Dynamic optimization of metabolic networks coupled with gene expression

Alexander Bockmayr

Freie Universität Berlin

Joint Work with S. Waldherr and D. Oyarzún

Research Center MATHEON

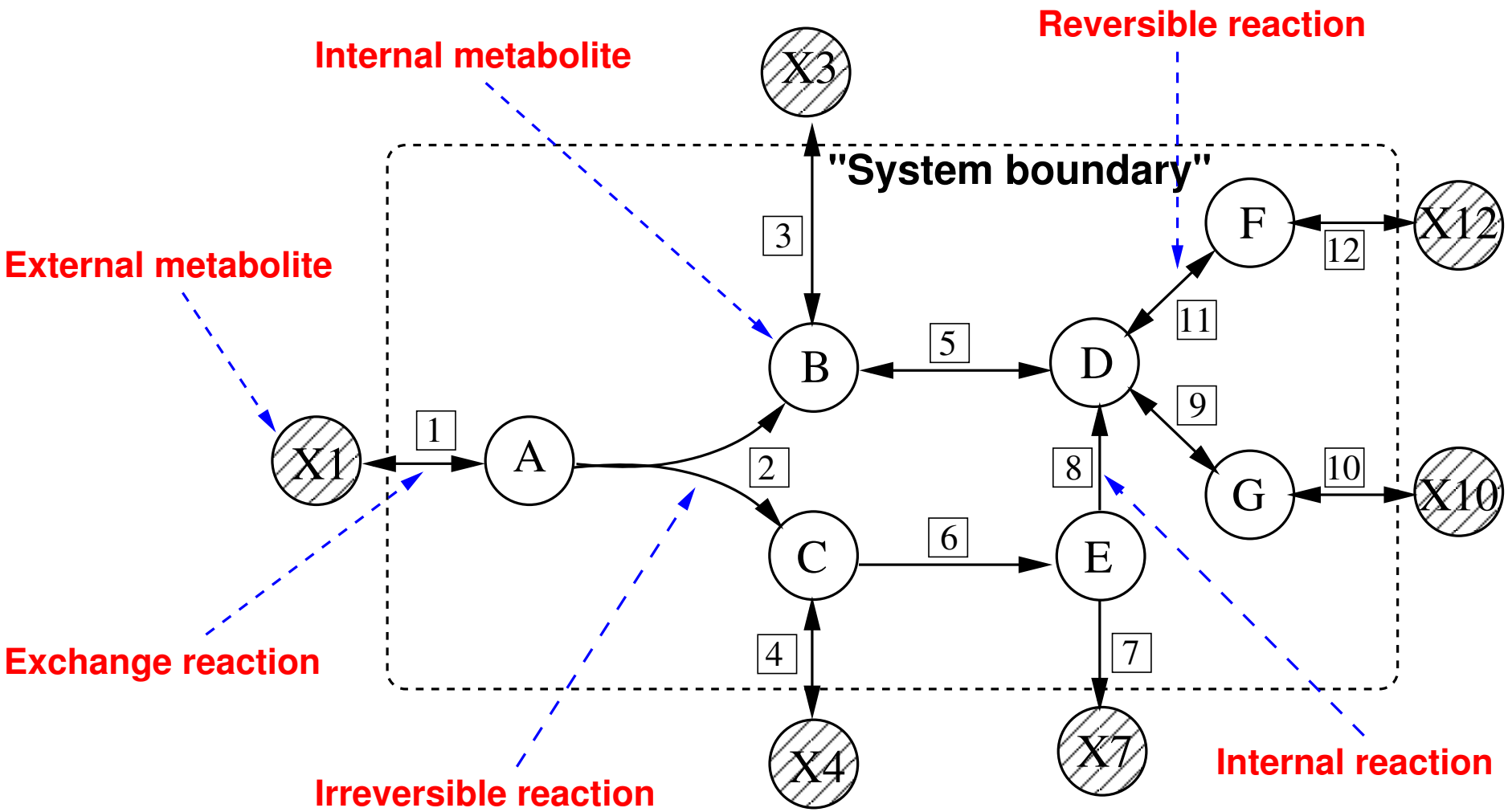
Mathematics for key technologies

ECMath

Einstein Center

for Mathematics Berlin

Toulouse, 14 October 2016

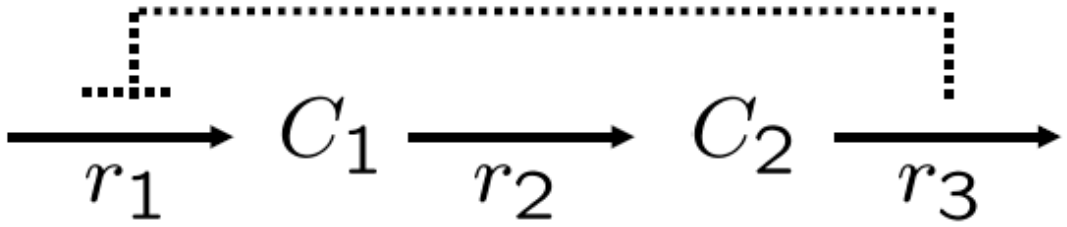

Internal metabolite
Reversible reaction
X3
"System boundary"
3
F
12
X12
External metabolite
11
B
5
D
9
1
X1
A
2
8
G
10
X10
C
6
E
4
7
Exchange reaction
X4
X7
Irreversible reaction
Internal reaction

- Stoichiometric matrix $S \in^{m \times n}$
  - Rows $\rightsquigarrow$ internal metabolites $i = 1, \ldots, m$
  - Columns $\rightsquigarrow$ internal and exchange reactions $j = 1, \ldots, n$
  - S_{ij}: stoichiometric coefficient of reactant i in reaction j
- Set of irreversible reactions Irr
- Metabolic model $\mathcal{M} = (S, Irr)$

- Metabolites i and reactions j
- $C_i(t)$: metabolite concentrations at time t
