon correct inference by utilizing the calcium indicator dynamics. Hence, with inference, information is gained. The use of the calcium indicator dynamics is added to recover the neural activity. As a result, the neural dynamics through time are accurate. In addition, noise is removed during inference according to biologically relevant criteria based on the calcium-binding properties (rather than using imaging processing criteria, which are unrelated to the biological process of fluorescence emission and have a greater chance of being part of the signal).

Finding the correct time evolution of the neural dynamics allows many advanced analyses to retrieve accurate findings. Such analyses include comparing two brain areas by testing how their neural activity correlates among them, or with behavior, and using more advanced correlation-based analyses such as PCA and NMF. Since they are based on correlations,their analyses are invariant under different absolute values, while being sensitive to the timing of relative changes between parameters (whether these are neural spiking rates or behavior measures). This further strengthens the need to find the accurate dynamics of the neural activity though time. In addition,

it renders the wide-field inability to find the absolute magnitude of neural activity a less significant drawback.

The main contribution of this paper is that it presents pioneering work that considers and phrases the inference problem arising from wide-field recordings. It gives an exact, complete solution to the phrased inference problem. To the best of our knowledge, less than a handful of previous attempts were made to test inference of multiple neurons' activity from their generated fluorescence (Wekselblatt et al. 2016, O'Rawe et al. 2023, Stern et al. 2020). Out of them, only one designed an adapted method to retrieve (and not estimate), in a timely manner, a continuous inferred spiking rate signal (Wekselblatt et al. 2016). We compare our new solution's performance to it. In addition, we verify that our inference solution is practical by testing it on real data. We suggest a novel approach for parameter-tuning. Many wild-field recordings supply a vast amount of data, which we exploit by comparing the data's dynamical properties with the inferred spiking rate dynamical properties to find the best parameter for the inference. To our knowledge, this is a unique new parameter-tuning approach for inference problems.

This paper hence includes the following content: First, we phrase the mathematical challenge that arises from the need to infer the total spiking rate generated by multiple neurons from their collectively generated single fluorescence trace. We then solve the problem and provide proof of our solution. We compare our solution performances to a previously used approach. We do so by using synthetic self-generated data, as well as real recorded data that includes fluorescence (recorded using the wide-field method) in parallel to spikes. Last, we consider a typical experiment in which only fluorescence is recorded and discuss our original parameter tuning analysis with respect to this scenario.

2 Model

We start by examining how single neuron spikes contribute to a recorded fluorescence trace. Multiple recent papers have considered this problem (Jewell & Witten 2018, Jewell et al. 2019, Friedrich et al. 2017, Pnevmatikakis et al. 2016). They had laid the foundation for inferring spike times from a fluorescence trace associated with a single neuron's activity (typically recorded in two-photon cameras; as originally phrased by (Vogelstein et al. 2009)). These studies used the dynamics of the calcium indicator concentration c, its relation to the observed fluorescence y, and its dependency on spike occurrences s to phrase the following auto-regression model:

$$\begin{aligned} &y_t^i = \beta_0^i + \beta_1 c_t^i + \epsilon_t^i, \qquad \epsilon_t^i \sim_{\text{ind.}} \left(0, (\sigma^i)^2\right), \quad t = 1, \ldots, T, \\ &c_t^i = \gamma c_{t-1}^i + s_t^i, \quad t = 2, \ldots, T, \end{aligned} \tag{1}$$

In the above model, the observed fluorescence of a neuron i at timepoint t , y_t^i, is a result of the neuron's calcium concentration at that timepoint, c_t^i, up to some relation factor β_1, a constant shift, β_0, and some noise ϵ_t^i, which is assumed to be Gaussian and i.i.d.. In other words, the fluorescence is a linear readout of the calcium state up to some noise. The calcium itself at each timepoint, c_t^i is equivalent to its value in the previous timepoint, c_{t-1}^i, up to a decay constant, $\gamma \in (0, 1)$ and the additional number of spikes emitted at that time s_t^i. In other words, the calcium decays exponentially over time according to its decay constant, and "jumps" with each spike emission.

We model the relation between a fluorescence trace that results from the activity of many neurons by summing the equations above over all neurons. This summation relies on the assumption that the camera remains in its linear dynamic range, i.e., the arriving light continues to impact the recorded signal in a linear way, allowing us to record the fluorescence generated by multiple

neurons as the linear sum of the fluorescence generated by each neuron separately. We also assume that all neurons have the same calcium indicator dynamics with equal decaying constant gamma, and similar relation to fluorescence emission. We hence receive the following auto-regression model for a fluorescence trace generated by many neuron's activity:

$$\begin{aligned} &\overline{y}_t = \overline{\beta}_0 + \overline{\beta}_1 \overline{c}_t + \overline{\epsilon}_t, \qquad \overline{\epsilon}_t \sim_{\text{ind.}} \left(0, \overline{\sigma}^2\right), \quad t = 1, \ldots, T, \\ &\overline{c}_t = \gamma \overline{c}_{t-1} + r_t, \quad t = 2, \ldots, T. \end{aligned} \tag{2}$$

where $\overline{y}_t \equiv \sum_{i=1}^N y_t^i$ is the total observed fluorescence from all N neurons at timepoint t. $\overline{\beta}_0 \equiv \sum_{i=1}^N \beta_0^i$, $\overline{c}_t \equiv \sum_{i=1}^N c_t^i$, and $\overline{\epsilon}_t \equiv \sum_{i=1}^N \epsilon_t^i$ with $\overline{\sigma}^2 = \sum_{i=1}^N \left(\sigma^i\right)^2$ are the total baseline fluorescence, total calcium at timepoint t, and total noise with total variance, respectively.

The spiking rate, $r_t = \sum_{i=1}^N s_t^i$, is the total number of spikes emitted by all N neurons whose activity contribute to the recorded fluorescence trace. Introducing the spiking rate into the model is a key change from the single-neuron model (1). The latter includes the number of spikes of a single neuron at each timepoint s_t^i, which is sparse (with many zero entries through time), with small single-digit numbers as its non-zero values. On the other hand, we expect the spiking rate to be a large number with gradual changes over time. Since both s_t^i and r_t are at the heart of the solution we seek, their different properties significantly impact the approach we take to the inference solution. We discuss this further as we phrase the optimization problem for our model.

We now chose $\overline{\beta}_1 = 1$ without loss of generality. This corresponds with multiplying the calcium by a constant factor and hence sets a scale between the spiking rate and the fluorescence. When we analyze parallel spiking rates and fluorescence recordings, we discuss how to scale the fluorescence to correspond with this choice.

3 Optimization

Following our model (2), we seek to find the underlying spiking rate r_t from a given fluorescence trace y_t recorded in the wide-field setting. We adopt (and adapt) the approach of formulating (and solving) an optimization problem for finding spikes from a fluorescence trace generated by a single neuron according to model (1) (Jewell & Witten 2018, Friedrich et al. 2017). Since in both cases the models, (2) and (1), hold similar structures, the naturally phrased optimization problem for single neurons fits model (2) with the use of its parameters. This leads to the minimization problem we wish to solve:

$$\begin{aligned} &\text{minimize}_{\overline{c}_1, \ldots, \overline{c}_T, \overline{\beta}_0} \left\{ \sum_{t=1}^T (\overline{y}_t - \overline{\beta}_0 - \overline{c}_t)^2 + \lambda P(\overline{c}_2 - \gamma \overline{c}_1, \overline{c}_3 - \gamma \overline{c}_2, \ldots, \overline{c}_T - \gamma \overline{c}_{T-1}) \right\} \\ &\qquad \text{subject to } \overline{c}_t \geq \gamma \overline{c}_{t-1}, \ t = 2, \ldots, T, \end{aligned} \tag{3}$$

where λ is a weight tuning parameter for the penalty P. For convenience, we rephrase (3) in terms of the spiking rate r_t for $t = 2, \ldots, T$. To this end, we define $r_1 \equiv \overline{c}_1$. We emphasize that this is strictly a definition; the spiking rate at $t = 1$ can not be recovered, and hence r_1 remains biologically meaningless.

Using this definition and our model (2), we can write the relation between the calcium and the spiking rate in a matrix form $r = D\overline{c}$, where D is a $T \times T$ full-rank matrix with 1's on its diagonal and $-\gamma$'s just below the diagonal. This leads to rephrasing of the optimization problem (3) as follows:

$$\text{minimize}_{r_1, \ldots, r_T, \overline{\beta}_0} \left\{ \|\overline{y} - \overline{\beta}_0 \mathbb{1}^\top - D^{-1} r\|^2 + \lambda P(r_2, \ldots, r_T) \right\} \text{ subject to } r_t \geq 0, t = 2, \ldots, T, \tag{4}$$

where $\mathbb{1}^\top$ is a column vector of length T with all elements equal to 1.

Proposition 1. *The optimization problem* (3) *is equivalent to* (4)*, in the sense that* $\hat{\beta}_0 = \tilde{\beta}_0$ *and* $\hat{r} = D\tilde{c}$*, where* $\hat{r} = (\hat{r}_1, \ldots, \hat{r}_T)^\top$ *and* $\hat{\beta}_0$ *solve* (4)*, and* $\tilde{c} = (\tilde{c}_1, \ldots, \tilde{c}_T)^\top$ *and* $\tilde{\beta}_0$ *solve* (3)*.*

The proof is given in Appendix A.

At this point, we choose the form of the penalty $P(r_2, \ldots, r_T)$. The form will dictate the properties of the solution r we find. Hence, to fit it best to wide-field imaging, we are required to consider the properties of the underlying spiking rate r in this case. The penalty that was chosen in the case of a fluorescence trace resulting from a single neuron activity is no longer relevant to us. This is because the properties of the number of spikes of a single neuron are vastly different than the properties of the average spiking rate of many neurons. The typical profile of the number of spikes per bin in a single neuron includes many zero entries and otherwise low positive numbers. To retrieve solutions with such properties, ℓ_0 penalty was chosen (Jewell & Witten 2018, Jewell et al. 2019) or ℓ_1 previously (Vogelstein et al. 2009, Pnevmatikakis et al. 2016). In our case, the spiking rate of many neurons is expected to be a positive number that continuously changes through time (either rapidly or slowly but always bounded; Figure 3ai,ii for example). We note that we can consider the spiking rate and the other parameters from our model (2) as the average quantities rather than the sum across the many neurons whose activity contributes to the recorded wide-field fluorescence. This change is straightforward but constructive. Since the number of neurons contributing to a single fluorescence trace in wide-field recordings is very large, the average number of spikes per neuron can be treated as a continuous number. To fit best its properties, continuously changing positive bounded values through time, we choose the ℓ_2 penalty through time $P\left((r_2 - r_1)^2, (r_3 - r_2)^2, \ldots, (r_T - r_{T-1})^2\right)$. This guarantees solutions for r with the desired properties. It also allows us to find and analyze an exact solution.

