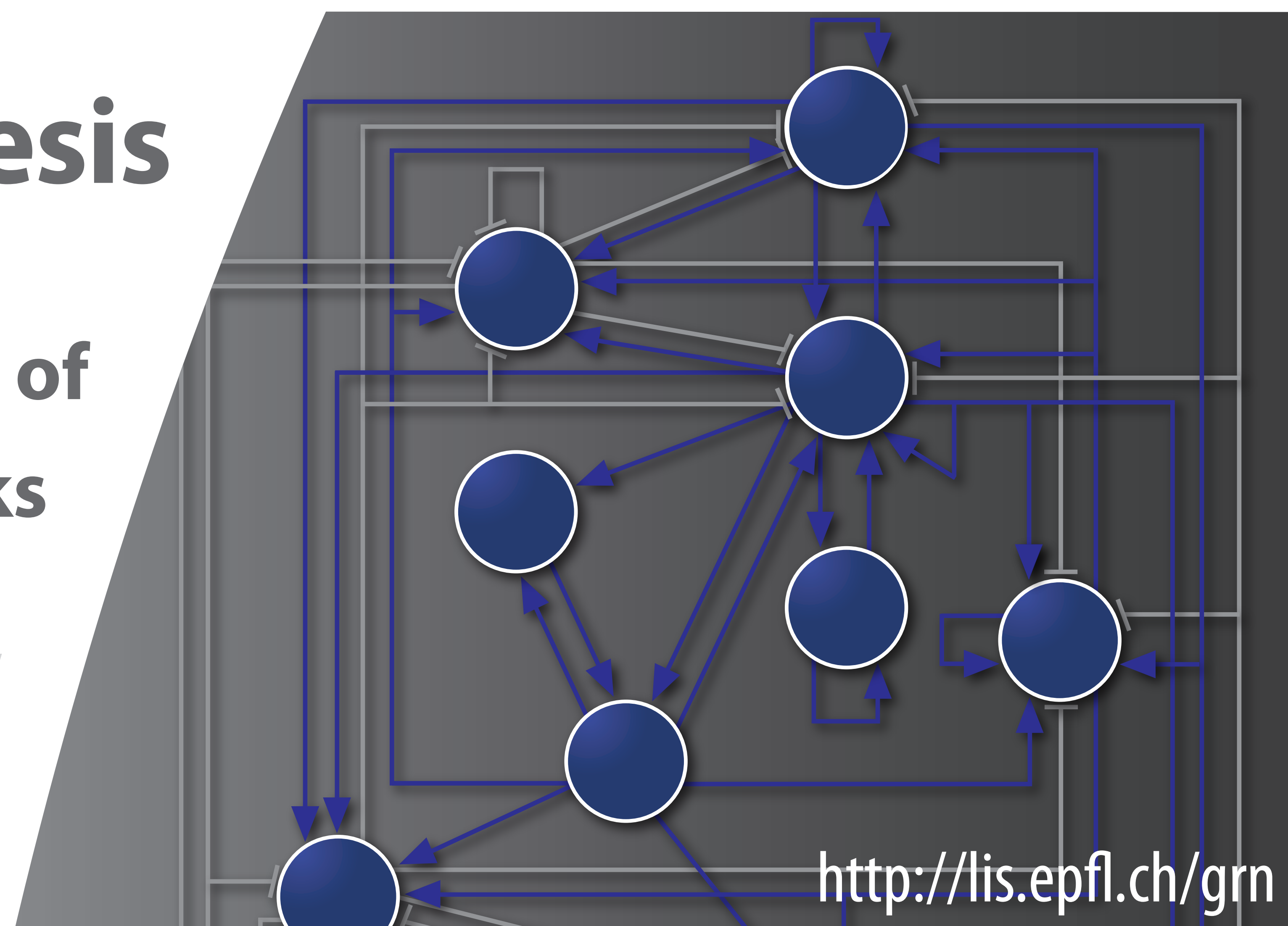


Scalable Reverse Engineering of Nonlinear Gene Networks

Thomas Schaffter, Prof. Dario Floreano
Microengineering
Laboratory of Intelligent Systems (LIS)
Ecole Polytechnique Fédérale de Lausanne (EPFL)
February 2008



<http://lis.epfl.ch/grn>

Sigmoid gene model

- » The present Reverse Engineering Algorithm allows the use of nonlinear gene models when a great number of existing algorithms use linear ones.
- » Genes in biological networks have nonlinear behaviours. Here we use a sigmoid model to approximate the variations of mRNA levels for each gene.

$$\frac{dx_i}{dt} = m_i \cdot \sigma \left(\sum_j w_{ij} x_j + b_i \right) - \lambda_i x_i$$

