

# Inferring structure of cortical neuronal networks from activity data: A statistical physics approach

Ho Fai Po <sup>a,\*</sup>, Akke Mats Houben <sup>b,c</sup>, Anna-Christina Haeb <sup>b,c</sup>, David Rhys Jenkins <sup>d,e</sup>, Eric J. Hill <sup>d,f</sup>, H. Rheinallt Parri <sup>d,e</sup>, Jordi Soriano <sup>b,c</sup> and David Saad <sup>a</sup>

<sup>a</sup>Department of Mathematics, Aston University, Birmingham B4 7ET, United Kingdom

<sup>b</sup>Departament de Física de la Matèria Condensada, Universitat de Barcelona, Barcelona E-08028, Spain

<sup>c</sup>Universitat de Barcelona Institute of Complex Systems (UBICS), Barcelona 08028, Spain

<sup>d</sup>College of Health and Life Sciences, Aston University, Birmingham B4 7ET, United Kingdom

<sup>e</sup>Aston Institute for Membrane Excellence, Aston University, Birmingham B4 7ET, United Kingdom

<sup>f</sup>Department of Chemistry, Loughborough University, Loughborough, Leicestershire LE11 3TU, United Kingdom

\*To whom correspondence should be addressed: Email: [h.po@aston.ac.uk](mailto:h.po@aston.ac.uk)

Edited By Derek Abbott

## Abstract

Understanding the relation between cortical neuronal network structure and neuronal activity is a fundamental unresolved question in neuroscience, with implications to our understanding of the mechanism by which neuronal networks evolve over time, spontaneously or under stimulation. It requires a method for inferring the structure and composition of a network from neuronal activities. Tracking the evolution of networks and their changing functionality will provide invaluable insight into the occurrence of plasticity and the underlying learning process. We devise a probabilistic method for inferring the effective network structure by integrating techniques from Bayesian statistics, statistical physics, and principled machine learning. The method and resulting algorithm allow one to infer the effective network structure, identify the excitatory and inhibitory type of its constituents, and predict neuronal spiking activity by employing the inferred structure. We validate the method and algorithm's performance using synthetic data, spontaneous activity of an *in silico* emulator, and realistic *in vitro* neuronal networks of modular and homogeneous connectivity, demonstrating excellent structure inference and activity prediction. We also show that our method outperforms commonly used existing methods for inferring neuronal network structure. Inferring the evolving effective structure of neuronal networks will provide new insight into the learning process due to stimulation in general and will facilitate the development of neuron-based circuits with computing capabilities.

**Keywords:** biological neuronal networks inference, neuronal-type classification, kinetic Ising model, generalized maximum likelihood, expectation-maximization algorithms

## Significance Statement

Inferring effective connectivity from firing patterns in neuronal networks is a long-standing problem, crucial for understanding how connectivity changes due to stimulation and plasticity occur. We introduce a probabilistic method to infer the excitatory/inhibitory type of neurons, links existence, and the effective coupling strengths in cortical neuronal networks from neuronal spiking recordings. We validate the efficacy of our method using both *in silico* and *in vitro* experiments, demonstrating that the effective structure revealed agrees with the true values and the key experimental observable such as modular organization. Additionally, the predicted neuronal activity using the inferred structures aligns with the observed data. The method will impact significantly on both fundamental studies in neuroscience and the development of neuron-based computing devices.

## Introduction

Revealing how cortical neuronal networks connectivity evolves in time, spontaneously or under stimulation, is a foundational question in neuroscience, in particular, understanding how cortical neurons learn from repeated stimulation through changes in topology and synaptic strengths (1–7). While there are many tools for investigating macroscopic brain activities and changes, such as functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG), and electroencephalography (EEG) (8–10), investigating the

microscopic changes which occur in neuronal tissues noninvasively remains a challenge. While *in vivo* interrogation of neuronal networks at the microscopic level remains difficult, *in vitro* techniques such as multielectrode array (MEA) (11) and calcium imaging (12, 13) facilitate the monitoring and stimulation of neuronal tissues at cellular resolution and open the way to greater understanding of the learning process. New developments in the application of neuron-based circuits with computing capabilities make the need to understand the exact relationship between learning and stimulation more urgent and relevant (14–17).

**Competing Interest:** The authors declare no competing interests.

**Received:** June 10, 2024. **Accepted:** December 11, 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of National Academy of Sciences. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Machine learning (ML) and artificial intelligence (AI) play an increasingly crucial role in our daily lives. However, training ML and AI systems require unsustainable computing power and energy consumption (18) and mostly lack the ability to adapt their structure in response to a changing situation. This has given rise to the search for alternative computing paradigms, in particular, the emerging field of biological computation, which aims at employing human neuronal networks (hNNs) as processing units in biological computing devices. These developments, such as using cortical brain organoids for nonlinear curve prediction (19) and employing cortical neuronal networks for decision-making in simulated gaming environments (20), have recently drawn the attention of both researchers and the general public. These breakthroughs point to the immense potential of hNNs in biological machine learning. The common belief is that plasticity and learning occur through appropriate stimulation in neuronal-based computing devices, such that both network topology and synaptic strengths evolve to structures that can carry out specific data/stimulation-driven tasks. However, the tools needed to investigate the evolving structure and support the understanding of how these hNNs-based devices operate are currently lacking. Hence, it is crucial to develop a principled inference tool that can reveal the effective cortical network structure, to better comprehend the mechanism that gives rise to task learning from stimulation.

Various methods have been employed to infer the effective neuronal network from its firing patterns. Commonly used techniques include generalized transfer entropy (GTE) (21), dynamic causal modeling, Granger causality (22), maximum entropy model (23), and generalized linear model (24). However, these methods have significant limitations. They can only measure directional causation between neurons and identify the existence of an effective connection by setting an appropriate but somewhat arbitrary threshold. These methods cannot find the excitatory or inhibitory type of neurons without manipulating the network through stimulation or channel blocking (21, 25), nor can they determine the model's effective synaptic strengths. As such, they are less suitable for studies requiring long-term monitoring and careful consideration of stimulation protocols as they may potentially affect network development as observed in the training of cortical neurons-based learning machines (19, 20).

Moreover, these methods often overlook the activities of nearby neurons and fail to capture interactions between multiple neurons, resulting in inaccurate inference. Additionally, methods such as GTE do not provide a probabilistic model for neuronal activity, making it difficult to predict or reproduce network activity using the inferred effective connectivity structure for validation or further investigations.

To fill these gaps, here we advocate mapping neuronal activities onto the kinetic Ising model of statistical physics as they share some common features, such as binary state of activity, unidirectional nonequilibrium and nonlinear dynamics, and multi-neuron interactions. Mapping neuronal activities onto the kinetic Ising model facilitates the inference of interneuron interactions (26, 27). Yet, inferring the kinetic Ising model structure and properties is challenging, and probabilistic methods have been developed in the statistical physics community for inferring the underlying directional interaction strengths from observation sequences (28–30). However, most methods have been derived for a simple model, where coupling strengths are Gaussian distributed with mean zero and small variance, which does not hold in biological neuronal networks, with the exception of (31) under specific conditions.<sup>a</sup> Thus, a principled probabilistic method that

can identify connectivity, synaptic strengths, and the excitatory/inhibitory type of each neuron is required.

In this paper, we introduce an algorithm that combines models from statistical physics, Bayesian inference, and probabilistic machine learning to infer the effective architecture of biological neuronal networks from firing patterns. Our proposed algorithm overcomes some of the limitations of existing methods and infers not only the effective connectivity from neuronal firing but also the neuronal characteristics (inhibitory/excitatory) and the existence of connections. Furthermore, unlike conventional methods, our algorithm provides a probabilistic model that facilitates the simulation of neuronal activities using the inferred architecture, which can be used for structure validation, prediction, and further investigations. We evaluate the performance of our algorithm using synthetically generated data, in silico neuronal network emulator data, and calcium imaging recordings of real in vitro cortical networks with patterned (32) and unpatterned substrates, demonstrating excellent agreement between the effective inferred model and the corresponding data.

## Model

### Kinetic Ising model

The kinetic Ising model (33) in statistical physics studies the activity of spins in a system with asymmetric coupling strengths. Sharing the common feature that spin (neuron) configuration is influenced by directional coupling (synaptic) strengths and the state of its neighbors, the kinetic Ising model is suitable for describing neuronal spiking activities. Here, we map the binary neuronal spiking activity onto the kinetic Ising model, which is a discrete-time nonequilibrium probabilistic structure. Consider a system of  $N$  neurons within an interacting neuronal network. We denote a discrete variable  $s_i^t = \pm 1$  when neuron  $i$  is spiking or silent at time step  $t$ , respectively, for  $i = 1, \dots, N$ . Previous works (21) suggest that considering interactions across multiple time intervals has minimal effect on the inference, so we define the transitional probability of neuron  $i$  at time  $t$ , given the neuronal activities at time  $t - 1$ , as

$$P(s_i^t | s^{t-1}, \mathbf{J}, H_i) = \frac{\exp\left[\left(H_i + \sum_j J_{ij} s_j^{t-1}\right) s_i^t\right]}{2 \cosh\left(H_i + \sum_j J_{ij} s_j^{t-1}\right)}, \quad (1)$$

where  $\mathbf{J} = \{J_{ij}\}_{ij}$  and  $J_{ij}$  denotes the synaptic strength from  $j$  to  $i$ , and  $H_i$  denotes the external local field acting on neuron  $i$ , which can be interpreted as the activeness of  $i$  when no signal is received from its neighbors. A positive (negative)  $J_{ij}$  represents the excitatory (inhibitory) strength of a signal sent from neuron  $j$  to neuron  $i$  when  $j$  spikes, while  $J_{ij} = 0$  indicates that neuron  $j$  is not effectively connected to  $i$ . Since the activities of neurons at time  $t$  depend only on activities in the previous time step and hence do not exhibit explicit same-bin interdependence, the probability of activities for all neurons is given by  $P(\mathbf{s}^t | \mathbf{s}^{t-1}, \mathbf{J}, \mathbf{H}) = \prod_i P(s_i^t | \mathbf{s}^{t-1}, \mathbf{J}, H_i)$ . In order to infer the neuronal type and effective link existence, we introduce two sets of latent variables,  $z_j = \pm 1$  representing the excitatory and inhibitory type of neuron  $j$ , respectively; and  $\phi_{ij} = \{1, 0\}$  indicating whether  $j$  is connected to  $i$  or not, respectively. One of the limitations of the kinetic Ising model is that it does not consider longer synaptic time delays, but as indicated in a number of studies (2, 21, 34), their influence is much weaker than that of single time step delays.

