

# Directed effective connectivity and synaptic weights of in vitro neuronal cultures revealed from high-density multielectrode array recordings

Chumin Sun<sup>1</sup>, K.C. Lin<sup>1</sup>, Yu-Ting Huang<sup>2,3</sup>, Emily S.C. Ching<sup>1\*</sup>, Pik-Yin Lai<sup>2</sup>, C.K. Chan<sup>3</sup>

**1** Department of Physics, The Chinese University of Hong Kong, Shatin, Hong Kong

**2** Dept. of Physics and Center for Complex Systems, National Central University, Chungli, Taiwan 320, ROC

**3** Institute of Physics, Academia Sinica, Taipei, Taiwan 115, ROC

\* ching@phy.cuhk.edu.hk

## Abstract

Studying connectivity of neuronal cultures can provide insights for understanding brain networks but it is challenging to reveal neuronal connectivity from measurements. We apply a novel method that uses a theoretical relation between the time-lagged cross-covariance and the equal-time cross-covariance to reveal directed effective connectivity and synaptic weights of cortical neuron cultures at different days in vitro from multielectrode array recordings. Using a stochastic leaky-integrate-and-fire model, we show that the simulated spiking activity of the reconstructed networks can well capture the measured network bursts. The neuronal networks are found to be highly nonrandom with an over-representation of bidirectionally connections as compared to a random network of the same connection probability, with the fraction of inhibitory nodes comparable to the measured fractions of inhibitory neurons in various cortical regions in monkey, and have small-world topology with basic network measures comparable to those of the nematode *C. elegans* chemical synaptic network. Our analyses further reveal that (i) the excitatory and inhibitory incoming degrees have bimodal distributions the excitatory and inhibitory incoming degrees have bimodal distributions, which are that distributions that have been indicated to be optimal against both random failures and attacks in undirected networks; (ii) the distribution of the physical length of excitatory incoming links has two peaks indicating that excitatory signal is transmitted at two spatial scales, one localized to nearest nodes and the other spatially extended to nodes millimeters away, and the shortest links are mostly excitatory towards excitatory nodes and have larger synaptic weights on average; (iii) the average incoming and outgoing synaptic strength is non-Gaussian with long tails and, in particular, the distribution of outgoing synaptic strength of excitatory nodes with excitatory incoming synaptic strength is lognormal, similar to the measured excitatory postsynaptic potential in rat cortex.

## Author summary

To understand how the brain processes signal and carries out its function, it is useful to know the connectivity of the underlying neuronal circuits. For large-scale neuronal networks, it is difficult to measure connectivity directly using electron microscopy

techniques and methods that can estimate connectivity from electrophysiological recordings are thus highly desirable. Existing methods focus mainly on estimating functional connectivity, which is defined by statistical dependencies between neuronal activities but the relevant direct causal interactions are captured by effective connectivity. Here we apply a novel covariance-relation based method to estimate the directed effective connectivity and synaptic weights of cortical neuron cultures from recordings of multielectrode array of over 4000 electrodes taken at different days in vitro. The neuronal networks are found to be nonrandom, small-world, excitation/inhibition balanced as measured in monkey cortex, and with feeder hubs. Our analyses further suggest some form of specialisation of nodes in receiving excitatory and inhibitory signals and the transmission of excitatory signals at two spatial scales, one localized to nearest nodes and the other spatially extended to nodes millimeters away, and reveal that the distributions of the average incoming and outgoing synaptic strength are skewed with long tails.

## Introduction

Unravelling synaptic connectivity and understanding the relationship between connectivity and dynamics and between network structure and function of neuronal circuits are actively pursued topics in neuroscience [1, 2]. In vitro neuronal cultures serve as simple yet useful experimental model systems for studying the relationship between connectivity and the rich dynamics observed [3–5]. One common technique used to record the activity of neurons in cultures is direct measurement of the electrical signals of neurons using multielectrode array (MEA) [6]. It is thus of great interest to estimate connectivity from MEA recordings. There is a distinction between functional and effective connectivity [7]. Functional connectivity is defined by statistical dependencies between the activities of neurons and is usually inferred based on the cross-covariance or cross-correlation of spike trains with the spikes first being detected from the recorded electrical signals of the electrodes of the MEA [8–10]. As statistical dependencies can arise from indirect interactions, functional connectivity does not necessarily reflect the direct causal interactions among the neurons. Direct causal interaction is referred to as effective connectivity, and is more relevant for studying the relationship between connectivity and dynamics and between network structure and function. Effective connectivity is used to be inferred by using some specific model of interactions and, as a result, it is sometimes said to be model-dependent. Reconstructing functional connectivity from measurements is already highly challenging and nontrivial especially for large-scale MEA with several thousands of electrodes. Existing methods thus focus predominantly on estimating functional connectivity.

For a class of networked systems having generic stationary dynamics modelled by a set of stochastic differential equations with directional interactions, it has been shown that the equal-time cross-covariance of the dynamics of the nodes does not carry sufficient information to recover the effective connectivity [11, 12], and the effective connectivity matrix can instead be extracted using the theoretical relation between the time-lagged cross-covariance and the equal-time cross-covariance [13]. Similar result has been found for systems with discrete-time dynamics [14]. This covariance-relation based method thus infers effective connectivity for a generic class of models and not only for a specific model. It has been validated by numerical simulations for weighted directed random and scale-free networks with different nonlinear dynamics and directional coupling including particularly the FitzHugh-Nagumo dynamics [15] with directional synaptic-like coupling [13, 16].

In this paper, we show how we adopt this covariance-relation based method to estimate directed effective connectivity of cultures of cortices of rat embryos from

recordings of large-scale MEA of over 4000 electrodes taken at different times, ranging from 11 to 66 days in vitro. This method makes use of the MEA recordings directly without the need to first detect spikes from the recordings. In addition, it can detect unidirectional and bidirectional connections equally well and further allows one to estimate the synaptic weights of the links, both excitatory and inhibitory. For cross-correlation based methods that estimate functional connectivity, inhibitory links pose additional challenges [10] and bidirectional connections are difficult to detect. As the validity of the covariance-relation based method has already been shown using numerical data, we focused to test how well the estimated directed effective connectivity can reproduce the measured spiking activity using a stochastic leaky-integrate-and-fire model. We carried out analyses of the reconstructed networks studying the basic network features and the distributions of incoming and outgoing degrees as well as the average incoming and outgoing synaptic strength.

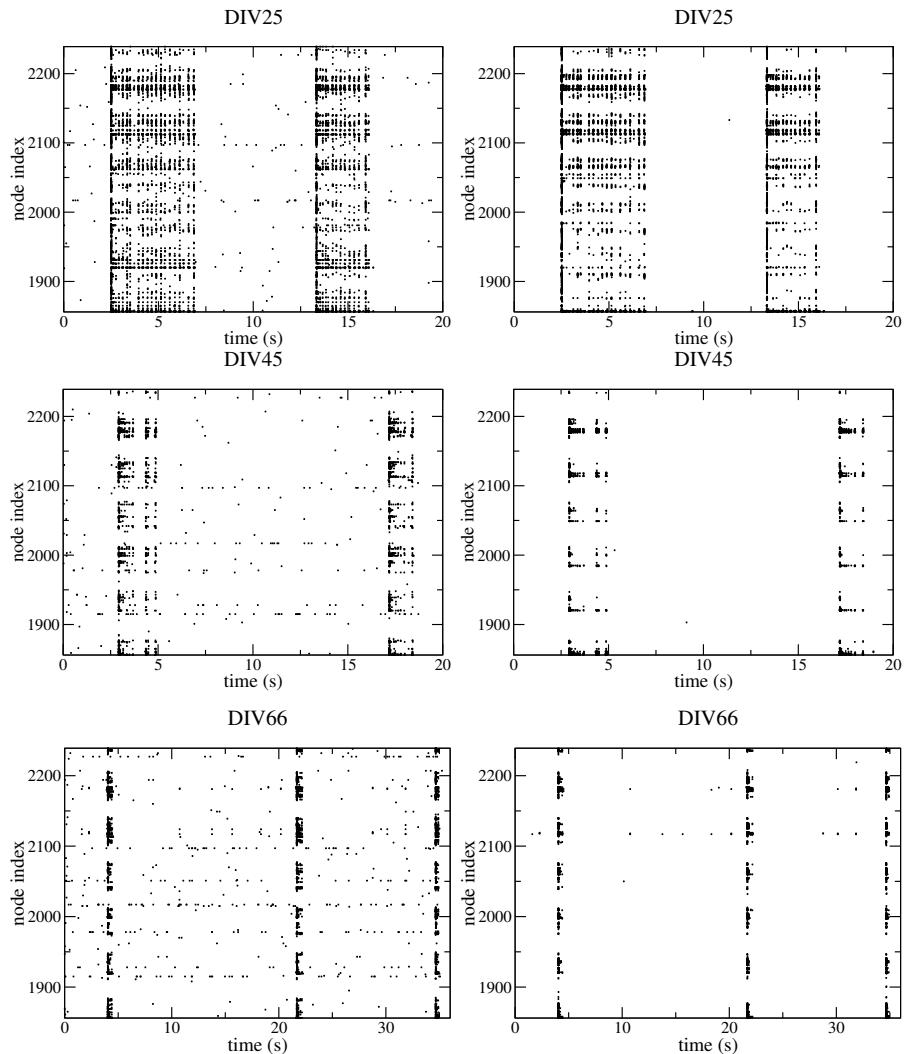
## Results

Cortices from embryonic day 17 (E17) Wistar rats were dissected and a total of about  $6 \times 10^4$  cells were plated on a 6mm by 6mm working area of a high-density MEA probe. The MEA probe has 4096 electrodes arranged on a 64 by 64 square grid in the central 2.67mm by 2.67mm active area. We recorded spontaneous activity at a sampling frequency of 7.06kHz at eight different Days In Vitro (DIV): 11, 22, 25, 33, 45, 52, 59 and 66. The electrode in the upper left corner was reserved for calibration purpose so 4095 electrodes recorded data. As in earlier work [8], we treated each electrode as a node of the neuronal network and estimated the  $4095 \times 4095$  directed and weighted effective connectivity matrix for all 8 different DIVs, referred to as DIV11, DIV22, DIV25, DIV33, DIV45, DIV52, DIV59 and DIV66, from the relation between the time-lagged covariance and equal-time covariance (c.f. Materials and methods). The validity of this covariance-relation based method in recovering the directed and weighted effective connectivity for random and scale-free networks with a generic model of dynamics has been verified using numerical simulations [13, 16]. In this work, we tested how well the estimated directed and weighted effective connectivity matrix can reproduce the measured spiking activity. Then we analysed the basic features of the reconstructed neuronal networks and calculated the distributions of degree and synaptic strength.