We derive the optimization problem for spiking rate inference from fluorescence in wide-field recordings:

$$\text{minimize}_{r_1,\ldots,r_T,\overline{\beta}_0} \left\{ \|\overline{y} - \overline{\beta}_0 \mathbb{1}^\top - D^{-1} r\|^2 + \lambda \sum_{t=3}^{T} (r_t - r_{t-1})^2 \right\} \text{ subject to } r_t \geq 0, t = 2, \ldots, T, \quad (5)$$

4 Solution

At first glance, the optimization problem (5) we wish to solve is simply a direct sum of two ℓ_2 expressions, which can be naively approached by equating to zero the derivatives with respect to β_0 and r and finding their minimum (as Least Mean Squares). However, a more careful examination shows that the solution of r has to be non-negative. And moreover, for T measurements of the fluorescence, only $T-1$ spiking rates timepoints are retrieved (since they correspond to the change in the fluorescence through their dependency on the calcium; remember r_1 is not a spiking rate only a notation). In other words, we have one degree of freedom, which our solution has to "absorb" while maintaining its structure so it remains a meaningful spiking rate solution. In this section, we prove that this degree of freedom corresponds to the constant shift in the fluorescence being arbitrary (originally a result of the $\Delta F/F$ procedure). This results in a solution for the spiking rate that is also invariant under a constant shift. We use this shift degree of freedom to ensure that the spiking rate solution is non-negative as required.

We start by simplifying the optimization problem.

Proposition 2. *The pair* $(\tilde{r}, \tilde{\beta}_0)$ *is a solution to problem* (5) *if and only if* $\tilde{\beta}_0 = \frac{1}{T}\mathbb{1}^\top(\overline{y} - D^{-1}\tilde{r})$ *and* $\tilde{r}$ *is a solution to*

$$\text{minimize}_{r_1,\ldots,r_T} \left\{ \|\tilde{y} - Ar\|^2 + \lambda \sum_{t=3}^{T} (r_t - r_{t-1})^2 \right\} \text{ subject to } r_t \geq 0,\ t = 2, \ldots, T, \tag{6}$$

where $A = \left(I - \frac{1}{T}\mathbb{1}\mathbb{1}^\top\right) D^{-1}$ *and* $\tilde{y} = \left(I - \frac{1}{T}\mathbb{1}\mathbb{1}^\top\right) \overline{y}$*, with* I *the* $T \times T$ *identity matrix and* $\mathbb{1}\mathbb{1}^\top$ *a* $T \times T$ *matrix of ones.*

We prove Proposition 2 in Appendix B.
Note that $\tilde{y}$ is the mean subtracted version of $\overline{y}$ and Ar is the column-wise mean subtracted version of $D^{-1}r$. This corresponds to the constant shift of the recorded fluorescence β_0 being arbitrary with respect to the spiking rate. In the following section, we explicitly prove it makes the spiking rate invariant under a constant shift.

We now solve

$$\text{minimize}_{r_1,\ldots,r_T} \left\{ \|\tilde{y} - Ar\|^2 + \lambda \sum_{t=3}^{T} (r_t - r_{t-1})^2 \right\} \text{ subject to } r_t \geq 0,\ t = 2, \ldots, T. \tag{7}$$

According to proposition 2 we can solve 5 by simply solving 7 to obtain the spiking rate and obtain β_0 by its closed form expression. Finding β_0 allows us to reconstruct the noise-removed version of the original recorded fluorescence with its shift, which is useful for testing the results.

To solve (7) it is constructive to present the following, simpler optimization problem, which includes the same ℓ_2 expressions but has no constrains on r:

$$\text{minimize}_{r_1,\ldots,r_T} \left\{ \|\tilde{y} - Ar\|^2 + \lambda \sum_{t=3}^{T} (r_t - r_{t-1})^2 \right\} \tag{8}$$

This problem can be solved analytically by computing $\nabla_r f(r) = 0$, with $f(r) = \|\tilde{y} - Ar\|^2 + \lambda \sum_{t=3}^{T} (r_t - r_{t-1})^2$. For convenience purposes we define Z as follows:

$$Z_{i,j} = \begin{cases} -1, & \text{if } |i-j| = 1 \text{ and } i + j > 3 \\ 2, & \text{if } i = j \text{ and } 2 < i < T \\ 1, & \text{if } i = j = 2 \text{ or } i = j = T \\ 0, & \text{otherwise} \end{cases}. \tag{9}$$

Note that $\sum_{t=3}^{T} (r_t - r_{t-1})^2 = r^\top Z r$.

Now f takes the form $f(r) = \|\tilde{y} - Ar\|^2 + \lambda r^\top Z r$ and $\nabla_r f(r) = 0$ yields:

$$-2A^\top(\tilde{y} - Ar) + 2\lambda Z r = 0. \tag{10}$$

The solution to (10), and hence to problem (8), takes the form

$$r = \left(A^\top A + \lambda Z\right)^+ A^\top \tilde{y} \tag{11}$$

where $\left(A^\top A + \lambda Z\right)^+$ is the pseudo-inverse of $A^\top A + \lambda Z$. Note that we have transformed the degree of freedom in the shift subtracted fluorescence and mean subtracted $D^{-1}r$ columns into the pseudo inverse. In the following section we find the explicit form of this degree of freedom.

Proposition 3. *Let $\hat{r}$ denote a solution to* (8)*. Then $\tilde{r} = \hat{r} + d\mathbb{1} + \tilde{r}_{m1}$ also solves* (8) *where $\tilde{r}_{m1}^\top = (r_{m1}, 0, 0, \ldots, 0)$ is a T-length vector with all entries equal zero except for its first entry which satisfies*

$$r_{m1} = -d\frac{\sum_{j=1}^{T} A_{1j}}{A_{11}} \tag{12}$$

We prove proposition 3 in Appendix C. Note that d can be any constant number. This proposition implies that problem (8) (and the original problem (5) as a result) is not strictly convex. It has many solutions which differ by shifting the spiking rate $r_2, \ldots, r_T$ by a constant and compensating for this shift by changing the value of r_1 (which is biologically meaningless; and β_0 accordingly). This concludes our proof that the degree of freedom in the problem translates to the spiking rate solution being invariant under a shift.

Solution (11) to problem (8) and propositions 2 and 3 lead to Algorithm 1 for solving (5).

Algorithm 1 Continuously-Varying Spiking Rate Inference: Solving (5)

Given the signal, $\overline{y}$, and the calcium decay in a time step, γ, do:

1. Calculate $\tilde{y} = \left(I - \frac{1}{T}\mathbb{1}\mathbb{1}^\top\right)\overline{y}$, and $A = \left(I - \frac{1}{T}\mathbb{1}\mathbb{1}^\top\right)D^{-1}$.
2. Calculate $r^{unconstrained} = \left(A^\top A + \lambda Z\right)^+ A^\top \tilde{y}$.
3. Find $d = -\min(0, r_2^{unconstrained}, r_3^{unconstrained}, \ldots, r_T^{unconstrained})$.
4. Let $r = r^{unconstrained} + d\mathbb{1} + \tilde{r}_{m1}$,
   with $\tilde{r}_{m1}^\top = (r_{m1}, 0, 0, \ldots, 0)$ and $r_{m1} = -d\frac{\sum_{j=1}^{T} A_{1j}}{A_{11}}$.
5. Let $\overline{\beta}_0 = \frac{1}{T}\mathbb{1}^\top(\overline{y} - D^{-1}r)$.

5 Results

5.1 Evaluation

Here, we evaluate the performances of our Continuously-Varying algorithm (Con-Var; Algorithm 1). To this end, we simulated spiking rate traces (see example in Figure 1ai, ivory line; see Appendix D for details). We imitated wide-field recording of fluorescence traces resulting from the simulated spiking rate activity according to the calcium indicator dynamics in our model (2). This included transforming the simulated spiking rate traces into calcium traces according to $c = rD^{-1}$ (a direct result of (2), see example in Figure 1aii) and adding Gaussian-independent noise and some constant shift to the calcium to receive the recorded fluorescence (Figure 1aiii). For the calcium indicator dynamics, we choose $\gamma = 0.95$, which resembles real calcium indicator dynamics (see section 5.3).