## Adapting Ising model to living neuronal networks

Based on the mathematical model defined, we introduce prior distributions to align with real-world neuronal network properties. For clarity, we use  $p(\cdot)$  to denote prior distributions for the corresponding variables, while  $P(\cdot)$  denotes the probability more generally. We define

$$p(z_j) = \gamma \delta_{z_j, +1} + (1 - \gamma) \delta_{z_j, -1}; \quad (2)$$

$$p(\phi_{ij} | a) = \delta_{\phi_{ij}, 1} \theta_{ij} e^{-a l_{ij}} + \delta_{\phi_{ij}, 0} (1 - \theta_{ij} e^{-a l_{ij}}); \quad (3)$$

$$p(J_{ij}) = \sum_{z_j} \sum_{\phi_{ij}} p(J_{ij} | \phi_{ij}, z_j) p(\phi_{ij}) p(z_j); \quad (4)$$

$$p(H_i) = \frac{e^{-\frac{(H_i - \mu_H)^2}{2v_H}}}{\sqrt{2\pi v_H}}, \quad (5)$$

where  $\gamma \in [0, 1]$  represents the proportion of excitatory neurons in the network;  $\theta_{ij}$  is the prior probability for the existence of a link from neuron  $j$  to neuron  $i$ , which can be adjusted for convergence assistance in networks with predefined or dictated connectivity (32, 35), or set to 1 otherwise;  $a \in \mathbb{R}^+$  is a controlling parameter that accounts for the decay in connection probability; and  $l_{ij} = l_{ji}$  the Euclidean distance between  $i$  and  $j$ , thus  $e^{-a l_{ij}}$  reflects the exponential decay in connection probability with distance;  $\mu_H$  and  $v_H \in \mathbb{R}$  represent the mean and variance of the distribution for  $\mathbf{H}$ , reflecting the distribution of neuron inherent activeness. The distribution of  $J_{ij}$  comprises a mixture of distributions since it is conditioned on whether there exists a link between two neurons and whether the type of the interaction is inhibitory or excitatory; the different cases may be characterized by different parameters. We define the conditional probabilities  $p(J_{ij} | 0, z_j) = \delta(J_{ij}) \approx \mathcal{N}(0, \epsilon)$  for small  $\epsilon$  (disconnected case) and  $p(J_{ij} | 1, z_j) = \mathbb{1}_{z_j J_{ij} > 0} e^{-(\ln z_j J_{ij} - \mu_j^z)^2 / (2v_j^z)} / \sqrt{2\pi v_j^z}$  for  $\mu_j^z, v_j^z \in \mathbb{R}^+$  and  $z_j = \pm 1$ , where  $\mathbb{1}$  is an indicator function ensuring all connections from  $j$  are of the same sign, and  $z_j J_{ij}$  follows a log-normal distribution to ensure the probability of coupling strengths of inhibitory or excitatory neurons sum to 1. The compact notation merely deals with the positive (excitatory) and negative (inhibitory) interaction strengths within the same expression by allowing for different distribution characteristics. For brevity, we denote  $\boldsymbol{\rho} = \{\gamma, a, \mu_H, v_H, \mu_j^\pm, v_j^\pm\}$  as the set of all defined hyperparameters (parameters that control the distributions of  $\mathbf{J}$  and  $\mathbf{H}$ ). By combining the distributions we define the evidence function as

$$\ln P(\mathbf{s} | \mathbf{J}, \mathbf{H}, \boldsymbol{\rho}) + \ln p(\mathbf{H} | \boldsymbol{\rho}) + \ln p(\mathbf{J} | \boldsymbol{\rho}). \quad (6)$$

In principle, the related probability of Eq. 6 should be integrated over  $\mathbf{J}$  and  $\mathbf{H}$  to obtain the marginalized probability  $P(\mathbf{s} | \boldsymbol{\rho})$  but is approximated at the peak of the posterior ("Maximum A Posteriori," MAP) (36).

## Connectivity network inference

We extend the kinetic Ising model by introducing two sets of latent variables and prior distributions for all variables to align with the nature of realistic biological neuronal networks. The Methods section and the [Supplementary Material](#) provide details on how the (latent) variables and hyperparameters can be evaluated in a principled way, enabling us to interpret the neuronal type from  $z_j$ , link existence through  $\phi_{ij}$ , and effective synaptic

strength from  $J_{ij}$ . In general, we are applying the generalized maximum-likelihood (GML) approach (whereby parameters and hyperparameters are iteratively determined, also termed the evidence procedure) and the variational expectation-maximization (EM) algorithm (36) to infer  $\mathbf{J}$ ,  $\mathbf{H}$ , and  $\boldsymbol{\rho}$ .

## Results

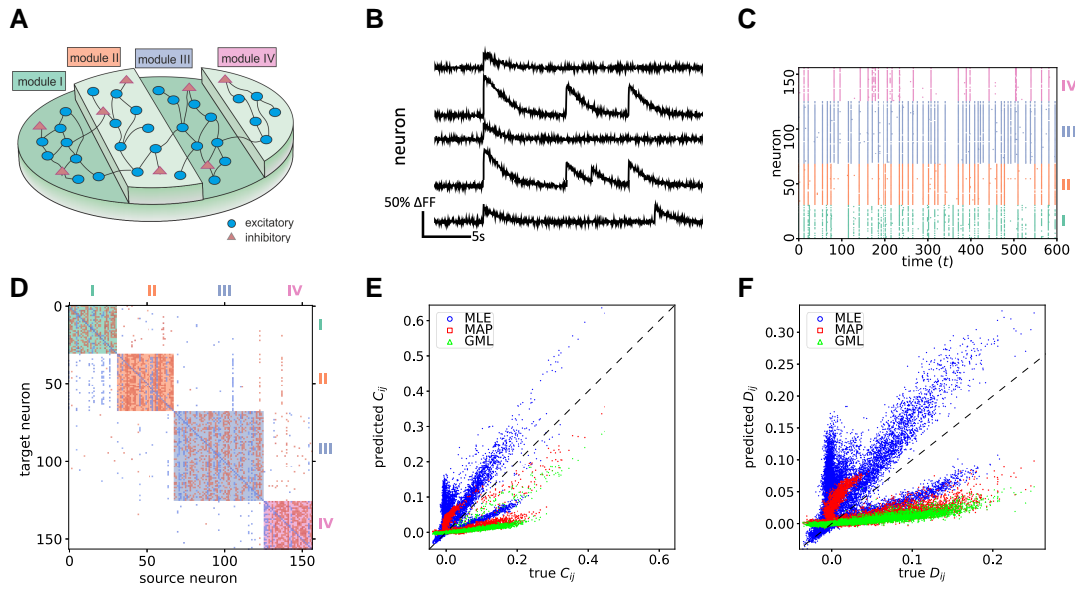
To validate the efficacy of our model and inference algorithm, we perform tests using three types of data: synthetic data generated by the kinetic Ising model, in silico data from computational models, and in vitro data from biological cortical neuronal activity recordings.

In the [Supplementary Material](#), we discuss and examine the properties of different algorithms for kinetic Ising model inference, comparing them with our proposed method in detail, and show that the proposed method performs significantly better than the naïve mean-field and maximum-likelihood approaches. In our GML inference for the kinetic Ising model, coupling strengths are not zero-mean and of small variance, and more importantly, the method allows one to infer the existence of connections between neurons. Here, we focus on the results obtained from the in silico model with patterned substrates as well as the two realistic in vitro cortical neuronal activities. One of these in vitro activities is of a homogeneous network with potassium stimulation, while the other focuses on spontaneous activities in a neuronal network with patterned substrates.

### In silico experimental data

Since validating the accuracy of neuronal-type classification and link existence is very costly and extremely difficult for in vitro experiments, we first test our algorithm on emulated in silico data since ground-truth topology is known. We generate an in silico neuronal network in which the culture is lying over a striped topographical patterned substrate (32) as illustrated in Fig. 1A. The patterned substrate facilitates high connection density between neurons located on the same stripe but allows for lower connection density across stripes, which is an interesting feature that can be expressed and tested using our method. Spontaneous neuronal activity is then generated on this network. We note that connectivity within a stripe is so strong that each one effectually shapes a "module" of highly interacting neurons. The detailed methodology of how the data are generated is described in the Methods section.

Figure 1B shows representative fluorescence traces of spontaneous neuronal activity generated in silico, and Fig. 1C shows the raster plot of activity for the entire network, with colors indicating the respective module neurons belong to. Using the activity as input, we infer the effective structure  $\mathbf{J}$  using the Maximum-Likelihood Estimator (MLE), MAP, and GML. Their methodology is described in detail in the Methods section and [Supplementary Material](#). For the prior distributions imposed in MAP and the evidence approximation, we first fix the hyperparameter  $a = 0.1$ , the fraction of excitatory neurons,  $\gamma = 0.8$  and  $\theta_{ij} = 0.5^{m_d}$ , where  $m_d$  is the number of stripes that separate neurons  $i$  and  $j$ . The values employed for the other hyperparameters used both for MAP inference and as initial values for the GML inference, are obtained using  $\mathbf{J}$  and  $\mathbf{H}$  inferred by MLE. The value of  $a$  reflects the fact that neurons are less likely to be connected if they are far away from each other;  $\gamma$ , the ratio between excitatory and inhibitory neurons is already statistically known; while for  $\theta_{ij}$ , we exploit available information about the physical structure and the value of 0.5 is chosen arbitrarily. We tested the performance of the inference methods using other



**Fig. 1.** A) A three-dimensional sketch illustrating the in silico experimental setup. The neurons ( $N = 156$ ) are lying on stripes (modules I to IV) of patterned substrates, which suppress cross-connections between different stripes. B) Example in silico traces of five neurons. C) The raster plot displays neuronal activity inferred from the in silico traces, with each neuron's color corresponding to the module it belongs to. D) The inferred structure  $\mathbf{J}$  of the in silico model obtained using GML and represented by a connectivity matrix. The background color is a mixture of two colors, with each color corresponding to the module of the source and target neurons. Each entry corresponds to the coupling strength  $J_{ij}$ . A positive (negative) strength colored in red (blue) refers to excitatory (inhibitory) signals sent from source ( $i$ ) to target ( $j$ ) neurons. E) The predicted equal time covariance  $\mathbf{C}$  against the true values evaluated from data. F) The predicted delayed time covariance  $\mathbf{D}$  against the true values evaluated from data.

values of  $\alpha$ ,  $\gamma$ , and  $\theta_{ij}$  and the results are similar. This is because when the number of samples is large enough, the effect of the priors is suppressed. Table S1 specifies all parameters, their type, description, and the values they acquire for clarity.

