bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Rapid learning of neural circuitry from holographic ensemble stimulation enabled by model-based compressed sensing

Marcus A. Triplett[1,2,3,4,*,†], Marta Gajowa[5,*], Benjamin Antin[1,2,3,4], Masato Sadahiro[5], Hillel Adesnik[5], and Liam Paninski[1,2,3,4]

September 15, 2022

[1]Mortimer B. Zuckerman Mind Brain Behavior Institute, Columbia University, NY
[2]Grossman Center for the Statistics of Mind, Columbia University, NY
[3]Center for Theoretical Neuroscience, Columbia University, NY
[4]Department of Statistics, Columbia University, NY
[5]Department of Molecular and Cell Biology, University of California, Berkeley, CA
[*] These authors contributed equally
[†] Correspondence: marcus.triplett@columbia.edu

Abstract

Discovering how neural computations are implemented in the cortex at the level of monosynaptic connectivity requires probing for the existence of synapses from possibly thousands of presynaptic candidate neurons. Two-photon optogenetics has been shown to be a promising technology for mapping such monosynaptic connections via serial stimulation of neurons with single-cell resolution. However, this approach is limited in its ability to uncover connectivity at large scales because stimulating neurons one-by-one requires prohibitively long experiments. Here we developed novel computational tools that, when combined, enable learning of monosynaptic connectivity from high-speed holographic neural ensemble stimulation. First, we developed a model-based compressed sensing algorithm that identifies connections from postsynaptic responses evoked by stimulation of many neurons at once, considerably increasing the rate at which the existence and strength of synapses are screened. Second, we developed a deep learning method that isolates the postsynaptic response evoked by each stimulus, allowing stimulation to rapidly switch between ensembles without waiting for the postsynaptic response to return to baseline. Together, our system increases the throughput of monosynaptic connectivity mapping by an order of magnitude over existing approaches, enabling the acquisition of connectivity maps at speeds needed to discover the synaptic circuitry implementing neural computations.

Introduction

The structure of synaptic connectivity is central to how the brain processes information, stores long-term memories, and facilitates cognition. To uncover such connectivity, two-photon optogenetic stimulation has

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

emerged as a promising technology due to its ability to flexibly probe neurons with single-cell resolution while monitoring postsynaptic currents (PSCs) using whole-cell recordings [Packer et al., 2012, Baker et al., 2016, Shemesh et al., 2017, Hage et al., 2022]. Yet existing optogenetic circuit mapping techniques have been limited to probing connectivity from small numbers of neurons that must be slowly stimulated one-by-one, and therefore require aggregating small-scale maps across experiments to obtain large-scale maps of connectivity [Hage et al., 2022]. By contrast, an ideal monosynaptic connectivity mapping technique would enable large numbers of synapses to be identified at high speed within a single experimental session. This would provide two crucial advantages. First, each experiment would produce a more comprehensive and representative map of neural circuitry, rather than having to pool together smaller experiments that each provide just a partial view of connectivity. Second, high speed identification of synapses would enable mapping experiments to take place *in vivo*, where connectivity can be related to functional information collected in the same experiment but where, compared to *in vitro* preparations, experiment time is especially limited due to greater recording instability.

Multiple strategies could serve to advance optogenetic circuit mapping towards this state, although each introduces experimental and computational challenges. The simplest strategy to improve the throughput of a connectivity mapping experiment is to increase the rate at which stimulation switches between neurons. This is primarily determined by the refresh rate of the holographic spatial light modulator (SLM) if using advanced light sculpting techniques, or, in the case of laser scanning approaches, just the time required for a neuron to integrate photocurrent until it elicits an action potential [Rickgauer and Tank, 2009]. However, while this approach allows a mapping experiment to be completed in a shorter time frame, naively stimulating too quickly confounds postsynaptic measurements because the membrane conductance of the postsynaptic neuron will not have sufficient time to return to baseline conditions before the next stimulus is applied. The speed at which a mapping experiment can feasibly take place therefore becomes limited by the ability to computationally demix PSC waveforms that overlap in time.

Simulation studies predict that connectivity mapping could also be greatly accelerated through the use of compressed sensing, where, in principle, sparse connectivity could be reconstructed from few measurements provided that stimulation is applied to ensembles of randomly selected neurons at once [Hu and Chklovskii, 2009, Fletcher et al., 2011, Mishchenko and Paninski, 2012, Shababo et al., 2013, Draelos and Pearson, 2020]. However, existing compressed sensing algorithms neglect the complex biophysics intrinsic to mapping experiments, and therefore yield biased estimates of synaptic connectivity. In particular, using compressed sensing to estimate connectivity requires knowledge of which neurons elicited spikes in response to stimulation and when, since otherwise a connectivity inference algorithm will erroneously average over postsynaptic measurements on stimulation trials where no presynaptic spikes actually occurred [Hu and Chklovskii, 2009]. Increasing mapping throughput via ensemble stimulation is therefore limited by the ability to correctly infer the successful optogenetic triggering of individual presynaptic spikes when only observing postsynaptic responses.

To enable high-speed, large-scale connectivity mapping experiments, we developed new computational tools for inferring connectivity from holographic stimulation of neural ensembles. While existing optogenetic circuit mapping techniques perform slow single-neuron stimulation, our tools allow experiments to proceed with rapid holographic ensemble stimulation, minimizing downtime of the holographic SLM since evoked PSCs are isolated in time, deconfounded of spontaneous PSCs, and cleaned of electrical noise by a computational

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

demixing procedure post-hoc. To infer connectivity, we developed a model-based compressed sensing algorithm that simultaneously estimates synaptic weights, presynaptic spikes, and presynaptic photoactivabilities while accounting for biophysical constraints and any residual spontaneous PSCs not already eliminated by the demixing procedure. To validate these tools, we performed connectivity mapping experiments using the recently engineered family of fast, potent ChroME2.0 opsins [Sridharan et al., 2022] together with two-photon holographic light sculpting [Pégard et al., 2017], and routinely mapped connectivity from hundreds of presynaptic candidates within 680x680x100 μm^3 volumes of cortex. By combining rapid ensemble stimulation with computational demixing and model-based compressed sensing, we reduced the stimulation time required to reconstruct connectivity by an order of magnitude over existing approaches, allowing large numbers of synapses to be quickly mapped within individual experimental sessions.

To facilitate the use of our techniques, we provide an open access software toolbox and web-based notebook implementation that researchers can straightforwardly use and validate in their own mapping experiments. While our analysis shows that the highest mapping throughput is obtained using holographic ensemble stimulation, our techniques also increase the accuracy and speed of more widely accessible mapping approaches based on conventional single-target stimulation, and therefore can enable fast and accurate connectivity mapping in a broad range of laboratories.

Results

Optogenetic circuit mapping framework

Accurately characterizing the existence and strength of a synapse requires precise control over the initiation of presynaptic action potentials. Ideally, each stimulus would evoke just a single presynaptic spike since this provides a direct measurement of synaptic charge transfer. To this end, we combined the family of potent soma-targeted ChroME opsins [Mardinly et al., 2018, Sridharan et al., 2022] with a scanless computer-generated holography system named 3D-SHOT [Pégard et al., 2017] (Figure 1a). We expressed ChroME2f in parvalbumin-expressing (PV) or somatostatin-expressing (SST) neurons by viral transfection with an adeno-associated virus [Sridharan et al., 2022]. Alternatively, in some experiments, ChroME2s was expressed in pyramidal neurons instead. We calibrated laser powers and illumination time such that brief (3-5 ms) periods of stimulation almost always led to either 0 or 1 action potential, with minimal instances of multiple action potentials resulting from a single pulse (Figure S1). Next, we confirmed through a similar process that we could reliably evoke action potentials when holographically stimulating ensembles of neurons at once (Figure S1).

Having established the high spatiotemporal precision of our holographic stimulation technique here and in previous work [Pégard et al., 2017, Bounds et al., 2021, Sridharan et al., 2022], we then developed a computational system for processing optogenetic data and inferring synaptic connectivity. The first component of our system seeks to extract precise measurements of optogenetically-evoked PSCs that have been corrupted by three notable factors: (1) electrical noise arising from the electrode that records intracellular currents from the voltage-clamped neuron, which increases the variability of the observed currents and can require additional repetitions of each presynaptic stimulus to overcome; (2) spontaneous currents that, depending on their timing relative to the optogenetic stimulus, can obscure optogenetically-evoked PSCs or increase the number of false-positive connections; and (3) postsynaptic responses from preceding or subsequent trials when the

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

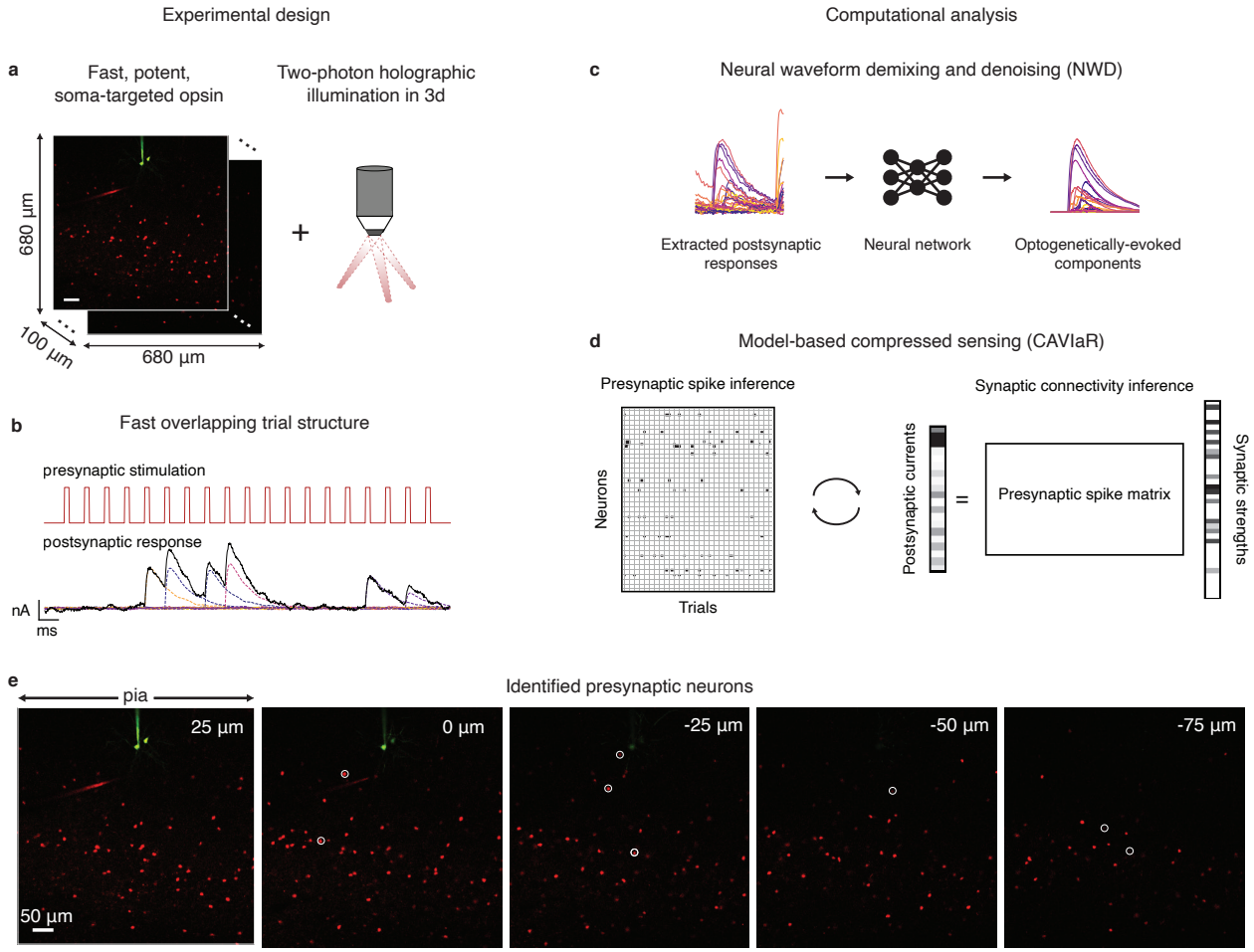


Figure 1: Overview of proposed high-throughput optogenetic circuit mapping framework. a, A fast, potent, and soma-targeted opsin (here ChroME2f) is expressed in a pool of candidate presynaptic neurons while intracellular currents are monitored from a voltage-clamped postsynaptic neuron. Two-photon holography is used to elicit spikes in ensembles of presynaptic candidates, increasing the number of putative synapses tested compared to single-target stimulation. Scale bar, 50 μm. **b**, To further increase throughput, stimulation is applied at high speeds (e.g., 30−50 Hz), such that the postsynaptic membrane conductance is not guaranteed to return to baseline conditions (i.e., trials are "overlapping" in time). This allows more synapses to be tested per minute, but complicates downstream identification of their existence and strength unless the evoked PSCs are subsequently demixed. **c**, A neural network is trained to isolate the components of the postsynaptic response that are plausibly due to the optogenetic stimulus, allowing high-fidelity measurements of postsynaptic current even at very high stimulation frequencies. **d**, The "demixed" postsynaptic responses evoked by holographic ensemble stimulation are used in a model-based compressed sensing algorithm (CAVIaR) to reconstruct the underlying connectivity. Unlike traditional compressed sensing techniques, compressive mapping of synaptic circuitry requires inference of any optogenetically elicited presynaptic spikes (left) as well as the synapses themselves (right). CAVIaR alternates between inference of presynaptic spikes and synaptic connectivity, in addition to a number of other pertinent biophysical variables (see main text). **e**, Presynaptic neurons (white circles) identified via the proposed pipeline in an example experiment mapping connectivity from presynaptic PV neurons to a postsynaptic pyramidal neuron.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

interstimulus interval is shorter than the PSC decay time, which are prevalent when attempting to perform mapping experiments at very high speeds. To simultaneously eliminate these factors, we developed a deep neural network architecture for demixing and denoising PSC waveforms (Figure 1b,c) in a process that we call neural waveform demixing (NWD). The NWD network attempts to isolate optogenetically evoked PSCs stimulus-by-stimulus, such that confounding synaptic currents not driven by the optogenetic stimulus are "subtracted out" by the network and the resulting currents cleaned of electrical noise. Consequently, the NWD network allows experiments to proceed with very short interstimulus intervals since evoked PSCs from previous and subsequent trials are subtracted out and the initial baseline current reset to zero (see below for more details).

The second component of our system is a model-based compressed sensing algorithm (Figure 1d). The key challenge for a compressive approach to connectivity mapping is to robustly infer presynaptic spikes from PSCs evoked by ensemble stimulation despite a multitude of unobserved sources of biophysical variability. For example, the illumination intensity required to elicit a presynaptic spike varies unpredictably between neurons because opsin is expressed differentially throughout a neural population, regardless of whether the opsin construct is delivered using a viral or transgenic targeting strategy. In a mapping experiment, this relationship between light and presynaptic spike probability cannot be observed from the presynaptic neuron itself, but must be inferred through electrophysiological measurements from the postsynaptic neuron. This inference problem is in turn complicated by the fact that PSCs can occur spontaneously, that postsynaptic responses are corrupted by electrical noise, that synaptic transmission can fail despite successful elicitation of a presynaptic spike, and that potentially tens of neurons are stimulated at a time. Accounting for the power-dependence of presynaptic spike initiation for individual neurons when only observing postsynaptic responses to ensemble stimulation is therefore computationally challenging, and constitutes only one of the inference problems that must be solved to accurately reconstruct connectivity using a compressed sensing approach.

To resolve these challenges, we embedded the compressed sensing step in a hierarchical Bayesian statistical model. Our model relates patterns of ensemble stimulation to electrophysiological measurements from the postsynaptic cell while, at the same time, accounting for variability in photoactivability and synaptic transmission across the opsin-expressing population. We developed a variational inference technique called CAVIaR (coordinate-ascent variational inference and isotonic regularization) to learn posterior distributions over the model parameters (Methods). CAVIaR identifies the existence and strength of individual synapses from ensemble stimulation, thereby testing for many more potential synaptic partners per trial and substantially increasing the rate at which cortical circuits can be mapped.

Neural waveform demixing allows mapping experiments to proceed rapidly

The NWD network has a sequential U-Net architecture [Ronneberger et al., 2015, Falk et al., 2019] that uses one dimensional convolutional filters to learn representations of the input signal at increasing levels of temporal compression. This allows the network to integrate information across the entire PSC trace, such that confounding synaptic currents preceding or following an admissible PSC "initiation zone" (typically selected to be a window of 3-12 ms following stimulation) can be accurately subtracted and the baseline current reset to zero (Figure 2a). Since the network is trained to generate noise-free, demixed PSCs from noisy inputs, NWD also markedly improves the overall signal quality (Figure 2). Furthermore, unlike computationally intensive

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

algorithms for deconvolving intracellular currents [Merel et al., 2016], NWD can isolate the PSC evoked by optogenetic stimulation with just a simple forward pass through the network, with the resulting demixed PSCs usable in any connectivity inference algorithm and in real time.

Performing optogenetic connectivity mapping experiments at high speed poses a trade-off: the experiment can be completed much more quickly, but measurements of the postsynaptic response are much more confounded by currents from previous and subsequent trials. For this reason, a typical rate of stimulation for a connectivity mapping experiment is $\sim$10 Hz (i.e., stimulating a neuron approximately every 100 ms, much longer than the decay time of a typical PSC [Printz et al., 2021, Hage et al., 2022]), which we take as a baseline rate in later method comparisons. To characterize the improvement in experiment speed-up resulting from the application of NWD, we simulated connectivity mapping experiments at full 20 kHz sampling resolution with varying synaptic connection densities and stimulation frequencies (Figure 2b, Figure S2). In a sparsely connected circuit, the primary action of NWD is to eliminate spontaneous synaptic currents and reduce electrical noise, leading to just a marginal improvement in signal fidelity compared to the raw PSC traces (Figure 2b, 0.1 connection density example). However, as synapses become more prevalent (as in the case of e.g. PV-pyramidal connections [Hage et al., 2022]) the occurrence of postsynaptic events greatly increases, with many confounding synaptic currents elicited from previous trials. Application of NWD therefore leads to a substantial improvement in the signal fidelity of the postsynaptic response at high stimulation frequencies (Figure 2b, 0.2 to 0.4 connection density examples). Similarly, the fidelity of the postsynaptic response closely depends on the number of simultaneously stimulated neurons and the background rate of spontaneous synaptic currents (Figure 2b, middle and right panels). However, NWD largely eliminates this dependence, and thus reduces experimental constraints on the duration of interstimulus intervals since confounding synaptic currents are overwhelmingly nullified. Stimulation can therefore be increased to rates closer to the refresh rate of the SLM. In practice, this enables us to stimulate much faster than the $\sim$40 ms decay time typical of PSC traces in our preparations.

We applied NWD to each of our experimental settings, including demixing of inhibitory PSCs from PV-pyramidal and SST-pyramidal mapping experiments under holographic single-target and ensemble stimulation (Figure 2d-e), as well as demixing of excitatory PSCs from pyramidal-pyramidal mapping experiments (Figure 2f). We found that we could improve the performance of the NWD network by tuning it to the time constants of either inhibitory-pyramidal or pyramidal-pyramidal synaptic currents (Figure S3). In each case the NWD network led to a dramatic reduction in confounding synaptic currents and electrical noise, in agreement with our simulated results (Figure 2b).

Ensemble stimulation can test for many synapses simultaneously

We sought to capitalize on recent technological developments in opsin engineering [Baker et al., 2016, Shemesh et al., 2017, Mardinly et al., 2018, Sridharan et al., 2022], holographic light sculpting [Papagiakoumou et al., 2010, Pégard et al., 2017, Papagiakoumou et al., 2020, Adesnik and Abdeladim, 2021] and fundamental mathematical results on efficient signal reconstruction in sparse settings [Donoho, 2006, Candes and Tao, 2006, Candes et al., 2006] by developing a statistical method for inferring synaptic connectivity from holographic ensemble stimulation. This was based on the logic that pairing rapid stimulation and computational demixing with the ability to test for many potential synapses with each hologram could dramatically speed

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

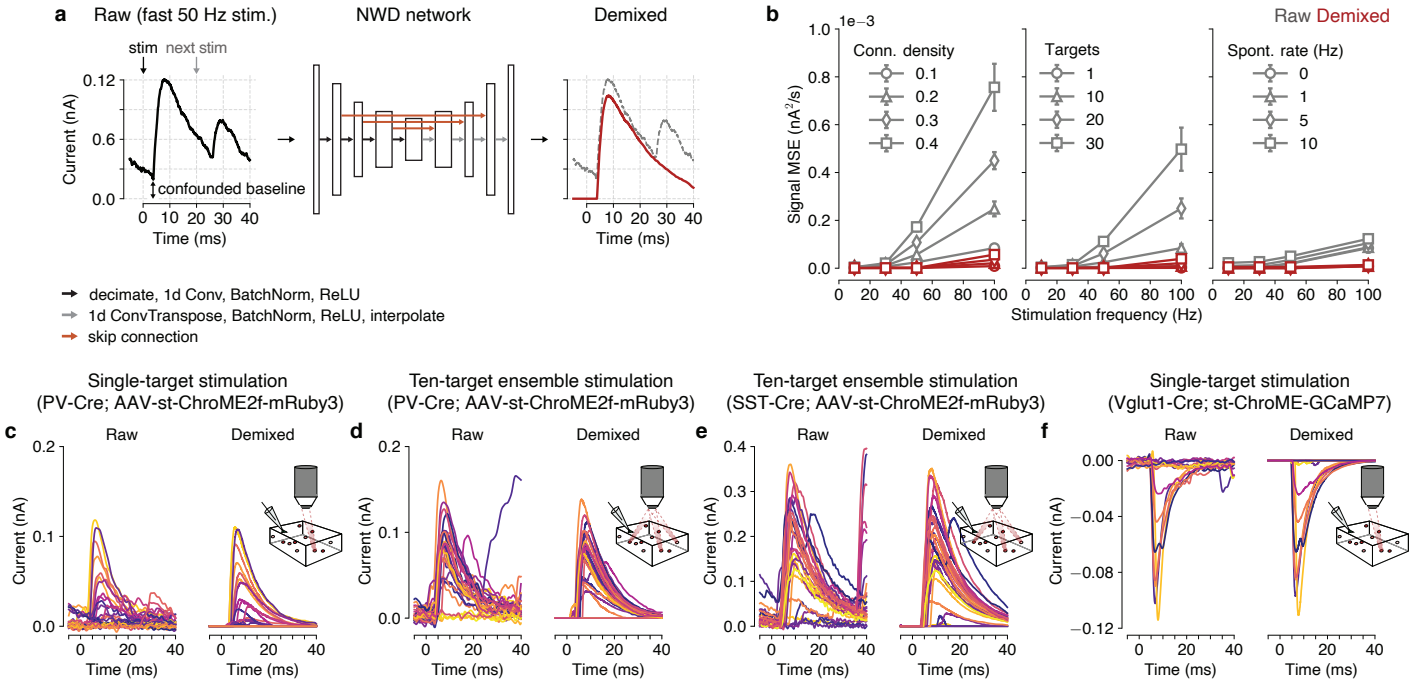


Figure 2: Simultaneous demixing and denoising of optogenetically evoked postsynaptic currents. a, Neural waveform demixing overview. Raw electrophysiological traces (left) are confounded by synaptic currents from previous and subsequent trials during fast stimulation, in addition to spontaneous synaptic currents. The NWD network (middle) is trained to isolate the PSC waveform associated with the optogenetic stimulus, resulting in a demixed trace (right, red) with the confounding synaptic currents subtracted and baseline current (at 0 ms) adjusted to 0 nA. Note that a 5 ms "pre-stimulus" window is provided to the network as additional context. **b**, Performance of NWD in simulation as a function of stimulation frequency under increasing connection density (left), number of simultaneously targeted cells (middle) and rate of spontaneous activity (right). In all cases NWD leads to a substantial improvement in signal fidelity (measured by the mean square error of the raw vs demixed trace compared to the ground truth evoked PSC). Errors are evaluated across all conditions using a single trained network. **c, d**, Examples of NWD applied to optogenetically evoked inhibitory PSCs under single target (c) and ten target (d) holographic stimulation of PV neurons expressing ChroME2f. Panel (d) shows responses to stimulating ensembles containing the neuron in panel (c). **e**, Same as (d), but for stimulation of SST neurons. **f**, same as (c), but for stimulation of pyramidal neurons expressing ChroME using the Ai203 mouse line [Bounds et al., 2021].

