Design of the First Neuronal Connectomics Challenge: From Imaging to Connectivity

Isabelle Guyon, Demian Battaglia, Alice Guyon, Vincent Lemaire, Javier G. Orlandi,
Bisakha Ray, Mehreen Saeed, Jordi Soriano, Alexander Statnikov, Olav Stetter

Abstract— We are organizing a challenge to reverse engineer the structure of neuronal networks from patterns of activity recorded with calcium fluorescence imaging. Unraveling the brain structure at the neuronal level at a large scale is an important step in brain science, with many ramifications in the comprehension of animal and human intelligence and learning capabilities, as well as understanding and curing neuronal diseases and injuries. However, uncovering the anatomy of the brain by disentangling the neural wiring with its very fine and intertwined dendrites and axons, making both local and far reaching synapses, is a very arduous task: traditional methods of axonal tracing are tedious, difficult, and time consuming. This challenge proposes to approach the problem from a different angle, by reconstructing the effective connectivity of a neuronal network from observations of neuronal activity of thousands of neurons, which can be obtained with state-of-the-art fluorescence calcium imaging. To evaluate the effectiveness of proposed algorithms, we will use data obtained with a realistic simulator of real neurons for which we have ground truth of the neuronal connections. We produced simulated calcium imaging data, taking into account a model of fluorescence and light scattering. The task of the participants is to reconstruct a network of 1000 neurons from time series of neuronal activities obtained with this model. This challenge is part of the official selection of the WCCI 2014 competition program.

I. INTRODUCTION

Understanding the brain structure and some of its alterations caused by disease, is key to accompany research on the treatment of epilepsy, Alzeimer's disorder and other neuropathologies [1], as well as gaining understanding of the general functioning of the brain and its learning capabilities. At the neuronal level, recovering the exact wiring of the brain (connectome) including nearly 100 billion neurons, having on average 7000 synaptic connections to other neurons, is a daunting task. Traditional neuroanatomic methods of axonal tracing cannot scale up to very large networks. Could there be alternative methods for recovering neuronal network structures from patterns of neuronal activity? Today's cutting edge neurophysiology multi-electrode recording tools are capable of recording (and even stimulating) of the order of 100 neurons [2]. Optical imaging of neuronal activity using fluorescent calcium indicator molecules (calcium imaging)

provide a tool to increase the number of neurons recorded by three orders of magnitude (see [3], for a review). Recently, researchers have been able to record activity of the brain of a zebrafish embryo in 80% of its 100,000 neurons [4]. There is an ever increasing set of tools to image the brain with various fluorescent reporters [5]. The rapid improvements of optogenetics methods for shutting down or switching on neurons by shining laser light already allows stimulating up to 1000 neurons (see Refs. [6, 7] for reviews). Mathematical algorithms capable of discovering network structures are faced with the challenge of solving a new inverse problem: recover the neuronal network structure of a living system given the observation of a very large number of neurons, with the possibility of stimulating a subset of them. Monitoring changes in effective connectivity patterns of a network in action during behavior promises to advance our understanding of learning and intelligence.

Challenges in bioinformatics to reverse engineer gene networks such as DREAM[1] and "SBV improver"[2] do not have yet their parallel in neuroscience. We are organizing a challenge to stimulate the advancement of research on neuronal network structure reconstruction algorithms from neurophysiological data, including causal discovery methods. This challenge makes use of realistic simulations of real networks of neurons observed via calcium fluorescence recordings. The winning methods will be used to analyze real data recorded from *in vitro* cultures of neurons (Figure 1) and the proposed structure will be verified in part by experiments conducted in a wetlab.
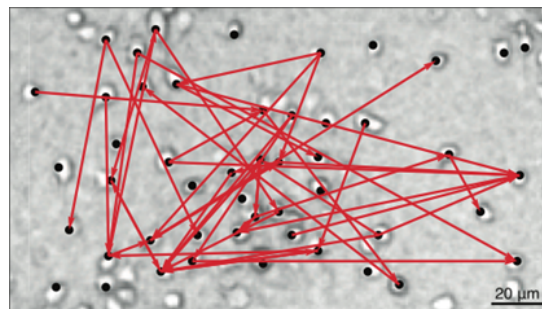

Fig. 1. Network reconstruction of *in vitro* neuronal cultures. White circular objects are neurons. Arrows are "effective" connections between neurons inferred from the analysis of their fluorescent calcium traces, as described in Ref. [8].

Isabelle Guyon is with ChaLearn, 955 Creston Road, Berkeley, CA94708, USA (email: guyon@chalearn.org); Demian Battaglia is with the University of Marseille, France; Alice Guyon is with the CNRS, Sophia Antipolis, France; Vincent Lemaire is with Orange, France; Javier Orlandi and Jordi Soriano are with the University of Barcelona, Spain; Mehreen Saeed is with FAST, National University of Computer and Emerging Sciences of Pakistan; Alexander Statnikov is with New York University, USA; Olav Stetter is with the Max Planck Institute for Dynamics and Self-Organization, Germany; Bisakha Ray is with New York University, USA.