The inferred connectivity matrix of  $\mathbf{J}$  using GML is shown in Fig. 1D, where the background mask is the mixture of colors corresponding to the source and target neurons, while inhibitory and excitatory transmissions are colored in blue and red, respectively. We note that the diagonal entries are mostly inhibitory; this does not necessarily correspond to physical connections but may reflect the fact that neurons are less likely to spike again after firing due to their refractory period. Another key feature is that except for the diagonal, signals in the same column share the same state, which corresponds to the inferred excitatory/inhibitory type of the neurons; this is validated by comparing the inference results to the true in silico model.

The most appropriate measures of success when contrasting inferred (“infer”) and ground-truth (“true”) topologies are the positive predictive value (PPV) and negative predictive value (NPV), or precision of neuronal type  $P(z_{\text{true}} | z_{\text{infer}})$ , compared with the prior-based random guess, where excitatory type is the positive case and inhibitory the negative one, as Table 1 shows. While being less relevant due to the biased nature of the variables, the true positive rate (TPR, sensitivity) and the true negative rate (TNR, specificity),  $P(z_{\text{infer}} = +/ - | z_{\text{true}} = +/ -)$  are also presented alongside the random guess, for completeness. Overall predictive performance is summed over both cases with the respective probabilities. Notably, no existing method can infer the excitatory and inhibitory type of neurons from single spontaneous activity recordings without interference, such as channel blocking. With a significant improvement over a prior-based random guess, our method identifies well the individual neuronal type.

Figure 1D reveals a strong clustering effect for connectivity within modules, indicating that two neurons within the same

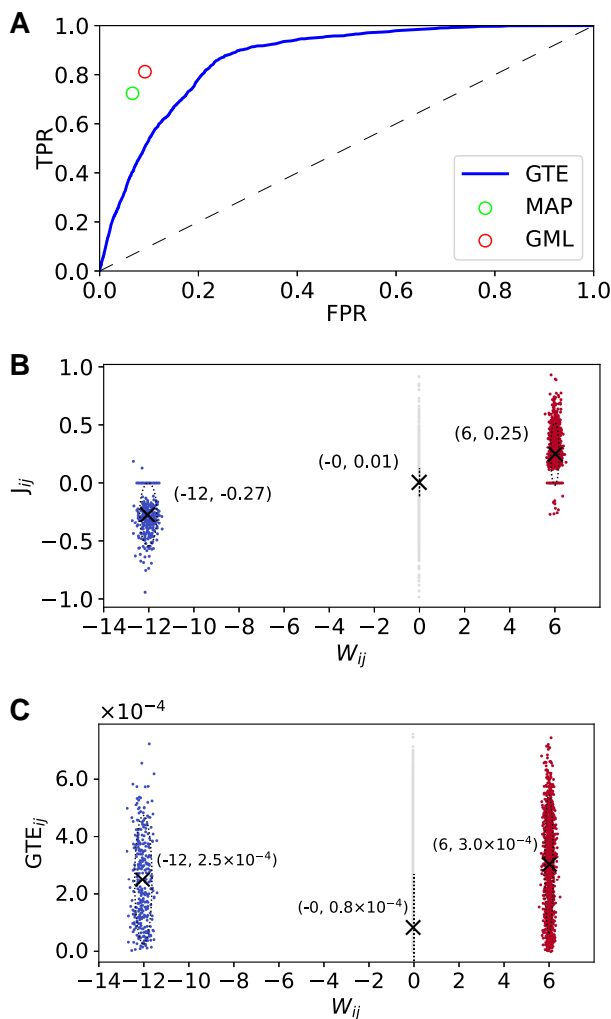
**Table 1.** Success measures in identifying neuron type, TPR (sensitivity), TNR (specificity), and PPV of the in silico model study.

Measure	TPR/TNR	PPV/NPV	Random
Excitatory	0.88	0.96	0.8
Inhibitory	0.84	0.64	0.2
Overall	n/a	0.87	0.68

module have a higher probability of being connected compared with two neurons from different modules. This suggests that the GML captures the structure of the topographical substrate. We then test the performance of inferring existing links using MAP, GML, and GTE (21). Since GTE is based on an assigned transfer entropy threshold value for identifying links, we plot the complete receiver operating characteristic (ROC) curve for all threshold values, of TPR against FPR of identifying nonzero links as shown in Fig. 2A. The TPR and FPR of finding nonzero links using GML are 54% and 2.6%, respectively, and 57% and 3.3% for MAP, as indicated by the red and green circles, respectively. These values exceed the ROC curve generated by GTE, indicating that our method offers a higher TPR than GTE at the same FPR level, or a lower FPR at the same TPR level. We remark that, although MAP has a higher TPR than GML, its FPR is also higher. Additionally, we observe that the difference between MAP and GTE at the same FPR is lower than that of GML, which emphasizes the importance of hyperparameter optimization. It is essential to point out that our approaches detect link existence in a principled manner, eliminating the need for heuristic threshold decisions, and rendering the inference results more reliable.

A key feature that our method offers is the inference of the effective synaptic (coupling) strength between neurons, which is crucial for understanding the learning process in both neuroscience and cortical neuron-based learning machines. For example,





**Fig. 2.** A) The ROC curve plotting the TPR against the false positive rate (FPR) of identifying effective links for in silico experiments using GTE (continuous blue line). The TPR and TFR using our MAP and GML are marked as green and red nodes respectively. The dashed diagonal line in A refers to a random guess in this case. B and C) The scatter plot of inferred coupling strength  $J_{ij}$  and GTE against the true synaptic strength of the in silico model, respectively. The inhibitory, excitatory, and nonexistent links are colored in blue, red, and gray, respectively.

observing changes in synaptic strengths over extended periods allows one to study neuronal network plasticity due to stimulation and facilitates the design of efficient stimulation protocols in cortical neurons-based learning devices. We note that the neuronal spiking mechanism of the in silico emulator follows the Izhikevich model (37), which is distinct from the kinetic Ising model so that there is no direct mapping between the inferred ( $J_{ij}$ ) and true synaptic strength ( $W_{ij}$ , actual values used in emulator simulations). To compare the capability of our method with existing techniques, we plot the GML inferred weights  $J_{ij}$  and GTE values of each link against the true synaptic strengths  $W_{ij}$  in Fig. 2B and C, respectively. In the true model, the mean value of the inhibitory synaptic strength is  $-12$ , considerably higher than the excitatory synaptic strength's mean value of  $6$ . GTE as a specific manifestation of transfer entropy only offers nonnegative scores and therefore cannot directly distinguish inhibitory and excitatory links.

Figure 2C shows the mean value of GTE in the inhibitory group to be  $2.6 \times 10^{-4}$ , which is lower than the value in the excitatory group,  $3 \times 10^{-4}$ —contradicting the true values. On the other hand, as

Fig. 2B shows, the mean value of the inferred  $J_{ij}$  for the inhibitory connections is  $-0.27$ , which is of higher magnitude than that of excitatory connections,  $0.25$ . These results suggest that the inferred  $J_{ij}$  using our method agrees with the true model synaptic strengths ( $W_{ij}$ ). The GTE values also exhibit a significant variance compared with the magnitudes of the mean values. Notably, the mean value of the inferred  $J_{ij}$  for the missing links is very close to zero with a small variance, indicating that although some of the links are classified as nonzero, incorrectly, the majority are overwhelmingly close to zero. This false positive classification may result from the Gaussian approximation of the delta function in the prior probability  $p(J_{ij})$ . This suggests that the inferred effective coupling strength using our GML method agrees with the true model.

Another advantage of our method is that one can adopt the inferred structure  $\mathbf{J}$  and  $\mathbf{H}$  for generating artificial data through Monte Carlo simulations to predict neuronal activity or validate how well the inferred structure describes the true model by comparing quantities of interest, such as the equal time covariance  $\mathbf{C} = \{ \langle s_i^t s_j^t \rangle_t \}_{i,j}$ , and the delayed time covariance  $\mathbf{D} = \{ \langle s_i^t s_j^{t-1} \rangle_t \}_{i,j}$ . Thus, we employed the inferred structures using MLE, MAP, and GML to generate artificial neuronal activity and evaluate the predicted equal time covariance  $\mathbf{C}$  and delayed time covariance  $\mathbf{D}$  to its true values, as Fig. 1E and F shows. We can see that the predicted  $\mathbf{C}$  and  $\mathbf{D}$  values closely align with the true values for all methods, suggesting that the kinetic Ising model effectively explains in silico neuronal activities.

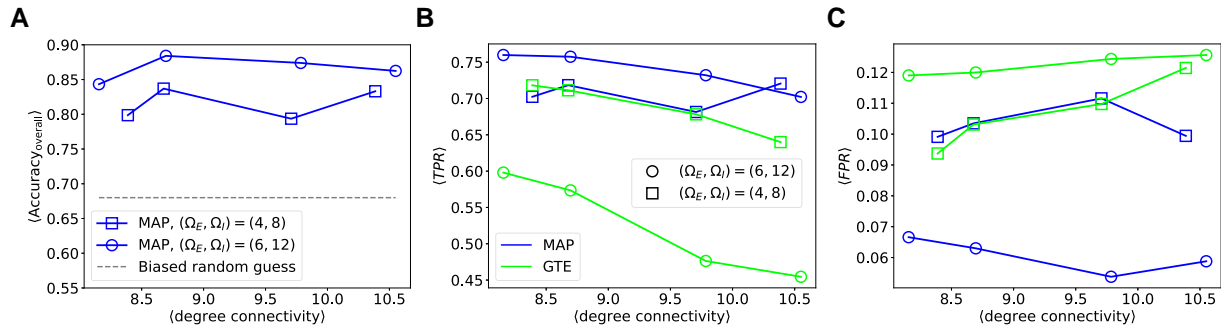
Our GML approach demonstrates strong performance in inferring the neuronal types and link existence as well as the ability for activity prediction. Notably, by comparing the results between MLE and our GML approach, we observe that GML exhibits a gentler slope and deviates more from the perfect prediction manifested by the  $y = x$  line. However, our results still demonstrate a strong overall agreement. This difference may be due to the fact that MLE does not impose any restrictions on the choice of  $J_{ij}$ , aiming to provide the best possible description of the data. On the other hand, in GML, the prior probability  $p(J_{ij} | 0, z_j) \mathcal{N}(0, \epsilon)$  acts as a regularization term and constrain  $J_{ij}$  to have the same sign for each fixed  $j$ . As a result, the inferred  $J_{ij}$  values are lower than those obtained through MLE, leading to an underestimation of  $\mathbf{D}$ . Nonetheless, this approach allows for accurate predictability of neuronal types and effective links, which is more important in neuroscience research.

