An Analytic Solution for Spiking Rate Inference

Merav Stern[1,*]

[1]The Center for Theoretical Studies in Physics and Biology, The Rockefeller University, New York, NY, 10065, USA
[*]mstern@rockefeller.edu

Abstract

Many cognitive functions involve multiple brain areas that simultaneously process, distribute, and share information. Adequately capturing such distributed brain-wide activity can be achieved through wide-field imaging techniques, which enable the simultaneous recording of brain activity from a wide field of view at a high rate. However, the wide field of view imposes limitations on the spatial resolution. As a result, each fluorescence trace captured by each camera pixel in this wide-field setup reflects the combined calcium-generated fluorescence of many neurons' activities. Additionally, calcium indicators, which convert neural activity into light emissions, distort the neural activity by their dynamics. The inherent noise in recordings, combined with the low spatial resolution and the distorted dynamics by the calcium indicators, makes it a particularly challenging mathematical problem to infer underlying neural activity from recorded fluorescence in wide-field imaging. To date, there has not been a rigorously studied analytic solution for this inference problem in the wide-field setting. In this work, we phrase the inference problem that arises from wide-field recordings as an optimization problem and provide an analytic solution to it. To ensure the robustness of our findings and establish a solid foundation for application, we rigorously verify our solution using real data. Furthermore, we propose a novel approach for the optimization problem parameter-tuning. Beyond recovering the neural dynamics, our inference method will enable future studies to conduct more accurate, correlation-based analyses of brain-wide activity.

1 Introduction

Optogenetic tools utilize calcium indicators to bind neuron-initiated spikes with fluorescence emission. The resulting fluorescence can then be captured by a camera, making neural activity visible and hence easier to detect.

In the wide-field setting, a vast amount of fluorescence emitted from neural activity is captured by a standard one-photon camera across a large field of view (Cardin et al. 2020). A typical example of such a wide-field view is the entire dorsal surface area of a mouse brain (Silasi et al. 2016). The wide-field setting also enables fast, simultaneous recordings across the whole surface (Gilad 2024). Consequently, wide-field recordings can reliably depict the time evolution of fluorescence signals across all recorded brain areas. However, due to the large field of view, the resolution is limited; each pixel captures a light trace that reflects the combined activity of thousands of neurons. In turn, the vast amount of fluorescence captured in each pixel allows the recordings to be made through the skull of animals such as mice (Silasi et al. 2016). This technique leaves the brain

intact and makes wide-field imaging minimally invasive. Because of its low invasiveness, wide-field imaging can easily be used to record awake-behaving animals across multiple sessions over several days, enabling to relate and compare recordings from all sessions (Musall et al. 2018, MacDowell & Buschman 2020). The long list of advantages of the wide-field method contributes to its popularity as a recording technique, despite its limited spatial resolution.

Inferring the total neural activity contributing to each recorded fluorescence trace in the wide-field setting is a significant mathematical challenge. The limited spatial resolution, a distinct characteristic of the wide-field method, combined with the presence of noise, necessitates a specialized solution to extract such underlying total neural activity. Many methods 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 demonstrate in this work, the mathematical challenge that stems from the need to infer the total spiking activity of many neurons is unique and solutions designed for the single-neuron scenario do not hold, requiring the development of a new tailored inference solution.

Other studies in neuroscience propose inference to find the average spiking rate of a single neuron following repeated stimuli (Ganmor et al. 2016) or to analyze the distribution of spiking rates (assuming that the neural temporal activity has already been deconvolved) (Wei et al. 2020). However, similar to inference methods for identifying single-neuron spike times, these approaches do not account for the characteristics of fluorescence dynamics driven by multiple spiking neurons, leaving the need for inference techniques suitable for fluorescence data recorded in wide field settings in place.

Beyond neuroscience, inference is a common challenge across many fields, and numerous mathematical methods have been developed (Chaudhuri et al. 2014), either to refine existing approaches or to address new problems (Engl & Groetsch 1987, Vogel 2002). A prominent example is image processing (Wipf & Zhang 2014), where issues such as blurs, distortions, reflections, or missing data require corrections achieved through inference, typically in the form of an iterative algorithm (Bertero et al. 2021). Each type of degradation necessitates a method tailored to its specific features and constraints, often resulting in distinct mathematical formulations (Hansen 2010). Even when methods appear similar, practical solutions can differ substantially because of these constraints. The same holds for temporal inference of spiking activity. In this work, we demonstrate that the characteristics of activity recorded in wide-field imaging differ from single-neuron dynamics. By utilizing these differences and while still adopting an optimization framework from the single neuron case, we show that the problem can be solved analytically under wide-field constraints. Although deriving this solution is nontrivial, it provides important advantages: analytic results are comprehensive, computationally efficient, and yield exact global solutions.

The relation between the neural activity and the dynamics of the calcium indicator, which is responsible for the recorded fluorescence, is known in calcium imaging (Zhang et al. 2023, Chen et al. 2013, for example). Nevertheless, since calcium levels integrate a neuron's spiking activity, recovering the neural spiking activity from fluorescence is fundamentally a derivative operation. The presence of noise makes such a direct inverse, derivative-based operation impractical, as it amplifies the noise to a level that obscures the underlying neural activity. Hence, naive approaches for inferring neural activity from fluorescence recordings by directly inverting calcium dynamics are ineffective and require more in-depth mathematical solutions.

To a first approximation, the calcium concentration depicted by fluorescence reflects the volume of recent neural activity. However, since neural activity contains a wealth of information about brain functions, it is frequently the detailed information we seek. Therefore, accurately inferring temporal changes in neural activity beyond first-order measurements is important.

But even after making inferences in a wide-field setting, information about neural activity remains limited. Specifically, the absolute count of spikes per second and the total number of neurons responsible for these spikes remain unknown. This limitation occurs because the recorded fluorescence is measured in arbitrary units that do not reflect the volume of brain activity contributing to the fluorescence signal. Furthermore, measurements in the wide-field setting lack known amplitude spiking events for comparison with the fluorescence (unlike the single spikes observed in two-photon recordings).

By using inference methods that leverage calcium indicator dynamics, relative changes in neural activity can be retrieved from wide-field fluorescence data. Inferred fluorescence thus provides a more accurate representation of the temporal evolution of neural activity than raw fluorescence signals alone. Moreover, inference reduces noise based on biologically relevant criteria related to calcium-binding properties, rather than using generic image-processing filters that operate on visibility criteria. It is therefore possible that inference reduces the risk of misclassifying neural activity as noise (or vice versa) and better preserves the integrity of the spiking rate information embedded in the fluorescence data.

Identifying the correct time evolution of neural dynamics via inference can also improve the accuracy of results when conducting more advanced analyses of the data. Such analyses include comparing two brain areas by examining how their neural activities correlate with each other or with behavioral measures. Other correlation-based techniques, such as Principal Component Analysis (PCA) or Non-negative Matrix Factorization (NMF), can also benefit from corrected temporal dynamics achieved through inference. Since these techniques are based on correlations, their results are invariant under shifts in magnitude, meaning they do not rely on the absolute volume of spiking rate activity. Instead, they are sensitive to the timing of relative changes between variables, whether those variables are neural spiking rates or behavioral measures. This further emphasizes the need to determine the dynamics of neural activity over time accurately. In addition, it renders the wide-field techniques' inability to measure the absolute magnitude of neural activity a less significant drawback.

The main contribution of this paper is pioneering work that formulates and addresses the inference problem arising from wide-field recordings. We provide an analytic solution to the inference problem we phrase. To the best of our knowledge, only a few previous attempts have been made to infer multiple spiking neurons' activity from their generated fluorescence (Wekselblatt et al. 2016, O'Rawe et al. 2023, Stern et al. 2020). Among these, only one adapted a method to calculate (rather than estimate), in a timely manner, a continuous inferred spiking-rate signal (Wekselblatt et al. 2016). We evaluate our solution against the method of Wekselblatt et al. (2016) and a commonly used, Gaussian process deconvolution approach (Tobar et al. 2023), which we adapt here for temporal inference of spiking rates. For completeness, we also compare all methods to the naive approach of using the fluorescence signal as an estimate of neural activity. We use three datasets for testing. The first consists of fluorescence generated from simulated neural spiking rates using our model. The second dataset is a biophysically detailed volumetric model in which neurons have realistic morphologies. Calcium concentration is calculated based on its molecular properties, and fluorescence simulations account for how light propagates through tissue (Song et al. 2021). This dataset enables us to test inference performances in a realistic yet fully known setting. Third, we test on real wide-field recordings, demonstrating that our solution is practical and applicable to experimental data. Finally, we introduce a novel approach for parameter tuning. Wide-field recordings provide vast amounts of data, which we exploit for comparing the dynamical properties of the recorded signals with those of the inferred signals obtained using different parameters, thereby identifying optimal parameters. To our knowledge, this represents a

unique parameter-tuning strategy for inference problems.

This paper hence includes the following content: First, we phrase the mathematical challenge of inferring the total spiking rate generated by multiple neurons based on their collective fluorescence trace. We then solve the problem and provide proof of our solution. We compare the performance of our solution to commonly used inference methods, including those previously applied to wide-field imaging. We do so by using synthetic self-generated data, a biophysically detailed dataset, and real recorded data that includes fluorescence (recorded using the wide-field method) alongside spike count recordings. Last, we consider a typical experiment where only fluorescence is recorded and discuss our original parameter tuning analysis with respect to this scenario.

