CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Nature Precedings

Strategies for recovering exact structure of neural circuits with broadly targeted fluorescent connectivity probes

Yuriy Mishchenko

Columbia University, Department of Statistics, 1255 Amsterdam Ave, New York, New York, 10027

Abstract

We present a framework for reconstructing structure of complete neural circuits in the brain using collections of independent measurements of connectivity performed with existing anatomical or functional fluorescent probes, and designed to provide complementary information about neural circuit's structure by targeting slightly different its parts either in deterministic or stochastic succession. We discuss specific implementation of this procedure using synaptic fluorescent marker GRASP and Cre/Lox system Brainbow to collect ensemble of observations of the sets of synapses between stochastically labeled samples of neurons. By representing such measurements mathematically as weak constraints on circuit's connectivity matrix and by solving a constrained optimization problem, we are able to exactly deduce the wiring diagram in *C. Elegans* in an in-silico experiment from only ~10,000 measurements. This offers possibility for routinely reconstructing complete connectivity in smaller organisms, such as *C. Elegans*, using exclusively light microscopy instruments over the span of single weeks.

Nature Precedings : hdl:10101/npre.2009.2669.2 : Posted 2 Mar 2009

Nature Precedings : hdl:10101/npre.2009.2669.2 : Posted 2 Mar 2009

The problem of acquiring detailed and complete structure of neural circuits in the brain is the fundamental century-old challenge of neuroscience. Accurate knowledge of such structure is the key to testing many theoretical propositions about processing of information in the brain[1-3], but also such knowledge is essential for seeking the principles of brain organization experimentally[4-7]. Until now, the only approach with the potential to produce such data was neural circuit reconstruction using serial section electron microscopy (ssEM)[8-15]. Although the only existing wiring diagram for a complete neural circuit was indeed obtained with ssEM (i.e. the circuit of about 300 neurons and 6000 synapses in *C. Elegans*[8-10, 15]), this technique is known to be extremely labor intensive, slow and vulnerable to errors. The abovementioned circuit in *C. Elegans* took over a decade to complete, and yet its recent revision led to changes in nearly 10% of synapses affecting as many as 40% of all neurons[8, 9].

In this paper we argue that recent developments in fluorescent connectivity probes[16-22], genetic targeting techniques[23-28] and mathematical sparse data analysis[29-31] permit reconstructions of complete neural circuits using only light microscopy tools, thus, providing for a much cheaper, faster and more reliable alternative to ssEM. In particular, we show that in smaller organisms, such as *C. Elegans*, reconstructions of their complete wiring diagram may be acquired with such tools routinely over the span of single weeks. Unprecedented ability to produce such reconstructions fast, routinely and under variety of conditions should prove invaluable for neuroscience. E.g., by reconstructing a number of same neural circuit instances in a smaller model organism one may be able to directly observe conserved and variable structures there and relate them to behavior. Likewise, by reconstructing circuit in different mutants of the same organism one may be able to directly identify structural signatures of behavior anomalies, etc.

Below we present a systematic framework for such reconstructions based on collecting large numbers of simple measurements of connectivity performed with existing anatomical or functional fluorescent probes. Different such probes should be designed to target different parts of studied circuit in either deterministic or stochastic succession, thus, offering complementary information about circuit's structure. We recognize that each measurement typically will only provide limited amount of information; yet, strong statements about circuit's connectivity may be made from a collection of such measurements by systematically representing them as mathematical constraints on the connectivity matrix and then resolving these "weak" constraints to identify the most likely circuit configuration consistent with all observations. We specifically describe how such procedure may be implemented using Cre/Lox system Brainbow[24] and fluorescent synaptic marker GRASP[19], and show that in organism as large as *C. Elegans* complete reconstruction of its wiring diagram will require only ~10,000 measurements, entirely within reach of existing technologies. Described general paradigm may be applied with variety of other connectivity probes including functional probes with calcium[17] or voltage[16] sensitive dyes, circuit mapping with glutamate uncaging[18] or light-gated ion channel ChR2[21], transsynaptic viruses[22], etc., and genetic targeting techniques such as mosaic analysis MARCM[23] or MADM[25] and libraries of Gal4-lines[27].

Nature Precedings : hdl:10101/npre.2009.2669.2 : Posted 2 Mar 2009

RESULTS

Reconstruction of neural circuits using broad fluorescent connectivity probes

Our approach is based on accumulation of a large number of simple measurements of connectivity in a neural circuit, possibly performed in different animals, each represented as a constraint on its connectivity matrix. Mathematically, such constraints are described with a probe-dependent posterior likelihood function $P(C \mid O_i)$ for the connectivity matrix C given one observation O_i. E.g., below we discuss circuit reconstruction procedure based on observation of the set of synapses between two general populations of neurons, labeled with fluorescent synaptic marker GRASP, and the measurement of their combined size represented by combined fluorescence from all observed synaptic puncta (we assume that fluorescence from a labeled puncta may be viewed as a proxy for synapse's physical size). Each such measurement constrains C a bit – i.e. it specifies how many synapses may exist between two given populations of neurons. However, individually each such constraint is very weak – there may be many circuit configurations all consistent with single observation O_i. Nevertheless, by combining information from different measurements the uncertainty in C may be dramatically reduced. In particular, if sufficient number of measurements may be acquired, the connectivity matrix may be identified exactly by finding the maximum of full posterior likelihood on C given the measurements $\{O_i\}$, $P(C \mid \{O_i\}) \sim \prod P(O_i \mid C) P(C)$.

The central piece of our framework, therefore, is systematic collation of the information obtained about a neural circuit from multiple experiments. Because different such experiments may need to be carried out in different animals, essential to this framework is the condition that structures thus observed in different animals may be explicitly related. This implies description of the neural circuit in terms of certain units and connectivity among them that are *stereotypical* from animal to animal, i.e. such that may be separately identified in individual animals and corresponded across different animals. E.g., in simpler animals, such as *C. Elegans*, individual neurons are known to be identifiable. (In larger animals, however, identification of such units is not straightforward, although certain alternatives may exist[26].) One may take units to be individual neurons or, more generally, they may be defined as genetically, anatomically or functionally identified neural populations. Connections between units may be characterized by counts or combined size of synapses formed by their neurons, strength of post-synaptic excitation in one unit given activity in the other, correlation in neural activity, etc. In all above cases our framework will be applicable; the aim of the reconstruction being the stereotypical matrix of connection weights between such units.