- $v_j(t) = v_j(C(t), k)$: reaction rates, depending on kinetic law and kinetic parameters k
- S_{ij}: stoichiometric coefficient

$$\frac{dC_i}{dt} = \sum_{j=1}^{n} S_{ij} v_j \quad \text{or} \quad \frac{dC}{dt} = S \cdot v(C, k)$$

- System of ordinary differential equations (ODEs)



$$\begin{pmatrix} dC_1/dt \\ dC_2/dt \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} v_1(C,k) \\ v_2(C,k) \\ v_3(C,k) \end{pmatrix}$$

$$\begin{aligned} v_1(C,k) &= v_{m1}/(1+(C_2/k_i)^p) \\ v_2(C,k) &= v_{m2} \cdot C_1/(k_1+C_1) \\ v_3(C,k) &= v_{m3} \cdot C_2/(k_2+C_2) \end{aligned}$$

Which kinetic laws?

Which kinetic parameters?

- Steady-state assumption
  Assume metabolite concentrations C_i and reaction rates v_j are constant (over some time interval)
  $\rightsquigarrow$ steady-state flux vector $v \in \mathbb{R}^n$
- Stoichiometric constraints (mass balance):

$$\sum_{j=1}^{n} S_{ij} v_j = 0, \text{ for all } i = 1, \ldots, m$$

- Thermodynamic irreversibility constraints:

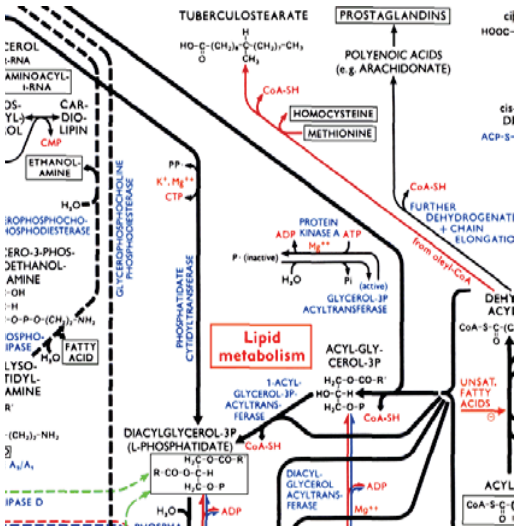
$$v_j \geq 0, \text{ if } j \text{ is irreversible}$$