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

up circuit mapping. Notably, ref. [Packer et al., 2012] established the technical feasibility of this experiment by using an SLM to split a two-photon laser beam into multiple beamlets targeting different neurons while confirming the simultaneous elicitation of presynaptic spikes in each target. We therefore created and tested a variety of connectivity inference algorithms, and ultimately found that a variational inference approach for a hierarchical Bayesian statistical model yielded state-of-the-art performance. Our model relates patterns of holographic stimulation to PSCs through a series of key latent variables: (1) optogenetic "power curves" that characterize the relationship between laser power and presynaptic spike probability due to variation in opsin expression and rheobase across neurons, (2) presynaptic spikes successfully elicited by photostimulation and transmitted to the postsynaptic neuron, and (3) synaptic weights that determine the amplitude of the resulting PSCs (Figure 3a). We learn posterior distributions over the latent variables using CAVIaR, an algorithm that iteratively updates the parameters of a variational approximation to the posterior such that the inferred presynaptic spikes respect the biophysical plausibility constraint of having isotonically increasing spike probabilities as a function of laser power (i.e., such that the probability of evoking a presynaptic spike increases with laser power on average, see Methods).

A key component of our experimental design was to randomly switch between three or more laser powers while mapping. Since CAVIaR infers presynaptic spikes, we could then estimate optogenetic power curves to determine whether the plausibility constraint was satisfied. When using our software, a user simply sets a parameter that imposes a minimum success rate (between 0 and 1) for evoking presynaptic spikes (given by the isotonically-constrained power curve at maximal power) and CAVIaR identifies putatively connected neurons meeting that criterion. However, PSCs can occur spontaneously, inflating the estimated presynaptic spike rate and increasing the number of false positives. We therefore simultaneously estimate the background rate of spontaneous currents and use this to adaptively adjust the plausibility criterion during inference. This allows CAVIaR to immediately disconnect spurious synapses for which spike counts happen to implausibly decrease with increasing power.

We compared CAVIaR to two other approaches: (1) compressive sampling matching pursuit (CoSaMP, [Needell and Tropp, 2009]), a highly scalable algorithm based on compressed sensing that was used in the first simulation study on compressive circuit mapping [Hu and Chklovskii, 2009], and (2) a variational inference technique for a spike-and-slab connectivity model similar to CAVIaR but that, among other methodological differences, does not impose isotonic regularization or account for spontaneous synaptic currents (CAVI-SnS, [Shababo et al., 2013]). We evaluated the ability of each method to reconstruct connectivity in simulations where the sparsity of the underlying connectivity varied, since compressed sensing-based techniques are known to be particularly sensitive to this parameter [Donoho, 2006]. We found that CAVIaR greatly outperformed conventional compressed sensing (Figure 3b, left vs right panels), primarily because the latter does not have a model for stochastic, power-dependent presynaptic spikes. Importantly, in line with earlier results [Hu and Chklovskii, 2009], we found that algorithms failing to account for this stochasticity will yield strongly biased synaptic weight estimates since they will average over postsynaptic measurements when no presynaptic spikes occurred. CAVIaR also outperformed CAVI-SnS, which transiently demonstrated high accuracy, but ultimately generated increasingly worse estimates of connectivity with increasing amounts of data since it lacked mechanisms to prevent spontaneous PSCs from being misidentified as false positives. We found that CAVIaR consistently achieved state-of-the-art performance when repeating the analysis with varying stimulus diversity, number of simultaneous targets, and rate of background spontaneous activity (Figure S6, Figure S7).

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

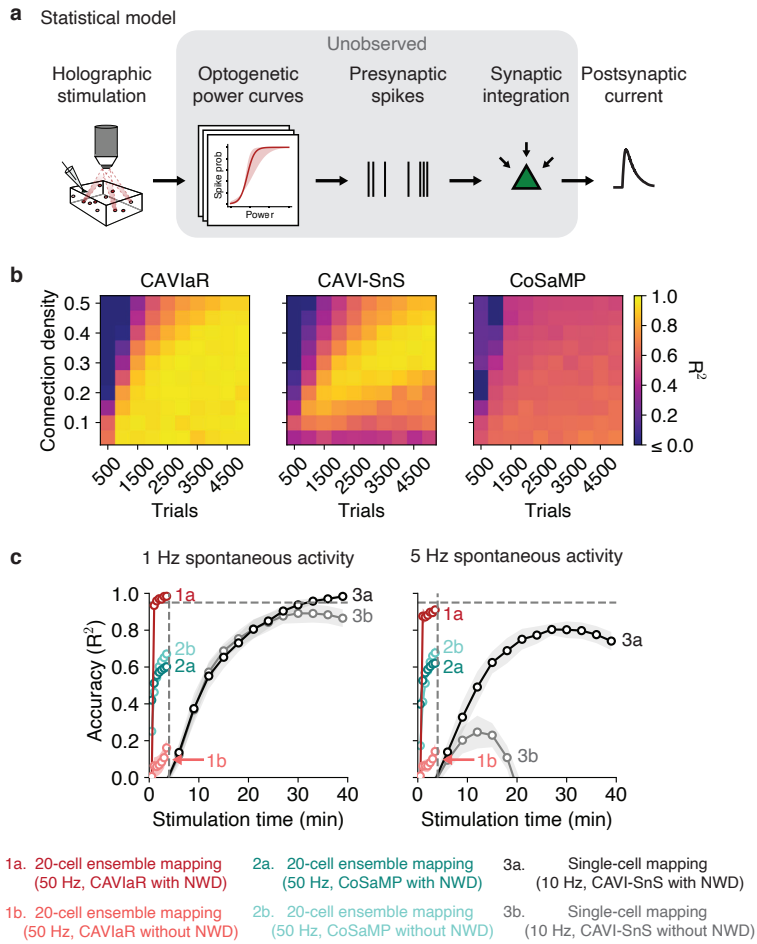


Figure 3: Inference of synaptic connectivity in simulation with CAVIaR. a, Statistical model relating patterns of holographic stimulation to measurements of postsynaptic current via latent optogenetic power curves, presynaptic spikes, and synaptic weights. **b**, Comparison of CAVIaR (ours, left) with CAVI-SnS (middle) and CoSaMP (right) for varying connection densities as a function of increasing amounts of experimental data. Simulated presynaptic candidate population size of N=1000 (hence, e.g., 4500 trials corresponds to an average of just 4.5 stimuli per neuron). All PSCs are demixed using NWD. Note that CoSaMP requires the number of desired connections to be specified in advance; we therefore provided CoSaMP with the true number of connections, such that the reported performance represents an upper bound. **c**, Order of magnitude improvement in mapping efficiency using NWD-accelerated circuit mapping with ensemble stimulation. Left, convergence speed of CAVIaR, CoSaMP, and CAVI-SnS with and without NWD, assuming a background rate of spontaneous PSCs of 1 Hz. Without NWD, the model-based methods CAVIaR and CAVI-SnS perform substantially more poorly than with NWD enabled. A higher rate of spontaneous activity (5 Hz, right panel) causes fast degradation of CAVI-SnS performance, but impacts CAVIaR and CoSaMP much less. Error bars show mean $\pm$ 1 s.d. over 20 simulations.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Next, we estimated the speed-up obtained by combining rapid ensemble stimulation with NWD and CAVIaR, and compared this against previously established techniques. Our simulations showed that even with large population sizes (e.g., mapping 1000 potential presynaptic neurons), the speed of connectivity mapping was optimal when stimulating ensembles of 20 neurons at a time (Figure S7) beyond which the accuracy of presynaptic spike inference gradually degraded. Our speed-up simulations were therefore performed using 20-cell stimulation. We anticipated that the background rate of spontaneous PSCs would also be a key factor influencing the ultimate accuracy of each connectivity inference method, since a high number of spontaneous PSCs that happen to fall within the stimulation window could mimic a synapse and lead to false positives. Single-target mapping at 10 Hz with CAVI-SnS required more than 30 minutes of stimulation to cross an accuracy of 0.95 in simulations with 1000 neurons, 10% connectivity, and 1 Hz spontaneous activity (Figure 3c, left panel; accuracy measured as the R^2 between the true and estimated synaptic weights). By comparison, NWD-enabled 50 Hz stimulation of 20-cell ensembles recovered synaptic connectivity at an accuracy exceeding 0.95 within just 30 seconds using CAVIaR, resulting in a rate of connectivity testing of more than 2000 neurons per minute in low spontaneous activity conditions. In the same time period CoSaMP obtained an accuracy of less than 0.6. Use of our computational tools is therefore predicted to enable a speed-up of more than an order of magnitude.

Increasing the rate of spontaneous activity (to 5 Hz, a more challenging regime) revealed several important behaviors of CAVIaR and CAVI-SnS (Figure 3c, right panel). Namely, the accuracy of CAVI-SnS both with and without NWD behaved non-monotonically (similar to Figure 3b, final R^2 at 5 Hz with NWD, 0.74; without NWD, <0), with connectivity estimates ultimately getting worse with additional experiment time due to the accumulation of spontaneous PSCs. However, CAVIaR with NWD obtained nearly the same accuracy when spontaneous PSCs occurred at 5 Hz as it did when they occurred at 1 Hz (final R^2 with spontaneous rate 1 Hz, 0.98; 5 Hz, 0.91) due to CAVIaR's built-in mechanisms for estimating and adapting to spontaneous PSCs. Moreover, the speed at which connectivity mapping is completed entails a lower number of spontaneous PSCs encountered. This suggests a principle whereby the rate of spontaneous activity sets a "clock" for mapping: if a mapping experiment takes too long to complete then spontaneous PSCs will eventually cause an accumulation of false positives. Connectivity mapping should therefore be performed as quickly as possible.

Validating CAVIaR using paired electrophysiological recordings *in vitro*

Having established the accuracy and speed of our connectivity mapping system in simulations, we next sought to validate our computational tools experimentally. We first wanted to confirm that the CAVIaR inferences are well-calibrated. Specifically, since compressive approaches to connectivity mapping are highly dependent on the accuracy of the presynaptic spike inference step [Hu and Chklovskii, 2009, Shababo et al., 2013], we wanted to ensure that the CAVIaR-inferred spikes were consistent with actual spikes. This would give credence to the ability of CAVIaR to generalize to new experiments. Notably, however, if two stimulated neurons share similar PSC waveforms but only one spiked, it would not be unreasonable to allocate the presynaptic spike to the non-spiking neuron for that particular trial. Nevertheless, by adopting a randomized experimental design, the accuracy of CAVIaR in simulation suggested that it could demarcate the contributions from each neuron provided that the spike inference was accurate on average across the experiment. We therefore performed multiple (n=4) paired recordings to obtain ground-truth spikes. We first used single-target stimulation to iden-

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

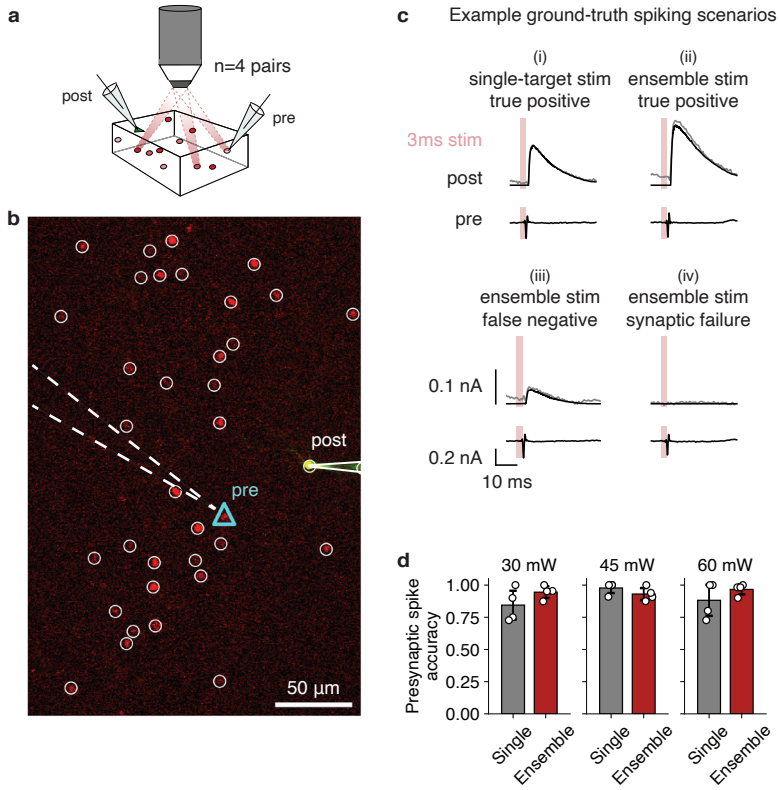


Figure 4: Validation of model calibration using paired recordings. **a**, Simultaneous whole-cell and cell-attached recordings provide ground-truth presynaptic spikes associated with postsynaptic responses during single-target and ensemble stimulation. **b**, Example image showing locations of presynaptic and postsynaptic neurons, among other segmented ROIs. **c**, Example scenarios from a paired recording associated with successful photoelicitation of presynaptic spikes. While trials (i) and (ii) resulted in correct predictions of presynaptic spiking by CAVIaR (i.e., are true positives), trials (iii) and (iv) were declared trials on which the stimulated neuron did not spike, either due to an unusually low-amplitude PSC or synaptic failure. Underlaid gray traces show raw PSC without NWD, black traces show PSC with NWD. Scales at bottom left apply to all subpanels. **d**, Performance of presynaptic spike inference over $n = 4$ paired recordings, stratified by single-target vs ensemble stimulation and laser power. Performance of presynaptic spike inference does not depend on power or single-target vs ensemble stimulation ($p > 0.05$ for all pairs, dependent t-test).

tify a likely presynaptic neuron in real time, then established a cell-attached patch clamp on the presynaptic neuron, and then randomly alternated between single-target and ensemble stimulation across multiple power levels while recording presynaptic spikes (Figure 4a,b). We confirmed that each of the neurons identified as connected using single-target stimulation following establishment of the patch clamp were also considered connected when the population was subsequently mapped using ensemble stimulation.

Paired recordings revealed a range of scenarios associated with ground-truth spiking (Figure 4c). This included true positives, where CAVIaR correctly inferred a presynaptic spike from a given PSC evoked by single-target or ensemble stimulation (scenarios (i) and (ii) in Figure 4c). This also included false negatives, however. For example, in some cases CAVIaR incorrectly estimated that an unusually small PSC did not arise by a stimulation-evoked spike of the patched presynaptic neuron but rather from another, more weakly connected neuron that was stimulated at the same time (scenario (iii)). False negatives also occurred due to synaptic failure, where the presynaptic neuron successfully elicited, but did not transmit, a spike (scenario (iv)). Despite the latter two challenging scenarios, CAVIaR achieved high accuracy in its overall estimates of presynaptic spikes with no significant degradation in its ability to infer presynaptic spikes for individual

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

neurons, even when stimulating ten neurons at once and over a range of laser powers (Figure 4d; all pair-wise differences not significant, $p > 0.05$, dependent t-test).

Inference of synaptic circuitry from ensemble stimulation across multiple cortical cell types

Next, we performed experiments mapping connectivity from ChroME2f-expressing PV and SST neurons using ten-target holographic ensemble stimulation performed at 30 Hz (Figure 5a). We demixed the optogenetically evoked PSCs using NWD and detected putative synapses using CAVIaR (Figure 5b, left columns). Single-target stimulation of the detected presynaptic neurons validated the presence of each synapse (Figure 5b, right columns). We then confirmed that the strengths of the synapses estimated from ensemble stimulation using CAVIaR were comparable to those estimated using single-target stimulation (Figure 5c, left columns; average differences in synaptic weight estimates for the shown examples: PV-pyramidal, 8.9%; SST-pyramidal, 20%).

Despite single-target confirmation of the presence and magnitude of the synapses identified using CAVIaR, we asked whether ensemble stimulation was systematically engaging "higher-order" or polysynaptic network effects. For example, PV neurons are known to provide high levels of mutual inhibition [Pfeffer et al., 2013], which could be triggered when stimulating many PV neurons at once. In this case, other PV neurons would be inhibited, potentially reducing the number of driven PV neurons and resulting in smaller inhibitory PSCs than would be expected by CAVIaR. On the other hand, SST neurons are known to primarily inhibit other (non-SST) cell types [Pfeffer et al., 2013], and are therefore less likely to influence postsynaptic responses through polysynaptic effects. In each case, the impact of such effects will depend closely on the sparsity and strength of the underlying connectivity.

To confirm that polysynaptic effects were not confounding our inferences, we performed an out-of-sample testing method that we refer to as leave-one-hologram-out cross-validation (LOHO-CV). Assuming that H different holograms are used in the experiment (where each hologram targets a different set of neurons), LOHO-CV works by running the CAVIaR algorithm on data corresponding to $H - 1$ holograms, such that the algorithm has never observed postsynaptic responses from simultaneous stimulation of the neurons targeted in the Hth hologram. Then, samples from the posterior distribution over the model parameters are used to predict the mean postsynaptic response to hologram H at three different power levels, a process that includes sampling over the model's uncertainty about whether spikes will be elicited by stimulation at any given laser power. The predicted response is compared against the held-out response, and the process continues with a different hologram selected to be held-out until the response to every hologram has been predicted.

LOHO-CV confirmed that CAVIaR achieved high accuracy in its predicted responses to novel holograms (Figure 5c, right panels). This indicated that CAVIaR did not miss synapses for which ensemble stimulation evoked large and reliable PSCs. Such synapses would manifest as large observed test responses, but with predicted responses of 0 (i.e., the test data would lie along the $y = 0$ axis in Figure 5c). Additionally, high LOHO-CV performance implied that CAVIaR did not make systematic errors arising due to higher-order network effects, which would manifest as consistent off-diagonal points in Figure 5c.

We also performed a small number of experiments (n = 2) mapping connectivity from ChroME2s-expressing pyramidal neurons onto patch-clamped ChroME2s-negative pyramidal neurons (Figure 5, right column). We considered whether the risk of polysynaptic effects was greater in this preparation because the stimulation of multiple pyramidal neurons could evoke suprathreshold excitatory input onto non-targeted neurons, driving

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

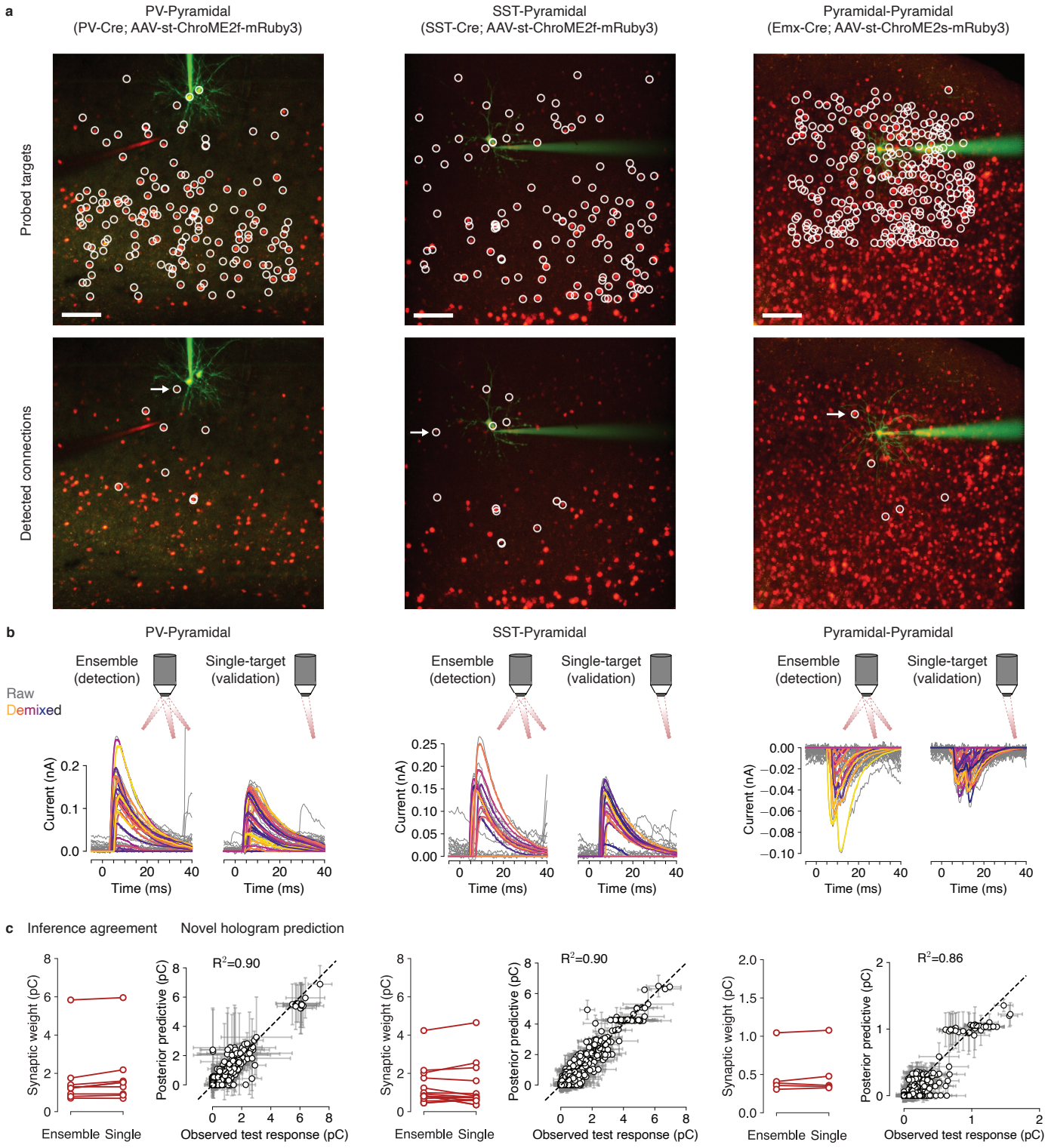


Figure 5: Model-based compressive inference and validation of synaptic connectivity from multiple cortical cell types. a, Z-projected visualization of probed targets (top) and monosynaptic connections detected across 5 planes using NWD and CAVIaR (bottom) for three different presynaptic cell types (PV, SST, Pyramidal). Scale bar, 50 μm. **b**, Example PSCs (corresponding to neurons marked by white arrows in (a)) evoked by holographic stimulation of ensembles containing a putatively connected neuron (left panels). Single-target stimulation of the putatively connected neuron validates the existence of the individual synapse (right panels). **c**, Agreement between synaptic weights identified using holographic ensemble stimulation and single-target stimulation (left panels, average differences in synaptic weight estimates 8.9%, PV-Pyramidal; 20%, SST-Pyramidal; 2% Pyramidal-Pyramidal). Prediction of postsynaptic responses to holographic stimulation of novel combinations of neurons (i.e., LOHO-CV; right panels). High coefficients-of-determination (R^2) indicate holographic stimulation is not engaging higher-order network effects that would not be recapitulated by the model. Error bars show mean $\pm$ 1 s.d.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

them to spike. However, we also reasoned that because stimulation is applied to 10 neurons at once, that pyramidal-pyramidal connectivity is known to be highly sparse [Hage et al., 2022], and that the resting potential of neurons in slices is low, polysynaptic effects could be negligible. Consistent with this reasoning, NWD and CAVIaR identified single-target-validated synapses and obtained high predictive accuracy using LOHO-CV (Figure 5a-c, columns labeled Pyramidal-Pyramidal; average difference in estimated synaptic weight for shown example, 2%).

Together, these results showed that, for the set of examples in Figure 5, NWD and CAVIaR enabled accurate recovery of synapses from fast ensemble stimulation without patterns of misestimation that would be consistent with the recruitment of confounding polysynaptic effects. However, we expect that additional tests will be required to ensure that accuracy remains unaffected by polysynaptic effects in more active networks such as with *in vivo* mapping experiments.