2 Model

We begin by examining how the spiking of an individual neuron contributes to a recorded fluorescence trace. Several recent studies have addressed this problem (Jewell & Witten 2018, Jewell et al. 2019, Friedrich et al. 2017, Pnevmatikakis et al. 2016). These studies have established a foundation for inferring spike times from fluorescence traces associated with a single neuron's activity (typically recorded in two-photon cameras; as originally phrased by (Vogelstein et al. 2009)). These studies utilized 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 model described above, the observed fluorescence of a neuron i at timepoint t , denoted as y_t^i, is determined by the neuron's calcium concentration at that timepoint, c_t^i up to some factor β_1, a constant shift, β_0, and noise ϵ_t^i, which is assumed to be Gaussian and i.i.d.. In other words, the fluorescence is a noisy linear readout of the calcium concentration. The calcium concentration at each timepoint, c_t^i, is the result of its decayed value from the previous timepoint, c_{t-1}^i, by a factor $\gamma \in (0, 1)$ and the number of spikes emitted at that time s_t^i. In other words, the calcium concentration decays exponentially over time according to its decay constant and "jumps" with each spike emission.

We model the dynamics of a fluorescence trace resulting from the activity of many neurons by summing the equations above over all these neurons. This summation relies on the assumption that the camera remains within its linear dynamic range, allowing the incoming light to affect the recorded signal in a linear manner. Hence, the total fluorescence is the linear sum of the fluorescence of the individual neurons. We also assume that all neurons exhibit the same calcium indicator dynamics, characterized by a similar decaying constant, γ, with similar relations to fluorescence emission. We hence receive the following auto-regression model for a fluorescence trace generated by the activity of many neurons:

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

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 $\overline{\sigma}^2 = \sum_{i=1}^N (\sigma^i)^2$ 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 at time t, by all N neurons (whose activity is contributing to the recorded fluorescence). Introducing the spiking rate into the model marks a significant change from the single-neuron model (1). With a single neuron, the number of spikes at each time point s_t^i is a sparse quantity, with many zero entries and occasional small single-digit values. On the other hand, with many neurons, we expect the spiking rate to be a larger number that changes gradually over time. Since both the number of spikes in the single-neuron model, s_t^i, and the spiking rate in the multi-neuron model, r_t, are at the heart of the solution we are pursuing, their different properties significantly influence the approach we take to the inference solution. We discuss this further as we phrase the optimization problem for our model.

It is likely that not all neurons contribute equally to the recorded fluorescence. In this case, each neuron i has its own relative contribution, which we denote as α^i. Consequently, the total fluorescence is given by $\overline{y}_t \equiv \sum_{i=1}^{N} \alpha^i y_t^i$. And we can calculate its dynamics by summing the dynamics of all individual neurons contributing to it, while weighting each neuron's dynamics (1) by its corresponding contribution α^i. This leads to the multi-neuron model in equation (2), where the spiking rate is given by $r_t = \sum_{i=1}^{N} \alpha^i s_t^i$. In this case, all other quantities are also expressed in terms of each neuron's relative contribution. The total baseline fluorescence is given by $\overline{\beta}_0 \equiv \sum_{i=1}^{N} \alpha^i \beta_0^i$, the total calcium at timepoint t is $\overline{c}_t \equiv \sum_{i=1}^{N} \alpha^i c_t^i$, and the total noise is $\overline{\epsilon}_t \equiv \sum_{i=1}^{N} \alpha^i \epsilon_t^i$ with its total variance $\overline{\sigma}^2 = \sum_{i=1}^{N} (\alpha^i \sigma^i)^2$. A key point relevant to this case is that the weighted rate remains a positive number that gradually changes over time, and its relation to the total fluorescence is still described by equation (2). We will further explore it as we examine the detailed biophysical Naomi dataset in Section 5.3.

It is worth considering multiplying equation (2) by a factor of $1/N$, which is equivalent to rescaling each α^i by $1/N$. In the special case where all neurons contribute equally, this corresponds to setting $\alpha^i = 1/N$ for all $i = 1, \ldots, N$, so that $\sum_{i=1}^{N} \alpha^i = 1$. Under this convention, the spiking rate and all other variables in equation (2) represent averages (weighted averages when the α's are unequal) rather than sums. This change is straightforward and constructive. Since the number of neurons contributing to a single fluorescence trace in wide-field recordings is very large, we can treat the (weighted) average spiking rate (which is equal to the average number of spikes per neuron in a time bin) as a continuous number. This perspective will be useful in Section 3 when we choose a penalty for the optimization problem.

We now choose $\overline{\beta}_1 = 1$ without loss of generality. This corresponds to scaling the calcium by a constant factor, thereby setting a fixed ratio between the spiking rate and the fluorescence. This choice is relevant when analyzing fluorescence and spiking rate recorded simultaneously, as discussed in Section 5.4. In this section, we will explain how to scale the fluorescence in accordance with this choice.

3 Optimization

We aim to determine the underlying spiking rate r_t from a fluorescence trace y_t recorded in a wide-field setting, utilizing our model (2). We adopt (and adapt) the approach of formulating and solving an optimization problem to find spikes from a fluorescence trace generated by a single neuron according to model (1) (Jewell & Witten 2018, Friedrich et al. 2017). Since both models, (2) and (1), share similar structures, the optimization problem for single neurons can be naturally applied to our model (2) with its corresponding parameters. This leads to the minimization

problem that we wish to solve:

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

where λ is a weight tuning parameter for the penalty P.

For convenience, we rephrase (3) in terms of the spiking rate r_t. To this end, we define $r_1 \equiv \overline{c}_1$. We emphasize that this is merely a definition; the spiking rate at $t = 1$ cannot be recovered, and hence r_1 remains biologically meaningless. Using this definition and our model (2), we can express the relation between the calcium concentration and the spiking rate in a matrix form as $r = D\overline{c}$, where D is a $T \times T$ full-rank matrix with 1's along the diagonal and $-\gamma$'s just below the diagonal. This allows us to rephrase the optimization problem in (3) as follows:

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

where $\mathbb{1}^\top$ is a column vector of length T with all its 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 Supporting Information A.

At this point, we need to determine the form of the penalty $P(r_2, \ldots, r_T)$. This choice will dictate the properties of the solution r that we derive. Hence, to best fit it for wide-field imaging, we need to consider the characteristics of the underlying spiking rate r in this context. The penalty that was chosen for a fluorescence trace resulting from the activity of a single neuron is no longer relevant. This is because the properties of the number of spikes of a single neuron differ significantly from the properties of the spiking rate of many neurons. The typical profile for the number of spikes per bin in a single neuron consists of many zero entries and otherwise features low positive numbers. To obtain solutions with such characteristics, ℓ_0 penalty has been chosen (Jewell & Witten 2018, Jewell et al. 2019) or ℓ_1 (Vogelstein et al. 2009, Pnevmatikakis et al. 2016). In our case, as discussed in Section 2, we can treat the spiking rate of many neurons as a continuous positive number. We also expect it to change gradually over time while remaining bounded (see Figure 4(a)i, ii, for example). To guarantee solutions for the spiking rate, r, that exhibit these desired properties, we choose the ℓ_2 penalty through time $P\left((r_3 - r_2)^2, (r_4 - r_3)^2, \ldots, (r_T - r_{T-1})^2\right)$. This penalty will also allow us to find and analyze an analytic solution for r.

We hence obtain the following optimization problem for inferring spiking rates from wide-field fluorescence 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) appears to be a direct sum of two ℓ_2 terms, which suggests it can be solved naively by setting its derivatives with respect to β_0 and r to zero and

finding their minimum (similar to the Least Mean Squares method). However, a more careful examination reveals that the solution for r must be non-negative. Moreover, with T fluorescence measurements, only $T-1$ time points for spiking rates can be determined, since they depend on changes in fluorescence through calcium dynamics (remember that r_1 is merely a notation and does not represent a spiking rate). In other words, we have one degree of freedom, which our solution must "absorb" while preserving its structure to yield a biologically meaningful spiking rate solution. In this section, we prove that this degree of freedom corresponds to an arbitrary constant shift in the fluorescence (a consequence of the $\Delta F/F$ procedure), which results in the spiking-rate solution also being invariant under such constant shift. We use this degree of freedom to ensure that the spiking rate solution is non-negative, as required.

We begin 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 Supporting Information 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 that it makes the spiking rate invariant under a constant shift.

We now solve the optimization problem (5) to infer the spiking rate from fluorescence recordings in a wide-field setting. According to Proposition (2), we can solve (5) by simply solving (6) to obtain the spiking rate, and then obtain β_0 from its closed-form expression. Finding β_0 enables us to reconstruct the noise-removed version of the original recorded fluorescence, along with its shift, which is helpful for testing the results.

