Summer 6-5-2012

Minimal Cut Sets as Computational Tool in Metabolic Engineering

Steffen Klamt
Max Planck Institute Magdeburg

Minimal Cut Sets as Computational Tool in Metabolic Engineering

Novel Theoretical Results and Their Applications

Steffen Klamt

Max Planck Institute for Dynamics of Complex Technical Systems
Magdeburg, Germany

Stoichiometric Approaches for Metabolic Engineering and Rational Strain Optimization

FBA-based optimization:

OptKnock (Maranas et al.) and variants: identify minimal set of knockouts that lead to coupling between biomass and product synthesis

→ bi-level optmization problems solved by Mixed Integer Programming

→ one solution

Elementary-modes-based optimization:

Minimal metabolic functionality (Trinh, Srienc et al.): identify knockout set that keeps only a small number of (optimal) elementary modes functional

→ one solution

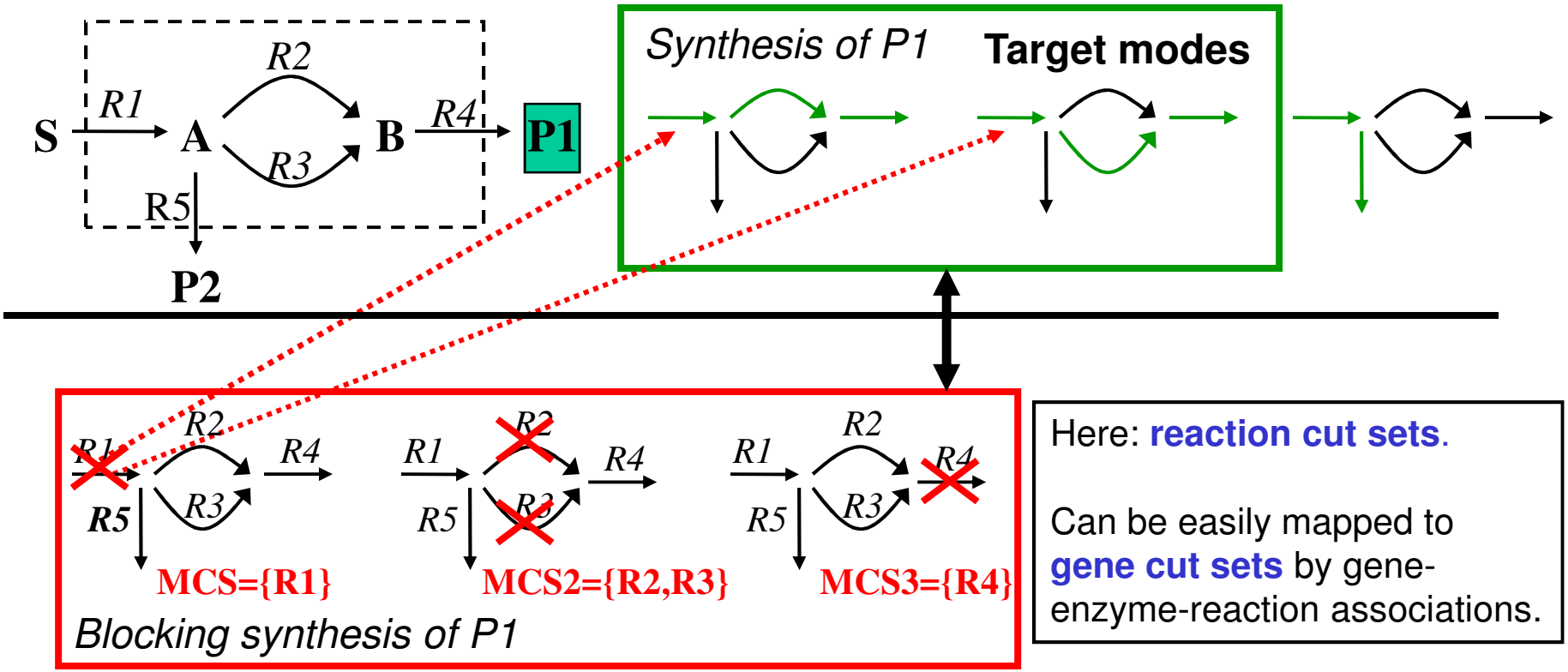
(Constrained) Minimal Cut Sets: enumeration of all minimal knockout sets that block *undesired* and keep *desired* behaviors.

→ based on elementary modes

→ all knockout solutions are generated

From Elementary Modes to Minimal Cut Sets

Elementary Modes: minimal functional units (pathways) that can operate in steady state



Here: **reaction cut sets**.

Can be easily mapped to **gene cut sets** by gene-enzyme-reaction associations.

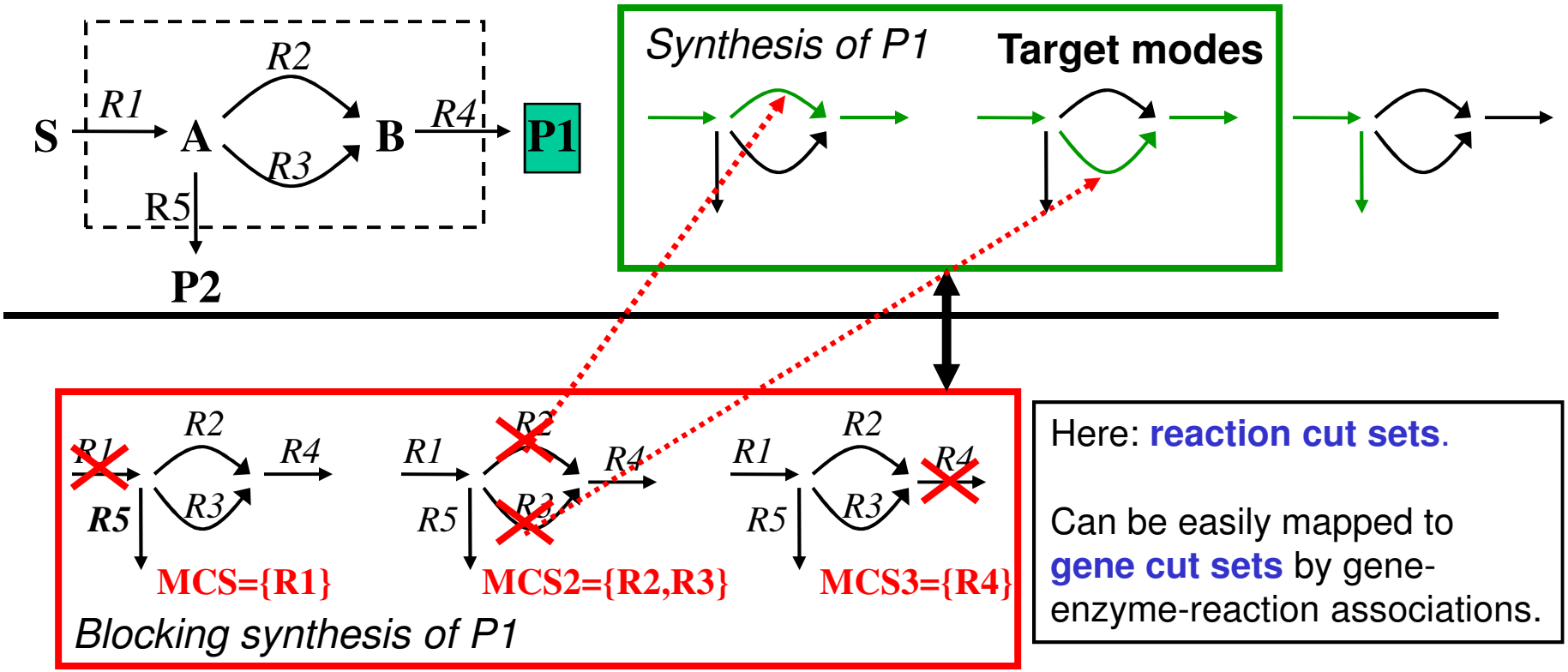
Minimal Cut Set (MCS): minimal set of cuts (knockouts) blocking certain functions in steady state

***MCSs** are the minimal hitting sets of the corresponding EMs*

Klamt S: BioSystems 2006, 83:233-237.

From Elementary Modes to Minimal Cut Sets

Elementary Modes: minimal functional units (pathways) that can operate in steady state



Here: **reaction cut sets**.

Can be easily mapped to **gene cut sets** by gene-enzyme-reaction associations.

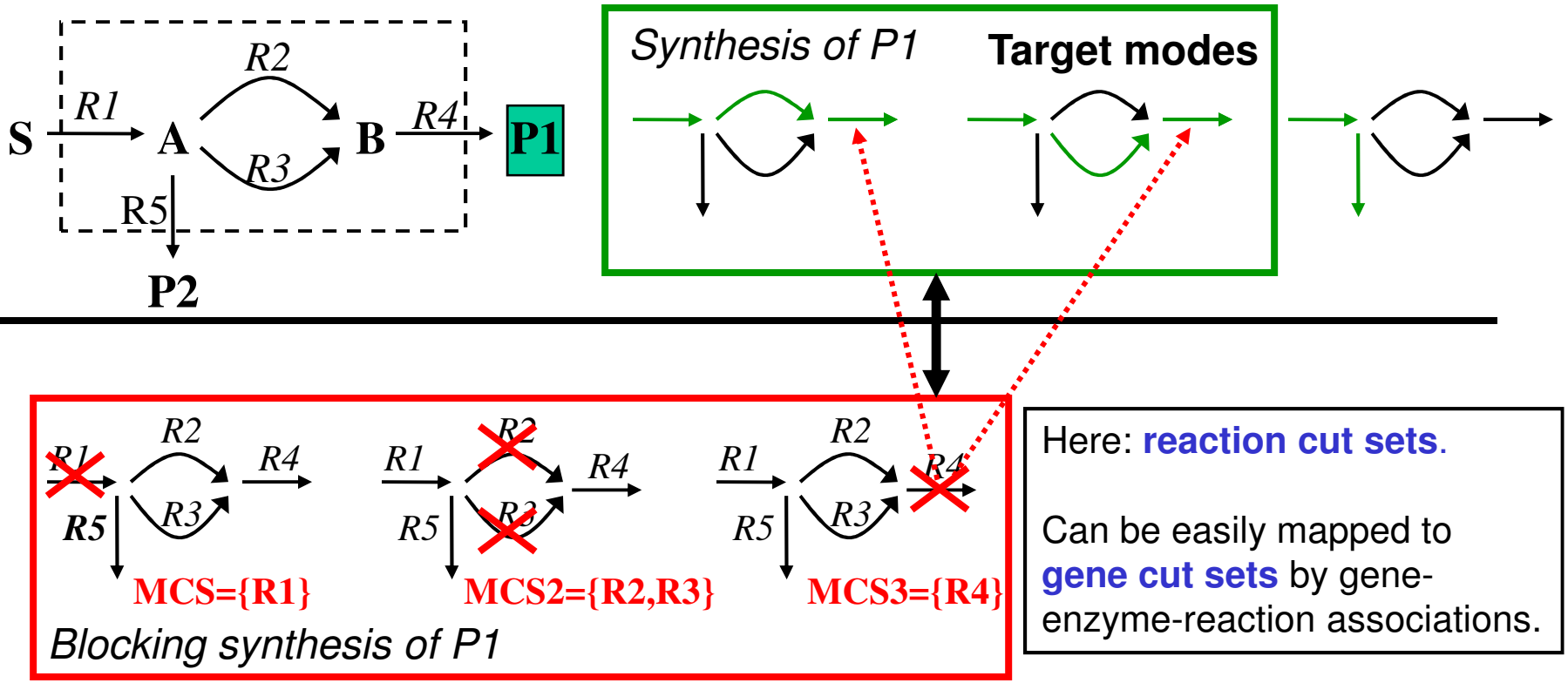
Minimal Cut Set (MCS): minimal set of cuts (knockouts) blocking certain functions in steady state

***MCSs** are the minimal hitting sets of the corresponding EMs*

Klamt S: BioSystems 2006, 83:233-237.

From Elementary Modes to Minimal Cut Sets

Elementary Modes (EM)*:* minimal functional units that can operate in steady state



Here: **reaction cut sets**.

Can be easily mapped to **gene cut sets** by gene-enzyme-reaction associations.

Minimal Cut Set (MCS)*:* minimal set of cuts (knockouts) blocking certain functions in steady state

***MCSs** are the minimal hitting sets of the corresponding EMs*

Klamt S: BioSystems 2006, 83:233-237.