Synaptic connectivity from hundreds of presynaptic candidates can be reconstructed in minutes

We repeated our experimental protocol to map connectivity from 1,016 PV neurons (following removal of duplicate ROIs, see Methods) across 7 slice experiments. We confirmed that the ability to detect synapses from ensemble stimulation using NWD and CAVIaR was robust, as demonstrated via LOHO-CV analysis (Figure 6a; mean R^2, 0.8), number of connections identified via ensemble stimulation that were also validated via single-target stimulation (i.e., true positives, Figure 6b; mean true positive rate, 0.99), and the average percentage difference in synaptic weight between ensemble-identified and single-target-validated connections (Figure 6c; mean percentage difference across experiments, 10%).

We then wanted to determine how efficient our protocol was for recovering connectivity as a function of continuous stimulation time (namely, without regard to experiment time devoted to pausing stimulation to measure access resistance, adjusting the seal of the patch pipette, etc.). To do so, we subsampled mapping data from the total set of available trials and, for each subset, applied NWD and CAVIaR to estimate connectivity. This allowed us to monitor the rate of convergence of CAVIaR to its final estimates in increments of tens of seconds. Note that because our principal aim was validation of the connectivity inferences themselves rather than a demonstration of speed-up specifically, each hologram was presented 10 times at random throughout the experiment. This experimental design provided a means to characterize performance via LOHO-CV analysis, but could only establish a lower bound on mapping speed because simulations demonstrate that optimal speeds are obtained only when unique holograms are used on every trial (Figure S7b). Notwithstanding this provision, we found that our mapping protocol converged to an accuracy of 0.95 with less than just 2-3 minutes of stimulation at 30 Hz (Figure 6d, population sizes ranging from 101 to 327), corresponding to an average mapping time of 484 ms per neuron (obtained by dividing experiment time by number of neurons, Figure 6e,f; median, 367 ms; minimum, 99 ms).

How do these mapping times compare with our predictions in simulation? We simulated experiments mapping 300 neurons using 10-target ensemble stimulation at 30 Hz and repeated each hologram 10 times (as in the experimental data). The rate at which connectivity mapping using ensemble stimulation converged to its final estimates was very similar in simulation and experiment (Figure 6g), confirming that our simulated results are predictive of performance in real experiments. We also simulated mapping using single-target

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

stimulation at 10 Hz as a baseline, which, by comparison with ensemble stimulation, required more than 12 minutes of stimulation time to recover connectivity (i.e., to cross an R^2 of 0.95, Figure 6g), demonstrating a substantial mapping speed-up.

To characterize the dependence of PV-pyramidal connection probability and strength on distance, we aggregated our results across mapping experiments. Out of 1,016 presynaptic candidate neurons, CAVIaR identified 51 neurons that were connected (Figure 6h). While connection strengths were largely uniform across distance (with a small number of stronger connections in close proximity to the postsynaptic neuron, Figure 6i), we found that the strongest connections within individual mapping experiments could be far from the postsynaptic neuron (Figure S8). Connection probability was highest at close distances (0.42 up to 42 μm from the postsynaptic neuron, reducing to 0.25 by 84 μm, Figure 6j), similar to previous reports [Hage et al., 2022]. Finally, the ability of our connectivity mapping system to screen large numbers of synapses within individual experiments allowed us to quantify variability between entire maps (Figure S9).

NWD and CAVIaR can be applied to mapping experiments performed "blind"

The mapping experiments in Figure 5 were performed using soma-targeted opsins with the constitutively fluorescent protein mRuby3 simultaneously expressed in cell nuclei. Opsin-expressing neurons could therefore be straightforwardly segmented from an image of mRuby3 expression and subsequently stimulated. However, in some cases opsin-expressing neurons do not also express a separate nuclear fluorophore for straightforward identification of photosensitive cell nuclei [Bounds et al., 2021]. We therefore considered whether our connectivity mapping system could be used in the case where neuron locations are not known *a priori* (i.e., when mapping is performed "blind"). In this context, the field of view is divided into a three-dimensional grid and locations are stimulated one grid point at a time [Wang et al., 2007, Baker et al., 2016, Naka et al., 2019, Bounds et al., 2021].

We first applied NWD to pyramidal-pyramidal connectivity mapping data collected via grid stimulation with transgenic mice expressing st-ChroME-GCaMP7s (Figure 7a). In this example the postsynaptic neuron expressed opsin, leading to a false positive region in the raw synaptic weight maps corresponding to photocurrents induced by direct stimulation (Figure 7a, white dashed square). Application of NWD led to two notable outcomes. First, because direct stimulation of an opsin-positive neuron elicits photocurrents with submillisecond latency [Sridharan et al., 2022], NWD largely subtracted the photocurrents from the electrophysiological traces because they violated the admissible conditions established during training of the network. This caused a suppression of the false positive region in the resulting synaptic weight map (Figure 7a, columns labelled "Demixed"). Second, electrical noise and ongoing spontaneous activity led to substantial background noise in the raw weight maps, blurring together neural regions of interest (ROIs). NWD overwhelmingly eliminated this noise, yielding sharper borders surrounding neural shapes and more clearly reflecting the underlying physiology.

We then applied CAVIaR to the demixed grid mapping data. CAVIaR allowed us to determine which pixels corresponded to connected neurons and the strengths of their corresponding synapses (Figure 7a, "CAVIaR" column). Furthermore, because CAVIaR estimates synaptic weights by identifying trials on which presynaptic spikes were elicited and successfully transmitted to the postsynaptic neuron, we could subsequently decouple synaptic weight from spike probability, resulting in even sharper edges surrounding neural ROIs. (Figure 7a,

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

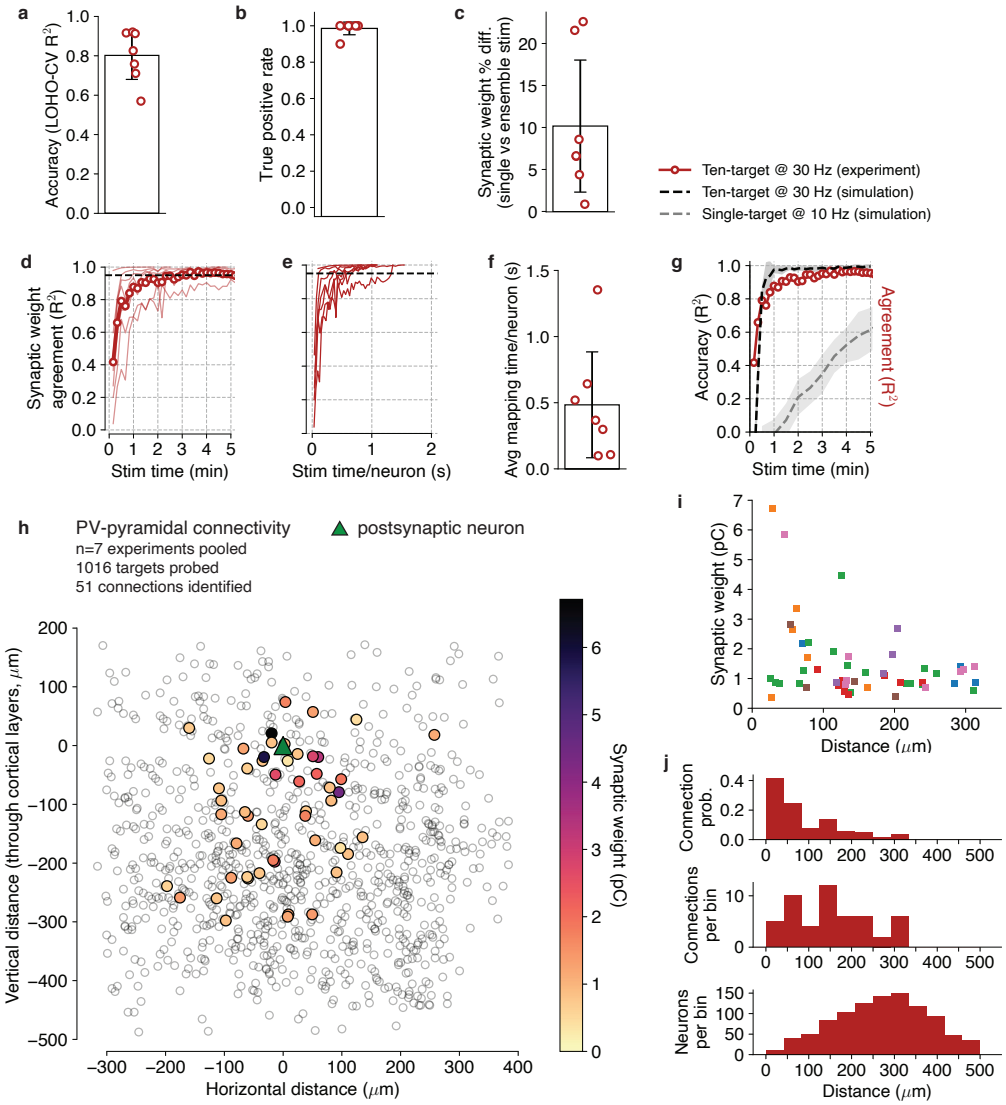


Figure 6: Summary of mapping performance across multiple (n=7) experiments. a, LOHO-CV results across experiments (mean R^2, 0.8 $\pm$ 0.12). **b**, Summary of true positive rates across experiments (mean, 0.99 $\pm$ 0.03). True positive defined as connection identified via ensemble stimulation that was also validated via single-target stimulation. **c**, Summary of average percentage differences between connections identified using single-target stimulation and ensemble stimulation (mean percentage difference, 10% $\pm$ 7.9%). **d**, Minutes time-scale reconstruction of synaptic connectivity *in vitro*, given by agreement (measured by R^2) between estimates at each time point and the end of the experiment. Faint lines show convergence of connectivity estimates for individual experiments. Dark line shows median at each timestep (open circles). **e**, Rate of convergence per neuron, obtained by normalizing stimulation times reported in panel (d) by population sizes. **f**, Average time taken to cross an agreement of 0.95, per neuron (mean time in seconds, 0.48 $\pm$ 0.4, median 0.37). **g**, Comparison of rates of convergence between compressive connectivity mapping in simulation (dashed black line) and experiment (open red circles). Close agreement between the two rates of convergence indicates simulated results are predictive of behavior in experiment. Single-target mapping at 10 Hz shown as baseline (dashed gray line). **h**, Z-projected map of connection strength across 7 pooled experiments. Colored circles indicate position of identified connections, open gray circles indicate position of all probed cells. Individual maps aligned to have postsynaptic neuron at location (0, 0). **i**, Connection strength vs distance. Color of each data point indicates experiment. **j**, Connection probability as a function of distance. Histograms display the ratio of connected to all probed cells (top), number of connected cells (middle), and number of probed cells (bottom). Histogram bin size, 42 μm.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

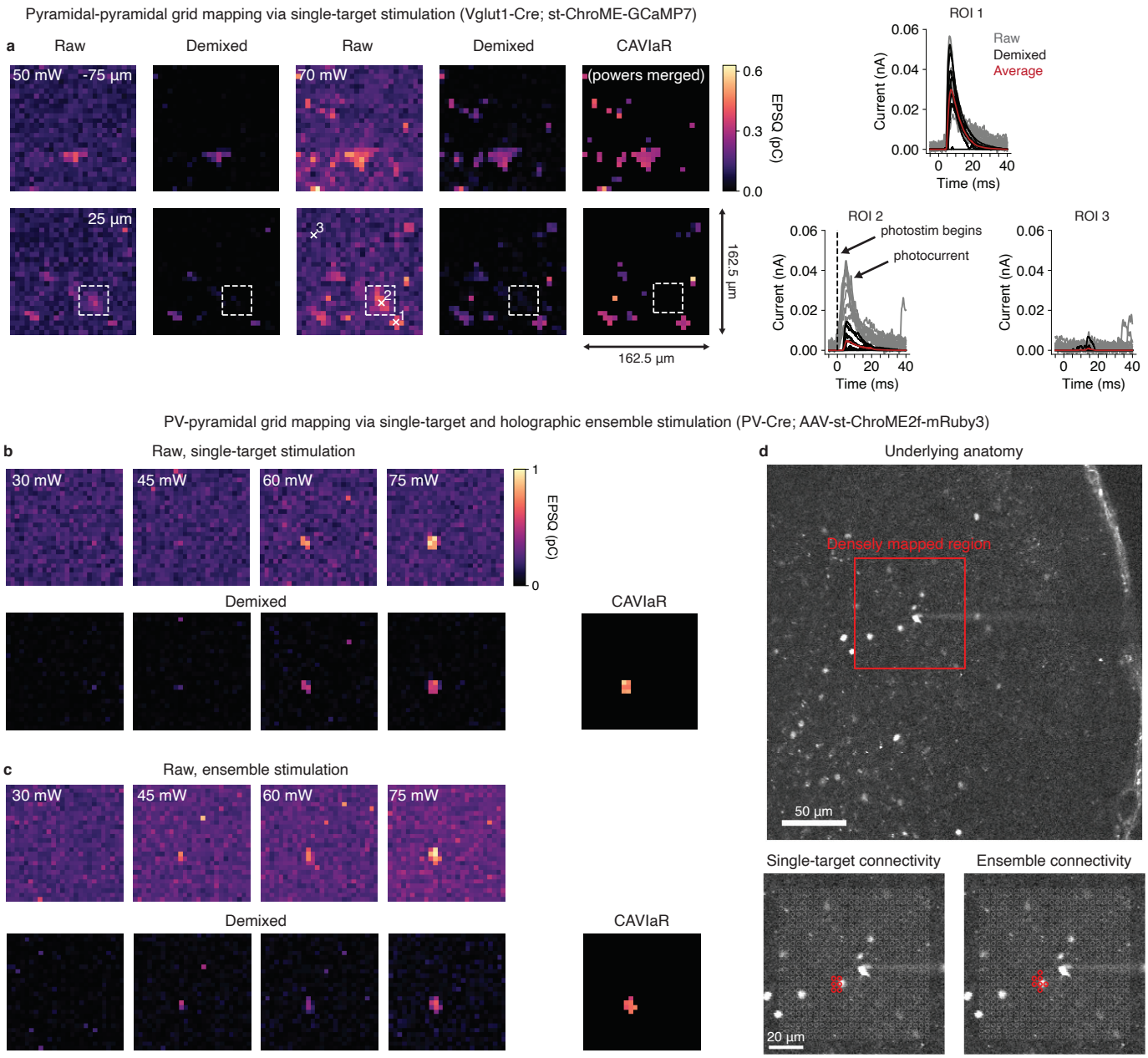


Figure 7: Demixing and connectivity inference for "blind" grid mapping experiments. **a**, Grid mapping of pyramidal-pyramidal connectivity using single-target stimulation at two powers. Two out of five planes are shown as an example. Note that maps of synaptic connectivity obtained by CAVIaR "merge" the multi-power demixed maps into a single map. Dashed white square shows region with suppression of direct photocurrents (ROI 2) resulting from stimulating the patched, opsin-expressing postsynaptic neuron. Right panels show the effect of NWD on individual PSC traces from three example ROIs (numbered white crosses on grid maps). **b**,**c**, Grid mapping of PV-pyramidal connectivity using both single-target (on plane 0 μm) and holographic ensemble stimulation (on five planes over 25 to -75 μm, separated by 25 μm each; only plane 0 μm shown for comparison) over four powers. Agreement between the two maps validates the use of CAVIaR in this regime. Each pixel in the raw maps shown in (c) obtained by averaging across all PSCs evoked by stimulation of an ensemble containing that pixel. **d**, Overlay of connected pixels (red circles) on underlying anatomy (obtained by imaging expression of mRuby) confirms that grid mapping using both single-target and ensemble stimulation correctly identifies an opsin-expressing neuron. Note that no such anatomical image is available for the Vglut1-Cre; st-ChroME-GCaMP7 mice (panel (a)) since the opsin construct does not include a constitutively fluorescent protein (e.g., mRuby).

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

compare 70 mW "Demixed" and "CAVIaR" columns). While the synaptic weights of pixels associated with putative neurons were approximately constant across the ROI, presynaptic spike probability decreased and excitatory PSC latency increased with increasing distance of the holographic stimulation site from the putative neuron centroid (Figure S10, Figure S11).

To confirm that putative neural ROIs obtained via grid stimulation correspond to actual opsin-expressing neurons, we repeated the grid mapping protocol using a ChroME2f construct with nuclear expression of mRuby3. This allowed us to directly overlay the stimulation grid on the underlying anatomical locations of opsin-expressing neurons (Figure 7d), which confirmed that the identified ROI indeed aligned with an opsin-expressing neuron. We hypothesized that the grid mapping protocol could be greatly accelerated by exploiting holographic stimulation of multiple grid locations at once, analogous to the results we obtained for mapping segmented nuclei. We performed 10-target stimulation of grid points at 30 Hz and visualized a "raw" connectivity map by averaging across the postsynaptic charges on trials for which a given grid point was stimulated (Figure 7c). We then applied NWD and CAVIaR to the ensemble mapping data, which identified the same ROI that was detected using single-target stimulation, and similarly aligned with the same opsin-expressing neuron in the anatomical image (Figure 7c,d; see Figure S11 for full, multi-plane analysis of grid mapping experiment). This confirmed that NWD and CAVIaR could also facilitate rapid connectivity mapping in the blind stimulation regime.

Discussion

Uncovering how neural computations are implemented in the cortex requires knowledge of the existence and strength of the underlying synaptic connections. We therefore developed a pair of complementary computational tools (NWD and CAVIaR) that together enable high-speed acquisition of large-scale synaptic connectivity maps using two-photon holographic optogenetics. Previous attempts at compressive connectivity mapping have only been demonstrated in simplified simulations [Hu and Chklovskii, 2009, Shababo et al., 2013, Draelos and Pearson, 2020], making our approach the first experimentally deployed implementation. This required considerable methodological extensions over previous algorithms to ensure that connectivity is accurately reconstructed in light of the substantial and unobserved sources of biophysical variability intrinsic to real holographic ensemble stimulation experiments. We found that use of NWD and CAVIaR provided a mapping speed-up of over an order of magnitude compared to existing approaches in realistic settings, granting the speed and scalability necessary to collect large-scale maps of connectivity within individual experimental sessions.

To implement and validate these tools in real mapping experiments we used cutting edge optogenetic technologies that offered several key advantages. Namely, since the ChroME2f opsin is ultrapotent with fast off-kinetics, we could titrate the pulse width and range of stimulation laser powers such that with high probability stimulation generated either 0 or 1 presynaptic spike(s), with minimal instances of more than one spike [Sridharan et al., 2022]. This facilitated tractable presynaptic spike inference because spikes could then be well-modelled by a latent power-dependent Bernoulli random variable. By comparison, potent opsins with slow off-kinetics (such as ChRmine [Marshel et al., 2019]) elicit larger and less predictible presynaptic spike counts that make connectivity inference much more computationally challenging as the combinatorial space of how presynaptic spikes can be allocated across the stimulated population grows rapidly, precluding efficient model

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

identifiability.

We also relied on the fact that ChroME2f expression is targeted to the soma through a fusion with the Kv2.1 tag [Mardinly et al., 2018, Sridharan et al., 2022], reducing the likelihood of triggering spikes in untargeted neurons by inadvertent stimulation of nearby neurites [Shemesh et al., 2017]. Moreover, we used two-photon holographic light sculpting to illuminate entire cells at once [Papagiakoumou et al., 2010, Hernandez et al., 2016, Pégard et al., 2017] and randomly ramped the stimulation laser power up and down across three or more powers, effectively mapping the system at multiple resolutions. Stimulating with low power shrinks the optical point spread function (PSF, notably in the axial dimension [Pégard et al., 2017]) and therefore minimizes the risk of off-target stimulation even further, but may fail to elicit spikes in neurons with low opsin expression or high rheobases. Notably, it cannot in general be known ahead of time whether a lack of optogenetically evoked PSCs was due to the stimulated neuron not being connected or because of a failure to photoelicit spikes, an ambiguity that can only be resolved through high power stimulation. On the other hand, stimulating with high power inflates the PSF, yielding a less spatially precise stimulus but eliciting spikes with high probability. We therefore elected to use a randomized design that balanced the trade-off between PSF size and power exploration, ultimately allowing us to map multiple principal inhibitory interneuron subtypes with high precision.

One potential concern with our proposed connectivity mapping approach is that it may engage mechanisms for short-term plasticity given our suggested high speed of stimulation and the fact that entire ensembles of neurons are stimulated at once. To help calibrate mapping protocols, we derived a simple combinatorial expression for the expected interstimulus interval (ISI, in seconds) for any given neuron (Methods). Assuming connectivity from N total neurons is mapped by stimulating a random ensemble of size R on each trial at a speed of f Hz, the mean ISI follows

$$\text{ISI} = \frac{N}{Rf}. \tag{1}$$

Thus, a typical experiment mapping ∼300 neurons with 10-target stimulation at 30 Hz yields an expected stimulation return time of 1 second. An additional concern associated with ensemble stimulation is the recruitment of higher-order network effects arising due to recurrent connectivity in the cortex. However, using LOHO-CV we confirmed that this was not a major factor influencing the quality of our inferences (Figure 5c), though we expect this to depend on the system and number of simultaneously illuminated targets. For example, optogenetic stimulation of even single CA1 pyramidal neurons (expressing an opsin via single-cell electroporation) appears to reliably recruit numerous other unstimulated CA1 pyramidal and interneurons [Geiller et al., 2022]. While it is likely that NWD will assist with deconfounding PSC measurements in this context (since such higher-order effects would be associated with longer latencies), more research will be required to determine the extent to which additional computational or experimental solutions will be necessary to enable high-speed mapping in the hippocampus.

We demonstrated the applicability of NWD and CAVIaR by mapping connectivity from three cell-types (PV, SST, and pyramidal neurons), using three ChroME variants (ChroME, ChroME2f, and ChroME2s), and using both targeted and blind experimental designs. In every case, application of our tools led to a substantial improvement in signal quality and speed of connectivity mapping. While in our experiments we typically mapped connectivity from several hundred neurons, this was only limited by factors outside of the influence of NWD and CAVIaR, such as the number of neurons targetable by our SLM or the rate at which phase masks could be computed. Our simulations predict that the scale and throughput of compressive connectivity mapping will

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

only grow with the adoption of experimental steps to increase the targetable field of view (e.g., by translating the microscope or SLM stimulation field), with larger commercially available SLMs, or with even more potent opsins enabling stimulation of more targets for the same laser power. Such advances would allow for precise, low-latency holographic control over larger neural ensembles or for mapping larger population sizes than those used here.

Our results indicate that NWD and CAVIaR can substantially reduce the cumulative time spent stimulating tissue. However, the speed of a mapping experiment can also be impacted by other factors, including the time required to identify and segment neurons, monitor the access resistance of the patch clamp, and compute hologram phase masks. Each of these factors similarly affects mapping performed using single-target stimulation and ensemble stimulation. While acquiring an image stack to identify presynaptic candidates is required for all experiments using targeted stimulation, measuring access resistance poses an additional trade-off between the speed of a mapping experiment and the desire to make regular quality checks. An ideal experimental design would therefore proceed with the longest periods of uninterrupted stimulation possible to yield the maximal benefits of NWD and CAVIaR, but with brief ($\sim$200 ms) automatic pauses included to check access resistance. The attending experimenter could then be prompted to address inadequate access only when necessary. In the case of phase mask generation, efficient computation is critical since otherwise experiment time could be dominated by this step. We expect that high-speed phase mask estimation methods, such as those based on deep learning [Eybposh et al., 2020, Eybposh et al., 2022], could play an increasingly important role in rapid connectivity mapping.

Combining NWD and CAVIaR with two-photon holographic optogenetics should enable new experiments where, with just a few minutes of rapid ensemble stimulation, one can obtain an extensive map of connectivity. This could allow for very high throughput collection and screening of synaptic connectivity maps from brain slices, or for *in vivo* applications where anatomical connectivity can be directly related to functional activity. Critically, *in vivo* experiments involve larger population sizes relative to *in vitro* experiments and are impacted by higher background rates of spontaneous activity and more prevalent polysynaptic effects. While further validation must be performed to ensure accuracy *in vivo*, the impact of each of these impediments is demonstrably reduced by our computational methods, implying that NWD and CAVIaR may be essential tools for high-precision mapping in this experimental regime.