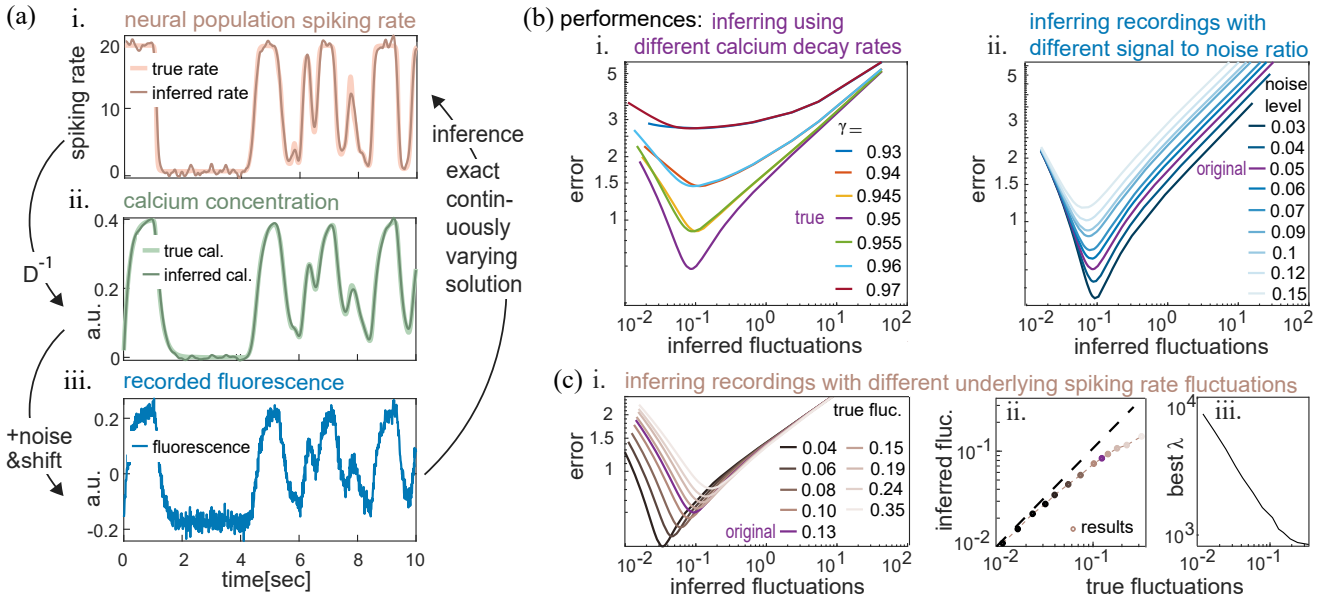

Figure 1: Evaluating the Continuously-Varying Algorithm (Con-Var; Algorithm 1). (a). An example: A simulated underlying spiking rate (i. ivory line) is convolved into calcium concentration (ii. light blue line) and transforms into recorded fluorescence (iii. blue line) using model (2). The fluorescence is then inferred back into spiking rate using Con-Var (i brown line). Note that the fluorescence units are presented at a 1/1000 scale to fit units of a typical camera setting. (b) i. The mean error (13) between simulated spiking rates and inferred spiking rates as a function of the mean inferred rate fluctuations (14). The inferred spiking rates were found by applying Con-Var on the fluorescence traces generated from the underlying spiking rates (as in (a)). For the inference Con-Var was applied many times, each time with a different γ (yielding the multiple lines) and λ (yielding changes in the inferred spiking rate fluctuations), with $0.93 \leq \gamma \leq 0.97$ and $10000 \leq \lambda \leq 1$. The mean was calculated over 10sec long 10000 traces generated as described in Appendix D. For generating calcium traces, we used underlying (true) $\gamma = 0.95$, and for the transformation to fluorescence we used $\overline{\sigma} = 0.05|\max\overline{y} - \min\overline{y}|$. ii. The mean error (13) between simulated spiking rates and inferred spiking rates as a function of the mean inferred rate fluctuations (14) for multiple underlying spiking rate datasets differentiated by the value of the noise $0.03|\max\overline{y} - \min\overline{y}| \leq \overline{\sigma} \leq 0.15|\max\overline{y} - \min\overline{y}|$. γ was held fixed at 0.95 for the inference, all other parameters are as in i. (c). i. The mean error (13) between simulated spiking rates and inferred spiking rates as a function of the mean inferred rate fluctuations (14) for multiple underlying spiking rate sets with different mean fluctuations (see Appendix D). γ was held fixed at 0.95 for the inference, all other parameters are as in (b)i. ii. Inferred spiking rate fluctuations as a function of the mean underlying (true) spiking rate fluctuations at the minimum error point retrieved for each true spiking rate fluctuations in i. iii. Values of best penalty λ, (which are the values of λ at the minimum error points in i), as as function of the true spiking rate fluctuations. The change in the best penalty value allows to fit the inferred spiking rate fluctuations to the true spiking rate fluctuations (shown in ii) and maintain minimal error with the same order of magnitude (in i).

We used Con-Var to infer the spiking rate from the simulated fluorescence recordings. We evaluated the quality of its inference by calculating the error:

$$\mathrm{err}(x, x^{est}) = \frac{1}{T-1} \sum_{t=2}^{T} |x_t - x_t^{est}|, \tag{13}$$

represented by the difference between the original simulated underlying spiking rate, $x = r$, and the inferred spiking rate $x^{est} = r^{est}$. We used the mean subtracted versions of both r_t and r_t^{est} for $t = 2, \ldots, T$ to evaluate (13) since we recall that the latter is invariant under a constant shift according to Proposition 3.

The amount of fluctuations in a trace is another useful quantity that helps us evaluate the quality of our inference. We quantify the fluctuations as follows,

$$\zeta(x) = \sum_{t=3}^{T} (x_t - x_{t-1})^2, \tag{14}$$

where x can be any time-dependent parameter (spiking rate r, calcium $\overline{c}$ or fluorescence $\overline{y}$). The fluctuations measure how vastly and frequently the signal changes through time. Note that the expression in equation (14), for $x = r$, is equal to the penalty in the optimization problem (5). Since a decrease or an increase in the weight tuning parameter λ, increases or decreases, respectively, the penalty value of a solution to (5), the fluctuations increase or decrease according to λ as well. Figure 1bi was generated by varying the penalty weight λ and the calcium decay constant, γ, we tested for inference. The results were plotted for each γ according to the fluctuations achieved with each λ (on the x-axis, with higher λ promoting lower fluctuation and vice versa).

For each calcium decay constant, γ, used for inference, there is an optimal fluctuations level and, accordingly, an optimal λ which retrieves a solution with minimal error. Con-Var succeeds in maintaining the fluctuations of the inferred spiking rate around the minimal error across a wide range of γ values tested (see Figure 1bi, x values around the minimum error for the various γ-plots). This suggests that a slight deviation in the calcium decay constant γ is not expected to change the optimal inference solution by a significant amount. Jitters in γ's value can be caused by neuron internal properties variability, high tissue density between neurons, significant changes in firing rate, and more. Our results indicate that despite these jitters the optimal inferred solution is maintained (at a very high proximity).

Importantly, Con-Var is consistent in the sense that it retrieves global minimum error using the same γ as was used for generating the calcium traces (Figure 1bi, purple line). Meaning it finds correctly the true underlying calcium decay dynamics. For this decay rate, γ, the error is also lower at any fluctuation level. This is a realization of the algorithm being convex, also with respect to its tuning parameters.

The signal-to-noise ratio can vary significantly between experiments. Therefore, it is essential to test the solution properties amid various noise levels. We changed SNR by changing the variance $\overline{\sigma}$ of the noise $\overline{\epsilon}$ with respect to the total range of the fluorescence. Although the minimal error of the optimal inferred solution increases with the noise, the fluctuation levels remain similar across many noise levels (see Figure 1bii; plots are given for SNR values ranging from 1/0.3 to 1/0.15; the minimal error over these SNRs doubles, and the fluctuations increase by about 1.5, despite the five-fold noise growth).

The above suggests that inferred solutions remain similar and maintain their properties amid possible noise variability. This robustness to noise is achieved because one by-product of Con-Var

is noise removal. For the underlying spiking rate to be inferred, the noise at each time point in the fluorescence recordings has to be estimated. The spiking rate is found "after" the removal of the estimated noise. Since Con-Var utilizes calcium dynamics for the inference, its noise removal is done by finding a fluorescence signal that is a calcium signal plus noise. The calcium signal obeys calcium indicator dynamics, which is biologically meaningful. Hence, Con-Var solutions, which are recorded-noise-free spiking rates, achieve robustness to noise since they are found with respect to a biological meaningful criterion.

Inferred solutions by Con-Var maintain different underlying spiking rate fluctuations (Figure 1cii) by correctly fitting the penalty λ at the optimal solution (Figure 1ciii). In other words, the inferred spiking rates fluctuation levels are similar to the underlying spiking rate fluctuations. They correctly follow changes in the fluctuations across an extensive range of underlying fluctuation levels. This allows proper fit of the inference to a large range of recording scenarios with wide range of underlying spiking rate fluctuations.

5.2 Comparison

Here, we compare the Continuously-Varying Algorithm (Con-Var; Algorithm 1) performances to the Lucy-Richardson algorithm (Lucy). Originally, Lucy was developed for astrophysics research purposes. It was designed to infer underlying light sources distorted by lenses (Lucy 1974). Recently, Lucy was modified to infer spiking rates distorted by the calcium dynamics in wide-field experiments (Wekselblatt et al. 2016). At their core, the modifications rely on considering the calcium dynamics as the relevant filter (in time rather than in space). The filter can be expressed as γ^t for $t = 1, \ldots, p$ where p is sufficiently large for capturing a calcium signal decay. To the best of our knowledge, this is the only algorithm in the published literature that has been used and tested for continuous spiking rate signal inference from wide-field imaging.

The main advantage of utilizing Lucy for wide-field imaging is an easy implementation access. Since the algorithm is commonly used in image processing, pre-written, and optimized code for it is available in most programming languages. The main drawback of Lucy is that, at its core, it is designed to solve a problem with different underlying assumptions compared to the model of recorded fluorescence generated by neural activity. For instance, Lucy assumes that the noise in the recordings follows a Poisson distribution, which is unlikely in fluorescence recordings. Our tests, which compare the error between the inferred and underlying spiking rate, reveal that these assumption discrepancies lead to various issues, most notably to the loss of correlation between the inferred and underlying spiking rate fluctuations, as described below.

To test Lucy and compare it to Con-Var, we inferred the same fluorescence traces we generated for Section 5.1, this time with Lucy. It is evident that the inferred spiking rate by Lucy fails to capture the full span of the underlying spiking rate (see Figure 2ai. brown line vs. ivory line). Since the fluorescence and the inferred spiking rate have arbitrary units of measurement, this failure is not a critical issue. It has no direct biological meaning that impacts further analysis of the results.