Elementary Modes, Minimal Cut Sets and Minimal Hitting Sets

Computation of Minimal Hitting Sets (here: for a given set of Target Modes):

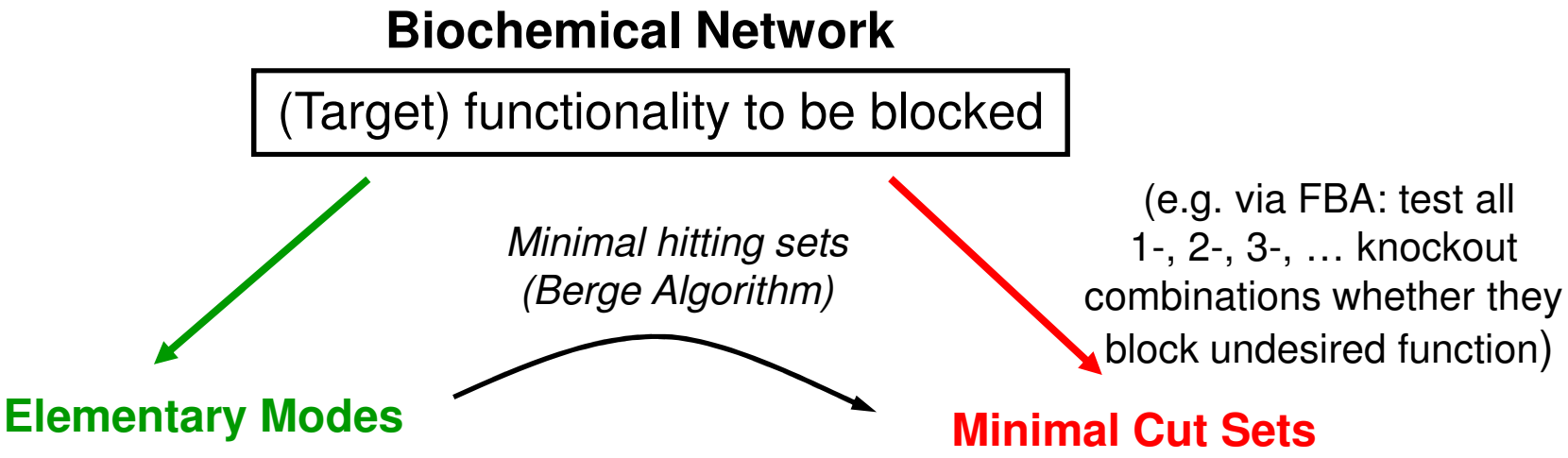
Many algorithms are known from hypergraph theory.

Algorithm of Claude Berge [1] performs well for metabolic networks [2].

[1] Claude Berge. (1989): Hypergraphs. Combinatorics of finite sets.(North-Holland, Amsterdam)

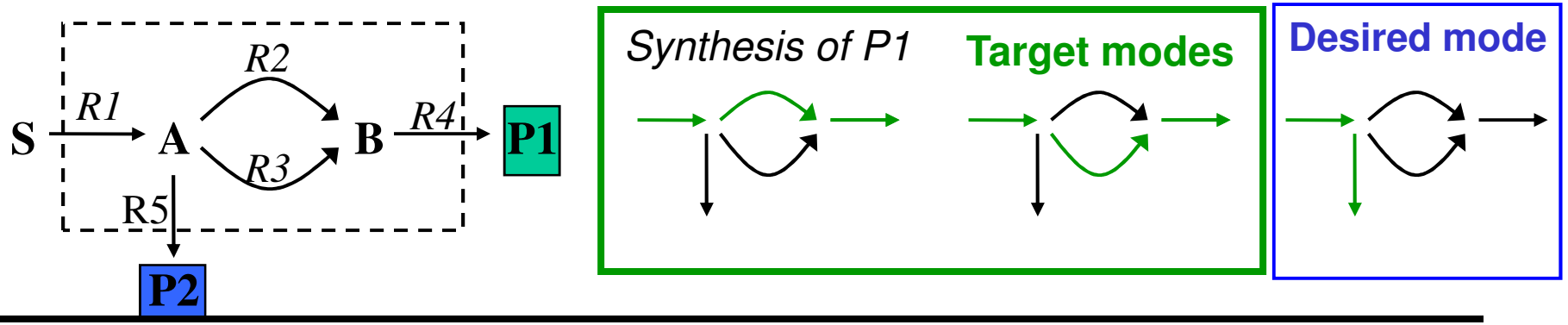
[2] Haus U, Klamt S, Stephen T. (2008) Computing knockout strategies in metabolic networks. J Comp Biol 15: 259-268.

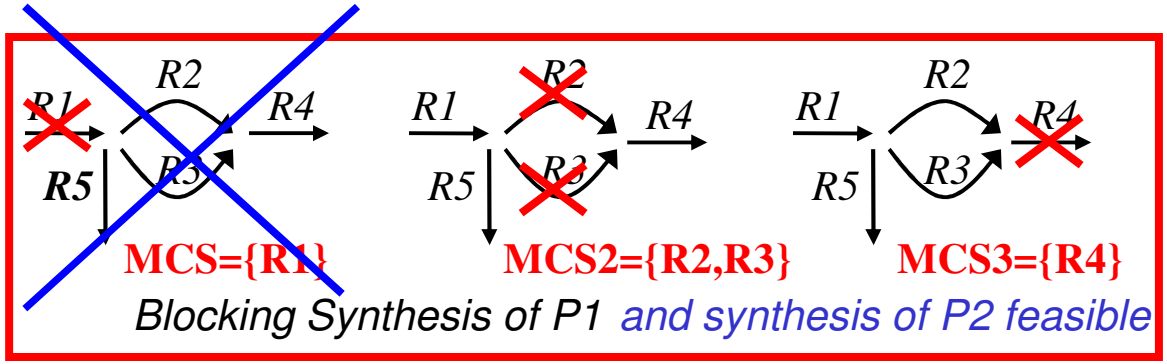
Elementary Modes, Minimal Cut Sets and Minimal Hitting Sets



(Constrained) Minimal Cut Sets for Rational Redesign of Metabolic Networks

Elementary Modes: minimal functional units (pathways) that can operate in steady state





Constrained MCSs:
- specify „desired modes“ (must not be hit by MCSs)

Constrained Minimal Cut Set (MCS): minimal cut sets blocking undesired while keeping desired metabolic behaviors.

Hädicke and Klamt (2011): Computing Complex Metabolic Intervention Stratgeies using Constrained Minimal Cut Sets. Metabolic Engineering, 13:204-213.

Generalization: Constrained MCSs

Constrained Minimal Cut Set (cMCSs) problem:

- set $\mathcal{T}$ of target EMs (to be blocked)
- set $\mathcal{D}$ of desired EMs
- n: minimal number of desired EMs in $\mathcal{D}$ that must not be hit by the MCSs

→ A constrained MCS **C** fulfills: $\mathbf{C} \cap \mathbf{T} \neq \emptyset \ \forall \mathbf{T} \in \mathcal{T}; \quad |\{\mathbf{D} \in \mathcal{D} \mid \mathbf{C} \cap \mathbf{D} = \emptyset\}| \geq n$

- algorithm for minimal hitting set calculation adapted for cMCSs:
  a) check on-the-fly whether MCS candidates keep desired modes (often faster)
  b) identify constrained MCSs from unconstrained MCSs during post-processing

- large variety of complex intervention problems can be conveniently formulated and solved by cMCSs

- many other methods (including OptKnock, Minimal Metabolic Functionality) can be reformulated as special cMCS problems

- all possible knockout strategies (also with higher cardinalities) can be found
  (... if the network is not too large)

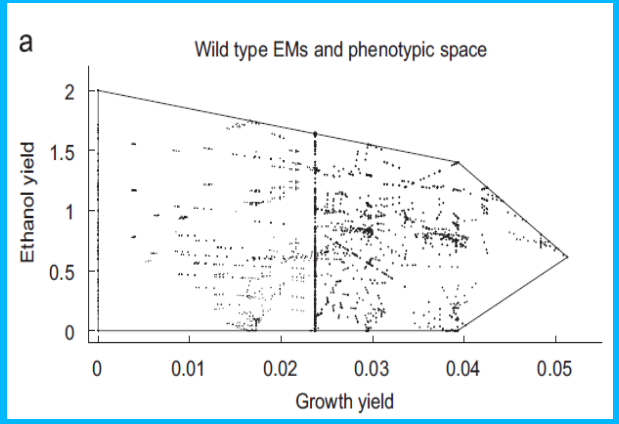
Hädicke and Klamt (2011): Computing Complex Metabolic Intervention Stratgeies using Constrained Minimal Cut Sets. Metabolic Engineering, 13:204-213.

Example: Coupled Biomass and Ethanol Synthesis in *E.coli*

Goal: search for interventions that lead to high (anaerobic) ethanol synthesis still enabling some formation of biomass

1) Compute elementary modes and inspect phenotypic space

(each dot = one mode)

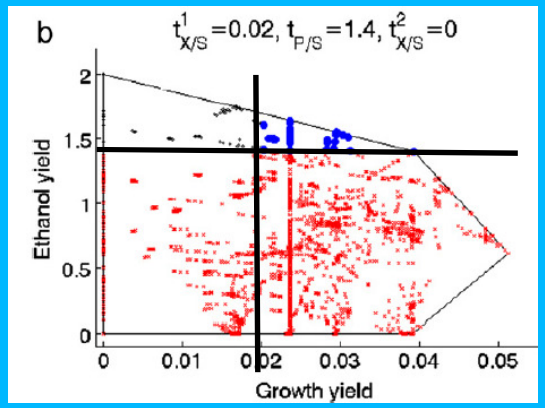


2) Specify target and desired modes

Red dots = target modes

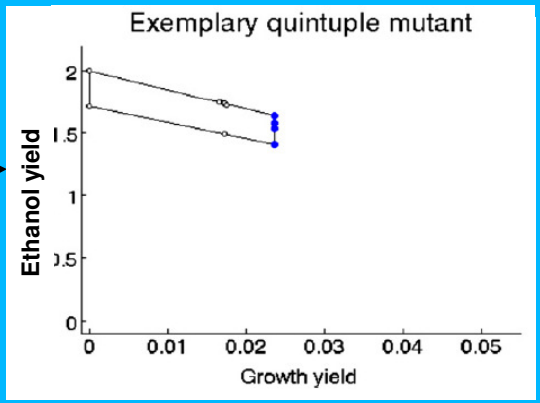
Blue dots = desired modes

Black dots = „don‘t care“



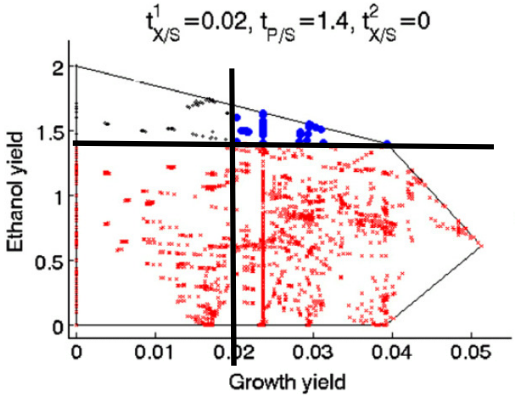
3) Compute constrained MCSs and implement the most practical

with n=1

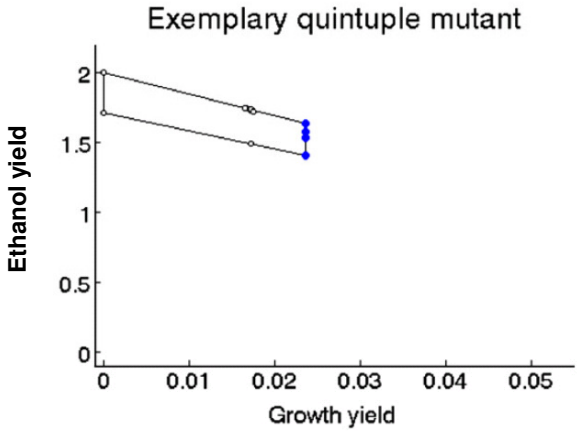


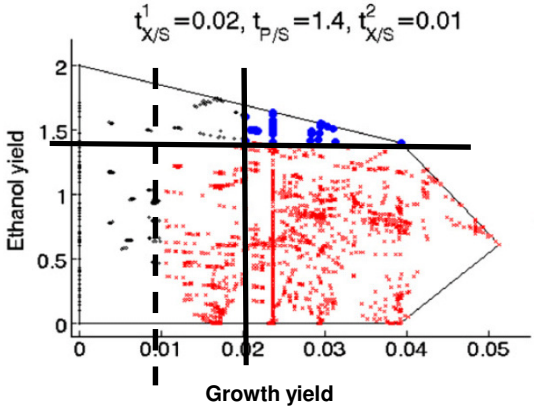
Example: Coupled Biomass and Ethanol Synthesis in *E.coli*