The above results show that our GML approach performs very well on data generated by the in silico model with patterned substrates. In the [Supplementary Material](#), we study the performance of our GML approach on another in silico homogeneous neuronal network with no patterned substrates. We show that even in a homogeneous network, which is harder for structure inference, as there are fewer constraints for the solution space of  $\mathbf{J}$ , our GML approach still performs very well. The results obtained by the GML for in silico data support the view that it is a highly suitable candidate for analyzing biological neuronal network data.

### Sensitivity analysis of in silico experiments

To further investigate the performance of our algorithms in different scenarios, we performed a sensitivity analysis on in silico neuronal networks where synaptic strengths follow different distributions, varying the degree-connectivity from sparse to dense. Results for networks with patterned substrates are shown below, while results for homogeneous networks are provided in the [Supplementary Material](#).

We examine two different cases where the synaptic strengths in the in silico networks vary. Specifically, we assume that the



**Fig. 3.** Sensitivity analysis of the performance of MAP and GTE on structural inference in *in silico* neuronal networks over stripe-patterned substrates with different synaptic weight distributions, varying over the average degree-connectivity. A) Average overall performance in neuronal-type classification, compared with a biased random guess. B) TPR of identifying effective links. The TPR of GTE is obtained from the ROC curve by fixing the FPR at the same value as MAP. C) FPR of identifying effective links. The FPR of GTE is obtained from the ROC curve by fixing the TPR at the same level as MAP. Results are averaged over five samples for a network of  $N = 156$  neurons.

synaptic strengths  $W_{ij}$  follow a Gaussian distribution with mean  $\Omega_E$  ( $\Omega_I$ ) and variance 0.05, for excitatory (inhibitory) neurons. Starting with a network over stripe-patterned substrates, where there are no connections between neurons (i.e. all axon lengths are initially zero), we gradually increase the degree-connectivity by growing the axon lengths, as detailed in the Methods section. At different connectivity levels, we take snapshots of the network structure to generate neuronal activities. We perform the analysis on an *in silico* neuronal network over stripes patterned substrates, consisting of  $N = 156$  neurons, averaged over five samples. We study two cases: strong synaptic strengths with  $(\Omega_E, \Omega_I) = (6, 12)$  and weak couplings with  $(\Omega_E, \Omega_I) = (4, 8)$ . Due to the high computational complexity of GML, we performed structural inference using MAP only for this sensitivity analysis.

As no existing method can perform neuronal-type classification, we compare our results with those from biased random guessing, as shown in Fig. 3A. The average overall accuracy of neuronal-type classification for both weak and strong synaptic strengths is significantly higher than random guessing, indicating strong performance. For effective connectivity identification, we compare our method with GTE. As discussed, GTE provides an ROC curve, not specific TPR or FPR values. To facilitate comparison, we fix the FPR for GTE at the same level as MAP to obtain the TPR value, and vice versa. Figure 3B and C shows the resulting TPR and FPR values, respectively. For the strong synaptic weights case, our method achieves higher TPR and lower FPR than GTE, outperforming GTE in both metrics. When synaptic weights are weak, our method shows only a slightly higher TPR and slightly lower FPR than GTE. This is due to weaker spiking activity and lower correlations between neurons, making structural inference more challenging. While MAP considers interactions among multiple neurons, which requires more data, GTE focuses on pairwise interactions only, allowing for quicker inference with less data. Nevertheless, our method outperforms GTE in both cases.

To further assess the predictive capability of our method, we generated predicted activities using Monte Carlo simulations based on the structures inferred by MAP and MLE for each sample, and compared them with the true activities by measuring the correlation between predicted and true delayed time covariances  $\mathbf{D}$  as shown in Fig. 4. It is worthwhile noting that correlation is a more appropriate success measure than other metrics like root-mean-squared error, since it measures the alignment between the true and predicted  $\mathbf{D}$ . As shown in Fig. 4, for strong synaptic strengths the average correlation between the predicted and true values of  $\mathbf{D}$  using MAP is higher than that of MLE when the

average degree-connectivity is low. For weak synaptic strengths, MAP performs significantly better than MLE across all degree-connectivities. Although MLE aims to find the optimal configuration for the couplings without restrictions, when the degree-connectivity is low or synaptic strengths are weak, neurons are relatively inactive, making it difficult for MLE to find good solutions. The restrictions in MAP make it easier to find a solution, yielding better activity predictions.

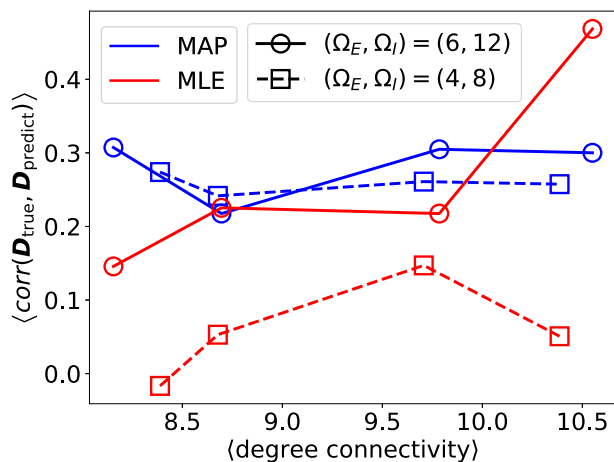
Our sensitivity analysis demonstrates that our algorithms do not only work effectively in single instances but also perform well across various connectivity and synaptic strength scenarios.

#### Experimental validation—in vitro network with patterned substrates

Having validated the efficacy of our GML approach to both synthetic data (see [Supplementary Material](#)) and emulator data, we now apply it to an *in vitro* rat cortical neuronal network grown on topographically patterned substrates. Data consisted of neuronal activity recording from calcium fluorescence imaging, processed to consider regions of interest (ROIs) as nodes in the neuronal network, as shown in Fig. 5A. Details of data acquisition and analysis are provided in the Methods section.

Using this experimental data, with representative traces in Fig. 5B and complete raster plot in Fig. 5C, we infer the effective structure and recreate the neuronal activity for validation. Similar to the *in silico* case, we group ROIs into modules according to the stripes patterning (Fig. 5A); the inferred coupling strengths are visualized in the connectivity matrix shown in Fig. 5D. The background colors indicate which module the source ROIs belong to. Entries colored in red (blue) indicate effective excitatory (inhibitory) connections from source ( $i$ ) to target ( $j$ ) ROIs. We observe dense connectivity between ROIs from the same stripe and sparse connections across stripes. Furthermore, the likelihood of connection decreases as the distance between ROIs increases. This agrees with the biological understanding that long-range connections are rare to minimize wiring cost, so that neurons preferentially connect to their neighbors, as observed in previous studies using similar patterned substrates (32).

While conventional approaches often yield sets of scores, like GTE, they lack a direct measure of estimation quality. Our probabilistic model-based inference, however, allows for the validation of inferred effective structures by reproducing neuronal activity using Monte Carlo simulation. We thus simulate activity using the estimated reconstructed effective network parameters



**Fig. 4.** Sensitivity analysis of the predictability performance on in silico neuronal networks with stripe-patterned substrates using MAP and MLE. The average correlation between predicted and true delayed time covariance  $\mathbf{D}$ , for strong and weak synaptic strengths, as a function of the average degree-connectivity. Results are averaged over five samples for a network of  $N = 156$  neurons.

$\mathbf{J}$  and  $\mathbf{H}$  and compare them with the experimental values, evaluating afterward the equal and delayed time covariance matrices  $\mathbf{C}$  and  $\mathbf{D}$ , as shown in Fig. 5E and F, respectively. Most of the predicted values are aligned with the dashed lines (exact match), meaning a good-quality prediction. However, one can see that for both  $\mathbf{C}$  and  $\mathbf{D}$ , some of the predicted values are very close to zero. This can be attributed to false positive errors generated due to the restrictions imposed on individual ROIs, such as having the same characteristics, e.g. being either excitatory or inhibitory and sharing the same synaptic connection sign. More realistically, within a single ROI, there might be neurons of different types, leading to some connections being erroneously decimated to zero. We anticipate this problem to be mitigated when smaller ROI sizes are considered and longer calcium recordings are used, which will be discussed in the next section.

In general, the GML inference method shows excellent agreement between the inferred effective network and the underlying biological structure, with predicted activities fitting well with the real data.

### Altered in vitro homogeneous network structure by modifying neuronal activity

To validate the algorithm further in a scenario of changing synaptic strengths, we implement experiments that neuronal connectivity is altered through plasticity, which is modulated via chemical stimulation.

Arguably, the simplest way to understand how neuronal network plasticity occurs due to stimulation is to compare the effective network structures of a neuronal network before and after stimulation. We apply our method to an in vitro primary homogeneous neuronal network comparing the situation in control and following exposure to elevated potassium chloride (KCl) stimulation, with the aim of depolarizing neurons and therefore increasing network activity. Neuronal activity data was also collected through calcium imaging, although here each ROI includes one neuron only. The video of the calcium imaging recording is available from (38). The details of data generation are provided in the Methods section.

We infer the effective structure using the neuronal activity data as input, with representative traces in Fig. 6B and complete raster plot in Fig. 6C, and then recreate the activity measures for validation. We first identify the neuronal type (of single neuron ROIs) as excitatory (red) and inhibitory (blue) as shown in Fig. 6A.

