Dimensionality reduction of calcium-imaged neuronal population activity

Tze Hui Koh[1,2], **William E. Bishop**[2,3,4], **Takashi Kawashima**[4,5], **Brian B. Jeon**[1,2], **Ranjani Srinivasan**[1,6], **Yu Mu**[7], **Ziqiang Wei**[4], **Sandra J. Kuhlman**[8,9], **Misha B. Ahrens**[4,†], **Steven M. Chase**[1,8,*,†], **Byron M. Yu**[1,8,10,*,†]

[1]Department of Biomedical Engineering, Carnegie Mellon University, PA

[2]Center for the Neural Basis of Cognition, PA

[3]Department of Machine Learning, Carnegie Mellon University, PA

[4]Janelia Research Campus, Howard Hughes Medical Institute, VA

[5]Department of Brain Sciences, Weizmann Institute of Science, Israel

[6]Department of Electrical and Computer Engineering, Johns Hopkins University, MD

[7]Institute of Neuroscience, Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, China

[8]Carnegie Mellon Neuroscience Institute, Carnegie Mellon University, PA

[9]Department of Biological Sciences, Carnegie Mellon University, PA

[10]Department of Electrical and Computer Engineering, Carnegie Mellon University, PA

Abstract

Calcium imaging has been widely adopted for its ability to record from large neuronal populations. To summarize the time course of neural activity, dimensionality reduction methods, which have been applied extensively to population spiking activity, may be particularly useful. However, it is unclear if the dimensionality reduction methods applied to spiking activity are appropriate for calcium imaging. We thus carried out a systematic study of design choices based on standard dimensionality reduction methods. We also developed a method to perform

deconvolution and dimensionality reduction simultaneously (Calcium Imaging Linear Dynamical System, CILDS). CILDS most accurately recovered the single-trial, low-dimensional time courses from simulated calcium imaging data. CILDS also outperformed the other methods on calcium imaging recordings from larval zebrafish and mice. More broadly, this study represents a foundation for summarizing calcium imaging recordings of large neuronal populations using dimensionality reduction in diverse experimental settings.

Introduction

Computations in the brain occur through the coordinated, time-varying activity of populations of neurons. Dimensionality reduction is a class of statistical methods commonly used for summarizing neural population activity [1, 2, 3]. It transforms high-dimensional neural recordings, such as spiking activity from a population of recorded neurons, into compact low-dimensional representations termed *latent variables*. These low-dimensional representations facilitate the investigation of how neural population activity varies over time, across experimental conditions, and across repeated experimental trials of the same condition. In particular, dimensionality reduction has been used to uncover neural mechanisms underlying decision making [4], motor control[5], learning [6], working memory [7], sensorimotor timing [8], attention [9], olfaction [10], speech [11], and more.

Dimensionality reduction has typically been applied to electrophysiological recordings. In the last decade, optical imaging has been widely adopted to record from large populations of neurons. Optical imaging has the ability to sample neurons densely within the field of view, track neurons over long periods of time, and label neurons by cell type or projection, among other advantages [12]. A leading type of optical imaging is calcium imaging, which uses calcium indicators to track the transient increase in intracellular calcium levels that accompanies electrical spiking activity [13]. These changes in calcium levels are then optically recorded via changes in fluorescence. Calcium imaging has the capability of imaging even the whole brain of some small animals (e.g., larval zebrafish) at single neuron resolution [14], albeit at a lower temporal resolution than electrical recordings.

With the increasing use of calcium imaging, many studies are now beginning to apply dimensionality reduction to calcium imaging recordings [15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26]. A critical question is whether the same dimensionality reduction methods previously applied to study spiking activity are also appropriate for calcium imaging recordings [27]. Here, we focus on the use of dimensionality reduction methods to extract single-trial time courses of latent variables, termed *neural trajectories*. This enables the study of how trial-to-trial differences in the time course of neural activity relate to trial-to-trial differences in perception, decision making, and behavior [28, 16, 29, 30, 31, 8]. We seek to understand if the neural trajectories extracted from calcium imaging recordings faithfully capture the shared, time-varying activity among the population of recorded neurons. A key property of calcium imaging is the slow, indicator-dependent decay of measured calcium levels after each spiking event [13] (Fig. 1). If ignored, this decay could introduce temporal correlations in the extracted neural trajectories that are not present in the underlying spiking activity of the recorded neurons. Thus, the neural trajectories that are extracted from calcium

imaged activity may be quite different from the neural trajectories that would be extracted from spiking activity. Deconvolution is a technique that aims to recover spiking activity from calcium imaging recordings [32]. However, deconvolution techniques do not, as yet, recover the underlying spikes exactly [33].

In this work, our central goals are to i) systematically study the appropriateness of dimensionality reduction methods for summarizing the time course of calcium imaging recordings, and ii) propose a dimensionality reduction method that is tailored for extracting neural trajectories from calcium imaging recordings in exploratory data analysis. We sought to address three questions. First, we asked if deconvolution should be used with dimensionality reduction when extracting neural trajectories, and if so, how it should be applied. Second, we asked how different experimental variables (e.g., the decay constant of the calcium indicator, the timescale of the latent time courses, and the number of imaged neurons) impact the ability to recover neural trajectories from calcium imaging recordings. Third, we asked if it is necessary for the dimensionality reduction method to employ a dynamics model for the latent variables, as such a model might enable the neural trajectory to be more cleanly separated from the time course of calcium decay.

We addressed these questions by comparing several approaches, including i) standard dimensionality reduction applied directly to the recorded fluorescence; ii) a two stage-method in which deconvolution is applied separately to each neuron's fluorescence trace to estimate spiking activity, then standard dimensionality reduction is applied to the estimated spiking activity; and iii) a unified method that we propose here (Calcium Imaging Linear Dynamical System, CILDS), which performs deconvolution and dimensionality reduction jointly. We first applied these methods to simulated fluorescence traces, in which we systematically varied several experimental variables over a wide range. We then applied these methods to calcium imaging recordings from the dorsal raphe nucleus of larval zebrafish and the primary visual cortex of mice. Across these settings, we found that CILDS outperformed the other methods. Compared to the other methods, CILDS was able to better peer through the calcium decay and apply the appropriate degree of temporal smoothing to the latent variables. Overall, our work provides a foundation for using dimensionality reduction to summarize the time course of calcium imaging recordings.

Results

Model overview

Our central goal is to develop dimensionality reduction methods appropriate for calcium imaging recordings to capture the shared, time-varying activity among the population of recorded neurons. To do so, we systematically compare three approaches.

For the first approach, we applied a standard dimensionality reduction method directly to the recorded fluorescence traces from calcium imaging (Fig. 2a, top). Here we chose to use a latent Linear Dynamical System (LDS), which is among the most basic methods for extracting neural trajectories. Conceptually, an LDS seeks to explain the temporal structure in the data using latent variables that vary smoothly over time.

Each time a neuron spikes, intracellular free calcium increases, then decays slowly over time. The calcium indicator kinetics influence the measured decay time, resulting in fluorescence traces whose intensity decays over hundreds of milliseconds to seconds, depending on the particular calcium indicator used [13]. This decay transient induces temporal correlations in the fluorescence measurements which are input to the LDS, which might attempt to capture these correlations in its latent variable estimates. This motivates our second approach, which first deconvolves each fluorescence trace separately, and then applies a LDS to the resulting estimated spiking activity. The deconvolution serves to remove a substantial portion of the calcium decay transient, producing activity traces similar to spike trains (or time varying firing rates). We term this two-stage method *deconv-LDS* (Fig. 2a, middle). Here we chose to deconvolve the fluorescence traces with OASIS [34, 32], which has been widely used in calcium imaging studies [23, 20, 35, 24].

In deconv-LDS, each neuron is deconvolved independently, and the stages of deconvolution and dimensionality reduction are performed sequentially. We asked if performing these stages jointly would lead to more accurate recovery of the latent variables (Fig. 2a, bottom). More specifically, we hypothesized that a sequential method (e.g., deconv-LDS) may inadvertently discard some of the shared activity amongst neurons due to the independent deconvolution of each neuron (Fig. 2b, left). Since the latent variables are intended to capture the shared spiking activity among neurons (and not the calcium decay, which is independent across neurons), it might be possible to better separate the calcium decay from the latent dynamics by considering all the neurons together, and performing the two stages of deconvolution and dimensionality reduction jointly. This allows the dimensionality reduction component to influence the deconvolution estimates, and vice versa, to more accurately estimate the ground truth latent variable than deconv-LDS (Fig. 2b, right). Thus, for our third method we developed a unified approach, *CILDS*, in which dimensionality reduction and deconvolution are performed jointly (Fig. 2a, bottom).

Deconvolution with dimensionality reduction

We first asked whether it is beneficial to first deconvolve fluorescence to estimate spiking activity in order to recover the underlying latent variables[27]. Indeed, multiple studies have applied dimensionality reduction to deconvolved spiking estimates [19, 20, 21, 22, 35, 23, 24, 36]. However, deconvolution is subject to particular statistical modeling assumptions (as is any statistical method) and usually does not recover the underlying spikes exactly. Therefore it is unclear how, or even if, deconvolution should be used with dimensionality reduction. We addressed this question by comparing the three approaches described above (Fig. 2a).

With calcium imaging recordings, the "ground truth" latent variables are unknown. Hence to directly compare each method's ability to extract latent variables, we designed a simulation framework in which we created known ground truth latent variables with smoothly-varying time courses. These latent variables were used to generate spike trains which, in turn, were used to generate fluorescence traces (see Methods). We then applied each of the three approaches to these simulated fluorescence traces to assess how accurately they recovered the ground truth latent variables. Examples of two combinations of experimental variables

are illustrated in Fig. 3a. In the simulations, we used 10 latent variables, of which two are shown.

We found that in both settings, CILDS outperformed the other two methods, returning more accurate estimates of the ground truth latent variables (Fig. 3b–c, points above the diagonal). By operating on the entire population of neurons together, CILDS was better able to separate the calcium transients from the shared activity among neurons (i.e., the latent variables) compared to deconv-LDS, which deconvolves the activity of each neuron individually, and LDS, which makes no attempt at this separation. Taken together, when extracting single-trial neural trajectories, one should use deconvolution jointly with dimensionality reduction, as in CILDS.

Impact of experimental variables on latent variable recovery

We next asked how different experimental variables affect the accuracy of the recovered latent variables. The simulation framework enables us to systematically vary experimentally relevant variables. These variables comprise four axes along which we can explore different combinations of values that might mimic a particular experimental paradigm. We varied the timescales of the latent time courses from 50ms to 5000ms, the number of neurons from 20 to around 100, the calcium decay timescales to match GCaMP6f, GCaMP6m, and GCaMP6s (fast, medium, slow) [13], and the variance of the added imaging noise that was independent of the calcium and spiking activity (Supplementary Table 1). We assessed how the accuracy of the dimensionality reduction methods changed as we systematically changed these variables.

As we increased the timescales of the latent variables, all dimensionality reduction methods improved their accuracy in estimating the ground truth latent variables (Fig. 3d). This occurs because with slower latent fluctuations, the latent variables become less independent across time. As a result, all methods can leverage future and past time points to better estimate the latent variables at the current time point. When the latent timescales are slow (order of seconds), the calcium indicator decay (which also has a timescale on the order of seconds) tends not to blur the underlying neural activity. Therefore a method that is not able to disambiguate between the time courses of the latent variables and the calcium decay (e.g., LDS) can still accurately recover the latent variables (see latent timescale of 5000 ms in Fig. 3d). However, at faster latent timescales (tens to hundreds of milliseconds) which reflect the timescales of many sensory, cognitive, and motor functions [37, 35], it is critical to use a method that accounts for the calcium decay, as both CILDS and deconv-LDS do (see latent timescale of 50 ms in Fig. 3d).

Next, when we increased the number of "recorded" neurons, all three methods improved in their ability to reconstruct the ground truth latent variables (Fig. 3e). This makes sense because each neuron provides a different, noisy view of the underlying latent variables. With more neurons, all methods are better able to "triangulate" the values of the latent variables.

We also varied the time constant of the GCaMP calcium indicator decay to match GCaMP6f, 7 6m, and 6s (from fast to slow) (Fig. 3f). All methods performed worse as the decay time constant increased. This occurred because the slower the calcium indicator is, the less

the resulting fluorescence signal resembles the original spike train, which increases the difficulty in disambiguating between the latent time courses and the calcium decay (Fig. 3f).