Testing different scenarios



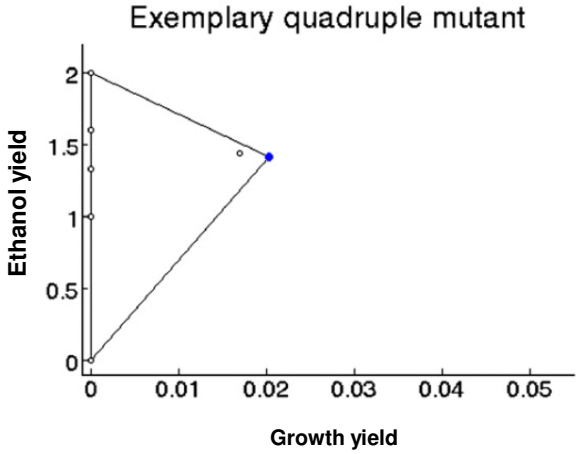
At least 5 knock-outs necessary





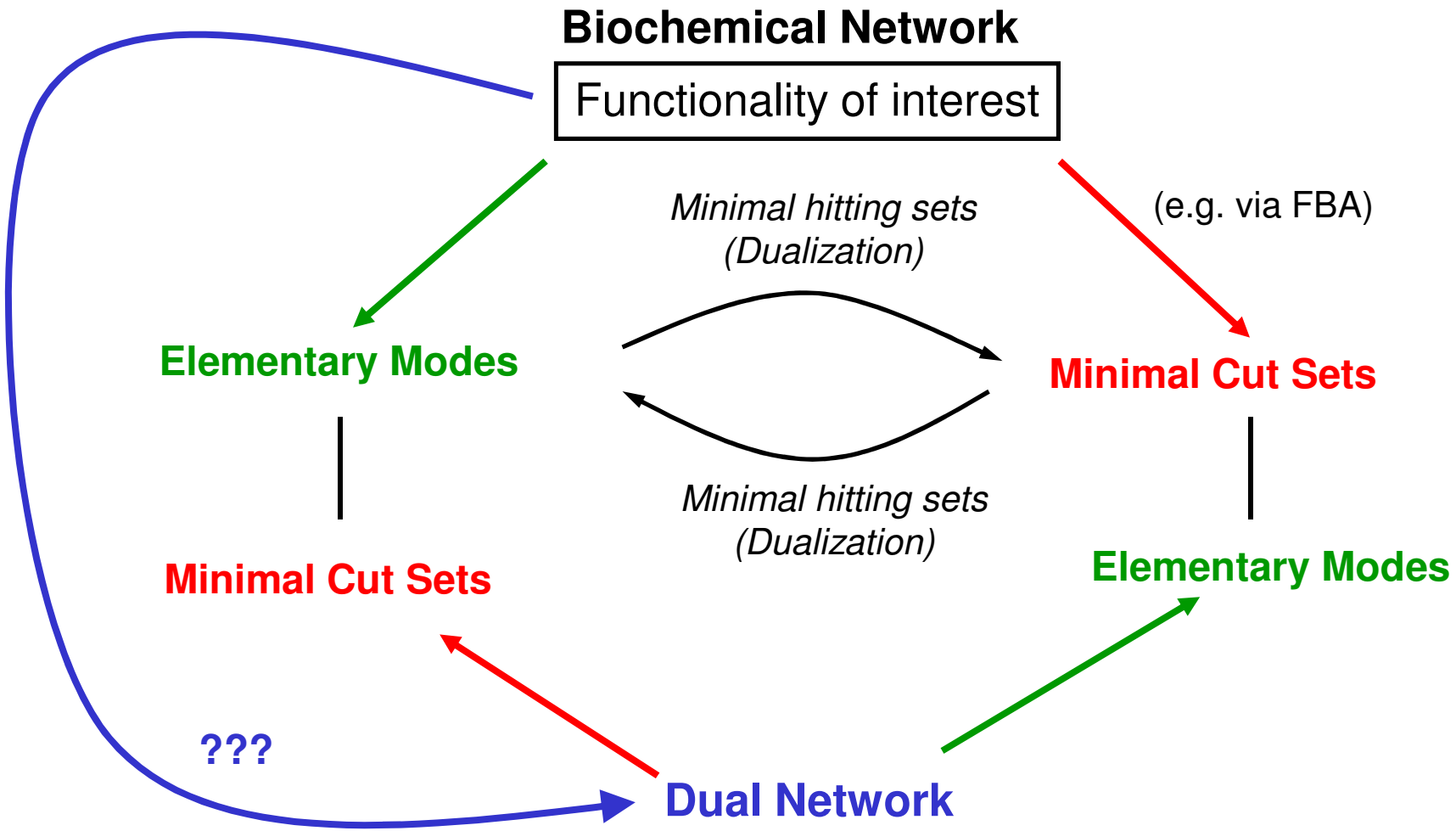
At least 4 knock-outs necessary

… smaller set of target modes
→ less knockouts required
→ but modes with low or no ethanol production exist (at low growth rates)



Hädicke and Klamt (2011): Computing Complex Metabolic Intervention Stratgeies using Constrained Minimal Cut Sets. Metabolic Engineering, 13:204-213.

Duality: Elementary Modes and Minimal Cut Sets



How to Get the Dual Network

- specify the set $\boldsymbol{T}$ of target flux vectors (to be blocked by MCS) by a vector $\mathbf{t}$:

$$\boldsymbol{T} = \{\mathbf{r}:\ \mathbf{Nr} = \mathbf{0},\ \mathbf{r}_{Irrev} \geq \mathbf{0},\ \mathbf{t}^T\mathbf{r} > 0\ \}$$

(e.g. $\mathbf{t}^T$=(0,0,0,1,0) if the 4-th reaction produces undesired product)

Define Inconsistent System S

$$\mathbf{Nr} = \mathbf{0}$$

$$\mathbf{I}_{Irrev}\mathbf{r}_{Irrev} \geq \mathbf{0}$$

$$\mathbf{t}^T\mathbf{r} \geq 1$$

$$\mathbf{Ir} = \mathbf{0}$$

Farkas Lemma
Theory of IISs
(Ryan et al.)

Consider System D

$$(\mathbf{I}\ -\mathbf{t}\ -\mathbf{I}^T_{irrev}\ \mathbf{N}^T)\begin{pmatrix}\mathbf{v}\\ w\\ \mathbf{z}\\ \mathbf{u}\end{pmatrix} = \mathbf{0};\quad \mathbf{z} \geq \mathbf{0}; w \geq 0$$

$\mathbf{N}_{dual}$ $\mathbf{r}_{dual}$

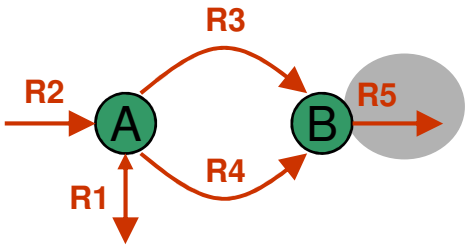
Minimal cut sets are the *irreducible inconsistent subsystems* (IISs) of **S** keeping the system inconsistent with a minimal subset of the equations **Ir**=**0**.

The IISs of **S** (corresponding to the MCSs) are the elementary modes of system D which are minimal in **v** and have support in *w*.

Ballerstein K, von KampA, Klamt S and Haus UU. (2012) Minimal cut sets in a metabolic network are elementary modes in a dual network. Bioinformatics 18: 381-387.

How to Get the Dual Network: Example

Primal Network



$$\mathbf{N} = \begin{array}{c} \begin{matrix} \mathbf{R1} & \mathbf{R2} & \mathbf{R3} & \mathbf{R4} & \mathbf{R5} \end{matrix} \\ \begin{pmatrix} -1 & 1 & -1 & -1 & 0 \\ 0 & 0 & 1 & 1 & -1 \end{pmatrix} \begin{matrix} \mathbf{A} \\ \mathbf{B} \end{matrix} \end{array}$$

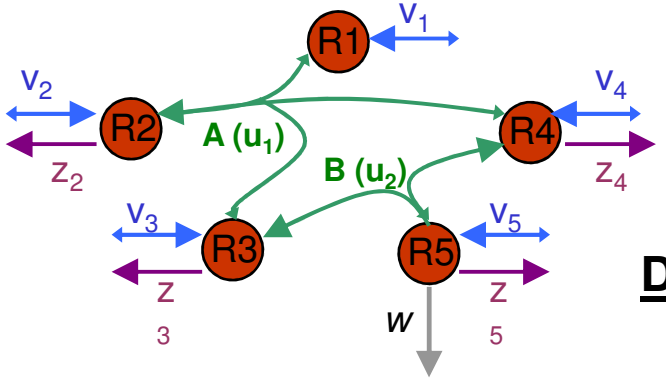
Reversible reaction: R1

Target reaction: R5

$\mathbf{t}^T\mathbf{r} = (0,0,0,0,1)\,\mathbf{r} > 0$

Target EMs (with R5):
- EM1={R1,R3,R5}
- EM2={R1,R4,R5}
- EM3={R2,R3,R5}
- EM4={R2,R4,R5}

MCSs for target reaction R5
- MCS1={R1,R2}
- MCS2={R3,R4}
- MCS3={R5}



Dual Network

$$\mathbf{N}_{dual} = (\mathbf{I}_{MCS} \; -\mathbf{t} \; -\mathbf{I}_{irrev} \; \mathbf{N}^T) =$$

	$\mathbf{I}_{MCS}$					$-\mathbf{t}$	$-\mathbf{I}^T_{irrev}$				$\mathbf{N}^T_A$		$\mathbf{N}^T_B$	
	v_1	v_2	v_3	v_4	v_5	w	z_2	z_3	z_4	z_5	u_1		u_2	
R1	1	0	0	0	0	0	0	0	0	0	−1		0	R1
R2	0	1	0	0	0	0	−1	0	0	0	1		0	R2
R3	0	0	1	0	0	0	0	−1	0	0	−1		1	R3
R4	0	0	0	1	0	0	0	0	−1	0	−1		1	R4
R5	0	0	0	0	1	−1	0	0	0	−1	0		−1	R5

Reversible reactions: $v_1 \ldots v_5$ and u_1, u_2 (former metabolites A and B in $\mathbf{N}^T$)

EMs with support in *w* and minimal support w.r.t reactions $v_1 \ldots v_5$:
- EM1={v_1,v_2,u_1,u_2,w}
- EM2={v_3,v_4,u_2,w}
- EM3={v_5,w}

MCSs hitting the three EMs above by cutting only $v_1,\ldots,v_5$:
- MCS1={v_1,v_3,v_5}
- MCS2={v_1,v_4,v_5}
- MCS3={v_2,v_3,v_5}
- MCS4={v_2,v_4,v_5}

*Ballerstein K, von Kamp A, Klamt S, Haus UU (2012): Minimal cut sets in a metabolic network are elementary modes in a dual network. Bioinformatics, **28**:381-387.*

Duality: Elementary Modes and Minimal Cut Sets

Implementation details:
- compress matrix **N** before dualization
- exploit that "**z**" reactions are parallel to "**v**" reactions
- the **u** reactions does not need to be processed
  (only interested in minimality of **v** reactions)

First results (α version):
- computation time similar to EM + Minimal Hitting Set calculation
  - but: still room for improvements

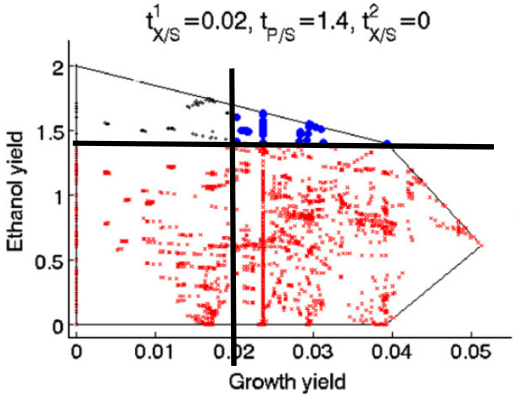
Generalization:
- incorporation of inhomogeneous constraints straightforward …

*Ballerstein K, von Kamp A, Klamt S, Haus UU (2012): Minimal cut sets in a metabolic network are elementary modes in a dual network. Bioinformatics, **28**:381-387.*

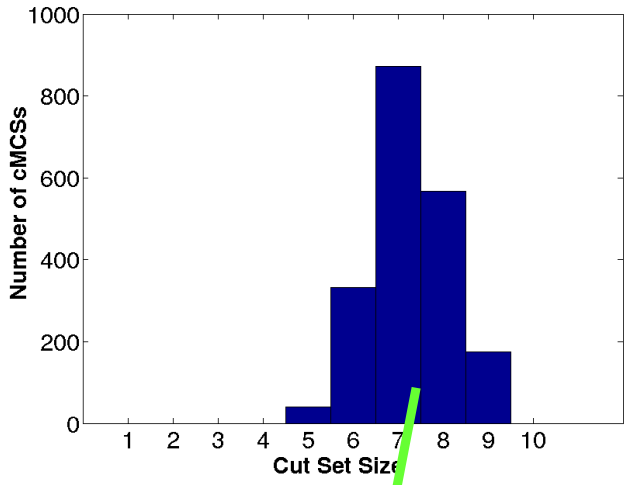
Coupled Biomass and Ethanol Synthesis in *E.coli* with Inhomogeneous Constraints

Homogeneous system:

Nr = 0; $r_{Irrev} \geq 0$



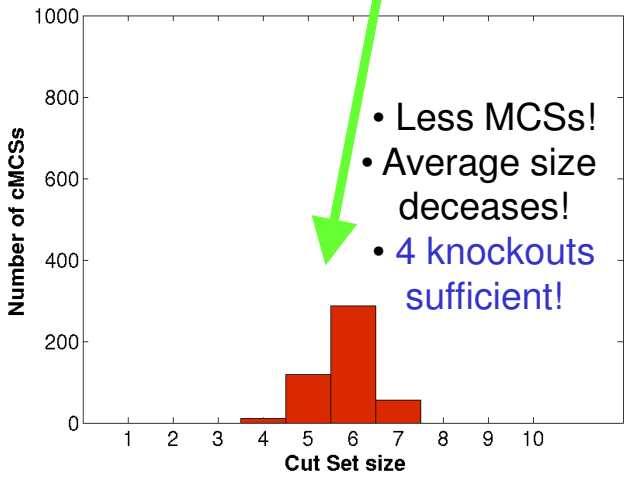
At least 5 knock-outs necessary



Inhomogeneous system (considering ATP maintenance)

Nr = 0;
$r_{Irrev} \geq 0$
$0 \leq r_{Gluc\_up} \leq 10$ mmol/(gDW*h);
$r_{ATPmaintenance} \geq 8.4$ mmol/(gDW*h)

(+ same thresholds for desired/target modes as above)



Computation of EMs / MCSs in Genome-Scale Networks ?!