[1] http://www.the-dream-project.org/
[2] https://www.sbvimprover.com/challenge-3/challenge

Our challenge makes available to the research community simulated fluorescence recordings of neuronal activity, in a format that does not require detailed understanding of neurobiology, and therefore will give an opportunity to computational intelligence scientists to make a contribution. Our simulated data [8] includes variations in dynamical regimes, as observed in living neuronal networks. In neuronal cultures, for instance, it is known that network activity may switch from a regime of synchronous activation (*bursts*) to another one of quieter, incoherent dynamics. This switching dynamics is a major challenge to network reconstruction, since directed "effective" connectivity can be very different during bursting and inter-burst phases [8]. It can bear a resemblance to the underlying anatomical synaptic connectivity only in selected dynamical regimes, in which causal influences reflect dominantly mono-synaptic interactions.

It is anticipated that our challenge will also help advance the state-of-the-art in inferring the structure of directed networks of units in general, beyond the neuronal connectomics applications. The goal of this challenge is to infer directed connections (synapses) of neuronal networks from patterns of neuronal activity. Such an oriented network can be thought of as a causal network. Neurons have complex temporal patterns of activity. The problem can therefore be thought of as a causal structure reconstruction problem from time series data. Other instances of such problems are found in genomics, climatology, epidemiology, engineering, and econometrics.

II. MOTIVATIONS

Although neuroanatomy is a very old science, connectomics is a relatively new, but fast emerging, field. If statistical knowledge about connectivity patterns has been long available (e.g. [9]), the aim of connectomics is to derive the detailed structure of whole large-scale neuronal systems. The first complete nervous system wiring diagram was accomplished with the 300 cell nervous system of a model organism: C. Elegans, a nematode worm, in the 1980's. It was deduced from reconstructions of electron micrographs of serial sections [10]. Partial connectomes for larger nervous systems including the fruit fly and the mouse have since been produced with a combination of neuroanatomical techniques [11]. But even the most advanced techniques for labeling individual neurons with distinguishable colors via a method called Brainbow [12] require a difficult tracing of neuron ramifications, and the resolution of optical microscopes is insufficient to reliably visualize synapses. Electron microscopy provides sufficient resolution, but no color coding, and yields voluminous amounts of data that is being analyzed very slowly with a combination of informatics methods and human labor. In 2012, a Citizen science project called EyeWire began attempting to crowdsource this task through an interactive game [11].

Another type of approach consists in reconstructing networks of interaction between neurons from patterns of activity to obtain an "effective topology". Inferring network topology from patterns of neuronal activity is not new. There have been active research efforts in the recent years to produce and analyze connectomic databases at the mesoscale and macroscale level, based on non-invasive imaging techniques of brain activity such as functional magnetic resonance imaging (fMRI), including the Human Connectome Project, led by the WU-Minn consortium. At the cellular level, the effort of reconstructing networks from neuronal activity can be traced back to a 2006 paper, already using the terminology of "effective topology", where only the in-degree of a neuron was estimated based on the simple logic "higher firing frequency = more inputs" [13]. The first major study using calcium imaging identified "hub neurons" with a simple cross-correlation approach [14]. This was followed by a major contribution by the Paninski group at Columbia also aiming at reconstructing the connectivity from calcium fluorescence imaging data [15, 16]. The idea was to first infer spike times as a Bayesian inverse problem, and then infer the *generalized linear model* (GLM) kernels — representing synaptic weights— of the supposed GLM for the neurons. Their work builds on a rigorous study of GLM models demonstrating reconstruction of spike data [17]. One criticism of the approach is that it was proven successful only with data generated with the same model as the model used for reconstruction. Real neurons are very diverse and an approach that is bound too tightly to a supposed model of the neurons may be plagued with artifacts. This motivated model-free network reconstruction techniques [8, 18].

Effective topology reconstruction is not *sensu stricto* the same concept as establishing a map of "anatomically correct" structural connections (actual synapses) because of a variety of reasons, including:

1) Some anatomical connections may be missed by the reconstruction because: (i) some synapses may be dormant (weak or inactive); (ii) some interactions may be invisible due to signal cancelation in feed-back circuits; (iii) two (or more) neurons may be overlapping and their signals merged.
2) Some spurious connections may be inferred because: (i) some effective connections may be relayed by invisible neurons and not correspond to actual synapses; (ii) some effective connections may result from artifacts of network reconstruction algorithms, which rely on data limited in time resolution; (iii) some effective connections may reflect real influences mediated by collective network properties, rather than by pairwise interactions.

However, algorithmic reconstruction approaches are far more scalable than anatomical axonal tracing and the hope is that, with improvements in imaging techniques, informatics tools, and theoretical understanding of observed neuronal dynamics, it will become possible to unravel connectomes of the nervous systems of large organisms. This motivates organizing this challenge.