Finally, we varied the amount of noise added to the fluorescence, which reflects imaging noise independent of calcium and spiking activity (Fig. 3g). As the variance of the noise increased, all three methods performed worse as expected, although the extent of the performance degradation differed across methods. As the fluorescence noise variance increased, CILDS continued to outperform the other two methods (Fig. 3g). This indicates that leveraging the population of neurons for simultaneous deconvolution and dimensionality reduction, as done by CILDS, provides statistical power to mitigate a loss in accuracy due to increased fluorescence noise (Fig. 2b).

Overall, CILDS performed as well or better than the other two methods in every simulated setting we tested (Fig. 3d–g, orange higher than purple and cyan, also see Supplementary Fig. 1 for additional combinations of simulation parameter settings and Supplementary Fig. 2 for how the results vary with mean firing rates). We additionally found that deconv-LDS usually outperformed LDS in accuracy of recovered latent variables, consistent with Wei et al., which applied PCA to trial-averaged activity [27] (Fig. 3d–g, purple higher than cyan). Using deconvolution is particularly important in regimes where the time scales of neural activity (i.e., the latent timescales) are faster than that of the calcium decay, which is the case for many commonly-studied brain functions.

Need for latent dynamical model in dimensionality reduction

All three dimensionality reduction methods considered so far explicitly attempt to extract latent variables that evolve smoothly over time via a latent dynamical model. We asked if including a latent dynamical model was necessary for the accurate recovery of latent variables.

To address this, we compared three methods: 1) our method that incorporates deconvolution and latent dynamics (CILDS), 2) a method we created that incorporates deconvolution but no latent dynamics (Calcium Imaging Factor Analysis, CIFA, see Methods), and 3) an off-the-shelf method that does not incorporate deconvolution or latent dynamics (Factor Analysis, FA). CIFA is identical to CILDS, except that CIFA does not have a latent dynamical model, as with conventional FA. Across a range of latent timescales, we measured the ability of each method to accurately recover the ground truth latent variables (Fig. 4a).

We found that when the latent timescales are fast, the calcium transient blurs the neural activity and deconvolution is necessary for more accurate latent variable recovery (Fig. 4a, left, consistent with Fig. 3d). This is why the methods that involve deconvolution (CILDS, CIFA) outperform the method that does not involve deconvolution (FA). CILDS outperforms CIFA in this regime because CILDS additionally smooths the latent variables over time via the latent dynamical model in a way that is determined by the data (see Methods).

When the latent timescales are slow, it becomes important that the latent variables are smoothed temporally (Fig. 4a, right), either through the use of a latent dynamical model

(CILDS) or by retaining the smoothing provided by the calcium indicator dynamics when they are not deconvolved away (FA). These methods outperform the method for which the latent variables do not have temporal smoothing (CIFA). Furthermore, the fact that CIFA does not have a latent dynamical model influences its estimate of the calcium decay constant. CIFA can attribute temporal smoothness in the fluorescence to only one possible source – calcium decay. For this reason, as the simulated latent timescale increases (left to right in Fig. 4b, Supplementary Fig. 3), CIFA erroneously attributes the slower varying fluorescence to a slower calcium decay, rather than to a longer latent timescale. By contrast, CILDS can attribute smoothness in the fluorescence to two possible sources – latent variables that evolve smoothly over time and calcium decay. As a result, as the simulated latent timescales increase (left to right in Fig. 4b, Supplementary Fig. 3), CILDS correctly attributes the slower varying fluorescence to the latent variables, and not to a longer calcium decay time constant.

In summary, we found that rather than a latent dynamical model being necessary per se, temporal smoothing of the latent variables is necessary for their accurate recovery, to a degree that is appropriate for the underlying latent or neural process. The trends that we observed for CILDS, CIFA, and FA are true across a variety of related methods (Supplementary Fig. 4). Overall, CILDS performs well in all regimes, regardless of whether the latent timescales are fast or slow. For this reason, we consider CILDS to be an excellent tool for exploratory data analysis for which we do not know a priori whether the underlying neural processes are fast or slow.

Application to calcium imaging recordings

To assess whether the advantages of CILDS also hold in real data, we applied each of the dimensionality reduction methods described above (CILDS, deconv-LDS, LDS, and CIFA) to calcium imaging recordings in two experimental contexts: larval zebrafish and mice. To emphasize the generality of our findings, these two experimental settings involve not only different animal species, but also different brain areas, behavioral tasks, and properties of the recorded fluorescence (see below). In these experiments, the ground truth latent variables are unknown. To quantify the accuracy of each method, we adopted a leave-neuron-out procedure used in previous studies [38, 39]. We estimate the latent variables using all-but-one neuron, and then assess how well these latent variables predict the recorded fluorescence of the held-out neuron (see Methods). This procedure assesses the relative ability of each method to identify a set of latent variables that captures the variability that is shared amongst the neurons. We used between 3 to 17 latent variables, depending on the dimensionality reduction method, for the zebrafish and 50 latent variables for the mice (see Methods).

The first experimental context involves larval zebrafish engaged in a "fictive swimming" moto-sensory gain adaptation task (Fig. 5a) [40]. Calcium imaging was performed on neurons expressing GCaMP6f in dorsal raphe nucleus (DRN) using light-sheet microscopy in three fish (19 to 22 neurons; mean: 20) (Fig. 5a). We applied each dimensionality reduction method to these recordings and assessed their performance using the leave-neuron-out prediction procedure (Fig. 5b). We found that CILDS more accurately predicted the fluorescence of the held-out neurons than the other methods (Fig. 5c–e), as quantified

by the correlation between the predicted and recorded fluorescence. Similar results were obtained when examining each fish individually (Supplementary Fig. 5), as well as for a larval zebrafish experiment in which hundreds of neurons across multiple brain areas were analyzed together (300 neurons) (Supplementary Fig. 6).

The second experimental context involves awake, head-fixed mice passively viewing static visual gratings (Fig. 6a) [41]. Two-photon calcium imaging was performed on neurons expressing GCaMP6f in the primary visual cortex (V1) of three mice (133 to 319 neurons; mean: 234.7). Comparing the raw recorded fluorescence of the two experimental contexts (Fig. 5a versus Fig. 6a), the neurons in the DRN of the larval zebrafish tend to exhibit slower fluctuations in fluorescence that are more correlated across neurons (Supplementary Fig. 7). In contrast, the neurons in mouse V1 exhibit faster changes in fluorescence, which are less correlated across neurons. We applied the same leave-neuron-out fluorescence prediction analysis from Fig. 5 to these mouse recordings (Fig. 6b). Despite the stark differences between the fish DRN and mouse V1 fluorescence traces, we again found that CILDS more accurately predicted the fluorescence of held-out neurons than LDS (Fig. 6c), deconv-LDS (Fig. 6d), and CIFA (Fig. 6e). These results were also true for each mouse individually (Supplementary Fig. 8).

Thus far, we have shown that CILDS extracts latent variables that more accurately predict the fluorescence of held-out neurons than the other methods. Another way to assess how meaningful the latent variables extracted by each method are is to measure how strongly they reflect external variables [42]. We thus performed a decoding analysis, whereby we classified the orientation and spatial frequency of the presented grating using the latent variables extracted by each of three methods (Fig. 6f). We found that CILDS and CIFA performed similarly, and both outperformed deconv-LDS (by 1.16 times) and LDS (by 1.43 times) (Fig. 6g). This demonstrates that the joint methods are better at capturing the shared activity changes among neurons that are relevant to the visual stimulus than the other methods. To understand why CILDS and CIFA performed similarly, recall from the simulations that the performance between CILDS and CIFA becomes more similar as latent timescales get faster (Fig. 4a). The latent timescales of these mouse recordings are at least as fast as the rate of change in the visual stimulus (every 250ms), as seen by the abrupt transients in the raw fluorescence traces (Fig. 6a). Furthermore, the higher decoding accuracy for deconv-LDS compared to LDS indicates that accounting for the calcium transients is important when the variable of interest changes on the timescale of tens to hundreds of milliseconds.

Taken together, the results based on calcium imaging recordings from two different recording regimes are consistent with what we identified in simulation. Namely, deconvolution should be used jointly with a dimensionality reduction method that provides temporal smoothing (as in CILDS), particularly if the neural process of interest changes on a timescale of tens to hundreds of milliseconds.

Discussion

This work focused on the question of which dimensionality reduction method is most appropriate for extracting single-trial neural trajectories from calcium imaging recordings. There are two other settings in which dimensionality reduction is commonly used. First, dimensionality reduction can be applied to analyze how trial-averaged activity differs across experimental conditions (e.g., ref. [17, 5, 4, 7]). In a recent study, Wei et al., applied principal components analysis (PCA) to trial-averaged electrophysiological recordings and calcium imaging with GCaMP6s [27]. They found important differences in the low-dimensional PCA trajectories obtained from electrophysiological recordings versus calcium imaging. This difference was mitigated by first deconvolving the calcium imaging recordings before applying dimensionality reduction, consistent with our findings. It may be possible to further improve the correspondence by applying CILDS to single-trial fluorescence recordings, then averaging the extracted low-dimensional neural trajectories across trials.

Second, dimensionality reduction is often used to analyze the trial-to-trial variability of neural population activity without time courses, i.e., using one time point or time window per trial (e.g., ref. [9, 43, 44, 45, 46]). In this case, there would be no information about how the calcium decays and so one would not be able to make use of a method that incorporates deconvolution. One might consider using CIFA, by analogy to the use of factor analysis to study the trial-to-trial variability of spike counts without time courses. However, it is important to note that, like CILDS, CIFA also requires multiple time points to be able to deconvolve, even though the latent variables in CIFA are independent from one time point to the next. If the original time series of calcium imaging is available, one can apply CILDS to the time series first, then average across the time points of the extracted latent variables. If the original time series of calcium imaging is not available, then a standard dimensionality reduction method such as factor analysis might be more suitable.

Previous studies have proposed methods for analyzing calcium imaging recordings that include latent variables and deconvolution in the same statistical model. Triplett et al. [47] developed a method to study the interaction between evoked and spontaneous activity using calcium imaging recordings in sensory systems. In their model, the latent variables represent activity fluctuations shared amongst neurons that are not explained by the sensory stimulus, where these activity fluctuations are defined to be spontaneous activity. Aitchison et al. [48] developed a method to infer spiking activity and neural connectivity from calcium imaging experiments that involve optogenetic stimulation. In their model, the latent variables represent shared activity amongst neurons that are not explained by the optogenetic stimulation or the activity of other neurons recorded simultaneously, and are intended to represent input from other brain areas. We developed CILDS for extracting latent variable time courses that summarize the population activity on individual experimental trials. In contrast to the two methods above which have more specific analysis goals, CILDS is general-purpose and well-suited for exploratory data analysis. For example, one might ask whether the neural activity can be separated based on behavioral context or sensory stimuli (Supplementary Fig. 9). This is akin to the use of methods such as LDS [49], GPFA [38], TCA [50], LFADS [42], and dPCA [51] for exploratory analysis of population spiking

activity, whose results can then lead to the use of methods with more specific goals (see ref. [52, 53, 54] for examples).

CILDS can be extended in the following ways. First, in this work, the observation model of CILDS is based on OASIS [34], which uses an autoregressive process to model the calcium decay. Here, we used an autoregressive order of one, which corresponds to an instantaneous rise in calcium after each spiking event, as was done in previous work [32, 55, 34]. Although this is a reasonable first approximation, when imaging rates are fast or the calcium indicator is slow, it may be desirable to use an autoregressive order greater than one to better capture the non-instantaneous rise in calcium [13]. Second, various deconvolution methods have been proposed, including MLSpike [56], and LZero [57]. Any innovations to the deconvolution methods beyond OASIS could, in principle, also be incorporated into the observation model of CILDS. Finally, different latent time series models can be used in the place of the linear dynamical system to achieve different analysis goals. For example, if one seeks only temporal smoothing in the latent variables without explicit dynamics, the linear dynamical system can be replaced with Gaussian processes [38]. If one seeks to extract richer dynamics, a nonlinear dynamical system such as a recurrent neural network[58, 36] can be used.

With the development of dimensionality reduction methods that are tailored for calcium imaging, such as CILDS, we can better leverage the statistical power of the populations of neurons recorded with calcium imaging. In addition, we can incorporate other advantages of calcium imaging, such as being able to obtain information about neuron type or knowledge about which neurons project to other brain areas. For example, one could use dimensionality reduction to understand how populations of different neuron types interact [59], or utilize information about where neurons project, coupled with dimensionality reduction, to understand how the projections contribute to the coordination of activity between brain areas [60, 61]. This can enable insights about neural population activity recorded using calcium imaging that go beyond what is currently possible with electrophysiology.