To solve (6), it is constructive to present the following, simpler optimization problem that retains the same ℓ_2 expressions but does not include any constraints on r:

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

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:

$$Z_{i,j} = \begin{cases} -1, & \text{if } |i-j| = 1 \text{ and } i+j > 3 \\ 2, & \text{if } i = j \text{ and } 2 < i < T \\ 1, & \text{if } i = j = 2 \text{ or } i = j = T \\ 0, & \text{otherwise} \end{cases}. \tag{8}$$

Note that $\sum_{t=3}^{T} (r_t - r_{t-1})^2 = r^\top Z r$. Therefore, f takes the form $f(r) = \|\tilde{y} - Ar\|^2 + \lambda r^\top Z r$ and $\nabla_r f(r) = 0$ yields the following:

$$-2A^\top(\tilde{y} - Ar) + 2\lambda Z r = 0. \tag{9}$$

The solution to (9), and hence to problem (7), takes the form

$$r = \left(A^\top A + \lambda Z\right)^+ A^\top \tilde{y} \tag{10}$$

where $\left(A^\top A + \lambda Z\right)^+$ is the pseudo-inverse of $A^\top A + \lambda Z$. Note that we have absorbed into this pseudo-inverse the degree of freedom that led us to use the mean subtracted version of the fluorescence, $\tilde{y}$, and the column-wise mean subtracted version of $D^{-1}r$, Ar. Next, we find the explicit form of this degree of freedom.

Proposition 3. *Let $\hat{r}$ denote a solution to (7). Then, for any constant number d, $\tilde{r} = \hat{r} + d\mathbb{1} + \tilde{r}_{m1}$ also solves (7) 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*

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

We prove proposition 3 in Supporting Information C. This proposition implies that problem (7) (and consequently the original problem (5)) is not strictly convex. It has multiple solutions that differ from each other by a constant shift in the spiking rate $r_2, \ldots, r_T$ and an adjustment, compensating for that shift, in r_1 (recall that r_1 has no biological meaning). As a result, the solution for β_0 changes accordingly to accurately relate the shifted spiking rate (through the calcium concentration) to the recorded fluorescence. This concludes our proof that the degree of freedom in the problem translates to the spiking rate solution being invariant under constant shifts.

Solution (10) to problem (7) 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 Variating Inference Method (Con-Var), as described in Algorithm 1. To this end, we simulated spiking rate traces (see example in Figure

1ai, ivory line; see Supporting Information D for details) and emulated the wide-field recording of fluorescence traces resulting from these spiking rates (following our model (2)). This process included transforming the simulated spiking rate traces into calcium traces according to $c = rD^{-1}$ (a direct result of (2), see example in Figure 1(a)ii) and adding Gaussian-independent noise and constant shift to the calcium traces to generate the recorded fluorescence (Figure 1(a)iii). For the calcium indicator dynamics, we selected $\gamma = 0.95$, which closely matches its observed values (see Section 5.4).

We used Con-Var to infer the spiking rates from the simulated fluorescence recordings. To assess the quality of the inference, we calculated the following error:

$$\mathrm{err}(x, x^{est}) = \frac{1}{T-1} \sum_{t=2}^{T} |x_t - x_t^{est}|, \tag{12}$$

defined as 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 assess the error (12). We chose this approach because the inferred spiking rate is invariant under a constant shift, as implied by Proposition 3.

The strength of fluctuations in a trace offers another useful measure for evaluating the quality of our inference. We quantify these fluctuations as follows:

$$\zeta(x) = \sum_{t=3}^{T} (x_t - x_{t-1})^2, \tag{13}$$

where x can be any time-dependent parameter (spiking rate r, calcium $\overline{c}$ or fluorescence $\overline{y}$). The fluctuations measure the extent and frequency with which the signal changes over time. Note that for $x = r$, the expression in equation (13) is equal to the penalty term in the optimization problem (5). Since the parameter λ controls the weight of this penalty, lowering λ weakens the penalty, allowing larger fluctuations in the inferred spiking rate, while increasing λ strengthens the penalty and suppresses fluctuations. Figure 1(b)i was generated by varying the penalty weight λ and the calcium decay constant, γ used for inference. The results were plotted for each value of γ according to the fluctuations observed with each λ (on the x-axis, with higher λ promoting lower fluctuations and vice versa).

For each calcium decay constant γ used for inference, there is an optimal level of fluctuations corresponding to an optimal λ that produces a solution with minimal error. Con-Var preserves this optimal level of fluctuations close to its value at the global minimum (lowest error) across a wide range of γ (see Figure 1(b)i; the x values of the minimum error points are similar across the different γ curves). This finding suggests that slight variations in the calcium decay constant γ are unlikely to have a significant impact on the optimal inference solution. These small jitters in γ's can occur due to several factors, including the intrinsic properties of neurons, high tissue density between neurons, and substantial changes in firing rates, among others. Our results indicate that, despite these jitters, the optimal inferred solution is maintained with a high degree of accuracy.

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

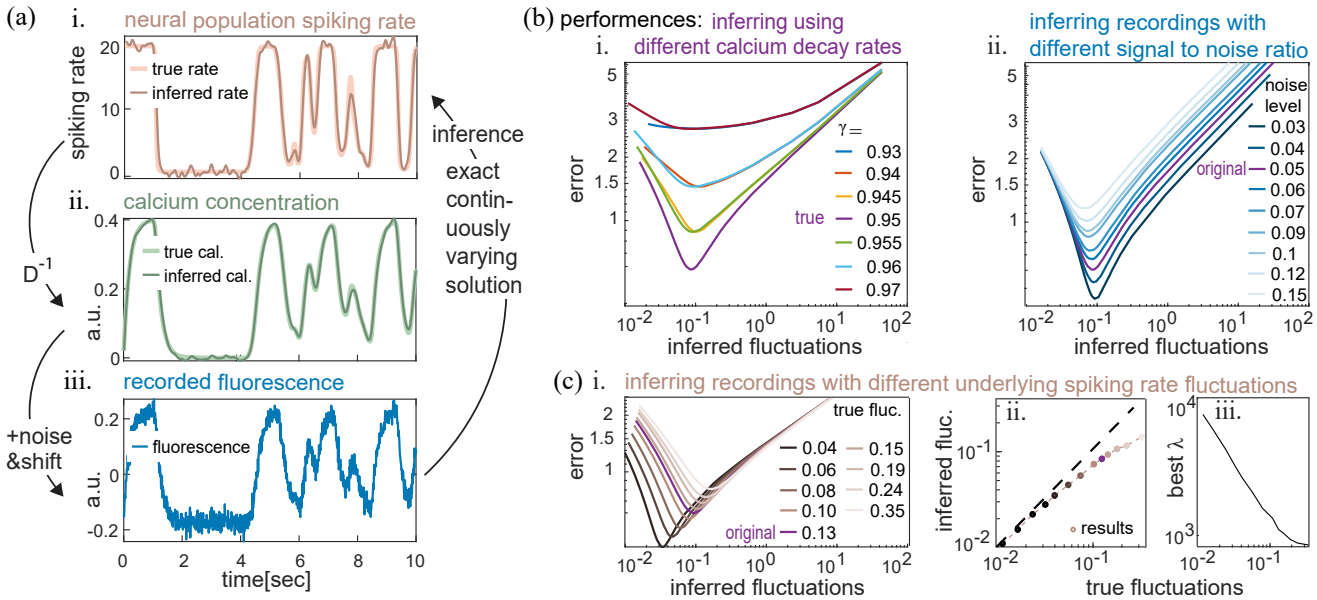

Figure 1: Evaluating the Continuously-Varying Inference, Algorithm 1 (Con-Var). (a). An example: A simulated underlying spiking rate (i. ivory line) is convolved into calcium concentration (ii. light blue line) and transformed into recorded fluorescence (iii. blue line) using our model (2). The fluorescence is then inferred back into spiking rate using Con-Var (i. brown line). Note that the fluorescence units are scaled by 1/1000 to match typical camera settings. (b) i. The mean error (12) between simulated and inferred spiking rates as a function of the mean inferred spiking rate fluctuations (13). The inferred spiking rates were obtained by applying Con-Var to fluorescence traces generated from simulated spiking rates (as in (a)). For inference, Con-Var was applied multiple times with different values of γ ($0.93 \leq \gamma \leq 0.97$, yielding multiple lines) and λ ($10000 \leq \lambda \leq 1$, yielding variations in the inferred spiking rate fluctuations). Means were taken over 10000 traces, each 10sec long, generated as described in Supporting Information D. Calcium traces were generated with $\gamma = 0.95$, and for the fluorescence we applied $\overline{\sigma} = 0.05|\max\overline{y} - \min\overline{y}|$. ii. The mean error (12) between simulated and inferred spiking rates as a function of the mean inferred rate fluctuations (13) for different datasets with different noise levels ($0.03|\max\overline{y} - \min\overline{y}| \leq \overline{\sigma} \leq 0.15|\max\overline{y} - \min\overline{y}|$). γ was held fixed at 0.95 for inference; all other parameters are as in i. (c). i. The mean error (12) between simulated and inferred spiking rates as a function of the mean inferred spiking rate fluctuations (13) for multiple simulated spiking rate sets with different mean (true) fluctuations (see Supporting Information D). γ was held fixed at 0.95 for inference; all other parameters are as in (b)i. ii. Mean inferred spiking rate fluctuations as a function of the mean simulated (true) spiking rate fluctuations at the minimum error point identified for each true spiking rate fluctuations in i. iii. Values of best penalty λ which yield the minimum error points in i., as a function of the true spiking rate fluctuations. The variation in the optimal penalty value allows for fitting the inferred spiking rate fluctuations to the true spiking rate fluctuations (shown in ii.) while maintaining a similar minimal error (in i).

