From Reaction Networks to Information Flow

Using Modular Response Analysis to Track Information in Signalling Networks

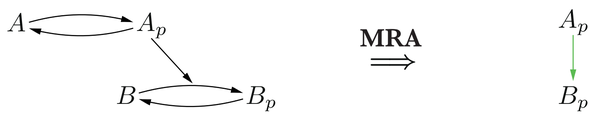
Pascal Schulthess*, Nils Blüthgen

Insitute for Pathology, Charité - Universitätsmedizin Berlin
Institute for Theoretical Biology, Humboldt University of Berlin

* Contact: pascal.schulthess@charite.de, Web: http://sys-bio.net

Motivation

Modular response analysis (MRA) aids in analysing the quantitative information transfer in signal transduction networks. The sensitivity of a target (e.g. transcription factor, protein) to an upstream stimulus (e.g. growth factor) can be determined by so-called global response coefficients.
In particular, modular response analysis allows to it generate networks of information flow from reaction networks.



Our aim was to develop an algorithm based on sparse matrix operations to perform MRA and conservation analysis that can handle large signalling networks. Using a monte-carlo method, we further apply this approach to determine whether a given species has positive, negative or no influence on any other species of the network.

Conservation Analysis

The dynamic behaviour of a biochemical reaction system is determined by

$$\frac{d}{dt}\boldsymbol{c}(t) = \boldsymbol{N}\boldsymbol{v}(\boldsymbol{c}(t))$$

with

$\boldsymbol{c} \in \mathbb{R}^m$	concentration vector
$\boldsymbol{N} \in \mathbb{R}^{m\times n}$	stoichiometric matrix
$\boldsymbol{v} \in \mathbb{R}^n$	reaction rate vector.

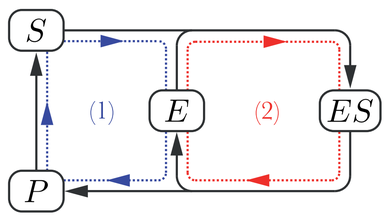


Figure 1 Example network with two conservation cycles

If conservation relations exist, one can separate the stoichiometric matrix by

$$\boldsymbol{N} = \begin{bmatrix} \boldsymbol{N_R} \\ \boldsymbol{N_0} \end{bmatrix}$$

with

$\boldsymbol{N_R} \in \mathbb{R}^{m_0\times n}$	linear independent metabolites
$\boldsymbol{N_0} \in \mathbb{R}^{(m-m_0)\times n}$	linear dependent metabolites

and

$$m_0 = rank(\boldsymbol{N}).$$

From the separation it follows that

$$\boldsymbol{N} = \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{\Lambda_0} \end{bmatrix} \boldsymbol{N_R} = \boldsymbol{\Lambda N_R}$$

from which the conservation relations can be determinded from

$$\boldsymbol{\Gamma} = \begin{bmatrix} -\boldsymbol{\Lambda_0} & \boldsymbol{I} \end{bmatrix}$$

with

$\boldsymbol{\Lambda} \in \mathbb{R}^{m\times m_0}$	link matrix
$\boldsymbol{\Gamma} \in \mathbb{R}^{(m-m_0)\times m_0}$	conservation matrix.

Numerical Conservation Analysis

To separate the stoichiometric matrix, and hence compute the link matrix and the conservation matrix one has to solve

$$\boldsymbol{N^T\Gamma^T} = 0$$

in practice. With increasing system size this becomes numerically expensive.
To approach this challenge we search for a LU Decomposition with full pivoting such that

$$\boldsymbol{PN^TQ = LU}$$

where the upper triangular matrix $\boldsymbol{U}$ can be partitioned as

$$\boldsymbol{U} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{M} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}$$

from which the link matrix, the reduced stoichiometric matrix and the conservation matrix follow to

$$\boldsymbol{\Lambda} = \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{M^T} \end{bmatrix}, \quad \boldsymbol{N_R} = \boldsymbol{\Lambda}^+\boldsymbol{N} \quad \text{and} \quad \boldsymbol{\Gamma} = \begin{bmatrix} -\boldsymbol{M}^T & \boldsymbol{I} \end{bmatrix}$$

respectively.