Methods

Dimensionality reduction methods

Here we describe mathematically the dimensionality reduction methods used in this work: LDS, deconv-LDS, CILDS, and CIFA. For the purposes of this work, we assume that the spatial footprint of each neuron has already been identified from the raw calcium imaging data (a procedure known as image segmentation[62, 55]), resulting in a fluorescence time course for each neuron. The dimensionality reduction methods presented here are applied to these fluorescence time courses.

Linear Dynamical System (LDS)

We first considered a standard dimensionality reduction method for summarizing the time course of spiking activity, a latent Linear Dynamical System (LDS), here applied to calcium imaging recordings. Let $y_t \in \mathbb{R}^{q \times 1}$ be a high-dimensional vector of fluorescence values recorded at time point t, where q is the number of neurons imaged simultaneously. The goal

is to extract a corresponding low-dimensional latent variable $z_t \in \mathbb{R}^{p \times 1}$ at each time point, where p is the number of latent dimensions ($p < q$). The observation model is

$$y_t = Az_t + b + \epsilon_t, \quad w_t \sim \mathcal{N}(0, R) \tag{1}$$

where $A \in \mathbb{R}^{q \times q}$ is the loading matrix that specifies how each neuron's activity is related to the latent variables, $b \in \mathbb{R}^{q \times 1}$ is an offset vector that accounts for constant background fluorescence, $R \in \mathbb{R}^{q \times q}$ is the observation noise covariance, and $t = 1, ..., T$. We constrained R to be diagonal, thereby capturing activity variability and imaging noise independent to each neuron. The time-evolution of the latent variables is described as a linear dynamical system

$$z_t = Dz_{t-1} + v_t, \quad v_t \sim \mathcal{N}(0, P) \tag{2}$$

$$z_1 \sim \mathcal{N}(h_1, G_1) \tag{3}$$

where $D \in \mathbb{R}^{p \times p}$ is the dynamics matrix that determines the timescale of the latent variables, $P \in \mathbb{R}^{p \times p}$ is the dynamics noise covariance, $h_1 \in \mathbb{R}^{p \times 1}$ and $G_1 \in \mathbb{R}^{p \times p}$ describe the mean and covariance of the latent variable at the first time point, and $t = 2, ..., T$. We constrained D, P, and G_1 here to be diagonal as a form of regularization, although a general LDS with these parameters unconstrained could also be used.

Equations (1), (2), and (3) together define the latent LDS. We fit the model parameters (A, b, R, D, P, h_1, G_1) using the expectation-maximization (EM) algorithm. To initialize the model parameters, we first performed factor analysis (FA) on y_t to obtain A, b, and R. We initialized D to be $0.999I$ (a stable system), which we found to work well in practice in the simulations. We ran the EM algorithm until convergence (defined as a log data likelihood increase of $< 10^{-6}$ or 1500 iterations, whichever came first). This maximum number of iterations was chosen heuristically, by noting empirically that the latent variables do not change significantly beyond this point.

Deconvolution - Linear Dynamical System (deconv-LDS)

Since fluorescence traces are an indirect measure of spiking activity, we also considered a two-stage approach, whereby we first deconvolve the fluorescence traces one neuron at a time to estimate the underlying spiking activity. We then applied a standard dimensionality reduction method, in this case LDS, to those deconvolved estimates. We refer to this two-stage method as deconv-LDS.

For the deconvolution stage of deconv-LDS, we used the Online Active Set method to Infer Spikes (OASIS), developed by Friedrich et al. [34], using their L_1 regularization. We also tested their L_0 regularization and found the L_1 regularization to work better for our datasets. As per Friedrich et al., this first order autoregressive model for OASIS is described for each neuron as

$$y_t = ac_t + b + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \sigma^2) \tag{4}$$

$$c_t = \gamma c_{t-1} + s_t \tag{5}$$

where y_t is the recorded fluorescence at at time t, c_t represents the calcium concentration at time t, ϵ_t captures imaging noise independent of the calcium and spiking activity, and s_t is the spiking activity. The parameter a relates the calcium concentration to fluorescence, b accounts for the baseline fluorescence, σ^2 captures the variance of the imaging noise, and γ specifies how quickly the calcium trace decays, which depends on the calcium indicator. Additionally, there is a hyperparameter in the OASIS model, minimum spike size, that sets the minimum value of s_t that would be identified. In this model, all variables are scalars and a is constrained to be non-negative. We initialized OASIS with γ values that are typical for the calcium indicators used [13] (see Supplementary Table 1), and allowed OASIS to optimize a, b, γ, σ^2 and the minimum spike size.

Applying deconvolution to the recorded fluorescence traces returned estimates of the time course of spiking activity for each neuron, s_t. We then used the estimated spiking activity of all the neurons as the observations $y_t \in \mathbb{R}^{q \times 1}$ for time points $t = 1, ..., T$ in the LDS model defined in equations (1)–(3).

Calcium Imaging Linear Dynamical System (CILDS)

CILDS unifies the approaches described above by allowing estimates of shared activity among neurons (i.e., the latent variables) to influence the estimates of deconvolved spiking activity, and vice versa. In other words CILDS performs deconvolution for all neurons and dimensionality reduction jointly, in a unified framework. This is in contrast to deconv-LDS, which deconvolves the activity of each neuron independently. With low-dimensional latent variables that are jointly estimated with the model of calcium decay, CILDS is better able to peer through the calcium decay to more clearly identify the shared activity among neurons, as compared to deconv-LDS and LDS applied directly on fluorescence.

Let $y_t \in \mathbb{R}^{q \times 1}$ be the high-dimensional vector of fluorescence traces recorded at each time point t, where q is the number of neurons imaged simultaneously. The goal is to extract a corresponding low-dimensional latent variable $z_t \in \mathbb{R}^{p \times 1}$ at each time point, where p is the number of latent dimensions ($p < q$). The observation model follows the multivariate form of equation (4) which was used for deconvolution

$$y_t = Bc_t + \epsilon_t, \quad v_t \sim \mathcal{N}(\mathbf{0}, R) \tag{6}$$

where $B \in \mathbb{R}^{q \times q}$ maps the calcium concentration to the recorded fluorescence, $R \in \mathbb{R}^{q \times q}$ is the fluorescence noise covariance, and $t = 1, ..., T$. We constrained B and R to be diagonal to allow each dimension of c_t to represent the calcium concentration of one neuron. B accounts for all experimental variables influencing the scale of the signal from each neuron, such as the amplification of the imaging system [32], and R accounts for fluorescence fluctuations

independent of calcium concentration. Here we omit the additive offset found in equation (4) without loss of generality due to the offset included in equation (7).

Similar to equation (5), the calcium decay for each neuron is described using a first order autoregressive model

$$c_t = \Gamma c_{t-1} + A z_t + b + w_t, \quad w_t \sim \mathcal{N}(\mathbf{0}, Q) \tag{7}$$

$$c_1 \sim \mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{V}_1) \tag{8}$$

where $\Gamma \in \mathbb{R}^{q \times q}$ captures the calcium decay, $A \in \mathbb{R}^{q \times p}$ is the loading matrix that describes how latent variable maps to calcium concentrations, $b \in \mathbb{R}^{q \times 1}$ is a constant vector, $Q \in \mathbb{R}^{q \times q}$ captures the spiking variability independent to each neuron, $\boldsymbol{\mu}_1 \in \mathbb{R}^{q \times 1}$ and $V_1 \in \mathbb{R}^{q \times q}$ describe the mean and variance of the calcium concentration at the first time point, respectively, and $t = 2, ..., T$. Our key innovation is to replace the spiking activity s_t from equation (5) with $A z_t + b + w_t$. By analogy to factor analysis, $A z_t$ describes the shared activity among neurons, and w_t describes the activity independent to each neuron. We constrain Γ, Q, and V_1 to be diagonal to prevent intermixing among neurons outside of the loading matrix. The form of Γ allows each neuron to have a different calcium decay as determined by the fitting procedure (see below), which can depend on the extent of calcium buffering within a cell and the calcium indicator used [32, 13, 34].

Similar to LDS, the low-dimensional latent variables z_t evolve over time according to a linear dynamical system

$$z_t = D z_{t-1} + v_t, \quad v_t \sim \mathcal{N}(\mathbf{0}, P) \tag{9}$$

$$z_2 \sim \mathcal{N}(h_2, G_2) \tag{10}$$

where $D \in \mathbb{R}^{p \times p}$ is the dynamics matrix, $P \in \mathbb{R}^{p \times p}$ is the dynamics noise covariance, $h_2 \in \mathbb{R}^{p \times 1}$ and $G_2 \in \mathbb{R}^{p \times p}$ are the mean and covariance of the latent variable at the first time point respectively, and $t = 3, ..., T$. We constrained D, P, and G_2 here to be diagonal as a form of regularization, although a model with these parameters unconstrained could also be used. Note that according to equation (7), z_2 is the first latent variable in the time series (i.e., there is no z_1).

Equations (6) – (10) define CILDS. The joint estimation of the parameters B, R, Γ, A, b, $Q, \boldsymbol{\mu}_1$, V_1, D, P, h_2, and G_2 allows CILDS to leverage the entire recorded neural population to perform deconvolution and estimate latent variables in a unified fashion. CILDS can be viewed as a special case of the standard LDS, where the parameters are constrained in a specific way. We fit CILDS using the EM algorithm, initialized using deconv-LDS run for 100 EM iterations. The EM algorithm was run until convergence, defined as a log data likelihood increase of $< 10^{-6}$ or 1500 iterations, whichever came first. The maximum number of iterations was chosen by noting empirically that the latent variables do not change

substantially beyond this point. The EM equations for CILDS are provided in Supplemental Information.

Calcium Imaging Factor Analysis (CIFA)

To evaluate the importance of incorporating a latent dynamical system in dimensionality reduction methods for calcium imaging, we developed a method (CIFA) identical to CILDS, with the only difference being that CIFA does not enforce latent dynamics. Like CILDS, CIFA also uses equations (6), (7), and (8). The only difference between CIFA and CILDS is that we replaced the latent dynamical system equations (9) and (10) with

$$z_t \sim \mathcal{N}(\mathbf{0}, I) \qquad (11)$$

for $t = 2, \ldots, T$. In other words, CIFA defines latent variables that are independent across time, whereas CILDS defines latent variables that evolve smoothly over time. Like CILDS, we fit CIFA using the EM algorithm, initialized using parameters from deconv-FA run for 100 EM iterations. The EM algorithm was run until convergence, defined as a log data likelihood increase of $< 10^{-6}$ or 1500 iterations, whichever came first. The EM equations for CIFA are the same as for CILDS (see Supplemental Information), without the equations estimating D, P, $\boldsymbol{h}_2$, G_2. Note that there is no loss of generality by setting the prior distribution of z_t to $\mathcal{N}(\mathbf{0}, I)$, compared to a general Gaussian distribution. Additionally, although CIFA has FA in its name, there is one key difference from FA. Whereas FA can be fit on data with no concept of time, CIFA requires a time series for deconvolution (equations (7) and (8)).

Simulation framework

We created a framework to simulate fluorescence recordings from calcium imaging for two reasons. First, in calcium imaging recordings, the ground truth latent variables are unknown. A simulation of fluorescence traces from known latent variables allows us to directly evaluate our dimensionality reduction methods by comparing the estimated latent variables with the ground truth latent variables. Second, we wanted a simulation framework in which we could systematically vary various experimentally relevant parameters to see their effects on the estimated latent variables. Specifically, we evaluated our ability to recover the ground truth latent variables as a function of the timescale of the latent variables, the number of neurons, the calcium decay rate, and the size of the fluorescence noise. The simulation procedure consists of first generating fluorescence traces from known latent variables while varying the experimentally relevant parameters listed above. Then, we applied each dimensionality reduction method to estimate latent variables from the simulated fluorescence traces. The estimated latent variables were then compared to the ground truth latent variables. The steps of this procedure are detailed below.