The signal-to-noise ratio (SNR) can vary significantly between experiments. Therefore, it is essential to test the properties of the solution 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 minimum error of the optimal inferred solution increases as SNR decreases (witn an incresead noise variance), the fluctuation levels remain similar across many noise levels (see Figure 1(b)ii; SNR values range from 1/0.3 to 1/0.15; over this range the minimum error doubles, and the fluctuations increase by about 1.5-fold, despite a five-fold increase in noise variance).

The above suggests that inferred solutions retain their properties amid possible noise variability. This robustness arises because Con-Var estimates the noise at each time point in the fluorescence recordings and removes it to derive the inferred spiking rate. Hence, noise reduction is a consequential outcome of the inference. Specifically, Con-Var represents fluorescence as a sum of a calcium signal and noise. It fits the calcium component according to a biologically meaningful criterion, the calcium dynamics, while the remainder of the fluorescence signal is identified as noise and removed. This procedure allows for the extraction of the calcium signal and, in turn, the inferred spiking rate. As a result, Con-Var solutions are spiking rates with reduced measurement noise and tend to be more robust because they are derived with respect to a biologically relevant criterion.

Inferred solutions by Con-Var exhibit fluctuations that closely resemble the underlying spiking rate fluctuations (Figure 1(c)ii). This similarity is conserved across an extensive range of fluctuation levels. Con-Var achieves this with an adjusted optimal penalty λ, which leads to the optimal solution (Figure 1(c)i,iii). In other words, the inferred spiking rate fluctuations correctly reflect the underlying spiking rate fluctuations. This enables a proper fit of the inference across various recording scenarios, accommodating a broad spectrum of underlying spiking rate fluctuations.

5.2 Comparison

Here, we compare the performances of the Continuously-Varying Inference (Con-Var), as detailed in Algorithm 1, with those of the Lucy-Richardson Algorithm (Lucy) and the Gaussian Process Deconvolution (GPDC). To provide a baseline for these comparisons, we also assess how Con-Var, Lucy, and GPDC perform against simply using the fluorescence signal itself as an inferred spiking rate.

Originally, the Lucy algorithm was developed for astrophysics to infer underlying light sources that have been distorted by lenses (Lucy 1974). Recently, it has been adapted to infer spiking rates from fluorescence traces that have been distorted by calcium dynamics in wide-field imaging experiments (Wekselblatt et al. 2016). At their core, the modifications rely on treating the calcium dynamics as the relevant filter (a temporal filter in this case rather than a spatial filter in the original scenario). This filter can be expressed as γ^t for $t = 1, \ldots, p$ where p is sufficiently large to capture the calcium signal decay. To the best of our knowledge, this is the only algorithm in the published literature that has been used and tested specifically for inferring continuous spiking rate signals from wide-field imaging.

A Gaussian process is a collection of variables, often representing consecutive measurements, that jointly follow a multivariate Gaussian distribution. As a mathematical framework, Gaussian processes offer a useful tool for modeling and inference, commonly applied to estimate missing data and reduce noise in fields such as signal processing and image reconstruction (Williams & Rasmussen 2006) In this work, we adapt Gaussian Process Deconvolution (GPDC) (Tobar et al. 2023) to infer spiking rates from wide field fluorescence recordings. The Gaussian Process Deconvolution assumes that the underlying signal, which in this case is the spiking rate, is a Gaussian process that has been convolved with a filter h. This filter represents here the calcium

dynamics, resulting in the observed fluorescence readout. We fit the filter h to match the one used in the adapted Lucy algorithm. Gaussian Process Deconvolution serves as a complementary method for inferring spiking rates; it is based on nonparametric Bayesian methods, whereas the Lucy relies on iterative algorithm, and Con-Var employs an analytical optimization approach.

The main advantage of utilizing Lucy for wide-field imaging is its easy and fast implementation. In contrast to our Con-Var solution, Lucy does not require parameter tuning. While its number of iterations can be adjusted, our results indicate that the default of 10 iterations consistently performs best across all datasets tested here. Lucy also benefits from pre-written and optimized code in most programming languages. Our Con-Var solution is analytic (yet involves a structured matrix inversion), making it comparable in efficiency to Lucy. GPDC, on the other hand, is computationally demanding and requires specifying the SNR in the dataset, which is available in simulations but typically unknown for real data.

The main drawback of both Lucy and GPDC is that, at their core, they rely on assumptions that do not necessarily match the characteristics of fluorescence dynamics generated by neural population activity. For instance, Lucy assumes a Poisson noise distribution, which may not fit the noise properties in fluorescence recordings. GPDC, in turn, places a stationary Gaussian process prior, which can be over-constraining. Our tests, which compare the error between the inferred and the underlying spiking rate, reveal that these discrepancies lead to various issues. Most notably, GPDC over-smooths and suppresses fast transients in the fluorescence data, leading to an error comparable to that achieved when using the fluorescence naively as the inferred spiking rate. For Lucy, the most notable issue is the loss of correlation between the fluctuations of the inferred and the underlying spiking rates, as described below.

To test Lucy and GPDC and compare them to Con-Var, we inferred the same fluorescence traces that we generated in Section 5.1, this time using Lucy and GPDC.

It is evident that the inferred spiking rate by Lucy does not capture the full span of the underlying spiking rate (Figure 2(a)iii, brown vs. ivory lines). This limitation is not critical, as both fluorescence and inferred rates are in arbitrary units and thus carry no direct biological meaning. Another possible weakness of Lucy is its inconsistency in estimating the calcium decay rate. It achieves minimal error with a decay constant γ close to, but not equal to, the true value of γ, leading to an incorrect inference of γ (Figure 2(b), true γ is marked by a purple dashed line).

The spiking rate inferred by Lucy is evidently also noisier compared to the underlying spiking rate (see Figure 2(a)iii brown vs. ivory lines), whereas inference by GPDC is evidently smoother (see Figure 2(a)v brown vs. ivory lines, and the discussion above). Inferred spiking rates by Con-Var take a balanced estimate, capturing the changes of the underlying rate more accurately (see Figure 2(a)i brown vs. ivory lines). These discrepancies, by Lucy and GPDC, create local dissimilarities between the underlying and the inferred spiking rates. At many timepoints, Lucy introduces changes in the inferred spiking rate when the underlying spiking rate is constant. GPDC introduces a constant or a slowly changing inferred spiking rate when the underlying spiking rate is more rapidly changing. As a result, the fluctuations inferred by Lucy remain consistently higher than those of the underlying rate, whereas the fluctuations inferred by GPDC remain consistently lower across a wide range of underlying fluctuation levels (Figure 2(e)ii). The inability of GPDC to follow changes in the underlying spiking rates results in error levels comparable to those obtained by using the fluorescence as the inferred spiking rate (Figure 2(b), green and black dashed lines, respectively). Lucy's consistent high fluctuations in its inference solutions suggest that it may be suitable for inferring fluorescence recordings generated by higher fluctuating underlying spiking rates.

The difficulty Lucy and GPDC have in fitting the fluctuations manifests both on average (Figure