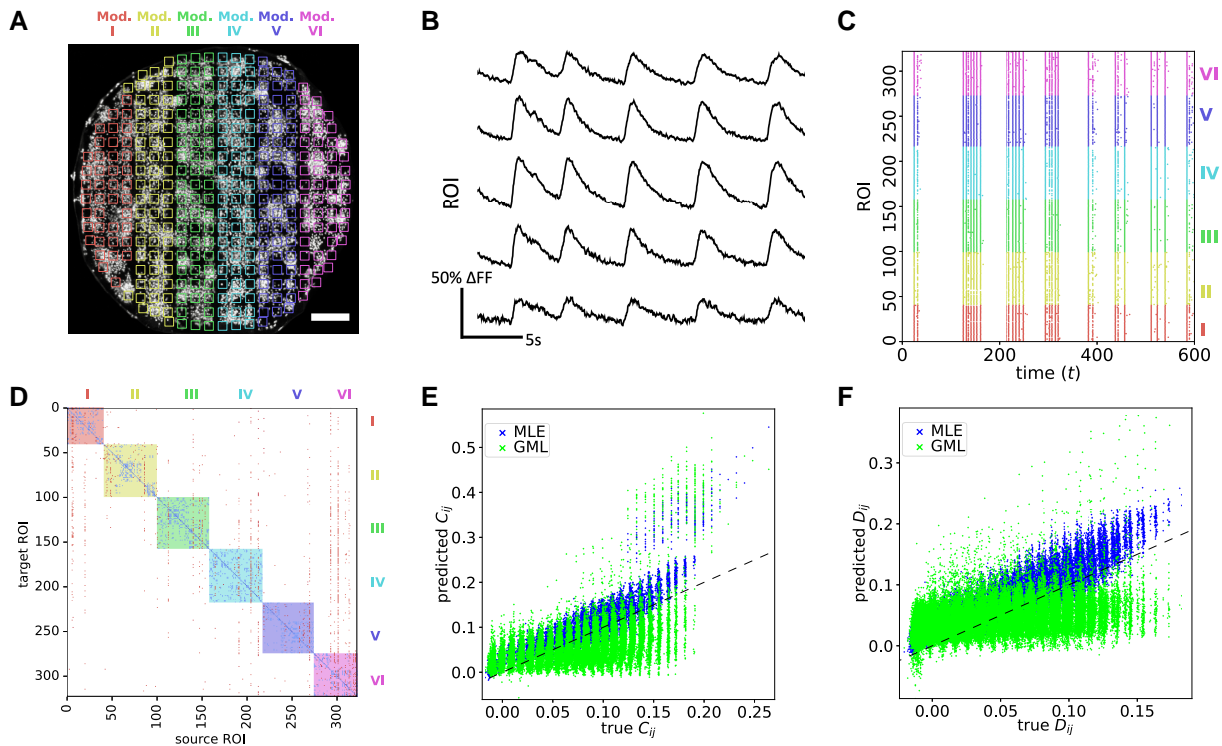
In particular, we simulate activity using the inferred  $\mathbf{J}$  and  $\mathbf{H}$ , then evaluate the equal and delayed time covariance  $\mathbf{C}$  and  $\mathbf{D}$  and compare them with the true covariance values, as shown in Fig. 6D and E, respectively. Similar to the case of the patterned substrate, some predicted values are very close to zero for both  $\mathbf{C}$  and  $\mathbf{D}$ . However, compared with Fig. 5E and F, we see that the predicted values align better with the dashed line, suggesting a more accurate prediction capability. We note that data shown in Fig. 5 and 6 are both acquired using calcium imaging approaches, but the resolution of the experiment in Fig. 6 enables the identification of single neuronal elements. This suggests that by significantly reducing the size of each ROI to consist of a single neuron, one can reduce errors with respect to multiple neurons ROIs.

In this section, we have studied two in vitro experiments providing different scopes. In the experiment with patterned substrates, we have studied a completely extended network; whereas in the experiment with a homogeneous network, we have studied neuronal activity modulated by chemical stimulation. While the samples and the experimental conditions are different, it is interesting to see that both covariance matrices show a much higher value under stimulation as the KCl stimulation affects the activity across the sample. A detailed study comparing the neuronal structure and synaptic weights before and after stimulation is underway and is beyond the scope of this work. Our results demonstrate that our methods do not only work effectively on synthetic in silico data but also perform well on more realistic neuronal networks.

## Discussion

Cortical neuronal network inference has long been an open question in neuroscience and is crucial for understanding the underlying mechanisms and properties of neuronal systems. Neuronal cultures are regarded as a promising living model to investigate a broad spectrum of technological challenges, from biologically inspired AI (14, 20) to efficient design of treatments for neurological disorders (39). Since the structural blueprint of neuronal connections is not easily accessible in a culture, nor their excitatory/inhibitory type, indirect techniques to infer such a blueprint have jumped into the front-line of computational neuroscience.

Here we introduced a novel and probabilistic algorithm based on statistical physics and Bayesian techniques for the effective structural inference of biological neuronal networks from activity data. The algorithm can not only infer the effective synaptic strengths between neurons but, more importantly, can identify the excitatory and inhibitory type of neurons as well as the effective connections between them in a probabilistic way, a capability that no other existing method can achieve. This is used in a principled way from single spontaneous recordings without additional interference to the culture such as stimulation. This capability goes beyond what most existing state-of-the-art methods can offer. Through synthetic, in silico and realistic in vitro experiments, we demonstrate that our algorithm: (i) outperforms existing methods in both synaptic strength inference and effective connections identification; (ii) achieves high accuracy in neuronal-type classification; (iii) exhibits good reproducibility in the inferred structure, justifying the reliability of the algorithm. In the [Supplementary Material](#), we also show that our algorithm



**Fig. 5.** A) The studied rat cortical neuronal network over striped a topographical pattern 2 mm in diameter, grouped into 400 regions of interest (ROIs). ROIs within different stripes are shown as squares with different colors. The scale bar of  $200\mu\text{m}$  is displayed in white in the figure. B) Example in vitro calcium imaging traces of five ROIs. C) The raster plot displays neuronal activity inferred from the calcium traces, with each ROI's color corresponding to the module it belongs to. D) The connectivity matrix of the effective structure inferred from the in vitro neuronal activity, using the proposed GML method. Each entry shows the coupling strength of the connections, a negative (positive) strength represents an inhibitory (excitatory) connection and is colored in blue (red). The background shade of the plot is the module that the ROI belongs to, with mixed colors where ROIs from two different stripes are connected. A strong clustering effect suggests a dense connectivity of ROIs within stripes and sparse connectivity across stripes. E and F) The predicted equal time covariance **C** and delayed time covariance **D** matrices, plotted against the true values evaluated from the data, respectively.

provides good inference results also in the absence of patterned substrates, in both in silico and in vitro studies.

The dynamics and functioning of neuronal networks are in large part determined by their connectivity and their evolving synaptic strengths. As such, the method is expected to have a direct impact on neuroscience, cortical neuron-based computing devices, and many other related biological and medical areas. For instance, studying the changes in the effective structure of neuronal cultures over time can lead to new theoretical understandings of how plasticity takes place in response to stimuli. Additionally, gaining insight into information processing and propagation in neuronal networks could greatly impact the development of artificial neuronal networks and neuromorphic computing, e.g. by understanding the importance of the ratio between excitatory and inhibitory neurons on functionality.

Revealing precise information about effective structure and neuronal types is essential for developing biological machine learning as it helps one to accurately represent the network dynamics, make predictions, and test hypotheses. Most importantly, it is critical for the design of stimulation learning protocols. Interdisciplinary research in this direction is underway.

## Methods

### Preprocessing and optimal time bin determination

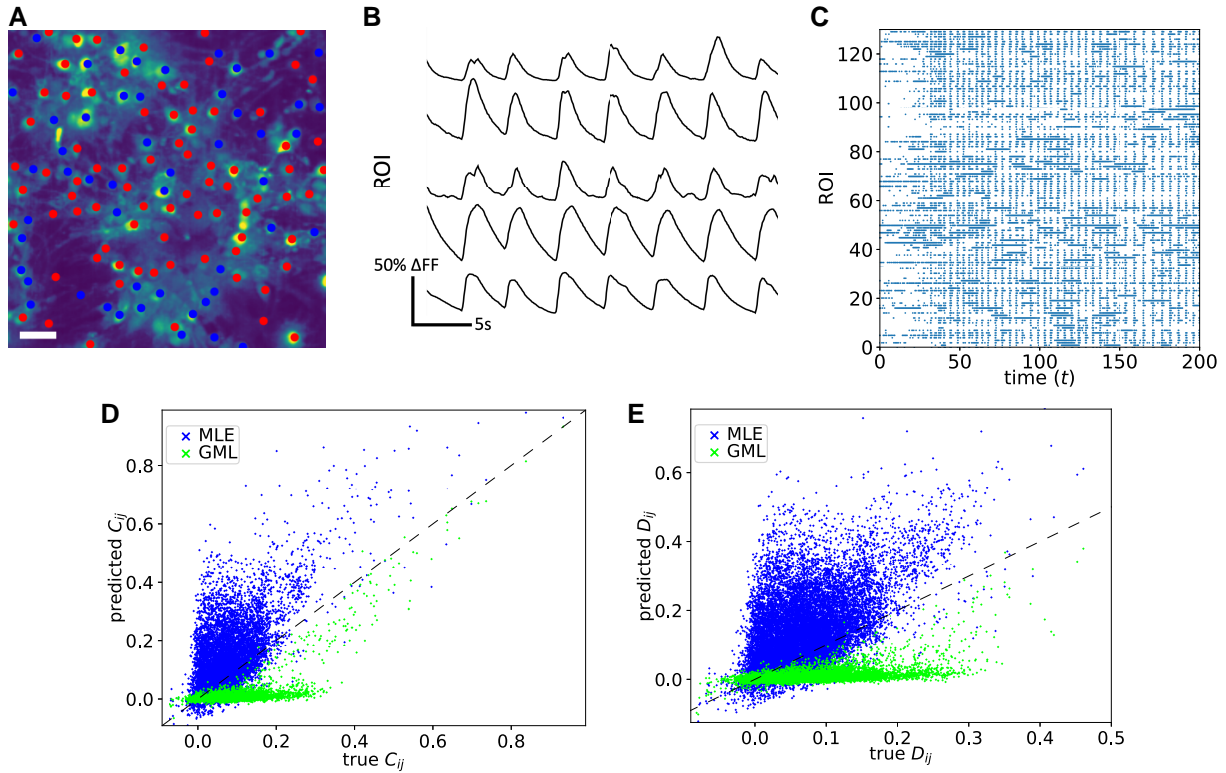
Since the mathematical model we introduce is based on discrete-time steps, one needs to binarize the neuronal activity before

carrying out the inference process. We note that the determination of the time bin  $\tau$  is crucial for inference as the binarized firing patterns can vary significantly with  $\tau$ . For instance, the delayed time covariance between neurons can vanish if  $\tau$  is either too large or small. To address this, we employed an information theory-based method to identify the optimal time bin size  $\tau^*$  that optimizes the total mutual information of the system (26, 27). The optimal bin size  $\tau^*$  is given by