Acknowledgements

The authors thank Kenneth Kay, Darcy Peterka, Ben Shababo, and Shizhe Chen for many insightful discussions and helpful suggestions. We thank Karthika Gopakumar for helping with animal husbandry and the opsin expression procedure. This work was funded by NIH awards 1RF1MH120680 and 1U19NS107613-01 to HA and LP. MAT, BA, and LP were supported by the Gatsby Charitable Foundation and NSF NeuroNex award 1707398.

Code availability

CAVIaR is implemented in Python using JAX [Bradbury et al., 2018], can be run on GPU, and is freely available at https://github.com/marcustriplett/circuitmap together with a PyTorch Lightning implementation of NWD.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

[Adesnik and Abdeladim, 2021] Adesnik, H. and Abdeladim, L. (2021). Probing neural codes with two-photon holographic optogenetics. *Nature Neuroscience*, 24(10):1356–1366.

[Ayer et al., 1955] Ayer, M., Brunk, H. D., Ewing, G. M., Reid, W. T., and Silverman, E. (1955). An empirical distribution function for sampling with incomplete information. *The annals of mathematical statistics*, pages 641–647.

[Baker et al., 2016] Baker, C. A., Elyada, Y. M., Parra, A., and Bolton, M. M. (2016). Cellular resolution circuit mapping with temporal-focused excitation of soma-targeted channelrhodopsin. *Elife*, 5:e14193.

[Bishop, 2006] Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*, volume 4. Springer.

[Blei et al., 2017] Blei, D. M., Kucukelbir, A., and McAuliffe, J. D. (2017). Variational inference: A review for statisticians. *Journal of the American statistical Association*, 112(518):859–877.

[Bounds et al., 2021] Bounds, H. A., Sadahiro, M., Hendricks, W. D., Gajowa, M., Oldenburg, I. A., Gopakumar, K., Quintana, D., Daigle, T., Zeng, H., and Adesnik, H. (2021). Multifunctional cre-dependent transgenic mice for high-precision all-optical interrogation of neural circuits. *bioRxiv*.

[Bradbury et al., 2018] Bradbury, J., Frostig, R., Hawkins, P., Johnson, M. J., Leary, C., Maclaurin, D., Necula, G., Paszke, A., VanderPlas, J., Wanderman-Milne, S., and Zhang, Q. (2018). JAX: composable transformations of Python+NumPy programs.

[Candes et al., 2006] Candes, E. J., Romberg, J. K., and Tao, T. (2006). Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 59(8):1207–1223.

[Candes and Tao, 2006] Candes, E. J. and Tao, T. (2006). Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE transactions on information theory*, 52(12):5406–5425.

[Donoho, 2006] Donoho, D. L. (2006). Compressed sensing. *IEEE Transactions on information theory*, 52(4):1289–1306.

[Draelos and Pearson, 2020] Draelos, A. and Pearson, J. (2020). Online neural connectivity estimation with noisy group testing. *Advances in Neural Information Processing Systems*, 33.

[Eybposh et al., 2020] Eybposh, M. H., Caira, N. W., Atisa, M., Chakravarthula, P., and Pégard, N. C. (2020). Deepcgh: 3d computer-generated holography using deep learning. *Optics Express*, 28(18):26636–26650.

[Eybposh et al., 2022] Eybposh, M. H., Curtis, V. R., Rodríguez-Romaguera, J., and Pégard, N. C. (2022). Advances in computer-generated holography for targeted neuronal modulation. *Neurophotonics*, 9(4):041409.

[Falk et al., 2019] Falk, T., Mai, D., Bensch, R., Çiçek, Ö., Abdulkadir, A., Marrakchi, Y., Böhm, A., Deubner, J., Jäckel, Z., Seiwald, K., et al. (2019). U-net: deep learning for cell counting, detection, and morphometry. *Nature methods*, 16(1):67–70.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

[Fletcher et al., 2011] Fletcher, A. K., Rangan, S., Varshney, L. R., and Bhargava, A. (2011). Neural reconstruction with approximate message passing (neuramp). In *NIPS*, pages 2555–2563.

[Geiller et al., 2022] Geiller, T., Sadeh, S., Rolotti, S. V., Blockus, H., Vancura, B., Negrean, A., Murray, A. J., Rózsa, B., Polleux, F., Clopath, C., et al. (2022). Local circuit amplification of spatial selectivity in the hippocampus. *Nature*, 601(7891):105–109.

[Gerchberg, 1972] Gerchberg, R. W. (1972). A practical algorithm for the determination of plane from image and diffraction pictures. *Optik*, 35(2):237–246.

[Hage et al., 2022] Hage, T. A., Bosma-Moody, A., Baker, C. A., Kratz, M. B., Campagnola, L., Jarsky, T., Zeng, H., and Murphy, G. J. (2022). Synaptic connectivity to l2/3 of primary visual cortex measured by two-photon optogenetic stimulation. *Elife*, 11:e71103.

[Hernandez et al., 2016] Hernandez, O., Papagiakoumou, E., Tanese, D., Fidelin, K., Wyart, C., and Emiliani, V. (2016). Three-dimensional spatiotemporal focusing of holographic patterns. *Nature communications*, 7(1):1–11.

[Hu and Chklovskii, 2009] Hu, T. and Chklovskii, D. (2009). Reconstruction of sparse circuits using multi-neuronal excitation (rescume). *Advances in Neural Information Processing Systems*, 22:790–798.

[Lee et al., 2020] Lee, J., Mitelut, C., Shokri, H., Kinsella, I., Dethe, N., Wu, S., Li, K., Reyes, E. B., Turcu, D., Batty, E., et al. (2020). Yass: Yet another spike sorter applied to large-scale multi-electrode array recordings in primate retina. *bioRxiv*.

[Mardinly et al., 2018] Mardinly, A. R., Oldenburg, I. A., Pégard, N. C., Sridharan, S., Lyall, E. H., Chesnov, K., Brohawn, S. G., Waller, L., and Adesnik, H. (2018). Precise multimodal optical control of neural ensemble activity. *Nature neuroscience*, 21(6):881–893.

[Marshel et al., 2019] Marshel, J. H., Kim, Y. S., Machado, T. A., Quirin, S., Benson, B., Kadmon, J., Raja, C., Chibukhchyan, A., Ramakrishnan, C., Inoue, M., et al. (2019). Cortical layer–specific critical dynamics triggering perception. *Science*, 365(6453):eaaw5202.

[Merel et al., 2016] Merel, J., Shababo, B., Naka, A., Adesnik, H., and Paninski, L. (2016). Bayesian methods for event analysis of intracellular currents. *Journal of Neuroscience Methods*, 269:21–32.

[Mishchenko and Paninski, 2012] Mishchenko, Y. and Paninski, L. (2012). A bayesian compressed-sensing approach for reconstructing neural connectivity from subsampled anatomical data. *Journal of computational neuroscience*, 33(2):371–388.

[Naka et al., 2019] Naka, A., Veit, J., Shababo, B., Chance, R. K., Risso, D., Stafford, D., Snyder, B., Egladyous, A., Chu, D., Sridharan, S., et al. (2019). Complementary networks of cortical somatostatin interneurons enforce layer specific control. *Elife*, 8:e43696.

[Needell and Tropp, 2009] Needell, D. and Tropp, J. A. (2009). Cosamp: Iterative signal recovery from incomplete and inaccurate samples. *Applied and computational harmonic analysis*, 26(3):301–321.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

[Packer et al., 2012] Packer, A. M., Peterka, D. S., Hirtz, J. J., Prakash, R., Deisseroth, K., and Yuste, R. (2012). Two-photon optogenetics of dendritic spines and neural circuits. *Nature methods*, 9(12):1202–1205.

[Pakman and Paninski, 2014] Pakman, A. and Paninski, L. (2014). Exact hamiltonian monte carlo for truncated multivariate gaussians. *Journal of Computational and Graphical Statistics*, 23(2):518–542.

[Papagiakoumou et al., 2010] Papagiakoumou, E., Anselmi, F., Bègue, A., De Sars, V., Glückstad, J., Isacoff, E. Y., and Emiliani, V. (2010). Scanless two-photon excitation of channelrhodopsin-2. *Nature methods*, 7(10):848–854.

[Papagiakoumou et al., 2020] Papagiakoumou, E., Ronzitti, E., and Emiliani, V. (2020). Scanless two-photon excitation with temporal focusing. *Nature Methods*, 17(6):571–581.

[Pégard et al., 2017] Pégard, N. C., Mardinly, A. R., Oldenburg, I. A., Sridharan, S., Waller, L., and Adesnik, H. (2017). Three-dimensional scanless holographic optogenetics with temporal focusing (3d-shot). *Nature communications*, 8(1):1–14.

[Pfeffer et al., 2013] Pfeffer, C. K., Xue, M., He, M., Huang, Z. J., and Scanziani, M. (2013). Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. *Nature neuroscience*, 16(8):1068–1076.

[Printz et al., 2021] Printz, Y., Patil, P., Mahn, M., Benjamin, A., Litvin, A., Levy, R., Bringmann, M., and Yizhar, O. (2021). Determinants of functional synaptic connectivity among amygdala-projecting prefrontal cortical neurons. *bioRxiv*.

[Rickgauer and Tank, 2009] Rickgauer, J. P. and Tank, D. W. (2009). Two-photon excitation of channelrhodopsin-2 at saturation. *Proceedings of the National Academy of Sciences*, 106(35):15025–15030.

[Ronneberger et al., 2015] Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer.

[Shababo et al., 2013] Shababo, B., Paige, B., Pakman, A., and Paninski, L. (2013). Bayesian inference and online experimental design for mapping neural microcircuits. *Advances in Neural Information Processing Systems*, 26:1304–1312.

[Shemesh et al., 2017] Shemesh, O. A., Tanese, D., Zampini, V., Linghu, C., Piatkevich, K., Ronzitti, E., Papagiakoumou, E., Boyden, E. S., and Emiliani, V. (2017). Temporally precise single-cell-resolution optogenetics. *Nature neuroscience*, 20(12):1796–1806.

[Sridharan et al., 2022] Sridharan, S., Gajowa, M. A., Ogando, M. B., Jagadisan, U. K., Abdeladim, L., Sadahiro, M., Bounds, H. A., Hendricks, W. D., Turney, T. S., Tayler, I., et al. (2022). High-performance microbial opsins for spatially and temporally precise perturbations of large neuronal networks. *Neuron*.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

[Wang et al., 2007] Wang, H., Peca, J., Matsuzaki, M., Matsuzaki, K., Noguchi, J., Qiu, L., Wang, D., Zhang, F., Boyden, E., Deisseroth, K., et al. (2007). High-speed mapping of synaptic connectivity using photostimulation in channelrhodopsin-2 transgenic mice. *Proceedings of the National Academy of Sciences*, 104(19):8143–8148.

Methods

Experimental methods

Animals

All experiments on animals were conducted with approval of the Animal Care and Use Committee of the University of California, Berkeley. In all experiments we attempted to use male and female mice equally. Mice used for experiments in this study were transgenic Emx-Cre, PV-Cre, or SST-Cre mice obtained by crossing the corresponding lines in-house with a wild type (CD-1 (ICR) white strain, obtained from Charles River). Mice were housed in cohorts of five or fewer in a reverse light:dark cycle of 12:12 hours, with experiments occurring during the dark phase.

Opsin expression method

We used a neonatal injection procedure to induce expression of ChroME2s (AAV9.CAG.DIO.ChroME-ST.P2A.H2B-mRuby6) in the visual cortex of Emx-Cre animals or ChroME2f (AAV9.CAG.DIO.Chrome2f.P2A.H2B-mRuby3.WPRE.SV43) in PV-Cre and SST-Cre animals. Both constructs expressed the mRuby3 fluorophore in opsin-positive cell nuclei. Young pups at P3 or P4 were anaesthetized by placing them on ice for approximately 3 minutes. Next, each animal was stabilized under the nanoliter-injector (WPI) and a small portion (30 nL/injection) of virus was injected directly in 3-5 places around V1 via the skin and skull and at 3-5 depths to target L2/3. After the procedure the animal was placed on a heating pad until it recovered. At the end of the procedure the injected litter was returned to their cage with their parents and housed together until reaching approximately 21 days of age.

3D-SHOT holography set up

All experiments were performed using the 3D-SHOT multiphoton holography setup (see [Mardinly et al., 2018] for details). Briefly, the setup was custom built around a commercial Sutter MOM (movable object microscope) platform (Sutter Instruments) and combined a 3D photostimulation path, a fast resonant-galvo two-photon raster scanning imaging path and a wide-field one-photon epifluorescence/IR (infrared) transmitted light imaging path. The stimulation and imaging beams were merged together using a polarizing beamsplitter placed before the microscope tube lens.

A femtosecond fiber laser was used for two-photon photostimulation (Monaco 1035-80-60; 1040 nm, 1 MHz, 300 fs, Coherent). The stimulation laser was directed onto a blazed diffraction grating (600 l/mm, 1000 nm blaze, Edmund Optics 49-570) for temporal focusing. In order to be able to utilize the total available laser power (60 W laser output), the beam was enlarged by a factor of 2.5 to prevent heat damage of the grating

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

surface. The spot on the grating was relayed onto a rotating diffuser where it formed a temporally focused spot. The rotating diffuser was used to both randomize the phase pattern imprinted on the temporally focused spot and to expand the beam in the direction orthogonal to the temporal focusing direction and fully fill the spatial light modulator (HSP 1920 1920x1152 pixels, Meadowlark Optics). The SLM plane was relayed through 4f systems to the back aperture of an Olympus 20x water immersion objective, resulting in custom 3D distribution of temporally-focused spots at the focus of the objective. Holographic phase masks were calculated using the iterative, in-house written and GPU-optimized Gerchberg-Saxton algorithm [Gerchberg, 1972] and the intensity distribution was corrected to accommodate for diffraction efficiencies.

The two-photon imaging path relied on a Ti:sapphire laser, Mai Tai (Spectra Physics), with external power control via Pockels cell (Conoptics, Inc). For fast raster scanning, the system was equipped with conjugated 8 kHz resonant galvo-galvo systems. The imaging path hardware was controlled by ScanImage software and custom Matlab code was used to control the spatial light modulator for targeted photostimulation and synchronize with imaging.

Epifluorescent one-photon excitation was via a Spectra X (Lumencor) light source filtered by an appropriate excitation filter set. For slice transillumination we used a 750 nm and IR diffuser. The image was collected using an Olympus 20x magnification water-immersion objective and a CCD camera and displayed on a screen enabling targeted patch clamping.

Slice electrophysiology

In vitro slice recordings were performed on 300 μm-thick coronal slices coming from 4–6-week-old animals expressing opsin in L2/3 of V1. During slicing the level of opsin expression was checked using a simple laser light to visualize the targeting of the opsin to V1 and the general brightness of the mRuby3 nuclear marker.

Whole-cell patch-clamp protocols were performed in ACSF (artificial cerebrospinal fluid) perfusion solution (in mM: NaCL 119, NaHCO3 26, Glucose 20, KCl 2.5, CaCl 2.5, MgSO4 1.3, NaH2PO4 1.3) in temperature-controlled (33°C) conditions. Patch pipettes (4-7 MΩ) were pulled from borosilicate glass filaments (Sutter Instruments) and filled with a Cesium (Cs2+)-based internal solution (in mM: 135 CeMeSO4, 3 NaCl, 10 HEPES, 0.3 EGTA, 4 Mg-ATP, 0.3 Na-GTP, 1 QX-314, 5 TEA-Cl, 295 mOsm, pH=7.45) also containing 50 μM Alexa Fluor hydrazide 488 or 594 dye (ThermoFisher Scientific). For loose-patch recordings the pipettes were filled with standard ACSF. Data was recorded at 20 kHz using a 700b Multiclamp Axon Amplifier (Molecular Devices). The headstage with the electrode holder (G23 Instruments) was controlled by a motorized micro-manipulator (MP285A, Sutter Instruments). All data was acquired and analyzed with custom code written in Matlab using the National Instruments Data Acquisition Toolbox.

Learning physiological point spread functions

An opsin-positive cell was loose-patched at various slice depths between 20-90 μm and a volume of tissue surrounding that cell was then probed using a dense grid of holograms. The total size of the grid was 65x65x75 μm, resulting in a 10x10x7 voxel grid. The interval between the centers of the hologram targets were 5 pixels (6.5 μm) in the x/y dimensions and 12.5 μm in the z dimension. The phase masks of the grid holograms were pre-computed and stored in memory before performing the measurement. Two versions of the physiological point spread function experiments were performed. One where each hologram was a single spot randomly

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

selected from the grid, and another where each hologram contained 10 spots: 1 spot randomly taken from the grid and 9 spots (either fixed or randomly placed) outside the grid.

Evoked spiking in response to 3-5 ms laser stimulation (30 Hz) across powers between 10-60 mW/spot were collected and analysed using a custom written script in Matlab. Each trial was randomly repeated 5-7 times. Evoked spiking was calculated by simple thresholding and the mean spiking probability across each grid point and powers was calculated for single and 10-spot ensemble experiments. A Gaussian fit to the per-cell normalized mean spiking probability data was applied to estimate the FWHM of the spiking physiological point spread function.

Whole-cell targeted mapping experimental protocol

First, a L2/3 opsin-negative cell of pyramidal shape was sealed on (below 40 μm of cortical depth). The tissue was then quickly imaged with approximately 40 frames at 4-5 planes spaced by 25 μm. The cell was always positioned to be in the second plane from the top within the collected stack. The position of the presynaptic candidates were automatically identified by an in-house algorithm detecting round shapes in a specified FOV due to the presence of mRuby3 fluorophore in the opsin-positive cell nuclei.

Next, we computed 10-20 different sets of holograms targeting ensembles of presynaptic candidates. Stable access to the cell was obtained during hologram computation. The cell properties (Rm and Rs) were checked and the response to one-photon photocurrent characterized. Following completion of hologram computation, presynaptic candidates were mapped using randomly interleaved experimental trials containing single-target holograms or samples from the set of 10-target holograms. Each hologram was repeated repeated 7-15 times (typically 10) across 3-4 laser powers. For example, a map containing 100 presynaptic candidates would contain data from 45 trials of 100 single-target holograms across 3 powers (15 trials/power condition) and 45 trials of 100 different 10-target holograms (100 presynaptic candidates split into 10 different 10-target holograms that were then arranged randomly in 10 sets) across 3 powers (15 trials/power condition), resulting in 495 total trials (45 single-target sets and 450 10-target sets). Importantly, the order of all holograms and powers within each repetition was randomized. The Rm and Rs parameters were controlled and frequently logged during the mapping protocol. Every 10 trials (or 20-180 seconds, depending on the number of presynaptic targets and the power condition), a series of hyperpolarizing steps was applied and the Rm and Rs logged. During offline analysis, trials when Rs significantly changed or were above 30 MOhm were excluded from the data. Experiments where no connections were identified were also excluded from the analysis in Figure 6.

Pair-patch experimental protocol

The pair-patch mapping protocol was similar to the single-patch protocol, but contained an additional 'pre-mapping' step. After establishing stable access to the opsin-negative pyramidal cell, all presynaptic candidates were probed with 10 trials of single-power, single-target holograms. A fast online analysis based on the z-score provided the coordinates of cells that responded to this simple connection screening procedure. One such identified cell would be approached and loose-patched with a second electrode while monitoring the stability of the postsynaptic recording. The main prerequisite for deciding which cell would be loose-patched (from a handful of putatively connected cells obtained via our simple screening procedure) was the ability to identify

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

the cell in the slice under IR illumination and its accessibility. While the loose patch was established, the surrounding area would be re-imaged to update the positions of the presynaptic candidates. Mapping would then be performed according to the above-described procedure, while recording both postsynaptic responses and presynaptic spiking.

Whole-cell grid mapping experimental protocol

Opsin-negative cells (or those expressing minimal amounts of opsin) were whole-cell patched at various slice depths between 30-90 μm. Cells were voltage-clamped at a holding potential of -70 mV. After establishing stable access to a cell, we acquired its response to one-photon pulses of increasing duration (1, 3, and 5 ms pulses at 10 Hz). A volume of tissue surrounding that cell would then be probed using a dense grid of holograms across 3-4 powers. The total size of the grid was 160x160x100 μm, resulting in a 25x25x4 voxel grid. The interval between the centers of the hologram targets were 5 pixels (6.5 μm) in the x/y dimensions and 25 μm in the z dimension. The phase masks of the grid holograms were pre-computed and stored in memory before performing the measurement. Recording parameters were frequently checked and logged by pausing the mapping protocol and probing the cell with a series of hyperpolarizing pulses.

Deduplication of segmented neurons

In some cases, cell nuclei appeared in more than one plane when inspecting the two-photon image stacks collected at the beginning of the experiment. If the centroids of neural ROIs appeared in very close proximity to each other (up to 7 μm), they were considered duplicates. To prevent these duplicates from biasing estimates of the number of presynaptic candidate neurons mapped, we removed them from the analysis after the mapping experiment completed. To do so, we selected the brightest ROI from the set of duplicates to retain and discarded the others. We also discarded clearly erroneous ROIs, such as those lying on the glass filament. To remove duplicate connections, we applied a similar procedure that made use of the computational methods (see Algorithm 4).

Computational methods

Neural waveform demixing network architecture and training

The NWD network is a sequential U-Net [Ronneberger et al., 2015] that forms compressed representations of 45 ms single-trial PSC traces via a “contraction path,” and generates a PSC waveform at the original trial length but with confounding synaptic currents removed via an “expansion path.” Each block in the contraction path consists of a 2x temporal decimation step, a 1d convolution, a batch-norm step, and a rectified linear activation function. Each block in the expansion path consists of a “transposed” 1d convolution (also known as a fractionally strided convolution, or as a pseudo deconvolution), a batch-norm step, a rectified linear activation function, and a 2x linear interpolation step. There are four contraction blocks and four expansion blocks, with the outputs of the contraction blocks submitted to the expansion blocks at the corresponding temporal resolution through skip connections as in the original U-Net architecture.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The NWD networks used in this study were trained using 50,000 simulated PSC traces. The simulated PSCs obeyed the following generative structure. First, a template PSC $\varrho(\,\cdot\,;\tau_r,\tau_d,\Delta)$ with parameters τ_r,τ_d,Δ took the form

$$\varrho(t;\tau_r,\tau_d,\Delta)=\left[\exp\left(\frac{t-\Delta}{\tau_d}\right)-\exp\left(\frac{t-\Delta}{\tau_r}\right)\right]\mathbb{1}_{[t\geq\Delta]}. \tag{2}$$

Every simulated PSC $c(t)$ was the sum of a random number of templates, accounting for the stimulation of multiple connected presynaptic cells,

$$c(t)=\sum_{j=1}^{J}\varrho(t;\tau_r^j,\tau_d^j,\Delta^j), \tag{3}$$

$$\tau_X^j\sim\text{Uniform}(\tau_X^{\text{min}},\tau_X^{\text{max}}) \tag{4}$$

$$\Delta^j\sim\text{Uniform}(\Delta^{\text{min}},\Delta^{\text{max}}) \tag{5}$$

$$J\sim p(J) \tag{6}$$

where $X\in\{r,\text{diff}\}$ with $\tau_d=\tau_r+\tau_{\text{diff}}$ to ensure $\tau_d>\tau_r$, and where $p(J)$ determined the probability of selecting $J\in\mathbb{N}$ (including the possibility of $J=0$). The PSC c_k for training example k was then given as a function of time t by

$$c_k(t)=c_k^{\text{prev}}(t)+c_k^{\text{tar}}(t)+c_k^{\text{next}}(t)+g_k(t)+\epsilon_k(t) \tag{7}$$

where c_k^{prev}, c_k^{tar} and c_k^{next} refer to PSC effects from the previous trial, the target trial, and the next trial(s). Note that c_k^{prev} is not the same as c_{k-1}^{tar} as each trial was simulated entirely independently. Here $g_k(t)$ represents temporally-correlated noise from a Gaussian process,

$$g_k\sim\text{Normal}(\mathbf{0},\mathbf{K}), \tag{8}$$

with the covariance matrix $\mathbf{K}$ defined by the radial basis function kernel with noise variance σ_{scale} and characteristic lengthscale ℓ_{gp}

$$\mathbf{K}_{t_1,t_2}=\sigma_{\text{scale}}\exp\left(\frac{-(t_1-t_2)^2}{2\ell_{\text{gp}}^2}\right). \tag{9}$$

Finally, $\epsilon_k(t)$ represents uncorrelated noise sampled from a univariate Gaussian,

$$\epsilon_k(t)\sim\text{Normal}(0,\sigma_{\text{noise}}^2) \tag{10}$$

where $\sigma_{\text{noise}}^2\sim U(\sigma_{\text{noise,min}}^2,\sigma_{\text{noise,max}}^2)$. The NWD network is trained to infer c_k^{tar} from c_k by minimizing the mean-squared error loss.