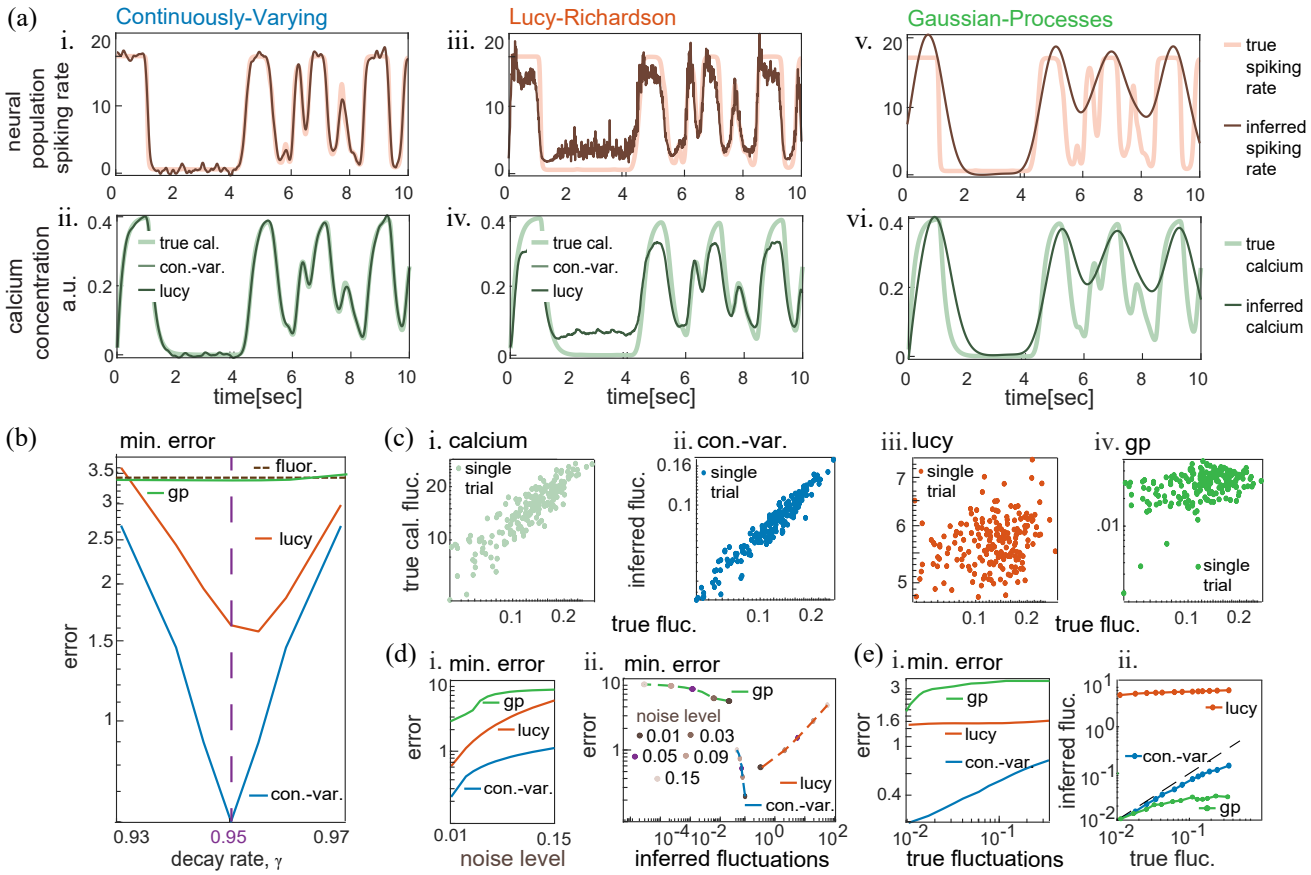

Figure 2: Comparing the Lucy-Richardson Algorithm (Lucy) and the Gaussian Process Deconvolution (GPDC) to the Continuously-Varying Inference, Algorithm 1 (Con-Var). (a) The same example as in Figure 1(a), including additional inference by Lucy and GPDC. (b) i. The mean error (12) between simulated (true underlying) spiking rates and inferred spiking rates as a function of the different γ's used for inference by Con-Var, Lucy and GPDC. For comparison, we also included the error when fluorescence (fluor) is naively considered as the inferred spiking rate. The dataset used in Figure 1(b)i was also utilized here. Similarly, the datasets in Figure 1(b)ii and Figure 1(c) where employed for (d), and (e), respectively. Con-Var results are represented by the minimum error achieved with each γ (minimum error of each line in Figure 1(b)i). The same applies to (d) and (e)i, where Con-Var minimum errors were derived from Figure 1(b)ii and Figure 1(c)i, respectively. (c) For 200 example traces, the relation between the following is presented: the true underlying spiking-rate fluctuation levels and i. the true calcium fluctuation levels, ii.-iv. the fluctuation levels of the spiking rate inferred by ii. Con-Var, iii. Lucy, iv. GPDC. The corresponding Pearson correlations (for the displayed examples) are: i. 0.88, ii. 0.91, iii. 0.55, iv. 0.35. Examples were generated using the best γ and λ (when applicable) for each inference method. (d) The mean error between simulated (true underlying) and inferred spiking rates as a function of: i. The noise level . ii. The mean inferred rate fluctuations (13) for the noise levels in i. (e)i. The mean error between simulated (true underlying) and inferred spiking rates for multiple underlying spiking rate datasets with different mean fluctuations. ii. The mean inferred spiking rate fluctuations as a function of the mean underlying spiking rate fluctuations. The black dash line marks $x = y$.

2(e)) and in individual traces (Figure 2(c)iii,iv). Our individual simulated calcium traces exhibit variable fluctuations that are highly correlated with those of the single spiking-rate traces from which they were generated (Figure 2(c)i, Pearson correlation 0.8846). In other words, the calcium data (and the fluorescence as a result) maintains the fluctuation structure of the underlying spiking rate. Lucy and GPDC infer spiking rate fluctuations that are less correlated with the underlying spiking rate fluctuations (Figure 2(c)iii,iv, Pearson correlations of 0.547 and 0.357, respectively). This is not inherent to inference; Con-Var successfully maintains the structure of fluctuations in the data even at the single trace level (Figure 2(c)ii. Pearson correlation 0.8936). This difference highlights a limitation of Lucy and GPDC, which is an outcome of their assumptions being less fitting for a wide-field setting. We hence find the fluctuations' correlation to be an informative metric for evaluating inference quality. We adopt its use, particularly for real data where the underlying spiking rate and thus the direct inference error are harder to estimate.

5.3 Anatomical and Optical (NAOMi) Simulated data

In this section, we evaluate inference performed by the Continuously-Varying Method (Con-Var; Algorithm 1), the Lucy–Richardson algorithm (Lucy), and Gaussian Process Deconvolution (GPDC). Here, we test these methods on fluorescence resulting from the activity of many neurons, generated using the Neural Anatomy and Optical Microscopy (NAOMi) simulator (Song et al. 2021). The advantage of using NAOMi lies in its supplying anatomically and biophysically detailed cortical datasets that are independent of our model assumptions (2), while still allowing full access to the ground truth by virtue of simulation.

The NAOMi simulator (Song et al. 2021) constructs a detailed volumetric model of cortex that includes vasculature and neurons, along with their somata, dendrites, and axons. It simulates neural spiking activity and its impact on the calcium concentration by accounting for calcium flux via voltage-dependent channel kinetics. It also simulates light emission and propagation by considering the binding of calcium to fluorescence-emitting molecules and the subsequent propagation of fluorescence through tissue (scattering, absorption, and reflection), which results in realistic optical readouts of neuronal activity.

We simulated a cortical-like volume using NAOMi, measuring $250 \times 250 \times 100\,(\mu m)^3$. This simulated volume contained 3,959 neurons, each spiking at an average rate of 10Hz. We "recorded" the simulated fluorescence resulting from this neural activity for the equivalence of half an hour and split the results into 60 traces of half a minute each. To emulate single-pixel imaging within a wide-field setup, we calculated the total fluorescence emitted from the entire volume, which is the sum of all photons that would be captured by NAOMi imaging simulator. [1]

Because neurons differ in their luminescence strength (sometimes by orders of magnitude), as reflected in NAOMi simulations, we needed a scheme to estimate the relative contribution of each neuron to the wide-field fluorescence trace. For each neuron i we defined a weight α^i equal to its baseline fluorescence, reflecting its approximate relative contribution. We used these weights to calculate the weighted spiking rate (see also the discussion in Section 2), $r_t = \frac{1}{N}\sum_{i=1}^{N} \alpha^i s_t^i$ where s_t^i is the spike train of neuron i. This weighting allowed Con-Var and Lucy to achieve successful inference (see Figure 3).

Con-Var consistently produced the lowest error, remaining below Lucy's minimum error with most decay constants, γ, and variable penalty weights, λ, tested (Fig. 3(c)i,ii). GPDC performed

[1] We are grateful to Alexander Song, the lead author of NAOMi, for his guidance in implementing a computationally efficient wide-field–like recording from NAOMi.

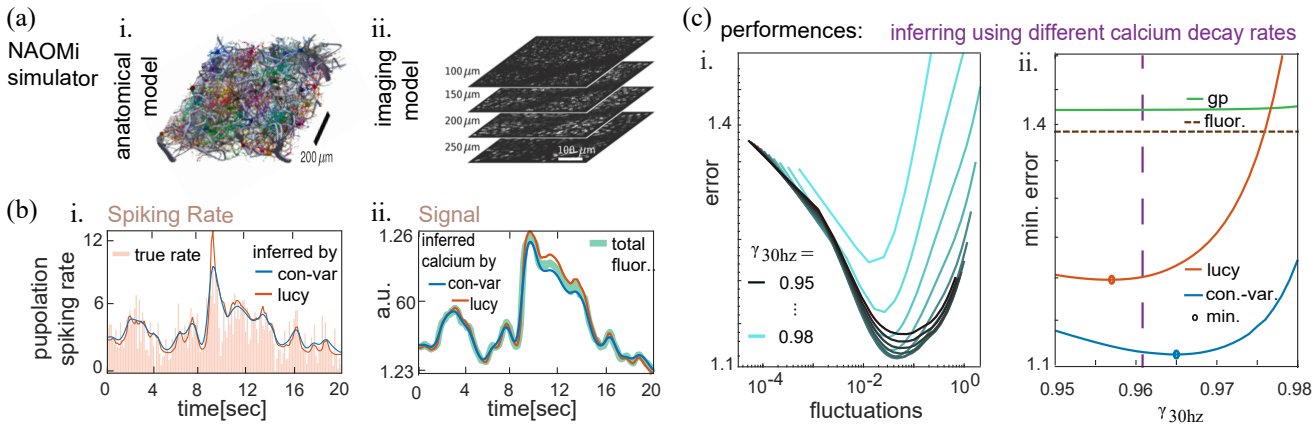

Figure 3: Evaluating inference with biophysical NAOMi simulations. (a) Adapted from NAOMi (Song et al., 2021) i. An illustration of a simulated cortical volume containing vasculature and neurons with dendrites and axons. ii. An illustration of simulated imaging of the activity from the volume in i. To simulate wide-field single-pixel fluorescence, we calculated the total fluorescence captured by the simulated imaging of the entire volume. (b) An example: i. A total (weighted) spiking rate calculated from a NAOMi simulation (ivory bars) and the inferred spiking rate by Con-Var (blue line) and Lucy (red line). ii. The resulting total fluorescence (green line) from the same spiking activity used in i. as simulated by the NAOMi imaging model. The inferred (weighted) calcium concentration by Con-Var (blue line) and Lucy (red line). (c) i. The mean error (12) between NAOMi simulated (weighted) spiking rates and Con-Var inferred spiking rates as a function of the mean inferred spiking rate fluctuations (13). For inference, Con-Var was applied multiple times with different values of γ ($0.95 \leq \gamma \leq 0.98$, yielding multiple lines) and λ ($1000 \leq \lambda \leq 0.01$, yielding variations in the inferred spiking rate fluctuations). ii. The mean error (12) between NAOMi simulated (weighted) spiking rates and the inferred spiking rates by Con-Var (blue line), Lucy (red line) and GPDC (green line), as a function of the different γs used for inference. Circles mark minimum errors achieved by Con-Var and Lucy. For comparison, we also included the error when fluorescence (fluor) is naively considered as the inferred spiking rate (dashed black line). Con-Var results are represented by the minimum error achieved with each γ (minimum error of each line in i.).