Generating fluorescence traces

To simulate fluorescence traces, we first drew p Gaussian processes (GP), where each GP has T time points at 1 ms time resolution. We denote the ith GP as $z_{i,:} \in \mathbb{R}^{1 \times T}$, where $i = 1, \ldots, p$. The GP allows us to specify the covariance $K_i \in \mathbb{R}^{T \times T}$ for the ith GP across the T time points as

$$\mathbf{z}_{i,:} \sim \mathcal{N}(\mathbf{0}, K_i), \text{ where the } (t_1, t_2) \text{ element of } K_i \text{ is} \tag{12}$$

$$K_i(t_1, t_2) = \sigma_{f,i}^2 \exp(-\frac{(t_1 - t_2)^2}{2\tau_i^2}) + \sigma_{n,i}^2 \cdot \delta_{t_1, t_2} \tag{13}$$

and $t_1, t_2 = 1, ..., T$. Here we chose the commonly used squared exponential covariance function. The squared exponential covariance is defined by its signal variance $\sigma_{f,i}^2 \in \mathbb{R}_+$, characteristic timescale $\tau_i \in \mathbb{R}_+$, and noise variance $\sigma_{n,i}^2 \in \mathbb{R}_+$. The Kroneker delta δ_{t_1, t_2} equals 1 if $t_1 = t_2$ and 0 otherwise. We set $\sigma_{f,i}^2 = 1 - \sigma_{n,i}^2$ so that the latent variable $\mathbf{z}_t \in \mathbb{R}^{p \times 1}$ at every time point has mean $\mathbf{0}$ and a variance I. The vector $\mathbf{z}_t$ comprises the tth time point from each of the p Gaussian processes. The noise variance $\sigma_{n,i}^2$ must be nonzero to ensure that K_i is invertible, hence for practical purposes we set $\sigma_{n,i}^2 = 10^{-9}$. We note that an LDS can also be used to introduce latent dynamics, but we chose to use a GP due to the ease by which we can generate stationary time series with specified time scales. Furthermore, using a GP introduced model mismatch for all of the dimensionality reduction methods, which enables a more meaningful comparison across methods.

Next, we projected the low-dimensional latent variables $\mathbf{z}_t$ into the high-dimensional neural space to obtain neural firing rates for each of q neurons at each time point $t = 1, ..., T$. We then imposed a rectifying nonlinearity applied element-by-element using $\log(1 + \exp(W\mathbf{z}_t + \boldsymbol{\mu}))$, where $W \in \mathbb{R}^{q \times p}$ is the loading matrix and $\boldsymbol{\mu} \in \mathbb{R}^{q \times 1}$ is a constant offset, to ensure that firing rates are non-negative. We generated binary spikes $\mathbf{s}_t \in \mathbb{R}^{q \times 1}$ at each time point using an inhomogeneous Poisson process with time-varying rates defined by the output of this rectifying nonlinearity. This allows us to assess how the dimensionality reduction methods, which assume Gaussian noise for mathematical tractability, would perform on realistic spiking variability that is not Gaussian. Finally, we obtained the calcium concentration $\mathbf{c}_t \in \mathbb{R}^{q \times 1}$ using a first order autoregressive model and fluorescence $\mathbf{y}_t \in \mathbb{R}^{q \times 1}$, as in Friedrich et al. 2017 [34]

$$\mathbf{c}_t = \Gamma \mathbf{c}_{t-1} + \mathbf{s}_t \tag{14}$$

$$\mathbf{y}_t = B\mathbf{c}_t + \mathbf{b} + \boldsymbol{\epsilon}_t, \boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, R) \tag{15}$$

where $\Gamma \in \mathbb{R}^{q \times q}$ determines how quickly the calcium decays, $B \in \mathbb{R}^{q \times q}$ relates the calcium concentration to the fluorescence, $\mathbf{b} \in \mathbb{R}^{q \times 1}$ is the baseline fluorescence, $R \in \mathbb{R}^{q \times q}$ is the imaging noise covariance, and $t = 1, ..., T$, at the same (1ms) resolution as the GP. We set Γ, B, and R to be diagonal. We specified Γ such that the decay constants approximately matched the decay constants of GCaMP6f, GCaMP6m, and GCaMP6s, found in Chen et al. [13] (See Supplementary Table 1). B represents experimental variables influencing the scale of the calcium signal of each neuron, such as the amplification of the imaging system. B is set as the identity matrix and $\mathbf{b}$ is set to be $\mathbf{0}$ in our simulations. We varied the signal-to-noise

ratio by varying R to simulate low, medium, and high noise regimes (Supplementary Table 1).

For each simulation run, we simulated $p = 10$ latent variables, where each latent dimension had the same latent timescale for ease of interpretation. Across different settings of experimentally relevant parameters, we explored a range of timescales $\tau \in \{50, 100, 200, 1000, 2000, 5000\}$ ms. To define loading matrices W that were realistic for spiking activity, we used electrophysiological recordings with 94 neurons [6]. Across runs, we tested $q \in \{20, 50, 94\}$ neurons. For the $q = 20$ and $q = 50$ cases, we randomly subsampled the electrophysiological recordings. To avoid a start-up transient at the start of every trial from the calcium model, we generated two long fluorescence traces for each neuron, each 6,000,000 time points long (1 ms resolution). One fluorescence trace was used for training, and the other for testing. We divided each trace into 100 trials. Each trial was 60 s long, comprising 60,000 time points. Finally, having generated these fluorescence traces at 1 kHz, we down-sampled the fluorescence rate to a more typical imaging rate of 40 Hz (2,400 time points per trial). The simulation parameters are summarized in Supplementary Table 1. See Practical Considerations below for how we generated GPs with 6, 000, 000 time points within the computer's memory constraints.

Evaluation of dimensionality reduction methods

We applied the four dimensionality reduction methods to the simulated fluorescence traces to evaluate how accurately each method reconstructed the ground truth latent variables. Rather than simulating fluorescence traces directly from the generative models of the dimensionality reduction methods, we simulated fluorescence traces in a way that introduced model mismatch, which is present when applying any statistical method to real-world data. We first generated the latent variables using Gaussian processes, rather than using an LDS. Then, we simulated spike trains based on the latent variables using an inhomogeneous Poisson process, rather than adding Gaussian noise to the firing rates. We believe the presence of model mismatch strengthens our results with simulated data.

We used two-fold cross validation in estimate the accuracy of latent variable recovery, so that the model parameters were fit using half of the data (100 training trials, each with 2,400 time points) and those parameters were then used to estimate the latent variables in the other half of the data (100 held-out trials). For all methods, the latent variables are only unique up to an arbitrary linear transformation. In order to compare the estimated and ground truth latent variables, we aligned them using the following procedure. First we split the estimated latent variables for each test fold into two further inner halves. We concatenated the trials over time, such that the estimated latent variables are defined as $\tilde{Z} \in \mathbb{R}^{p \times \tilde{T}}$, and the ground truth latent variables are defined as $Z \in \mathbb{R}^{p \times \tilde{T}}$, where $p = 10$ is the dimensionality of the latent variables and $\tilde{T} = 120,000$ for one inner half. We then applied linear regression to relate $\tilde{Z}$ and the Z corresponding to that inner half. This yielded a transformation matrix W, where $W = (Z\tilde{Z}')(\tilde{Z}\tilde{Z}')^{-1}$. Finally we used W to align the $\tilde{Z}$ corresponding to the other inner half to obtain the transformed estimated latent variables $\hat{Z} = W\tilde{Z}$, where $\hat{Z} \in \mathbb{R}^{p \times \tilde{T}}$. We performed the same procedure to align each inner half of each cross-validation fold. In

Figs. 3 and 4, we provide each dimensionality reduction method with the number of ground truth latent variables ($p = 10$) to be extracted. We also performed an analysis where we varied the number of extracted latent variables for a fixed number of ground truth latent variables (Supplementary Fig. 10).

We computed the accuracy of the ith transformed estimated latent variables

$$R^2 = 1 - \frac{\sum_{t=1}^{\tilde{T}} (z_t^{(i)} - \hat{z}_t^{(i)})^2}{\sum_{t=1}^{\tilde{T}} (z_t^{(i)} - \bar{z}^{(i)})^2} \tag{16}$$

where $z_t^{(i)} \in \mathbb{R}$ is the (i, t) entry of Z, and $\hat{z}_t^{(i)} \in \mathbb{R}$ is the (i, t) entry of $\hat{Z}$. The mean of the ith ground truth latent variable across time is defined as $\bar{z}^{(i)} \in \mathbb{R}$, where $i = 1, \ldots, p$. A larger R^2 means a better match between the estimated latent variables and the ground truth latent variables, based on the proportion of total variance (of the ground truth latent variables) explained. R^2 has an upper limit of 1, and any value below 0 indicates that the estimate is poorer than the using the ground truth mean. We repeated this process for each cross-validation fold and averaged the results across all p latent dimensions and folds. It is important to note that since the results are cross-validated, a dimensionality reduction method with more parameters will not necessarily outperform a method with fewer parameters.

Experimental data

Larval Zebrafish

Neurons were imaged from the dorsal raphe nucleus (DRN) of larval zebrafish while they engaged in a "fictive swimming" motosensory gain adaptation task [40]. Calcium imaging was performed using light-sheet microscopy at 30 Hz on a single plane of narrow area around the DRN. This was performed with $Tg(elavl3:GCaMP6f)^{jf1}$ fish expressing GCaMP6f in the cytosol [63]. We analyzed baseline corrected fluorescence recordings from three fish, the same data used in Kawashima et al. 2016[40]. In the task, the fish underwent an initialization period of 20 seconds to increase locomotor drive, a training period in which the fish attenuated their locomotor drive, and a delay period of 10 seconds which stopped the fish from swimming. Finally, there was a test period of 5 seconds to probe the extent to which the attenuated locomotor drive persisted throughout the delay period. There were three different training period lengths of 7, 15, or 30 seconds. Here we combined the data from the different training periods. We included only the 20-second initialization period and the first 7 seconds of the training period of each trial. Thus, the analyzed portion of each trial is nominally identical. For each fish, this yielded 15 trials, each 27 seconds long. We analysed 22, 19, and 19 neurons imaged from the DRN of each fish, respectively. All experiments presented in this study were conducted according to the animal research guidelines from NIH and were approved by the Institutional Animal Care and Use Committee and Institutional Biosafety Committee of Janelia Research Campus.

Mouse

Two-photon calcium imaging was performed in the binocular zone of V1 in awake head-fixed mice resting atop a floating spherical treadmill [41]. GCaMP6f was expressed in excitatory neurons and imaged at 15.5 Hz. We analyzed recordings from three mice (2 male, 1 female, age 42-80 days). These were homozygous Emx1cre mice (Jackson Laboratories, stock number 005628) crossed with homozygous Ai93/heterozygous Camk2a-tTA mice (Jackson Laboratories, stock number 024108). Experimental mice were heterozygous for all three alleles. Mice were positioned to passively view static sinusoidal gratings, without reward. There were 180 gratings presented, comprising 12 different orientations equally spaced with range $\{0 - 165\}°$ and 15 different spatial frequencies equally spaced with range $\{0.02 - 0.30\}$ cycles/°. Each "trial" was a 196.7 seconds long recording (3049 time points), comprising 4 presentations of each of the 180 possible gratings in random order. Each presentation lasted 250 ms without an intervening grey screen. The onset time of the first stimulus relative to the beginning of the recording was varied. This means that there is a short period of time recorded before the first stimulus is shown, and a short period of time recorded after the last stimulus is shown. The experiment comprised 15 trials for each mouse. We analysed 133, 252, and 319 neurons from V1 of mouse (labeled mouse 1-3, respectively). These mice correspond to mouse2317, mouse2320, and mouse2209 in the experiments. All experimental procedures were compliant with the guidelines established by the Institutional Animal Care and Use Committee of Carnegie Mellon University and the National Institute of Health, and all experimental protocols were approved by the Institutional Animal Care and Use Committee of Carnegie Mellon University.

Data analysis

Leave-neuron-out fluorescence prediction

We sought to compare the four dimensionality reduction methods using experimental data. In experimental data, ground truth latent variables are unknown, and so we could not use the same evaluation procedure of comparing estimated and ground truth latent variables as in the simulations. Furthermore, the cross-validated data likelihoods are not comparable across all methods. Hence to compare the four methods, we performed a leave-neuron-out fluorescence prediction test to determine which method best summarizes the neuronal activity with low-dimensional latent variables [38, 39]. The intuition is that a method that provides a better summary of the population activity using the latent variables would be better able to predict the activity of held-out neuronal fluorescence traces.

For the leave-neuron-out fluorescence prediction, we performed 5-fold cross-validation. We first split the trials into five equal-sized folds. For a given dimensionality reduction method, we fit the model parameters using four of the folds. With the remaining validation fold, we estimated the latent variables using all but one neuron, and then predicted the activity of that held-out neuron using those estimated latent variables. We did this for every neuron in the validation fold. In the same manner, we performed leave-neuron-out fluorescence predictions for the remaining folds and thus obtained predictions of the fluorescence activity for all the neurons at all time points.