The time constants of the training data were selected to match the PSCs recorded in experiments via an interactive simulator that overlaid the template waveform on experimental traces. We found that the NWD network performed best when trained separately on training data matched to either pyramidal-to-pyramidal mapping experiments, or PV-to-pyramidal mapping experiments, depending on the data being processed.

The NWD network intrinsically performs a denoising of the evoked PSCs since each target PSC in the training data was uncorrupted by the noise terms g_k and ϵ_k. We found that we could obtain further improvement

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

in PSC denoising by randomly incorporating 45 ms snippets of pure electrical noise taken from experiments during periods without stimulation. For such "negative" templates, the network was trained to produce an output of zeros.

Finally, we performed a minor correction step to guarantee that the output of the NWD network monotonically decayed after a predefined time [Lee et al., 2020]. Namely, after a user-specified time t_{monotone}, the network output N_t for all $t > t_{\text{monotone}}$ was adjusted by recursively setting $N_t = \min\{N_t, N_{t-1}\}$.

We implemented NWD using PyTorch Lightning. Networks were trained using stochastic gradient descent with a batch size of 64 and learning rate of 0.01 for 3000 epochs, which was sufficiently long for the optimizer to converge in our training regime.

Statistical model

We investigated a number of different approaches to connectivity inference, enumerated in the Supplementary Methods. Here we present the novel CAVIaR algorithm that we found achieved state-of-the-art performance.

CAVIaR aims to fit a statistical model to the demixed optogenetic data that relates patterns of holographic stimulation to the resulting measurements of postsynaptic current. Let $\mathbf{c}_k \in \mathbb{R}^T$ represent the PSC trace on trial k. In accordance with the trial structure used for NWD, we assume that $T = 900$ frames (or 45 ms), with photostimulation beginning at $t = 100$ (or 5 ms). To make use of compressed sensing-style techniques, we collapse the demixed, denoised PSC trace into a single number by integrating over the duration of the trial,

$$y_k = \int_0^T c_k^{\text{NWD}}(t)dt, \quad k = 1, \ldots, K, \tag{11}$$

where c_k^{NWD} represents the demixed trace on trial k. The measurements y_k therefore represent the total synaptic charge transfer resulting from the transmission of presynaptic spikes. Since the postsynaptic neuron is held under voltage clamp, we may model these measurements as a simple sum of optogenetically-evoked presynaptic spikes weighted by the strengths of the corresponding synapses,

$$y_k \sim \text{Normal}\left(\mathbf{w}^\top \mathbf{s}_{:,k}, \sigma^2\right), \qquad \sigma^{-2} \sim \text{Gamma}\left(t_{\text{sh}}, t_{\text{ra}}\right). \tag{12}$$

Here $\mathbf{w}$ is a vector of synaptic weights encoding the charge transfer resulting from a single spike, and $\mathbf{s}_{:,k}$ is the vector of presynaptic spikes on trial k. Note that, as described below, spike generation is highly stochastic, and hence s_{nk} is a latent variable. The synaptic weights are regularized by imposing a Gaussian prior,

$$w_n \sim \text{Normal}(u, b^2), \tag{13}$$

with hyperparameters u and b^2, and the variance σ^2 is intended to account for electrical noise from the recording electrode and variability in the amplitude of the evoked PSCs. Model parameters are listed in Table 1.

The variability of spike generation in our experimental data shows that the probability of generating a spike is not only different between cells, but also across power levels, and is therefore effectively impossible to know ahead of time. We propose a structured presynaptic model that can flexibly adjust to varying spike probabilities on a cell-by-cell basis. In particular, presynaptic spikes are assumed to be driven by holographic photostimulation according to a linear-nonlinear-Bernoulli model,

$$s_{nk} \mid \boldsymbol{\phi}_n \sim \text{Bernoulli}\left(f\left(\phi_n^0 I_{nk} - \phi_n^1\right)\right) \tag{14}$$

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Symbol	Interpretation	Hyperparameters	Posterior
$\mathbf{w}$	Connection strengths	u, b^2	$\boldsymbol{\mu}$, $\boldsymbol{\Omega}$
s_{nk}	Presynaptic spike	–	λ_{nk}
ϕ_n^0, ϕ_n^1	Presynaptic sigmoid coefficients	$\mathbf{v}$, $\mathbf{L}$	$\boldsymbol{\nu}_n$, $\boldsymbol{\Sigma}_n$
σ^2	Observation noise variance	t_{sh}, t_{ra}	θ_{sh}, θ_{ra}
z_k	Spontaneous synaptic current		
y_k	Observed response (integral of the PSC)		
I_{nk}	Laser power applied to neuron n on trial k		
N (n)	Number of neurons (neuron index)		
K (k)	Number of trials (trial index)		
f	Sigmoid function		

Table 1: Table of model parameters and their interpretations.

where $f(x) = 1/(1 + \exp(-x))$ is the sigmoid nonlinearity, I_{nk} denotes the power of the stimulation laser focused on cell n on trial k, and the sigmoid coefficients follow bivariate normal distributions

$$\boldsymbol{\phi}_n \sim \text{Normal}(\mathbf{v}_n, \mathbf{L}_n). \tag{15}$$

During inference we thus effectively perform a Bayesian logistic regression of each cell's inferred presynaptic spikes on the laser targets. This presynaptic model allows spike probabilities to adapt to variation in opsin expression via the sigmoid coefficient term ϕ_n^0, as well as to variation in rheobase via the intercept term ϕ_n^1.

Posterior inference

Variational approximation

The goal of CAVIaR is to infer the posterior distribution over the hidden variables, including the synaptic weights and presynaptic spikes. The posterior describes the distribution of parameter values that are consistent with the observed data, given our prior assumptions. Importantly, this provides a description of uncertainty in the inferred synaptic weights. It is not computationally feasible to compute the exact posterior distribution, but standard methods have been developed to approximate the posterior [Blei et al., 2017]. Here we modify the conventional coordinate-ascent variational inference algorithm by augmenting it with several steps that account for spontaneous synaptic currents and biophysical plausibility of the inferred spikes. To this end, we note that the posterior distribution over the latent variables factorizes as

$$p(\mathbf{w}, \mathbf{s}, \boldsymbol{\phi}, \sigma^2 \mid \mathbf{y}, \mathcal{I}) \propto p(\sigma^2) \prod_{n=1}^{N} p(w_n) p(\boldsymbol{\phi}_n) \prod_{k=1}^{K} p(s_{nk} \mid I_{nk}, \boldsymbol{\phi}_n) p(y_k \mid \mathbf{w}, \mathbf{s}_{:,k}, \sigma^2). \tag{16}$$

We approximate the posterior distribution $p(\mathcal{Z} \mid \mathbf{y}, \mathcal{I})$ by a variational model $q(\mathcal{Z})$ (where $\mathcal{Z}$ represents the set of latent variables) that obeys a similar factorization,

$$q(\mathbf{w}, \mathbf{s}, \boldsymbol{\phi}, \sigma^2) = q(\sigma^2 \mid \theta_{\text{sh}}, \theta_{\text{ra}}) q(\mathbf{w} \mid \boldsymbol{\mu}, \boldsymbol{\Omega}) \prod_{n=1}^{N} q(\boldsymbol{\phi}_n \mid \boldsymbol{\nu}_n, \boldsymbol{\Sigma}_n) \prod_{k=1}^{K} q(s_{nk} \mid \lambda_{nk}) \tag{17}$$

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

where the individual factors are $q(\mathbf{w} \mid \boldsymbol{\mu}, \boldsymbol{\Omega}) = \text{Normal}(\mathbf{w} \mid \boldsymbol{\mu}, \boldsymbol{\Omega})$, $q(s_{nk} \mid \lambda_{nk}) = \text{Bernoulli}(s_{nk} \mid \lambda_{nk})$, $q(\boldsymbol{\phi}_n \mid \boldsymbol{\nu}_n, \boldsymbol{\Lambda}_n) = \text{Normal}(\boldsymbol{\phi}_n \mid \boldsymbol{\nu}_n, \boldsymbol{\Sigma}_n)$, and $q(\sigma^{-2} \mid \theta_{\text{sh}}, \theta_{\text{ra}}) = \text{Gamma}(\sigma^{-2} \mid \theta_{\text{sh}}, \theta_{\text{ra}})$.

Given the above approximation, coordinate-ascent variational inference seeks to perform an update of each factor $q(\mathcal{Z}_i)$ one-by-one for all $\mathcal{Z}_i \in \mathcal{Z}$, with the idea being that each update moves the approximate posterior $q(\mathcal{Z})$ closer to the true posterior $p(\mathcal{Z} \mid \mathbf{y}, \mathcal{I})$ in the sense of the Kullback-Leibler divergence (KL-divergence) [Blei et al., 2017]. We take a variety of approaches to updating each factor depending on the relative tractability of the update. The complete algorithm is given in Algorithm 1.

Inference of synaptic weights

First, it can be shown that the optimal variational update for the synaptic strengths $\mathbf{w}$, conditional on all other latent variables $\mathcal{Z} \setminus \mathbf{w}$, obeys the equation

$$q(\mathbf{w} \mid \boldsymbol{\mu}, \boldsymbol{\Omega}) \propto \exp \mathbb{E}_{q(\mathcal{Z} \setminus \mathbf{w})} \left[\ln p(\mathbf{y}, \mathcal{Z} \mid \mathcal{I}) \right], \tag{18}$$

see, e.g., [Bishop, 2006]. Since both the prior on $\mathbf{w}$ and the observations are Gaussian-distributed, one can evaluate Equation 18 analytically and solve for $\boldsymbol{\mu}$ and $\boldsymbol{\Omega}$ by "completing the square", yielding the block update

$$\boldsymbol{\Omega} = \left(\frac{\theta_{\text{sh}}}{\theta_{\text{ra}}} \sum_{k=1}^{K} (\mathbf{D}_k + \mathbf{M}_k) + \frac{1}{b^2} \mathbf{I}_n \right)^{-1}, \quad \boldsymbol{\mu} = \boldsymbol{\Omega} \left(\frac{\theta_{\text{sh}}}{\theta_{\text{ra}}} \sum_{k=1}^{K} y_k \boldsymbol{\lambda}_{:,k} + \frac{1}{b^2} u \right) \tag{19}$$

where $\mathbf{D}_k = \text{diag}\left(\lambda_{1,k}(1-\lambda_{1,k}), \ldots, \lambda_{Nk}(1-\lambda_{Nk})\right) \in \mathbb{R}^{N \times N}$ is diagonal and $\mathbf{M}_k = \boldsymbol{\lambda}_{:,k} \boldsymbol{\lambda}_{:,k}^{\top} \in \mathbb{R}^{N \times N}$. Inspecting the update for $\boldsymbol{\mu}$, one sees that a neuron's synaptic weight is only determined by the magnitude of the postsynaptic responses on trials for which that neuron actually elicited a spike. However, those spikes are not directly observed, and must themselves be inferred from the data.

Inference of presynaptic spikes

Given a measurement y_k resulting from the photostimulus $I_{:,k}$ and our current estimates of synaptic strengths, electrical noise, and *a priori* presynaptic spike probabilities, we must decide which of the targeted cells actually elicited a spike. In early trials, this depends more on the prior (i.e., the presynaptic model) rather than on the likelihood, and this balance shifts in favor of the likelihood the more experimental data that is collected. However, unlike in Equation 18, the optimal variational update for the presynaptic spike probabilities λ_{nk} cannot be solved for analytically. Instead, one can recognize ([Bishop, 2006]) that the KL-divergence from the approximate to the true posterior, $\text{KL}(q \| p)$, can be decomposed into the sum

$$\text{KL}(q \| p) = \ln p(\mathbf{y}) - \mathcal{L}(\mathcal{Z}) \tag{20}$$

where $\ln p(\mathbf{y}) = \ln \int p(\mathbf{y}, \mathcal{Z}) d\mathcal{Z}$ is the logarithmic model evidence and where

$$\mathcal{L}(\mathcal{Z}) = \mathbb{E}_q \left[\ln p(\mathbf{y}, \mathbf{w}, \mathbf{s}, \boldsymbol{\phi}, \sigma^2) \right] + \mathcal{H}_q \left[\mathbf{w}, \mathbf{s}, \boldsymbol{\phi}, \sigma^2 \right], \tag{21}$$

with $\mathcal{H}_q[x]$ the Shannon entropy of variable x under density q. Since the KL-divergence is always non-negative, the inequality $\mathcal{L}(\mathcal{Z}) \leq \ln p(\mathbf{y})$ always holds, and therefore $\mathcal{L}(\mathcal{Z})$ is a lower bound on the logarithmic model evidence (referred to as the "evidence lower bound"). By maximizing the evidence lower bound as a function

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

of the variational model parameters, $\mathcal{L}(\mathcal{Z})$ approaches $\ln p(\mathbf{y})$, and hence drives KL$(q\|p)$ towards zero. If the bound is tight, i.e., if $\mathcal{L}(\mathcal{Z}) = \ln p(\mathbf{y})$, the KL-divergence is zero and the variational approximation to the posterior becomes exact. To this end, differentiating $\mathcal{L}$ and solving for λ_{nk} yields the update

$$\ln\left(\frac{\lambda_{nk}}{1-\lambda_{nk}}\right) = \mathbb{E}_{q(\boldsymbol{\phi})}\left[\ln\left(\frac{f_{nk}}{1-f_{nk}}\right)\right] - \frac{\theta_{\text{sh}}}{2\theta_{\text{ra}}}\left\{-2y_k\mu_n + 2\mu_n\sum_{j\neq n}\mu_j\lambda_{jk} - \left(\mu_n^2+\beta_n^2\right)\right\}$$

where we have defined $f_{nk} = f\left(\phi_n^0 I_{nk} - \phi_n^1\right)$. The expectation in the above term is analytically intractable due to multiple nonlinearities. We thus make a Monte Carlo approximation to the expectation

$$\mathbb{E}_{q(\boldsymbol{\phi})}\left[\ln\left(\frac{f_{nk}}{1-f_{nk}}\right)\right] \approx \frac{1}{M}\sum_{m=1}^{M}\ln\left(\frac{f_{nk}[m]}{1-f_{nk}[m]}\right) \tag{22}$$

with $f_{nk}[m] = f\left(\phi_n^0[m]I_{nk} - \phi_n^1[m]\right)$, and for each pair $(\phi_n^0[m], \phi_n^1[m])^\top \sim q(\boldsymbol{\phi}_n \mid \boldsymbol{\nu}_n, \boldsymbol{\Lambda}_n)$. As discussed below, the sigmoid coefficients are sampled from a *truncated* multivariate normal distribution. However, sampling from truncated multivariate normal distributions is computationally demanding and remains an open area of research [Pakman and Paninski, 2014]. Instead, the terms ϕ_n^0 and ϕ_n^1 are sampled independently from their truncated marginal distributions.

Inspecting the updates for λ_{nk}, one notices that each update depends on all other neurons in the population, but is independent of all other trials. Thus, we update the spike probabilities for all trials simultaneously, but perform this one neuron at a time (where the order of the neuron updates is randomized every iteration). Immediately following the inference of a neuron's spikes, we evaluate their biophysical plausibility as a function of laser power. It is well known from patch-clamp studies with opsin-expressing neurons that the probability of photoeliciting an action potential should increase monotonically with laser power [Sridharan et al., 2022]. We therefore perform an isotonic regression through the inferred spike rate using the pool adjacent violators algorithm (PAVA) [Ayer et al., 1955].

Concretely, let $K_n^p = \{k \in \{1,\ldots,K\} : I_{nk} = p\}$ be the set of all trials on which neuron n was stimulated with power p, and let

$$\bar{\lambda}_n^p = \frac{1}{|K_n^p|}\sum_{k\in K_n^p}\lambda_{nk} \tag{23}$$

denote the average inferred spike probability for neuron n at power p. PAVA computes an isotonic regression function $\hat{F}_n$ by solving the optimization problem

$$\text{minimize}\sum_{p\in\mathcal{P}}(\bar{\lambda}_n^p - \hat{F}_n(p))^2 \;\text{ such that } \hat{F}_n \text{ is non-decreasing,} \tag{24}$$

where $\mathcal{P}$ is the set of non-zero powers used in the experiment. This provides us with an isotonically-constrained optogenetic "power curve" $\hat{F}_n$ that we use to judge whether or not the putative synapse connecting neuron n to the patched neuron is sufficiently plausible. In particular, we require that at the maximal power used in the experiment the inferred probability of successfully initiating and transmitting a spike is greater than some threshold $\theta_{\text{PAVA}} \in [0,1]$, which we refer to as the "min spike rate at max power" (MSRMP). However, some fraction of the inferred spikes may result from spontaneous synaptic currents, which can lead to false positives if left unaccounted for. We therefore infer the rate of spontaneous currents (λ_{spont}, described below) and add this

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

to the user-defined threshold θ_{PAVA} to adaptively adjust the plausibility criterion. Neurons with $\hat{F}_n(\max_k I_{nk}) < \theta_{\text{PAVA}} + \lambda_{\text{spont}}$ (i.e., whose adjusted maximal spike rate is below θ_{PAVA}) are considered unconnected and have their synaptic weights and inferred spike probabilities set to zero.

We also use a "masking" procedure to prevent small, noise-driven postsynaptic measurements from causing spurious spike inferences. To do this, we evaluate a test statistic τ on each trace c_k. If $\tau(c_k) < \tau_{\text{min}}$ we force $\lambda_{nk} = 0$ for all $n = 1, \ldots, N$. The statistic τ is typically chosen to be the sample autocorrelation.

Inference of presynaptic sigmoid coefficients

The presynaptic sigmoid coefficients $\boldsymbol{\phi}$ (governing the presynaptic models) are updated using Laplace's method. Namely, rather than solving for the optimal variational approximation or optimizing the evidence lower bound, we learn a posterior distribution over each $\boldsymbol{\phi}_n$ by making a second-order Taylor approximation about the posterior mode, with the posterior covariance matrix given by the inverse of the Hessian matrix appearing in the Taylor series. In some cases we encountered pathological spike inference when the sigmoid coefficients $\boldsymbol{\phi}_n$ became negative, and hence we enforce a positivity constraint in the mode-finding algorithm described below. Moreover, the Monte Carlo approximation described in Equation 22 depends closely on the non-negativity of the sampled variates. Thus we truncate the Laplace approximation to the non-negative real line,

$$q(\boldsymbol{\phi}_n \mid \boldsymbol{\nu}_n, \boldsymbol{\Sigma}_n) = \text{TruncNorm}(\boldsymbol{\phi}_n \mid \boldsymbol{\nu}_n, \boldsymbol{\Sigma}_n) \tag{25}$$

with $\nu_n^i > 0$ for $i \in \{0, 1\}$ and supp $q(\boldsymbol{\phi}_n \mid \boldsymbol{\nu}_n, \boldsymbol{\Sigma}_n) = \mathbb{R}^2_{>0}$.

Note that, while PAVA is a non-parametric estimator of the spike rate, its use differs from the presynaptic spike model which takes a parametric form. The presynaptic model is used to evaluate the prior probability of photoeliciting a spike, whereas PAVA is used to evaluate the plausibility of the inferred spikes.

We use Newton's method with a log barrier and backtracking line search to identify the posterior mode $\boldsymbol{\nu}_n$. To this end, define the objective function with barrier penalty t as

$$\Psi_n^t = -\,\mathbb{E}_{q(\mathbf{s}_n)}\left[\sum_{k=1}^{K} \ln p(s_{nk} \mid \boldsymbol{\phi}_n, I_{nk}) + \ln p(\boldsymbol{\phi}_n \mid \mathbf{v}, \mathbf{L})\right] - \frac{1}{t}\sum_{i=0}^{1} \ln(\phi_n^i),$$

where we average over the spike uncertainty. Note that the presynaptic models depend only on the spikes $\mathbf{s}$, and are independent conditional on $\mathbf{s}$, allowing us to compute Laplace approximations in parallel across cells.

The mode-finding algorithm starts with an initial t. We then solve for the mode by iterating the Newton steps

$$\boldsymbol{\nu}_n^{(t)} \leftarrow \boldsymbol{\nu}_n^{(t)} - \kappa \mathbf{H}_{nt}^{-1} \mathbf{J}_{nt} \tag{26}$$

until convergence, where $\mathbf{J}$ and $\mathbf{H}$ are respectively the Jacobian and Hessian of the presynaptic model for neuron n with barrier sharpness t

$$\mathbf{J}_{nt} = \nabla_{\boldsymbol{\phi}_n} \Psi_n^t, \quad \mathbf{H}_{nt} = \nabla\nabla_{\boldsymbol{\phi}_n} \Psi_n^t. \tag{27}$$

The stepsize κ is adaptively selected using a standard backtracking line search rule. We then increase t to sharpen the log barrier and repeat as required. Typically we only sharpen t two or three times as this proved sufficient for our data.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Inference of observation noise variance

To infer the observation noise variance, we use the same approach as with the update rule for the synaptic weights (Equation 18). In particular, noting that

$$q(\sigma^{-2} \mid \theta_{\text{sh}}, \theta_{\text{ra}}) \propto \exp \mathbb{E}_{q(\mathcal{Z} \setminus \sigma^{-2})} \left[\ln p(\mathbf{y}, \mathcal{Z} \mid \mathcal{I}) \right],$$

and recognizing that the gamma prior is conjugate to the Gaussian likelihood, one obtains the variational update

$$\theta_{\text{sh}} = t_{\text{sh}} + \frac{K}{2}, \quad \theta_{\text{ra}} = t_{\text{ra}} + \frac{1}{2} \sum_{k=1}^{K} \mathbb{E}_q \left[\left(y_k - \mathbf{w}^\top \mathbf{s}_{:,k} \right)^2 \right]$$

where the expectation above can be evaluated analytically, though we leave it in the above form for legibility.