We further focus on the use of GRASP and Brainbow genetic constructs for implementation of the above general paradigm. GRASP is recently developed genetic construct which uses so called split-GFP to fluorescently label synapses between two cells[19]. Split-GFP is a molecule of Green Fluorescent Protein (GFP) broken into two fragments which themselves do not fluoresce but recover fluorescence if brought into proximity and allowed to reassemble. In GRASP, split-GFP is genetically tailored to endogenous synaptic proteins and separately expressed in two cells. At the location of synaptic junctions such split-GFP reassemble, thus, rendering selected synapses visible with light microscope[19]. Brainbow is a genetic construct which uses Cre/Lox recombination

system to stochastically express a set of fluorescent proteins (FP) in different cells[24]. In one Brainbow version, FP sequences in genome are flanked with inversely oriented loxP-sites and exposed to action of Cre-recombinase. With Cre, such sequences begin to randomly flip until a distribution of direct and inverted orientations is established. Only those cells where FP sequence assumed direct orientation are able to transcribe it. Livet et al. report that each FP typically expresses in random 50% of cells[24]. We use this Cre/Lox system with GRASP to express GRASP stochastically (Figure 1). Each stochastic such expression pattern, therefore, allows to fluorescently label a collection of synaptic puncta between two sets of pre- and post-synaptic neurons. To identify which cells participate in each pattern we need an additional helper construct. For that, we may bundle split-GFP sequence with a sequence for nuclei-targeted FP inside loxP bracket (Figure 1A), thus, making sure to always express split-GFP together with associated nuclei-targeted FP and allowing identification of expression patterns using thus labeled cell-nuclei.

With the above construct, therefore, one is able to produce a variety of GRASP expression patterns within single genetic line and, thus, observe synapses corresponding to different random samples of pre- and post-synaptic neurons. Because all synaptic puncta still have the same color, only the total count or combined size of observed puncta in each such single animal constitutes the available information (assuming no prior knowledge about position or size of synapses). The measurement O_i, therefore, is such total count or size of all visible puncta from one animal. Such measurements O_i along with the identities of corresponding pre- and post-synaptic cells are the pieces of information we propose to collect. From a sufficiently large collection of these, as we show below, complete wiring diagram may be recovered.

Nature Precedings : hdl:10101/npre.2009.2669.2 : Posted 2 Mar 2009

Determining circuit structure from collection of Brainbow-guided GRASP probes

To understand how complete wiring diagram may be reconstructed from the above data consider the following scenario. Let O be a single observation of the total fluorescence from GRASP in one animal. On average $\mathrm{E}[O] \approx Sf^2$, where S is the total size of all synapses in the circuit and f=0.5 is the probability for one neuron to express GRASP in one trial with Brainbow. Now, consider two neurons A and B that are strongly connected with each other, and let's pay attention to the trials where neurons A and B simultaneously express GRASP. If we are careful, we will notice that in such trials measurements O will be on average higher than $\mathrm{E}[O] \approx Sf^2$. This is because in each such trial a very large synapse between A and B will be invariably observed unlike in the rest of the measurements O. The stronger the connection between A and B, the stronger the deviation will be. In this way the information about connectivity between all neurons gets encoded in the fluctuations of measurement O in relation to GRASP expression patterns. The connection strength between any two neurons A and B may be calculated from such data specifically as

$$C_{AB} = \mathrm{E}[O \mid \mathrm{AB}] + \mathrm{E}[O \mid \bar{A}\bar{B}] - \mathrm{E}[O \mid A\bar{B}] - \mathrm{E}[O \mid \bar{A}B]. \quad (1)$$

Here $\mathrm{E}[O \mid \mathrm{AB}], \mathrm{E}[O \mid \bar{A}\bar{B}], \mathrm{E}[O \mid A\bar{B}], \mathrm{E}[O \mid \bar{A}B]$ are the triggered averages given A and B simultaneously express GRASP, neither A nor B express GRASP, etc. Although Eq.(1) is straightforward and computationally simple, it is not practical requiring a huge number of

trials before convergence (Figure 2A and C). Here we only discuss it to illustrate the principles behind encoding of information in Brainbow + GRASP.

In reality, circuit structure may be calculated from far smaller number of such observations. We first observe that each measurement O_i may be viewed as a linear constraint on the connectivity matrix C,

$$O_i \approx \sum_{A \in PRE(i)} \sum_{B \in POST(i)} C_{AB} \,, \tag{2}$$

where summation is over the sample $PRE(i)$ of neurons expressing pre-synaptic GRASP construct in animal i and the sample $POST(i)$ of neurons expressing post-synaptic construct (Figure 1B). Eq.(2) is a simple representation of the information gained by observing the set of identically labeled synapses between two populations of neurons, assuming no additional knowledge about synapses' positions or sizes. (Here we intentionally assume no such detailed prior, although where it is available one may write stronger constraints and further speedup the reconstruction.)

Second key observation is that the connectivity matrix is sparse – e.g. in *C. Elegans* only 2,000 neuronal pairs are connected out of total 80,000 possibilities[8], and in larger animals connectivity gets progressively sparser. Sparseness is a powerful additional constraint; in particular, it can be rigorously shown that sparse matrix may be reconstructed *exactly* from only $\sim N_p \log N$ of *nearly arbitrary* measurements (2)[30]. N_p here is the number of nonzero elements in the matrix and N is its size. Sought reconstruction is almost always the smallest l_1-norm solution of the full linear problem (2). l_1-norm in our case corresponds simply to the total strength of all synapses in the circuit S. Therefore, having acquired M observations $\{O_i\}$, the circuit's connectivity matrix may be found by solving for matrix C satisfying all constraints (2) with the smallest total synaptic strength, and such solution will converge to the true connectivity matrix with the exact answer attained at $M \sim N_p \log N$ observations. Minimizing l_1-norm under set of linear constraints (2) is a standard linear-programming problem; it is computationally tractable. In Methods we describe few algorithms for solving it efficiently.