After predicting the leave-neuron-out fluorescence for each neuron using each dimensionality reduction method, we computed the Pearson's correlation coefficient between the predicted fluorescence and the ground truth fluorescence for every neuron and every trial. For each neuron, we then use the correlation values obtained for all the trials to apply a paired two-tailed t-test (across 15 values per neuron for the larval zebrafish DRN recordings, and 10 values per neuron for the mouse V1 recordings). We indicated the neurons for which the difference in leave-neuron-out fluorescence prediction correlation values between CILDS and each of the other methods was statistically significant ($p < 0.05$; Figs. 5, 6, Supplementary Figs. 5, 6, 8). For the fish (Fig. 5), we found that CILDS outperforms LDS (59% of red points above the diagonal), deconv-LDS (100% of red points above the diagonal), and CIFA (93% of red points above the diagonal). For the mice (Fig 6), we found that CILDS outperforms LDS (68% of red points above the diagonal), deconv-LDS (84% of red points above the diagonal), and CIFA (99% of red points above the diagonal).

Performing the leave-neuron-out prediction procedure requires selecting the latent dimensionality p for each method. The classic way of finding the optimal latent dimensionality is through nested cross-validation. Nested cross-validation uses, in this case, the 5-fold cross-validation described earlier as the outer folds. Within each outer fold, the training portions are then used to perform an inner 4-fold cross-validation to determine the latent dimensionality that maximizes the cross-validated data likelihood. We used this latent dimensionality to estimate the model parameters using the same four training folds. These model parameters were then used for the leave-neuron-out prediction procedure of the remaining (validation) outer fold. Note that although the cross-validated data likelihood is not comparable across methods, it is comparable across different latent dimensionalities for the same method. To illustrate the model selection process, we show how the cross-validated data likelihood varies with dimensionality for both fish and mice, which is representative of how the cross-validation procedure is carried out for each outer fold (Supplementary Fig. 11).

For the larval zebrafish DRN recordings, we selected the latent dimensionality for each dimensionality reduction method (LDS: 17, deconv-LDS: 17, CILDS: 7 to 17, CIFA: 3 to 5) using nested cross-validation, as described above. For the mouse V1 recordings, nested cross-validation would have required us to fit the model 25 times, for which the running time would have been prohibitively long. For this reason, we selected dimensionality for the mouse V1 leave-neuron-out analysis by assessing the cross-validated decoding performance. We found that the performance of all dimensionality reduction methods increased with latent dimensionality across the range of dimensionalities tested (5-50). Thus we used a latent dimensionality of 50 for all methods for cross-validated leave-neuron-out prediction in mouse V1.

Decoding analysis

Another way to assess how meaningful the extracted latent variables are is by decoding external variables from the latent variables [42]. For the larval zebrafish recordings, there is no moment by moment behavior that can be decoded from the neural activity. In the mouse

recordings, we can decode the orientation and spatial frequency of the grating stimuli. To begin, we applied the dimensionality reduction methods to the mouse recordings. We then applied a linear Gaussian Naïve Bayes classifier [64] to the extracted latent variables

$$P(z|C_k) = \frac{1}{(2\pi)^{p/2}|\sum|^{1/2}}\exp[-(z - \mu_k)'\sum^{-1}(z - \mu_k)] \tag{17}$$

where $z \in \mathbb{R}^{p \times 1}$ is the latent variable averaged across the time points in a 250 ms window corresponding to a given stimulus C_k, $k = 1, …, 180$. The parameters of the classifier are μ_k, the mean of the latent variables corresponding to class k, and $\sum$, the covariance of the latent variables across trials. We constrain $\sum$ to be diagonal and the same across all classes. The parameters μ_k and $\sum$ are fit by maximizing the likelihood of the training data using 5-fold cross-validation.

We then used the parameters found from the training data to classify the held-out data according to

$$\hat{y} = \arg\max_{y} P(C_k = y|z) \tag{18}$$

where $\hat{y}$ is the predicted class label (1,…,180). We then computed the accuracy of the predicted class labels against the true class labels (chance level is $\frac{1}{180}$). A higher accuracy indicates that the latent variables are better at capturing the shared modulations among neurons that are relevant to the visual stimulus. For latent dimensionalities {5, 10, 20, 30, 40, 50}, we fit each dimensionality reduction method using all data. Then, we performed 5-fold cross-validation in the decoding stage.

To decode visual stimuli, there are two considerations. First, visual information takes time to arrive in the visual cortex, hence there is a need to shift the window of neural activity relative to the stimulus presentation. Second, there is an additional time delay introduced by the calcium indicators. Unlike deconv-LDS and CILDS, LDS does not attempt to remove the calcium decay. Hence, we expected a longer latency for LDS than for deconv-LDS and CILDS. To determine the appropriate window, we considered a range of time lags and evaluated their cross-validated classification accuracy. We found that the best cross-validated accuracy was obtained for a 4 time point (260 ms) shift for LDS, and a 3 time point (190 ms) shift for deconv-LDS and CILDS. Thus we used these time lags to report the classification accuracy of our decoding analysis (Fig. 6g).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Data Availability

Larval zebrafish DRN data are available at https://doi.org/10.6084/m9.figshare.21646682 (ref. [65]). Mouse V1 data are available at https://doi.org/10.12751/g-node.wc3f8g (ref. [66]) Source Data for Figs. 3–6 are available for this Article.

References

[1]. Cunningham JP & Yu BM Dimensionality reduction for large-scale neural recordings. Nature Neuroscience 17, 1500–1509, 10.1038/nn.3776 (2014). [PubMed: 25151264]

[2]. Gallego JA, Perich MG, Miller LE & Solla SA Neural manifolds for the control of movement. Neuron 94, 978–984, 10.1016/j.neuron.2017.05.025 (2017). [PubMed: 28595054]

[3]. Gao P & Ganguli S On simplicity and complexity in the brave new world of large-scale neuroscience. Current Opinion in Neurobiology 32, 148–155, 10.1016/jxonb.2015.04.003 (2015). Large-Scale Recording Technology (32). [PubMed: 25932978]

[4]. Mante V, Sussillo D, Shenoy KV & Newsome WT Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature 503, 78–84, 10.1038/nature12742 (2013). [PubMed: 24201281]

[5]. Churchland MM et al. Neural population dynamics during reaching. Nature 487, 51–56, 10.1038/nature11129 (2012). [PubMed: 22722855]

[6]. Sadtler PT et al. Neural constraints on learning. Nature 512, 423–426, 10.1038/nature13665 (2014). [PubMed: 25164754]

[7]. Murray JD et al. Stable population coding for working memory coexists with heterogeneous neural dynamics in prefrontal cortex. Proceedings of the National Academy of Sciences 114, 394–399, 10.1073/pnas.1619449114 (2017). https://www.pnas.org/content/114/2/394.full.pdf.

[8]. Wang J, Narain D, Hosseini EA & Jazayeri M Flexible timing by temporal scaling of cortical responses. Nature Neuroscience 21, 102–110, 10.1038/s41593-017-0028-6 (2018). [PubMed: 29203897]

[9]. Cohen MR & Maunsell JHR A neuronal population measure of attention predicts behavioral performance on individual trials. Journal of Neuroscience 30, 15241–15253, 10.1523/JNEUROSCI.2171-10.2010 (2010). https://www.jneurosci.org/content/30/45/15241.full.pdf. [PubMed: 21068329]

[10]. Mazor O & Laurent G Transient dynamics versus fixed points in odor representations by locust antennal lobe projection neurons. Neuron 48, 661–673, 10.1016/j.neuron.2005.09.032 (2005). [PubMed: 16301181]

[11]. Bouchard KE, Mesgarani N, Johnson K & Chang EF Functional organization of human sensorimotor cortex for speech articulation. Nature 495, 327–332, 10.1038/nature11911 (2013). [PubMed: 23426266]

[12]. Peron S, Chen T-W & Svoboda K Comprehensive imaging of cortical networks. Current Opinion in Neurobiology 32, 115–123, 10.1016/j.conb.2015.03.016 (2015). Large-Scale Recording Technology (32). [PubMed: 25880117]

[13]. Chen T-W et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature 499, 295–300, 10.1038/nature12354 (2013). [PubMed: 23868258]

[14]. Ahrens MB, Orger MB, Robson DN, Li JM & Keller PJ Whole-brain functional imaging at cellular resolution using light-sheet microscopy. Nature Methods 10, 413–420, 10.1038/nmeth.2434 (2013). [PubMed: 23524393]

[15]. Briggman K, Abarbanel H & Kristan W Optical imaging of neuronal populations during decision-making. Science (New York, N.Y.) 307, 896–901, 10.1126/science.1103736 (2005). [PubMed: 15705844]

[16]. Harvey CD, Coen P & Tank DW Choice-specific sequences in parietal cortex during a virtual-navigation decision task. Nature 484, 62–68, 10.1038/nature10918 (2012). [PubMed: 22419153]

[17]. Ahrens MB et al. Brain-wide neuronal dynamics during motor adaptation in zebrafish. Nature 485, 471–477, 10.1038/nature11057 (2012). [PubMed: 22622571]

[18]. Kato H, Chu M, Isaacson J & Komiyama T Dynamic sensory representations in the olfactory bulb: modulation by wakefulness and experience. Neuron 76, 962–975, 10.1016/j.neuron.2012.09.037 (2012). [PubMed: 23217744]

[19]. Daie K, Goldman M & Aksay E Spatial patterns of persistent neural activity vary with the behavioral context of short-term memory. Neuron 85, 847–860, 10.1016/j.neuron.2015.01.006 (2015). [PubMed: 25661184]

[20]. Morcos AS & Harvey CD History-dependent variability in population dynamics during evidence accumulation in cortex. Nature Neuroscience 19, 1672–1681, 10.1038/nn.4403 (2016). [PubMed: 27694990]

[21]. Makino H et al. Transformation of cortex-wide emergent properties during motor learning. Neuron 94, 880–890.e8, 10.1016/j.neuron.2017.04.015 (2017). [PubMed: 28521138]

[22]. Driscoll LN, Pettit NL, Minderer M, Chettih SN & Harvey CD Dynamic reorganization of neuronal activity patterns in parietal cortex. Cell 170, 986–999.e16, 10.1016/j.cell.2017.07.021 (2017). [PubMed: 28823559]

[23]. Stringer C, Pachitariu M, Steinmetz N, Carandini M & Harris KD High-dimensional geometry of population responses in visual cortex. Nature 571, 361–365, 10.1038/s41586-019-1346-5 (2019). [PubMed: 31243367]

[24]. Rumyantsev OI et al. Fundamental bounds on the fidelity of sensory cortical coding. Nature 580, 100–105, 10.1038/s41586-020-2130-2 (2020). [PubMed: 32238928]

[25]. Pashkovski SL et al. Structure and flexibility in cortical representations of odour space. Nature 583, 253–258, 10.1038/s41586-020-2451-1 (2020). [PubMed: 32612230]

[26]. Nieh E et al. Geometry of abstract learned knowledge in the hippocampus. Nature 595, 80–84, 10.1038/s41586-021-03652-7 (2021). [PubMed: 34135512]

[27]. Wei Z et al. A comparison of neuronal population dynamics measured with calcium imaging and electrophysiology. PLOS Computational Biology 16, e1008198, 10.1371/journal.pcbi.1008198 (2020). [PubMed: 32931495]

[28]. Afshar A et al. Single-trial neural correlates of arm movement preparation. Neuron 71, 555–564, 10.1016/j.neuron.2011.05.047 (2011). [PubMed: 21835350]

[29]. Kiani R, Cueva CJ, Reppas JB & Newsome WT Dynamics of neural population responses in prefrontal cortex indicate changes of mind on single trials. Current Biology 24, 1542–1547, 10.1016/j.cub.2014.05.049 (2014). [PubMed: 24954050]

[30]. Gouvêa TS et al. Striatal dynamics explain duration judgments. eLife 4, e11386, 10.7554/eLife.11386 (2015). [PubMed: 26641377]

[31]. Kaufman MT, Churchland MM, Ryu SI & Shenoy KV Vacillation, indecision and hesitation in moment-by-moment decoding of monkey motor cortex. eLife 4, e04677, 10.7554/eLife.04677 (2015). [PubMed: 25942352]

[32]. Vogelstein JT et al. Fast nonnegative deconvolution for spike train inference from population calcium imaging. Journal of Neurophysiology 104, 3691–3704, 10.1152/jn.01073.2009 (2010). 10.1152/jn.01073.2009. [PubMed: 20554834]

[33]. Theis L et al. Benchmarking spike rate inference in population calcium imaging. Neuron 90, 471–482, 10.1016/j.neuron.2016.04.014 (2016). [PubMed: 27151639]

[34]. Friedrich J, Zhou P & Paninski L Fast online deconvolution of calcium imaging data. PLOS Computational Biology 13, e1005423, 10.1371/journal.pcbi.1005423 (2017). [PubMed: 28291787]

[35]. Runyan CA, Piasini E, Panzeri S & Harvey CD Distinct timescales of population coding across cortex. Nature 548, 92–96, 10.1038/nature23020 (2017). [PubMed: 28723889]

[36]. Zhu F et al. A deep learning framework for inference of single-trial neural population dynamics from calcium imaging with subframe temporal resolution. Nature Neuroscience 25, 1724–1734, 10.1038/s41593-022-01189-0 (2022). [PubMed: 36424431]

[37]. Murray JD et al. A hierarchy of intrinsic timescales across primate cortex. Nature Neuroscience 17, 1661–1663, 10.1038/nn.3862 (2014). [PubMed: 25383900]

[38]. Yu BM et al. Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. Journal of neurophysiology 102, 614–635, 10.1152/jn.90941.2008 (2009). [PubMed: 19357332]

[39]. Pei FC et al. Neural latents benchmark '21: Evaluating latent variable models of neural population activity. In Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2) (2021).