$$\tau^* = \underset{\tau}{\operatorname{argmax}} \left[ \left( \frac{T}{\tau} - 1 \right) \sum_{i \neq j} I_{\tau}(\mathbf{s}_i, \mathbf{s}_j) \right], \quad (7)$$

where  $I_{\tau}(\mathbf{s}_i, \mathbf{s}_j)$  is the mutual information between  $\mathbf{s}_i^t$  and  $\mathbf{s}_j^{t-\tau}$ . The idea of mutual information is straightforward:  $I_{\tau}(\mathbf{s}_i, \mathbf{s}_j)$  measures the discrepancy between the joint probability  $P(\mathbf{s}_i^t, \mathbf{s}_j^{t-\tau})$  and the factorized probability  $P(\mathbf{s}_i^t)P(\mathbf{s}_j^{t-\tau})$  where activity of  $i$  and delayed activity of  $j$  are assumed to be independent. Thus, a higher  $I_{\tau}(\mathbf{s}_i, \mathbf{s}_j)$  suggests stronger correlation between neurons  $i$  and  $j$ . Thus,  $\tau^*$  maximizes the total mutual information, indicating that the neurons are least likely to be independent of each other and more information about their co-dependence can be extracted. Intuitively,  $\tau^*$  can be understood as the average effective reaction time, starting from when a neuron spikes, the spike is transmitted through the synapse and until the target neuron responds. Using  $\tau^*$  evaluated in Eq. 7 to discretize the neuron firing times into distinct time steps that provide the observed data  $\mathbf{s}$ , which is then ready for the inference process.





**Fig. 6.** A) The in vitro primary cortical homogeneous neuronal network analyzed, grouped into 131 regions of interest (ROIs). The red and blue dots correspond to the regions of interest (ROIs) that are classified as excitatory or inhibitory, respectively, where each includes one neuron only. The scale bar of  $25\mu\text{m}$  is displayed in white in the figure. B) Example in vitro calcium imaging traces of five ROIs. C) The raster plot displays neuronal activity inferred from the calcium traces of each ROI. D and E) The predicted equal and delayed time covariance matrices generated by the inferred structure using Monte Carlo simulation, **C** and **D**, respectively, plotted against the true values evaluated from the data.

## Inference algorithm

Unlike the basic kinetic Ising model of statistical physics, inferring the effective structure from neuronal activities faces two major challenges: (i)  $\mathbf{J}$  is not Gaussian with small variance due to the existence of different neuron types; (ii) the connectivity between neurons is affected by multiple factors, including neuronal distance and the patterned substrates (e.g. PDMS stamps). These factors make it difficult to successfully adapt established approaches (28, 29, 40–45) directly to this problem.

To tackle these challenges based on the defined mathematical model, we utilize the generalized maximum-likelihood (GML) (36) technique, also known as evidence approximation and MAP estimation. This combination enables us to jointly infer optimal hyperparameters, latent variables, and the effective network structure in a principled manner. While the detailed derivation is available in the [Supplementary Material](#), we provide a general overview of the method here.

The aim of this algorithm is to infer the effective structure parameters  $\mathbf{J}$  and  $\mathbf{H}$ ; this is supported by the optimal  $\rho$  that maximize the evidence function  $\ln P(\mathbf{s}|\rho)$ . The optimal values of the latent variables also contribute important insight, determining the type of each neuron (excitatory/inhibitory— $\mathbf{z}$ ) and the existence of links ( $\phi$  taking the value  $\{0, 1\}$  for each link). Our approach involves a two-layered EM algorithm for the estimation of hyperparameters and (latent) variables as illustrated in Fig. 7 comprising a “macro” and a “micro” EM algorithms, the latter is being used to perform the M-step of the macro EM algorithm. In the macro E step, the posterior of the effective structure variables,  $\mathbf{J}$  and  $\mathbf{H}$ , is evaluated. To make the algorithm tractable we focus on the most likely values, determined using gradient descent and the derivatives  $\frac{\partial}{\partial \rho} \ln P(\mathbf{s}|\rho)$  and

$\frac{\partial}{\partial \mathbf{H}} \ln P(\mathbf{s}|\rho)$ , given the hyperparameters  $\rho$  found in the macro M step. The macro M step maximizes the expected complete-data log-likelihood of the evidence function with respect to the hyperparameters  $\rho$ , to determine the optimal values  $\rho^*$  through the secondary EM process; the micro EM process iterates between the expectation of the probability distributions of the latent variables  $\phi_{ij}$  and  $z_j$ , and maximization of the hyperparameters  $\rho$ .

Parameters obtained by the macro E step are treated as fixed in the macro M step. In the micro E step, we use the current hyperparameters  $\rho_{\text{old}}$  (the notation old refers to values obtained from the maximization step in the micro EM iteration part) and the current effective structure  $\mathbf{J}^*$  and  $\mathbf{H}^*$  to evaluate the posterior probabilities  $P_{\text{old}}(\phi_{ij})$  and  $P_{\text{old}}(z_j)$  of the latent variables  $\phi_{ij}$  and  $z_j$ . Specifically, the probability distributions of the latent variables are given by

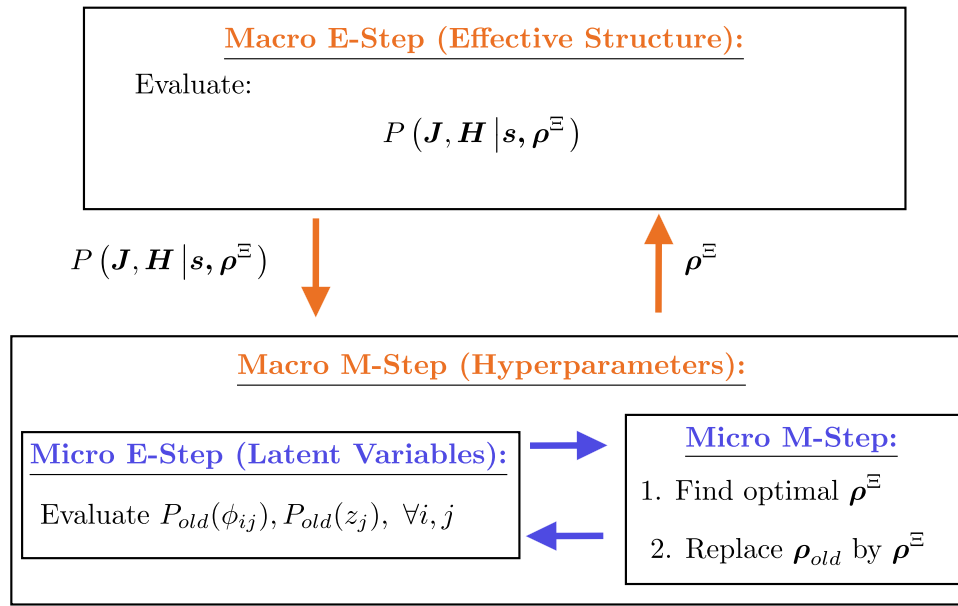
$$P_{\text{old}}(z_j) \propto p(\{\mathbf{J}_{ij}^*\}_i | z_j, \rho) p(z_j | \rho), \quad (8)$$

$$P_{\text{old}}(\phi_{ij}) = \sum_{z_j} P_{\text{old}}(\phi_{ij} | z_j) P_{\text{old}}(z_j). \quad (9)$$

While in the micro M step, one uses  $P_{\text{old}}(\phi_{ij})$  and  $P_{\text{old}}(z_j)$  evaluated in Eqs. 8 and 9 to determine the optimal hyperparameters that maximize the expected complete-data log-likelihood

$$\begin{aligned} \mathcal{Q} = & \ln P(\mathbf{s} | \mathbf{J}, \mathbf{H}) + \ln p(\mathbf{H} | \bar{\rho}) \\ & + \sum_j \sum_{z_j} \sum_{\{\phi_{ij}\}_i} p(\{\phi_{ij}\}_i, z_j | \{\mathbf{J}_{ij}^*\}_i, \rho) \ln p(\{\mathbf{J}_{ij}^*\}_i, \{\phi_{ij}\}_i, z_j | \rho). \end{aligned} \quad (10)$$

In particular, the hyperparameters  $\gamma, \mu_j^+, v_j^+, \mu_j^-, v_j^-, \mu_H, v_H$  are obtained by setting the corresponding derivatives of  $\mathcal{Q}$  to zero, while the decay parameter  $a$  is estimated using gradient descent.



**Fig. 7.** A sketch of the proposed GML-based approximation algorithm. The effective connectivity and the optimal hyperparameters are jointly evaluated by incorporating two applications of the EM algorithm. The posterior distributions of  $\mathbf{J}$  and  $\mathbf{H}$  are evaluated in the macro E step, while the hyperparameters values are evaluated in the macro M step, which has an internal nested EM algorithm of its own, representing the interplay between the latent variable values and their distributions. In the micro EM algorithm: the posterior distributions of  $\phi_{ij}$  and  $z_j$  are evaluated in the micro E step, while the optimal hyperparameters are evaluated in the micro M step, constituting jointly the macro M step that feeds back into the macro E step.

In general, the inference and optimization framework consists of iteratively executing the macro E and M steps until convergence. Within each macro M-step, the hyperparameters are evaluated by iteratively conducting the micro E step and M step until convergence. Finally, one decides on the structure parameters, neuronal type, and link existence by choosing the highest probability states. Notably, for a more sensible starting point, the initial conditions of the hyperparameters  $\rho$  can be evaluated using the inferred  $\mathbf{J}$  and  $\mathbf{H}$  by using MLE. Additionally, it is worth mentioning that if rapid convergence is preferred over absolute accuracy, the number of macro EM iterations can be limited to one, which effectively results in MAP estimation.

## In silico data generation

Emulated neuronal data have been generated with a spiking neuronal network model previously used to model the network growth and activity of biological neuronal cultures (2, 46). The existing network growth model has been adapted to incorporate the effect of inhomogeneous environments on the network connectivity, thus making the emulated data replicate experimental calcium-recorded results on biological neuronal cultures in inhomogeneous environments (32).

Briefly, network growth is modeled by placing  $N$  neurons in a nonoverlapping manner on a surface and modeling axon growth from each neuron by concatenating line segments, in which segment  $i$  is placed with a random angle  $\phi_i = \phi_{i-1} + \sigma_\phi \mathcal{N}(0, 1)$  with respect to the previous segment  $i-1$ . Once an axon segment of neuron  $j$  is placed within a radius  $r_{\text{soma}} \approx 7.5 \mu\text{m}$  of another neuron  $i$ , a connection  $W_{ij}$  is made with a probability  $\alpha$ . The strength of the connection is drawn from a Gaussian distribution with a mean and standard deviation depending on the neuron type. Inhibitory neurons make up 20% of the network, the remainders are excitatory.

