Designing optimal cell factories: integer programming couples elementary mode analysis with regulation

Molecular Networks B SS13

Jonas Ibn-Salem

25.04.13

Freie Universität Berlin

Paper and authors

Jungreuthmayer and Zanghellini *BMC Systems Biology* 2012, **6**:103
http://www.biomedcentral.com/1752-0509/6/103

BMC
Systems Biology

METHODOLOGY ARTICLE **Open Access**

Designing optimal cell factories: integer programming couples elementary mode analysis with regulation

Christian Jungreuthmayer[1,2] and Jürgen Zanghellini[1,2*]

source: biotec.boku.ac.at

source: biotec.boku.ac.at

[1] Austrian Centre of Industrial Biotechnology, Vienna, Austria

[2] Department of Biotechnology, University of Natural Resources and Life Sciences, Vienna, Austria

Overview

1. Background and Motivation
2. Theory
3. Illustrative example
4. Including regulation
5. (Optimizing metabolic functionality)
6. Results of realistic example
7. Discussion
8. Summary and conclusions

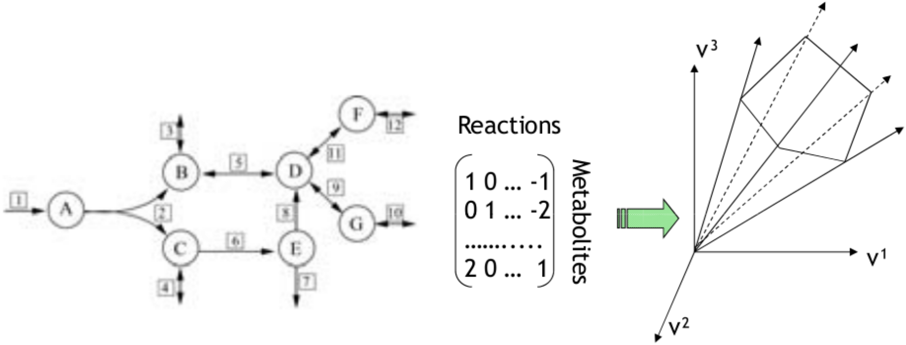
Background and Motivation

- microbes can be used for metabolic engineering e.g. ethanol production in *E. coli*
- Methabolic networks are available for some model organisms

Gaol

Gene deletion strategie to optimize the production efficiency of strains.

Metabolic network



source A.Bockmayer lecture slides Algorithms in Systembiology SS12
Model of metabolic network with

- m internal metabolites
- n reactions
- $S \in \mathbb{R}^{m \times n}$ stoichiometric matrix
- $\hat{v} \in \mathbb{R}^n$ flux vector

fig?

Steady state assumption

Steady state assumption

Metabolite concentrations and reaction rates are constant.

$$S \cdot \hat{v} = 0 \tag{1}$$

Steady state flux coen:
$C = \{v \in \mathbb{R}^n | Sv = 0, v_i \geq 0, i \in Irr\}$

Elementary mode analysis

- For $v \in \mathbb{R}^n$, the support of v is defined as $supp(v) = \{i \in \{1, ..., n\} | v_i \neq 0\}$
- $v \in C$ is an *elementary flux mode* if $supp(v)$ is minimla w.r.t $\subseteq$ i.e, if there is no $v' \in C, v' \neq 0$ with $supp(v') \subsetneq supp(v)$.

Definition

An elementary mode (EM) is a minimal and indivisible set of reactions that operates under steady state conditions, while obeying all (ir-)reversibility constraints.

Network of minimal functionality (NMF)

Theory: Elementary mode analysis

$e = e(\hat{e})$ is the binary representation of a EM e.

$$e_i := e(\hat{e}_i)) = \begin{cases} 1 & \text{if } \hat{e}_i \neq 0 \\ 0 & \text{if } \hat{e}_i = 0 \end{cases} \tag{2}$$

$e^T v \leq e^T e = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} e_i =: ||e||$

Binary linear program (BLP)

Group all q elementary modes:

$$G := (e_1, ..., e_r)^T \qquad \textit{goal matrix} \text{ with desirable EM}$$

$$H := (e_{r+1}, ..., e_{r+s})^T \qquad \textit{helper matrix} \text{ tolerated EM}$$

$$K := (e_{r+s+1}, ..., e_{r+s+t})^T \qquad \textit{kill matrix} \text{ unwanted EM}$$

where $q = r + s + t$

Binary linear program (BLP)

$max\ ||x||$

s.t.