[40]. Kawashima T, Zwart MF, Yang C-T, Mensh BD & Ahrens MB The serotonergic system tracks the outcomes of actions to mediate short-term motor learning. Cell 167, 933–946.e20, 10.1016/j.cell.2016.09.055 (2016). [PubMed: 27881303]

[41]. Jeon BB, Swain AD, Good JT, Chase SM & Kuhlman SJ Feature selectivity is stable in primary visual cortex across a range of spatial frequencies. Scientific reports 8, 15288–15288, 10.1038/s41598-018-33633-2 (2018). [PubMed: 30327571]

[42]. Pandarinath C et al. Inferring single-trial neural population dynamics using sequential autoencoders. Nature Methods 15, 805–815, 10.1038/s41592-018-0109-9 (2018). [PubMed: 30224673]

[43]. Churchland M et al. Stimulus onset quenches neural variability: A widespread cortical phenomenon. Nature neuroscience 13, 369–78, 10.1038/nn.2501 (2010). [PubMed: 20173745]

[44]. Williamson RC et al. Scaling properties of dimensionality reduction for neural populations and network models. PLOS Computational Biology 12, e1005141, 10.1371/journal.pcbi.1005141 (2016). [PubMed: 27926936]

[45]. Umakantha A et al. Bridging neuronal correlations and dimensionality reduction. Neuron 109, 2740–2754.e12, 10.1016/j.neuron.2021.06.028 (2021). [PubMed: 34293295]

[46]. Huang C et al. Circuit models of low-dimensional shared variability in cortical networks. Neuron 101, 337–348.e4, 10.1016/j.neuron.2018.11.034 (2019). [PubMed: 30581012]

[47]. Triplett MA, Pujic Z, Sun B, Avitan L & Goodhill GJ Model-based decoupling of evoked and spontaneous neural activity in calcium imaging data. PLOS Computational Biology 16, e1008330, 10.1371/journal.pcbi.1008330 (2020). [PubMed: 33253161]

[48]. Aitchison L et al. Model-based Bayesian inference of neural activity and connectivity from all-optical interrogation of a neural circuit. In Guyon I et al. (eds.) Advances in Neural Information Processing Systems, vol. 30 (Curran Associates, Inc., 2017).

[49]. Kao JC et al. Single-trial dynamics of motor cortex and their applications to brain-machine interfaces. Nature Communications 6, 7759, 10.1038/ncomms8759 (2015).

[50]. Williams AH et al. Discovering precise temporal patterns in large-scale neural recordings through robust and interpretable time warping. Neuron 105, 246–259.e8, 10.1016/j.neuron.2019.10.020 (2020). [PubMed: 31786013]

[51]. Kobak D et al. Demixed principal component analysis of neural population data. eLife 5, e10989, 10.7554/eLife.10989 (2016). [PubMed: 27067378]

[52]. Archer EW, Koster U, Pillow JW & Macke JH Low-dimensional models of neural population activity in sensory cortical circuits. In Ghahramani Z, Welling M, Cortes C, Lawrence N & Weinberger KQ (eds.) Advances in Neural Information Processing Systems, vol. 27 (Curran Associates, Inc., 2014).

[53]. Lakshmanan KC, Sadtler PT, Tyler-Kabara EC, Batista AP & Yu BM Extracting low-dimensional latent structure from time series in the presence of delays. Neural Computation 27, 1825–1856, 10.1162/NECO_a_00759 (2015). https://direct.mit.edu/neco/article-pdf/27/9/1825/946601/neco_a_00759.pdf. [PubMed: 26079746]

[54]. Elsayed G & Cunningham J Structure in neural population recordings: An expected byproduct of simpler phenomena? Nature Neuroscience 20, 10.1038/nn.4617 (2017).

[55]. Pnevmatikakis E et al. Simultaneous denoising, deconvolution, and demixing of calcium imaging data. Neuron 89, 285–299, 10.1016/j.neuron.2015.11.037 (2016). [PubMed: 26774160]

[56]. Deneux T et al. Accurate spike estimation from noisy calcium signals for ultrafast three-dimensional imaging of large neuronal populations in vivo. Nature Communications 7, 12190, 10.1038/ncomms12190 (2016).

[57]. Jewell S & Witten D Exact spike train inference via ℓ_0 optimization. The Annals of Applied Statistics 12, 10.1214/18-AOAS1162 (2017).

[58]. Prince LY, Bakhtiari S, Gillon CJ & Richards BA Parallel inference of hier-archical latent dynamics in two-photon calcium imaging of neuronal populations. bioRxiv 10.1101/2021.03.05.434105 (2021). https://www.biorxiv.org/content/early/2021/03/08/2021.03.05.434105.full.pdf.

[59]. Bittner SR et al. Population activity structure of excitatory and inhibitory neurons. PLOS ONE 12, 1–27, 10.1371/journal.pone.0181773 (2017).

[60]. Chen JL, Voigt FF, Javadzadeh M, Krueppel R & Helmchen F Long-range population dynamics of anatomically defined neocortical networks. eLife 5, e14679, 10.7554/eLife.14679 (2016). [PubMed: 27218452]

[61]. Semedo JD, Gokcen E, Machens CK, Kohn A & Yu BM Statistical methods for dissecting interactions between brain areas. Current Opinion in Neurobiology 65, 59–69, 10.1016/jxonb.2020.09.009 (2020). Whole-brain interactions between neural circuits. [PubMed: 33142111]

[62]. Pachitariu M et al. Suite2p: beyond 10,000 neurons with standard two-photon microscopy. bioRxiv 10.1101/061507 (2017). https://www.biorxiv.org/content/early/2017/07/20/061507.full.pdf.

[63]. Freeman J et al. Mapping brain activity at scale with cluster computing. Nature methods 11, 10.1038/nmeth.3041 (2014).

[64]. Santhanam G et al. Factor-analysis methods for higher-performance neural prostheses. Journal of Neurophysiology 102, 1315–1330, 10.1152/jn.00097.2009 (2009). 10.1152/jn.00097.2009. [PubMed: 19297518]

[65]. Kawashima T & Ahrens MB Calcium imaging in the dorsal raphe nucleus of zebrafish. 10.6084/m9.figshare.21646682.v6 (2022).

[66]. Jeon B & Kuhlman S Responses of V1 excitatory neurons to full-field grating. G-Node. 10.12751/g-node.wc3f8g (2022).

[67]. Koh T kohth/cilds: cilds, 10.5281/zenodo.7388544 (2022).

[68]. Buchanan EK et al. Penalized matrix decomposition for denoising, compression, and improved demixing of functional imaging data. bioRxiv 10.1101/334706 (2019). https://www.biorxiv.org/content/early/2019/01/21/334706.full.pdf.
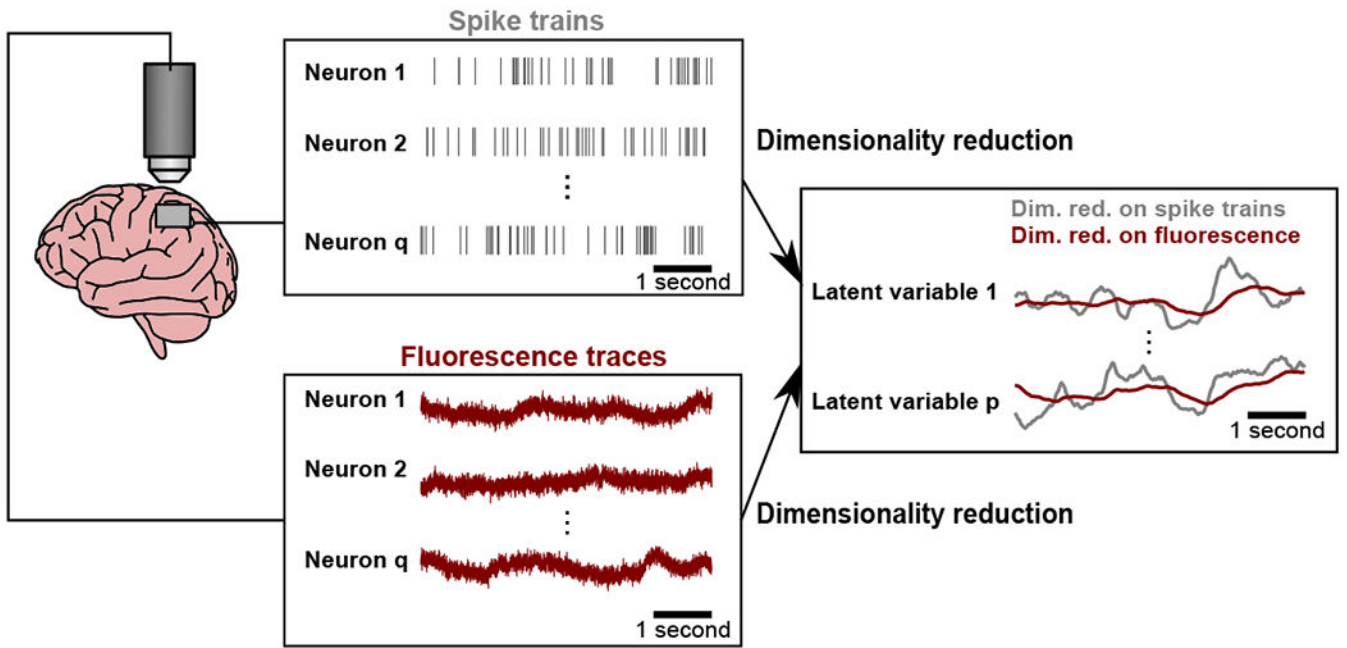
Figure 1: Dimensionality reduction of calcium imaging recordings.
A key property of calcium imaging is the slow decay of the measured fluorescence (left panel, maroon) after each spiking event (left panel, grey). If ignored, the calcium decay could introduce temporal correlations in the estimated latent variables (right panel, maroon), where those temporal correlations would not be present had we estimated the latent variables from the underlying spike trains (right panel, grey).
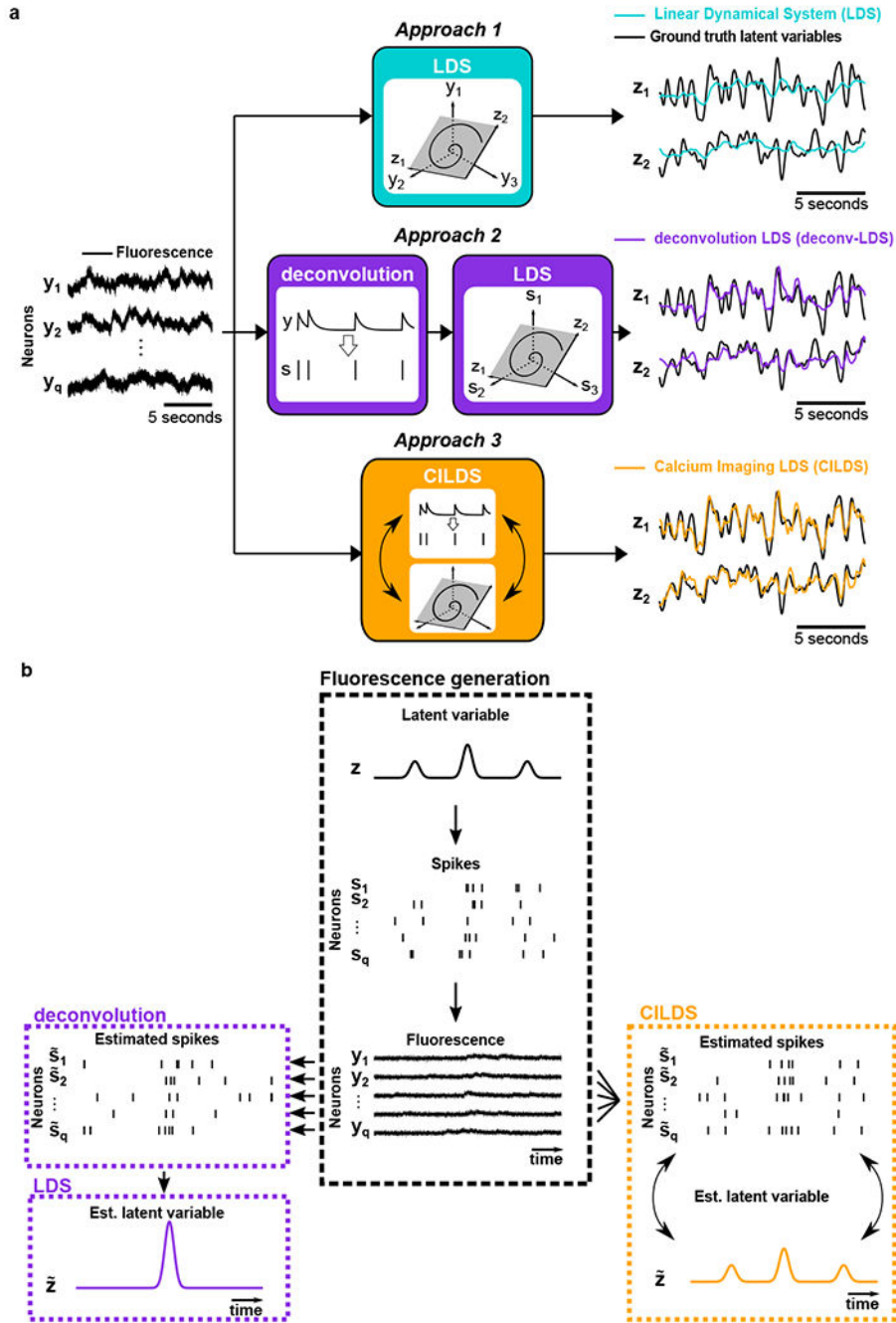
Figure 2: Comparison of three classes of dimensionality reduction methods.
(a) Each of the three classes of methods was applied to the simultaneously-recorded fluorescence of a population of neurons (y_1, y_2, ...y_q) to extract latent variables. Top, Approach 1: a standard dimensionality reduction method (e.g., LDS) applied directly to calcium imaging recordings, extracting corresponding low-dimensional latent variables at each time point (illustrated here with two dimensions, z_1 and z_2). Middle, Approach 2: deconvolution is applied separately to each neuron's fluorescence trace to estimate its underlying spiking activity (s_1, s_2, ..., s_q). A standard dimensionality reduction method (e.g., LDS) is then applied