### Testing how well the estimated directed effective connectivity can reproduce the measured spiking activity

We applied the Precision Timing Spike Detection (PTSD) algorithm [17] to the MEA recordings to obtain the measured spiking activity. We used the stochastic leaky integrate-and-fire (LIF) model (c.f. Materials and methods for the computer model) of 4095 neurons coupled according to the reconstructed directed and weighted effective connectivity matrix to simulate the spiking activity. In the LIF model, which is a simple, intuitive and most widely used spiking neuron model [18], the membrane potential for each neuron changes in time due to the inputs from its pre-synaptic neurons and when it reaches a threshold, the neuron generates an action potential or a spike and the membrane potential is reset to the resting state. We added a noise to the usual LIF model [19] acting on each neuron to drive the spontaneous spiking activity of interest. For each neuron, we simulated the spiking activity using the measured spiking activity of all its pre-synaptic neurons as inputs. We carried out the simulations for DIV25, DIV45, and DIV66 for a time window of observation  $T_{obs}$  ranging between 18.9s to 35.9s. In Fig 1, we compared the raster plots of the simulated spiking activity with the measured spiking activity. For a close-up comparison, the plots focus on the 384

nodes corresponding to the electrodes in the 6 middle rows of the 64 by 64 MEA. An outstanding feature of the measured spiking activity is the collective bursts of activity by almost the entire network. The duration of these network bursts decreases as the age of the cultures increases from 25 to 66 DIV. The simulated spiking activity of the reconstructed networks captures such network bursts well.



**Fig 1.** Comparison of the measured spiking activity (left panel) with the simulated spiking activity using the reconstructed network (right panel) for DIV25, DIV45 and DIV66.

A significant fraction of nodes participated in the network bursts but some nodes did not. We denoted nodes that have no measured spiking activity in  $T_{obs}$  as null-nodes and checked how well the simulated results managed to retrieve these null-nodes. As a control, we randomly shuffled the reconstructed network with the distributions of the incoming and outgoing degrees kept intact and assigned the synaptic weights randomly from a Gaussian distribution of zero mean and unit variance and repeated the simulations. We compared how well the simulated results from the reconstructed networks and the control networks managed to retrieve the null-nodes by calculating the sensitivity, which measures the fraction of null-nodes that is correctly retrieved, and the

precision, which measures the fraction of nodes simulated to be null-nodes are indeed measured null-nodes. Denote the number of null-nodes that have null or nonzero simulated spiking activity by  $N_{TP}$  or  $N_{FN}$  and the number of non null-nodes with null or nonzero simulated spiking activity by  $N_{FP}$  or  $N_{TN}$ . The sensitivity (SEN) and precision, which is also known as the positive predictive value (PPV) are defined by

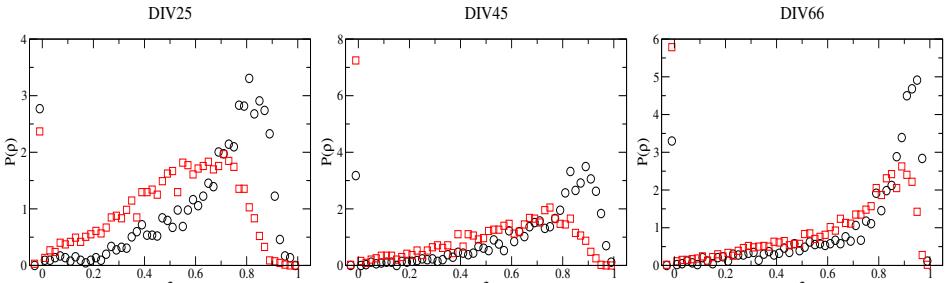
$$\text{SEN} = \frac{N_{TP}}{N_{TP} + N_{FN}}, \quad \text{PPV} = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (1)$$

The results obtained are presented in Table 1. Both SEN and PPV are much higher for the reconstructed networks than for the control randomly-shuffled networks. Thus the reconstructed networks outperform the control randomly reshuffled networks in retrieving the null nodes.

DIV	Network	$N_{TP}$	$N_{FN}$	$N_{FP}$	$N_{TN}$	SEN	PPV	$f$
25	reconstructed	222	370	236	3267	0.38	0.48	0.32
	control	39	553	146	3357	0.066	0.21	0.059
45	reconstructed	1171	453	755	1716	0.72	0.61	0.47
	control	347	1277	297	2174	0.21	0.54	0.14
66	reconstructed	781	490	656	2168	0.61	0.54	0.58
	control	97	1174	119	2705	0.076	0.45	0.35

**Table 1.** The sensitivity and precision or positive predictive value of the reconstructed and control randomly-shuffled networks in retrieving null-nodes.  $f$  is the fraction of the  $N_{TN}$  nodes that have correlation coefficient  $\rho \geq 0.8$ .

We further studied the correlation of the measured and simulated spiking activity at tens of millisecond timescale. We calculated the Pearson correlation coefficient  $\rho$  between the measured and simulated spiking activity using a sliding window of  $w = 28.4$  ms for each of the  $N_{TN}$  nodes with non-zero measured and simulated spiking activity. The distributions of the correlation coefficient  $P(\rho)$  are shown in Fig 2. The fraction  $f$  of these  $N_{TN}$  nodes with  $\rho \geq 0.8$  is significantly larger for the simulated results using the reconstructed networks (see Table 1) than those using the control randomly shuffled networks.



**Fig 2.** Distributions of the correlation coefficient  $P(\rho)$  for all the  $N_{TN}$  nodes for DIV25, DIV45 and DIV66. Black circles are results for the simulations using the reconstructed network and red squares are results for simulations using the control randomly-shuffled network.

All these results further support the validity of the covariance-relation based method in estimating directed effective connectivity. In the next subsections, we show results of the basic features and the distributions of degrees and synaptic strength of the reconstructed neuronal networks.

## Basic network features

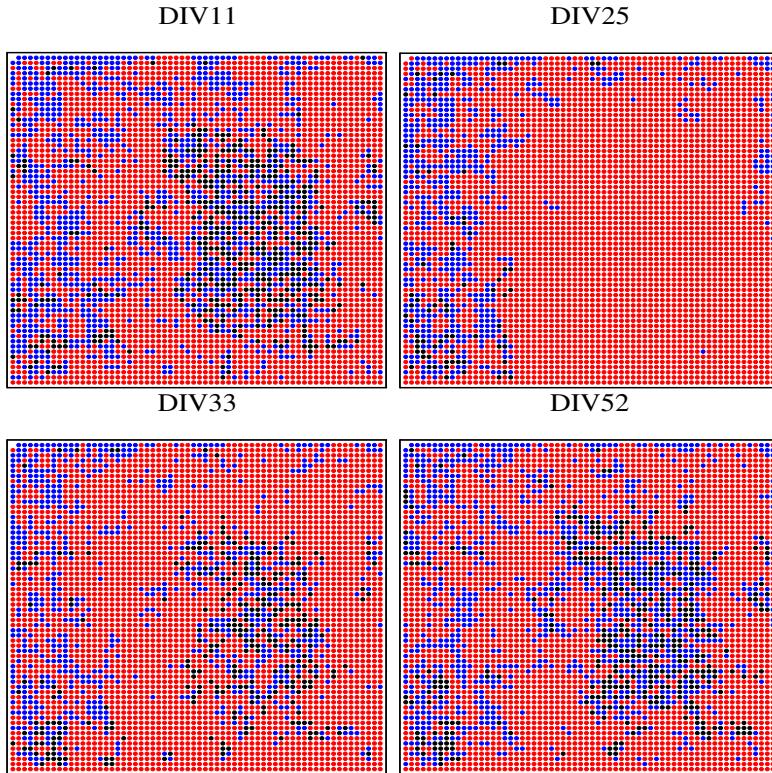
We calculated several basic features of the neuronal networks for all the 8 DIVs (see Table 2). The connection probability  $p$  of a network of  $N$  nodes with  $N_L$  links is defined by  $p = N_L/[N(N - 1)]$ . We found that  $p$  ranges from 1.1-1.9%, confirming that the neuronal networks are sparse. This value of  $p$  is comparable to that of the chemical synapse network of *C. elegans* [20]. Most of the connections are unidirectional as expected. For a random network of the same number of nodes and the same connection probability  $p$ , the expected number of bidirectionally connected pairs is given by  $N(N - 1)p^2/2$ . The number of bidirectionally connected pairs is indeed comparable to this expected number for the control randomly shuffled networks but exceeds four times this expected number for the reconstructed networks. This observation of an over-representation of bidirectional connections as compared to a random network in the reconstructed neuronal networks echoes previous reports for local regions of the rat cortex [21–24]. Thus neuronal networks are highly nonrandom. As a comparison, we calculated the number of bidirectional connections in the functional connectivity estimated by using the cross-covariance based methods [10] for DIV25, DIV45 and DIV66 and found none.