similarly to simply treating the fluorescence as the inferred spiking rate (Figure 3(c) ii, green, and dashed black lines, respectively) and in practice recovered an almost constant spiking rate due to its over-smoothing (see also the discussions in Sections 5.2 and 5.4) and the low variability of spiking rates in the NAOMi simulation.

Both Con-Var and Lucy captured the small fluctuations present in the simulations (Fig. 3(b)), with Lucy producing more fluctuating solutions, consistent with earlier results (Section 5.2, Figure 2(a)iii and Figure 2(e)ii), which in this case aligns more accurately with the data. Unfortunately, NAOMi simulations show relatively consistent fluctuation levels in the weighted spiking rate traces they generate (with only about a 5% variance across traces). This consistency hinders our ability to use NAOMi for evaluating how effectively inference methods track fluctuations across different scales. Additionally, this consistency leads to weak correlations (approximately 0.2) between the underlying weighted spiking rate and calcium or inferred spiking rates. These correlations primarily reflect noise rather than the ability (or inability) of the inference methods to follow fluctuations across scales.

Since NAOMi directly simulates calcium flux, binding, and fluorescence emission, it lacks an underlying decay rate γ. Song et al. (2021) estimated the effective decay constant for NAOMi, in the setting we employed, which simulates GCaMP6f kinetics (at a sampling rate of 30Hz), as $\gamma = 0.961$ (Fig. 3(c)ii purple dashed line). This finding aligns with the decay rate value reported for GCaMP6f by Chen et al. (2013), who also indicated approximately 10% deviations (Song et al. (2021) did not provide an estimate for these deviations in NAOMi). Our inference achieved minimal error with similar decay constants within this margin, estimating γ as approximately 0.957 with Lucy and 0.965 with Con-Var.

Overall, the results with NAOMi simulations are consistent with those obtained using our model in Section 5.2.

5.4 Recorded Data

Next, we tested the performances of the Continuously-Varying Inference (Con-Var, as detailed in Algorithm 1), the Lucy-Richardson Algorithm (Lucy) and the Gaussian Process Deconvolution (GPDC) on real recorded data. We used simultaneous recordings of wide-field fluorescence and spike counts. Unfortunately, it is currently not technically feasible to record the exact underlying neural activity alongside its corresponding fluorescence trace in a wide-field setting. As a result, we do not have the precise underlying spiking rate to test the inference methods. However, we hold an estimate of it. The dataset we use includes spikes recorded from the spatial location of the fluorescence origin from Clancy et al. (2019) Supplementary figure 4c; see Supporting Information E for more details about this dataset. Clancy et al. (2019) verified that the spikes in the dataset decently represent the underlying 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 discuss the implications of this as we tune the inference parameters for Con-Var.

Prior to testing the algorithms on the data, we rescaled the recorded fluorescence traces so that the assumption of $\overline{\beta}_1 = 1$ made in (2) holds. Since both the spiking rate r and the fluorescence $\overline{y}$ are provided in the dataset, $\overline{\beta}_1$ is implicitly given as well. We calculate its value using equation (27) with the same value of γ that we use for inference; see Supporting Information F. For inference, we take the recorded fluorescence multiplied by $\frac{1}{\overline{\beta}_1}$ (then the assumption $\overline{\beta}_1 = 1$ holds). This adjustment allows us to directly compare the (mean-subtracted) inferred spiking rate with the (mean-subtracted) recorded spiking rate.

We inferred the data using Con-Var with various calcium decay constants (γ values) and different penalties (λ values). As discussed in Section 5.1, any change in the penalty affects the fluctuations in the inferred spiking rate. We plotted the error against fluctuations for each γ (see Figure 4(b)i). The results indicate that Con-Var remains convex with respect to the error for both parameters, γ and λ, similar to the findings in the simulated data (Figure 1(b)i and Figure 3(c)i). For low fluctuations, the error lines are shallow. This is encouraging because it suggests that minor deviations in either the calcium decay constant γ or the penalty λ have a minor impact on the inferred spiking rate solution. This aspect 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 in the dataset we are examining. Similarly, Chen et al. noted a comparable value with a 10% deviation across cells, meaning $\gamma = 0.97 \pm 0.003$. These deviations may originate from natural cell variability, including differences in physical properties and significant variations in spiking rates (Chen et al. 2013). In wide-field recordings, deviations can also stem from the presence of non-neuronal tissue, such as glial cells, or from fluorescence that originates from axons or dendrites rather than

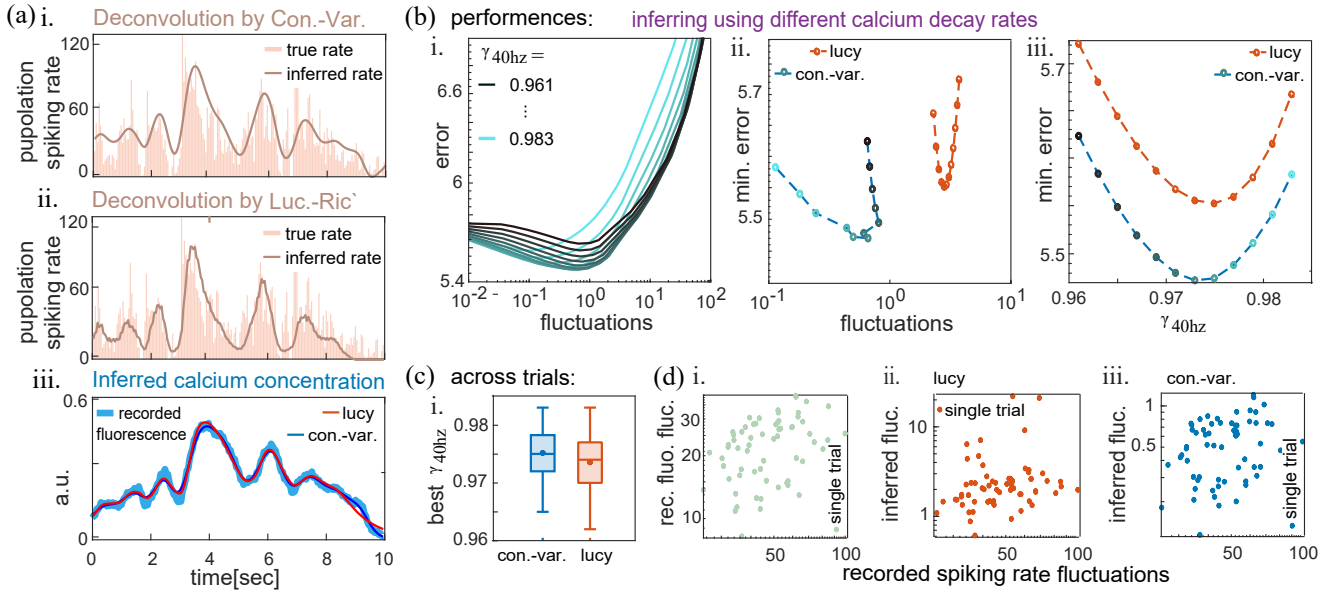

Figure 4: Evaluating algorithm performances on parallel spike counts and wide field fluorescence recordings. (a) An example of an underlying spiking rate (i. ivory spike counts) and the inferred spiking rate using Con-Var (i. brown line) or Lucy (ii. brown line). For inference, the fluorescence recorded in parallel was used (iii. light blue). (b) The mean error (12) between the recorded spiking rates and inferred spiking rates by Con-Var as a function of the mean inferred rate fluctuations (13). 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 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 corresponding i. recorded fluorescence fluctuations, ii. Lucy inferred spiking rate fluctuations, iii. Con-Var inferred spiking rate fluctuations. The Pearson correlations are: i. 0.375 ii. 0.138 iii. 0.265. Best γ and λ according to (b) were used for inference.

somas. Both Con-Var and Lucy yield minimum average errors when using decay constant values within the range reported for single cells ($\gamma = 0.973$ and $\gamma = 0.974$, respectively, Figure 4(b)iii). This suggests that the calcium decay rate observed in single cells is a reasonable estimate for the wide field setting as well.

