

# Reconstruction of Biological Networks through Gene Silencing and Overexpression

Russell Howes<sup>1</sup>, Neil Dalchau<sup>2</sup>, Jorge Gonçalves<sup>2</sup>, Sean Warnick<sup>1,\*</sup>

1. Information and Decision Algorithm Laboratories, Computer Science Department  
Brigham Young University, Provo, Utah

2. Department of Engineering, University of Cambridge, UK

\*E-mail: sean@cs.byu.edu

## 1 Network Reconstruction

Many methods that exist for reconstructing biological networks take into account assumptions of sparsity or already known relationships among states of the network. Here we wish to completely characterize the relationships among observed states and inputs in a network without making excessive assumptions about structure. In this work we demonstrate that a series of inducible overexpression or silencing experiments equal to the number of observed states will completely characterize system structure.

Dynamical structure functions, as explained in [1], describe the causal relationships among observed states and inputs in a network. The dynamical structure  $(Q, P)$  of a linear system is related to that system's transfer function by

$$(I - Q)G = P \tag{1}$$

where  $Q$  represents relationships among the observed states (basically, the information we look for, i.e., the dynamical network) and  $P$  represents effects of the inputs on the observed states.

In [1] we showed that input-output data alone is insufficient for the task of network reconstruction even if a system is linear and time-invariant. There we characterized precisely the additional information needed to obtain network structure without making further assumptions. Here we show how this information can be obtained through a series of typical biological experiments. Assuming nothing is known about the internal structure of the system (i.e.  $Q$ ), such experiments are designed in a way that the structure of  $P$  is diagonal. As explained in [1], with enough experiments  $Q$  can be solved uniquely from (1).

## 2 Overexpression and Silencing

Gene overexpression and silencing are two methods of biological system modification. Overexpression may be constitutive or inducible, through introduction of a transgene into the host which is specifically designed to increase the abundance of the desired transcript. Expression may also be down-regulated, via mutation or inhibition, such as RNA silencing. The target specificity of these methods allows the control of gene expression, without directly affecting other genes in the network. Such methods can be modeled by introducing extra states to the linear(ized) system, which can be controlled externally.

Silencing of a gene  $i$  will drive the corresponding RNA  $x_i$  near to zero, independently of its upward regulators, by some interaction with a typically hidden variable that is externally

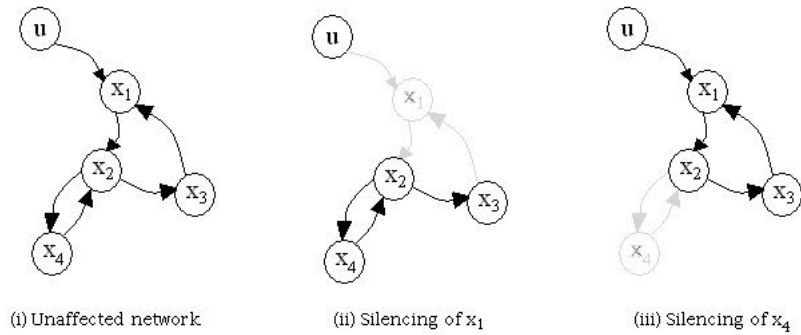


Figure 1: A sample network (i) and the modified networks that result from silencing of certain genes in the network (ii,iii) Each silencing (or overexpression) experiment reveals information about system structure that cannot be obtained from input-output data alone—for example, we note that (ii) will result in a zero transfer function. This implies the input  $u$  only works directly on  $x_1$ .

controlled. The result is that the vector of  $P$  corresponding to this experiment will have one and only one nonzero entry (the  $i$ th entry). As for  $Q$ , row  $i$  will all be zero since no measured state can affect  $x_i$ . Such experiments provide new information about the elements of  $Q$  (see Figure 1.)

A gene  $j$  can be inducibly overexpressed with a similar modification to the original system. Basically, an externally controlled typically hidden variable turns the gene on, independently of its upward regulators. The implications for recovering  $Q$  and  $P$  are similar to those for silencing of a gene.

### 3 Reconstructing Networks from Overexpression or Silencing Experiments

We discuss why silencing and overexpression experiments are preferable to mutations and other experiments, explain how this work relates to other results ([2], [3], [4], [5], [6], [7], [8]) and outline future applications and modifications of this theory for nonlinear systems.

## 4 References

### References

- [1] J. Gonçalves, S. Warnick, R. Howes, “*Necessary and Sufficient Conditions for Dynamical Structure Reconstruction of LTI Networks*,” IEEE Conference on Decision and Control (December 2007).
- [2] E.D. Sontag, A. Kiyatkin, and B.N. Kholodenko, “*Inferring dynamic architecture of cellular networks using time series of gene expression, protein and metabolite data*,” IEEE Bioinformatics, 20(12):1877-1886, 2004.
- [3] D. Di Bernardo, T. Gardner, J. Collins, “*Robust Identification of large genetic networks*,” Pac. Symp. Biocomput. 2004, 9, 486-497.

- [4] N. Friedman, “*Inferring Cellular Networks Using Probabilistic Graphical Models,*” *Science*, 6 Feb 2004, pp. 799-805.
- [5] K.-H. Cho, S.-M. Choo, S.H. Jung, J.-R. Kim, H.-S. Choi and J. Kim, “*Reverse engineering of gene regulatory networks,*” *Systems Biology, IET*, May 2007, 149-163.
- [6] Tegner, et al. Reverse engineering gene networks: Integrating genetic perturbations with dynamical modeling. *Proc. Natl. Acad. Sci. USA*, 2003, 100, p.5944-5949.
- [7] K. Gadkar, R. Gunawan, F. Doyle, “*Iterative approach to model identification of biological networks,*” *BMC Bioinformatics* 2005, 6:155.
- [8] B. Kholodenko, E. Sontag, et al. “*Untangling the wires: A strategy to trace functional interactions in signaling and gene networks,*” *PNAS* vol 99 no. 20, pp. 12841-12846.