Neuronal dynamics is modeled using the Izhikevich model neuron (37), with added synapse dynamics

$$\frac{dP_i}{dt} = -\frac{P_i}{\tau_P} + \beta R_i \delta(v_i - v_{th}) \quad (11)$$

$$\frac{dR_i}{dt} = \frac{1 - R_i}{\tau_R} - \gamma R_i \delta(v_i - v_{th}). \quad (12)$$

The quantity  $\sum_j W_{ji} P_i$  represents the postsynaptic potential induced by the neuron  $i$ , and  $R_i$  is the corresponding synaptic neurotransmitter reserve.

## In vitro data generation—experimental methods

### PDMS topographical reliefs

Topographical patterns were generated by pouring liquid polydimethylsiloxane (PDMS) on specially designed printed circuit board molds shaped as parallel tracks  $300 \mu\text{m}$  wide,  $70 \mu\text{m}$  high and separated by  $200 \mu\text{m}$  (32). PDMS was cured at  $100^\circ\text{C}$  for 2 h, separated from the mold, and perforated with sterile punchers to set 4 PDMS cylinders 2 mm in diameter and 0.5 mm high that contained the inverse topographical pattern of the mold. The 4 cylinders were then evenly distributed on a glass coverslip 13 mm in diameter, autoclaved, and coated with PLL. Flat PDMS substrates were also prepared to investigate the impact of topography.

### Preparation of in vitro neuronal cultures

*Rat primary cultures for patterned networks on PDMS substrates—* Sprague-Dawley rat primary neurons (Charles River Laboratories, France) from embryonic cortices at days 18–19 of development were used in all experiments. Manipulation and dissection of the embryonic cortices were carried out under ethical order B-RP-094/15–7125 of the Animal Experimentation Ethics Committee (CEEA) of the University of Barcelona and in accordance with the regulations of the Generalitat de Catalunya (Spain). Dissection was carried out identically as described in (32). Briefly, cortices were dissected in ice-cold L-15 medium

(Gibco), enriched with 0.6% glucose and 0.5% gentamicin (Sigma-Aldrich). Brain cortices were first isolated from the meninges and then mechanically dissociated by repeated pipetting. The resulting dissociated neural progenitors were plated on a set of precoated Poly-L-Lysine (PLL, 10 mg/mL, Sigma-Aldrich) PDMS topographical substrates in the presence of plating medium, which ensured both the development of neurons and glial cells. A density of a half cortex per 1.3 cm<sup>2</sup> (glass coverslip area) was seeded. This step corresponded to day in vitro (DIV) 0. Two hours after plating, cells were transduced with adeno-associated viruses bearing the genetically encoded calcium fluorescence indicator GCaMP6s under the Synapsin-I promoter, so that only mature neurons expressed the indicator. At DIV 5 the proliferation of glial cells was restricted by incorporating 0.5% FUDR in the culture medium for two more days. From DIV 7 onwards, cells were maintained in a minimum essential medium supplemented with horse serum (Sigma). This medium was changed periodically every 3 days. Cultures were incubated at 37°C, 5% CO<sub>2</sub>, and 95% humidity.

**Mouse primary neuronal cultures for network inference**—13 mm glass coverslips (VWR) were sterilized in 70% Ethanol for 30 min. Coverslips were then transferred to a biological safety cabinet and dried completely before coating for 2 h with 0.02% Poly-L-Ornithine (Sigma). This was washed once with sterile ddH<sub>2</sub>O, and then 20 µg/mL murine laminin was added overnight. P0-P2 C57BL/6 mice were used for the network inference study. All animal procedures were approved by Aston University Bioethics Committee and performed in accordance with the United Kingdom Animals Scientific Procedures act of 1986 and current EU legislation. The 3 Rs, replacement, refinement, and reduction were considered for planning all animal procedures. The experiment was carried out at Aston University. Cortices were dissected in ice-cold HBSS (Gibco) containing 1% Penicillin/streptomycin (P/S). Meninges were removed before dissection of both cortices, which were then each cut into eight pieces for enzymatic digestion. Cortices were transferred into 37°C prewarmed HBSS containing 25 U/mL papain (Sigma), 2 µg/mL DNase (Sigma) and L-Cysteine (Sigma). These were incubated at 37°C for 30 min, gently moving and rocking the tube every 7.5 min. Papain solution was removed, and washed twice for 5 min each, at 37°C in MEM (Gibco) + 10% Horse serum (Sigma) + Glutamax (Gibco) + 1% P/S. Cortices were then mechanically dissociated with glass, fire-polished pipettes, to produce homogenous cell suspension. Cells were passed through a 70 µm cell strainer (Appleton Woods) to remove any remaining meninges or large clusters of cells or debris. Cells were counted, and seeded onto Poly-L-Ornithine/MuLAM coated glass coverslips at a density of 400,000 cells/cm<sup>2</sup>. After 2 h, media was replaced with Neurobasal (Gibco) + 1% B27 (Gibco) + 1% Glutamax (Gibco) + 1% P/S. Cells were fed every 4 days with a half media change.

### Intracellular calcium fluorescence imaging

Rat neuronal cultures shaped as Ø 2 mm PDMS discs allowed the monitoring of the whole network along development. Spontaneous neuronal network activity was recorded using wide-field fluorescence microscopy in combination with the GCaMP6s indicator. Although the networks contained both neurons and glia, only neurons were visualized. Recordings were carried out at DIV 17 for 15 min on a Zeiss Axiovert C25 inverted microscope equipped with a high-speed camera (Hamamatsu Orca Flash 4.0) in combination with an optical zoom. Recordings were carried out at room temperature with the camera software Hokawo 2.10 at 33 frames per second (fps), 8-bit grayscale format, and a size of 1,024 × 1,024 pixels.

For mouse neuronal cultures, cells were loaded with 10 µM Fluo4-AM in DMSO (Invitrogen) for 40 min at 37°C. Coverslips were then transferred onto an upright Nikon FN1 microscope and images were acquired using a Crest Optics XLight V3 spinning disk confocal and a Teledyne Photometrics Kinetix high-speed camera, and in an area of 300 × 300 µm<sup>2</sup>. The setup was controlled through Micro-Manager (47). Cultures were perfused with 37°C heated Artificial CSF (aCSF) as a control, and an increase of 2.5 mM KCl to increase baseline activity. Cultures were settled for 5 min before recording commenced at 10 Hz for 10 min. aCSF solution containing the following (in mM): NaCl 120, NaHCO<sub>3</sub> 25, KCl 2, KH<sub>2</sub>PO<sub>4</sub> 1.25, MgSO<sub>4</sub> 1, and CaCl<sub>2</sub> 2. aCSF chemicals were obtained from Sigma-Aldrich (48).

### Data analysis

Calcium fluorescence recordings were analyzed with the NETCAL software (49, 50) run in MATLAB in combination with custom-made packages. To analyze the data, and as described in (32), Regions of Interest (ROIs) were first laid on the area covered by each culture. For rat primary cultures, ROIs were shaped as a 20 × 20 grid centered at the culture and extending its entire 2 mm circular shape, while for mouse primary cultures ROIs corresponded to individual neurons. Next, the average fluorescence trace within each ROI was extracted as a function of time, corrected from drifts, and normalized. Sharp peaks in the fluorescence signals revealed neuronal activations, which were detected using the Schmitt trigger method, finally leading to a binarized time series of neuronal activity in which “1” indicated the presence of neuronal activity and “0” its absence.

## Immunocytochemistry

This technique was used to identify the position of neuronal cell bodies in culture and extract their fluorescence trace with precision. Neuronal cultures were fixed for 20 min with 4% PFA (Sigma) at room temperature. After washing with PBS, the samples were incubated with a blocking solution containing 0.03% Triton (Sigma) and 5% normal donkey serum (Jackson ImmunoResearch) in PBS for 45 min at room temperature. To visualize the neuronal nuclei, the samples were incubated with primary antibodies, against the neuronal marker NeuN (M1406, Sigma), diluted in blocking solution, and incubated overnight at 4°C. Cy3-conjugated secondary antibody against rabbit (711-165-152, Jackson ImmunoResearch) was diluted in blocking solution and incubated for 90 min at room temperature. Then, cultures were rinsed with PBS and mounted using DAPI-fluoromount-G (ShouthernBiotech). Immunocytochemical images were acquired on a Zeiss confocal microscope.

## Note

<sup>a</sup> This method is accurate only when coupling strengths are relatively small or when the activation function is smooth.

## Supplementary Material

Supplementary material is available at PNAS Nexus online.

## Funding

This research is supported by the European Union Horizon 2020 research and innovation program under Grant No. 964877 (project NEU-CHiP). J.S. also acknowledges support from grants

PID2022-137713NB-C22 and PLEC2022-009401, funded by MCIU/AEI/10.13039/501100011033 and by ERDF/EU, and by the Generalitat de Catalunya under grant 2021-SGR-00450. The authors acknowledge the support of Aston University Biomedical Facility and Aston Institute for Membrane Excellence (AIME) for the purpose of providing infrastructure support within the College of Health and Life Sciences.

## Author Contributions

H.F.P. and D.S. designed research, derived algorithm, created synthetic data, and analyzed all data. A.C.H., D.R.J., H.R.P., E.J.H., and J.S. performed in vitro experiments; A.M.H. and J.S. performed in silicon experiments. H.F.P., A.M.H., A.C.H., D.R.J., H.R.P., E.J.H., J.S., and D.S. wrote the paper.

## Preprints

This manuscript is available as a preprint at: <https://arxiv.org/abs/2402.18788>.

## Data Availability

All data presented in this paper are available from <https://doi.org/10.17036/researchdata.aston.ac.uk.00000635>. This includes the Python file containing the derived algorithm, as well as both the in silico and in vitro neuronal firing data being studied.