- EM / MCSs calculation only possible in networks with moderate size
  - → useful for models of central metabolism
  - → for many products: suitable interventions lie often in central metabolism

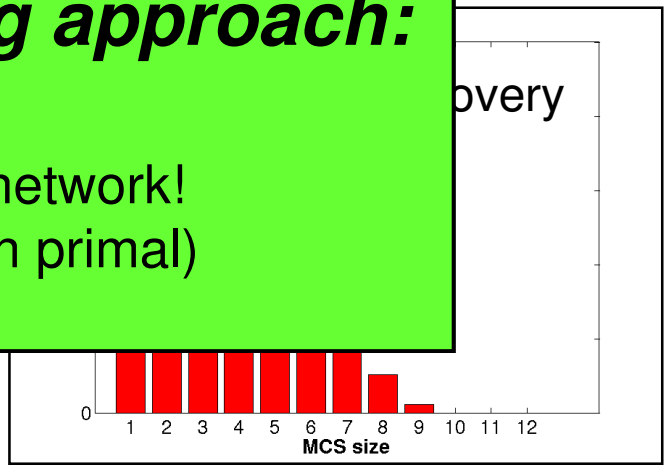
- possible solution for large-scale applications: EM / MCS sampling

  - → approximate combinatorial properties of EMs by computing a re… …leta)
  - → co…

First test … …overy
sufficient t…
smaller siz…
h…

Alternative sampling approach:

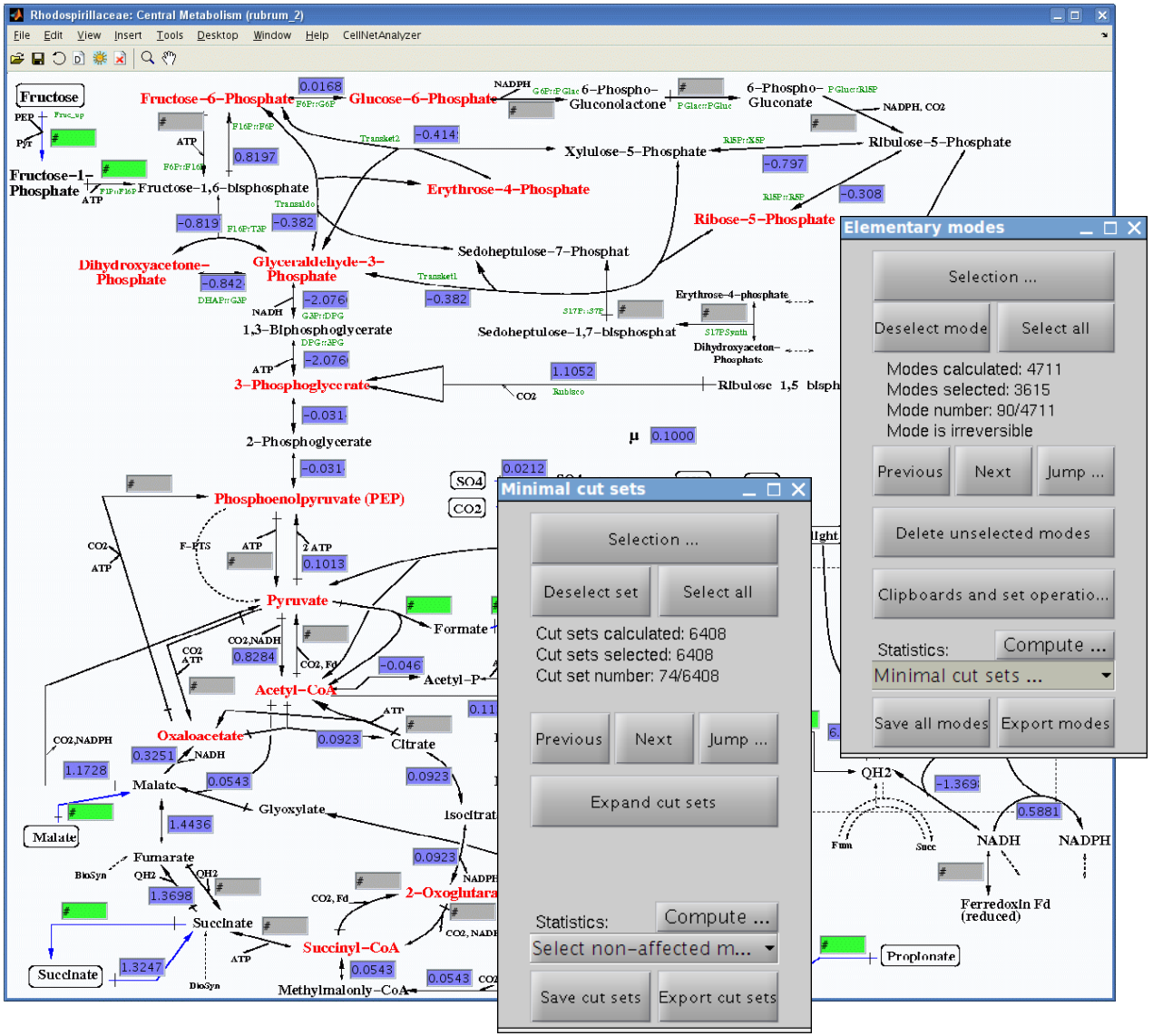
Sample EMs in dual network!
(= sample of MCSs in primal)



CellNetAnalyzer

Large variety of methods for computing and analysing EMs and (c)MCSs.

Experimental work based on cMCSs engineering strategies underway.



Conclusions

- **Constrained Minimal Cut Sets**: *blocking undesired* while *keeping desired* metabolic behaviors.
  → very flexible and convenient approach for enumerating intervention strategies for metabolic engineering.

- **Duality** between EMs and MCSs offers new *computational* and *conceptual* perspectives for studying functions and (re)design strategies in metabolic networks.

- **Inhomogeneous constraints** can be considered.

- **Sampling approaches** for large-scale networks.

- Algorithms for (c)MCSs computation and analysis implemented in ***CellNetAnalyzer***.

- Similar approach for regulatory/signaling networks: **Minimal Intervention Sets**.

Samaga R, von Kamp A and Klamt S (2010) Computing combinatorial intervention strategies and failure modes in signaling networks. *Journal of Computational Biology* **17**:39-53.

Acknowledgements

• Oliver Hädicke
(MPI Magdeburg)

• Axel von Kamp
(MPI Magdeburg)

• Kathrin Ballerstein (ETH Zurich)
• Utz-Uwe Haus (ETH Zurich)

• Christoph Kaleta (University Jena)

Funding:

Magdeburg Centre for
Systems Biology (by BMBF)

CellNetAnalyzer

http://www.mpi-magdeburg.mpg.de/projects/cna/cna.html


MAX-PLANCK-GESELLSCHAFT


Centre for Systems
MaCS
Biology
Magdeburg

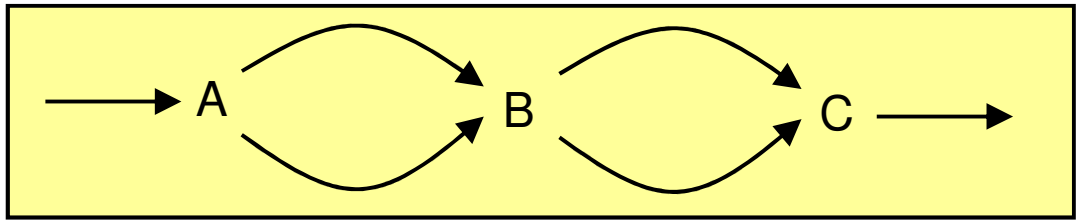
THANK YOU!


MAX-PLANCK-GESELLSCHAFT

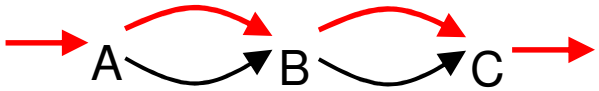

Centre for Systems Biology
MaCS
Magdeburg

Collection

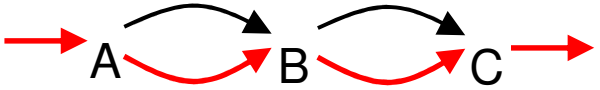
Elementary-Modes Sampling for Computing MCS in Large Networks



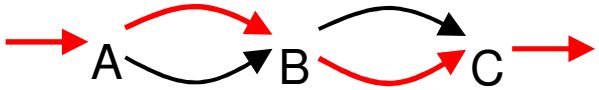
EM1



EM2



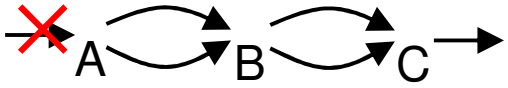
EM3

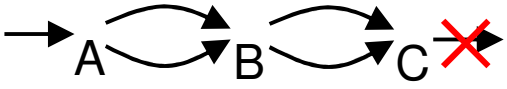


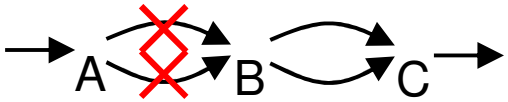
EM4

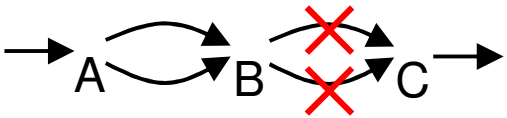
MCSs for intervention goal; block synthesis of C (via minimal hitting sets)!

4 MCSs

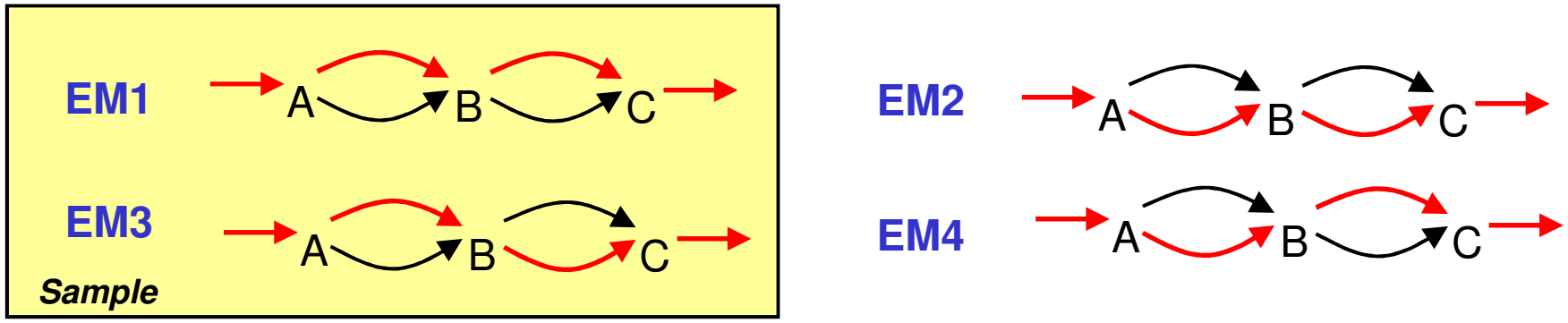




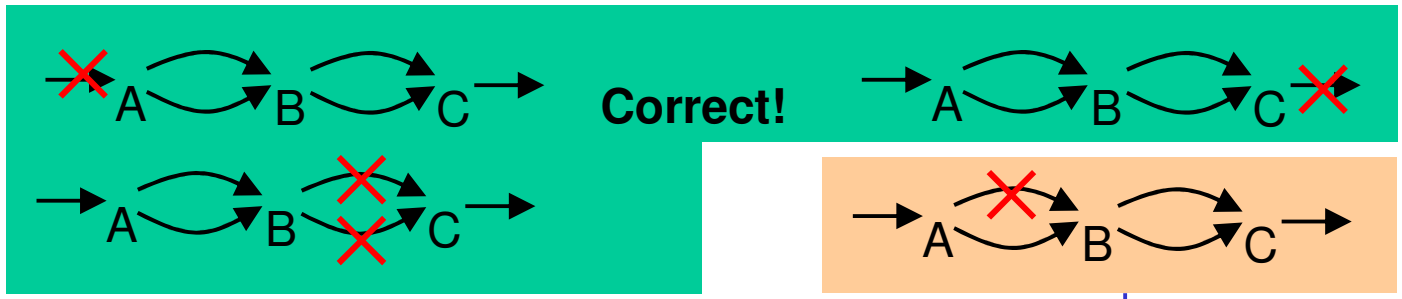




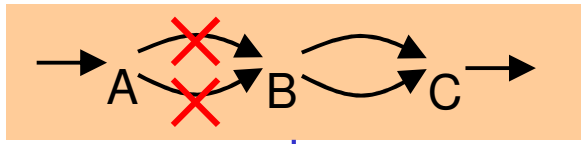
Elementary-Modes Sampling for Computing MCS in Large Networks



MCSs for Sample



Not a correct (full) MCSs!
… but can easily be excluded by FBA in postprocessing!



Missed MCSs!
Cannot be found by this sample!