The spiking rate inferred by Lucy is evidently much noisier compared to the spiking rate inferred by Con-Var (see Figure 2ai brown line vs. ivory and light blue lines). This creates local dissimilarities between the underlying and the inferred spiking rate dynamics. At many timepoints, the underlying spiking rate remains constant while the spiking rate inferred by Lucy increases or decreases. As a result, the fluctuations of the inferred spiking rate by Lucy are higher and remain high (with similar value) across many underlying spiking rate fluctuations (Figure 2e). The inability of Lucy to generate solutions that follow the underlying fluctuations is a significant

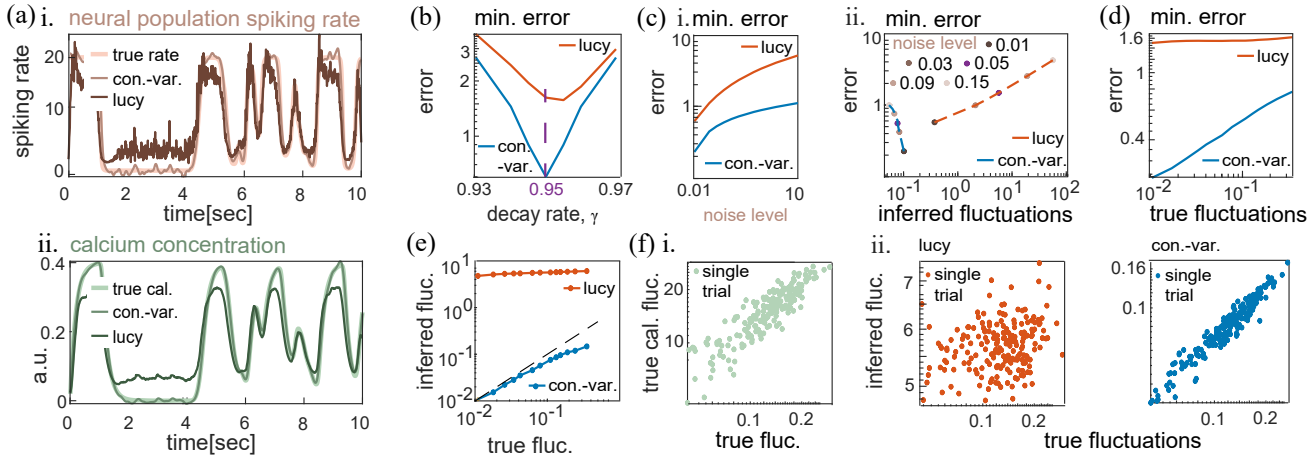

Figure 2: Comparing the Lucy-Richardson Algorithm (Lucy) to the Continuously-Varying Algorithm (Con-Var; Algorithm 1). (a) The same example as in Figure 1), with the additional inference by Lucy. (b) i. The mean error (13) between simulated spiking rates and inferred spiking rates as a function of the different γs used for inference. The same simulated underlying spiking rate dataset as in Figure 1bi was inferred here by Lucy. Similarly, the dataset in Figure 1bii was inferred by Lucy for plots (c)i,ii and the dataset in Figure 1c was used for plots (d) and (e). Con-Var results are represented here by the minimum error achieved with each γ (min. error of each line in Figure 1bi). The same holds for plot (c) with con-var. minimum error taken from Figure 1bii and plots (d) and (e) Con-Var minimum error taken from Figure 1ci. (c) The mean error (13) between simulated underlying spiking rates and inferred spiking rates as a function of: i. The noise level (same as Figure 1bii). ii. The mean inferred rate fluctuations (14) for the noise levels in i. (d) The mean error (13) between underlying simulated spiking rates and inferred spiking rates for multiple underlying spiking rate datasets with different mean fluctuations. (e) The mean inferred spiking rate fluctuations as a function of the mean underlying spiking rate fluctuations. The black dash line marks $x = y$. (f) Relations between the followings (for 200 example traces): i. The true underlying calcium fluctuation levels and the true underlying spiking rate fluctuation levels. ii. The true underlying calcium fluctuation levels and the inferred by Lucy spiking rate fluctuation levels. iii. The true underlying calcium fluctuation levels and the inferred by Con-Var spiking rate fluctuation levels. Pearson correlations (considering the displayed examples): i. 0.88 ii. 0.55 iii. 0.91. Examples were generated with best γ and λ for each inference method.

weakness of Lucy. For much higher values of underlying spiking rate, it is possible that Lucy could infer the fluctuations correctly. However, this is a very restricting constraint on the possible underlying spiking rate traces Lucy can infer with correct fluctuations. Notably, this possible well-fit-fluctuations window does not agree with the fluctuation levels we see in experiments (Figure 3di).

The inability to fit the fluctuations on average also manifests in single traces. Our simulated single spiking rate traces exhibit variable fluctuation levels. The single calcium trace fluctuations are highly correlated with the fluctuations of the single spiking rate traces from which they were generated, (Figure 2fi, Pearson correlation 0.8846). In other words, the calcium data (and the fluorescence, not shown) maintains the fluctuations' structure in the spiking rate data. However, the inference by Lucy retrieves spiking rate fluctuations which are poorly correlated with the underlying spiking rate fluctuations (Figure 2fii., Pearson correlations 0.3570). This is not a mandate

of inference; Con-Var successfully maintains the fluctuations' structure in the data (Figure 2fiii. Pearson correlation 0.8936). It is a core difference between Con-Var and Lucy and a realization of the assumptions of Lucy not fitting the wide-field setup.

Another significant drawback of Lucy is its inconsistency concerning finding the correct calcium dynamics. It often yields the best results (minimal error between underlying and inferred spiking rate) with a calcium decay constant γ that is close to, but not the accurate one we set for the calcium dynamics, leading to an incorrect inference of γ (Figure 2b, correct γ is marked by purple dashed line) and skewed inference of the spiking rates. This limitation further degrades the algorithm's fit for spiking inference in wide-field imaging.

5.3 Recorded Data

Next, we tested the performances of the Continuously-Varying Algorithm (Con-Var; Algorithm 1) and the Lucy-Richardson algorithm (Lucy) on real recorded data. We used simultaneous recordings of wide-field fluorescence and spikes. Unfortunately, recording the exact neural activity and simultaneously its generated fluorescence trace in the wide-field setup is currently not possible (technically). Hence, we do not have the exact underlying spiking rate to test the algorithms. However, we hold an estimate of it. The dataset we use includes spikes recorded from the spacial location of the fluorescence origin, Clancy et al. (2019) Supplementary figure 4c; see Appendix E for more details about this dataset. We verified that the spikes in the dataset represent well (in good correlation) the exact spiking rates that drive the recorded fluorescence. For our study purposes, this implies that the spiking rate is also a noisy measurement (of the exact underlying spiking rate). We will return to discuss its implications as we tune the inference parameters of Con-Var.

Prior to testing the algorithms on the data, we rescaled the recorded fluorescence traces such that the assumption of $\overline{\beta}_1 = 1$ we made in (2) will hold. Since both the spiking rate r and the fluorescence $\overline{y}$ are given in the dataset, $\overline{\beta}_1$ is implicitly given as well. We find its value with equation (28) using the same value of γ as we use for inference; see Appendix F. We use the recorded fluorescence multiplied by $\frac{1}{\overline{\beta}_1}$ for inference (then the assumption $\overline{\beta}_1 = 1$ holds). This allows us to compare the (mean substructed) inferred spiking rate directly with the (mean substructed) recorded spiking rate.

We inferred the data using Con-Var with various calcium decay constant γ values and various penalty, λ levels. As discussed in Section 5.1, any change in penalty results in a change in the fluctuations of the inferred spiking rate. We plotted the error across fluctuations for each value of γ we used, Figure 3bi. The results show that Con-Var remains convex with respect to the error for both parameters γ and λ, Figure 3bi, similar to the simulated data results, Figure 3bi. For low fluctuations, the error lines are shallow. This is very encouraging since it implies that small deviations in the calcium decaying constant γ or penalty λ, have only a small impact on the inferred spiking rate solution. This is important since the exact true underlying calcium decay constant is unknown to us. Clancy et al. (2019) reported $\gamma = 0.97$ for single neurons for the dataset in hand. Chen et al. reported a similar value with 10% deviation across cells, meaning $\gamma = 0.97 \pm 0.003$. These deviations originate from natural cell variability, including physical properties and significant changes in spiking rates (Chen et al. 2013). In wide-field recordings, deviations can also be caused by a substantial presence of non-neuronal tissue, like glial cells. Our Con-Var yields minimal average error with $\gamma = 0.973$ while Lucy yields minimal error with $\gamma = 0.974$, Figure 3biii.

We tested how robust is our estimate of the underlying γ by looking at all the γs that minimize the error in each of the 65 segments of the data (see Appendix E). Our analysis shows that the