We further tested the robustness of our estimate for the calcium decay rate γ by examining all γ values that minimize the error across each of the 65 segments of the data (see Supporting Information E). Our analysis shows that the error between the recorded and inferred spiking rates is minimized with $\gamma = 0.9749 \pm 0.0042$ when inferred using Con-Var, and $\gamma = 0.9731 \pm 0.0047$ when inferred using Lucy (where $\pm$ indicates the standard deviation). The calculated ranges (means and standard deviations) fall within the known single-cell range.

When we applied GPDC to this data, it produced an error significantly lower than that observed when using fluorescence as the inferred spiking rate (see Supporting Information Section H, Figure

7(b)). This differs from the results on simulated data (Figures 2(b) and 3(c)ii), where GPDC achieved a similar error to fluorescence. The difference is likely due to real recordings being much noisier than simulated, highlighting the importance of performing inference on real data. Yet, GPDC's error remains higher than what is achieved with Lucy or Con.-Var., likely because it captures slow changes only (see Supporting Information Section G, Figure 7(a)), which is consistent with results from simulations (see Figures 2(b) and 3(c)ii). As a result, GPDC does not provide additional insights for estimating the calcium decay rate in a wide-field setting.

The fluctuations of the inferred spiking rates by Lucy are higher than those inferred by Con-Var (Figure 4(b)ii). This result is similar to what was obtained from the simulated data (Figure 2(e)ii). The range of the inferred spiking rates by Lucy is often lower than those inferred by Con-Var (Supporting Information H Figure 7(a)), similar to what was observed from the simulated data (Figure 2(a)iii vs. Figure 2(a)i).

The structure of fluctuations in the data, i.e., the correlations between the recorded fluorescence fluctuations and the recorded spiking rate fluctuations, is limited by noise and 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 4(d)i, Pearson correlation of 0.375). And yet, similar to the results on simulated data, Lucy struggles to maintain the existing structure of fluctuations in the data (Figure 4(d)ii Pearson correlations of 0.138) while Con-Var better maintains some of it (Figure 4(d)iii Pearson correlations of 0.265).

5.5 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-neuron data. According to our model, equation (2), the calcium decay rate for the calcium concentration of many neurons follows the decay rate as that of single neurons. Our results in Section 5.4 confirm that; they demonstrate that the calcium decay in a wide field setting can be estimated from the single-neuron calcium dynamics.

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 noisy fluorescence recordings, it does not 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 the two tuning methods for λ in the following sections.

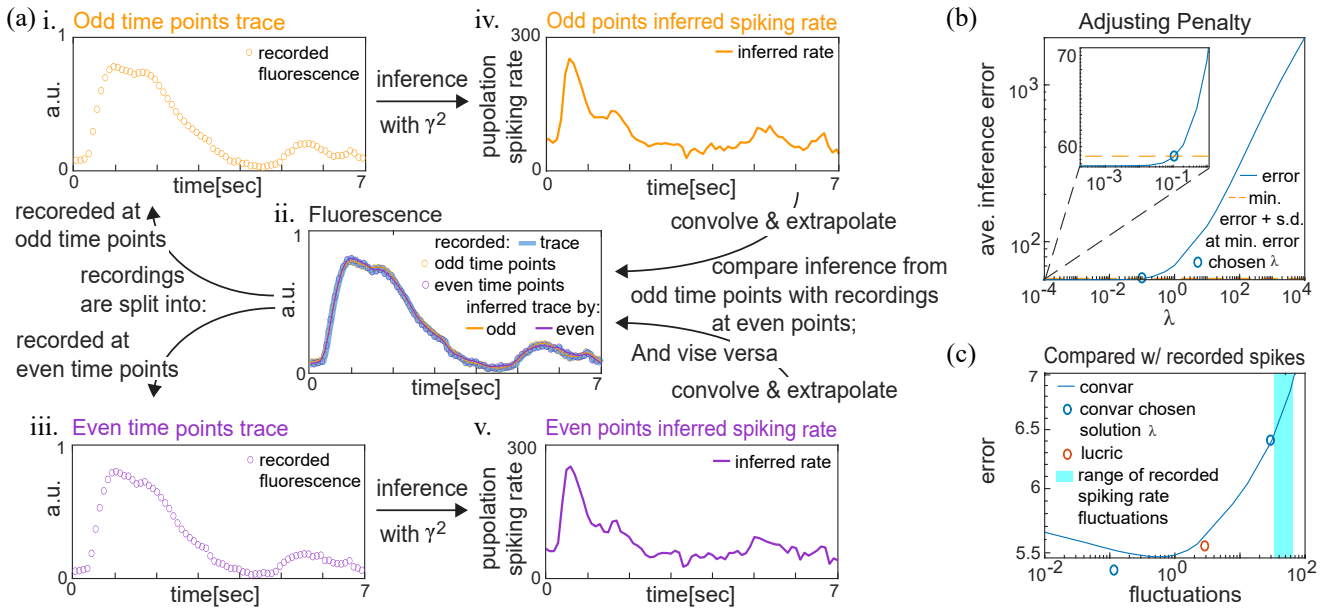

Figure 5: Evaluating the penalty λ by a cross validation scheme. (a) An example of a fluorescence trace (ii light blue), separated into an odd (i) and even (iii) indexed time points. Inferred from them, by Con-Var, are the odd and even indexed time points spiking rate (iv,v respectively). Their convolution to calcium signals are compared with the recorded fluorescence at the even and odd indexed time points (respectively, see Supporting Information G for the complete algorithm). This process is repeated with different λ values. (b) The average error (12) 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 minimum error achieved plus the standard deviation at the minimum error (orange dashed line). (c) The mean error (12) between the recorded spiking rates and inferred spiking rates as a function of the mean inferred rate fluctuations (13). The plot is taken from Figure 4(b) 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.

5.5.1 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 Supporting Information 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 5(a)i,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, and the noise in the recording is assumed to be independent (in 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 5(a)ii). The penalty λ is chosen to take its maximum value that still gives an error less than the minimum error plus the standard error deviation at the minimum error point (Figure 5(b)).

We applied Algorithm 2 (Supporting Information G) on the fluorescence recordings from Clancy

et al. (2019). The results were satisfactory (Figure 5(c) blue circle) but did not minimize the error as much 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 5(c) 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.

5.5.2 Choosing Penalty by Maintaining Signal Properties

We offer here an alternative, a novel approach, to tune the penalty λ. It includes an analysis of the results, which can also be used in perplexing cases to evaluate the data. This approach leverages large datasets by examining the relative amount of a dynamic quantity each segment in the data contains. It suits wide-field experiments because a large amount of fluorescence data is typically recorded in each such experiment. The core idea behind this approach is to find a penalty that best maintains the relative amount of a fundamental dynamical quantity prior to and post-inference. The logic behind it is that the chosen dynamical 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 rate fluctuations, on average (see Figure 6(a) for an example) and vice versa for lower fluctuations. We hence 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 that noise. With too much signal inferred as noise, much of the signal is lost (Figure 6(a)iii); with too little, the noise is inferred as signal (Figure 6(v)iii), 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 does not meet our needs. At some middle value of λ the correlation has a local maximum. At this point, most of the dynamical information is conserved, while most of the noise is removed from the signal. This is our chosen λ.

Our data indicates that the finding-local-maximum scheme (as described above) yields a solution that is very close to the best available option (Figure 6(d). 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. This allows it to evaluate the inferred spiking rate while considering both noise resources. Hence, it fits our data (recall that the recorded spike counts in Clancy et al. (2019) are only correlated with the true underlying spiking rate of the recorded fluorescence). In different experimental setups, additional noise in the underlying spiking rate may arise from situations where only a few traces are available to represent an average response or to compare across multiple experiments, among other factors.