In-silico reconstruction of complete wiring diagram in *C. Elegans* using Brainbow-guided GRASP probes

We performed a test of the above approach using in-silico model for Brainbow + GRASP experiment in *C. Elegans*, using actual complete wiring diagram known from electron microscopy[8]. In this model, we also considered noise factors that are likely to affect real-life experiment: biological variability (i.e. connectivity changes from one animal to another), observation noise (i.e. errors in assessing combined puncta fluorescence in each animal) and errors in identifying cells in GRASP expression pattern (see Methods). Assuming different noise levels, we generated collections of 500 to 12,000 in-silico Brainbow + GRASP measurements and attempted recovery of *C. Elegans* wiring diagram from these.

Obtained reconstructions were inspected visually as matrices (Supplementary Figure 1) and as scatter plots showing reconstructed vs. actual connection weights (Figures 2C-D and Supplementary Figure 2). Quantitatively, we characterized reconstructions quality with the

Nature Precedings : hdl:10101/npre.2009.2669.2 : Posted 2 Mar 2009

correlation coefficient r^2 between reconstructed and actual connection weights. We observed that in *C. Elegans* reasonable reconstructions of its wiring diagram were obtained already from 5,000-10,000 measurements (Figure 2B). Reconstructions were sensitive to noise in observations O_i and cell-id errors, thus, imposing stringent requirements on both (Figures 3B-C). Degradation from biological variability was less substantial (Figure 3A). Changing the fraction f of neurons expressing GRASP in each trial allowed to improve reconstruction robustness to observation noise (Figure 3D), but not cell-id errors or biological variability (data not shown).

DISCUSSION

We argue that reconstruction of complete wiring diagram in an animal as large as *C. Elegans* may be attainable using the above approach and existing optical and genetic tools in a short amount of time. Given small size (100 μm diameter and 1000 μm long) and fast development (2 -3 days), 10,000 *C. Elegans* may be incubated rapidly on a single Petri-dish. In fact, modern *C. Elegans* phenotype screens already routinely pay attention to populations that big. The measurement of the combined fluorescence from GRASP may be quickly obtained with a low resolution snapshot of each animal. Identities of the neurons expressing GRASP may be determined from a high resolution 3D scan of nuclei-bound fluorescence with a computer algorithm such as in Long et al.[32] Such scans may be performed in about 1 minute or less with an off-the-shelf light microscopes, assuming 10MHz acquisition rate and voxel size of $(0.5\mu m)^3$. Given all of the above, such reconstructions of complete circuit in *C. Elegans* appear plausible to complete in about one week. The only essential prerequisite for such reconstructions is the creation of Brainbow + GRASP transgenic animal.

Described general framework may be applied to neural circuits in larger organisms such as *Drosophila*, zebrafish, leech, mouse, etc., with different connectivity probes such as calcium indicators, neuronal tracers, ChR2 or transsynaptic viruses, and different targeting techniques such as MARCM, MADM or UAS/Gal4. Importantly, stochasticity of the expression patterns is not a necessary feature of our approach. On the contrary, connectivity may be equally well recovered from nearly arbitrary sets of expression patterns, e.g. such as produced with libraries of Gal4-lines. Furthermore, ability to target individual or small groups of neurons, e.g. with UAS/Gal4, is also not necessary here. With broad probes we are able to constrain the circuit effectively in just $\sim N_p \log N$ time, comparable to $\sim N_p$ time that would be necessary to determine connectivity with precise probes, such as whole-cell patches, by directly probing connections between individual neurons when it is *already known* which pairs should be tested. Finally, although the connectivity matrix here was by construction nonnegative, one can still map inhibitory or other types of connections with this approach by designing measurements targeting exclusively inhibitory synapses or employing arrays of distinct markers.

Although we show that in *C. Elegans* described approach should be successful with existing tools, applications in larger circuits will require meeting a number of challenges. In larger circuits unfavorable size scaling should be overcome: circuits get physically bigger, and so imaging gets harder, while the complexity gets higher and so more measurements

Nature Precedings : hdl:10101/npre.2009.2669.2 : Posted 2 Mar 2009

are needed. Transgenic models allowing for multiple connectivity probes, such as multi-color GRASP, is the likely answer to this problem. Improvement of the reconstruction algorithm aimed at reduction of the number of measurements may also be possible. While Candes et al. discuss *exact* reconstruction of sparse matrices with theoretical performance bound[30], faster algorithms may be possible if certain error is tolerated or if detailed priors are available. Even greater challenge is determination of the probe expression patterns in different animals. In larger animals it is not known how to perform such determination or even what stereotypical units of connectivity might be. The progress in designing transgenic systems allowing for cell-type identification, or systems where probes may be driven in specific classes of cells, e.g. using UAS/Gal4, may be advantageous. Finally, it may be possible for the reconstruction algorithm to rely to much lesser degree on the prior knowledge of expression patterns than what is described here.

Nature Precedings : hdl:10101/npre.2009.2669.2 : Posted 2 Mar 2009

METHODS

Computational reconstruction problem for Brainbow + GRASP. Specific computational reconstruction problem for a collection of M Brainbow + GRASP measurements, taking into account sparseness of the objective matrix, is formulated as following linearly constrained l_1-optimization:

$$\min\left\|C\right\|_{l_1} = \min \sum_{A=1}^{N}\sum_{B=1}^{N}\left|C_{AB}\right|, \text{ subject to} \tag{3a}$$

$$O_i = \sum_{A=1}^{N}\sum_{B=1}^{N}\alpha_A(i)\beta_B(i)C_{AB},\ i=1...M, \tag{3b}$$

$$\alpha_A(i) = \begin{cases} 1, A \in PRE(i) \\ 0, A \notin PRE(i) \end{cases}, \tag{3c}$$

$$\beta_B(i) = \begin{cases} 1, B \in POST(i) \\ 0, B \notin POST(i) \end{cases}. \tag{3d}$$

$\alpha_A(i)$ and $\beta_B(i)$ are the indicator functions for the sets $PRE(i)$ and $POST(i)$, and describe GRASP expression patterns in different animals. By construction, these are determined independently in each animal, e.g., from nuclei-targeted fluorescence. Problem (3) corresponds to maximizing $P(C\,|\,\{O_i\}) \sim \prod P(O_i\,|\,C)P(C)$ assuming normal distribution of noise in observations and its constant variance and exponential sparseness prior on the connectivity $P(C) \sim \exp\left(-\lambda\sum\left|C_{AB}\right|\right)$, $\lambda \to 0$. Optimization problem (3) is a standard linear-program and vast literature exists about solving it efficiently[33]. For larger N the following approximate but less demanding algorithm is advantageous. If total synaptic strength S is known in advance (in our case $S \approx \mathrm{E}[O]/f^2$), the solution to (3) may be found as the intersection of two convex sets – $\left\|C\right\|_{l_1} = S$ and the hyperplane of linear constraints (3b)[29]. By Candes et al.[30], for sufficiently large M such intersection contains single point. Such intersection may be found even for problems of very large size by alternately projecting on these two sets[34].