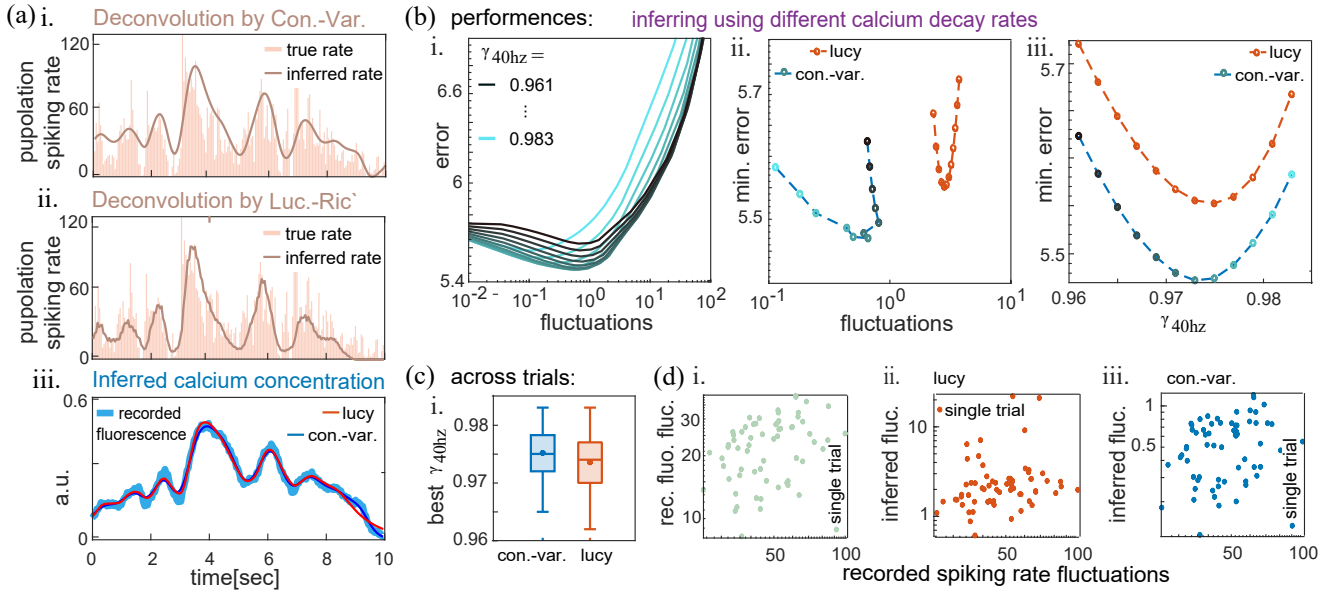

Figure 3: Evaluating algorithm performances on parallel spiking and wide field fluorescence recordings. (a) An example of an underlying spiking rate (i. ivory spike counts) and the inferred spiking rate using Con-Var (i. brown line) or Lucy (ii. brown line). For the inference, the fluorescence recorded in parallel was used (iii. light blue). (b) The mean error (13) between the recorded spiking rates and inferred spiking rates by Con-Var as a function of the mean inferred rate fluctuations (14). The inferred spiking rates were found by applying Con-Var on the fluorescence traces recorded in parallel. Con-Var was applied many times, each time with a different γ (yielding the multiple lines) and λ (yielding changes in the inferred spiking rate fluctuations), with $0.961 \leq \gamma \leq 0.983$ and $10000 \leq \lambda \leq 0.001$. ii. The same as i, including Lucy for inference (red line). Results presented for Con-Var (blue line) include only minimum error achieved by Con-Var with each γ (blue circle shades fit the lines in i from whom the minimum error value was taken). iii. Same as ii, but as function of γ used for inference. (c). The medians (lines inside), means (circles), quartiles (boxes) and outliers (whiskers) of all data segments' best γ yielding minimal error between underlying and inferred spiking rate for inference using Con-Var (blue) or Lucy (red). (d) The recorded spiking rate fluctuations and their: i. Recorded fluorescence fluctuations ii. Inferred by Lucy spiking rate fluctuations iii. Inferred by Con-Var spiking rate fluctuations. Pearson correlations for the plots: i. 0.375 ii. 0.138 iii. 0.265. Best γ and λ according to (b) were used for the inference.

error between the recorded and inferred spiking rates is minimized best with $\gamma = 0.9749 \pm 0.0042$ if inferred with Con-Var or $\gamma = 0.9731 \pm 0.0047$ if inferred with Lucy, where the $\pm error$ is the standard deviation. The calculated range retrieved by both the Lucy and Con-Var falls within the known single-cell range. Con-Var results suggest the calcium dynamic is slower (with higher decaying constant mean value). Results are shown in Figure 3c (in circles) where similar results by the median values are plotted as well (lines in boxes). Note that when we calculated the mean error first, and best γ for the mean error, as in Figure 3biii, it is Lucy's inferred solutions that suggest slower dynamics (higher best γ). This is because parameter values skew the error values, and hence, the estimate of best γ is biased. Last, we note that the deviations of best γ for inference by Lucy are larger, indicating more sensitivity to the particular instances of the data.

Similar to the results on simulated data, the fluctuations of the inferred spiking rates by Lucy are higher than those of the inferred spiking rates by Con-Var, Figure 3bii.

The fluctuations' structure in the data, i.e., the correlations between the recorded fluorescence fluctuations and the recorded spiking rate fluctuations, is limited by the noise and by the partial representation of the true underlying spiking rate compared to the recorded spiking rate (as explained above). This manifests in limited fluctuation correlations in the data, Figure 3di, Pearson correlation of 0.375. And yet, similar to the results on simulated data, Lucy breaks the existing fluctuation correlations, Figure 3dii Pearson correlations of 0.138. In contrast, Con-Var maintains the fluctuations' structure within the data better, Figure 3diii Pearson correlations of 0.265.

5.4 Analyzing Recorded Fluorescence

All the above analyses directly compared the inferred spiking rate with an underlying spiking rate, either recorded or simulated. However, in a typical wide-field experiment, only fluorescence data is recorded. Hence, for our algorithm to be practical, we need to be able to use it, including tuning its parameters and testing it, without using underlying spiking data. Here, we do just that. We tune and test our Continuous-Varying Algorithm by using only fluorescence traces. We then compare our results to the previous section. In other words, in this section, we ignore the recorded spiking rates but eventually utilize the information they carry to evaluate performances.

First, we tune the calcium decay rate γ. To do so we use its known value from single neurons data. According to our model , equation (2), the calcium decay rate for many neurons' calcium concentration is equal to the calcium decay rate for single neurons' calcium concentration, equation (1). Our results in Section 5.3 show that, indeed, the best γ for the Con-Var and Lucy is within the range of single neurons' γ.

Our main challenge is to find penalty λ to use in the Con-Var algorithm without relying on minimizing the error between the inferred and underlying spiking rate. We survey here two methods for doing so. The first method was proposed by Jewell & Witten (2018). It divides the fluorescence trace into a training set and a testing test. While this method fits well with noise removal from the recordings of the fluorescence, it doesn't account for noise in the spiking rate. We present a second, novel approach to fit the penalty λ. It takes advantage of the vast amount of data available. It maximizes the preservation of data properties (prior and post inference) in relation to other portions of the data. For spiking rate inference, we use the fluctuation levels of the inferred spiking rate (relative to the distribution of the fluorescence fluctuations) as the conserved property. We discuss both tuning methods for λ in the following sections.

5.4.1 Choosing Penalty by Cross Validation

We summarize here the scheme for finding the penalty λ as described in Jewell & Witten (2018), Appendix B. For this manuscript's coherence, we also bring the complete algorithm according to this scheme (Jewell & Witten (2018) Algorithm 3) in Appendix G.

In short, the algorithm divides the fluorescence trace into two. One part consists of recordings made at even indexed time points and the other at odd indexed time points (Figure 4ai,iii). The two parts become the training set and the testing set of each other. The fluorescence is assumed to change slowly compared to the recording rate at the noise in the recording is assumed to be independent through time (equation 2). Hence, the average of two consecutive points in one trace (odd, for example) after inference and convolution back to calcium can be compared with the appropriate point of the other trace (even, for example) before inference (Figure 4aii). The penalty λ is chosen to be the maximal λ value, which still gives an error that is less than the minimum error plus the error standard deviation at the minimum error point (Figure 4b).

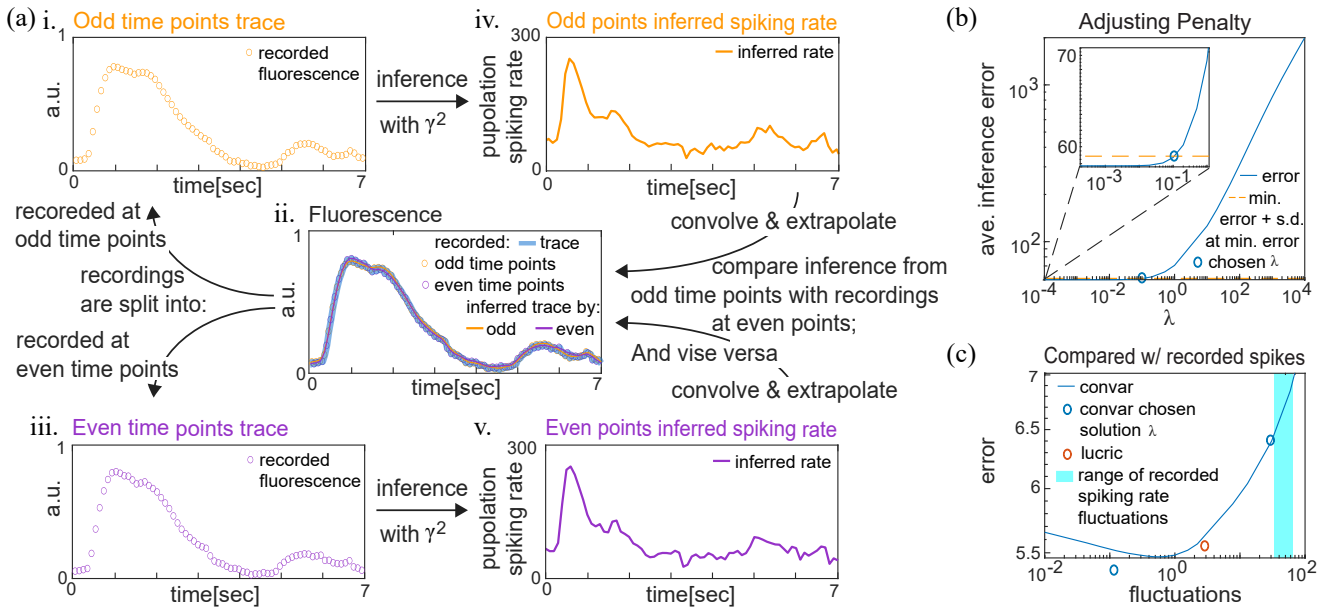

Figure 4: Evaluating the penalty λ by a cross validation scheme, Jewell & Witten (2018). (a) An example of a fluorescence trace (ii light blue), separated into an odd (i) and even (iii) indexed time points. Inferred from them, by Con-Var, are the odd and even indexed time points spiking rate (iv,v respectively). Their convolution to calcium signals are compared with the recorded fluorescence at the even and odd indexed time points (respectively, see Appendix G for the complete algorithm). This process is repeated with different λ values. (b) The average error (13) between the recorded fluorescence at even or odd indexed time points and the convolved calcium of spiking rate inferred from odd or even fluorescence indexed time points (respectively), as a function of λ. The chosen penalty (blue circle) is the maximum λ with error (blue line) smaller than the minimal error achieved plus the standard deviation at the minimal error (orange dashed line). (c) The mean error (13) between the recorded spiking rates and inferred spiking rates as a function of the mean inferred rate fluctuations (14). The plot is taken from Figure 3b with $\gamma = 0.973$. Varying λ yields the changes in the inferred spiking rate fluctuations and we marked in blue circle the cross-validation scheme chosen λ in (b). Red circle marks the solution given by Lucy for comparison.