Due to the use of sparse matrices, we were able to reduce not only memory usage and storage but also lower computation time. With this algorithm it is now possible to handle large scale signalling networks in a reasonable amount of time.

Modular Response Analysis

The unscaled elasticity coefficient matrix

$$\boldsymbol{\epsilon} = \left.\frac{\delta\boldsymbol{v}(\boldsymbol{c}(t))}{\delta\boldsymbol{c}(t)}\right|_{\bar{\boldsymbol{c}}}$$

gives the sensitivities of all local reaction rates to perturbations in all species concentrations.
The dependencies among the species are described by the Jacobian matrix which follows to

$$\boldsymbol{J} = \boldsymbol{N\epsilon}$$

and

$$\boldsymbol{J_R} = \boldsymbol{N_R\epsilon\Lambda}$$

for the full system and the system reduced with the help of the conservation relations, respectively.

After expressing the elasticities as scaled elasticities such that

$$\tilde{\boldsymbol{\epsilon}} = \boldsymbol{\epsilon}\left.\frac{\boldsymbol{c}(t)}{\boldsymbol{v}(\boldsymbol{c}(t))}\right|_{\bar{\boldsymbol{c}}}$$

the local and global response matrices follow to

$$\tilde{\boldsymbol{r}} = \boldsymbol{N_R}\tilde{\boldsymbol{\epsilon}}\boldsymbol{\Lambda}$$

with

$$\tilde{\boldsymbol{r}}_{i,i} = -1 \qquad \forall i \in \{1, 2, ..., n\}$$

and

$$\tilde{\boldsymbol{R}} = \tilde{\boldsymbol{r}}^{-1},$$

respectively.

The local response matrix determines direct information flow between species while the global response matrix describes information flows between species over several intermediates. For both the local and the global response matrix the influence between species can be categorised by

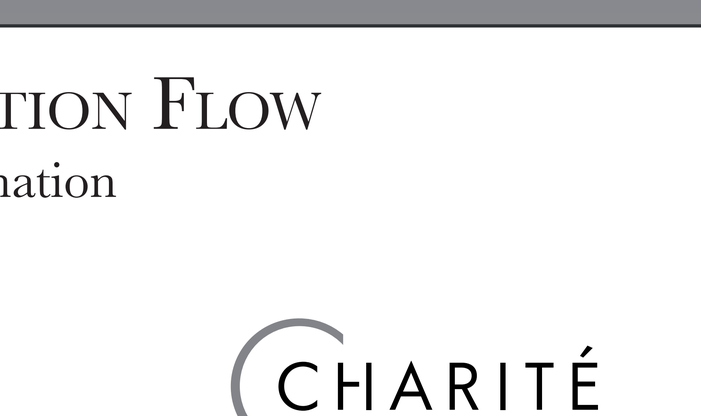
(k,l)-th entry < 0	inactivation of k by l
(k,l)-th entry > 0	activation of k by l
(k,l)-th entry $= 0$	no (direct) influence between k and l

with

$$k, l \in \{1, 2, ..., m_0\}.$$

Wnt Pathway Information Flow

Depending on the presence of a Wnt stimulus the pathway has two different readouts. If Wnt is not present the so-called destruction complex labels Beta-catenin for ubiquitination. If Wnt is present the destruction complex is inhibited by the active form of Dishevelled which allows Beta-catenin to translocate into the nucleus and trigger transcription.