In-silico model for circuit reconstruction using Brainbow + GRASP in *C. Elegans*. In-silico model for *C. Elegans* circuit reconstruction using Brainbow + GRASP is set as follows. Complete wiring diagram for *C. Elegans* is available from electron microscopy in the form of integer counts of distinct synapses between different pairs of neurons C_{AB}[8], as first documented by White et al.[15] Thus, instead of measuring combined size of synaptic puncta, we here adopt the convention that O_i represents the count of synapses labeled in each Brainbow + GRASP animal. To account for possible biological variability, i.e. variability in the connection weights from animal to animal, we replace such connection weights in each individual animal as in $C_{AB} \to C_{AB}(1-a_b)+P[a_b C_{AB}]$, where $P[x]$ is Poisson-distributed random number with mean x. Parameter a_b controls the degree of expected variability: a_b=0 corresponds to no variability at all, and a_b=1 corresponds to the case where synapses are formed completely at random with C_{AB} synapses on average. Next, we assume that samples of neurons *PRE* and *POST* are formed with Brainbow, i.e. each neuron is led to express pre- and post-synaptic constructs of GRASP with constant probability f=0.5. Corresponding to the above C_{AB}, and chosen *PRE* and *POST* expression patterns, for each animal we record the number of selected synapses O_i. We add Gaussian noise to this count to model observation errors as follows: $O_i \to O_i(1+N[0,a_o])$. $N[0,a_o]$ is Normally-distributed random number with zero mean and variance a_o. We assume that identities of cells in GRASP expression patterns are recovered with computer algorithm from associated nuclei-bound fluorescence. Possible error in such identification is modeled by all-to-all shuffling of cell-identities in a small fraction a_i of randomly chosen neurons. This is a rather pessimistic model: in reality some cells will be better identified than others and the confusions will not be all-to-all but only with the nearby cells. Thus, Figure 3C should be viewed as an upper bound on these error-rates and actual experiments may tolerate this kind of errors better. Finally, connectivity matrix is calculated from thus produced collection of noisy measurements $\{O_i, PRE(i), POST(i)\}$ by solving Eqs.(3) with the method of alternate projections above.

Nature Precedings : hdl:10101/npre.2009.2669.2 : Posted 2 Mar 2009

ACKNOWLEDGEMENTS

The author would like to acknowledge essential discussions with Shiv V. Vitaladevuni, Veit Elser, Max Nikitchenko and Liam Paninski. This work was partially carried out at Janelia Farm Research Campus, HHMI and Aspen Center for Physics.

FIGURE LEGENDS

Figure 1: Reconstructing neural circuits with Brainbow + GRASP. A) Pre- and post-synaptic split-GFP (sGFP) are bundled with sequences for two nuclei-targeted FP (nFP), and separately flanked with inverted loxP. The construct allows producing variety of GRASP expression patterns within single genetic line, while nFP allows identifying populations of cells thus probed in different animals. B) In each Brainbow + GRASP trial two random sets of neurons express pre-synaptic (red) and post-synaptic (green) sGFP. Whenever a synapse is present between any two from pre/post-synaptic neurons, sGFP reassemble and create fluorescent puncta (yellow). Combined fluorescence strength of all puncta is measured, corresponding mathematically to a constraint on the connection

weights at the intersection of selected pre-synaptic rows and post-synaptic columns. When large number of such measurements is obtained, connectivity may be recovered from the fluctuations in O in relation to GRASP expression patterns. Three trials are shown for illustration.

Figure 2: In-silico reconstruction of complete connectivity in *C. Elegans* using Brainbow + GRASP, based on the actual wiring diagram known from electron microscopy[8]. (A) Quality of the reconstructed connectivity matrix with triggered-average method, as measured by the correlation coefficient r^2 between reconstructed and actual connection weights. (B) Quality of the reconstruction using the method of alternate projections. (C) Example of best reconstruction using triggered-averaging with M=300K measurements, r^2=0.4. (D) Example of reconstruction using the method of alternate projections for M=8K measurements, r^2=0.9.

Figure 3: Impact of different noise factors on the reconstruction using Brainbow + GRASP in *C. Elegans*, as measured by the correlation coefficient r^2 between reconstructed and actual connection weights. (A) Biological variability is of lesser concern with good reconstructions obtained even when synaptic weights in different animal are Poisson random with given mean. (B) Cell-identification errors have significant impact on reconstructions implying that the fraction of misidentified neurons should be kept below 3-5%. (C) Observation noise has significant impact on reconstructions implying that the measurements should be obtained with relative error better than 3%. (D) Impact of observation noise may be effectively controlled by reducing expression frequency f: larger amount of observation noise may be tolerated for f=0.1-0.2. Changing f has no impact in case of biological variability or cell-id errors (data not shown).

Nature Precedings : hdl:10101/npre.2009.2669.2 : Posted 2 Mar 2009

References

1. Broome, B. M., Jayaraman, V. & Laurent, G. Encoding and decoding of overlapping odor sequences. *Neuron* **51**, 467-482 (2006).
2. Jones, L. M., Fontanini, A., Sadacca, B. F., Miller, P. & Katz, D. B. Natural stimuli evoke dynamic sequences of states in cortical ensembles. *PNAS* **104**, 18772-7 (2007).
3. Rabinovich, M. I., Huerta, R., Varona, P. & Afraimovich, V. S. Transient cognitive dynamics, metastability, and decision making. *PLoS Comput. Biol.* **4**, e1000072 (2008).
4. Chalasani, S. H. et al. Dissecting a circuit for olfactory behavior in Caenorhabditis elegans. *Nature* **450**, 63-70 (2007).
5. Dunn, N. A., Lockery, S. R., Pierce-Shimomura, J. T. & Conery, J. S. A neural network model of chemotaxis predicts functions of synaptic connections in the nematode Caenorhabditis elegans. *J. Comp. Neurosci.* **17**, 137-47 (2004).
6. Gray, J. M., Hill, J. J. & Bargmann, C. I. A circuit for navigation in Caenorhabditis elegans. *PNAS* **102**, 3184-91 (2005).
7. Song, S., Sjostrom, P. J., Reiql, M., Nelson, S. & Chklovskii, D. B. Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biol.* **3**, e68 (2005).