We focus on reconstructing networks from so-called "observational data" which means the recording of cells let to evolve according to their own dynamics, without intervention of the experimentalists. This is in contrast with

"interventional data" obtained by stimulating neurons with external means (electrical, optical, or chemical). Forcing given network nodes to assume given states, disconnecting them from their natural influences from other network nodes, is the basis of the experimental methodology in causal inference. It is the only reliable way to unambiguously unravel causal relationships (directed network connections) from node activity. However, conducting proper interventional experiments is costly, technically difficult, and sometimes unethical or impossible. Optogenetics [6] is one of the most promising methods because it offers the possibility of intervening simultaneously on hundreds of neurons. Optogenetic paradigms use genetic techniques to induce the expression of light-activated ion channels into a living organism such that focused shining of light can trigger action potentials in the targeted neurons [19]. But, the apparatus is complicated and expensive. Additionally, intervening on neurons puts stress on them and cannot be done extensively without damaging them. Moreover, intervention on the network can change it by modifying the strengh of the connections by plasticity processes (such as long term potentiation or depression). For these reasons, algorithms that unravel neuronal network structures from purely observational data will remain important as standalone methods or in conjunction with interventions to prepare them or guide them. We believe that this is feasible, particularly because from the perspective of complex dynamical systems the distinction between "observational" and "interventional" data (as made in the causal discovery literature) is blurred by the fact that neuronal networks generate in a sense their own self-organized set of experimental interventions by spontaneously bursting. Nonetheless, it is our intention to also investigate how the methods developed by the participants using observational data can be validated or complemented using interventional data. This could lay the basis for a new challenge of neuronal structure reconstruction blending observational and interventional data.

III. BRIEF OVERVIEW OF NETWORK RECONSTRUCTION METHODS

There is a rich literature on methods of network structure reconstruction from observed time series, not only stemming from research in neuroscience, but also machine learning and econometrics, which have fueled the area of causal inference from temporal data with numerous novel techniques [20]. Briefly, despite the 20th century rise to prominence of statistics, initially intended to resolve causal quandaries in agricultural and industrial process refinement, the field of statistical causal inference is relatively young. Although its pioneers have received wide praise (Clive Granger receiving the Nobel Prize and Judea Pearl receiving the ACM Turing Award) the methods they have developed are not yet widely known and are still subject to refinement. Even though one of the least controversial necessary criterion of establishing a cause-effect relationship is temporal precedence, many causal inference algorithms do not require time information and establish possible causal relations among observations

on other grounds, based on conditional independence testing [21], or, more recently, based on statistics of the joint distribution of pairs of variables[3]. The work of Clive Granger, built upon the 20th century development of time series modeling in engineering and economics, with some input from physiology, lead to a framework which admittedly does not allow us to identify causality unequivocally, but has received a lot of attention because of the simplicity of the method and practical successes obtained in econometrics and neuroscience [20].

The basic idea behind Granger causality to test whether observations of time series of two variables A and B are symptomatic of an underlying process "A causes B" rather than "B causes A", is to fit various predictive models A(present time) and B(present time) as a function of A(past times) and B(past times). Clues are obtained if A can be predicted better from past values of A and B rather than from A itself, but B cannot be predicted better from past values of A and B rather than from B itself. Numerous improved methods have been derived, incorporating, for instance, frequency domain analysis in lieu of time domain analyses [22]. One recent idea is to add contemporaneous values of B to predict A and vice versa to take into account instantaneous causal effect, due for instance to insufficient time resolution [23]. In neuroscience, simple linear autoregressive (AR) models underlying Granger causality do not capture well the complexity of neuronal signals. A non-linear version of Granger causality called Transfer Entropy [24], which reduces to Granger causality for simple AR models [25] is gaining popularity [26, 8].

It is well known that causal relationships can be confounded: the fact that A and B are correlated or co-variant does not mean that A and B are in a causal relationship: there may be a third common cause C. A typical way of alleviating the problem of false positive causal relationships is to perform conditional independence tests. If A and B are independent given C, the existence of a direct causal relationship between A and B is ruled out and the remaining possibilities are $A \rightarrow C \rightarrow B$, $B \rightarrow C \rightarrow A$ or $A \leftarrow C \rightarrow B$. However, one of the greatest challenges that network structure reconstruction methods have to face is the curse of dimensionality. With the explosion of the number of variables it becomes quickly impractical to reliably conduct conditional independence tests, which require a number of samples exponential in the number of variables jointly tested. Moreover, it is practically never possible to record all the neurons of a network (with fluorescence methods for instance, some neurons may be invisible or not marked). Hence it is likely that one would violate the assumption of "causal sufficiency" (namely that no neuron that influences two observed neurons is unobserved), which often made by methods relying on conditional independence tests.

Another approach, which is not limited to statistics of pairs of variables, is to use score-based methods, by performing

[3]For a bibliography, see http://webdav.tuebingen.mpg.de/causality/.

a search in the space of all possible architectures, guided by an objective function assessing the goodness of signal reconstruction (possibly penalized to favor sparse connectivity). Such methods include Bayesian approaches such as Dynamical Causal Modeling (DCM) [27], which compare data generating models formulated in terms of differential equations (modeling the dynamics of hidden states in the nodes of a probabilistic graphical model), where conditional dependencies are parameterized in terms of directed effective connections. Other related methods include L1 and/or L2 penalized regression methods [28].