In order to obtain the information flow diagram we assume that the scaled elasticity matrix is not completely known, i.e. the values can vary due to the lack of experimental data, but the signs are determined. Thus, we sample the elasticity coefficients to analyse which interactions are activating, inhibiting or not strictly determined. Furthermore, conservation and modular response analysis was carried out and the information flow model follows to

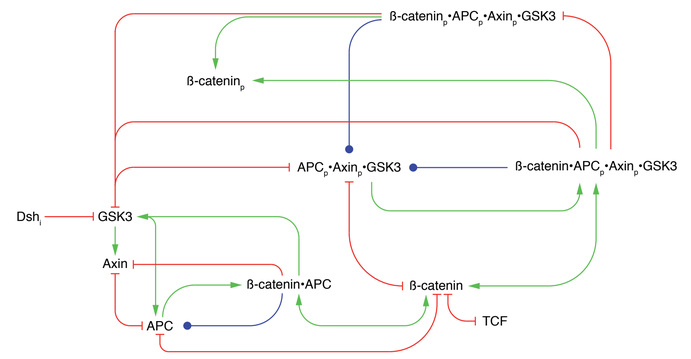


Figure 2: Information flow diagram of the Wnt pathway with inhibitions, activations and varying interactions.

Sign Distribution after Sampling

After performing conservation analysis the given scaled elasticity coefficients were sampled by random numbers between 0 and 1 while preserving the signs. The fractions of the entries in the local response matrix that are always zero, always positive, always negative and changing between positive and negative sign over all samples were ploted in Figure 3.

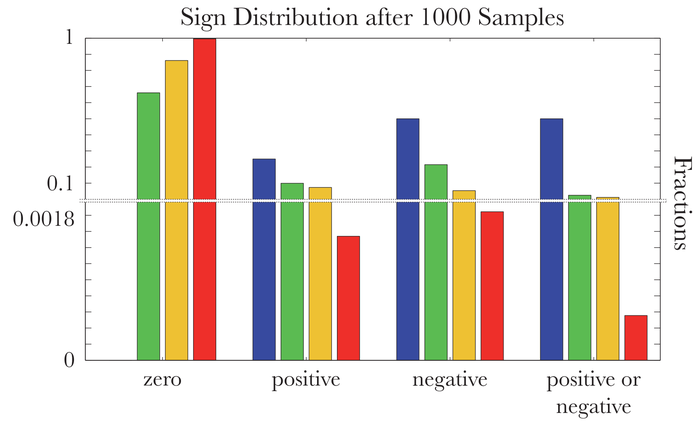


Figure 3: Sign distribution in the local response coefficients after sampling the scaled elasticity coefficients 1000 times for the following 4 models: Example model with 4 species and 3 reactions (cf. Figure 1), Wnt model with 15 species and 17 reactions (cf. Figure 2), Mapk model with 97 species and 148 reactions and Reactome model with 6232 species and 3652 reactions.

Conclusions

Conservation and modular response analysis

Based on sparse matrix operations we developed an algorithm to carry out not only conservation analysis but also perfom modular response analysis to aid in the understanding of signal transduction networks. Formerly hidden network properties can henceforth be revealed effectively even for large and complex systems.

Sampling experiments

In conjunction with the conservation and modular response analysis, the sampling of the usually weakly determined scaled elasticity coefficients unveiled a large definiteness of the interactions in signaling networks. More than 99% of the signs of the local response matrix could be identified uniquely from the structure of the reaction network, i.e. even without knowing the reaction network in detail its information flow can be calculated.

Vallabhajosyula et al. - Bioinformatics (2006) vol. 22 (3) pp. 346-53
Lee et al. - PLoS Biol (2003) vol. 1 (1) pp. E10
Bruggeman et al. - J Theor Biol (2002) vol. 218 (4) pp. 507-20
Kholodenko et al. - Proc Natl Acad Sci USA (2002) vol. 99 (20) pp. 12841-6
Schoeberl et al. - Nat Biotechnol (2002) vol. 20 (4) pp. 370-5
Kholodenko et al. - FEBS Lett (1997) vol. 414 (2) pp. 430-4
Reder. - J Theor Biol (1988) vol. 135 (2) pp. 175-201
Vastrik et al. - Genome Biol (2007) vol. 8 (3) pp. R39

P.S.: Thanks a lot to the GENESYS network for travel support.