Inference of spontaneous synaptic currents

Finally, we estimate spontaneous synaptic currents z_k using soft-thresholding. The idea is that, given a numerical tolerance ϵ (where typically $\epsilon = 0.05$) and positively rectified residuals

$$e_k := [y_k - \boldsymbol{\mu}^\top \boldsymbol{\lambda}_{:,k}]_+ \tag{28}$$

we apply a soft-thresholding function S with penalty γ defined by

$$S(e, \gamma) := \begin{cases} 0 & \text{if } e \leq \gamma \\ e - \gamma & \text{otherwise.} \end{cases} \tag{29}$$

We then iteratively shrink the penalty γ until the norm of the residual data comprises no more than ϵ of the norm of the observed data; i.e., until

$$\frac{\sum_{k=1}^{K} (y_k - \boldsymbol{\mu}^\top \boldsymbol{\Lambda}_{:,k} - z_k)^2}{\sum_{k=1}^{K} y_k^2} \leq \epsilon \tag{30}$$

where z_k is obtained via the soft-thresholded residuals. However, we also require that $(z_1, \ldots, z_K)$ be approximately orthogonal to $\boldsymbol{\lambda}_n$ for all n, and apply the masking procedure as noted in the spike inference step above, leading to the spontaneous synaptic current estimator

$$z_k = \begin{cases} 0 & \text{if } \sum_{n=1}^{N} \lambda_{nk} > \theta_{\text{orthog}} \text{ or } \tau(c_k) < \tau_{\text{min}} \\ S(e_k, \gamma) & \text{otherwise} \end{cases} \tag{31}$$

where $\theta_{\text{orthog}} \approx 0$, but is not exactly zero to allow for numerical imprecision. The spontaneous events are constrained to be approximately orthogonal to the inferred spikes to prevent variability in PSC amplitude causing false positives.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Algorithm 1: Coordinate-ascent variational inference and isotonic regularization (CAVIaR)

input: PSC traces $\mathbf{c}$, stimulus information $\mathcal{I}$, PAVA threshold θ_{PAVA}, spontaneous penalty backtracking scalar α, soft orthogonality threshold θ_{orthog}, minimal test statistic τ_{min}, number of iterations *iters*

initialise $\lambda_{nk} \leftarrow 1$ for all n, k such that $I_{nk} > 0$ and $\tau_{\text{test}}(\mathbf{c}_k) \geq \tau_{\text{min}}$
$\lambda_{\text{spont}} \leftarrow 0$ // initialize spontaneous rate to 0
$i \leftarrow 1$
while $i \leq iters$ **do**
  update $q(\mathbf{w} \mid \boldsymbol{\mu}, \boldsymbol{\Omega}) \propto \exp \mathbb{E}_{q(\mathcal{Z} \setminus \mathbf{w})} [\ln p(\mathbf{y}, \mathcal{Z} \mid \mathcal{I})]$ // variational solution for synaptic weights
  for $n = 1, \ldots, N$ **do** // infer spikes via Monte Carlo ELBO solution
    sample $\boldsymbol{\phi}_n[m] \sim q(\boldsymbol{\phi}_n \mid \boldsymbol{\nu}, \boldsymbol{\Sigma})$ for $m = 1, \ldots, M$
    for $k = 1, \ldots, K$ **do**
      $\lambda_{nk} \leftarrow \operatorname{argmax}_{\lambda_{nk}} \textsc{Elbo}(\lambda_{nk} \mid \{\boldsymbol{\phi}_n[m]\}_{m=1}^{M})$
    end
    $\hat{F}_n \leftarrow \text{PAVA}(\mathcal{I}, \boldsymbol{\lambda}_n)$ // estimate optogenetic power curve
    if $\hat{F}_n(\max_k I_{nk}) < \theta_{PAVA} + \lambda_{spont}$ **then** // check plausibility criterion
      $\mu_n \leftarrow 0$, $\boldsymbol{\lambda}_n \leftarrow \mathbf{0}$
    end
  end
  for $n = 1, \ldots, N$ **do** // Laplace approx of receptive field posterior
    $\boldsymbol{\nu}_n, \boldsymbol{\Sigma}_n \leftarrow \textsc{ReceptiveFieldLaplace}(\mathbf{s}_n, \boldsymbol{\phi}_n, \mathbf{v}, \mathbf{L})$
  end
  update $q(\sigma^{-2} \mid \theta_{\text{sh}}, \theta_{\text{ra}}) \propto \exp \mathbb{E}_{q(\mathcal{Z} \setminus \sigma^{-2})} [\ln p(\mathbf{y}, \mathcal{Z} \mid \mathcal{I})]$ // variational solution for noise precision
  while $\|\mathbf{y} - \boldsymbol{\mu}^\top \boldsymbol{\Lambda} - \mathbf{z}\|_2^2 / \|\mathbf{y}\|_2^2 \geq \epsilon$ **do** // begin spontaneous PSC inference
    for $k = 1, \ldots, K$ **do**
      if $\sum_{n=1}^{N} \lambda_{nk} \leq \theta_{orthog}$ *and* $\tau_{test}(\mathbf{c}_k) \geq \tau_{min}$ **then**
        $z_k \leftarrow S([y_k - \boldsymbol{\mu}^\top \boldsymbol{\lambda}_{:,k}]_+, \gamma)$ // soft-threshold residual with penalty γ
      end
    end
    $\gamma \leftarrow \alpha\gamma$ // shrink penalty
  end
  $\lambda_{\text{spont}} \leftarrow \frac{1}{K} \sum_{k=1}^{K} \mathbb{1}_{[z_k \neq 0]}$ // update spontaneous rate
  $i \leftarrow i + 1$
end
$\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\Lambda}, \boldsymbol{\phi}, \mathbf{z} \leftarrow \textsc{FNScan}(\boldsymbol{\mu}, \mathbf{z})$ // scan resulting spontaneous PSCs for potential false negatives

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Algorithm 2: RECEPTIVEFIELDLAPLACE

for $n = 1, \ldots, N$ **do**
  for $t = 1, \ldots, T_{max}$ **do**
    $\kappa \leftarrow 1$ // Newton stepsize
    $\Psi_n(\boldsymbol{\phi}_n) = -\mathbb{E}_{q(s_n|\boldsymbol{\lambda}_n)}\left[\sum_{k=1}^{K} \ln p(\mathbf{s}_{nk} \mid \boldsymbol{\phi}_n, I_{nk}) + \ln p(\boldsymbol{\phi}_n \mid \mathbf{v}, \mathbf{L})\right] - \frac{1}{\alpha_{\text{barrier}}} \sum_{i=0}^{1} \ln(\phi_n^i)$
    $\mathbf{J} = \nabla_{\boldsymbol{\phi}_n} \Psi_n$, $\mathbf{H} = \nabla\nabla_{\boldsymbol{\phi}_n} \Psi_n$
    $\mathbf{d} = \mathbf{H}_n^{-1} \mathbf{J}_n$ // search direction
    while $\Psi_n(\boldsymbol{\phi}_n + \kappa \mathbf{d}) > \Psi_n(\boldsymbol{\phi}_n) + \alpha_{backtrack}\, \kappa\, \mathbf{J}_n^\top \mathbf{d}$ **do** // backtrack
      $\kappa \leftarrow \beta_{\text{backtrack}}\, \kappa$
    end
    $\boldsymbol{\phi}_n \leftarrow \boldsymbol{\phi}_n - \kappa \mathbf{d}$ // make step
  end
  $\boldsymbol{\nu}_n \leftarrow \boldsymbol{\phi}_n$
  $\boldsymbol{\Sigma}_n \leftarrow \mathbf{H}_n^{-1}$
end
$q(\boldsymbol{\phi} \mid \boldsymbol{\nu}_n, \boldsymbol{\Sigma}_n) = \text{TruncNormal}(\boldsymbol{\nu}_n, \boldsymbol{\Sigma}, 0, \infty)$ // truncate support to $(0, \infty)$

Algorithm 3: FNSCAN

input: Synaptic weights $\boldsymbol{\mu}$, spontaneous synaptic currents $\mathbf{z}$, stimulus information $\mathcal{I}$, PAVA threshold θ_{PAVA}

$S_{\text{disc}} \leftarrow \{n \in \{1, \ldots, N\} : \mu_n = 0\}$ // initialize pool of candidate neurons
while $|S_{disc}| > 0$ **do**
  for $n = 1, \ldots, N$ **do** // collect spontaneous PSC indices aligning with neuron stim times
    $\text{spont}_n \leftarrow \{k \in \{1, \ldots, K\} : z_k \neq 0 \text{ and } I_{nk} > 0\}$
  end
  $n^* \leftarrow \operatorname{argmax}_n |\text{spont}_n|$ // select neuron with most coincidental spontaneous PSCs
  $\hat{F}_{n^*}^{\text{spont}} \leftarrow \text{PAVA}\left((I_{n^*k})_{k \in \text{spont}_n}, (z_k)_{k \in \text{spont}_n}\right)$ // estimate putative power curve
  if $\hat{F}_{n^*}^{spont}(\max_k I_{n^*k}) \geq \theta_{PAVA}$ **then**
    $\mu_{n^*} \leftarrow \text{mean}(\{z_k : k \in \text{spont}_{n^*}\})$ // neuron passes PAVA criterion, declare connected
    $\beta_{n^*} \leftarrow \text{s.e.m.}(\{z_k : k \in \text{spont}_{n^*}\})$
    for $k \in spont_{n^*}$ **do**
      $\lambda_{n^*k} \leftarrow 1$ // declare spontaneous PSC to be spike from neuron n^*
      $z_k \leftarrow 0$ // remove spontaneous PSC from vector $\mathbf{z}$
    end
  end
  $S_{\text{disc}} \leftarrow S_{\text{disc}} \setminus \{n^*\}$ // remove n^* from pool of disconnected neurons
end

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Post-hoc scan for false negatives

Occasionally the CAVIaR algorithm will erroneously declare a neuron to be disconnected, either due to a failure to meet the PAVA-based plausibility criterion in the earliest iterations of the algorithm, or due to the biological data violating the assumptions of the underlying statistical model. To correct for this, following inference we perform a post-hoc scan for potential false negatives. The idea is to reconnect neurons that CAVIaR originally declared to be disconnected by checking whether spontaneous PSCs coinciding with the stimulation of a selected neuron could constitute valid postsynaptic responses when the rigidity of the statistical model is relaxed. Since the overwhelming bulk of connections are already identified by the first CAVIaR pass, missed connections are comparatively rare and thus we found a simple greedy algorithm to be effective.

The false-negative scanning algorithm is given in Algorithm 3. Briefly, the algorithm begins by collecting all neurons S_{disc} declared to be disconnected by the first CAVIaR pass. It then selects the neuron $n^* \in S_{\text{disc}}$ with the greatest number of coincidental spontaneous PSCs, and checks if assigning these PSCs to neuron n^* would satisfy the PAVA criterion. If so, neuron n^* is declared connected, with the posterior distribution of the reconnected neuron's parameters determined by sample statistics of the corresponding spontaneous PSCs. Neuron n^* is then removed from S_{disc} (whether reconnected or not), and the algorithm repeats until $S_{\text{disc}} = \emptyset$.

Inference of canonical postsynaptic current waveforms

Once we have inferred the presynaptic spike matrix $\mathbf{\Lambda}$, we can obtain accurate PSC waveforms $\mathbf{r}_n \in \mathbb{R}^T$ using ridge regression. Collecting the waveforms in the rows of a matrix $\mathbf{R}$, they can be obtained simultaneously by solving the non-negative L_2 problem

$$\hat{\mathbf{R}} = \underset{\mathbf{R} \geq \mathbf{0}}{\operatorname{argmin}} \left\{ \|\mathbf{C} - \mathbf{\Lambda}^\top \mathbf{R}\|_F + \gamma \|\mathbf{R}\|_F \right\} \tag{32}$$

where $\mathbf{C} \in \mathbb{R}^{K \times T}$ is the matrix of PSC traces $(\mathbf{c}_1, \ldots, \mathbf{c}_K)$, $\gamma > 0$ is the ridge penalty, and $\|\cdot\|_F$ is the Frobenius norm. Note that by using the spike matrix $\mathbf{\Lambda}$ instead of the optogenetic "design matrix" (where each element of the matrix determines which neurons were merely stimulated as in [Hu and Chklovskii, 2009], rather than which spiked as a result of photostimulation), the estimated waveforms are much less biased by trials in which neurons were not photoactivated.

Removal of duplicated connections

We use the ridge-inferred canonical PSC waveforms to merge putative duplicate connections (Algorithm 4). To do this, we partition the total set of neurons into groups whose constituent members are located close-by in space and have similar PSC waveforms. We then use a two-photon image of the opsin expression to select the brightest member from each group as a representative.

Leave-one-hologram-out cross-validation

We assess the accuracy of the CAVIaR inferences using leave-one-hologram-out cross-validation (LOHO-CV, Algorithm 5). Let $\mathcal{H}$ represent the complete set of hologram designs; i.e., $h \in \mathcal{H}$ determines which neurons will be targeted for stimulation, regardless of laser power. LOHO-CV proceeds by selecting $h \in \mathcal{H}$, fitting

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Algorithm 4: REMOVEDUPLICATES

input: Postsynaptic current traces, matrix of inferred spikes, ridge penalty γ, waveform error tolerance ϵ_ω, distance error tolerances $\epsilon_{d_{xy}}$, ϵ_{d_z}, Cartesian coordinates of each neuron in 3d-space, two-photon image of opsin expression

1 estimate canonical waveforms $\{\boldsymbol{\omega}_n\}_{n=1}^N$ using ridge regression with penalty γ, PSC traces, and inferred spikes

2 partition neurons into groups $\mathcal{G}_i$ such that for each $n, m \in \mathcal{G}_i$ one has

$$\|\boldsymbol{\omega}_n - \boldsymbol{\omega}_m\|^2 \leq \epsilon_\omega, \|(x_n, y_n) - (x_m, y_m)\|^2 \leq \epsilon_{d_{xy}}, \text{ and } (z_n - z_m)^2 \leq \epsilon_{d_z} \tag{33}$$

3 select group representative $g_i^* \in \mathcal{G}_i$ as the neuron with the highest pixel brightness in the two-photon image of opsin distribution

CAVIaR to the PSC traces and stimuli corresponding to all other holograms $\mathcal{H} \setminus \{h\}$, and then sampling from the posterior predictive distribution to obtain an estimate of the postsynaptic response at multiple power levels. Note that since we use multiple samples from the posterior predictive, we obtain a distribution of possible responses, which provides error bars for the predictions in Figure 5. Similarly, each ten-target hologram from the experimental data is repeated ten times, providing error bars for the observations in Figure 5.

Simulated circuit mapping experiments

We used simulated data to characterize the performance of the techniques we tested. To ensure the accuracy of this characterization, rather than sampling directly from the generative model that we propose, we added several layers of biophysical realism. In particular, we sampled noisy synaptic currents that were first demixed using NWD before being supplied as input to the connectivity inference algorithms. This way we could test for the combined accuracy of NWD and connectivity inference.

The simulated data was generated as follows. First, sigmoid parameters ϕ_n^X were sampled uniformly from $U(\phi_{\text{min}}^X, \phi_{\text{max}}^X)$ for $X \in \{0, 1\}$. Then, presynaptic spikes s_{nk} were sampled from a linear-nonlinear-Bernoulli model

$$s_{nk} \sim \text{Bernoulli}\left(\sigma\left(\phi_n^0 I_{nk} - \phi_n^1\right)\right) \tag{34}$$

where the laser power on each trial was randomly selected from a discrete set matched to the experimental data. Each neuron had a canonical PSC transient $h(t, \Delta_{nk}; \tau_r^n, \tau_d^n)$ that took a form similar to that used for training the NWD networks. The unnormalized transient $\tilde{h}_n$ was defined by

$$\tilde{h}(t, \Delta_{nk}; \tau_r^n, \tau_d^n) = \left[\exp\left(-\frac{t - \Delta_{nk}}{\tau_d^n}\right) - \exp\left(-\frac{t - \Delta_{nk}}{\tau_r^n}\right)\right] \mathbb{1}_{[t \geq \Delta_{nk}]} \tag{35}$$

where τ_r^n and τ_d^n are rise and decay time constants and Δ_{nk} represents the spike delay (i.e., the combined spike photoelicitation and transmission latencies) for neuron n on trial k. The transient was then normalized to take integral 1,

$$h(t, \Delta_{nk}; \tau_r^n, \tau_d^n) = \tilde{h}(t, \Delta_{nk}; \tau_r^n, \tau_d^n) / \int \tilde{h}_n(t', \Delta_{nk}; \tau_r^n, \tau_d^n) dt', \tag{36}$$

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Algorithm 5: LOHO-CV

input: PSCs $\mathbf{c}$, stimulus information $\mathcal{I}$, holograms $\mathcal{H}$, number of posterior samples J

for $h \in \mathcal{H}$ **do**
  estimate variational parameters $\boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\nu}, \boldsymbol{\Sigma}$ from $\{(\mathbf{c}_k, I_{:,k}) : I_{:,k}$ is not an instance of $h\}$ with CAVIaR
  for $j = 1, \ldots, J$ **do**
    for $n = 1, \ldots, N$ **do** // begin ancestral sampling from posterior predictive distribution
      sample $(\phi_n^0, \phi_n^1) \sim$ TruncNorm$(\boldsymbol{\nu}_n, \boldsymbol{\Sigma}_n)$
      sample $w_n \sim$ Normal$(\mu_n, \beta_n^2)\mathbb{1}_{[\mu_n \neq 0]} + \delta_0 \mathbb{1}_{[\mu_n = 0]}$
      for *each laser power* p **do**
        sample $s_n^p \sim$ Bernoulli$(\sigma(\phi_n^0 h_n^p - \phi_n^1))$, where h_n^p is the power delivered to neuron n assuming hologram h is applied with power p
      end
    end
    for *each laser power* p **do**
      set $y^p[j] \leftarrow \mathbf{w}^\top \mathbf{s}^p$
    end
  end
  predict response to hologram h at power p as $\frac{1}{J}\sum_{j=1}^{J} y^p[j]$
end

such that the synaptic weight for that neuron (w_n, defined as the synaptic charge transfer below) was preserved when multiplying the weight by the PSC.

The spike delays were laser power-dependent so that, in accordance with experimental data for the opsins used in this study [Sridharan et al., 2022], neurons initiated and propagated spikes faster if stimulated at higher powers. Concretely, for laser power I_{nk} targeted at neuron n on trial k, the spike time was sampled from a right-shifted gamma distribution,

$$\Delta_{nk} \sim \text{Gamma}\left(\frac{\alpha}{I_{nk}^2}, \beta, \Delta_{\text{min}}\right) \tag{37}$$

where

$$\text{Gamma}(x \mid a, b, t) = \frac{b^a}{\Gamma(a)}(x-t)^{a-1}\exp(-b(x-t))\mathbb{1}_{[x>t]}. \tag{38}$$

Note that since the mean of the shifted gamma distribution is $t + a/b$, the optogenetically evoked spikes followed, in expectation, an inverse-square dependence of time on power, consistent with the physics of two-photon absorption [Papagiakoumou et al., 2020]:

$$\mathbb{E}[\Delta_{nk}] = \Delta_{\text{min}} + \frac{\alpha}{\beta I_{nk}^2}. \tag{39}$$

The synaptic weights w_n are intended to model the total synaptic charge transfer resulting from the transmission of a presynaptic spike to the postsynaptic neuron. We sampled the weights in a way that reflects the typical observations from our experiments. Namely, of the neurons that were chosen to be synaptically connected, a small number of them were strongly connected (in our simulations this was 20% of the connected

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

population, though the precise value did not notably impact our results) and a large number of them were more weakly connected (80%). Specifically, for a given connectivity rate $\alpha \in (0, 1)$, a subset of $\lceil \alpha N \rceil$ neurons were randomly selected as being connected to the postsynaptic neuron. Then, if a neuron n was chosen to be strongly connected, its weights were sampled as

$$w_n \sim U(w_{\text{min}}^{\text{strong}}, w_{\text{max}}^{\text{strong}}) \tag{40}$$

where $w_{\text{min}}^{\text{strong}}$ and $w_{\text{max}}^{\text{strong}}$ respectively represent the lower and upper bounds of the uniform distribution. If n was weakly connected then its weights were sampled as

$$w_n \sim \text{Exp}(w_{\text{mean}}^{\text{weak}}, w_{\text{min}}^{\text{weak}}) \tag{41}$$

where the exponential distribution above is in its two-parameter, right-shifted form

$$\text{Exp}(x \mid \lambda, \Delta) = \frac{1}{\lambda} \exp\left(-\frac{1}{\lambda}(x - \Delta)\right) \mathbb{1}_{[x \geq \Delta]} \tag{42}$$

with $\Delta \geq 0$.

We used two approaches to generating PSC traces. Either we directly simulated individual 45 ms trials for each $k = 1, \ldots, K$, or we simulated continuous circuit mapping experiments at 20 kHz sampling resolution (matched to the experimental data) that lasted for tens of minutes. Using the latter approach we could very closely mimic the exact contribution of confounding synaptic currents arising from stimulation at very high frequencies. Simulations performed in this continuous manner were subsequently restructured into the usual 45 ms snippets of activity.

In the first "trial-wise" approach to simulating data, the presynaptic spikes, synaptic weights, and PSC kernels were used to generate the postsynaptic responses y_k as

$$y_k = \int \left(w_n m_{nk} s_{nk} h_n(t) + h_k^{\text{spont}}(t) + g_k(t) + \epsilon_k(t) \right) dt. \tag{43}$$

Here m_{nk} is a multiplicative noise term, $h_k^{\text{spont}}(t)$ is a spontaneous PSC term, $g_k(t)$ is a temporally correlated noise term, and $\epsilon_k(t)$ is an additive noise term.

The multiplicative noise term m_{nk} accounted for the fact that, in our experimental data, the precise amplitudes of PSC transients were observed to vary from trial to trial. We sampled m_{nk} from a log-normal distribution,

$$m_{nk} \sim \text{LogNormal}(0, \sigma_{\text{mult}}^2), \tag{44}$$

such that the median postsynaptic charge transfer following a presynaptic spike still took the value w_n. The spontaneous term h_k^{spont} was either a PSC as in Equation 36 with random time constants, amplitudes, and spike times, or the zero vector, depending on the probability of spontaneous events. Temporally correlated noise, $\mathbf{g}_k \in \mathbb{R}^T$, was sampled from a Gaussian process,

$$\mathbf{g}_k \sim \text{Normal}(\mathbf{0}, \mathbf{K}) \tag{45}$$

where $\mathbf{K}$ was defined by the radial basis function kernel,

$$\mathbf{K}_{t_1,t_2} = \sigma_{\text{scale}} \exp\left(\frac{-(t_1 - t_2)^2}{2\ell_{\text{gp}}^2}\right) \tag{46}$$

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

for $t_1, t_2 = 1, \ldots, T$. Finally, the additive noise $\epsilon_k(t)$ was sampled independently and identically from a zero-mean Gaussian, $\epsilon_k(t) \sim \text{Normal}(0, \sigma^2_{\text{noise}})$. Due to high computational tractability this trial-wise approach was used to generate the heatmaps in Figure 3.

The second "continuous experiment" approach to simulation was much more computationally demanding. Simulating a 30 minute experiment at 20 KHz requires 36,000,000 time points, making, for example, the use of Gaussian processes prohibitive. Our approach was to generate vectors encoding spike times for each neuron and convolve this with the corresponding neuron's PSC kernel while batching over blocks of time. To this end, an experiment of length T was evenly subdivided into trials according to the stimulation frequency f. For each such trial, we sampled spikes, spike latencies, and multiplicative noise for each neuron as described above. Then, using this collection of variables, we defined spike vectors $\boldsymbol{\zeta}_n \in \mathbb{R}^T$ for $n = 1, \ldots, N$ by setting

$$\zeta_n(\lceil t_k + \Delta_{nk} \rceil) = w_n m_{nk} s_{nk}, \quad \text{for each } k = 1, \ldots, K, \tag{47}$$

where t_k is the timebin at which the kth trial begins, and $\zeta_n(t) = 0$ at all other timebins. Similarly, we generated spontaneous spike vectors $\{\boldsymbol{\zeta}_j^{\text{spont}}\}_{j=1}^J$ at a specified rate λ_{spont} (in Hz), where each $\boldsymbol{\zeta}_j^{\text{spont}}$ encodes the activation of a single spontaneous PSC with waveforms $\mathbf{h}_j^{\text{spont}}$ defined by randomly sampled time constants, amplitudes, and onset times. The full-length postsynaptic measurement vector was then obtained by convolving the spike vectors with the PSC waveforms and summing across neurons and spontaneous inputs,

$$\mathbf{c} = \sum_{n=1}^{N} \mathbf{h}_n * \boldsymbol{\zeta}_n + \sum_{j=1}^{J} \mathbf{h}_j^{\text{spont}} * \boldsymbol{\zeta}_j^{\text{spont}} + \boldsymbol{\epsilon}. \tag{48}$$

Here the noise follows a first-order autoregressive process

$$\epsilon(t) \sim \text{Normal}(\gamma \epsilon(t-1), \sigma^2_{\text{noise}}) \tag{49}$$

with autoregressive coefficient $\gamma \in (0, 1)$, which was suitably scalable for continuous simulated experiments.

The postsynaptic responses y_k used by the connectivity inference algorithms were extracted from $\mathbf{c}$ by setting

$$y_k := \int_{t_k - t_{\text{pre}}}^{t_k + t_{\text{post}}} c^{\text{NWD}}(t) dt \tag{50}$$

where t_k represents the beginning of the k-th trial, and $t_{\text{pre}} = 100$ and $t_{\text{post}} = 800$. Note that as the rate of stimulation increases, the interstimulus-interval decreases, and the trial windows increasingly overlap in time. This leads to a confounding of the observed PSCs as in Figure 2, which the NWD network must suppress.