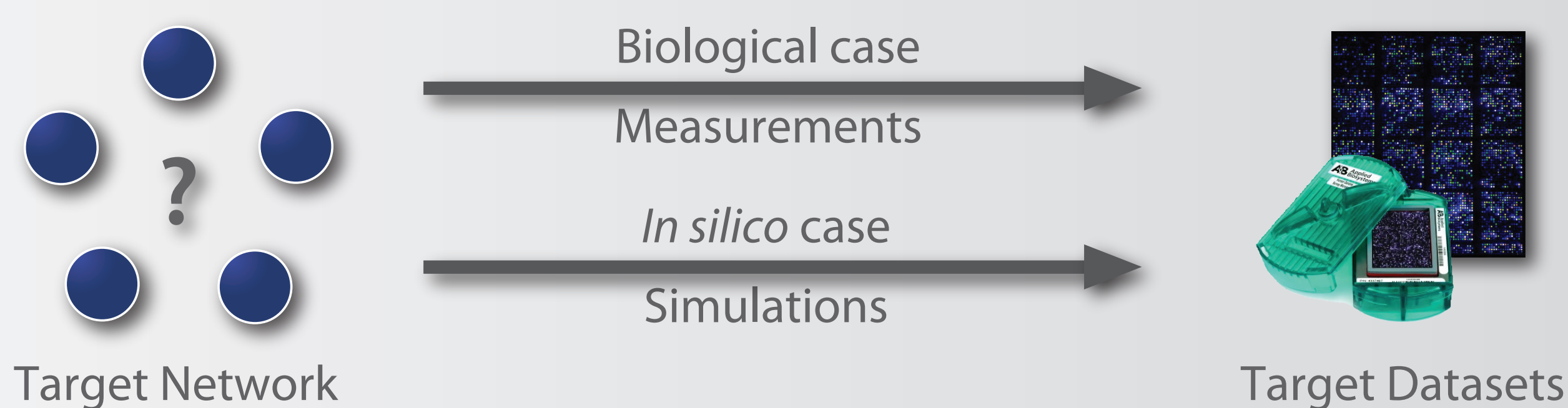
σ : sigmoidal activation function $\sigma(y) = 1 / (1 + e^{-y})$
 x_i : expression level of gene i
 w_{ij} : regulatory influence of gene j on gene i
 m_i, b_i, λ_i : max. transcription rate, bias, and decay rate of gene i

Abstract

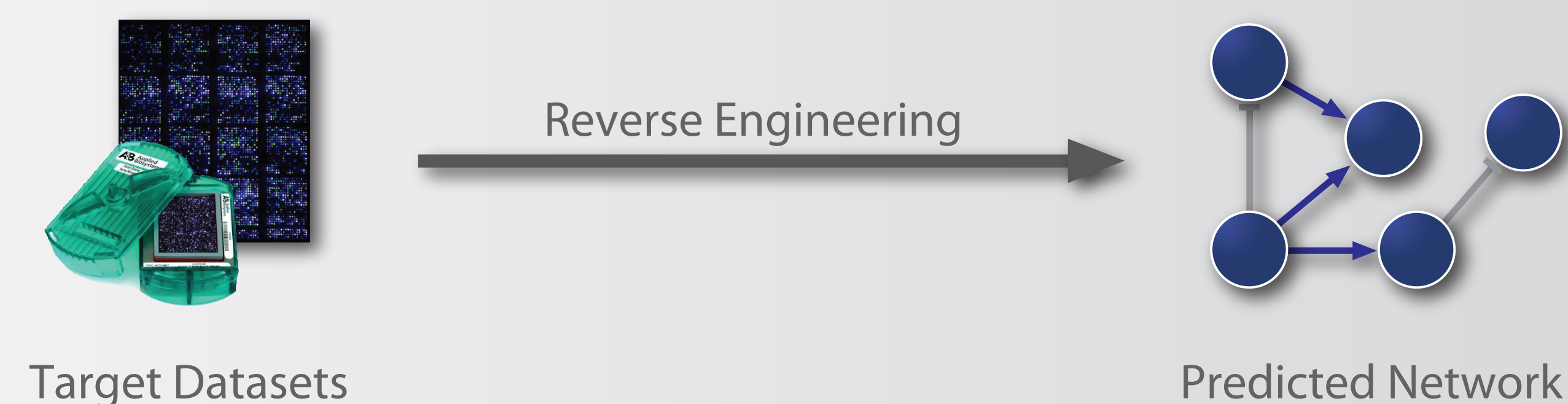
Current advanced molecular biology techniques provide gene expression (mRNA) levels of selected genes of an organism. In such aggregate, the mRNA level of one gene is mediated by the presence of specific proteins and metabolites produced by other genes. Gene Regulatory Networks (GRNs) include such structures, where a node (gene) is linked to others through gene-gene interactions. The developed algorithm is a new state-of-the-art process to infer GRNs, i.e. to find the gene-gene relationships from experimental mRNA levels. Knowledge of such interactions and so having possibility to act and control genes is important in biotech and pharmaceutical industries.