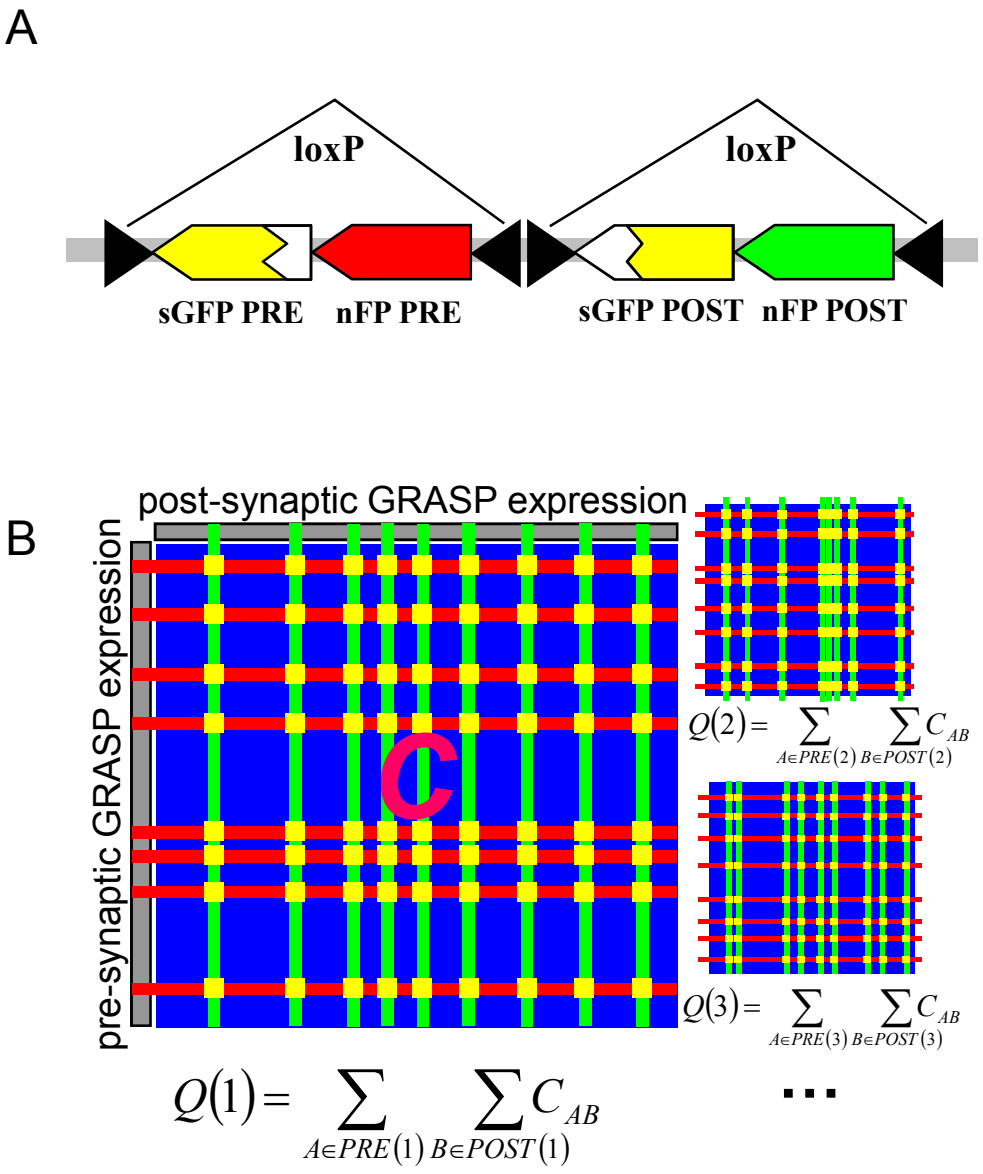
8. Wormatlas: a database of behavioral and structural anatomy of Caenorhabditis elegans at http://www.wormatlas.org.
9. Chen, B. L., Hall, D. H. & Chklovskii, D. B. Wiring optimization can relate neuronal structure and function. *PNAS* **103**, 4723-28 (2006).
10. Hall, D. H. & Russell, R. L. The posterior nervous system of the nematode Caenorhabditis elegans: Serial Reconstruction of identified neurons and complete pattern of synaptic interactions. *J. Neurosci.* **11**, 1-22 (1991).
11. Jain, V. et al. Supervised learning of image restoration with convolutional networks. International Conference on Computer Vision, Rio de Janeiro, Brazil, 2007.
12. Jurrus, E. et al. Axon tracking in serial block-free scanning electron microscopy. MICCAI 2006 workshop., Copenhagen, 2006.
13. Macke, J. et al. Contour-propagation algorithms for semi-automated reconstruction of neural processes. *J. Neurosci. Methods* **167**, 349-57 (2008).
14. Mishchenko, Y. Automation of 3D reconstruction of neural tissue from large volume of conventional Serial Section Transmission Electron Micrographs. *J. Neurosci. Methods* In Press; doi:10.1016/j.jneumeth.2008.09.006.
15. White, J., Southgate, E., Thomson, J. N. & Brenner, S. The structure of the nervous system of the nematode caenorhabditis elegans. *Phil. Trans. Royal Soc. London. Series B, Biological Sciences* **314**, 1-340 (1986).
16. Baker, B. J. et al. Imaging brain activity with voltage- and calcium-sensitive dyes. *Cell. Mol. Neurobiol.* **25**, 245-82 (2005).
17. Ikegaya, Y., Le Bon-Jego, M. & Yuste, R. Large-scale imaging of cortical network activity with calcium indicators. *Neurosci. Research* **52**, 132-8 (2005).
18. Bureau, I., Shepherd, G. M. & Svoboda, K. Precise development of functional and anatomical columns in the neocortex. *Neuron* **42**, 789-801. (2004).
19. Feinberg, E. H. et al. GFP Reconstitution Across Synaptic Partners (GRASP) defines cell contacts and synapses in living nervous systems. *Neuron* **57**, 353-63 (2008).
20. Micheva, K. D. & Smith, S. J. Array tomography: a new tool for imaging the molecular architecture and ultrastructure of neural circuits. *Neuron* **55**, 25-36 (2007).
21. Petreanu, L., Huber, D., Sobczyk, A. & Svoboda, K. Channelrhodopsin-2-assisted circuit mapping of long-range callosal projections. *Nat. Neurosci.* **10**, 663-8 (2007).
22. Wickersham, I. R., Finke, S., Conzelmann, K. K. & Callaway, E. M. Retrograde neuronal tracing with a deletion-mutant rabies virus. *Nat. Methods* **4**, 47-9 (2007).
23. Lee, T. & Luo, L. Mosaic analysis with a repressible cell marker (MARCM) for Drosophila neural development. *Trends Neurosci.* **24**, 251-4 (2001).
24. Livet, J. et al. Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* **450**, 56-62 (2007).
25. Luo, L. Fly MARCM and mouse MADM: genetic methods of labeling and manipulating single neurons. *Brain Res. Rev.* **55**, 220-7 (2007).
26. Luo, L., Callaway, E. M. & Svoboda, K. Genetic dissection of neural circuits. *Neuron* **57**, 634-60 (2008).
27. Phelps, C. B. & Brand, A. H. Ectopic gene expression in Drosophila using GAL4 system. *Methods* **4**, 367-79 (1998).