Example for EM / MCS Sampling: Compute MCSs that Block Growth in *E. coli*

Via full enumeration of EMs

1) Compute all EMs (21592 EMs)

↓

2) Compute MCSs (via hitting set algorithm)

Via a sample of EMs (10%)

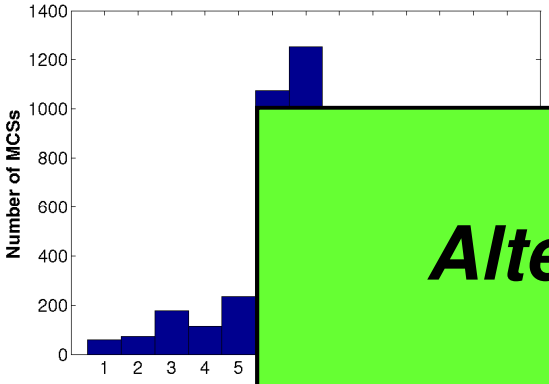
1) Sample EMs (2160 EMs = 10%)

↓ (e.g. EFMEnumerator; Kaleta et al.)

2) Compute MCSs (via hitting set algorithm)

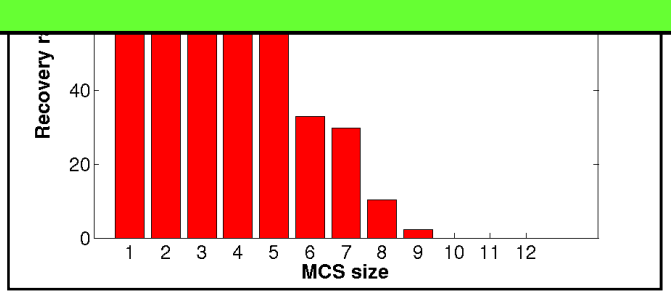
↓

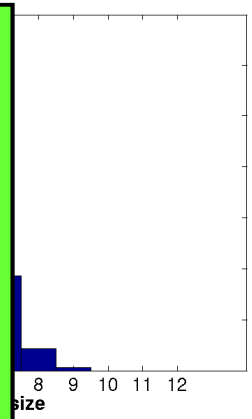
3) Postprocessing: discard false MCSs via FBA



Alternative approach:

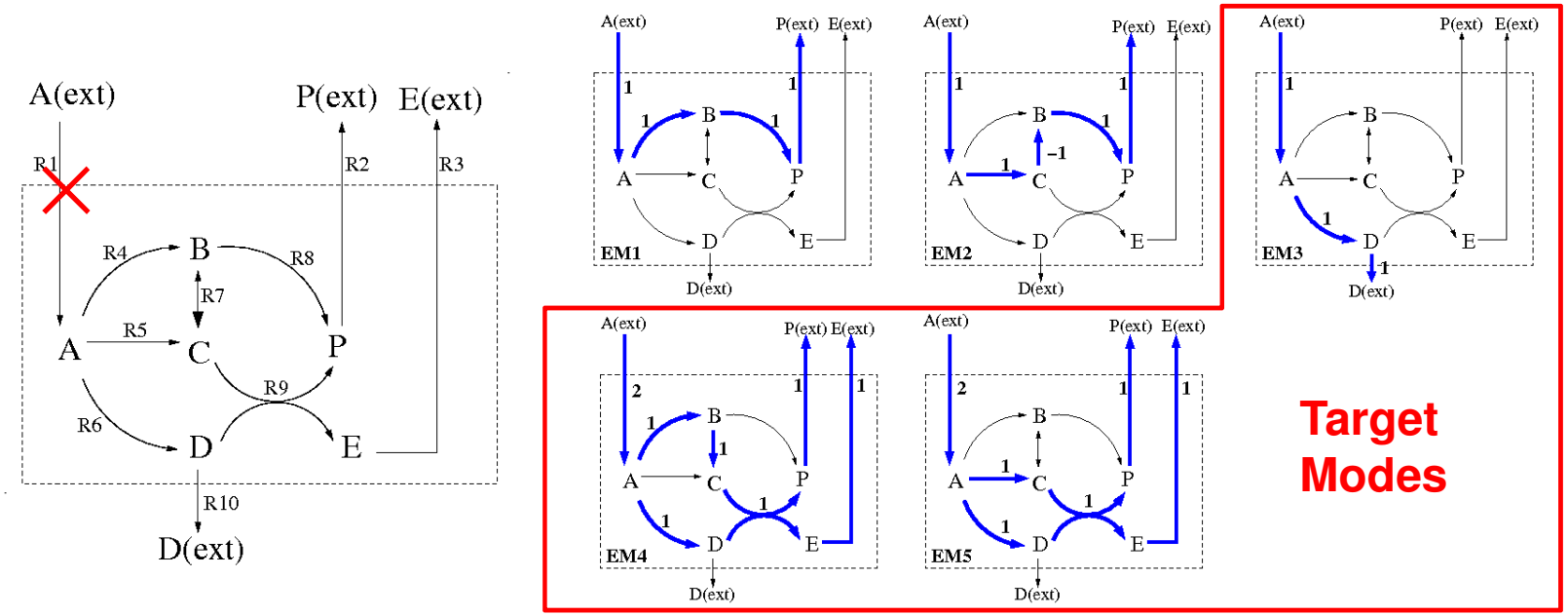
Sample EMs in dual network!
(= sample of MCSs in primal)





Constrained Minimal Cut Sets: Limitation of "Simple" Minimal Cut Sets

Intervention goal: repress non-optimal production routes for P



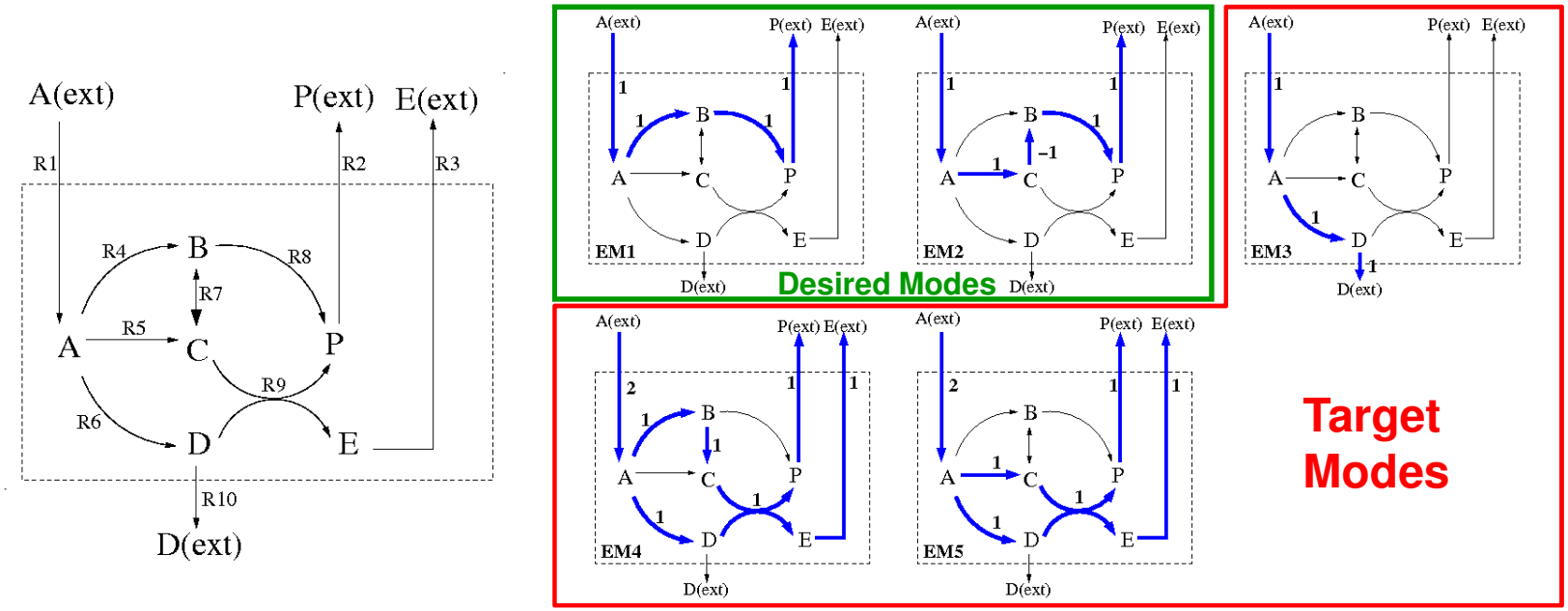
MCS1={R6}, MCS2={R1}, MCS3={R2, R10}, MCS4={R9,R10}, MCS5={R3,R10}, MCS6={R4,R5,R10}, MCS7={R5,R7,R10}

→ **Some MCSs (e.g. {R1}) induce side effects and disable functions we want to keep!**

→ **Constrained MCSs (cMCSs):** define set **T** of target modes *and* set **D** of desired modes and specify minimum number of modes in **D** to be preserved.

Generalization: Constrained MCSs

Intervention goal: repress non-optimal production routes for P



Target modes: T={EM3,EM4,EM5} (all non-optimal pathways for synthesis of P)

Desired modes: D={EM1,EM2} (all optimal pathways for synthesis of P}

Minimum number of desired modes to be saved: *n*=1

→ **MCS1={R6}, MCS2={R9,R10}, MCS3={R3,R10}, MCS4{R5,R7,R10}**

Minimum number of desired modes to be saved: *n*=2

→ **MCS1={R6}, MCS2={R9,R10}, MCS3={R3,R10}**

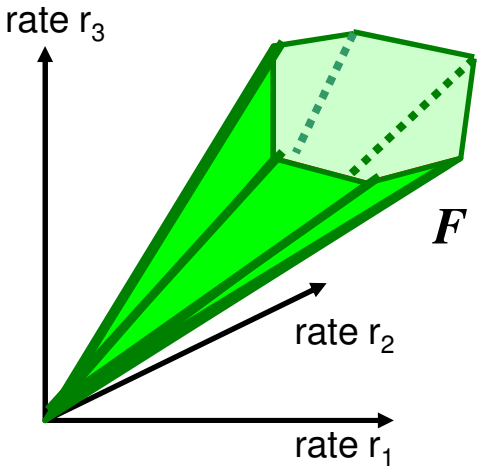
Computation of EMs / MCSs in Genome-Scale Networks ?!

- EM / MCSs calculation only possible in networks with moderate size

  → typically used in models of central metabolism
  → for many products: suitable interventions lie often in central metabolism (as also suggested/found by genome-scale methods)

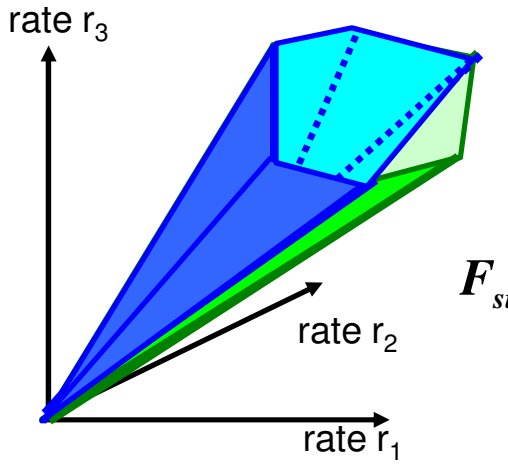
- possible solution for other applications: EM / MCS sampling …

Flux cone *F* spanned by set *E* of (all) EMs:



$$\boldsymbol{F} = \sum_{e_i \in \boldsymbol{E}} \alpha_i \mathbf{e}_i$$

Sub-cone spanned by (sampled) subset *S* of EMs:



$$\boldsymbol{S} \subset \boldsymbol{E}$$

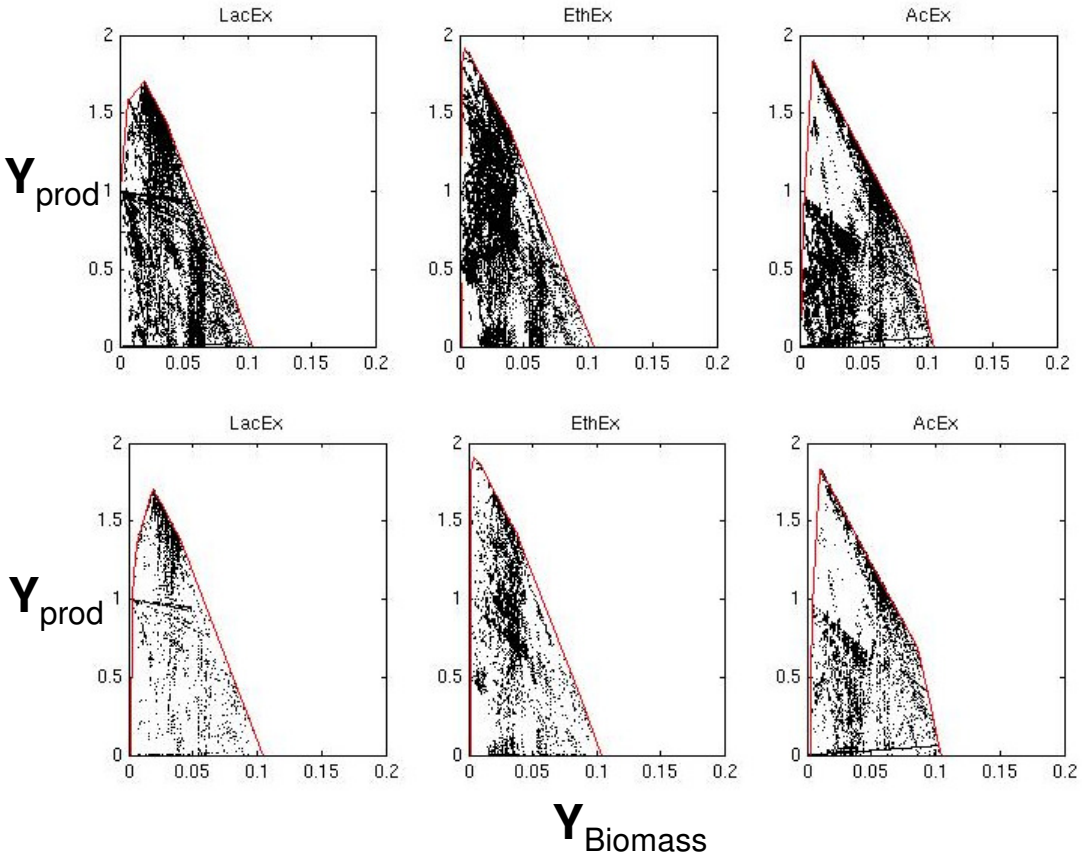
$$\boldsymbol{F}_{sub} = \sum_{e_i \in \boldsymbol{S}} \alpha_i \mathbf{e}_i$$

Elementary-Modes Sampling for MCSs Calculation

Example: 2D-yield space of EMs for central metabolism of E.coli
(with **EFMEvolver**; Kaleta et al.)