Process

A. Collect experimental steady-states and/or time series data from the target network to infer (gene expression levels).



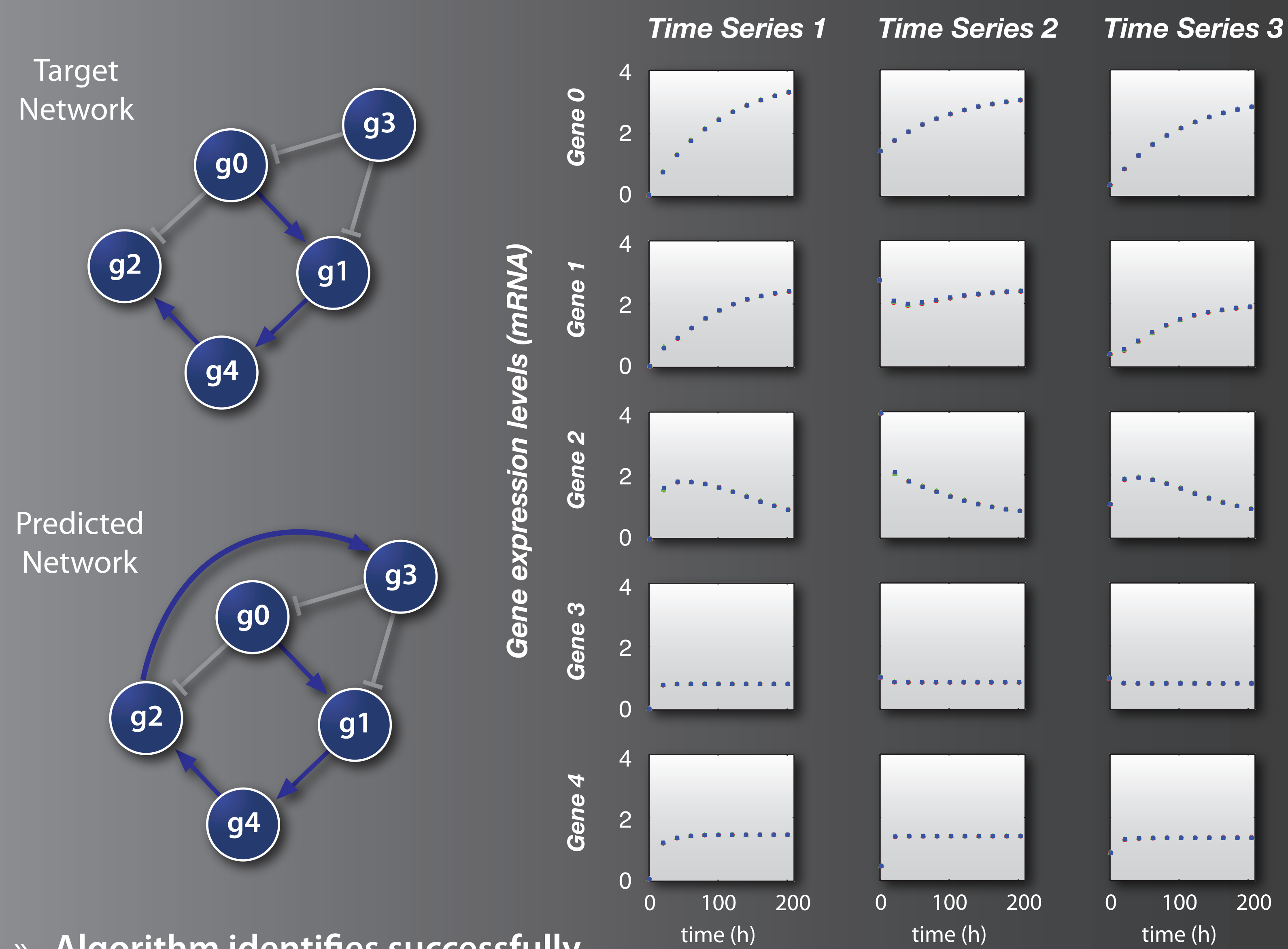
B. Infer the topology (gene-gene interactions) of the target regulatory network from experimental datasets.



In the biological case, target networks are unknown. Therefore *in silico* benchmarks are used to assess efficiency of the algorithm.

Results

» Reverse engineering of an *in silico* 5-genes regulatory network



» Algorithm identifies successfully

presences and inhibitory and excitatory types

of the 6 gene-gene interactions of the target regulatory network.

» 1 false positive interaction is found.

Indeed, if the problem is underdetermined (not enough input data), there exist several different networks that produce identical, output data.