We applied Algorithm 2 on the fluorescence recordings from Clancy et al. (2019). The results were good (Figure 4c blue circle) but didn't minimize the error as possible. We note that this method's chosen λ does yield a spiking rate solution with fluctuations that fit, on average, the underlying recorded spiking rate fluctuations (Figure 4c light blue area). Hence, the relative noise this approach separates from the calcium signal manages to retrieve inferred underlying spiking rate with correctly maintained dynamical properties.

5.4.2 Analyzing and Tuning Penalty by Maintaining Signal Properties for the Case of Noisy Underlying Spiking Rate

We offer here an alternative, a novel approach, for tuning the penalty λ. It includes an analysis of the results, which can also be used in perplexing cases to evaluate the data. It takes advantage of large datasets by looking at the relative amount of some dynamic quantity each segment from the data holds. It fits wide-field experiments since they typically make much fluorescence available. The core idea behind this approach is to find a penalty that best maintains the relative amount

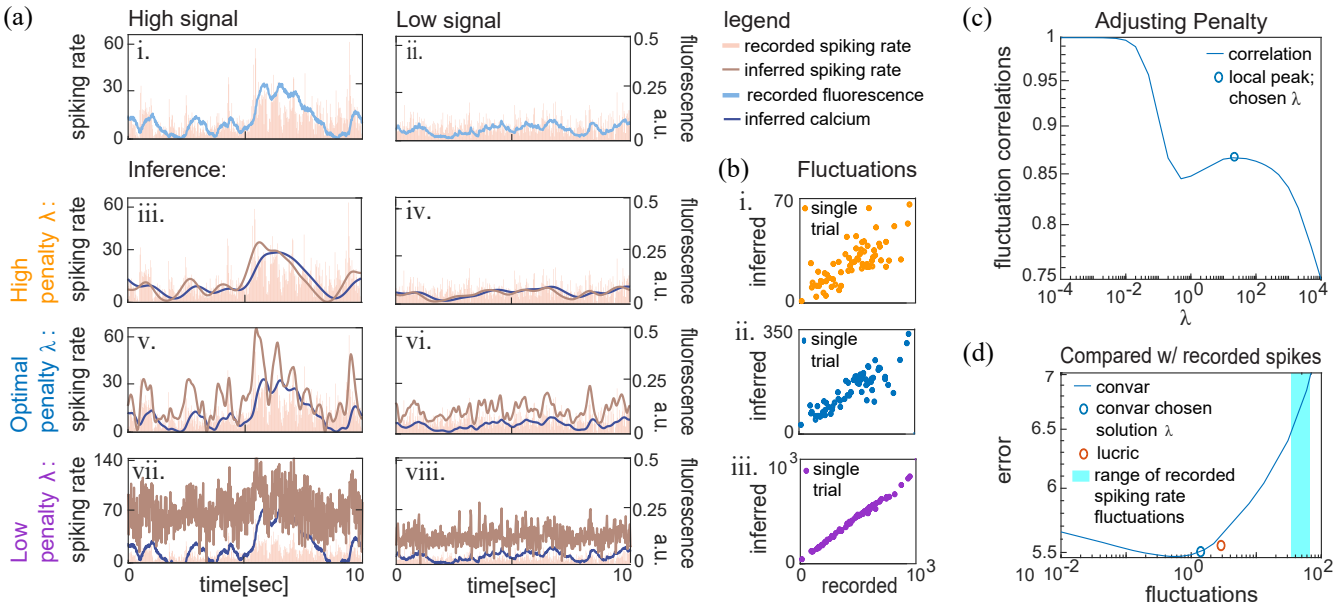

Figure 5: Analyzing and estimating the penalty λ according to fluctuation correlations. (a) i,ii. Two examples of recorded fluorescence (light blue) and the parallel recorded spiking rates (ivory). iii-vii. The inferred spiking rate (brown) from the two fluorescence examples using iii,iv. high, v,vi. medium and vii,viii.low penalty values and their resulting convolved calcium (blue). The optimal penalty succeeds best at differentiating between the high and low signal examples. (b) Scatter plots of all 65 fluorescence segment fluctuations as recorded and their corresponding calcium segment fluctuations as calculated from the inferred spiking rates convolved back to calcium. Pearson correlation for the plots: i. 0.753 ii. 0.867 iii. 0.997. (c) The correlation between the recorded fluorescence fluctuations and the fluctuations of the convolved calcium from the inferred spiking rate, as a function of λ. The chosen penalty (blue circle) maximizes, locally, the correlation. (d) Same as Figure 4c with the blue circle representing the local minimum solution as marked in (c).

of a fundamental dynamic quantity prior to and post-inference. The logic behind it is that the chosen dynamic quantity reflects a characteristic of the underlying signal. Although its absolute amount may be altered by inference, the relation between amounts of it in different segments is preserved. In our case, the fluctuation level is the primary dynamical quantity we've been using. We expect segments with higher fluctuating fluorescence to have higher fluctuating spiking rates inferred from them compared to other segments' inferred spiking rates, on average, (see Figure 5a for an example) and vice versa for lower fluctuations. We further expect the inference to maintain the relations between the fluctuation levels prior to and post-inference.

Maintaining the correct fluctuations amount order (and ratio) across segments can be achieved best if the fluctuations of the spiking rate receive their proper weight. This translates directly to adjusting the penalty λ. This also implies finding the correct signal-to-noise ratio in the fluorescence trace and, by inference, removing the noise. With too much signal inferred as noise, much of the signal is lost (Figure 5aiii); with too little, the noise is inferred as signal (Figure 5viii), and the result is an overly noisy inference. We wish to maximize the information in the spiking rate (our underlying signal) while still maximizing the noise removal. This corresponds to correctly maintaining the dynamic quantity (here the fluctuations) ratio post inference.

To choose λ, we hence calculate the correlation between the fluctuations of different segments prior to and post-inference (after being convolved back to calcium). We repeat this with many

values of λ. With zero penalty (no noise removal), the restoration after inference (and convolution) is perfect, and the correlation is 1. However, this preserves all the noise in the inference and doesn't meet our needs. At some middle value of λ the correlation has local maxima. At this point, the maximal dynamic information is conserved, with most possible noise being removed from the signal. This is our chosen λ.

Our data shows this finding-local-maximum scheme (as described above) results in a solution that is very close to being the best available. Note that because the fluctuation is an average measure of a segment and not a point-by-point direct comparison (as done in the cross-validation scheme), it considers noise that can be present in both the fluorescence recordings and the underlying spiking rates. Hence, it fits our data; recall the Clancy et al. (2019) underlying spiking rate is only correlated with the true underlying spiking rate of the recorded fluorescence. Meaning, the recorded spiking rate is a noisy representation of the true underlying rate. Noise from the fluorescence recording or noise from the partial fit of the underlying spiking rate are not distinguishable by our inference. However, we can correctly evaluate the inferred spiking rate while considering both noise resources. In other setups, other noise resources in the underlying spiking rate can stem from a few traces required to represent some average response or comparison across multiple experiments and so on.

6 Discussion

Many algorithms exist to correct and analyze spatial distortions and features in wide-field recordings. They include methods for correcting small motions in the recorded plain or camera (Reddy & Chatterji 1996, for example), denoising (Buchanan et al. 2018, for example), defining regions of interest (Saxena et al. 2020), and more. While they provide useful tools to reduce the noise and analyze the data, they are not capable of, nor aspired to, correct distortions in the temporal domain. In other words, after treating spatial distortions, it remains a challenge to portray the neural activity through time correctly. This is because the fluorescence intensity through time is a transformation, in the temporal domain, of the original neural activity magnitude due to the calcium dynamics. Less than a handful of attempts were made to estimate the correct temporal neural dynamics from the fluorescence (O'Rawe et al. 2023, Wekselblatt et al. 2016, Peters et al. 2021, Stern et al. 2020). Here, we offer pioneering work in which we solve exactly the inference problem of recovering continuously varying spiking rate dynamics from the fluorescence recordings.

We compare our new approach, the Continuously-Varying Algorithm 1, to the Lucy-Richardson algorithm (with its adaptation to the wide-field setting (Wekselblatt et al. 2016); all other inference methods mentioned above retrieve only an estimate (O'Rawe et al. 2023, Peters et al. 2021, Stern et al. 2020)). First, we show that our solution is more accurate than the Lucy-Richardson algorithm in all the scenarios we test (various SNR and underlying fluctuations in synthetic data, Figure 2b and real data, Figure 3bii; in both figures Con-Var reaches lower error values than Lucy). Second, we show that our solution is more consistent compared to the Lucy-Richardson algorithm in the sense that it retrieves the lowest error with the correct underlying calcium decay dynamics, Figure 2b. And third, our solution maintains the initial data structure best post inference Figures 2fiii & 3diii, fluctuation correlations between the inferred spiking rate and fluorescence fluctuations remain similar to those between the underlying spiking rate and fluorescence fluctuations in the data).

Limitations of the solution we bring here, as well as all previous attempts to infer the spiking

rate from wide-field recordings, include their disregard for the calcium rise time. All methods consider only the calcium decay time. Although the rise time is significantly faster, it is not instant. Calcium rise times span from a few single milliseconds to a few dozen milliseconds, depending on the calcium indicator. A simple workaround to this limitation is to use time bins longer than the rise time. Such time bins effectively make the rise time immediate (within the bin). We leave the consideration of the rise time within the inference, and hence the ability to work with even finer time resolution, to future work.

The solution we bring here also requires parameter tuning. While this increases the initial effort in setting up the inference for each setup, it ensures that our Continuously-Varying, exact inference solution is adaptable; it is broadly applicable to many setups (including a variety of animals, brain regions, calcium indicators, and more). We study two different approaches for tuning the penalty, including a novel scheme for the unique case of noisy spiking rate (with respect to the fluorescence or some other defined criterion). We find that after tuning, even when the underlying spiking rate data is unavailable (as in a typical wide-field experiment), the result is satisfying and surpasses other possible non-tuning-requiring approaches like the Lucy-Richardson algorithm. Hence, we find that the possible challenges in tuning are negligible compared to the gains. Our optimization problem approach allows to consider the inference dependency on the calcium dynamics shape, the bounded spiking rate amplitudes, and the noise expected randomness (equation 2). All of it leads to our approach success in minimizing the error in inference and maintaining the structure in the data. An additional advantage of our approach is its ability to maintain highly similar spiking rate solutions when parameters jitter, which happens during experiments, as we discuss in Section 5.3.