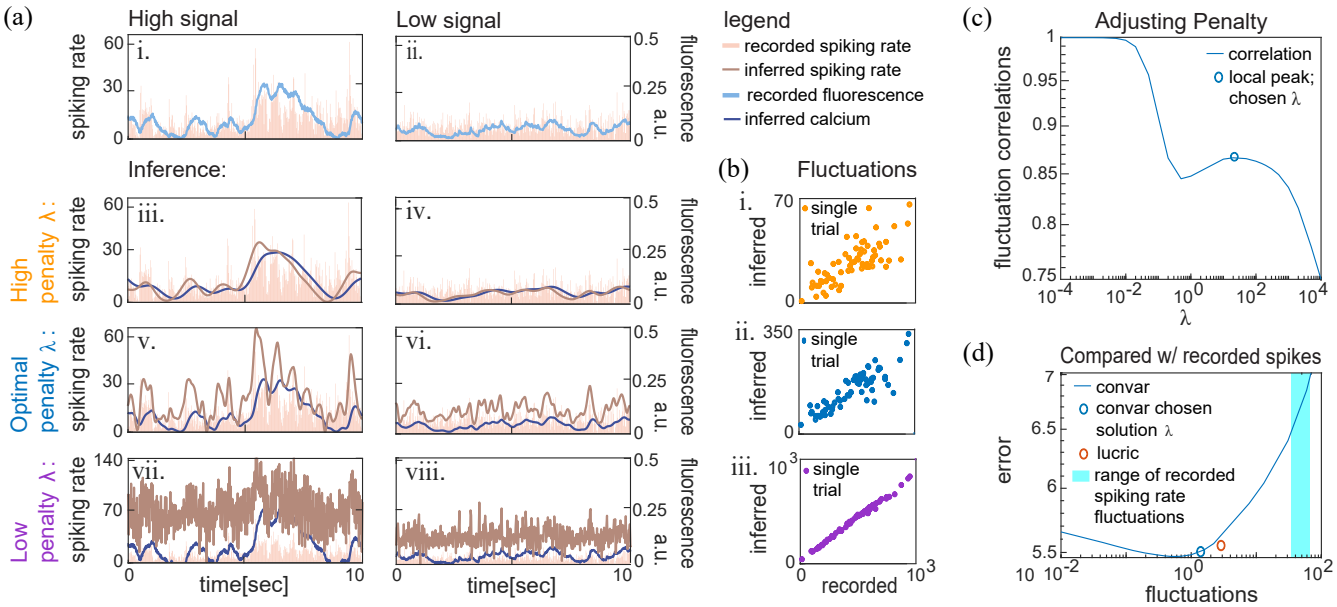

Figure 6: Analyzing and estimating the penalty λ according to fluctuation correlations. (a) i,ii. Two examples of recorded fluorescence (light blue) and the parallel recorded spiking rates (ivory). iii-vii. The inferred spiking rate (brown) from the two fluorescence examples using iii,iv. high, v,vi. 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 5(c) with the blue circle representing the local minimum solution as marked in (c).

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 present an analytical solution to the inference problem of recovering continuously varying spiking rate dynamics from the fluorescence recordings obtained in a wide field setting.

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) and the Gaussian Process Deconvolution (Tobar et al. 2023) (adapted here to wide-field); Other inference methods

in the published literature for wide-field recordings produce only rough estimates of the spiking rates (O'Rawe et al. 2023, Peters et al. 2021)).

First, we demonstrate that our solution yields lower error than the Lucy-Richardson algorithm and the Gaussian Process Deconvolution in the scenarios we tested, including various SNR levels and underlying fluctuations in synthetic data (Figure 2(b) and Figure 3(c)ii) and real data (Figure 4(b)ii). Second, we demonstrate that our solution is more consistent compared to the Lucy-Richardson algorithm and the Gaussian Process Deconvolution in the sense that it retrieves the lowest error with the correct underlying calcium decay dynamics (Figure 2(b)). Our Continuously-Varying solution also best maintains fluctuation correlations among the methods we tested (Figures 2(f)iii and Figure 4(d)iii).

Limitations of the solution we present 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 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 a 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 allows our Continuously-Varying solution to be adaptable, designed for the possibility of being 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 competitive with other possible non-tuning-requiring approaches, such as the Lucy-Richardson algorithm. Overall, we find that the Continuously-Varying solution's lower error and better fluctuation correlation are encouraging for addressing the potential challenges in tuning.

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 5.4.

We envision our inference solution becoming part of the "pipeline" for 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 recommend using our inference solution after spatial distortion corrections and before any dynamics-related analysis, such as dimensional reduction, finding regions of interest, PCA, NMF, and mixed linear models. With these steps, the inference is expected to perform best, leading to more accurate analysis results altogether.

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

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

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

$$\overline{c}_t = \sum_{t'=1}^{t} \gamma^{t-t'} r_{t'}. \tag{14}$$

Since $\overline{c} = D^{-1}r$, we find from (14) that the entries of D^{-1} are given by

$$\left(D^{-1}\right)_{i,j} = \begin{cases} 0, & \text{if } i < j \\ 1, & \text{if } i = j \\ \gamma^{i-j}, & \text{if } i > j \end{cases}. \tag{15}$$

From (15) we find the sum of the elements in each D^{-1} row i,

$$\sum_{j=1}^{T}\left(D^{-1}\right)_{i,j} = \sum_{j=1}^{T}\gamma^{j-1} = \frac{1-\gamma^i}{1-\gamma}, \tag{16}$$

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

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

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,

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

where the last step was calculated using (16) and (17). 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,

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

The structure of A and D^{-1} column sums (17) also gives the following,

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

From (20) and (19) we have,

$$\frac{\sum_{j=1}^{T} A_{i,j}}{A_{i,1}} = -\frac{\gamma}{(1-\gamma)}. \tag{21}$$

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 3. 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,

$$\|\tilde{y} - A\tilde{r}\|^2 = \|\tilde{y} - (A\hat{r} + Ad\mathbb{1} + A\tilde{r}_{m1})\|^2 = \|\tilde{y} - A\hat{r}\|^2. \tag{22}$$

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

$$\sum_{t=3}^{T} (\tilde{r}_t - \tilde{r}_{t-1})^2 = \sum_{t=3}^{T} (\hat{r}_t - \hat{r}_{t-1})^2. \tag{23}$$

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; for Figure 1a,b we used $g = 5.7$ that generated 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 traces, we subtracted their 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):

$$\tilde{y} = \overline{\beta}_1 Ar. \tag{24}$$

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

$$\underset{\overline{\beta}_1}{\text{minimize}}\{\|\tilde{y} - \overline{\beta}_1 Ar\|^2\} \tag{25}$$

yielding

$$(Ar)^\top \tilde{y} = \overline{\beta}_1 (Ar)^\top Ar, \tag{26}$$

and hence the solution

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

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 (27) 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 (15). 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 Jewell 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*}

H Figure 4 Supplementary

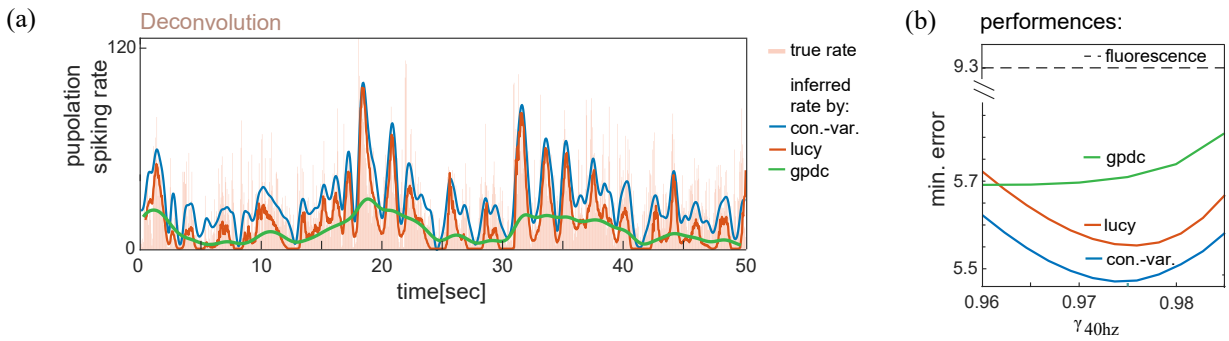

Figure 7: Evaluating method performances on parallel spike counts and wide field fluorescence recordings. (a) An example of an underlying spiking rate (ivory spike counts) and the inferred spiking rate using Con-Var (blue line), Lucy (red line) or GPDC (green line). (b) The mean error (12) between the recorded spiking rates and and the inferred spiking rates by Con-Var (blue line), Lucy (red line) and GPDC (green line), as a function of the different γs used for inference. We included the error when fluorescence is naively considered as the inferred spiking rate (dashed black line). Con-Var results are represented by the minimum error achieved with each γ (minimum error of each line in main text Figure 4(b)i.)

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.

Bertero, M., Boccacci, P. & De Mol, C. (2021), *Introduction to inverse problems in imaging*, CRC press.

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.

Chaudhuri, S., Velmurugan, R. & Rameshan, R. (2014), 'Blind deconvolution methods: A review', *Blind Image Deconvolution: Methods and Convergence* pp. 37–60.

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.

Engl, H. W. & Groetsch, C. (1987), 'Inverse and ill-posed problems', *Notes and Reports in Mathematics in Science and Engineering* .

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

Ganmor, E., Krumin, M., Rossi, L. F., Carandini, M. & Simoncelli, E. P. (2016), 'Direct estimation of firing rates from calcium imaging data', *arXiv preprint arXiv:1601.00364* .

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

Hansen, P. C. (2010), *Discrete inverse problems: insight and algorithms*, SIAM.

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.

Song, A., Gauthier, J. L., Pillow, J. W., Tank, D. W. & Charles, A. S. (2021), 'Neural anatomy and optical microscopy (naomi) simulation for evaluating calcium imaging methods', *Journal of neuroscience methods* **358**, 109173.

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.

Tobar, F., Robert, A. & Silva, J. F. (2023), 'Gaussian process deconvolution', *Proceedings of the Royal Society A* **479**(2275), 20220648.

Vogel, C. R. (2002), *Computational methods for inverse problems*, SIAM.

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.

Wei, X. et al. (2020), 'A zero-inflated gamma model for deconvolved calcium imaging traces. arxiv preprint 450', *arXiv preprint arXiv:2006.03737* **451**.

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.

Williams, C. K. & Rasmussen, C. E. (2006), *Gaussian processes for machine learning*, Vol. 2, MIT press Cambridge, MA.

Wipf, D. & Zhang, H. (2014), 'Revisiting bayesian blind deconvolution', *Journal of Machine Learning Research (JMLR)* .

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.