» Present algorithm needs only about one minute to infer 5-genes networks when others take 6-8 hours.

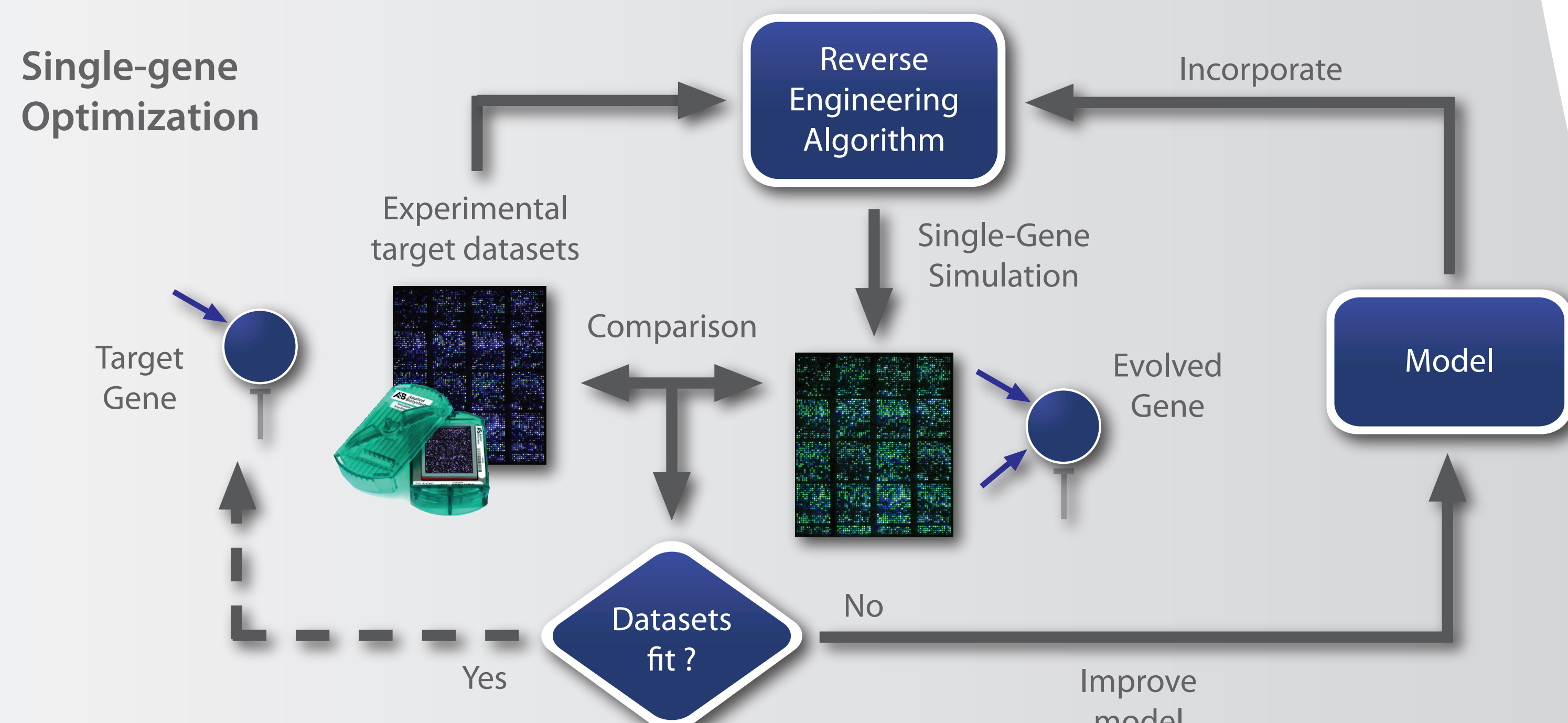
» Current application of this new algorithm to a 50-genes network given as challenge during the second DREAM conference (New York, 2007).

Selected publications

- » Marbach, D., Mattiussi, C. and Floreano, D. (2008) Replaying the Evolu-Tape: Biomimetic Reverse Engineering of Gene Networks. *Ann NY Acad. To appear*
- » Mendes, P. (2005) Developing strategies for systems biology. *In Virginia Bioinformatics Institute 2005 Scientific Report.*

Reverse Engineering Algorithm

- » mRNA levels are described by a system of N coupled differential equations (N genes). System hard to solve if N large! (dimension $\sim N^2$)
- » Idea: Divide the system into N independent equations of dim $\sim N$. System solved "gene-by-gene" (N incoming interactions to find per gene)
- » Multiple optimizations of each gene and selection of the best one(s).



» Goal: Find genes that generate the same data as the target ones. Use Covariance Matrix Adaptation - Evolution Strategy (CMA-ES) to generate the gene's parameters.