Another possible remedy to the problem of confounding, which attacks the problem of the curse of dimensionality from a different angle, is to recourse to conditioning on the average activity of the population of nearby neurons rather than on combinations of single neurons, then rely only on statistics of joint activity of pairs of neurons [8, 29]. The promising results of the "Cause-Effect Pairs" Challenge that we recently organized demonstrates that we can go a long way to infer causal relationships from pairs of variables, without conditioning on other variables.[4] The Area under the ROC Curve (AUC) of the top ranking participants exceeded 0.8, on a combination of real and artificial data (an AUC of 0.5 is obtained for random guesses and the perfect score is 1). The methods used by the participants are model-free. They exploit features of the joint distribution of two variables, some of which are derived from information theoretic principles. The predictions are made with pattern recognition algorithms trained on thousands of examples of cause-effect pairs. The challenge was limited to data samples not time-ordered. For this reason, we are planning a new Cause-Effect Pairs Challenge for Time-Series Data, which will be held in conjunction with the challenge described in this paper. The hope is that such techniques could be applied with success to our new network structure reconstruction challenge, possibly even reaching better performance after preprocessing by conditioning on average activity of nearby neurons and by exploiting the availability of time ordering of samples.

IV. Challenge design

Each scientific discipline has its favorite approach, not necessarily reflecting better match of techniques to domains, but rather historical tradition. Standard benchmarks are needed to foster scientific progress, but the design of a good benchmark, which is not biased in favor a particular model or approach, is not trivial. Our team includes neuroscientists and seasoned challenge organizers. This allowed us to define a number of tasks geared towards deriving conclusive and statistically significant results.

A. Protocol sketch

The goal of the challenge is to predict the directed connections (synapses) of a neuronal network, given calcium fluorescence imaging recordings of the activity of each

[4]http://clopinet.com/isabelle/Projects/NIPS2013/

neuron in this network. We devised tasks of progressive difficulty. In this first challenge organized for WCCI 2014, the participants' task will be to reconstruct the structure of a medium size network (1000 neurons). We are planning a follow-up challenge on a much larger network of 80,000 neurons. The motivation behind having a first challenge on a smaller size network is to lower the barrier of entry. The follow up challenge will then force the participants to scale their method up to address a more computationally challenging task.

During a development period, the participants will experiment with sample data generated from several "training networks" of sizes 50-500 neurons, spanning a variety of topologies and dynamical behaviors, using a realistic simulator of real neuronal networks. The architecture of the "training networks" will be revealed to the participants. Other networks will be used to generate "validation" and "test" data, for which the truth values of the network connections will be hidden to the participants. The goal of the challenge is to predict the unknown network connections. The validation data will be used for practice purposes during the development period. The final ranking and the selection of the winners will be based on the test data. The workflow of data generation and challenge participation is outlined is (Figure 2). More details on the data are provided in Section V.

Submissions are made via a web-based platform provided by Kaggle.com. During the development period, the participants will make practice submissions using the validation data. On-line immediate feed-back on performance will be provided and posted on a public leaderboard. This should stimulate participation. The submissions using final data will also be made via the on-line platform. The results on final data will be kept on a private leaderboard visible only to the organizers until the challenge ends and the final rankings are revealed.

We provide a brief tutorial, baseline software based on the Transfer Entropy (TE) method [8], and a sample submission produced with the software.

B. Evaluation metrics

In a network of n neurons, there are n^2 potential directed connections (including self-connections). For each "ordered pair" (i, j) of neurons among n^2 possible combinations, the problem is to determine whether there is a connection $i \to j$. We can there formulate the network structure reconstruction problem as a classification problem:

- Classify (i, j) as 1 (positive class) if neuron i has a connection to neuron j.
- Classify (i, j) as -1 (negative class) otherwise.

Note that a "connection" refers to an effective connection, which may correspond to one or more synapses (physical connections), since a neuron commonly makes several synapses with another neuron.

The score provided by the participants for each ordered pair of neurons will be understood as a discriminant value.
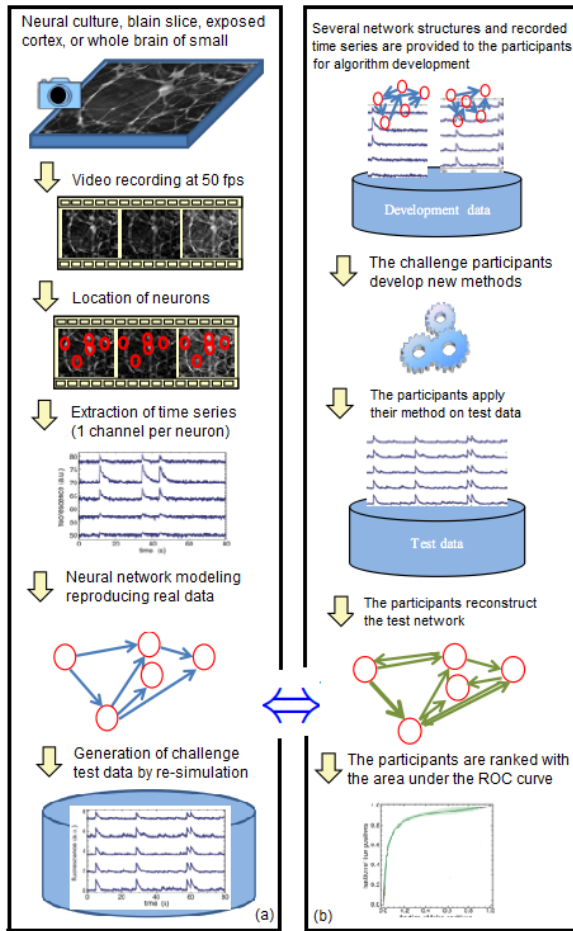
Fig. 2. Data production and challenge workflows: (a) Data production, (b) Challenge, (⇔ predictions are compared to the true connections).