Combinatorics of expected interstimulus intervals

Let N represent the total number of potential presynaptic neurons, R the size of the ensemble, and f the rate of stimulation (in Hz). Assuming neurons are chosen uniformly at random, the probability of selecting neuron $n \in \{1, \ldots, N\}$ is

$$\frac{\binom{N-1}{R-1}}{\binom{N}{R}} = \frac{R}{N}. \tag{51}$$

Hence, one must stimulate N/R times on average to return to the same neuron. If stimulation occurs at f Hz, this takes N/Rf seconds.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary information: Rapid learning of neural circuitry from holographic ensemble stimulation enabled by Bayesian model-based compressed sensing

Symbol	Interpretation	Value
τ_r^{min}	Rise time constant minimum	10
τ_r^{max}	Rise time constant maximum	40
$\tau_{\text{diff}}^{\text{min}}$	Difference between rise and decay time constants minimum	150
$\tau_{\text{diff}}^{\text{max}}$	Difference between rise and decay time constants maximum	340
$\Delta_{\text{tar}}^{\text{min}}$	PSC onset time minimum (target trial)	160
$\Delta_{\text{tar}}^{\text{max}}$	PSC onset time maximum (target trial)	400
$\Delta_{\text{prev}}^{\text{min}}$	PSC onset time minimum (previous trial effects)	-400
$\Delta_{\text{prev}}^{\text{max}}$	PSC onset time maximum (previous trial effects)	159
$\Delta_{\text{next}}^{\text{min}}$	PSC onset time minimum (next trial effects)	400
$\Delta_{\text{next}}^{\text{max}}$	PSC onset time maximum (next trial effects)	899
σ_{scale}	GP variance	0.045
ℓ_{gp}	GP lengthscale	45
$\sigma_{\text{noise,min}}^2$	Noise variance minimum	0.001
$\sigma_{\text{noise,max}}^2$	Noise variance maximum	0.02

Table S1: Parameters used for training an NWD network for inhibitory-to-excitatory mapping experiments. Assumes timesteps measured at 20 kHz.

Symbol	Interpretation	Value
τ_r^{min}	Rise time constant minimum	10
τ_r^{max}	Rise time constant maximum	40
$\tau_{\text{diff}}^{\text{min}}$	Difference between rise and decay time constants minimum	60
$\tau_{\text{diff}}^{\text{max}}$	Difference between rise and decay time constants maximum	120
$\Delta_{\text{tar}}^{\text{min}}$	PSC onset time minimum (target trial)	160
$\Delta_{\text{tar}}^{\text{max}}$	PSC onset time maximum (target trial)	400
$\Delta_{\text{prev}}^{\text{min}}$	PSC onset time minimum (previous trial effects)	-400
$\Delta_{\text{prev}}^{\text{max}}$	PSC onset time maximum (previous trial effects)	159
$\Delta_{\text{next}}^{\text{min}}$	PSC onset time minimum (next trial effects)	400
$\Delta_{\text{next}}^{\text{max}}$	PSC onset time maximum (next trial effects)	899
σ_{scale}	GP variance	0.045
ℓ_{gp}	GP lengthscale	45
$\sigma_{\text{noise,min}}^2$	Noise variance minimum	0.001
$\sigma_{\text{noise,max}}^2$	Noise variance maximum	0.02

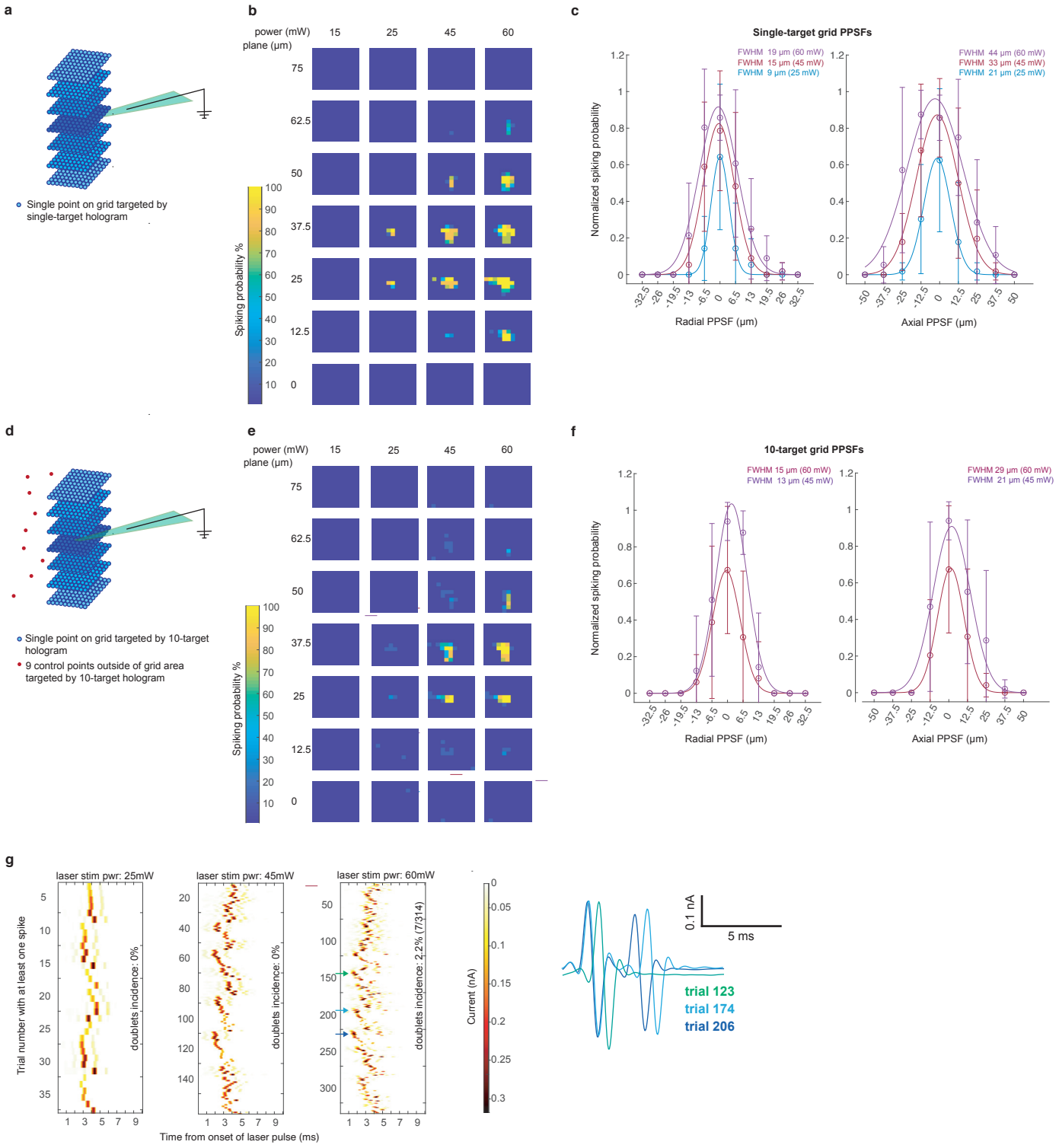
Table S2: Parameters used for training an NWD network for excitatory-to-excitatory mapping experiments. Assumes timesteps measured at 20 kHz.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Symbol	Interpretation	Default value
ϕ^0_{min}	Presynaptic sigmoid coefficient (0) minimum	0.2
ϕ^0_{max}	Presynaptic sigmoid coefficient (0) maximum	0.25
ϕ^1_{min}	Presynaptic sigmoid coefficient (1) minimum	10
ϕ^1_{max}	Presynaptic sigmoid coefficient (1) maximum	15
$\tau_{r,\text{min}}$	Rise time constant minimum	10
$\tau_{r,\text{max}}$	Rise time constant maximum	40
$\tau_{\Delta,\text{min}}$	Difference between rise and decay time constants minimum	250
$\tau_{\Delta,\text{max}}$	Difference between rise and decay time constants maximum	300
α	Rate parameter for gamma-distributed spike times	10^4
β	Shape parameter for gamma-distributed spike times	15
Δ_{min}	Minimum spike time	60
$w^{\text{strong}}_{\text{min}}$	Strong synaptic weight minimum	20
$w^{\text{strong}}_{\text{max}}$	Strong synaptic weight maximum	40
$w^{\text{weak}}_{\text{mean}}$	Weak synaptic weight mean (of the unshifted distribution)	4
$w^{\text{weak}}_{\text{min}}$	Weak synaptic weight minimum	5
ℓ_{gp}	GP lengthscale	50
σ_{scale}	GP noise variance	4^{-3}

Table S3: Table of default parameters for simulation studies.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure S1: Physiological point spread functions (PPSFs) for holographic single-target and ensemble stimulation. **a**, Illustration of multi-plane grid of stimulation sites used to obtain single-target PPSFs. A loose-patched, opsin-positive cell was positioned to be in the center plane of the grid (37.5 μm). Planes separated by 12.5 μm. **b**, Single-cell example of mean spike probability when stimulating individual points on grid (7 repetitions per power). **c**, Mean radial (left panel) and axial (right panel) PPSF for a population of opsin-expressing PV cells. Error bars indicate standard deviation over n=8 cells. **d**, Illustration of multi-plane grid of stimulation sites used to obtain 10-target ensemble PPSFs. In these experiments, one point on the grid is targeted by the hologram while the remaining 9 "control" targets are outside of the grid area and kept fixed (i.e., do not change from stimulus to stimulus). **e**, Single-cell example of mean spike probability when stimulating using a 10-target hologram (with one target being the opsin-positive cell; 7 repetitions per power). Panels (a) and (c) display results from the same cell. **f**, Mean radial (left panel) and axial (right panel) PPSF for a population of opsin-positive PV cells when stimulating 10 targets at once. Error bars indicate standard deviation over n=7 cells. **g**, Incidence of multiple action potentials (i.e., doublets) in response to a single laser pulse. Trials shown are those where the patched cell spiked at least once. Postsynaptic neuron same as in panel (b). 10 ms time windows from the start of a 3 ms laser pulse are displayed. Exemplary traces displayed on the right are marked by arrows.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

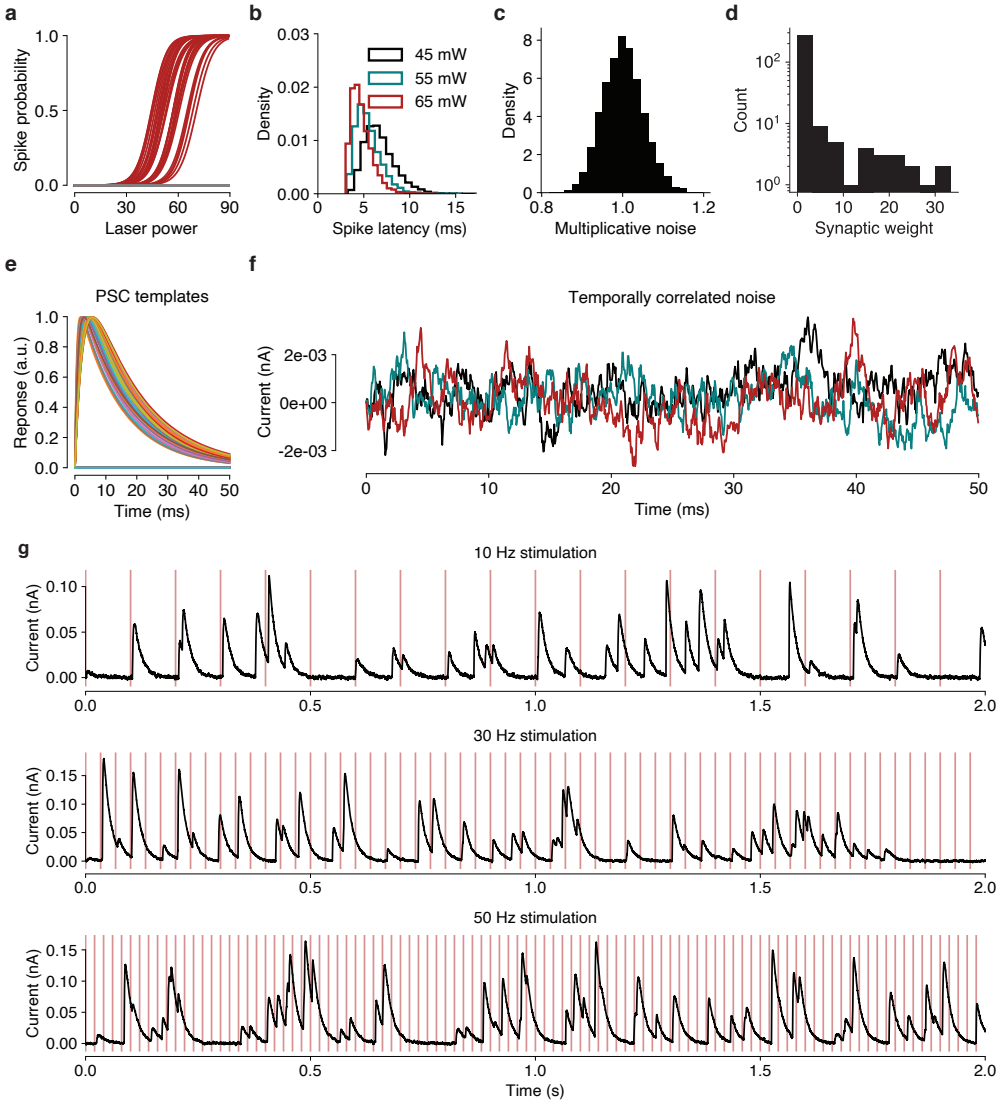


Figure S2: Components of an example simulated connectivity mapping experiment. **a**, Presynaptic spike probabilities as a function of laser power. Connected neurons shown in red, disconnected neurons shown in gray. **b**, Example histograms (shown as normalized densities) of presynaptic spike latencies for three typical laser powers. Average spike and transmission times decrease quadratically with increasing power. **c**, Example log-normal distribution of multiplicative noise terms that induce trial-to-trial variability in PSC amplitude. **d**, Distribution of synaptic weights, shown on log scale. **e**, Example PSC templates. Disconnected neuron templates represented by the zero vector. **f**, Three example noise processes sampled from a first-order autoregressive process. **g**, Example connectivity mapping experiment under increasing rates of stimulation. Simulation has 300 neurons with 10% connection probability and spontaneous synaptic currents occuring at 5 Hz.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

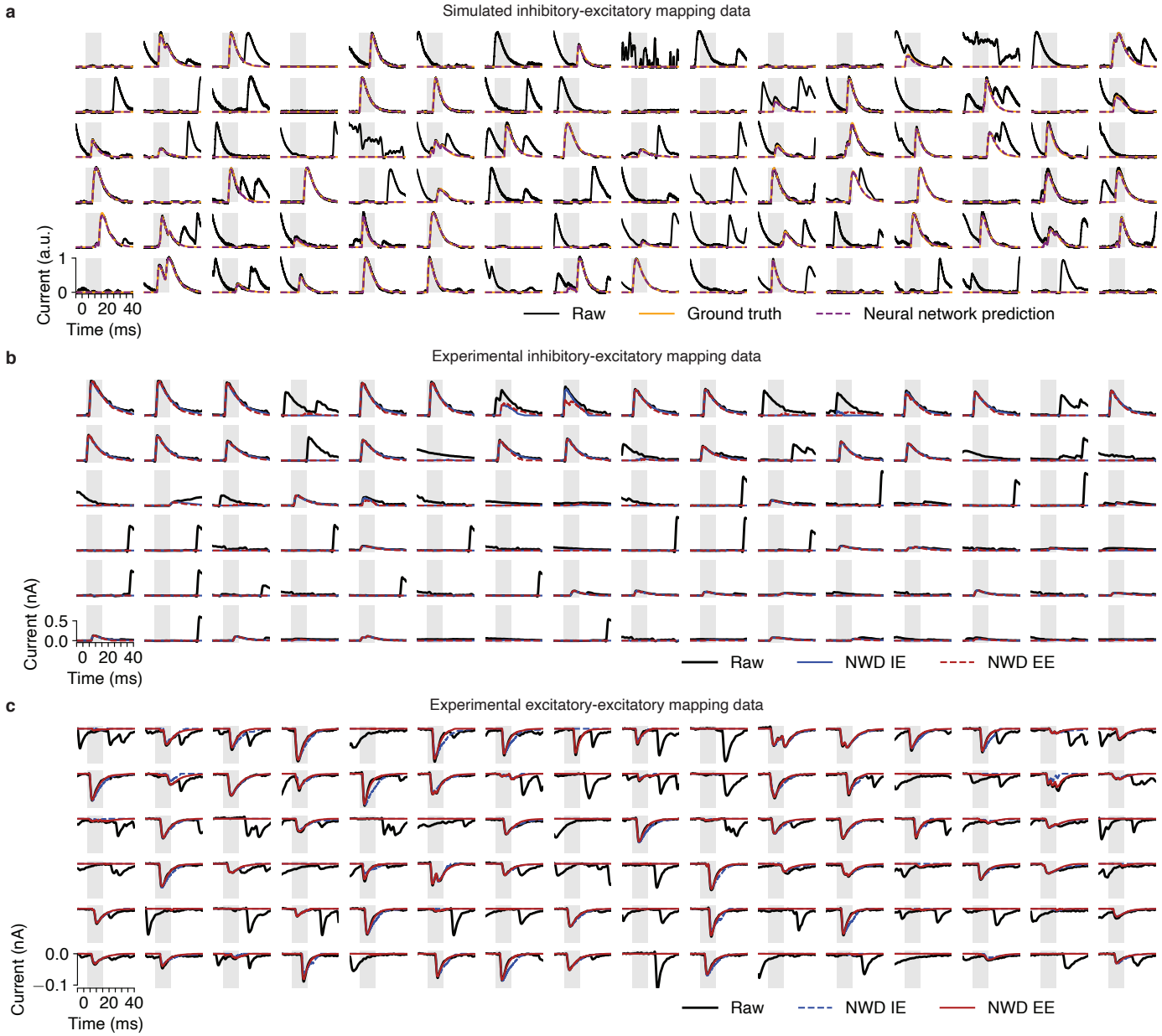


Figure S3: **a**, NWD network performance on simulated test data. Solid black line shows raw data given as input to the neural network. Orange line shows target trace. Dashed purple line shows prediction from the NWD network. **b**, Performance of NWD trained on simulated data matched to IE mapping experiments (solid, blue). For comparison, NWD trained on simulated data matched to EE mapping experiments shown as dashed red line. **c**, Performance of NWD trained on simulated data matched to EE mapping experiments (solid, red). For comparison, NWD trained on simulated data matched to IE mapping experiments shown as dashed blue line.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

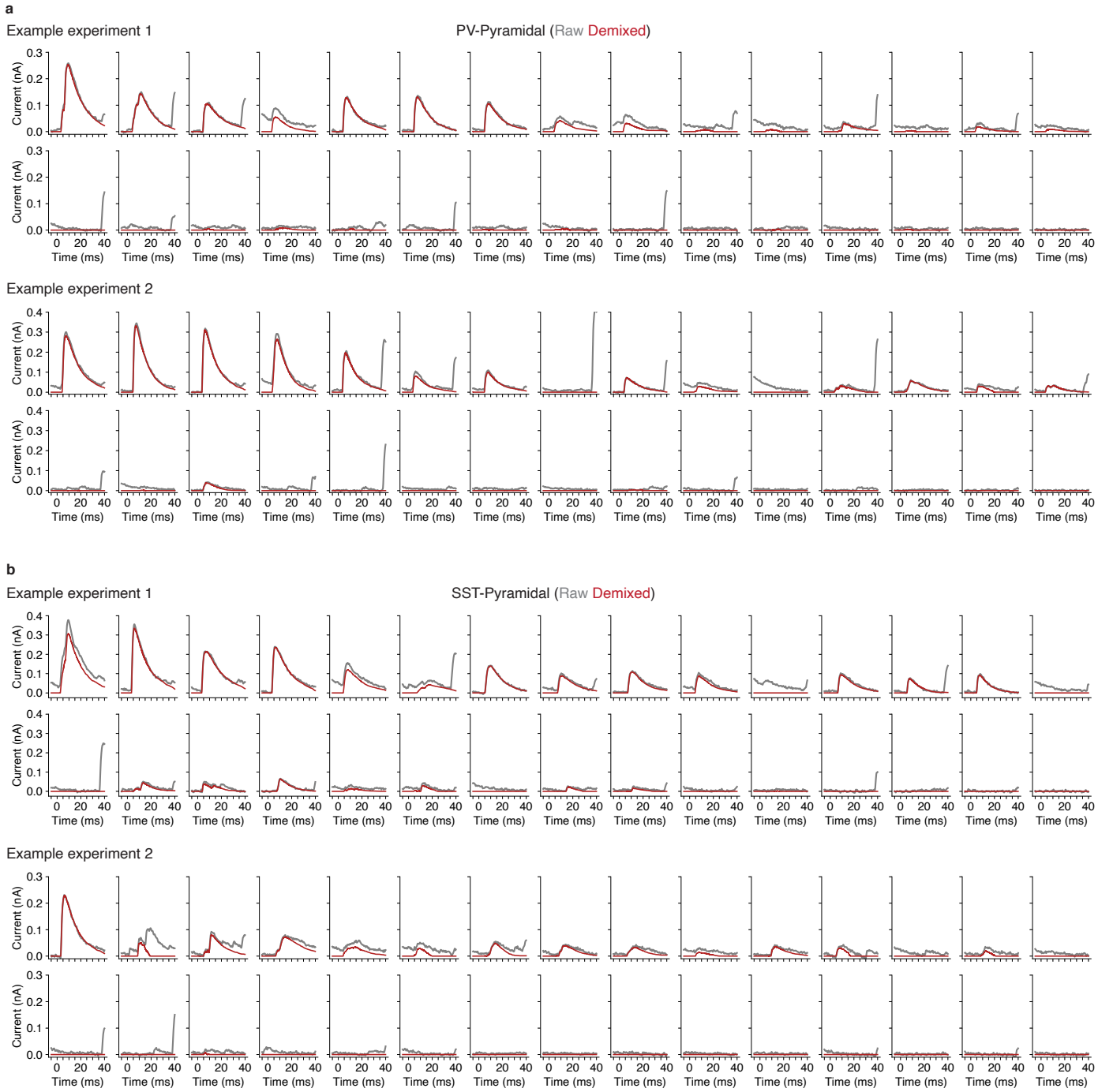


Figure S4: Example application of NWD to individual PSCs evoked by holographic ensemble stimulation in PV (**a**) and SST (**b**) to pyramidal mapping experiments. PSCs selected uniformly at random and sorted by magnitude of postsynaptic response.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