to the estimated spiking activity to extract latent variables (z_1 and z_2). Bottom, Approach 3: A unified method (e.g., CILDS) that takes calcium imaging recordings as input and performs deconvolution and dimensionality reduction simultaneously to extract the latent variables (z_1 and z_2). **(b)** Cartoon depicting the intuition behind the difference between Approaches 2 and 3. Center column: a latent variable z (representing, for example, common input) is used to generate spike trains which, in turn, are used to generate fluorescence traces. Left column: Deconvolution is performed neuron by neuron (Approach 2, deconv-LDS), then an LDS is applied to the estimated spiking activity to extract latent variables. Right column: A unified method (Approach 3, CILDS) is applied to all neurons together to dissociate the calcium transients from the underlying shared spiking activity among neurons (i.e., the estimated latent variable). This is done by jointly performing deconvolution and dimensionality reduction, as illustrated by the the double arrows. Note that the estimated spiking activity is depicted here as spike trains for visual clarity, even though they are in fact continuous-valued time courses.
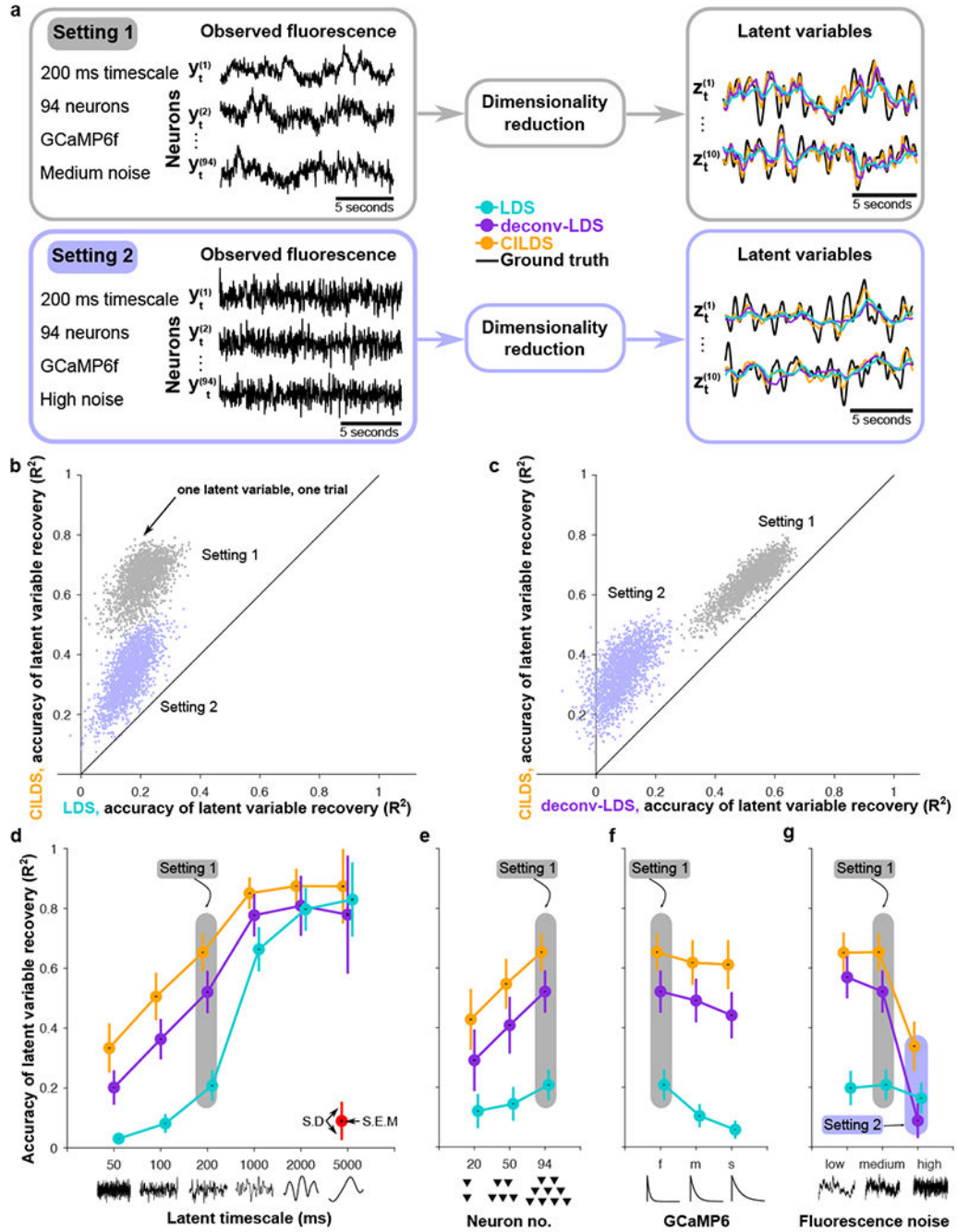
**Figure 3: Accuracy of latent variable recovery in simulation.**

(a) Example simulated fluorescence traces (left panels) and estimated latent variables (right panels) for two combinations of experimental variables. Setting 1 corresponds to a latent timescale of 200ms, 94 neurons, calcium decay corresponding to GCaMP6f, and medium fluorescence noise (see Methods). Setting 2 is the same as Setting 1, but with high fluorescence noise. Each of the three dimensionality reduction approaches introduced in Fig. 2 (LDS, cyan; deconv-LDS, purple; CILDS, orange) is applied to the simulated fluorescence traces. The latent variables extracted by each method are then compared to the

ground truth latent variables (black). **(b-c)** Accuracy of latent variables estimated by CILDS versus that of (b) LDS and (c) deconv-LDS. Accuracy is measured by the R^2 between each of the estimated and ground truth latent variables. Each point represents one latent variable on one trial, with a total of 2000 points for each setting, comprising 200 trials and 10 latent variables per trial. **(d-g)** Mean accuracy of latent variable recovery, as the (d) latent timescale, (e) number of neurons, (f) GCaMP6 indicator decay time constant, and (g) fluorescence noise level was varied. In each panel (d-g), one of the experimental variables was varied, while the other three variables were held constant at the Setting 1 values. The common point across the four panels is Setting 1 (shaded gray). Setting 2 (shaded purple) only appears in panel (g) because panels (d)-(f) correspond to medium rather than high fluorescence noise. The R^2 for other combinations of experimental variables are shown in Supplementary Fig. 1. Colored error bars indicate standard deviation, and black error bars indicate standard error across $n = 2000$ latent variables (see Methods). The points are horizontally offset for visual clarity.
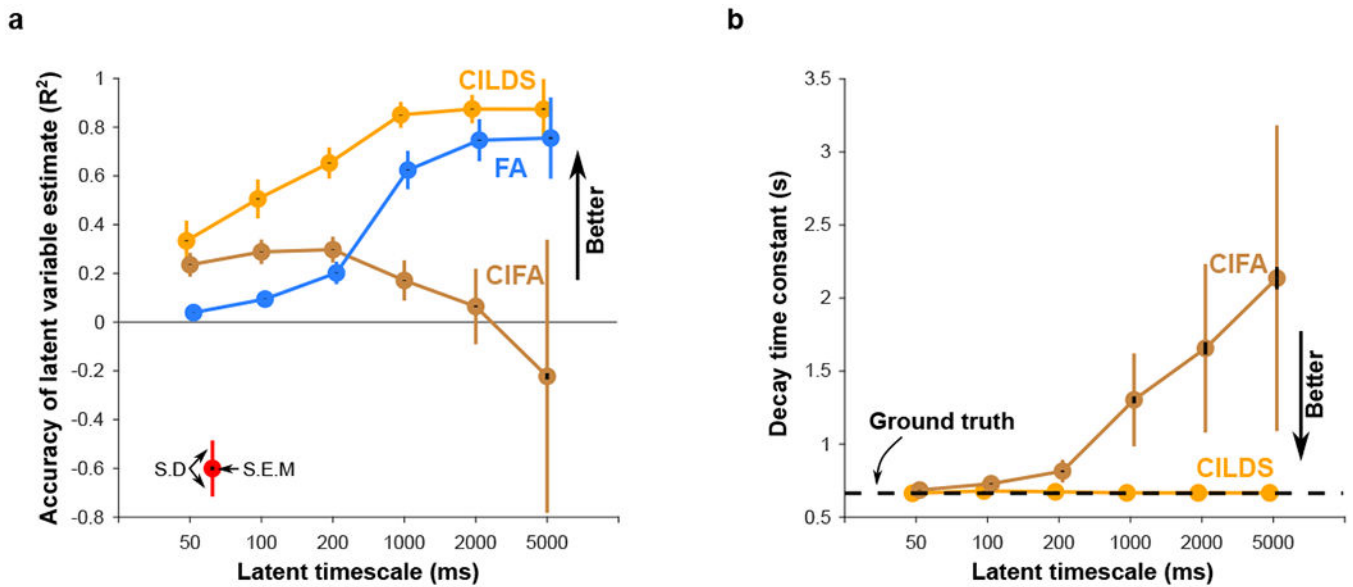
Figure 4: Comparison between CILDS and methods that do not include a latent dynamical model.