The classification performance will be evaluated with the area under the ROC curve (AUC). The AUC is equivalent to the area under the curve obtained by plotting sensitivity (success rate of the positive class) against specificity (success rate of the negative class), by varying a threshold on the prediction values to determine the classification result. The AUC on test data will be used for ranking the participants.

Approximate error bars can be computed to evaluate the test set size necessary to get good performance estimation [30]. For a test network of $n = 1000$ neurons, a decision on $N = n^2 = 10^6$ potential connections must be made. N is the number of "examples" to be classified. However, we have sparsely connected networks, so only about $f = 10\%$ of the examples belong to the positive class. In reference [31], we adapted the error bar formula to the case of imbalanced classes: $E \simeq \sqrt{p(1-p)/(2fN)}$. Where p is the anticipated balanced success rate (BAC), the average of the sensitivity and the specificity. The BAC is equivalent to the AUC for the optimal sensitivity and specificity tradeoff. Based on the preliminary work of our collaborators [8], we anticipate a success rate of the order of BAC=0.9. This leads to an error bar of: $E \simeq \sqrt{0.9(1-0.9)/(2 \times 0.1 \times 10^6)} \simeq 10^{-3}$.

Having a global figure of merit such as the AUC is convenient to summarize results and rank the participants.

However, it does not provide a fine analysis of the strengths and weaknesses of the methods. In post challenge analyzes, we will make use of additional metrics. Specifically, we will threshold the scores of the participants at the level of the real probability of connection known from the literature (taking the top 12% as reconstructed links). We will then compute the fraction of correctly identified motifs of triples. Of particular interest are the triples $\{A, B, C\}$ with true structure $A \rightarrow B$ and $A \rightarrow C$ with no link $B \rightarrow C$ (co-causation) and the triples with true structure $A \rightarrow B$ and $B \rightarrow C$ and no link $A \rightarrow C$ (indirect causation). The study of motifs is an important step towards understanding the functionality of larger circuits [32]. We may also include higher order motifs of interest based on a literature survey.

Our test networks are designed such that sub-networks have various degrees of node clustering. It will also be interesting to study the correlation between clustering coefficients in the sub-networks that generated the data and in the reconstructed connectivity.

C. Prizes

Anybody complying with the rules of the challenge, with the exception of the organizers, will be eligible to enter the challenge. However, the participants barred from receiving compensations from US sources by US export regulations will be barred from winning prizes. The participants will be allowed to remain anonymous, except if they want to claim a prize. The will have to be organized in mutually exclusive teams. The participants will make only one final submission for the test network. After the challenge is over and before the prizes are attributed, (i) All participants will be asked to fill out fact sheets describing their methods if they want to be ranked in the challenge; (ii) To qualify for prizes, the top three ranking participants will have to make their source code available under a popular OSI license (http://opensource.org/licenses). The organizers will verify the reproducibility of the results by executing the code using the datasets of the challenge. (iii) To qualify for prizes, the top three ranking participants will have to contribute a paper to the proceedings.

We will attribute 3000 USD in prizes to the winners, donated by ChaLearn:

- First place: 500 USD and 1000 USD travel award.
- Second place: 250 USD and 750 USD travel award.
- Third place: 100 USD and 400 USD travel award.

The travel awards may be used for one of the workshops organized in conjunction with the challenge. The award money will be granted in reimbursement of expenses including airfare, ground transportation, hotel, or workshop registration. Reimbursement is conditioned on (i) attending the workshop, (ii) making an oral presentation of the methods used in the challenge, and (iii) presenting original receipts and boarding passes.