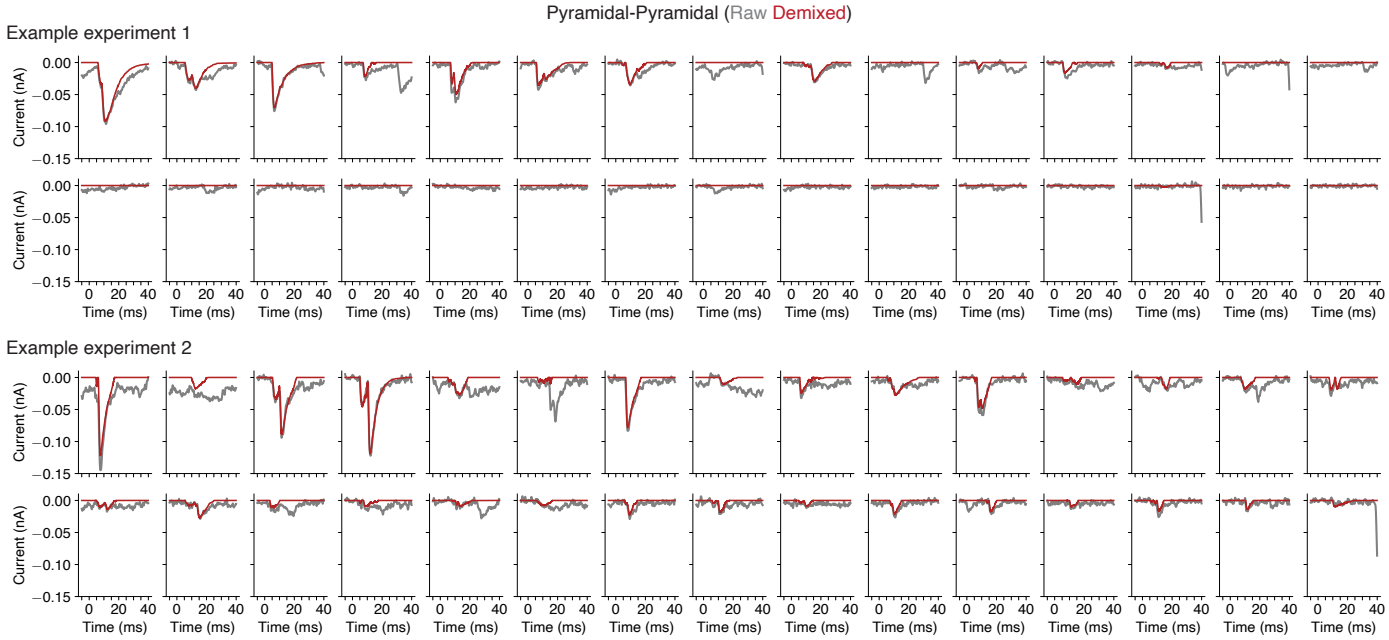


Figure S5: Example application of NWD to individual PSCs evoked by holographic ensemble stimulation in a pyramidal-pyramidal mapping experiment. PSCs selected uniformly at random and sorted by magnitude of postsynaptic response.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

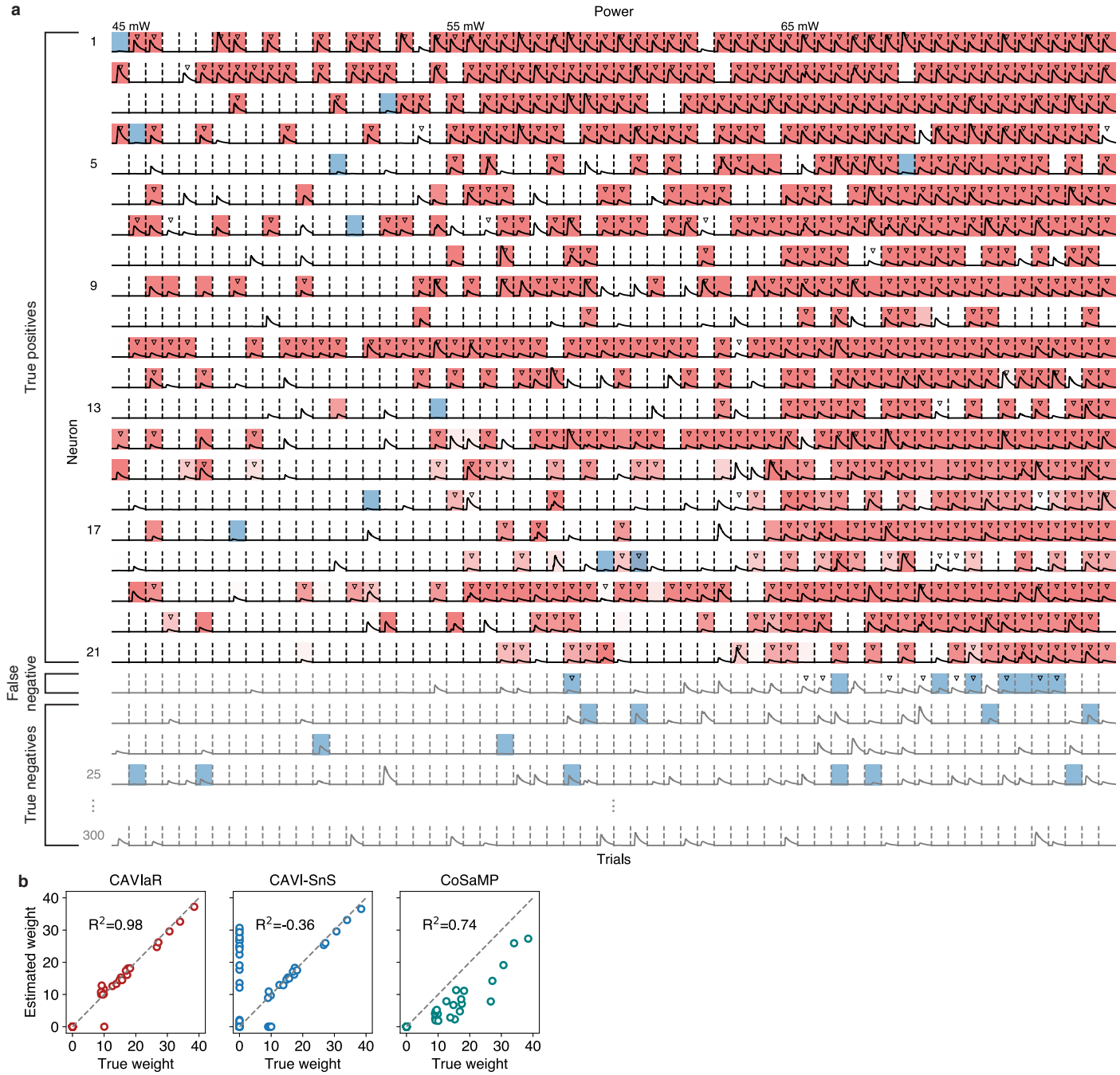


Figure S6: Large-scale classification of presynaptic spikes, spontaneous synaptic currents, and synaptic connectivity in a 3 minute simulated connectivity mapping experiment with 300 neurons mapped using 10-target ensemble stimulation at 10 Hz. Spontaneous PSCs occur at a rate of 5 Hz. **a**, "Checkerboard" visualization of CAVIaR model inference. Each row shows sample of PSCs evoked by stimulating the listed neuron across multiple powers. In this example, 9 neurons in addition to the listed neuron are stimulated on each trial. Shaded red cells indicate detected presynaptic spike for listed neuron. Shaded blue cells indicate detected spontaneous PSC. Triangles indicate ground-truth presynaptic spikes. Accuracy of inferred presynaptic spikes for this simulation, 94.3%. Traces shown in gray indicate neurons that the model declared disconnected. Out of 300 neurons mapped, one putative connection was a false negative (neuron 22) due to CAVIaR incorrectly assigning its PSCs to other stimulated neurons or to spontaneous activity. **b**, Comparison of connectivity inference accuracy between CAVIaR, CAVI-SnS, and CoSaMP.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

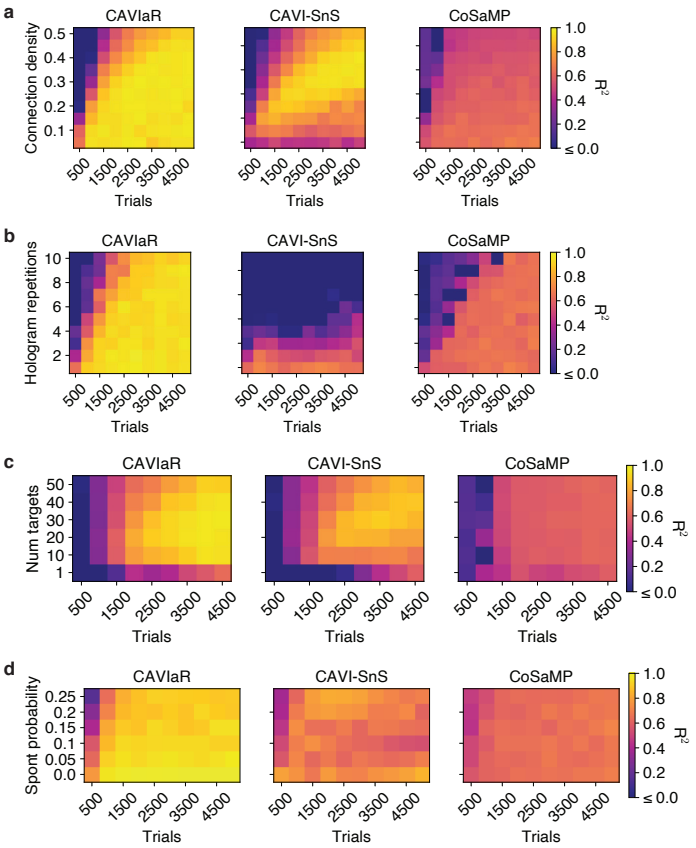


Figure S7: Performance and phase transitions for three connectivity inference techniques for a population of 1000 candidate presynaptic neurons. **a**, Performance as a function of underlying density of synaptic connectivity (connection probability). **b**, Performance as a function of stimulus diversity (hologram repetitions). **c**, Performance as a function of the number of simultaneously targeted neurons. **d**, Performance as a function of probability of spontaneous PSC lying within admissible PSC initiation window.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

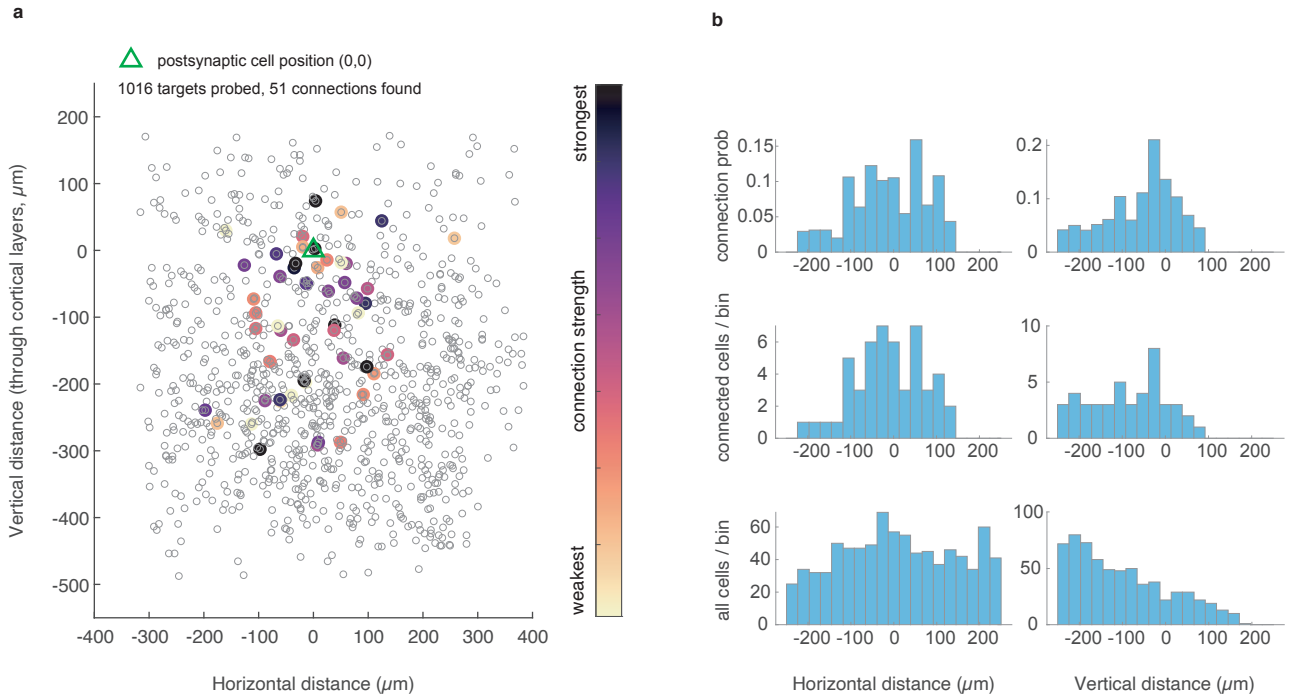


Figure S8: Connection strength and probability (normalized within each experiment, c.f. Figure 6). **a**, Summary of connection strengths across experiments (Z-projected, n=7 experiments). Map obtained by normalizing connection strengths from weakest to strongest within each experiment, then overlaying on a common map. This shows that, for a given experiment, the strongest connection can be far from the postsynaptic neuron. **b**, Histograms showing the connection probability, number of connected cells in each bin (26 μm wide), and number of probed cells in each bin respectively (top to bottom). Histograms separated into horizontal and vertical (i.e., translaminar) dimensions.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

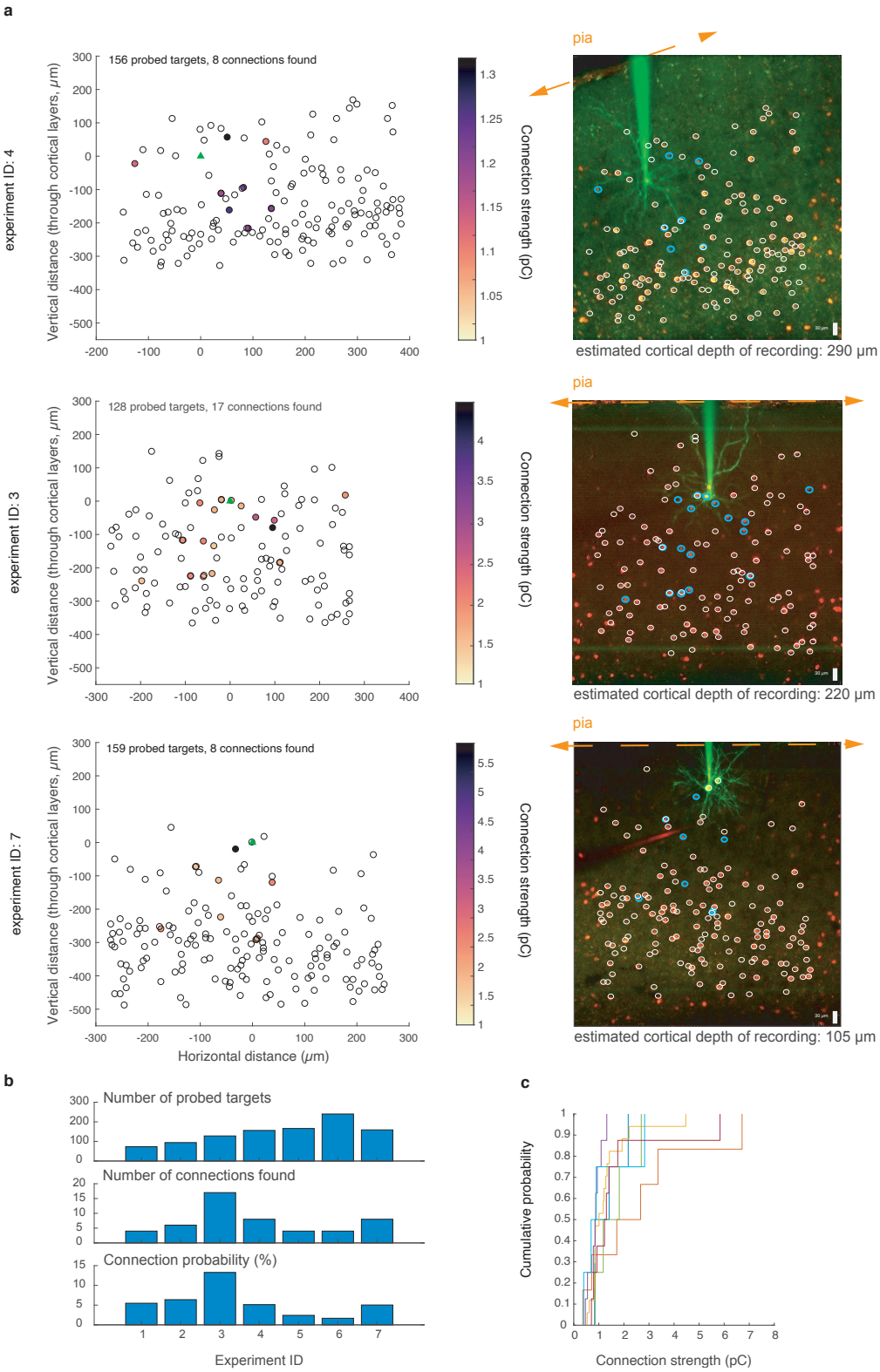


Figure S9: Example connectivity maps from individual experiments. **a**, Three example Z-projected maps for postsynaptic cells recorded at different depths of L2/3 in the cortex. Left panels display positions of probed (empty circles) and connected (colored circles) cells. The green triangle marks the position of postsynaptic cell. Right panels display max-projected, contrast-enhanced images of the mapped volumes. White circles indicate position of probed cells, cyan circles indicate detected connections. Postsynaptic cell perfused with Alexa dye. **b**, Summary of number of probed targets, connections found, and connection probability across experiments. **c**, Cumulative distribution functions of synaptic strength across experiments.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

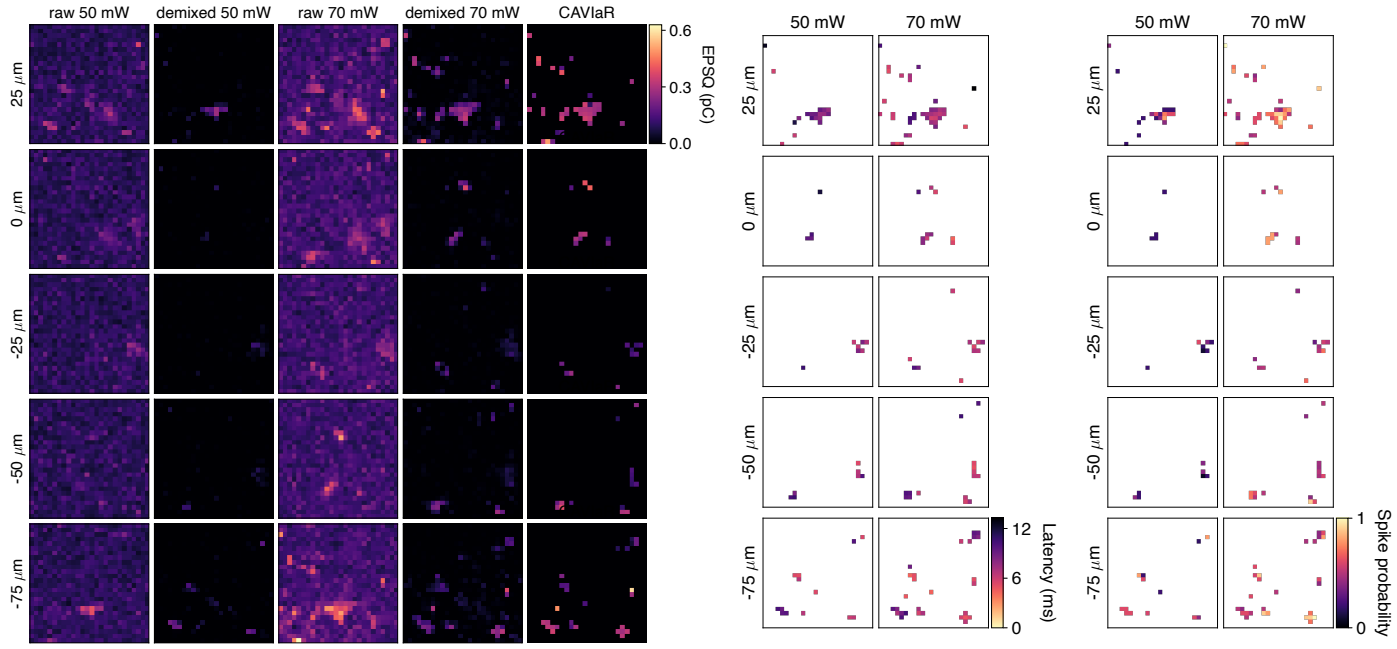


Figure S10: Demixing and denoising of optogenetic grid mapping data using NWD and CAVIaR. Example shows pyramidal-pyramidal connectivity mapping (Vglut1-Cre; st-ChroME-GCaMP7). Latencies (middle two columns) estimated by first computing canonical PSC waveforms from the inferred spike matrix using ridge regression, then reporting the time (in ms) required to cross 10% of the PSC amplitude (Methods). Note the increase in spike latency and decrease in spike probability with increasing distance from the ROI centroid at maximal power. Postsynaptic cell is on plane 0 μm.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507926; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

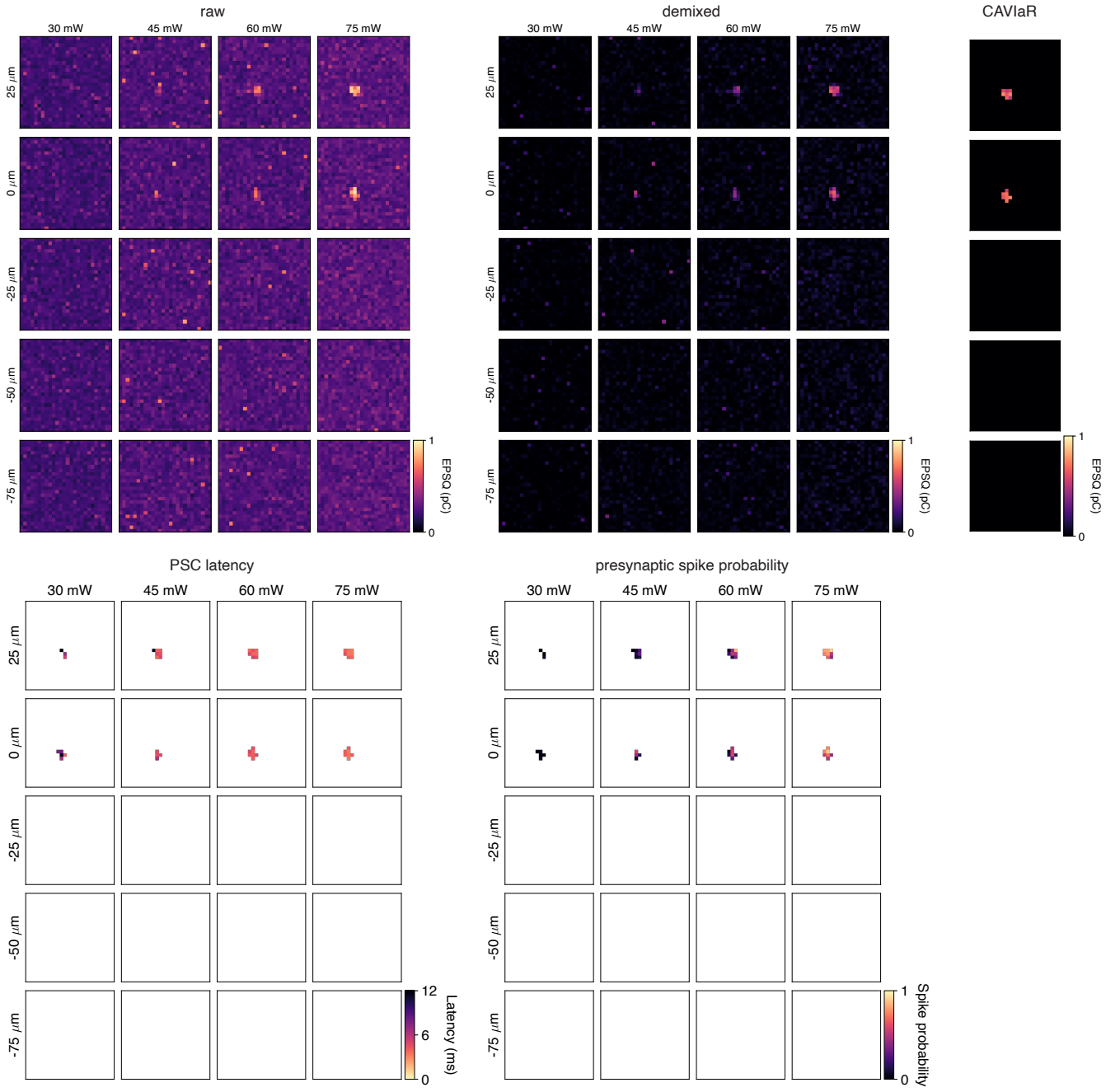


Figure S11: Demixing and denoising of optogenetic grid mapping data using NWD and CAVIaR. Example shows PV-pyramidal connectivity mapping (PV-Cre ChroME2f). Postsynaptic cell is on plane 0 μm.