Nature Precedings : hdl:10101/npre.2009.2669.2 : Posted 2 Mar 2009

28. Wu, J. S. & Luo, L. A protocol for mosaic analysis with a repressible cell marker (MARCM) in Drosophila. *Nat. Protocols* **1**, 2583-9 (2006).
29. Candes, E. J. & Romberg, J. Practical signal recovery from random projections. SPIE Conference on Wavelet Applications in Signal and Image Processing XI, 2005.
30. Candes, E. J., Romberg, J. & Tao, T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory* **52**, 489-509 (2006).
31. Herman, G. T. & Kuba, A. *Discrete Tomography: Foundations, Algorithms, and Applications.* (eds. Herman, G. T. & Kuba, A.) (Birkhauser, Boston, 1999).
32. Long, F., Peng, H., Liu, X., Kim, S. & Myers, E. Automatic recognition of cells (ARC) for 3D images of C. Elegans. in Lecture Notes in Computer Science: Research in Computational Molecular Biology pp. 128-139 (Springer Berlin, 2008).
33. Vanderbei, R. J. *Linear Programming: Foundations and Extensions*. (Springer-Verlag, New York, 2001).
34. Bregman, L. M. The method of successive projection for finding a common point of convex sets. *Soviet Math. Dokl.* **6**, 688-92 (1965).

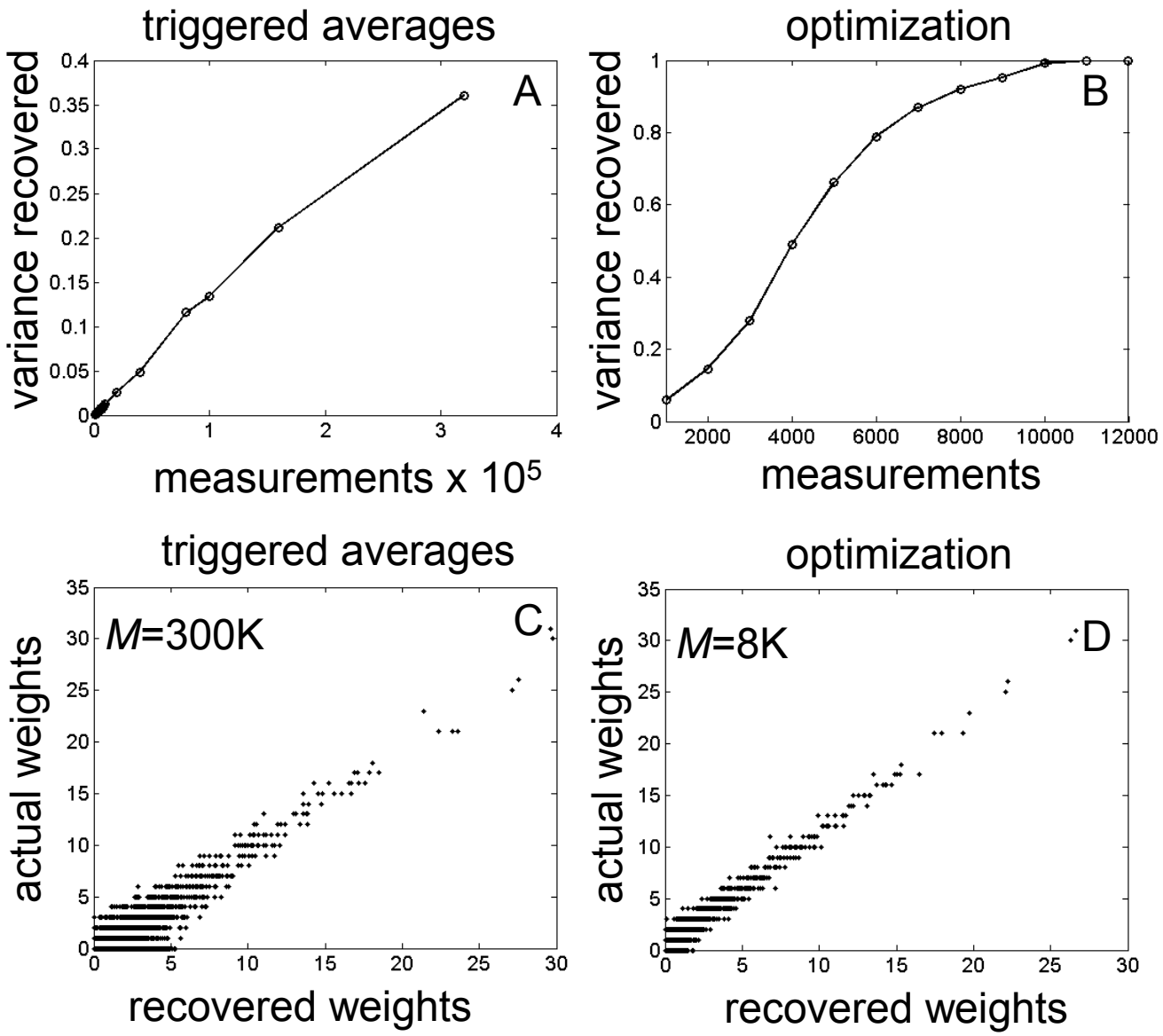
Nature Precedings : hdl:10101/npre.2009.2669.2 : Posted 2 Mar 2009