V. DESCRIPTION OF THE DATA AND SUBMISSION FORMAT

We are using simulated data supplied by a realistic simulator of real neurons and a model of calcium fluorescence

recording, providing data closely resembling real recordings of cultured neurons, while providing unequivocal ground truth of synaptic connections. The data were generated using a simulator extensively studied and validated [8]. The dynamic behavior of the neurons was adjusted to reproduce collective properties of real networks of cultured neurons. The model also simulates limitations and defects of the imaging technology (calcium fluorescence): limited time resolution (not allowing to separate individual spikes) and light scattering artifacts (by which the activity of given neuron influences the measurements of nearby neurons). Reconstructing networks from artificially generated data can be thought of as a futile mathematical exercise. This is why we have moved away from reconstructing data generated by simple models such as Bayesian networks or Structural Equation Models making over-simplifying assumptions of linearity and Gaussianity. The data simulator that we are using integrates realistic scenarios at three levels [8]:

- **Network structure**: At the network architecture level we use topologies as close as possible to natural topologies. We use connectivity models taking into account neuron proximity and enforcing a realistic degree of node clustering. We hide a certain number of neurons, accounting for the fact that only about 80% of neurons are visible in real experimental data. This makes the problem of structure reconstruction considerably more difficult.
- **Neuron models:** We use leaky integrate-and-fire models of spiking neurons, as implemented by the NEST simulator (Neuronal Simulation Technology, see http://www.nest-initiative.org/). The dynamic regimes reproduce faithfully experimentally observed neuronal recordings. However, in this first challenge, we consider only excitatory synapses as an experimentally meaningful simplification and further restrict ourselves to constant value synapses. More complexity will be added in the follow up challenge.
- **Fluorescence model:** We simulate the calcium fluorescence time series taking into account time averaging and light scattering effects. The signals generated closely resemble real data.

For development purposes, we will provide simulated neuronal recordings for various network sizes in the range 50-500 neurons, generated under varying conditions of connectivity and exhibiting a range of dynamical behavior. The truth values of the network connection will be supplied for such "training networks", which the participants will use to make predictions on the submission platforms for practice purposes.

Our final test set will consist of data from a network of 1000 real neurons simulating cultures neurons. The architecture of the network will be designed such that sub-networks exhibit a range of node clustering to test the robustness of the reconstruction algorithms. The simulated neuronal recordings will be obtained with the same neuronal simulator used to generate the training data. The truth values of the network

connections will not be provided to the participants, since the goal of the challenge is to uncover them. The validation set will be contructed similarly as the test set, but with a different simulated network.

We are providing to the participants data consisting of simultaneous recordings of neurons in networks of:

- n neurons (n varying from 50 to 1000)
- sparsely connected (with an average probability of connection of 0.12)
- coming as tables with neurons in columns and time samples in lines
- for each network we supply 180k time samples (1h of recording at 50fps)
- we supply the position of the neuron in two dimensions (emulating a neuronal culture).

To generate a rich set of training networks, we varied the parameters of the network construction and dynamical simulation. The data include:

- 2 networks with the same parameter sets as the validation and test data,
- a network with higher/lower sparsity,
- a network with higher/lower rate of external inputs,
- a network with higher/lower camera noise,
- a network with higher/lower maximum clustering index.

The contestants will return in a text file a vector of predictions suitable to be scored by the competition platforms. Each line will be associated with one ordered pair of neurons (i, j) and will provide a score that neuron i influences neuron j via a direct synaptic connection.

VI. Baseline software and baseline results

We prepared a software kit written in Matlab to help the participants getting started, which we are making publicly available. The kit includes:

- Data visualization software showing a movie of simulated calcium fluorescence.
- Network visualization software.
- Time series visualization software.
- Baseline network reconstruction software.

The baseline network reconstruction method, which we implemented, is described in details in [8]. It is based on Generalized Transfer Entropy (GTE), which is an extension of Transfer Entropy first introduced by Schreiber [24], a measure that quantifies predictive information flow between stationary systems evolving in time. It is given by the Kullback-Leibler divergence between two models of a given time series, conditioned on a given dynamical state of the system, which in the case of fluorescence signals corresponds to the population average. Transfer Entropy captures linear and non-linear interactions between any pair of neurons in the network and is model-free, i.e., it does not require any a priori knowledge on the type of interaction between neurons.

Apart from GTE, we have also provided the implementation of cross correlation and two information gain (IG) measures based on entropy and gini for network reconstruction. Cross correlation gives best results when there are

zero time delays, which reduces it to a simple correlation coefficient measure. Hence, all these methods treat the data as independent instances/points in space instead of time series data.

Another module that we have added to our software kit is a supervised learner, which extracts features from a network whose ground truth values are known and builds a simple linear classifier for learning whether a connection is present between two neurons or not. Currently, the features extracted are GTE, correlation, information gain using gini and information gain using entropy.

We ran the sample code on the six small networks that we have made available to the participants. Each network contains 100 neurons and their ground truth values are also made publicly available. Table I gives the results on all of the afore mentioned methods for all the six networks. As we can see GTE outperforms all other methods for almost all the networks.

TABLE I

AUC FOR DIFFERENT METHODS WITH 100 NEURONS

Small networks	GTE	Correlation Coefficient	IG (gini)	IG (entropy)	Trained Predictor
network 1	0.643	0.5224	0.5217	0.5246	0.5384
network 2	0.7036	0.5789	0.5854	0.5789	0.6011
network 3	0.7848	0.6995	0.7153	0.6964	0.7514
network 4	0.8280	0.8081	0.7790	0.7648	0.8290
network 5	0.8749	0.7061	0.6186	0.7062	0.8375
network 6	0.9156	0.8816	0.8382	0.8563	0.9206

Sample code was generated in C++. The conditioning level (CL) for correlation, cross-correlation, and Granger causality was 0.25 and for GTE was 0.10 and 0.15. Table II gives the results for all of the afore mentioned methods for all the six networks.