We envision our inference solution becoming part of the "pipeline" in all wide-field imaging experiment data analyses. We implement it such that its output, the inferred spiking rate solution, has the same structure as the input, the fluorescence recording. Since the inference is in the temporal domain and the data structure remains untouched, it can be added to the pipeline, in principle, at any stage. Its location in the pipeline will impact noise removal (a shared goal of multiple processing stages). We leave the comparison of the different pipeline results for future studies. We do recommend using our inference solution post spatial distortion corrections and before any dynamics-related analysis, such as dimensional reduction, finding regions of interest, PCA and NMF, mixed linear models, etc. With these steps, the inference is expected to perform best, and to lead to most accurate analysis results overall, since imaging-related distortions would be annihilated before its use, and the dynamical-related analysis will enjoy the dynamics corrected by it.

We make our code available https://github.com/meravstr/Wide-Field-Inference-2024 .

A Proof of Proposition 1

Proof. The result follows by noticing that for the one-to-one mapping $r = D\overline{c}$, the optimization problems (3) and (4) are identical. □

B Proof of Proposition 2

Lemma 1. *Let* $A = PD^{-1}$, $\tilde{y} = P\overline{y}$ *with* $P = I - \frac{1}{T}\mathbb{1}\mathbb{1}^\top$ *and* $\overline{\beta}_0 = \frac{1}{T}\sum_{t=1}^{T}(\overline{y}_t - (D^{-1}r)_t)$. *Then,* $\|\overline{y} - \mathbb{1}\overline{\beta}_0 - D^{-1}r\|^2 = \|\tilde{y} - Ar\|^2$.

Proof of Lemma 1. Since $P^\top P = P$ and $P\mathbb{1} = 0$, it follows that

$$\|\overline{y} - \mathbb{1}\overline{\beta}_0 - D^{-1}r\|^2 = \left(\overline{y} - 1_T\overline{\beta}_0 - D^{-1}r\right)^\top \left(I - \frac{1}{T}\mathbb{1}\mathbb{1}^\top + \frac{1}{T}\mathbb{1}\mathbb{1}^\top\right)\left(\overline{y} - \mathbb{1}\overline{\beta}_0 - D^{-1}r\right)$$
$$= \left(\overline{y} - \mathbb{1}\overline{\beta}_0 - D^{-1}r\right)^\top P^\top P\left(\overline{y} - \mathbb{1}\overline{\beta}_0 - D^{-1}r\right) + \left(\overline{y} - \mathbb{1}\overline{\beta}_0 - D^{-1}r\right)^T \frac{1}{T}\mathbb{1}\mathbb{1}^\top\left(\overline{y} - \mathbb{1}\overline{\beta}_0 - D^{-1}r\right)$$
$$= \|\tilde{y} - Ar\|^2 + \frac{1}{T}\|\mathbb{1}^\top\overline{y} - \mathbb{1}^\top\mathbb{1}\overline{\beta}_0 - \mathbb{1}^\top D^{-1}r\|^2 = \|\tilde{y} - Ar\|^2.$$

The last equality follows from the fact that $\mathbb{1}^\top\overline{y} - \mathbb{1}^\top\mathbb{1}\overline{\beta}_0 - \mathbb{1}^\top D^{-1}r = 0$. □

Proof of Proposition 2. Taking the derivative of (5) with respect to $\overline{\beta}_0$ and setting it equal to zero, we find that $\overline{\beta}_0 = \frac{1}{T}\sum_{t=1}^T(\overline{y}_t - (D^{-1}r)_t)$. The result follows from Lemma 1. □

C Proof of Proposition 3

Lemma 2. *For any $i = 1, \ldots, T$ the following equality holds* $\frac{\sum_{j=1}^T A_{i,j}}{A_{i,1}} = \frac{\sum_{j=1}^T A_{1,j}}{A_{1,1}}$

Proof of Lemma 2. Equation (2) implies that

$$\overline{c}_t = \sum_{t'=1}^{t} \gamma^{t-t'} r_{t'}. \tag{15}$$

Since $\overline{c} = D^{-1}r$, we find from (15) that the entries of D^{-1} are given by

$$\left(D^{-1}\right)_{i,j} = \begin{cases} 0, & \text{if } i < j \\ 1, & \text{if } i = j \\ \gamma^{i-j}, & \text{if } i > j \end{cases}. \tag{16}$$

From (16) we find the sum of the elements in each D^{-1} row i,

$$\sum_{j=1}^{T}\left(D^{-1}\right)_{i,j} = \sum_{j=1}^{T}\gamma^{j-1} = \frac{1-\gamma^i}{1-\gamma}, \tag{17}$$

where the last step was calculated using the expression for the sum of a geometric series. Similarly for the columns,

$$\sum_{j=1}^{T}\left(D^{-1}\right)_{j,i} = \sum_{j=1}^{T-i+1}\gamma^{j-1} = \frac{1-\gamma^{T-i+1}}{1-\gamma}. \tag{18}$$

Recall that $A = \left(I - \frac{1}{T}\mathbb{1}\mathbb{1}^\top\right)D^{-1}$, with I the identity matrix and $\mathbb{1}\mathbb{1}^\top$ a matrix of ones, is the column means subtracted version of D^{-1}. Hence,

$$\sum_{j=1}^{T} A_{i,j} = \sum_{j=1}^{T}\left(D^{-1}\right)_{i,j} - \sum_{j=1}^{T}\frac{1}{T}\sum_{k=1}^{i}\left(D^{-1}\right)_{k,j} = \frac{1-\gamma^i}{1-\gamma} - \sum_{j=1}^{T}\frac{1}{T}\frac{1-\gamma^{T-j+1}}{1-\gamma}, \tag{19}$$

where the last step was calculated using (17) and (18). We evoke once again the expression for the sum of a geometric series, and rearrange the result to receive from the above equation the following,

$$\sum_{j=1}^{T} A_{i,j} = -\frac{\gamma}{(1-\gamma)}\left(\gamma^{i-1} - \frac{1}{T}\frac{1-\gamma^T}{1-\gamma}\right). \tag{20}$$

The structure of A and D^{-1} column sums (18) also gives the following,

$$A_{i,1} = -\left(D^{-1}\right)_{i,1} - \frac{1}{T}\sum_{k=1}^{i}\left(D^{-1}\right)_{k,1} = \gamma^{i-1} - \frac{1}{T}\left(\frac{1-\gamma^T}{1-\gamma}\right). \tag{21}$$

From (21) and (20) we have,

$$\frac{\sum_{j=1}^{T} A_{i,j}}{A_{i,1}} = -\frac{\gamma}{(1-\gamma)}. \tag{22}$$

This implies that $\frac{\sum_{j=1}^{T} A_{i,j}}{A_{i,1}}$ is independent of i, including $i = 1$, and therefore one can write for every i, $\frac{\sum_{j=1}^{T} A_{i,j}}{A_{i,1}} = \frac{\sum_{j=1}^{T} A_{1,j}}{A_{1,1}}$

□

Proof of Proposition 2. Lemma 2 implies that $Ad\mathbb{1}+A\tilde{r}_{m1} = 0$ for any d constant number and $\tilde{r}_{m1}$ a T-length vector with all entries equal zero except for its first entry which satisfies $r_{m1} = -d\frac{\sum_{j=1}^{T} A_{1j}}{A_{11}}$. Therefore,

$$\|\tilde{y} - A\tilde{r}\|^2 = \|\tilde{y} - (A\hat{r} + Ad\mathbb{1} + A\tilde{r}_{m1})\|^2 = \|\tilde{y} - A\hat{r}\|^2. \tag{23}$$

By construction of $\tilde{r}$ and $\tilde{r}_{m1}$, $\tilde{r}$ also satisfies

$$\sum_{t=3}^{T}(\tilde{r}_t - \tilde{r}_{t-1})^2 = \sum_{t=3}^{T}(\hat{r}_t - \hat{r}_{t-1})^2. \tag{24}$$

The convexity of problem (5) ensures that if $\hat{r}$ solves it, and $\tilde{r}$ maintains the value of its objective, as the previous two equations implies, than $\tilde{r}$ is also a solution to this problem.

□

D Simulating spiking rate traces

We generated continuous spiking rate traces by integrating the following neural network $\frac{dx}{dt} = -x + gJ\tanh(x)$ where x is a vector of size $N = 1000$, J is an $N \times N$ matrix with i.i.d. entries $J_{i,j} \sim N(0, N)$, and we varied g between 3.1 and 11.4 for the desired aimed fluctuations span $0.04 \ldots 0.35$ (Figure 1c; $g = 5.7$ was used to generate Figure 1a,b) with 0.13 fluctuations level. We integrated with steps $dt = 0.05$ where each step represents 10ms of neural activity. These choices led to chaotic dynamics (Sompolinsky et al. 1988). To ensure non-negativity of the rate trace, we subtracted out the minimum activity.

E Clancy et al. Supplemental Figure 4 Dataset

Clancy et al. (2019) recorded wide-field fluorescence and, in parallel, spikes originating from the exact physical location as the fluorescence. The recording was done in V1. It includes a single fluorescence trace and a single spike counts trace, each with 130000 measurement points, taken at a 40hz rate (portraying almost an hour of activity). This dataset was obtained to verify that the fluorescence reflects the neural activity (to first order). Clancy et al. (2019) concluded it does so; see Clancy et al. (2019) Supplemental Figure 4c. They have reported to us a calcium decay of $\gamma = 0.973$ (does not appear in their manuscript) by directly recording single neuron fluorescence traces and their spikes, from the same experiment preparation.