Figure 1



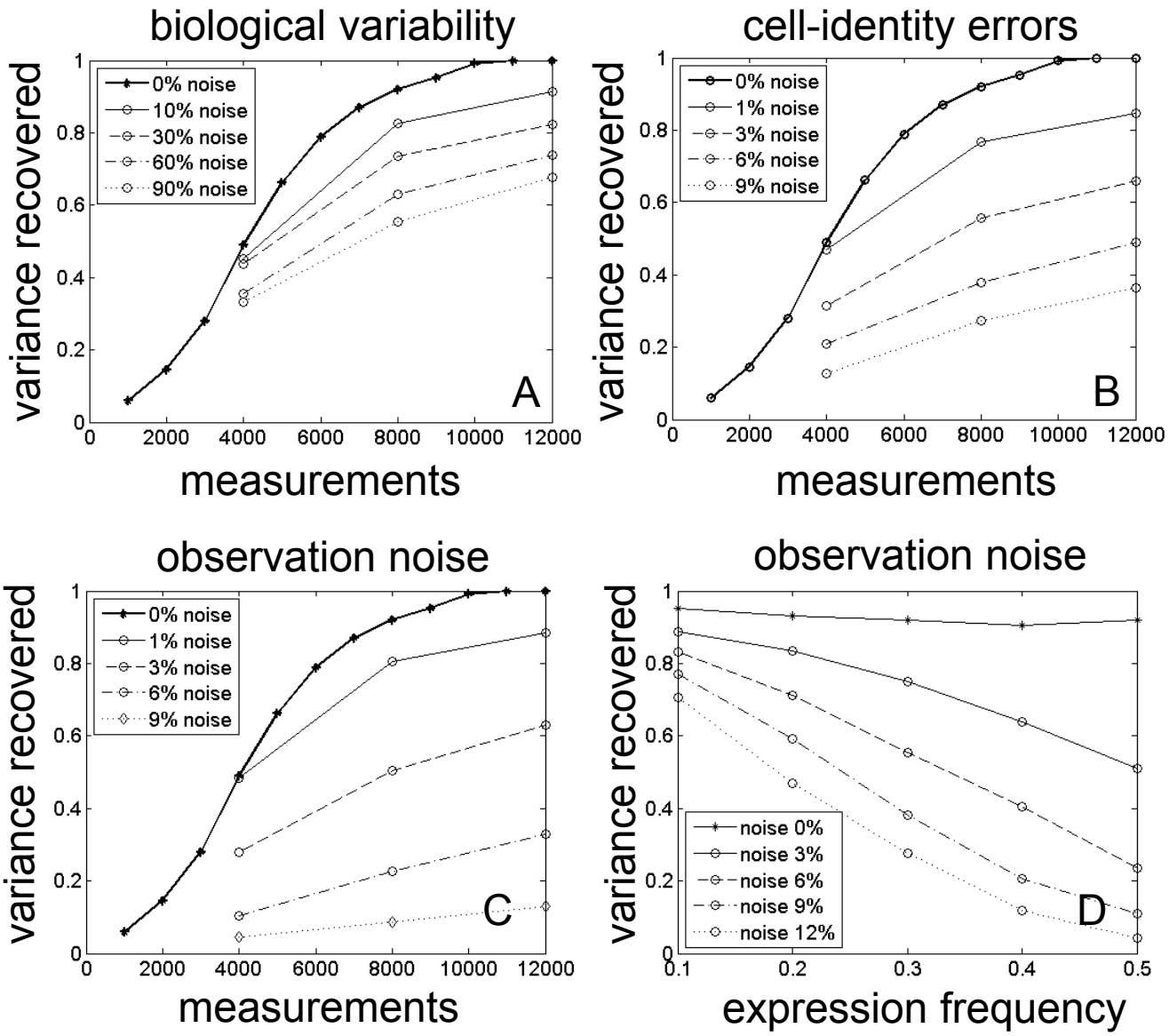
Nature Precedings : hdl:10101/npre.2009.2669.2 : Posted 2 Mar 2009

Figure 2



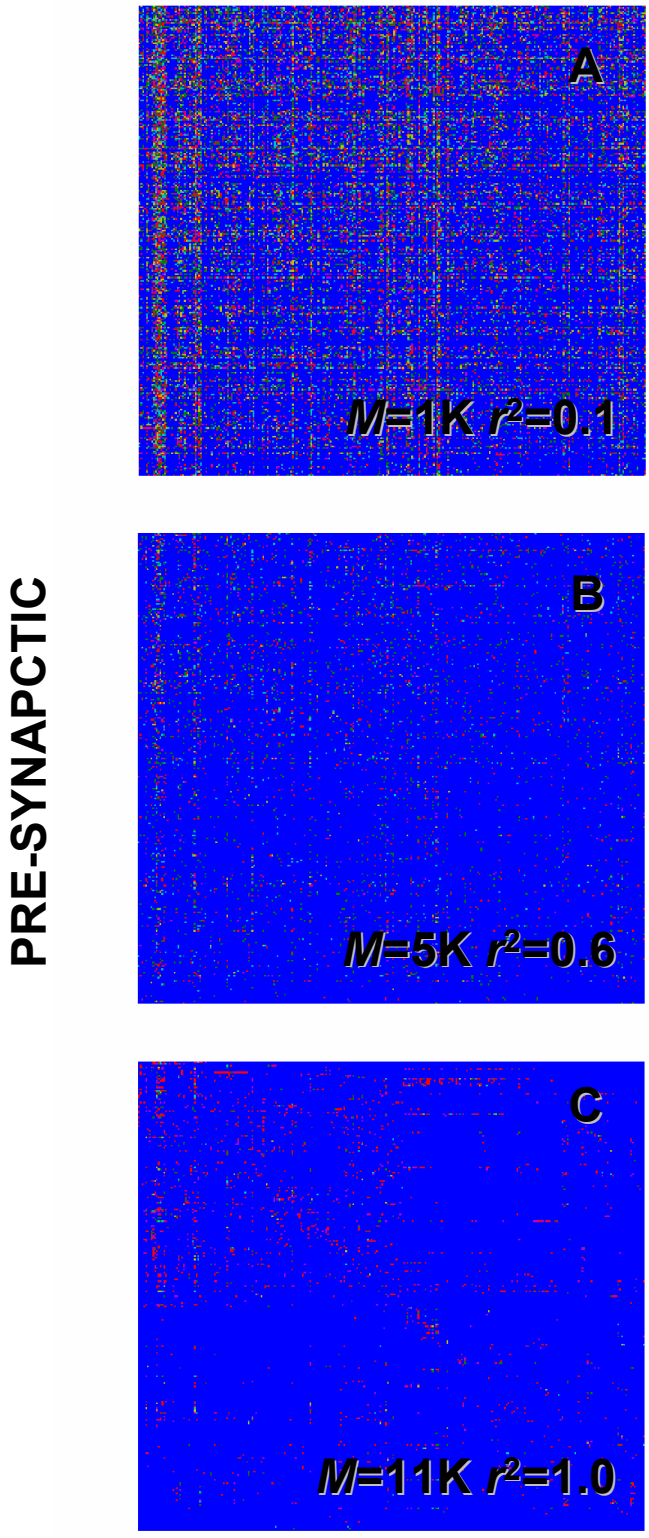
Nature Precedings : hdl:10101/npre.2009.2669.2 : Posted 2 Mar 2009