	DIV11	DIV22	DIV25	DIV33	DIV45	DIV52	DIV59	DIV66	C. elegans
$p$ (%)	1.2	1.9	1.4	1.5	1.1	1.7	1.4	1.5	2.8
$r_B$	5.9	5.5	10.7	5.7	6.5	4.21	4.0	4.4	—
$f_E$	0.62	0.80	0.84	0.75	0.62	0.66	0.48	0.57	—
$f_I$	0.27	0.13	0.14	0.18	0.21	0.24	0.32	0.28	—
$f_{SCC}$	0.88	0.93	0.98	0.92	0.81	0.90	0.77	0.83	0.85
$l$	4.0	3.7	3.7	3.9	4.1	3.7	3.8	3.8	3.5
CC	0.26	0.36	0.38	0.30	0.25	0.28	0.18	0.22	0.22
SWI	13.1	11.3	17.1	12.7	14.4	10.4	7.9	8.8	2.3

**Table 2.** Basic network measures of the reconstructed networks for the 8 DIVs including the connection probability  $p$ , the ratio  $r_B$  of the number of bidirectionally connected pairs to the expected number  $N(N - 1)p^2/2$  for a random network with  $p$ , the fractions  $f_E$  and  $f_I$  of excitatory and inhibitory nodes, the fraction  $f_{SCC}$  of nodes that form the strongly connected component, the characteristic path length  $l$ , the average clustering coefficient CC and the small world index SWI. When available, the corresponding results for the chemical synapse network of *C. elegans* [20] are included for comparison.

Each node in the neuronal network is inferred to be excitatory or inhibitory according to the sign of the synaptic weights of the outgoing links or has no detectable outgoing links. The spatial distribution of the excitatory and inhibitory nodes and nodes with no detectable outgoing links are shown in Fig 3 for several DIVs. As shown in Table 2, the fraction  $f_E$  of excitatory nodes first increases as DIV increases and peaks at DIV25 then generally decreases as DIV further increases while the fraction  $f_I$  of inhibitory nodes first decreases and reaches a minimum at DIV22 and then generally increases as DIV further increases. Moreover,  $f_I$  ranges from 0.13 to 0.31, which is comparable to the measured fractions (0.15-0.30) of inhibitory neurons in various cortical regions in monkey [25]. The fraction  $f_{SCC}$  of nodes that form the largest strongly connected component exceeds 70% for all the 8 DIVs. The characteristic path length  $l$ , which is the average shortest path length for each pair of nodes in the strongly connected component, is about 4, and the average local clustering coefficient (CC) measuring the average connection probability for nodes with different outgoing degree ranges from 0.18 – 0.38. These values are comparable to those of the chemical synapse network of *C. elegans*. Furthermore, the neuronal networks at all the 8 DIVs have

small-world topology, with small-world index (SWI) substantially greater than one. 170



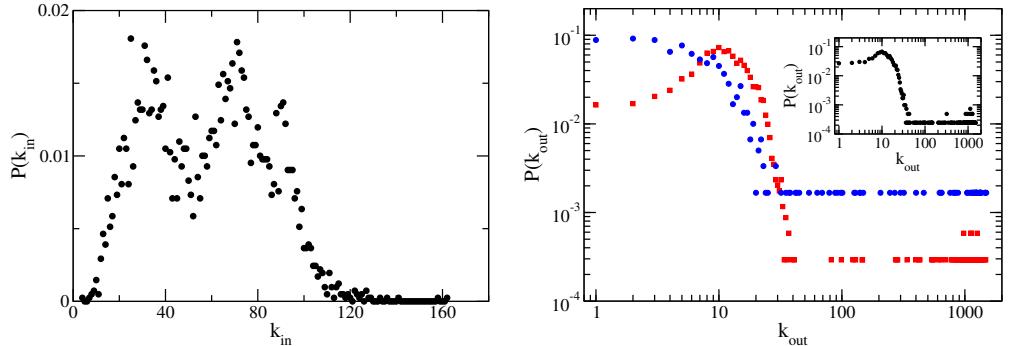
**Fig 3.** Spatial distribution of excitatory nodes (red), inhibitory nodes (blue) and nodes with no detectable outgoing links (black) for DIV11, DIV25, DIV33 and DIV52.

## Distributions of incoming and outgoing degrees 171

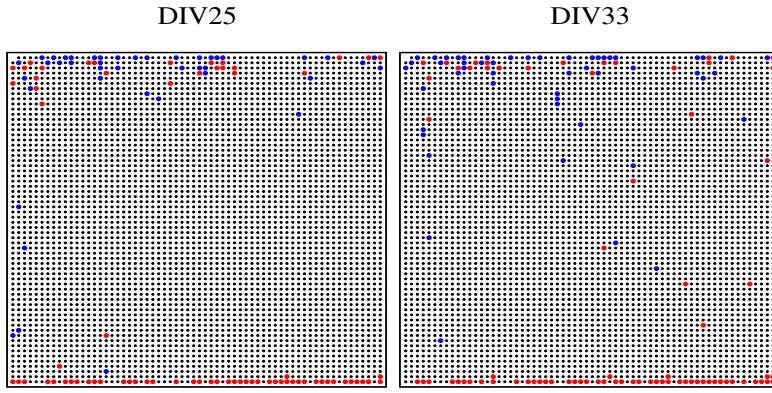
We calculated both the incoming degree  $k_{in}$  and the outgoing degree  $k_{out}$  for each node, and  $k_{in} \neq k_{out}$  as the neuronal networks are directed. The distributions of the incoming and outgoing degrees are qualitatively the same for all the 8 DIVs and the results for DIV25 are shown in Fig 4. A striking feature of the distribution of the incoming degree is its approximate bimodal feature, which is in contrast to the degree distributions of the chemical synapse network of *C. elegans* [20] and the scale-free distribution found in the functional connectivity [10]. For undirected networks, there have been studies indicating that the robustness against both random failures and attacks can be optimized by having a bimodal degree distribution [26, 27]. 172  
173  
174  
175  
176  
177  
178  
179  
180

For DIV25, the outgoing degree of most of the nodes is below 30, which is smaller than the average incoming degree, and a small fraction of nodes have exceptionally large  $k_{out}$  ( $k_{out} > 1000$ ). This is the case for both excitatory and inhibitory nodes. Thus the hubs of the neuronal networks are excitatory and inhibitory feeders. The spatial location of these feeder hubs for DIV25 and DIV33 is shown in Fig 5. Interestingly, the feeder hubs locate mostly on the top and bottom rows of the MEA probe. 181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192

To investigate further the approximate bimodal feature of  $P(k_{in})$ , we calculated separately the distributions of excitatory and inhibitory incoming degrees  $k_{in}^+$  and  $k_{in}^-$  with incoming links of positive and negative  $w_{ij}$  respectively. As shown in Fig 6, both  $P(k_{in}^+)$  and  $P(k_{in}^-)$  are clearly bimodal and have a smaller mode of larger  $k_{in}^+$  or  $k_{in}^-$ . Interestingly, a much reduced fraction of nodes in the mode of larger  $k_{in}^+$  belong to the mode of larger  $k_{in}^-$ , and similarly a much reduced fraction of nodes in the mode of larger



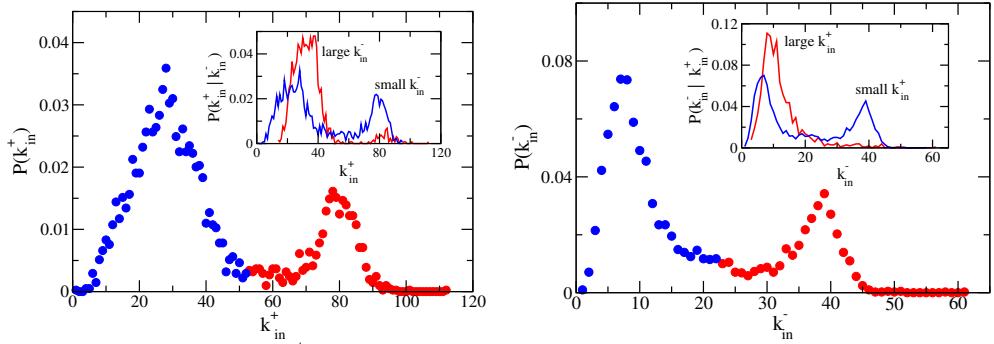
**Fig 4.** Left panel: Distribution  $P(k_{in})$  of the incoming and outgoing degrees for DIV25. Right panel: Distribution  $P(k_{out})$  of the outgoing degrees for excitatory (red) and inhibitory nodes (blue) for DIV25. In the inset, we show the overall distribution of  $k_{out}$  for all the nodes.



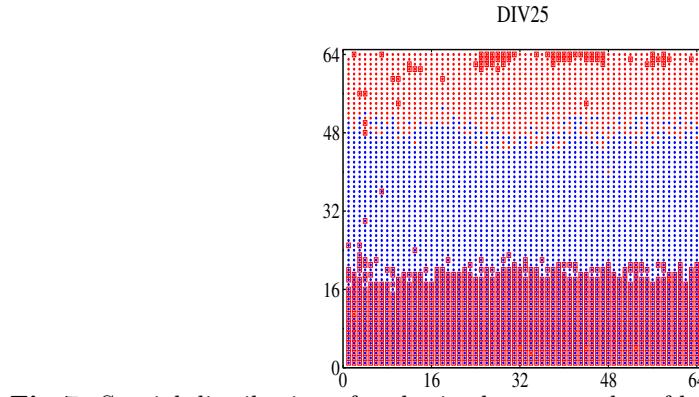
**Fig 5.** Spatial distribution of excitatory (red) and inhibitory (blue) feeder hubs with  $k_{out} > 1000$  for DIV25 and DIV33.