Since the calcium indicator rise time is about 40ms, which is longer than the original time bin of the recording (of 25ms), we worked with a two-point average of the data (making the rise time shorter than the 50ms after-average-bin). This also significantly improved the signal-to-noise ratio. Our analyzed signal hence has a 20hz rate. We fitted the decay rate to γ^2 for the algorithms accordingly. We further divided the single long fluorescence and spike count traces into 65 traces each; this way, we generated 65 data "trials," each with 1000 measurement points (50sec long), resembling a possible typical neuroscience experiment.

Recall that recording spike counts originating from the exact location of the fluorescence does not guarantee a true underlying spiking rate of the fluorescence, only an estimate. While Clancy et al. (2019) already verified that this estimate is reliable, Our results also support this conclusion. This can be seen visually in Figures 3ai,ii,ii & 5a where examples of the recorded fluorescence, their inference, and the recorded spiking rate dynamics highly agree. It is also seen in Figure 3di, where fluorescence fluctuations are reasonably correlated with the recorded spiking rate fluctuations, meaning the recorded spikes indeed drive the fluorescence recorded activity.

F Estimating $\overline{\beta}_1$ from the data

To evaluate $\overline{\beta}_1$ we first allow it to take any arbitrary value in equation (2) (and do not set it to be equal to one). We repeat the steps of writing the model in equation (2) in its matrix form and subtracting its parameter means. We receive the following relation (up to noise contributions):

$$\tilde{y} = \overline{\beta}_1 A r. \tag{25}$$

To find $\overline{\beta}_1$ we minimize the squared error of the above relation w.r.t. $\overline{\beta}_1$,

$$\underset{\overline{\beta}_1}{\text{minimize}}\{\|\tilde{y} - \overline{\beta}_1 A r\|^2\} \tag{26}$$

yielding

$$(Ar)^\top \tilde{y} = \overline{\beta}_1 (Ar)^\top Ar, \tag{27}$$

and hence the solution

$$\overline{\beta}_1 = (Ar)^\top \tilde{y} \left((Ar)^\top Ar\right)^+, \tag{28}$$

where $\left((Ar)^\top Ar\right)^+$ is the pseudo inverse of $(Ar)^\top Ar$.

To estimate $\overline{\beta}_1$ from data, we calculate $\overline{\beta}_1$ according to equation (28) using the recorded spiking rate r and the recorded fluorescence $\overline{y}$ (recorded in parallel), with $\tilde{y} = P\overline{y}$, $Ar = PD^{-1}r$, $P = I - \frac{1}{T}\mathbb{1}\mathbb{1}^\top$ and D^{-1} as in equation (16). If multiple segments of recorded data exist from the same experiment or setup (the more the merrier) we average across all $\overline{\beta}_1$ calculated from the different segments, while excluding outliers, to give a single value $\overline{\beta}_1$ estimate.

G Choosing Penalty λ

Adopted from Clancy et al. (2019), Algorithm 3.

Algorithm 2 A cross-validation scheme for choosing λ

Given the fluorescence signal $\overline{y}$, the calcium decay γ, and a set of potential penalties $\lambda_1 \dots \lambda_M$, do:

1. Repeat twice:
   (a) Assign odd timesteps of the signal to the training set, $\overline{y}^{train}$, and even timesteps to the test set, $\overline{y}^{test}$, for the first repeat, and vice versa for the second repeat.
   (b) With each $m = 1 \dots M$ do:
      i. Solve r_m^{train} with Con-Var (Algorithm 1) using $\overline{y}^{train}$, γ^2 and λ_m.
      ii. Convolve $c_m^{train} = D^{-1} r_m^{train}$.
      iii. Average adjacent values in order to obtain predictions on the test set $\hat{c}_m^{test} = \frac{c_{m;1:T/2-1}^{train} + c_{m;2:T/2}^{train}}{2}$
      iv. Calculate the test set MSE for current repeat i, $ycMSE_m^i = \frac{2}{T} \sum_{t=1}^{T/2} (\overline{y}^{test} - \hat{c}^{test})^2$
2. Average the test set MSE over repeats for each m yielding $\overline{ycMSE}_m$.
3. Find $\hat{m} = \text{argmin}_m \{\overline{ycMSE}_m\}$
4. Calculate the standard error of the test set MSE over repeats, for each m, $se(ycMSE)_m = \sqrt{\frac{\left(\overline{ycMSE}_m^1 - \overline{ycMSE}_m\right)^2 + \left(\overline{ycMSE}_m^2 - \overline{ycMSE}_m\right)^2}{2}}$
5. Find $m* = \max\{m : \overline{ycMSE}_m \leq \overline{ycMSE}_{\hat{m}} + se(ycMSE)_{\hat{m}}\}$
6. Return λ_{m*}

References

Berens, P., Freeman, J., Deneux, T., Chenkov, N., McColgan, T., Speiser, A., Macke, J. H., Turaga, S. C., Mineault, P., Rupprecht, P. et al. (2018), 'Community-based benchmarking improves spike rate inference from two-photon calcium imaging data', *PLoS computational biology* **14**(5), e1006157.

Buchanan, E. K., Kinsella, I., Zhou, D., Zhu, R., Zhou, P., Gerhard, F., Ferrante, J., Ma, Y., Kim, S. H., Shaik, M. A. et al. (2018), 'Penalized matrix decomposition for denoising, compression, and improved demixing of functional imaging data', *BioRxiv* p. 334706.

Cardin, J. A., Crair, M. C. & Higley, M. J. (2020), 'Mesoscopic imaging: shining a wide light on large-scale neural dynamics', *Neuron* **108**(1), 33–43.

Chen, T.-W., Wardill, T. J., Sun, Y., Pulver, S. R., Renninger, S. L., Baohan, A., Schreiter, E. R., Kerr, R. A., Orger, M. B., Jayaraman, V. et al. (2013), 'Ultrasensitive fluorescent proteins for imaging neuronal activity', *Nature* **499**(7458), 295.

Clancy, K. B., Orsolic, I. & Mrsic-Flogel, T. D. (2019), 'Locomotion-dependent remapping of distributed cortical networks', *Nature Neuroscience* p. 1.

Friedrich, J., Zhou, P. & Paninski, L. (2017), 'Fast online deconvolution of calcium imaging data', *PLoS Computational Biology* **13**(3), e1005423.

Gilad, A. (2024), 'Wide-field imaging in behaving mice as a tool to study cognitive function', *Neurophotonics* **11**(3), 033404–033404.

Jewell, S., Hocking, T. D., Fearnhead, P. & Witten, D. (2019), 'Fast nonconvex deconvolution of calcium imaging data', *Biostatistics* .
URL: *https://doi.org/10.1093/biostatistics/kxy083*

Jewell, S. & Witten, D. (2018), 'Exact spike train inference via ℓ_0 optimization', *The Annals of Applied Statistics* **12**(4), 2457.

Lucy, L. B. (1974), 'An iterative technique for the rectification of observed distributions', *The Astronomical Journal* **79**, 745.

MacDowell, C. J. & Buschman, T. J. (2020), 'Low-dimensional spatiotemporal dynamics underlie cortex-wide neural activity', *Current Biology* **30**(14), 2665–2680.

Musall, S., Kaufman, M. T., Gluf, S. & Churchland, A. K. (2018), 'Movement-related activity dominates cortex during sensory-guided decision making', *BioRxiv* p. 308288.

O'Rawe, J. F., Zhou, Z., Li, A. J., LaFosse, P. K., Goldbach, H. C. & Histed, M. H. (2023), 'Excitation creates a distributed pattern of cortical suppression due to varied recurrent input', *Neuron* **111**(24), 4086–4101.

Pachitariu, M., Stringer, C. & Harris, K. D. (2018), 'Robustness of spike deconvolution for neuronal calcium imaging', *Journal of Neuroscience* **38**(37), 7976–7985.

Peters, A. J., Fabre, J. M., Steinmetz, N. A., Harris, K. D. & Carandini, M. (2021), 'Striatal activity topographically reflects cortical activity', *Nature* **591**(7850), 420–425.

Pnevmatikakis, E. A., Soudry, D., Gao, Y., Machado, T. A., Merel, J., Pfau, D., Reardon, T., Mu, Y., Lacefield, C., Yang, W. et al. (2016), 'Simultaneous denoising, deconvolution, and demixing of calcium imaging data', *Neuron* **89**(2), 285–299.

Reddy, B. S. & Chatterji, B. N. (1996), 'An fft-based technique for translation, rotation, and scale-invariant image registration', *IEEE transactions on image processing* **5**(8), 1266–1271.

Saxena, S., Kinsella, I., Musall, S., Kim, S. H., Meszaros, J., Thibodeaux, D. N., Kim, C., Cunningham, J., Hillman, E. M., Churchland, A. et al. (2020), 'Localized semi-nonnegative matrix factorization (locanmf) of widefield calcium imaging data', *PLoS computational biology* **16**(4), e1007791.

Silasi, G., Xiao, D., Vanni, M. P., Chen, A. C. & Murphy, T. H. (2016), 'Intact skull chronic windows for mesoscopic wide-field imaging in awake mice', *Journal of Neuroscience Methods* **267**, 141–149.

Sompolinsky, H., Crisanti, A. & Sommers, H.-J. (1988), 'Chaos in random neural networks', *Physical Review Letters* **61**(3), 259.

Stern, M., Shea-Brown, E. & Witten, D. (2020), 'Inferring the spiking rate of a population of neurons from wide-field calcium imaging', *BioRxiv* pp. 2020–02.

Vogelstein, J. T., Watson, B. O., Packer, A. M., Yuste, R., Jedynak, B. & Paninski, L. (2009), 'Spike inference from calcium imaging using sequential monte carlo methods', *Biophysical Journal* **97**(2), 636–655.

Wekselblatt, J. B., Flister, E. D., Piscopo, D. M. & Niell, C. M. (2016), 'Large-scale imaging of cortical dynamics during sensory perception and behavior', *Journal of neurophysiology* **115**(6), 2852–2866.

Zhang, Y., Rózsa, M., Liang, Y., Bushey, D., Wei, Z., Zheng, J., Reep, D., Broussard, G. J., Tsang, A., Tsegaye, G. et al. (2023), 'Fast and sensitive gcamp calcium indicators for imaging neural populations', *Nature* **615**(7954), 884–891.