Full enumeration

Sampling (20%)

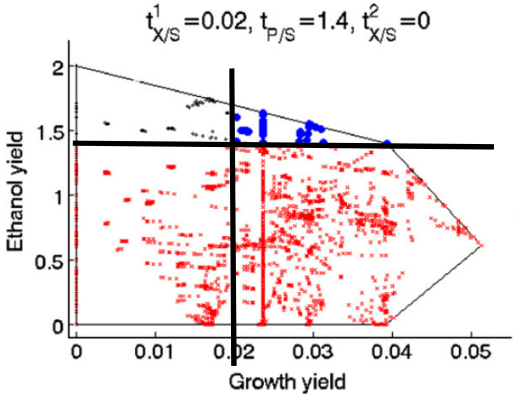


→ CASOP for genome-scale networks (**CASOP-GS**) underway (Kaleta group)

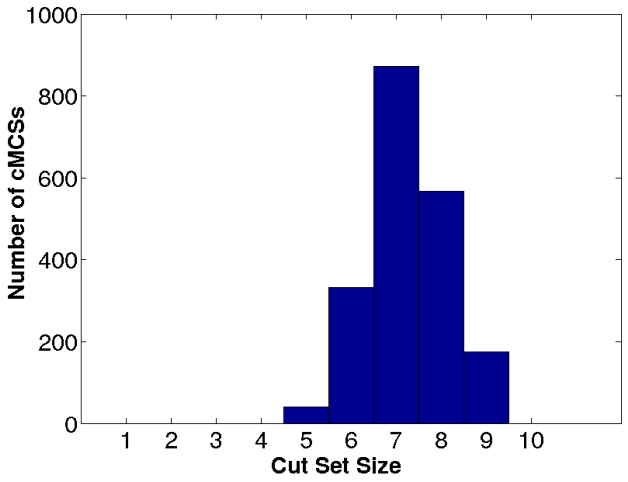
Inhomogenous Constraints: Coupled Biomass and Ethanol Synthesis in *E.coli*

Homogeneous system:

Nr = 0; $r_{Irrev} \geq 0$



At least 5 knockouts necessary



Inhomogeneous sys… nce)

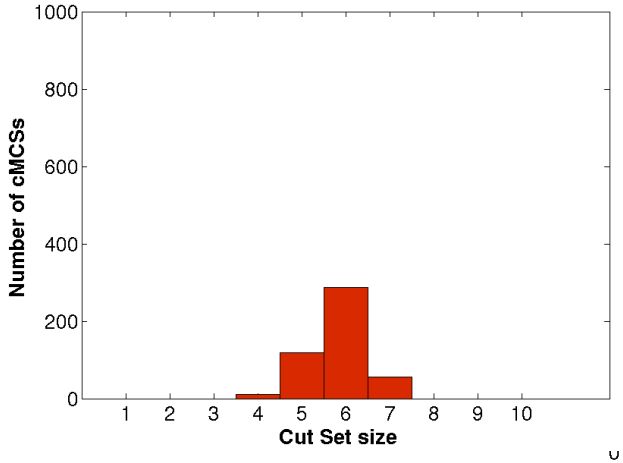
Shift!

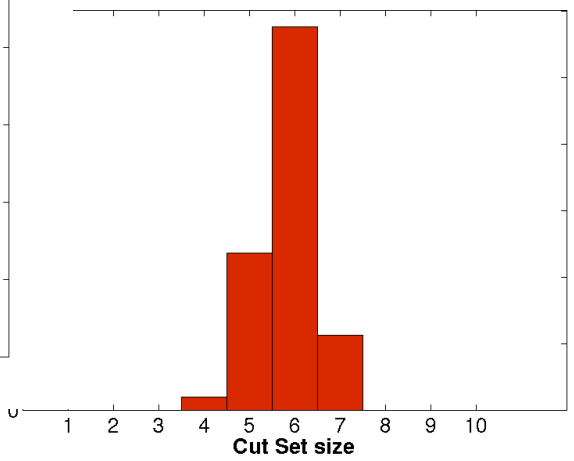
Nr = 0;

$r_{Irrev} \geq 0$

$0 \leq r_{Gluc\_up} \leq 10$ mmol/(g

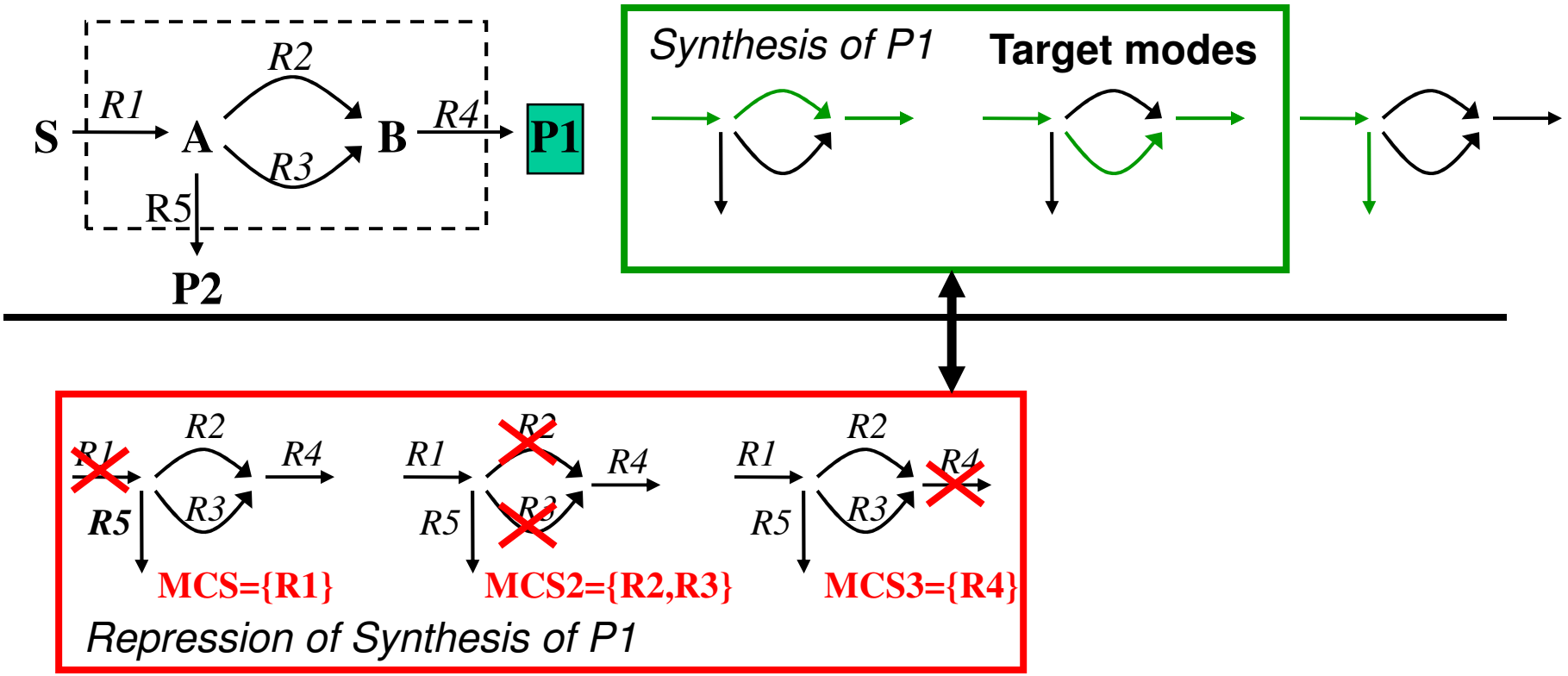
$r_{ATPmaintenance} \geq 8.4$ mmo





From Elementary Modes to Minimal Cut Sets

Elementary Modes: minimal functional units (pathways) that can operate in steady state



Minimal Cut Set (MCS): minimal set of cuts (knockouts) blocking a certain network function in steady state

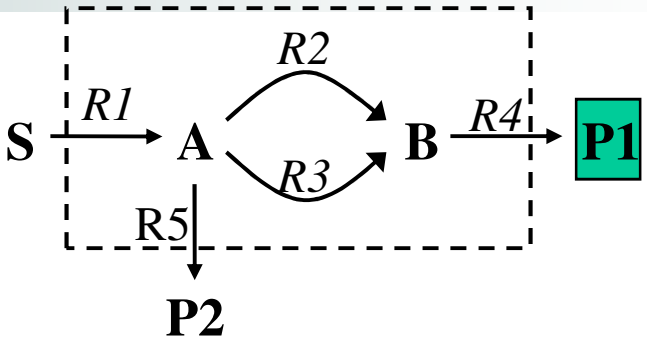
***MCSs** are the minimal hitting sets of the corresponding target EMs*

Klamt S: BioSystems 2006, 83:233-237.

Elementary Modes, Minimal Cut Sets and Minimal Hitting Sets

For now we identify an EM **e** (a vector) by its support **E** (which is a set):

E := supp(**e**) = {i: $e_i \neq 0$}.

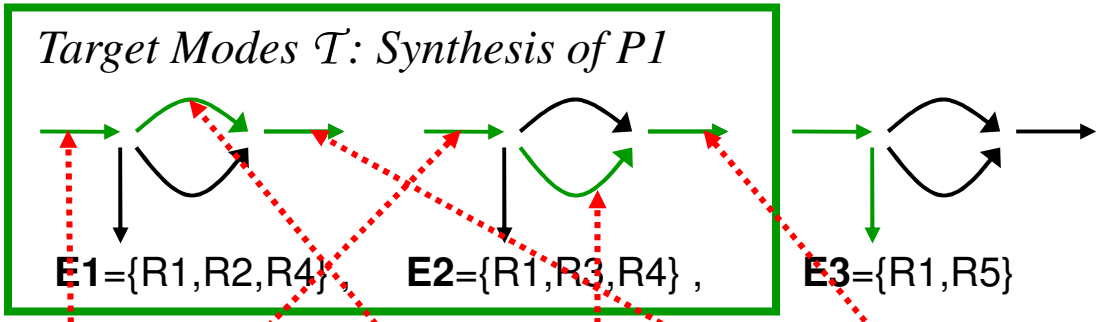


Set $\mathcal{T}$ of target modes:

$\mathcal{T}$ = {**E1**,**E2**}

= {{R1,R2,R4} , {R1,R3,R4}}

→ $\mathcal{T}$ is a family of (reaction) subsets (= undirected hypergraph)

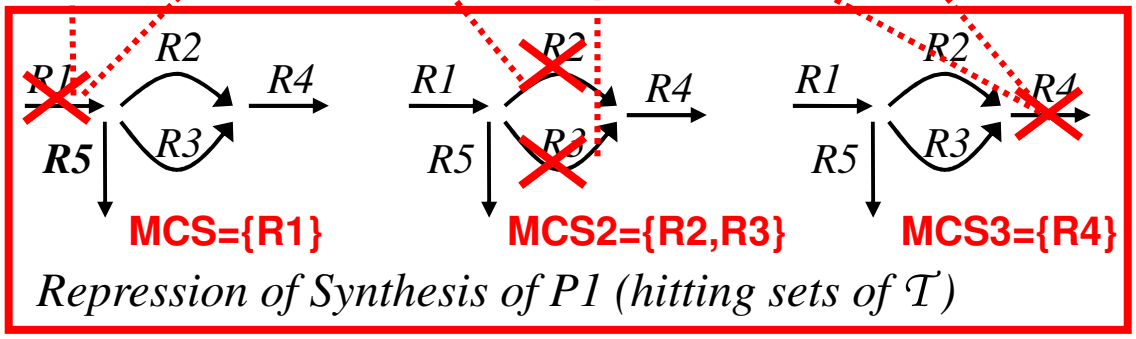


A minimal hitting set **C** of $\mathcal{T}$ hits all subsets (EMs) of $\mathcal{T}$, i.e.

$\mathbf{C} \cap \mathbf{E} \neq \emptyset$, $\forall\ \mathbf{E} \in \mathcal{T}$

and no subset of **C** does so.

A minimal hitting set hits all target EMs and knocking it out thus yields an MCS blocking all target EMs.



Klamt S: BioSystems 2006, 83:233-237.