$k_{in}^-$  belong to the mode of larger  $k_{in}^+$  showing that nodes avoid to exhibit both large excitatory and large inhibitory incoming degrees. We show the spatial distribution of nodes in the two modes of the two distributions (see Fig 7). Nodes in the mode of larger  $k_{in}^+$  occupy mostly the upper-quarter region while nodes in the mode of larger  $k_{in}^-$  occupy mostly the lower-quarter region showing clearly that nodes avoid having both large excitatory and large inhibitory incoming degrees. This suggests some form of specialization of nodes in receiving excitatory and inhibitory signals.

We show the distributions of the physical length of excitatory and inhibitory links in Fig 8. Both distributions have a broad peak at  $L \approx 60d = 2.5\text{mm}$ , where  $d = 42\mu\text{m}$  is the separation between electrodes. Interestingly, there is an additional peak in the length distribution of excitatory links at  $L \approx d$ . This suggests that excitatory signal is transmitted at two spatial scales, one localized scale with transmission limited to nearest nodes situated within several tens of microns and another spatially extended scale with transmission to nodes at millimeters away whereas inhibitory signal is mostly transmitted at the longer millimeter range. We will see later that strong excitation is mainly transmitted at the localized spatial scale.



**Fig 6.** Distributions  $P(k_{in}^+)$  and  $P(k_{in}^-)$  of excitatory and inhibitory incoming degrees for DIV25. The distributions are bimodal with a smaller mode of larger  $k_{in}^+$  ( $k_{in}^+ \geq 53$ ) or larger  $k_{in}^-$  ( $k_{in}^- \geq 23$ ) (red) and a larger mode of smaller  $k_{in}^+$  or smaller  $k_{in}^-$  (blue). In the insets, we show the conditional distributions of  $P(k_{in}^+|k_{in}^-)$  and  $P(k_{in}^-|k_{in}^+)$  for nodes in the two modes.



**Fig 7.** Spatial distribution of nodes in the two modes of larger  $k_{in}^+$  (solid red circles) and smaller  $k_{in}^-$  (solid blue circles) and in the two modes of larger  $k_{in}^-$  (open red squares) and smaller  $k_{in}^-$  (all symbols without open red squares) for DIV25.

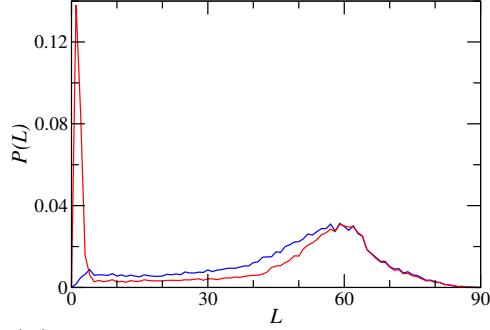
## Distributions of incoming and outgoing synaptic strength

We define the average incoming synaptic strength  $s_{in}(i)$  and the average outgoing synaptic strength  $s_{out}(i)$  for every node  $i$  by

$$s_{in}(i) = \frac{\sum_j w_{ij}}{k_{in}(i)}, \quad s_{out}(i) = \frac{\sum_j w_{ji}}{k_{out}(i)}, \quad (2)$$

For excitatory nodes  $s_{out} > 0$  whereas for inhibitory nodes  $s_{out} < 0$ . For nodes with no detectable outgoing links,  $s_{out} = 0$ . For each node, regardless of whether it is excitatory, inhibitory or has no detectable links,  $s_{in}$  can take either sign. Synaptic weights are difficult to be inferred from MEA recordings and results on the distributions of  $s_{in}$  and  $s_{out}$  are generally lacking in previous studies.

As shown in Fig 9, the distribution of  $s_{out}$  is different for excitatory and inhibitory nodes, and for excitatory nodes, the distribution is further different for nodes with positive and negative average incoming synaptic strength  $s_{in}$ . For excitatory nodes with  $s_{in} > 0$ ,  $s_{out}$  has an approximately lognormal distribution. This result is in accord with the general findings of lognormal distribution of synaptic weights in both in vitro and in vivo studies [28]. For excitatory nodes with  $s_{in} < 0$  and for inhibitory nodes, the distributions of  $|s_{out}|$  have similar functional forms that deviate from lognormal but are



**Fig 8.** Distribution  $P(L)$  of the physical length  $L$  of the excitatory (red) and inhibitory (blue) links for DIV25 with  $L$  measured in units of  $d = 42\mu\text{m}$ .

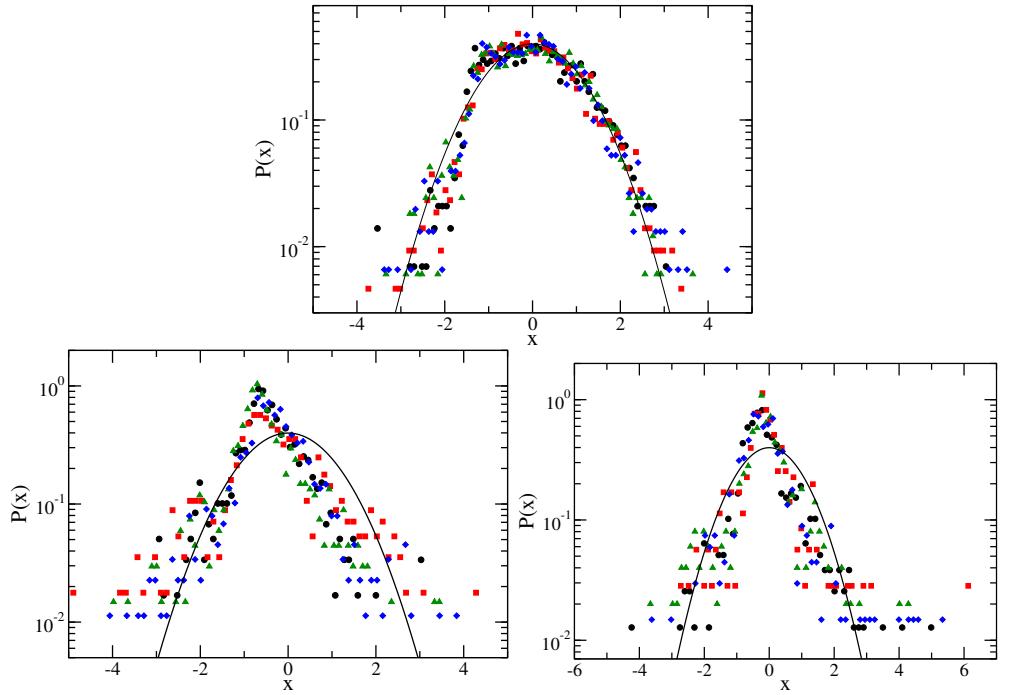
better approximated by slightly asymmetric log-exponential distributions.

The standardized distributions of  $|s_{in}|$  have a universal shape for all DIVs, regardless of whether  $s_{in}$  is greater or less than zero (see Fig 10). These distributions are not lognormal but are better approximated by a highly asymmetric log-exponential distribution. Thus the distributions of both the average incoming and outgoing synaptic strength are non-Gaussian and skewed with long tails. This shows that a small fraction of the nodes have dominantly strong synaptic strength and that the synaptic strength cannot be represented by the mean values.

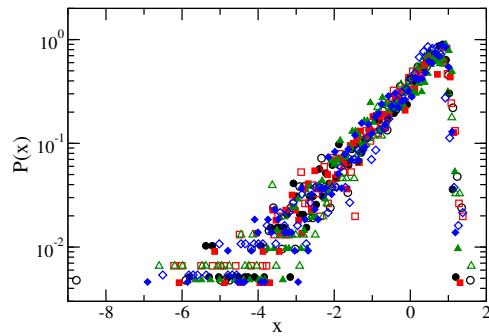
We show the excitatory and inhibitory links of the reconstructed neuronal network for DIV25 in Fig 11. For clarity, we show only the strongest links. We calculated the mean and standard deviation of  $\ln|w_{ij}|$  separately for excitatory and inhibitory links and show only those links with  $\ln|w_{ij}|$  being at least the mean value plus 2.5 times the standard deviation. We note the interesting result that the strongest excitatory links mostly have length of  $L \approx d$ , much shorter than the strongest inhibitory links which mostly have  $L$  of several tens of  $d$  of the order of mm. This indicates that strong excitatory signals are mostly transmitted at the localized scale while strong inhibition is mostly long range. Moreover, the strongest inhibitory links are dominated by links from two hubs and the strongest links, excitatory or inhibitory alike, are mainly towards excitatory nodes. We further checked that the shortest links of  $L = d$  are mostly excitatory links towards excitatory nodes and have larger synaptic weight on average.