Here we compare the performance of three methods at recovering ground truth latent variables in simulation: one with deconvolution and latent dynamics (CILDS), one with de-convolution but no latent dynamics (CIFA), and one with no deconvolution and no latent dynamics (FA). The simulation parameters are GCaMP6f with 94 neurons and medium noise, as in Fig. 3d. **(a)** Accuracy of latent variable recovery for CILDS (orange), CIFA (brown), and FA (blue) across a wide range of latent timescales. Note that the R^2 can be less than zero because these results are cross-validated. The CILDS curve shown here is the same as in Fig. 3d. **(b)** Mean calcium decay time constant estimated using CILDS (orange) and CIFA (brown) for different simulated latent timescales. FA does not estimate a calcium decay time constant. The dashed black line indicates the ground truth decay time constant. In both panels, coloured error bars indicate standard deviation, and black error bars indicate standard error across n=2000 latent variables (see Methods).
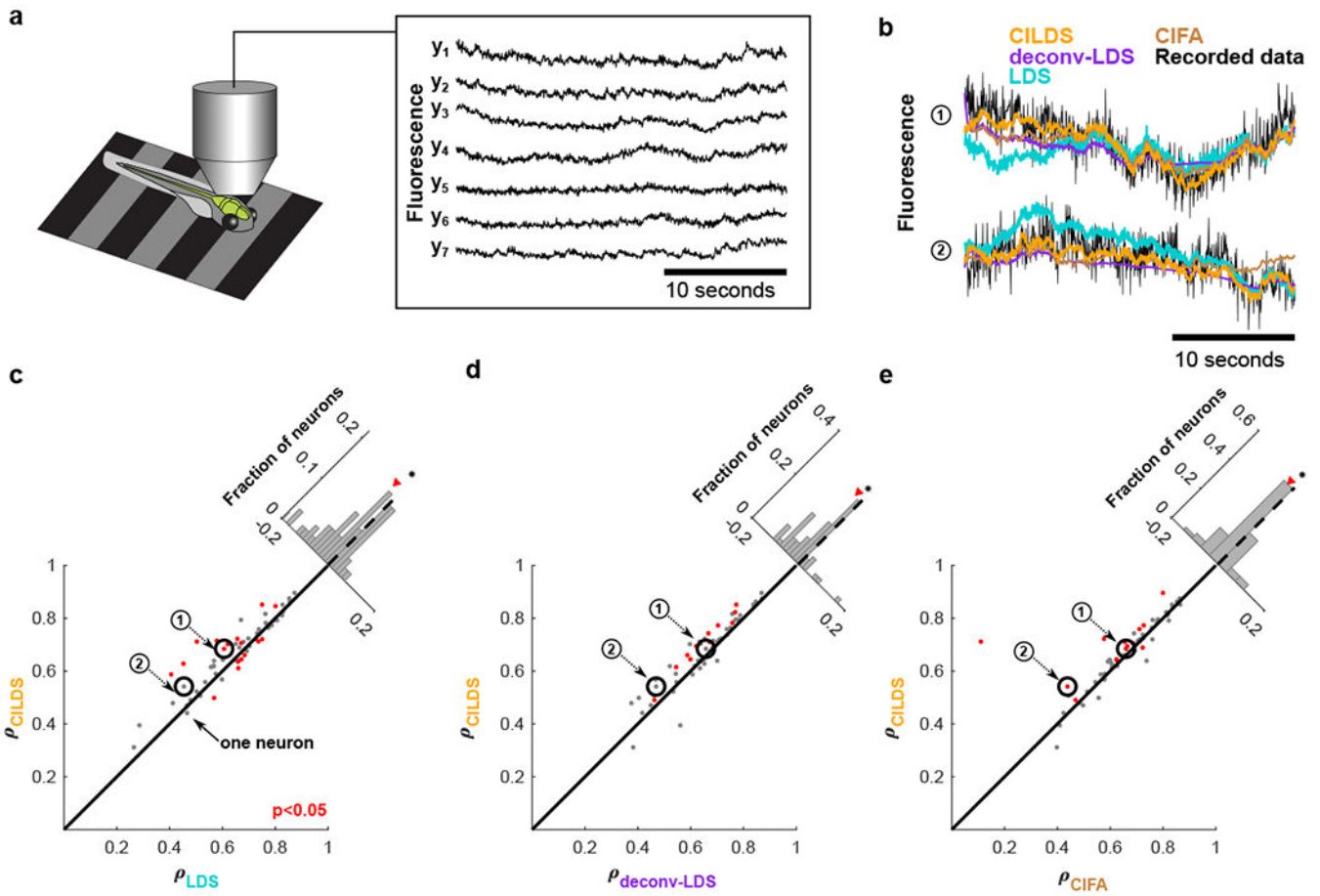
Figure 5: Performance comparison on larval zebrafish DRN recordings.
(a) Two-photon calcium imaging using GCaMP6f at 30Hz was performed on three larval
zebrafish in a virtual reality environment. Shown are representative fluorescence traces
from seven of the imaged neurons. **(b)** Example recorded fluorescence traces (black) and
leave-neuron-out predicted fluorescence using CILDS (orange), deconv-LDS (purple), LDS
(cyan), and CIFA (brown). **(c-e)** Correlation between the recorded fluorescence and the
leave-neuron-out predicted fluorescence for CILDS versus each of the other methods. Each
point represents one neuron, where the correlation is computed for each trial (27 seconds
long), then averaged across all 15 trials. Diagonal histograms show the paired difference in
performance between CILDS and one of the other methods, as indicated. The correlation is
higher for CILDS than (c) LDS ($p = 0.0001$, $n = 60$ neurons, paired two-tailed t-test across
the population of neurons, black asterisk indicating statistical significance), (d) deconv-LDS
($p = 5.94 \times 10^{-5}$, $n = 60$ neurons), and (e) CIFA ($p = 0.046$, $n = 60$ neurons). Note that
the histograms are zoomed-in for visual clarity, and therefore the ends of the histograms are
not shown. The numbered points (black circles) correspond to the examples shown in panel
b. Red points indicate a statistically significant difference per neuron between CILDS and
the other method being compared using a paired two-tailed t-test across trials ($p < 0.05$, see
Methods). Note that the threshold used for the t-test means that we might expect 5% of the
neurons to appear significant even if the effect is not real.
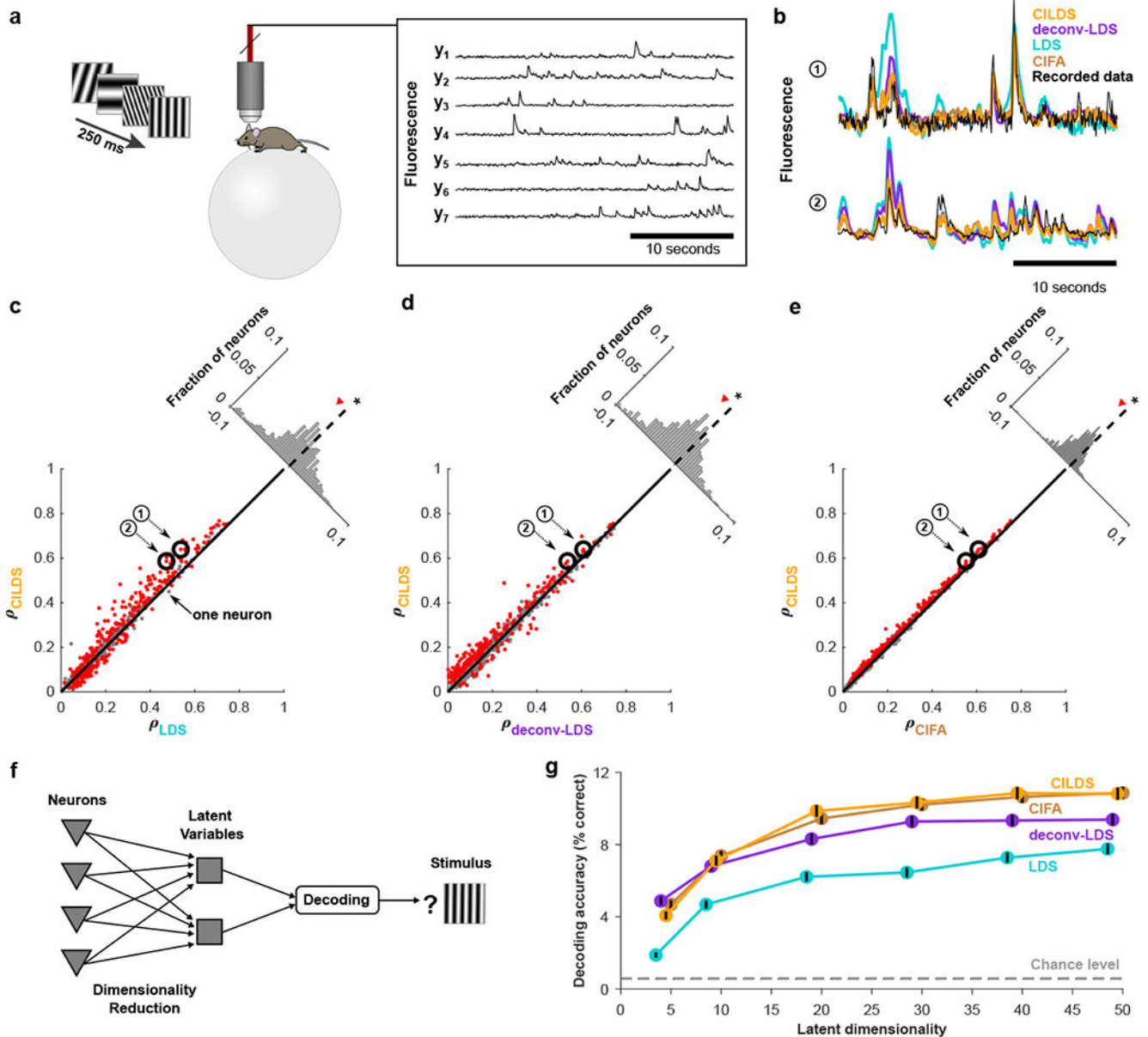
Figure 6: Performance comparison on mice V1 recordings.
(a) Two-photon calcium imaging performed on awake mice viewing static gratings with
different spatial frequencies and orientations (180 total stimuli) using GCaMP6f at 15.5Hz.
Shown are representative fluorescence traces from seven of the imaged neurons. **(b)**
Example segment of recorded fluorescence traces (black) and leave-neuron-out predicted
fluorescence traces using CILDS (orange), deconv-LDS (purple), LDS (cyan), and CIFA
(brown). **(c-e)** Correlation between the recorded fluorescence and the leave-neuron-out
predicted fluorescence for CILDS versus each of the other methods. Each point represents
one neuron, where the correlation is computed for each trial (196.7 seconds long) then
averaged across all 15 trials. Diagonal histograms show the paired difference in performance
between CILDS and one of the other methods, as indicated. The correlation is higher for

CILDS than (c) LDS ($p = 5.04 \times 10^{-17}$, $n = 704$ neurons, paired two-tailed t-test across the population of neurons, black asterisk indicating statistical significance), (d) deconv-LDS ($p = 1.7 \times 10^{-27}$, $n = 704$ neurons), and (e) CIFA ($p = 3.19 \times 10^{-81}$, $n = 704$ neurons). Note that the histograms are zoomed-in for visual clarity, and therefore the ends of the histograms are not shown. The numbered points (black circles) correspond to the examples shown in panel b. Red points indicate a statistically significant difference per neuron between CILDS and the other method being compared using a paired two-tailed t-test across trials ($p < 0.05$, see Methods). Note that the threshold used for the t-test means that we might expect 5% of the neurons to appear significant even if the effect is not real. **(f)** Flow diagram depicting decoding of visual stimuli using low-dimensional latent variables, which are obtained by applying a dimensionality reduction method to the recorded fluorescence traces. **(g)** Classification accuracy of the visual stimulus based on latent variables extracted using CILDS (orange), deconv-LDS (purple), and LDS (cyan). Classification was performed using a Gaussian Naive Bayes decoder, where the number of latent variables extracted by each dimensionality reduction method was systematically varied (horizontal axis). There were 180 total gratings (with different orientations and spatial frequencies) shown during the experiment, so the chance classification accuracy is 1/180 (gray dashed). The decoding window was 250 ms, which is the duration of stimulus presentation. Black error bars around the mean indicate 95% confidence intervals (Bernoulli process).