Generalization: Constrained MCSs

Constrained Minimal Cut Set (cMCSs) problem:

- set $\mathcal{T}$ of target EMs (to be blocked)
- set $\mathcal{D}$ of desired EMs
- n: minimal number of desired EMs in $\mathcal{D}$ that must not be hit by the MCSs

→ A constrained MCS **C** fulfills: $\mathbf{C} \cap \mathbf{T} \neq \emptyset \; \forall \mathbf{T} \in \mathcal{T}; \quad |\{\mathbf{D} \in \mathcal{D} \mid \mathbf{C} \cap \mathbf{D} = \emptyset\}| \geq n$

Generalized Version:

→ set $\mathcal{T}$ of target EMs (to be blocked)
→ sets $\mathcal{D}_1 \ldots \mathcal{D}_k$ of desired EMs
→ $n_1 \ldots n_k$: minimal number of desired EMs in $\mathcal{D}_1 \ldots \mathcal{D}_k$ that must not be hit by the MCSs

→ A constrained MCS **C** fulfills: $\mathbf{C} \cap \mathbf{T} \neq \emptyset \; \forall \mathbf{T} \in \mathcal{T}; \quad |\{\mathbf{D} \in \mathcal{D}_i : \mathbf{C} \cap \mathbf{D} = \emptyset\}| \geq n_i \;\; \forall i$

Hädicke and Klamt (2011): Computing Complex Metabolic Intervention Stratgeies using Constrained Minimal Cut Sets. Metabolic Engineering, 13:204-213.

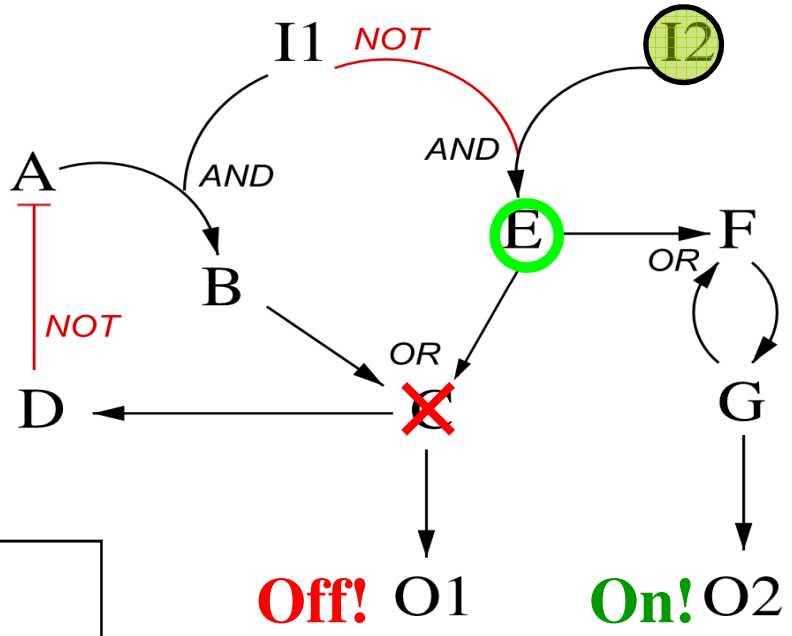
Counterpart of MCSs in Regulatory / Signaling Networks: Minimal Intervention Sets

Minimal intervention set (MIS): minimal combination of knockouts (off) and constitutive activations (on) that induce a desired response (defined by intervention goal)

Intervention Goal:
Inactivation of O1 (0)
Activation of O2 (1)
Side Constraint:
I2 is active (1)

MIS1={**I1=off, C=off**}
MIS2={**C=off**, **E=on**}

Example



Applications:
- target identification
- fragile points in the network
- diagnosis: what causes an observed effect

Samaga R, von Kamp A and Klamt S (2010) Computing combinatorial intervention strategies and failure modes in signaling networks. *Journal of Computational Biology* **17**:39-53.

CASOP:

Computational Approach for Strain Optimization Aiming at High Productivity

CASOP: Computational Approach for Strain Optimization Aiming at High Productivity

Motivation:

- existing approaches consider only knockouts, not overexpression candidates
- existing approaches for metabolic engineering aim at optimizing **product yield**:

$$Y^{P/S} \; [mol\ Product / mol\ Substrate]$$

- However, in practice, the (specific) **productivity** r_P is often the relevant performance parameter to be optimized:

$$\text{Specific Productivity}: \quad r_P = Y^{P/S} \cdot r_S \quad [mol\ Product\ /\ (gDW \cdot h)]$$

$$r_S: \text{ Substrate uptake rate } [mmol\ Substrate\ /\ (gDW \cdot h)]$$

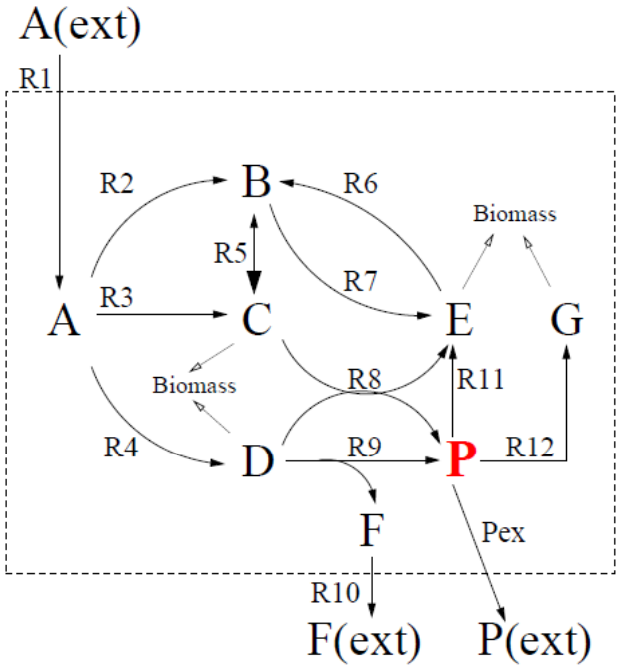
Optimizing productivity: search for optimal trade-off between high yield and high capacity (low-yield pathways may have a high capacity!)

→ CASOP: *stoichiometric* approach based on elementary-modes analysis suggesting knockout + overexpression candidates for optmizing productivity

CASOP: Computational Approach for Strain Optimization Aiming at High Productivity

Example network

P is the product of interest



R1 : Aext => A
R2 : A => B
R3 : A => C
R4 : A => D
R5 : B <=> C
R6 : E => B
R7 : B => E
R8 : C + D => E + P
R9 : 3 D => 2 F + P
R10: F => F(ext)
R11: P => E
R12: P => G
Pex: P => P(ext)
μ : 4 C + D + 4 E + G => Biomass

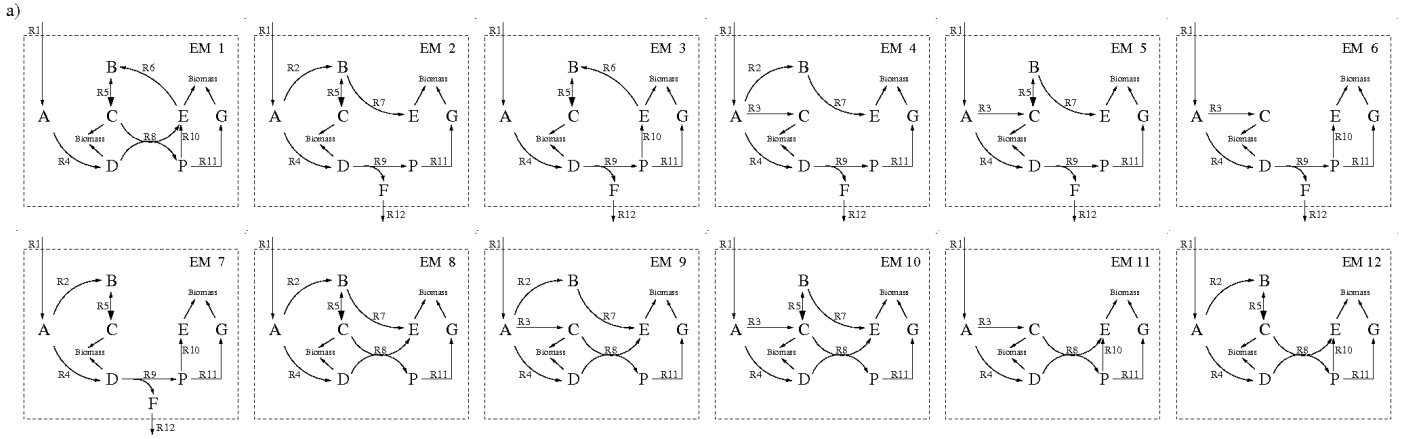
First step: add a pseudo-reaction *Rx* „consuming“ *product* P(ext) and *biomass:*

$$Rx : (1-\gamma)\, Biomass + \gamma\, \alpha_P P(ext) \rightarrow X \quad , \quad \gamma \in [0,1]$$ („proportion factor“)

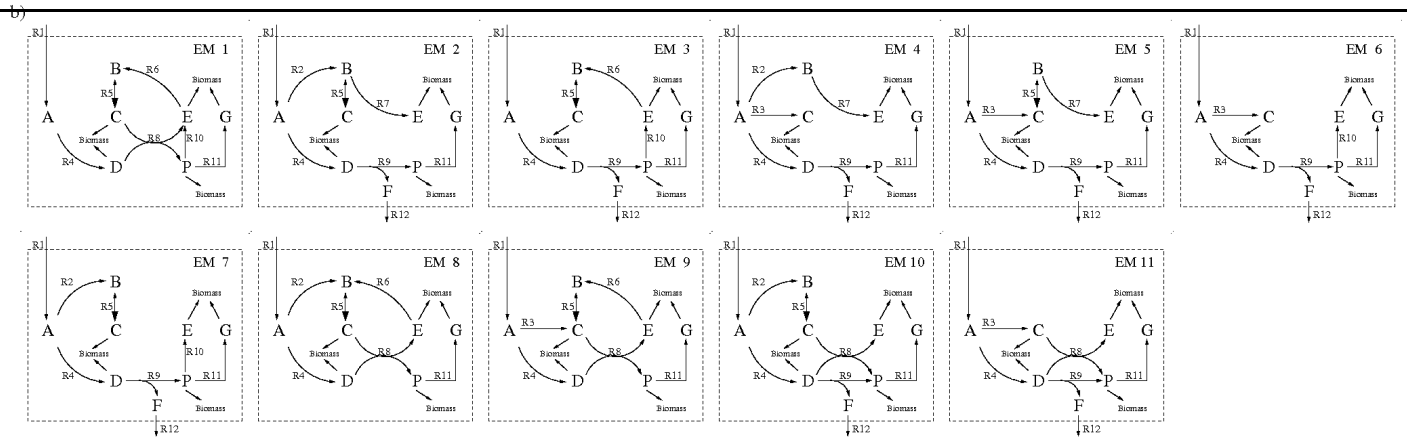
Increasing γ simulates the change in stoichiometric precursor demand from pure biomass production (γ = 0) via combined production of biomass and product (e.g. γ = 0.5) to exclusive production of P (γ = 1).

Compute Elementary Modes (EMs) for Different γ

γ = 0



γ = 0.9 (biomass synthesis still possible)



→ different spectra of EMs (pathways) for different **γ**
→ analyze these differences for each reaction …

Elementary-Modes Weights, Reaction Importances and Reaction Rankings

EM weight for EM *i* in scenario γ:

$$\nu_{i,\gamma} = \frac{(Y_i^{X/S})^k}{\sum_{i=1}^{n(\gamma)} (Y_i^{X/S})^k}$$

- relates the yield of an EM to sum of yields of all EMs
- adjustable parameter *k*: large *k* weight yield-optimal and EMs stronger; *k*=0→ each EM has same weight)
- sum of all weights is 1

Reaction importance measure for reaction *j* in a given scenario γ:

$$\omega_\gamma(r_j) = \sum_{i \in EM(r_j)} \nu_{i,\gamma}$$

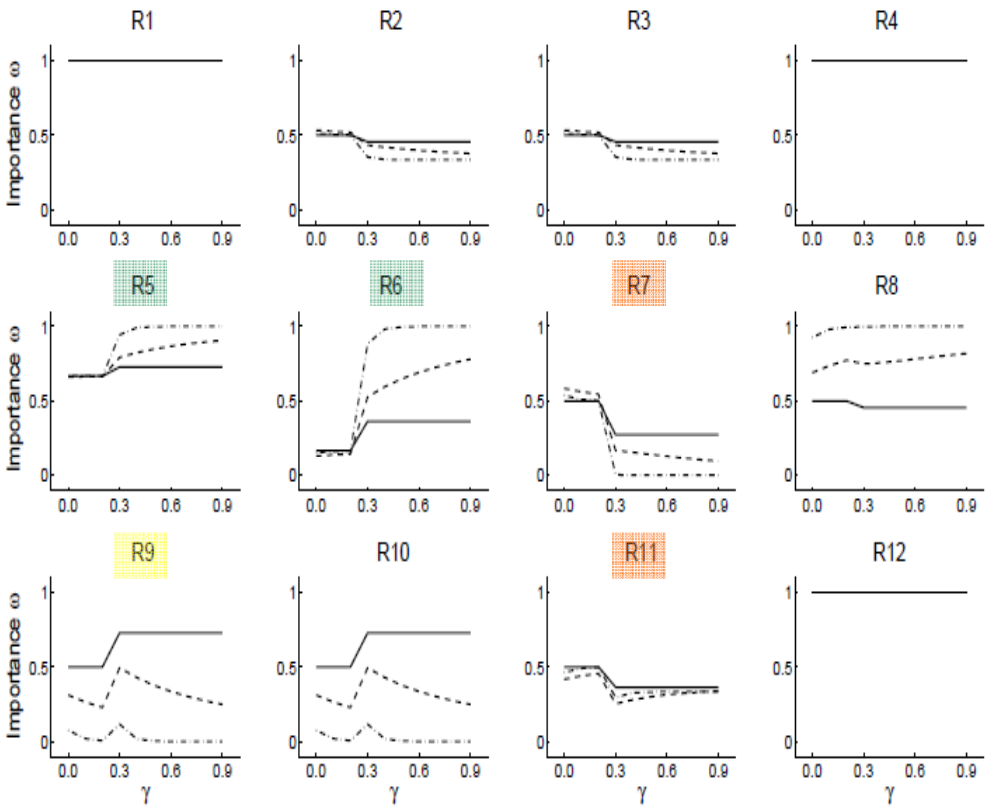
- sum of all EM weights in which reaction *j* partcipates
  → considers importance of *j* with respect to yield (weight) + flexibility/capacity
- in the range of [0,1]; e.g. 1 for essential reactions