### Comparison of the partial network reconstructed using only measurements from a subset of electrodes and the corresponding subnetwork

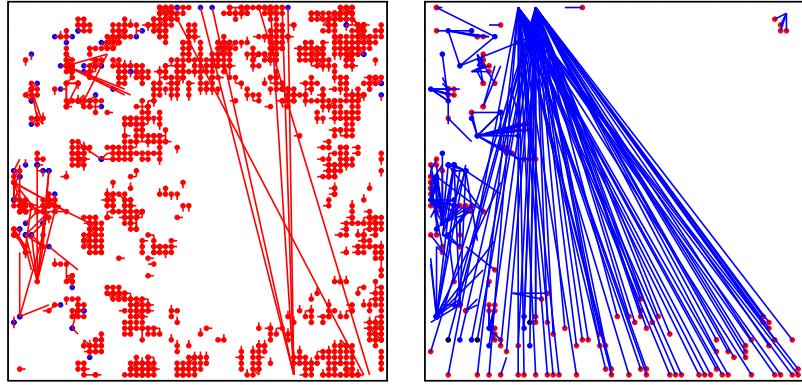
As the working area of the MEA probe is more than 4 times the active electrode area, many cells were outside the active electrode area. These cells could form synaptic connections with cells within the electrode area but their spontaneous activities were not recorded. Thus there are missing or hidden nodes in the reconstructed neuronal networks. For bidirectional networks, it has been shown that the presence of hidden nodes that are randomly missed out has no adverse effects on the reconstruction of the network of the measured nodes [29]. The effects of hidden nodes on the reconstruction of directed effective connectivity for general directed networks have yet to be worked out. Thus we studied the effects of hidden nodes in the present problem by reconstructing a partial network using recordings from a subset of 45 by 45 electrodes in the central region of the active electrode area and compared this partial network with the subnetwork of the corresponding 2025 nodes of the reconstructed network using recordings from all 4095 electrodes. We found that the partial network captures



**Fig 9.** Top panel: Distributions  $P(x)$  of  $x = [\ln(s_{out}) - \mu]/\sigma$  for excitatory nodes with  $s_{in} > 0$ . Bottom panel: Distributions  $P(x)$  of  $x = [\ln |s_{out}| - \mu]/\sigma$  for (left panel) excitatory nodes with  $s_{in} < 0$  and for (right panel) inhibitory nodes.  $\mu$  and  $\sigma$  are the mean and standard deviation of  $\ln(s_{out})$  respectively. DIV11 (black circles), DIV33 (red squares), DIV52 (green triangles), and DIV66 (blue diamonds) and the solid line is a standardized Gaussian distribution with zero mean and unit variance.

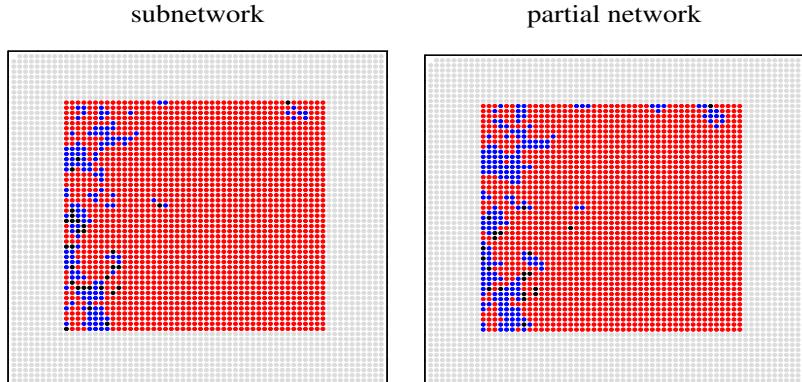


**Fig 10.** Distributions  $P(x)$  of  $x = [\ln(|s_{in}|) - \mu]/\sigma$  for nodes with  $s_{in} < 0$  (open symbols) and  $s_{in} > 0$  (solid symbols) for DIV11 (black circles), DIV33 (red squares), DIV52 (green triangles), and DIV66 (blue diamonds).  $\mu$  and  $\sigma$  are the mean and standard deviation of  $\ln |s_{in}|$  respectively.



**Fig 11.** Left panel: The strongest excitatory (left panel) and inhibitory links (right panel) of the reconstructed neuronal network for DIV25 with  $\ln |w_{ij}| \geq \text{mean} + 2.5$  standard deviation, with the mean value and standard deviation of  $\ln |w_{ij}|$  calculated separately for excitatory and inhibitory links. Each excitatory (red) link points from an excitatory node to either an excitatory (red) or inhibitory node (blue). Similarly, each inhibitory (blue) link points from an inhibitory node to either an excitatory (red), inhibitory (blue) or node with no detectable outgoing links (black).

correctly 99.8% of the non-existent links and 83.0% of the links in the subnetwork. We compare the spatial distribution of the excitatory, inhibitory nodes and nodes with no detectable links in the subnetwork (which is the same as the central region of the plots shown in Fig 3) and the partial network in Fig 12. There are differences but a great resemblance can be seen.

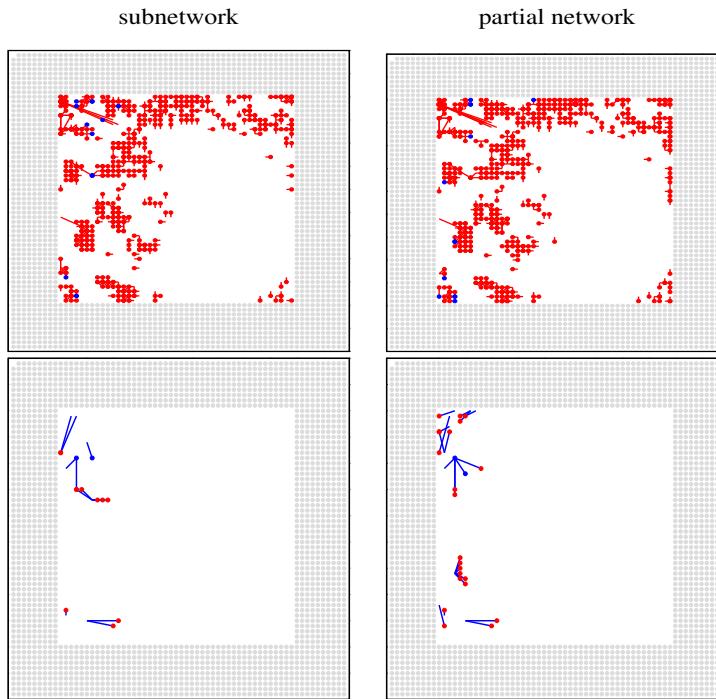


**Fig 12.** Spatial distribution of excitatory nodes (red), inhibitory nodes (blue) or nodes with no detectable outgoing links (black) for DIV25 in the subnetwork of the central  $45 \times 45$  electrodes (left) and the partial network reconstructed by using only measurements of a subset of these 2025 electrodes (right).

We further compare the strongest excitatory and inhibitory links of the partial network and the subnetwork using the same criterion defined for strongest links in Fig 11. As can be seen in Fig 13, the partial network can capture the strongest excitatory links in the subnetwork very well. Though there is a larger error for the partial network in obtaining the strongest inhibitory links, the partial network manages to give a correct qualitative picture of the spatial distribution of these strongest inhibitory links. These results hence support that the directed effective connectivity among the measured nodes in the reconstructed networks is not much affected by the missing signal or hidden nodes due to the neural cells lying outside the active electrode

area of the MEA probe.

274



**Fig 13.** Top panel: The strongest excitatory links of the subnetwork (left) and the partial network (right) for DIV25. Bottom panel: The strongest excitatory links of the subnetwork (left) and the partial network (right) for DIV25.

## Discussion

275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296

Revealing directed connectivity of neuronal networks from electrophysiological data recorded by large-scale MEA is a challenging inverse problem. Existing methods focus on functional connectivity which measures statistical dependencies of the activities of neurons. To address questions regarding connectivity and dynamics and network structure and functions of neuronal networks, the relevant quantity should, however, be effective connectivity which measures direct causal interactions and not functional connectivity. As it is natural to infer effective connectivity using a specific dynamical model, it is commonly thought that effective connectivity is model-dependent. This further adds to the difficulty to generally investigate directed effective connectivity of neuronal networks. One established technique for estimating functional connectivity makes use of the cross-correlation of spiking activity of neurons. It was demonstrated theoretically that for a generic class of models and not just for a specific model, the effective connectivity matrix is contained in a theoretical relation between the time-lagged cross-covariance and the equal-time cross-covariance [13]. Making use of this relation, a method has been developed to recover effective connectivity [13] and this method was validated for models with different dynamics and coupling using numerical data [13, 16]. This covariance-relation based method involves a calculation of a principal matrix logarithm, which is very sensitive to noise in the data. When we applied it directly to the MEA recordings, we obtained an unphysical complex matrix for the supposed effective connectivity matrix. A complex matrix was also found when this covariance relation was used directly to estimate directed effective connectivity of a

cortical network of 68 regions from fMRI recordings and this motivated the development of a Lyapunov optimization procedure [11]. Here, we showed that the unphysical complex matrix problem can be avoided by first applying a moving average filter to the MEA recordings to reduce random noise (c.f. Materials and methods for the method of reconstruction). Moreover, we extended the originally proposed clustering analysis using a Gaussian mixture model [13] to include also a single-Gaussian fit with outliers. We made additional assumptions that the neuronal networks are sparse and outgoing links of any node, if exist, have to be all excitatory or all inhibitory. We have shown that this new covariance-based method, the original proposed method together with the modifications, can estimate the directed and weighted effective connectivity matrix of neuronal networks from recordings taken by MEA of over 4000 electrodes. Specifically, the simulated spiking activity using the estimated directed and weighted effective connectivity in a stochastic LIF model can capture the spontaneous network bursts of the measured spiking activity and retrieve the null-nodes in the time-window of observation reasonably well.