$$
\begin{aligned}
e_g^T x &= ||e_g|| && g \in \{1, ..., r\} \\
e_h^T x &\leq ||e_h|| && h \in \{r+1, ..., r+s\} \\
e_k^T x &\leq ||e_k|| - 1 && k \in \{r+s+1, ..., r+s+t\}
\end{aligned}
$$

$$\Delta_{min} = n - ||x||$$

Alternative solutions

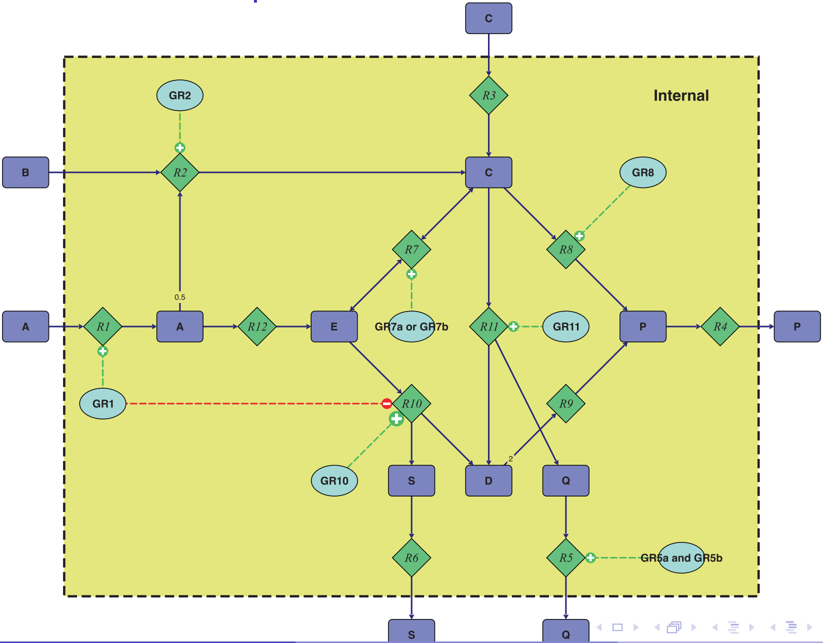
- The BLP may have no or a finite number of solutions.
- Iteratively exclude existing solutions $x^{(j)}$ by adding constraints:

$$\sum_{i \in B} x_i \leq |B| - 1$$
$$\sum_{i \in N} x_i \geq 1$$

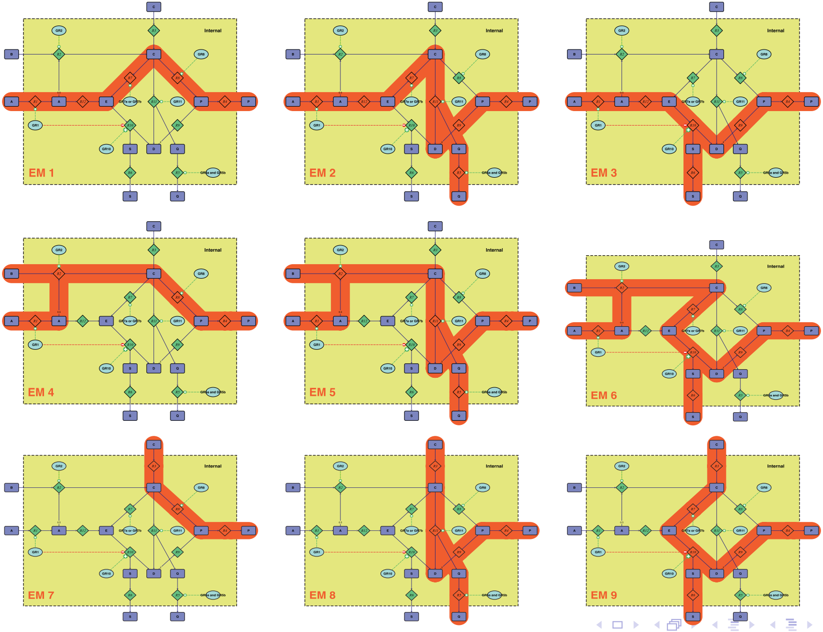
where

$$B = \{i | x_i^{(j)} = 1\}$$
$$N = \{i | x_i^{(j)} = 0\}$$

Illustrative example



Elementary modes



List of elementary modes

	EM flux vector, $\hat{e}_i$											
	R1	**R2**	**R3**	**R4**	**R5**	**R6**	**R7**	**R8**	**R9**	**R10**	**R11**	**R12**
EM 1	1.0	0.0	0.0	1.0	0.0	0.0	1.0	1.0	0.0	0.0	0.0	1.0
EM 2	1.0	0.0	0.0	0.5	1.0	0.0	1.0	0.0	0.5	0.0	1.0	1.0
EM 3	1.0	0.0	0.0	0.5	0.0	1.0	0.0	0.0	0.5	1.0	0.0	1.0
EM 4	0.5	1.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0
EM 5	0.5	1.0	0.0	0.5	1.0	0.0	0.0	0.0	0.5	0.0	1.0	0.0
EM 6	0.5	1.0	0.0	0.5	0.0	1.0	-1.0	0.0	0.5	1.0	0.0	0.0
EM 7	0.0	0.0	1.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0
EM 8	0.0	0.0	1.0	0.5	1.0	0.0	0.0	0.0	0.5	0.0	1.0	0.0
EM 9	0.0	0.0	1.0	0.5	0.0	1.0	-1.0	0.0	0.5	1.0	0.0	0.0

List of binary elementary modes

	Binary representation, e_i, of EM flux vector, $\hat{e}_i$													
	R1	**R2**	**R3**	**R4**	**R5**	**R6**	**R7**	**R8**	**R9**	**R10**	**R11**	**R12**	$\|\|e_i\|\|$	
$G =$	1	0	0	1	0	0	1	1	0	0	0	1	5	$= \|g\|$
	1	0	0	1	1	0	1	1	1	0	1	1	8	
	1	0	0	1	0	1	0	0	1	1	0	1	6	
$K =$	1	1	0	1	0	0	0	1	0	0	0	0	4	$= \|k\|$
	1	1	0	1	1	0	0	0	1	0	1	0	6	
	1	1	0	1	0	1	1	0	1	1	0	0	7	
	0	0	1	1	0	0	0	1	0	0	0	0	3	
$H =$	0	0	1	1	1	0	0	0	1	0	1	0	5	$= \|h\|$
	0	0	1	1	0	1	1	0	1	1	0	0	6	

BLP for example network

$$max \sum_{i=1}^{12} x_i$$

subject to

$$x_1 + x_4 + x_7 + x_8 + x_{12} = 5$$

$$x_3 + x_4 + x_8 \leq 3$$
$$x_3 + x_4 + x_5 + x_9 + x_{11} \leq 5$$
$$x_3 + x_4 + x_6 + x_7 + x_9 + x_{10} \leq 6$$
$$x_1 + x_4 + x_5 + x_7 + x_8 + x_9 + x_{11} + x_{12} \leq 7$$
$$x_1 + x_4 + x_6 + x_9 + x_{10} + x_{12} \leq 5$$

$$x_1 + x_2 + x_4 + x_8 \leq 3$$
$$x_1 + x_2 + x_4 + x_5 + x_9 + x_{11} \leq 5$$
$$x_1 + x_2 + x_4 + x_6 + x_7 + x_9 + x_{10} \leq 6$$

Minimal cut sets

	w_1				
i	**minimal cut set**			f_i	
1	R2	R9		9.0	
2	R2	R5	R6	8.0	
3	R2	R5	R10	8.0	*
4	R2	R6	R11	8.0	
5	R2	R10	R11	8.0	*

Biological feasibility

Problem:

- Distinguish chemical from genetic interventions
- Gene-enzyme-reaction mapping not always known.

Solution:

- add weights to objective function

$$max||w^T x||$$

- low weights for uptake reactions
- high weights for reactions with missing genetic information

Example

$$w_2^T = (0.1,\ 0.1,\ 0.1,\ 99,\ 1,\ 99,\ 2,\ 1,\ 99,\ 1,\ 1,\ 99)$$

Minimal cut sets wights wights

	w_1					w_2				
i	**minimal cut set**			f_i		**minimal cut set**			f_i	
1	R2	R9		9.0		R2	R5	R10	301.2	*
2	R2	R5	R6	8.0		R2	R10	R11	301.2	*
3	R2	R5	R10	8.0	*	R2	R9		204.2	
4	R2	R6	R11	8.0		R2	R5	R6	203.2	
5	R2	R10	R11	8.0	*	R2	R6	R11	203.2	

Introduction: Fixed Point Analysis

Given the system of differential equations:

$$y' = f(y)$$

Definition

A *fixed point* y^* is defined by $f(y^*) = 0$.

- Solve the equation $f(y) = 0$
- Analyse eigenvalues of the Jacobian at fixed points.

Now: System with *controle parameter* μ.

$$y' = f(y, \mu)$$

How does μ influence the number, location and stability of fixed points?