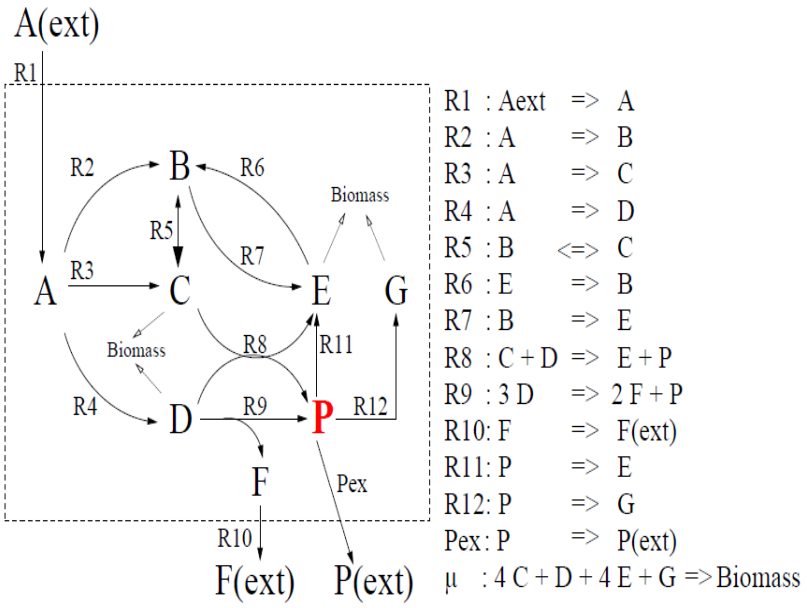
Rank reactions based on the importances:

Relative to wild type: $Z(r_j) = \omega_{0.9}(r_j) - \omega_0(r_j)$ $\quad Z \in [-1,1]$

high Z-score: overexpression candidate
low Z-score: knockout (KO) candidate

Example Network





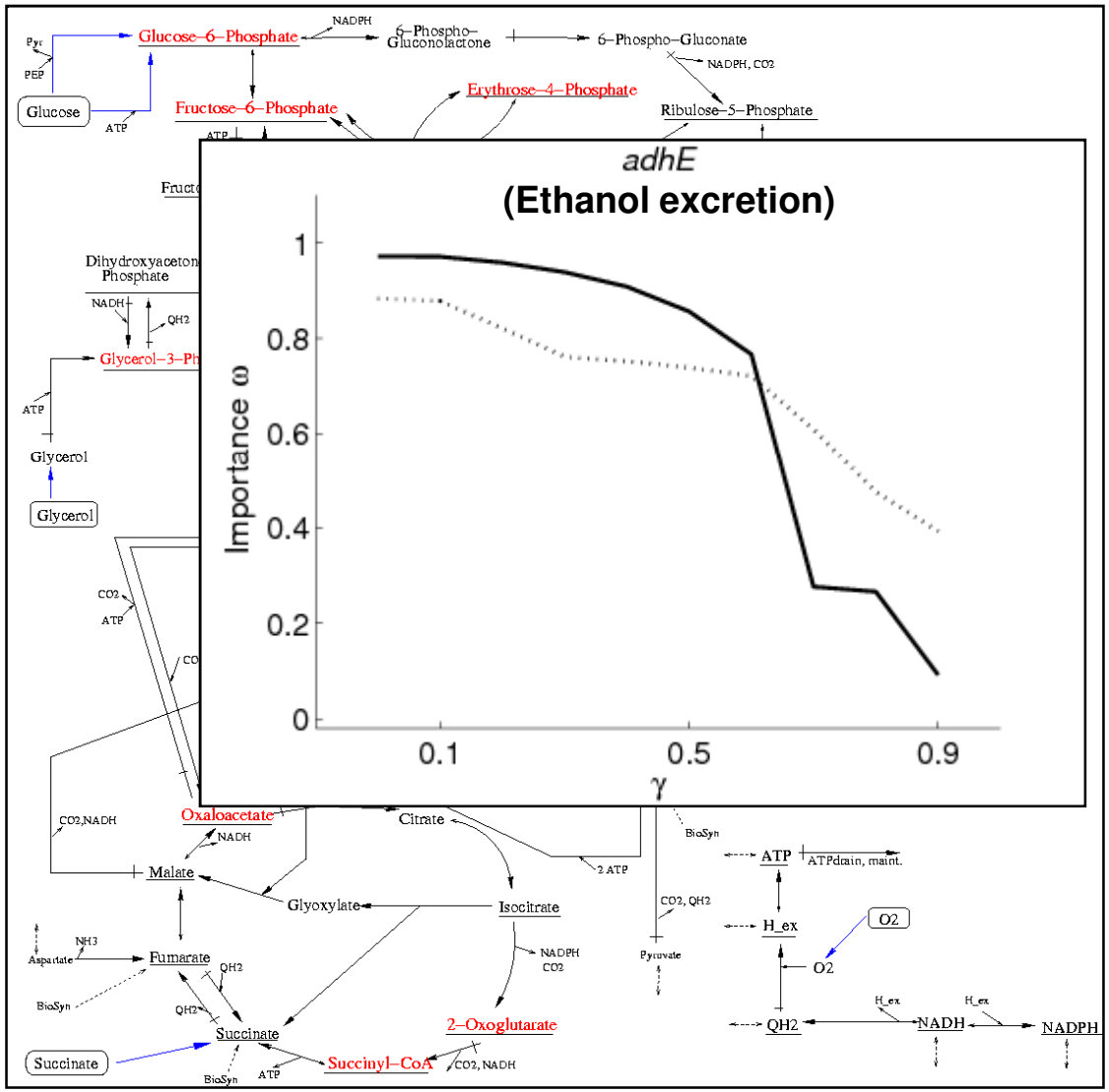
k=0

k=2

k=10

- R7, R11: for all k: ω deacreses when increasing γ → KO candidates
- R6: for all k: ω increases when increasing γ → overexpression candidates
- R9: overexpression candidate for low *k*, KO candidate for high *k*

Realistic Example: Succinate Production with *E.coli*



- model of central metabolism
- substrate: glucose; anaerobic conditions (fermentation)

Highest ranked KO candidates:

- Ethanol excretion
- Acetate excretion
- Lactate excretion
- Glucose PTS
- Pyruvate kinase
- Malic enzyme
- Succinate dehydrogenase
- Pyruvate formiate lyase

Highest ranked overexpression candidates

- G6P-kinase
- fumarate reductase
- PEP synthase
- Malate Dehydrogenase
- PEP carboxylase
- „CO2 uptake“
- glyoxylate shunt

Applied in practice

Identification of Potential Excess or Undersupply of Co-factors and Small Metabolites

- cofactors: NADH, ATP, CoA ... (also applicable to other (small) metabolites: CO2, ...)
- for a cofactor/metabolite of interest:

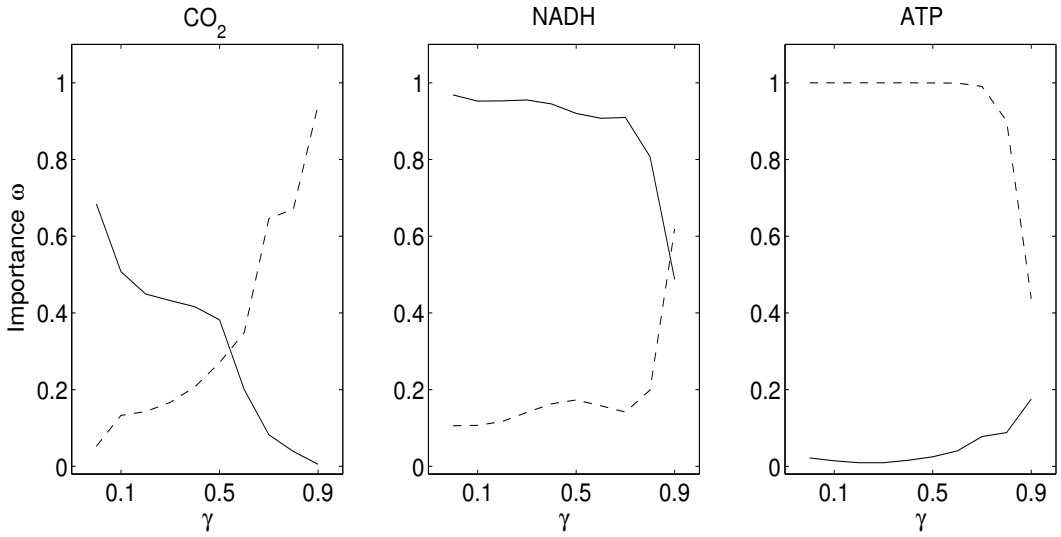
  insert (separately) artifical reactions consuming/supplying the cofactor/metabolite

  NADH → NAD (electron sink) / NAD → NADH (electron source)

  ATP → ADP + Pi (energy sink) / ADP + Pi → ATP (energy source)

- apply CASOP and check the importances of the artificial reactions!

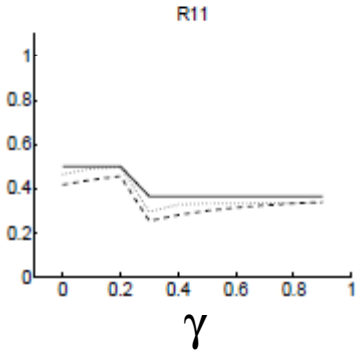
Anaerobic succinate production in *E.coli* (*k*=2)



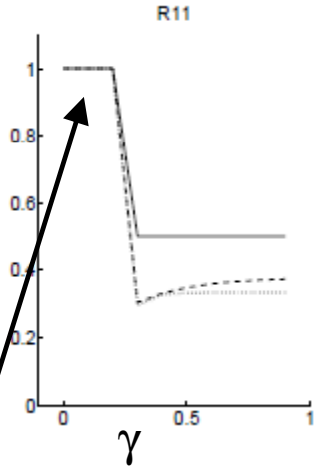
Multiple Interventions

- Select intervention based on reaction ranking (e.g. Knockout)
- Select all EMs where the reaction knocked-out is not used and recalculate elementary-modes weights, reaction importances + ranking
- Select next intervention

k=0
k=2
k=10



Knockout of R7



R11 is a promising KO-candidate in the next iteration: Knocking out of R11 will lead to coupling of biomass synthesis and product formation

How to Get the Dual Network

- define intervention goal (target reactions / target flux vectors) by $\mathbf{t}^T\mathbf{r} > 0$

- wanted: MCSs *C* such that for all **r** satisfying

$$(1)\ \mathbf{Nr} = \mathbf{0} \quad (2)\ \mathbf{r}_{Irrev} \geq \mathbf{0} \quad (3)\ r_i = 0,\ i \in C$$

it holds that $\mathbf{t}^T\mathbf{r} = 0$

Define Inconsistent System **S**

$$\mathbf{Nr} = \mathbf{0}$$
$$\mathbf{r}_{Irrev} \geq \mathbf{0}$$
$$\mathbf{t}^T\mathbf{r} \geq 1$$
$$\mathbf{Ir} = \mathbf{0}$$

Farkas Lemma
Theory of IISs
(Ryan et al.)

Consider System **D**

$$(\mathbf{I}\ \ -\mathbf{t}\ \ -\mathbf{I}_{irrev}\ \ \mathbf{N}^T)\begin{pmatrix}\mathbf{v}\\ w\\ \mathbf{z}\\ \mathbf{u}\end{pmatrix} = \mathbf{0};\quad \mathbf{z} \geq \mathbf{0}; w \geq 0$$

$\mathbf{N}_{dual}$ $\mathbf{r}_{dual}$

We are interested in *irreducible inconsistent subsystems* (IISs) of **S** keeping the system inconsistent with a minimal subset of equations from **Ir=0** (i.e. with a minimal (cut) set of reactions set to zero).

The IISs of **S** (corresponding to the MCSs) are the elementary modes of system D which are minimal in **v** and have support in *w*.

Ballerstein, von Kamp, Klamt, Haus: Minimal cut sets in a metabolic network are elementary modes in a dual network. Submitted.