Figure 3



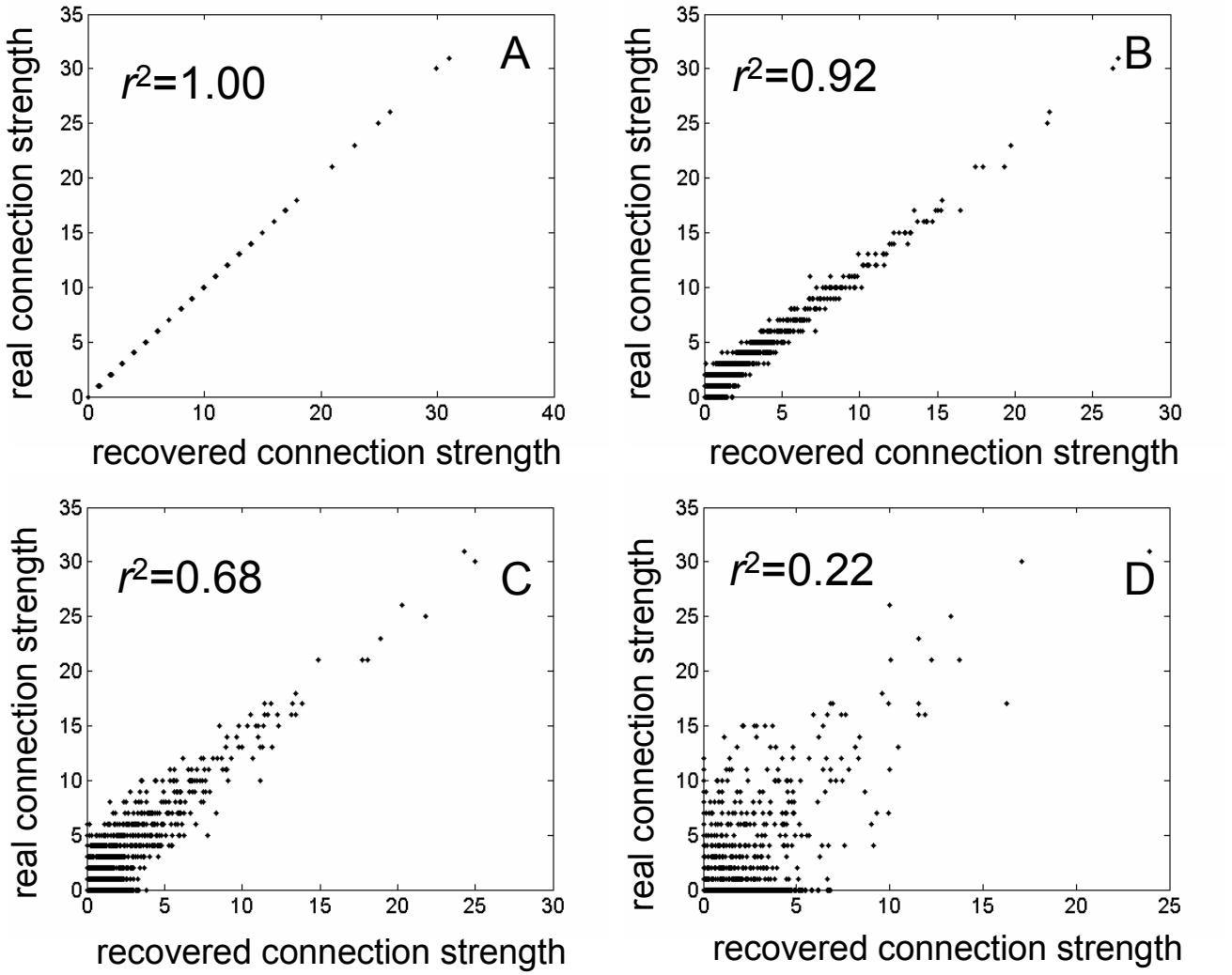
Nature Precedings : hdl:10101/npre.2009.2669.2 : Posted 2 Mar 2009

Nature Precedings : hdl:10101/npre.2009.2669.2 : Posted 2 Mar 2009



Supplementary Figure 1: (A-C) Reconstructed connectivity matrix for M=1,000, M=6,000 and M=11,000 trials, and corresponding r^2, under noiseless conditions.

Nature Precedings : hdl:10101/npre.2009.2669.2 : Posted 2 Mar 2009



Supplementary Figure 2: Scatter plots of reconstructed vs. actual connection strengths. (A) Noiseless reconstruction with M=11,000 trials gives the exact answer. (B) Noiseless reconstruction with M=8,000 is very close to the exact answer. (C) Noisy reconstruction with connection strengths from trial-to-trial given by fully Poisson random and M=8,000 is very close to the exact answer. (D) Noisy reconstruction with 6% of injected observation noise and M=8,000 is not very useful.