## References

- Maeda E, Robinson HP, Kawana A. 1995. The mechanisms of generation and propagation of synchronized bursting in developing networks of cortical neurons. *J Neurosci.* 15(10):6834–6845.
- Orlandi JG, Soriano J, Alvarez-Lacalle E, Teller S, Casademunt J. 2013. Noise focusing and the emergence of coherent activity in neuronal cultures. *Nat Phys.* 9(9):582–590.
- Eytan D, Marom S. 2006. Dynamics and effective topology underlying synchronization in networks of cortical neurons. *J Neurosci.* 26(33):8465–8476.
- Wagenaar DA, Pine J, Potter SM. 2006. An extremely rich repertoire of bursting patterns during the development of cortical cultures. *BMC Neurosci.* 7(1):11. doi:10.1186/1471-2202-7-11.
- Cohen E, Ivenshitz M, Amor-Baroukh V, Greenberger V, Segal M. 2008. Determinants of spontaneous activity in networks of cultured hippocampus. *Brain Res.* 1235:21–30. (Brain Oscillations in Cognition and Cognitive Disorders). doi:10.1016/j.brainres.2008.06.022.
- Pasquale V, Massobrio P, Bologna LL, Chiappalone M, Martinoia S. 2008. Self-organization and neuronal avalanches in networks of dissociated cortical neurons. *Neuroscience.* 153(4):1354–1369.
- Tetzlaff C, Okujeni S, Egert U, Wörgötter F, Butz M. 2010. Self-organized criticality in developing neuronal networks. *PLoS Comput Biol.* 6(12):1–18. doi:10.1371/journal.pcbi.1001013.
- Buzsáki G, Draguhn A. 2004. Neuronal oscillations in cortical networks. *Science.* 304(5679):1926–1929. doi:10.1126/science.1099745.
- Berger H. 1929. Über das elektroencephalogramm des menschen. *Archiv für Psychiatrie und Nervenkrankheiten.* 87(1):527–570.
- Logothetis NK. 2008. What we can do and what we cannot do with fMRI. *Nature.* 453(7197):869–878.
- Maccione A, et al. 2010. Experimental investigation on spontaneously active hippocampal cultures recorded by means of high-density meas: analysis of the spatial resolution effects. *Front Neuroeng.* 3:1294. doi:10.3389/fneng.2010.00004.
- Grienberger C, Konnerth A. 2012. Imaging calcium in neurons. *Neuron.* 73(5):862–885.
- Kim TH, Schnitzer MJ. 2022. Fluorescence imaging of large-scale neural ensemble dynamics. *Cell.* 185(1):9–41.
- Sumi T, et al. 2023. Biological neurons act as generalization filters in reservoir computing. *Proc Natl Acad Sci U S A.* 120(25):e2217008120.
- Moriya S, Yamamoto H, Hirano-Iwata A, Kubota S, Sato S. Quantitative analysis of dynamical complexity in cultured neuronal network models for reservoir computing applications. In: 2019 *International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2019. p. 1–6.
- Zanini G, Parodi G, Chiappalone M, Martinoia S. 2023 Dec. Investigating the reliability of the evoked response in human iPSCs-derived neuronal networks coupled to micro-electrode arrays. *APL Bioeng.* 7(4):046121. doi:10.1063/5.0174227.
- Parodi G, Brofiga M, Pastore VP, Chiappalone M, Martinoia S. 2023. Deepening the role of excitation/inhibition balance in human iPSCs-derived neuronal networks coupled to MEAs during long-term development. *J Neural Eng.* 20(5):056011.
- Strubell E, Ganesh A, McCallum A. Energy and policy considerations for deep learning in NLP. In: Korhonen A, Traum D, Márquez L, editors. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Florence, Italy, 2019 Jul. p. 3645–3650. doi:10.18653/v1/P19-1355.
- Cai H, et al. 2023. Brain organoid computing for artificial intelligence. *bioRxiv* 530502. <https://doi.org/10.1101/2023.02.28.530502>, pages 2023–02, preprint: not peer reviewed.
- Kagan BJ, et al. 2022. In vitro neurons learn and exhibit sentence when embodied in a simulated game-world. *Neuron.* 110(23):3952–3969.
- Orlandi JG, Stetter O, Soriano J, Geisel T, Battaglia D. 2014. Transfer entropy reconstruction and labeling of neuronal connections from simulated calcium imaging. *PLoS One.* 9(6):e98842.
- Friston KJ. 2011. Functional and effective connectivity: a review. *Brain Connect.* 1(1):13–36.
- Olsen VK, Whitlock JR, Roudi Y. 2024 May. The quality and complexity of pairwise maximum entropy models for large cortical populations. *PLoS Comput Biol.* 20(5):1–30. doi:10.1371/journal.pcbi.1012074.
- Song W, Cajigas I, Brown EN, Giszter SF. 2015. Adaptation to elastic loads and BMI robot controls during rat locomotion examined with point-process GLMs. *Front Syst Neurosci.* 9:62. doi:10.3389/fnsys.2015.00062.
- Ito S, et al. 2011. Extending transfer entropy improves identification of effective connectivity in a spiking cortical network model. *PLoS One.* 6(11):e27431.
- Terada Y, Obuchi T, Isomura T, Kabashima Y. Objective and efficient inference for couplings in neuronal networks. In: Bengio S, Wallach H, Larochelle H, Grauman K, Cesa-Bianchi N, Garnett R, editors. *Advances in neural information processing systems*. Vol. 31. Curran Associates, Inc, 2018. [https://proceedings.neurips.cc/paper\\_files/paper/2018/file/03cf87174debacc689c90c34577b82f-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2018/file/03cf87174debacc689c90c34577b82f-Paper.pdf).
- Terada Y, Obuchi T, Isomura T, Kabashima Y. 2020. Inferring neuronal couplings from spiking data using a systematic procedure with a statistical criterion. *Neural Comput.* 32(11):2187–2211.
- Roudi Y, Hertz J. 2011. Mean field theory for nonequilibrium network reconstruction. *Phys Rev Lett.* 106(4):048702.



- 29 Mézard M, Sakellariou J. 2011. Exact mean-field inference in asymmetric kinetic Ising systems. *J Stat Mech.* 2011(07):L07001.
- 30 Lombardi F, Pepić S, Shriki O, Tkačik G, De Martino D. 2023. Statistical modeling of adaptive neural networks explains co-existence of avalanches and oscillations in resting human brain. *Nat Comput Sci.* 3(3):254–263. doi:[10.1038/s43588-023-00410-9](https://doi.org/10.1038/s43588-023-00410-9).
- 31 Dahmen D, Bos H, Helias M. 2016. Correlated fluctuations in strongly coupled binary networks beyond equilibrium. *Phys Rev X.* 6(3):031024.
- 32 Montalá-Flaquer M, et al. 2022. Rich dynamics and functional organization on topographically designed neuronal networks in vitro. *iScience.* 25(12):105680. doi:[10.1016/j.isci.2022.105680](https://doi.org/10.1016/j.isci.2022.105680).
- 33 Fredrickson GH, Andersen HC. 1984. Kinetic Ising model of the glass transition. *Phys Rev Lett.* 53(13):1244.
- 34 Battaglia D, Witt A, Wolf F, Geisel T. 2012 Mar. Dynamic effective connectivity of inter-areal brain circuits. *PLoS Comput Biol.* 8(3):1–20. doi:[10.1371/journal.pcbi.1002438](https://doi.org/10.1371/journal.pcbi.1002438).
- 35 Yamamoto H, et al. 2023. Modular architecture facilitates noise-driven control of synchrony in neuronal networks. *Sci Adv.* 9(34):eade1755.
- 36 Bishop CM, Nasrabadi NM. 2006. *Pattern recognition and machine learning*. Vol. 4. Springer.
- 37 Izhikevich EM. 2003. Simple model of spiking neurons. *IEEE Trans Neural Netw.* 14(6):1569–1572. doi:[10.1109/TNN.2003.820440](https://doi.org/10.1109/TNN.2003.820440).
- 38 Po HF, et al. 2024. Research Data - Inferring Structure of Cortical Neuronal Networks from Firing Data: A Statistical Physics Approach. [Dataset] Aston University. doi: [10.17036/researchdata.aston.ac.uk.00000635](https://doi.org/10.17036/researchdata.aston.ac.uk.00000635)
- 39 Carola G, et al. 2021. Parkinson's disease patient-specific neuronal networks carrying the LRRK2 G2019S mutation unveil early functional alterations that predate neurodegeneration. *NPJ Parkinsons Dis.* 7(1):1–14. doi:[10.1038/s41531-021-00198-3](https://doi.org/10.1038/s41531-021-00198-3).
- 40 Decelle A, Zhang P. 2015. Inference of the sparse kinetic Ising model using the decimation method. *Phys Rev E.* 91(5):052136.
- 41 Aguilera M, Moosavi SA, Shimazaki H. 2021. A unifying framework for mean-field theories of asymmetric kinetic Ising systems. *Nat Commun.* 12(1):1197.
- 42 Zeng H-L, Aurell E, Alava M, Mahmoudi H. 2011. Network inference using asynchronously updated kinetic Ising model. *Phys Rev E.* 83(4):041135.
- 43 Zhang P. 2012. Inference of kinetic Ising model on sparse graphs. *J Stat Phys.* 148:502–512.
- 44 Huang H, Kabashima Y. 2014. Dynamics of asymmetric kinetic Ising systems revisited. *J Stat Mech.* 2014(5):P05020.
- 45 Kappen HJ, Spanjers JJ. 2000. Mean field theory for asymmetric neural networks. *Phys Rev E.* 61(5):5658.
- 46 Alvarez-Lacalle E, Moses E. 2009. Slow and fast pulses in 1-D cultures of excitatory neurons. *J Comput Neurosci.* 26(3):475–493. doi:[10.1007/s10827-008-0123-5](https://doi.org/10.1007/s10827-008-0123-5).
- 47 Edelstein A, Amodaj N, Hoover K, Vale R, Stuurman N. 2010. Computer control of microscopes using  $\mu$ manager. *Curr Protoc Mol Biol.* 92(1):14.20.1–14.20.17. doi:[10.1002/0471142727.mb1420s92](https://doi.org/10.1002/0471142727.mb1420s92).
- 48 Butcher JB, et al. 2022. A requirement for astrocyte IP<sub>3</sub>R2 signaling for whisker experience-dependent depression and homeostatic upregulation in the mouse barrel cortex. *Front Cell Neurosci.* 16:905285. doi:[10.3389/fncel.2022.905285](https://doi.org/10.3389/fncel.2022.905285)
- 49 Orlandi JG, et al. 2017. NETCAL: an interactive platform for large-scale, NETWORK and population dynamics analysis of CALcium imaging recordings. Zenodo.
- 50 Orlandi R. 2023. Netcal. <https://github.com/orlandi/netcal>.