The balance between excitation and inhibition in the cortex is believed to play an important role in executing proper brain functions and disruption of such a balance may underlie the behavioural deficits that are observed in conditions such as autism and schizophrenia [30–34]. To study and gain insights of excitation/inhibition balance in neuronal networks, it is important that the methods estimating connectivity can detect excitatory and inhibitory links equally well. The cross-covariance based method for functional connectivity has a considerably weaker sensitivity for detecting inhibition and a recently introduced filtering approach [10] overcomes this problem but the detected excitation/inhibition ratio from MEA of 4096 electrodes tends to stabilize only when the recording time is of the order of 30 min. Our method can detect excitatory and inhibitory links equally well and we successfully detected both excitatory and inhibitory links from the same type of MEA using only a recording time of 5 min and identified nodes of the neuronal networks as excitatory, inhibitory or with no detectable outgoing links. The fraction  $f_I$  of inhibitory nodes is 0.13–0.31 for the 8 DIVs studied, which overlaps with the range 0.15–0.30 of measured fractions of inhibitory neurons in various cortical regions of monkey [25]. In this sense, the reconstructed neuronal networks are excitation/inhibition balanced as found in monkey cortex.

Neuronal networks are expected to be organized and should thus be highly nonrandom in order to transmit signals efficiently and to carry out their functions effectively. Similar to the chemical synapse network of *C. elegans* [20], the reconstructed neuronal networks have small-world topology, and it has been shown that small-world network allows efficient transmission of information [35]. The nonrandom feature of local cortical circuits has been well documented by an over-representation of bidirectional connected pairs as compared to a random network of the same connection probability [21–24]. In the reconstructed neuronal networks, the number of bidirectionally connected pairs was indeed found to exceed at least four times the expected number for the random network. Bidirectional connected pairs can play a direct role in forming feedback control in neuronal circuits. We estimated the functional connectivity for DIV25, DIV45 and DIV66 by using the cross-covariance based methods [10] and bidirectionally connected pairs are not found. Neuronal networks are also expected to be robust against random failure or targeted attacks. There were studies in network science indicating that optimization of the network topology of undirected networks against both random failures and attacks can be achieved by a bimodal degree distribution [26, 27]. Interestingly, the incoming degree distribution of the reconstructed neuronal networks is found to be bimodal. It is thus interesting to investigate whether a bimodal incoming degree distribution would optimize the robustness of directed networks against certain attacks.

The incoming and outgoing degrees of the neuronal networks have distinctive distributions. The outgoing degree distribution for both excitatory and inhibitory nodes has a long tail. The maximum incoming degree is of the order of 100 while the maximum outgoing degree is of the order of 1000. Thus hubs are feeders rather than receptors and there are both excitatory and inhibitory feeders locating mainly on the top and bottom rows of the MEA probe. The excitatory and inhibitory incoming degrees have bimodal distributions and nodes in the two modes are spatially segregated. The nodes in the modes of larger  $k_{in}^+$  and larger  $k_{in}^-$  occupy the upper-quarter and lower-quarter regions of the MEA probe respectively showing clearly that nodes do not exhibit both large excitatory and large inhibitory incoming degrees. These results suggest some form of differentiation of roles of the nodes in the neuronal circuits in receiving excitatory and inhibitory signals. The distribution of the physical length of excitatory links has two peaks indicating that excitatory signals are transmitted at two distinct spatial scales, one highly localized scale with transmission to nearby nodes within several tens of microns and another spatially extended scale with transmission to nodes at millimeters away. The distribution of the physical length of inhibitory links has only one broad peak at the spatially extended scale. We further found that strong excitatory signals are mostly transmitted locally to nearby excitatory nodes while strong inhibitory signals are mostly transmitted at the spatially extended millimeter scale from a few hubs. Long range inhibitory projections from GABAergic neurons have been demonstrated in adult rat hippocampus [36, 37] and the hub function of these neurons has been suggested [38, 39].

Incoming synaptic strength is difficult to measure directly and results on their distributions have been lacking in the literature. The present covariance-relation based method allows us to estimate the synaptic weights of all the links and, as a result, we can calculate the distributions of incoming and outgoing synaptic strength of the nodes. We found that these distributions are described by lognormal or log-exponential functional form (Fig 9 and 10) and they are in line with the ubiquitous lognormal-like distributions of many physiological and anatomical features at different levels in the brain [28]. The significance of these skewed non-Gaussian distributions with long tails is that a small fraction of nodes have dominantly strong synaptic strength suggesting the possibility that the bulk of the information flow occurring mostly through these nodes [28]. Understanding how this lognormal-like synaptic strength might be related to the spiking dynamics or the general relation between connectivity and dynamics will be a topic of great interest for future studies.

## Materials and methods

### Ethics statement

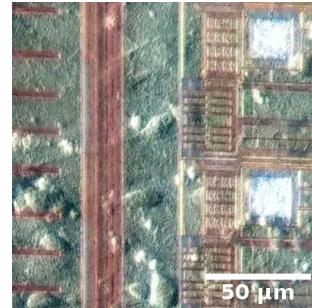
Cortices from embryonic day 17 (E17) Wistar rat were used in the experiments. The experimental protocol was evaluated and approved by the Institutional Animal Care and Use Committee of Academia Sinica (AS IACUC) with Protocol ID: 12-12-475.

### Neuronal cultures and experimental set-up

Tissues were dissected from 2 to 3 rats and digested with 0.125 % trypsin for 15 min at 37°C. A small drop (100  $\mu$ l) of cell suspension was plated on the working area of the complementary metal-oxide-semiconductor (CMOS)-based high density multielectrode array (HD-MEA), which was pre-coated with 0.1 % Poly-D-lysine (Sigma P6407) and 0.1 % adhesion proteins laminin (Sigma L2020) (see Fig 14). A total of about  $6 \times 10^4$  cells were plated on a working area of about 6 mm  $\times$  6mm of the HD-MEA chip

resulting on a cell density of about 1670 cells/mm<sup>2</sup>. After plating on the chip, cultures were filled with 1 ml of culture medium [DMEM (Gibco 10569) + 5% FBS (Gibco 26140) + 5% HS (Gibco 16050) + 1% PS (Gibco 15140)] and placed in a humidified incubator (5% CO<sub>2</sub>, 37°C). Half of the medium was replaced by Neurobasal medium supplemented with B27 [Neurobasel medium (Gibco. 21103) + 2% 50X B27 supplement (Gibco. 17504) + 200 μM GlutaMAX (Gibco 35050)] twice a week.

The HD MEA probe (HD-MEA Arena, 3Brain AG) has 4096 electrodes arranged in a 64 by 64 square grid. The size of each square electrode is 21 μm by 21 μm and the electrode pitch is 42 μm which gives an active electrode area of 2.67 mm by 2.67 mm. Neuronal activities from HD MEA were recorded with the recording device (BioCAM, 3Brain AG) and the associate software (BrainWave 2.0, 3Brain AG) at 7.06 kHz. One electrode was used for calibration purpose so there were 4095 electrodes that recorded signals. Spontaneous activities were recorded from 6 to over 60 days in vitro. Samples were placed into the recording device 10 min before the recording in order to prevent the effects of vibration. Each experimental session lasted 5 min and recorded in dark since the CMOS is a light active material.



**Fig 14.** Surface of high-density CMOS MEA. The line shows a scale of 50 μm.

## Method of reconstruction

We assumed that the activity  $x_i(t)$ , recorded by each of the electrodes, is stationary and considered a generic dynamical model in which the time evolution of the fluctuations around the average value,  $\delta x_i(t) = x_i(t) - \langle x_i(t) \rangle$ , is governed by a system of linear stochastic differential equations

$$\frac{d\delta x_i}{dt} = \sum_j Q_{ij} \delta x_j + \eta_i, \quad (3)$$

where  $\langle \dots \rangle$  denotes a time average,  $Q_{ij} = w_{ij}(1 - \delta_{ij}) + Q_{ii}\delta_{ij}$  are the elements of the interaction matrix  $\mathbf{Q}$  and  $\eta_i$  is a Gaussian white noise of zero mean with

$$\overline{\eta_i(t)\eta_j(t')} = D_{ij}\delta(t - t'). \quad (4)$$

When the activity  $x_j$  affects the activity  $x_i$ , there is a link from node  $j$  to node  $i$  with synaptic weight  $w_{ij}$  otherwise  $w_{ij} = 0$ . Thus  $w_{ij}$  give the directed and weighted effective connectivity of the neuronal network. We note that such a linear stochastic system of stationary fluctuating activity can arise even for systems whose activity obey nonlinear dynamics [13]. For such a model, the time-lagged covariance matrix  $\mathbf{K}(\tau)$  and the equal-time covariance matrix  $\mathbf{K}(0)$  [11, 13, 40, 41] of the measurements are related by the interaction matrix  $\mathbf{Q}$ :

$$\mathbf{K}(\tau) = \exp(\tau\mathbf{Q})\mathbf{K}(0) \quad (5)$$

The elements of  $\mathbf{K}(\tau)$  and  $\mathbf{K}(0)$  are defined by

$$K_{ij}(\tau) = \langle [x_i(t + \tau) - \langle x_i(t + \tau) \rangle][x_j(t) - \langle x_j(t) \rangle] \rangle \quad (6)$$

$$K_{ij}(0) = \langle [x_i(t) - \langle x_i(t) \rangle][x_j(t) - \langle x_j(t) \rangle] \rangle \quad (7)$$

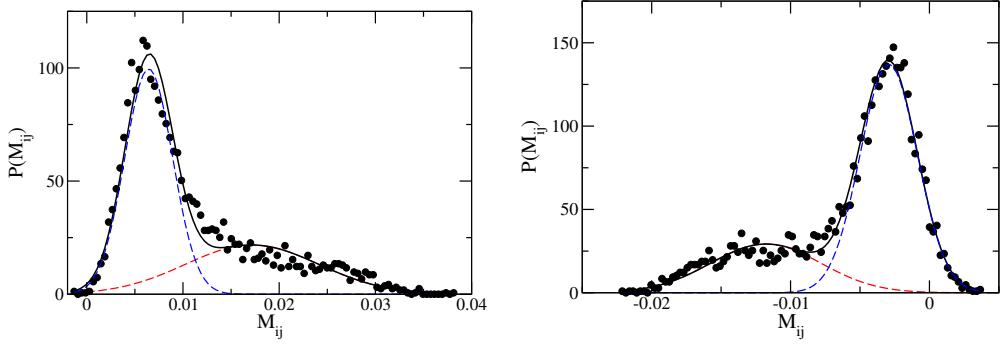
respectively. This relation implies that the off-diagonal elements of  $\mathbf{M} = \log(\mathbf{K}(\tau)\mathbf{K}(0)^{-1})/\tau$ , where  $\log$  is the principal matrix logarithm, would separate into two groups, one corresponds to  $w_{ij} = 0$  and another to  $w_{ij} \neq 0$ . It has been shown [13] that by clustering analysis of  $M_{ij}$  using a Gaussian mixture model,  $w_{ij}$  can be inferred. We extended this proposed clustering analysis to include also a single-Gaussian fit with outliers.