TABLE II

AUC FOR DIFFERENT METHODS WITH 100 NEURONS

Small networks	Correlation Coefficient	Cross Correlation	Granger Causality	GTE CL=0.10	GTE CL=0.15
network 1	0.6155	0.5591	0.4815	0.6451	0.6733
network 2	0.6602	0.5674	0.4877	0.7339	0.7223
network 3	0.7453	0.5862	0.4831	0.8115	0.7993
network 4	0.8165	0.5811	0.4924	0.8646	0.8479
network 5	0.8315	0.5725	0.4775	0.8857	0.8796
network 6	0.9242	0.7632	0.4818	0.9396	0.9379

Results obtained using Python code for correlation coefficient without and with discretization at -10, 10, and 20 and number of bins = 3 are summarized in Table III.

Using the sample code, we generated two benchmark submissions for the challenge on the validation set of 1000 neurons. One based on the correlation coefficient, giving an AUC of 0.873 and one using GTE which gives a slightly higher AUC of 0.880. The correlation coefficient takes a couple of minutes to run on a regular laptop computer. Although the ROC curve yields a relatively high AUC value, there's still a lot of room for improvement. Given that the real network is sparse, only the beginning of the ROC curve

TABLE III

AUC FOR DIFFERENT METHODS ON THE 6 SMALL NETWORKS WITH 100 NEURONS AND TWO WITH 1000 NEURONS

Small networks	Correlation Coefficient without discretization	Correlation Coefficent with discretization	Correlation Coefficent bins = 3
network 1	0.5665	0.6155	0.5357
network 2	0.5734	0.6602	0.5935
network 3	0.5923	0.7453	0.6971
network 4	0.5874	0.8165	0.7679
network 5	0.5784	0.8315	0.7066
network 6	0.7634	0.9242	0.8573
normal-1	0.6816	0.8808	0.6752
normal-2	0.6995	0.8764	0.7577

is really interesting, usually up to 10% false positives where currently only a 70% of the true positives are detected. After around a week of the launch of the challenge, participants had already crossed the benchmark scores and reached AUC above 0.90. We expect the participants to reach much higher values, hopefully above a 90% of true positives at the 10% false positive mark, corresponding to an AUC over 0.95.

VII. CONCLUSION

This challenge will push the computational intelligence community to scale up methods of network structure discovery. We will provide baseline methods and tutorial material to induce participation. Our proposed setting will allow computational intelligence scientists to make a contribution without having an in depth understanding of neuroscience. In post challenge analyses, the structures hypothesized from new real data by the winning methods will be verified by actual wetlab experiments. If accepted, this paper will be updated with preliminary challenge results before its publication. The results will be presented at the WCCI 2014 conference, where the winners will be awarded their prizes. The challenge platform will remain open beyond the end of the challenge as an on-going benchmark and the datasets will be made publicly available.

REFERENCES

[1] D. S. Bassett and E. T. Bullmore, "Human brain networks in health and disease." *Current opinion in neurology*, vol. 22, no. 4, pp. 340–7, Aug. 2009.

[2] F. O. Morin, Y. Takamura, and E. Tamiya, "Investigating neuronal activity with planar microelectrode arrays: achievements and new perspectives." *Journal of*

bioscience and bioengineering, vol. 100, no. 2, pp. 131–43, Aug. 2005.

[3] C. Grienberger and A. Konnerth, "Imaging calcium in neurons," *Neuron*, vol. 73, pp. 862–885, 2012.

[4] M. B. Ahrens, M. B. Orger, D. N. Robson, J. M. Li, and P. J. Keller, "Whole-brain functional imaging at cellular resolution using light-sheet microscopy," *Nature Methods*, vol. 10, no. 5, p. 413420, 2013.

[5] J. Patti and E. Y. Isacoff, "Measuring membrane voltage with fluorescent proteins," *Cold Spring Harbor Protocols*, vol. 2013, no. 7, pp. 606–613, 2013.

[6] L. Fenno, O. Yizhar, and K. Deisseroth, "The development and application of optogenetics." *Annual review of neuroscience*, vol. 34, pp. 389–412, Jan. 2011.

[7] A. Reiner and E. Y. Isacoff, "The Brain Prize 2013: the optogenetics revolution," *Trends Neurosci*, vol. 36, no. 10, pp. 557–560, 2013.

[8] O. Stetter, D. Battaglia, J. Soriano, and T. Geisel, "Model-free reconstruction of excitatory neuronal connectivity from calcium imaging signals." *PLoS Computational Biology*, vol. 8, no. 8, 2012.

[9] V. Braitenberg and A. Schüz, *Anatomy of the Cortex: Statistics and Geometry (Studies of Brain Function)*. Springer, 1991.