For the reconstruction, we only made use of measurements  $y_i(t)$ ,  $i = 1, 2, \dots, 4095$  taken during which all the 4095 electrodes were recording properly. The principal matrix logarithm is very sensitive to noise in the measurements and a calculation of  $\mathbf{K}_y(\tau)$  and  $\mathbf{K}_y(0)$  directly from these MEA recordings  $y_i(t)$  [as defined in Eqs (6) and (7) with  $x_i$  replaced by  $y_i$ ] would lead to the matrix  $\log(\mathbf{K}_y(\tau)\mathbf{K}_y(0)^{-1})$  having complex elements. This has also been found for fMRI measurements [11]. We avoided this problem by applying a moving average filter to the MEA recordings to reduce random noise. The moving average filter is a simple digital low-pass filter that is optimum for reducing random noise while retaining sharp transitions in the data [43]. Specifically, we took  $x_i(t) = [y_i(t) + y_i(t + \Delta)]/2$ , where  $\Delta$  is the sampling time interval for all the valid MEA recordings  $y_i(t)$ ,  $i = 1, 2, \dots, 4095$ , and calculated  $\mathbf{K}(\tau)$  and  $\mathbf{K}(0)$  from  $x_i(t)$  with  $\tau = \Delta$ . The resulting matrix  $\mathbf{M}$  was found to be real. Then we adopted the clustering method of Ref. [13] with several modifications to cluster  $M_{ij}$ 's for each fixed  $j$  into two groups, one corresponds to  $w_{ij} = 0$  and the other to  $w_{ij} \neq 0$ .

We assumed that the outgoing links of each node, if exist, can only be all excitatory with  $w_{ij} > 0$  or all inhibitory with  $w_{ij} < 0$ . This is in accord to neurons are either excitatory or inhibitory. We first fitted the values of  $M_{ij}$ 's with fixed  $j$  by a Gaussian mixture model of two components, namely by a sum of two Gaussian distributions, using MATLAB 'fitgmdist', which return the means,  $\mu_1$  and  $\mu_2$ , and standard deviations,  $\sigma_1$  and  $\sigma_2$ , of the two Gaussian distributions, and the relative proportion of the two components. We divided the results into three cases according to the separation between the two fitted Gaussian distributions: (i) the two distributions are well separated with  $|\mu_1 - \mu_2| > \sigma_1 + \sigma_2$ , (ii) the two distributions are very close to one another with  $|\mu_1 - \mu_2| < \max(\sigma_1, \sigma_2)$ , and (iii) the in-between case with  $\max(\sigma_1, \sigma_2) \leq |\mu_1 - \mu_2| \leq \sigma_1 + \sigma_2$ . For case (i), we continued with this two-Gaussian fit. For case (ii), we fitted the data again by a single Gaussian distribution instead. For case (iii), we chose either the two-Gaussian fit or the single-Gaussian fit according to the Bayesian information criterion (BIC) for fitting models selection [42, 44].

When using the two-Gaussian fit, we thus inferred the unconnected component for  $w_{ij} = 0$  as the component of a relative proportion greater than 0.6 based on the assumption that the neuronal cultures are sparse networks or as the component whose mean is closer to zero when none of the relative proportions exceeds 0.6. We obtained the posterior probability  $p$  of each data point belonging to the unconnected component using MATLAB 'cluster'. For each  $M_{ij}$ , if  $p > 0.5$  we inferred  $w_{ij} = 0$  otherwise, we inferred  $w_{ij}$  to be given by  $M_{ij} - \langle M_{ij}|w_{ij} = 0 \rangle_i$  and  $\langle M_{ij}|w_{ij} = 0 \rangle_i$  is the average of  $M_{ij}$  with inferred  $w_{ij} = 0$  over  $i$  for fixed  $j$ . Thus the synaptic weights  $w_{ij}$  of the outgoing links of each node  $j$  have the same sign. Depending on whether  $w_{ij}$  is greater than or less than zero, the node was inferred to be excitatory or inhibitory (c.f. Fig 15).

When using the single-Gaussian fit, we identified the data points that are well fitted by the Gaussian, denoted by  $P_G$  of mean  $\mu$ , with  $w_{ij} = 0$  and the outliers, which are data points deviate significantly from the fitted Gaussian, with  $w_{ij} \neq 0$ . For this



**Fig 15.** Typical distributions  $P(M_{ij})$  of  $M_{ij}$  (circles) with  $j$  fixed for an excitatory node in DIV66 (left panel) and an inhibitory node in DIV 11 (right panel) inferred by using a Gaussian-mixture model (solid black lines) of two components (blue and red dashed lines).

purpose, we let  $x_E$  and  $x_I$  be

$$x_E = \min_{x > \mu} \{P_K(x) = 3P_G(x)\} \quad (8)$$

$$x_I = \max_{x < \mu} \{P_K(x) = 3P_G(x)\} \quad (9)$$

where  $P_K$  is the distribution of  $M_{ij}$  obtained by the Kernel density estimation and calculated the number of data points ( $N_1$ ) satisfying  $M_{ij} > x_E$  and the number of data points ( $N_2$ ) satisfying  $M_{ij} < x_I$ . The outliers are the  $N_1 M_{ij}$ 's if  $N_1 > N_2$  or the  $N_2 M_{ij}$ 's if  $N_1 < N_2$ . The identified outliers were inferred to have  $w_{ij} = M_{ij} - \langle M_{ij} | w_{ij} = 0 \rangle_i$ . If both  $N_1 = N_2 = 0$ , then node  $j$  was inferred to have no detectable outgoing links.

## Code availability

The MATLAB program for clustering analysis of the calculated matrix  $\mathbf{M}$  to give the adjacency matrix elements  $A_{ij}$  and coupling  $w_{ij}$  can be found in the Open Science Framework (OSF) <https://osf.io/5dc4j/>

## Stochastic leaky-integrate-and-fire spiking neuron model

The dynamics of a stochastic leaky-integrate-and-fire (LIF) neuron is governed by the following equation

$$\frac{dV_i(t)}{dt} = -\frac{1}{\tau_m}[V_i(t) - V_r] + \frac{I_i^{syn}(t)}{C_m} + \text{noise} \quad (10)$$

where  $V_i(t)$  is the membrane potential of neuron  $i$ ,  $V_r$  is the resting potential,  $C_m$  and  $G_m$  are the membrane capacitance and membrane conductance,  $\tau_m = C_m/G_m$  is the decay constant, and  $I_i^{syn}(t)$  is the synaptic current input from the presynaptic nodes. When  $V_i(t)$  exceeds a threshold  $V_{th}$ , node  $i$  generates an action potential or spike at that instant and  $V_i$  is reset to  $V_r$ . The synaptic current  $I_i^{syn}(t)$  is given by

$$I_i^{syn}(t) = [V_E - V_i(t)]G_E(t) - [V_i(t) - V_I]G_I(t) \quad (11)$$

where  $V_E$  and  $V_I$  are the reversal potentials satisfying  $V_I \leq V_r < V_{th} \leq V_E$ , and  $G_E(t)$  and  $G_I(t)$  are the excitatory and inhibitory synaptic conductances given by

$$G_E(t) = G \sum_{j=1, w_{ij}>0} |w_{ij}| \sum_k e^{-\frac{(t-t_{j,k})}{\tau_s}} \Theta(t - t_{j,k}) \quad (12)$$

$$G_I(t) = G \sum_{j=1, w_{ij}<0}^N |w_{ij}| \sum_k e^{-\frac{(t-t_{j,k})}{\tau_s}} \Theta(t - t_{j,k}) \quad (13)$$

where  $G$  is a typical synaptic conductance,  $w_{ij}$  are the synaptic weights,  $\tau_s$  is the decay constant of the synaptic conductances,  $\Theta(t)$  is the Heaviside step function and  $t_{j,k}$  is the spiking time of the  $k$ -th spike of the presynaptic neuron  $j$ .

To simplify the notations, we worked with the dimensionless membrane potential, which is defined by

$$v_i(t) = \frac{V_i(t) - V_r}{V_{th} - V_r} \quad (14)$$

with  $0 \leq v_i \leq 1$  then the equations become

$$\frac{d}{dt} v_i(t) = -\frac{v_i}{\tau_m} + \beta i_i^{syn}(t) + \eta_i, \quad (15)$$

$$i_i^{syn} = \sum_j (\alpha_{ij} - v_i) |w_{ij}| \sum_k e^{-\frac{(t-t_{j,k})}{\tau_s}} \Theta(t - t_{j,k}) \quad (16)$$

where  $\beta = G/C_m$ ,

$$\alpha_{ij} = \begin{cases} v_E = \frac{V_E - V_r}{V_{th} - V_r} > 0, & w_{ij} > 0 \\ v_I = \frac{V_I - V_r}{V_{th} - V_r} < 0, & w_{ij} < 0, \end{cases} \quad (17)$$

and  $\eta_i$  is a Gaussian white noise of zero mean defined by Eq (4) with  $D_{ij} = \sigma^2$ . When  $v_i \geq 1$ ,  $v_i$  is reset to 0.

In the numerical simulations, we used  $\tau_m = 20$  ms,  $\tau_s = 2$  ms,  $v_E = 6$  and  $v_I = -2$  (which follow from  $V_r = -60$  mV,  $V_{th} = -50$  mV,  $V_E = 0$  mV and  $V_I = -80$  mV) [45] and  $\sigma = 2$ . We took  $\beta = 1000$  and  $\beta = 40$  respectively for the reconstructed networks and the control randomly-shuffled networks with standardized Gaussian synaptic weights. The time step  $\Delta t$  of the simulations was taken to be the sampling time interval  $\Delta t = 1/7060s \approx 0.142$  ms and we obtained the spiking activity of each node  $i$  using the measured spikes of all its presynaptic nodes (with  $w_{ij} \neq 0$ ) as inputs.

## References

1. Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci*. 2009; 10(4):186-198.  
doi:10.1038/nrn2575
2. Bassett DS, Sporns O. Network neuroscience. *Nat Neurosci*. 2017; 20(3):353-364.  
doi:10.1038/nn.4502
3. Eckmann JP, Feinerman O, Gruendinger L, Moses E, Soriano J, Tlusty T. The physics of living neural networks. *Phys Rep*. 2007; 449:54-76.  
doi:10.1016/j.physrep.2007.02.014

4. Feinerman O, Rotem A, Moses E. Reliable neuronal logic devices from patterned hippocampal cultures. *Nat Phys.* 2008; 4:967-973.
5. Soriano J, Casademunt J. Neuronal cultures: The brain's complexity and non-equilibrium physics, all in a dish. *Contrib Sci.* 2015; 11:225-235 doi:10.2436/20.7010.01.234
6. Obien MEJ, Deligkaris K, Bullmann T, Bakkum DJ, Frey U. Revealing neuronal function through microelectrode array recordings. *Front Neurosci.* 2015; 8:423. doi:10.3389/fnins.2014.00423
7. Friston KJ. Functional and Effective Connectivity: A Review. *Brain Connectivity* 2011; 1:13-36
8. Maccoine A, Garofalo N, Nieus T, Tedesco M, Berdondini L, and Martinoia, S. Multiscale functional connectivity estimation on low-density neuronal cultures recorded by high-density CMOS Micro Electrode Arrays. *J NeuroSci Methods* 2012; 207:161-171.
9. Poli D, Pastore VP, and Massobrio P. Functional connectivity in in vitro neuronal assemblies. *Front Neural Circuits* 2015; 9:57 doi:10.3389/fncir.2015.00057.
10. Pastore VP, Massobrio P, Godjoski A, Martinoia S. Identification of excitatory-inhibitory links and network topology in large-scale neuronal assemblies from multi-electrode recordings. *PLoS Comput Biol.* 2018; 14(8):e1006381. doi:10.1371/journal.pcbi.1006381.
11. Gilson M, Moreno-Bote R, Ponce-Alvarez A, Ritter P, Deco G. Estimation of Directed Effective Connectivity from fMRI Functional Connectivity Hints at Asymmetries of Cortical Connectome. *PLoS Comput Biol.* 2016; 12(3):e1004762. doi:10.1371/journal.pcbi.1004762
12. Tam HC, Ching ESC, Lai PY. Reconstructed networks from dynamics with correlated noise. *Physica A* 2018; 502:106-122.
13. Ching ESC, Tam HC. Reconstructing links in directed networks from noisy dynamics. *Phys Rev E* 2017; 95:010301(R).
14. Lai PY. Reconstructing Network topology and coupling strengths in Directed Networks of discrete-time dynamics. *Phys Rev E.* 2017; 95:022311.
15. FitzHugh R. Impulses and physiological states in theoretical models of nerve membrane. *Biophys J.* 1961; 1:445.
16. Tam HC. Reconstruction of Networks from Noisy Dynamics [Ph.D Thesis] The Chinese University of Hong Kong; 2017. Available from: <https://search.proquest.com/docview/2022313258?accountid=10371>
17. Maccione A, Gandolfo M, Massobrio P, Novellino A, Martinoia S, Chiappalone M. A novel algorithm for precise identification of spikes in extracellularly recorded neuronal signals. *J Neurosci Methods* 2009; 177:241-9.
18. Izhikevich EM. Which model to use for cortical spiking neurons? *IEEE Trans Neural Netw.* 2004; 15:1063-1070 doi:10.1109/TNN.2004.832719
19. Burkitt AN. A Review of the Integrate-and-fire Neuron Model: I. Homogeneous Synaptic Input. *Biol Cybern.* 2006; 95(1):1-19.

20. Varshney LR, Chen BL, Paniagua E, Hall DH, Chklovskii DB. Structural Properties of the *Caenorhabditis elegans* Neuronal Network. *PLoS Comput Biol*. 2011; 7(2):e1001066. doi:10.1371/journal.pcbi.1001066
21. Song S, Sjöström PJ, Reigl M, Nelson S, B Chklovskii DB. Highly Nonrandom Features of Synaptic Connectivity in Local Cortical Circuits. *PLoS Biology* 2005; 3(10):e350. doi:10.1371/journal.pbio.0030350.
22. Markram H, Lübke J, Frotscher M, Roth A, Sakmann B. Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. *J. of Physio.* 1997; 500.2:409-440.
23. Sjöström PJ, Turrigiano GG, Nelson SB. Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron* 2001; 32:1149-1164.
24. Perin R, Berger TK, Markram H. A synaptic organizing principle for cortical neuronal groups. *Proceedings of the National Academy of Sciences* 2011; 108(13): 5419-5425. doi:10.1073/pnas.1016051108
25. Hendry SH, Schwark HD, Jones EG, Yan J. Numbers and proportions of GABA-immunoreactive neurons in different areas of monkey cerebral cortex. *J Neurosci.* 1987; 7:1503-1519.
26. Valente AXCN, Sarkar A, Stone HA. Two-peak and three-peak optimal complex networks. *Phys Rev Lett* 2004; 92: 118702.
27. Tanizawa T, Paul G, Cohen R, Havlin S, Stanley HE. Optimization of network robustness to waves of targeted and random attacks. *Phys Rev E* 2005; 71: 047101.
28. Buzsaki G, Mizuseki K. The log-dynamic brain: how skewed distributions affect network operations. *Nature Neurosci.* 2014; 15:264-278.
29. Ching ESC, Tam PH. Effects of hidden nodes on the reconstruction of bidirectional networks. *Phys Rev E.* 2018; 98:062318.
30. Rubenstein JL, Merzenich MM. Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes Brain Behav* 2003; 2(5):255-67.
31. Levitt P, Eagleson KL, Powell EM. Regulation of neocortical interneuron development and the implications for neurodevelopmental disorders. *Trends Neurosci.* 2004; 27(7):400-6
32. Yizhar O, Fenno LE, Prigge M, Schneider F, Davidson TJ, O'Shea DJ et al. Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature* 2011; 477:171-178. doi:10.1038/nature10360
33. Lewis DA, Fish KN, Arion D, Gonzalez-Burgos G. Perisomatic inhibition and cortical circuit dysfunction in schizophrenia. *Curr Opin Neurobiol.* 2011; 21(6):866-72
34. Marin O. Interneuron dysfunction in psychiatric disorders. *Nat Rev Neurosci.* 2012; 13(2):107-20
35. Latora V, Marchiori M. Efficient Behavior of Small-World Networks, *Phys Rev Lett.* 2001; 87:198701.
36. Jinno S et al. Neuronal Diversity in GABAergic Long-Range Projections from the Hippocampus. *J Neurosci.* 2007; 27:8790

37. Wick ZC, Tetzlaff MR, Krook-Magnuson E. Novel long-range inhibitory nNOS-expressing hippocampal cells. *eLife* 2019; 8:e46816. doi:10.7554/eLife.46816
38. Buzsaki G, Geisler C, Henze DA, Wang XJ. Interneuron Diversity series: Circuit complexity and axon wiring economy of cortical interneurons. *Trends Neurosci.* 2004; 27:186.
39. Villette V et al. Development of Early-Born  $\gamma$ -Aminobutyric Acid Hub Neurons in Mouse Hippocampus From Embryogenesis to Adulthood. *J Comp Neurol.* 2016; 524:2440-2461. doi:10.1002/cne.23961
40. Prill RJ, Vogel R, Cecchi GA, Altan-Bonnet G, and Stolovitzky G. Noise-Driven Causal Inference in Biomolecular Networks, *PLoS ONE* 2015; 10:e0125777. doi.org/10.1371/journal.pone.0125777
41. Lünsmann BJ. Reconstruction of Physical Interactions in Stationary Stochastic Network Dynamics [MSc Thesis]. University Göttingen; 2015. Available from: <http://www.nld.ds.mpg.de/publications/mastersthesisreference.2015-06-10.5857567117>
42. Kass RE, Raftery AE. Bayes factors. *J Amer Stat Assoc.* 1995; 90(430):773-795.
43. Smith SW. The Scientist and Engineer's Guide to Digital Signal Processing. 2nd ed. San Diego: California Technical Publishing; 1999.
44. Raftery AE. Bayes factors and BIC: Comment on "a critique of the Bayesian information criterion for model selection". *Sociol Methods Res.* 1999; 27(3):411-427.
45. Nobukawa S, Maruo T, and Nishimura H. Spontaneous Activity Modeling in Spiking Neural Systems with Log-normal Synaptic Weight Distribution. *SICE Annual Conference* 2014.