[10] J. White, E. Southgate, J. N. Thomson, and S. Brenner, "The structure of the nervous system of the nematode c. elegans," *Philosophical transactions Royal Society London*, vol. 314, pp. 1–340, 1986.

[11] S. Seung, *Connectome: How the Brain's Wiring Makes Us who We are*, ser. A Mariner Book. Houghton Mifflin Harcourt, 2012.

[12] J. Livet, T. A. Weissman, H. Kang, R. W. Draft, J. Lu, R. A. Bennis, J. R. Sanes, and J. W. Lichtman, "Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system." *Nature*, vol. 450, no. 7166, pp. 56–62, 2007.

[13] D. Eytan and S. Marom, "Dynamics and effective topology underlying synchronization in networks of cortical neurons," *J Neurosci*, vol. 26, no. 33, pp. 8465–76+, 2006.

[14] P. Bonifazi, M. Goldin, M. A. Picardo, I. Jorquera, A. Cattani, G. Bianconi, A. Represa, Y. Ben-Ari, and R. Cossart, "GABAergic hub neurons orchestrate synchrony in developing hippocampal networks." vol. 326, no. 5958, pp. 1419–1424, 2009.

[15] J. T. Vogelstein, B. O. Watson, A. M. Packer, R. Yuste, B. Jedynak, and L. Paninski, "Spike inference from calcium imaging using sequential Monte Carlo methods." *Biophysical journal*, vol. 97, no. 2, pp. 636–655, 2009.

[16] Y. Mishchenko, J. T. Vogelstein, and L. Paninski, "A Bayesian approach for inferring neuronal connectivity from calcium fluorescent imaging data," *The Annals of Applied Statistics*, vol. 5, no. 2B, pp. 1229–1261, June 2011.

[17] W. Truccolo, U. T. Eden, M. R. Fellows, J. P. Donoghue, and E. N. Brown, "A Point Process Framework for Relating Neural Spiking Activity to Spiking History, Neural Ensemble, and Extrinsic Covariate Effects," *Journal of Neurophysiology*, vol. 93, no. 2, pp. 1074–1089, 2005.

[18] J. G. Orlandi, O. Stetter, J. Soriano, T. Geisel, and D. Battaglia, "Transfer entropy reconstruction and labeling of neuronal connections from simulated calcium imaging," 2013.

[19] T. Knöpfel, "Genetically encoded optical indicators for the analysis of neuronal circuits." *Nature reviews. Neuroscience*, vol. 13, no. 10, pp. 687–700, Oct. 2012.

[20] F. Popescu and I. Guyon, Eds., *Causality in Time Series*, ser. Challenges in Machine Learning. Microtome Publishing, 2013, vol. 5.

[21] J. Pearl, *Causality: Models, Reasoning, and Inference*. Cambridge University Press, 2000.

[22] G. Nolte and K. R. Mueller, "Localizing and estimating causal relations of interacting brain rhythms," *Frontiers in Human Neuroscience*, vol. 4, no. 209, 2010.

[23] A. Moneta and P. Spirtes, "Graph-based search procedure for vector autoregressive models," Laboratory of Economics and Management (LEM), Sant'Anna School of Advanced Studies, Pisa, Italy," LEM Papers Series, 2005.

[24] T. Schreiber, "Measuring information transfer," *Phys. Rev. Lett.*, vol. 85, pp. 461–464, Jul 2000.

[25] L. Barnett, A. B. Barrett, and A. K. Seth, "Granger Causality and Transfer Entropy Are Equivalent for Gaussian Variables," *Physical Review letters*, vol. 103, p. 238701, 2009.

[26] D. Battaglia, A. Witt, F. Wolf, and T. Geisel, "Dynamic effective connectivity of inter-areal brain circuits." *PLoS computational biology*, vol. 8, no. 3, Mar. 2012.

[27] K. J. Friston, L. Harrison, and W. Penny, "Dynamic causal modelling," *NeuroImage*, vol. 19, no. 4, pp. 1273–1302, Aug. 2003.

[28] S. Ryali, T. Chen, K. Supekar, and V. Menon, "Estimation of functional connectivity in fmri data using stability selection-based sparse partial correlation with elastic net penalty," *NeuroImage*, vol. 59, no. 4, pp. 3852 – 3861, 2012.

[29] P. N. Steinmetz and C. Thorp, "Testing for effects of different stimuli on neuronal firing relative to background activity," *Journal of Neural Engineering*, vol. 10, no. 5, 2013.

[30] I. Guyon, J. Makhoul, R. Schwartz, and V. Vapnik, "What size test set gives good error rate estimates?" *PAMI*, vol. 20, no. 1, pp. 52–64, 1998.

[31] I. Guyon, A. Saffari, G. Dror, and G. Cawley, "Agnostic learning vs. prior knowledge challenge," in *IEEE/INNS conference IJCNN 2007*, Orlando, Florida, August 12-17 2007.

[32] C. I. Bargmann and E. Marder, "From the connectome to brain function," *Nature Methods*, vol. 10, no. 6, pp. 483–490, May 2013.