$\rightsquigarrow$ system of linear equations and inequalities in $\mathbb{R}^n$

Set of all possible steady-state flux distributions

$$C = \{v \in^n | \; Sv = 0, \; v_i \geq 0, \; i \in Irr\}$$

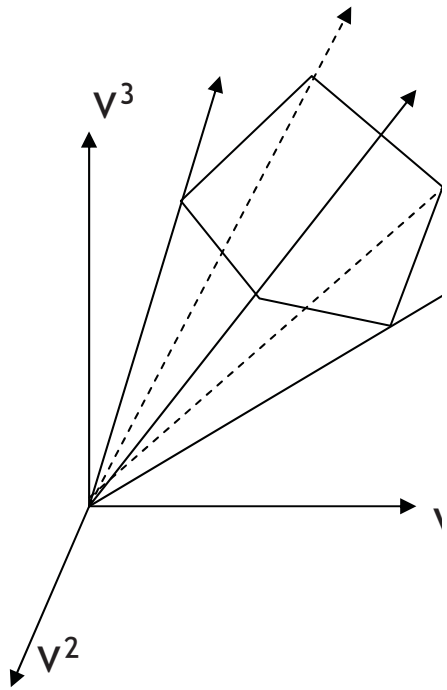
$\rightsquigarrow$ polyhedral cone



Reactions

$$\begin{pmatrix} 1 & 0 & \dots & -1 \\ 0 & 1 & \dots & -2 \\ \dots & \dots & \dots & \dots \\ 2 & 0 & \dots & 1 \end{pmatrix}$$

Metabolites



- Assume cellular behavior is determined by a certain biological objective.
- Determine a corresponding "best" flux distribution.
- Use mathematical optimization to predict phenotype.

- Simplest case: Linear programming (LP)

$$\max\{c^T x \mid Ax \leq b, x \in^n\}$$

- Flux balance problem (FBA) Varma/Palsson 94

$$\max\{c^T v \mid Sv = 0,\ l \leq v \leq u\}$$

Orth/Thiele/Palsson 10

- *E. coli* metabolism
- Genome-scale reconstruction (*i*JO1366)
- 1336 metabolites, 2251 reactions
- Objective function: biomass
- Glucose and oxygen uptake reactions
- Aerobic and anaerobic growth
- Software: e.g. COBRA Toolbox 2.0

Waldherr/Oyarzún/Bockmayr 15

$$Y \underset{\uparrow P}{\overset{V_y}{\longleftrightarrow}} X, \quad X \underset{\uparrow P}{\overset{V_x}{\longrightarrow}} \tilde{X}, \quad X \underset{\uparrow P}{\overset{V_p}{\longrightarrow}} P$$

- Molecular species
  - Extracellular nutrients and/or waste Y
  - Intracellular metabolites X
  - Macromolecules/enzymes P
- Reaction fluxes
  - Exchange reactions V_y
  - Internal metabolic reactions V_x
  - Biomass reactions V_p
- Stoichiometric matrices S^i_j (species i, reactions j)

- Mass balance

$$\dot{Y} = -S_y^y V_y$$
$$\dot{P} = S_p^p V_p$$
$$\dot{X} = S_y^x V_y + S_x^x V_x - S_p^x V_p$$

- Macromolecule production is slow: small ε
- Macromolecules are made from many components: large α

$$\dot{Y} = -S_y^y V_y$$
$$\dot{P} = \varepsilon S_p^p V_p$$
$$\dot{X} = S_y^x V_y + S_x^x V_x - \varepsilon\alpha S_p^x V_p$$

- Time-scale separation (using Tikhonov's theorem)

$$\dot{Y} = -S_y^y V_y,$$
$$\dot{P} = \varepsilon S_p^p V_p,$$
$$0 = S_y^x V_y + S_x^x V_x - \alpha \varepsilon S_p^x V_p.$$

- Exchange reactions and biomass production coupled via quasi steady-state constraint for intracellular metabolism.
- Model reduction

	Metabolism	+ Enzyme production
Steady state	Flux Balance Analysis (FBA) Varma/Palsson 94	Resource Balance Analysis (RBA) Goelzer et al. 11
Dynamic	Dynamic FBA (dFBA) Mahadevan et al. 02	Dynamic Enzyme Cost Analysis (deFBA) – This Talk –

- Goal: Determine metabolic fluxes maximizing a cellular objective.
- Network model: Metabolism & steady state

$$0 = S_y v_y + S_x v_x - S_{bm} v_{bm}$$

- Constraints: Bounds on fluxes

$$v_{i,min} \leq v_i \leq v_{i,max}$$

- Optimisation: Linear programming (LP)

$$\max_v v_{bm}$$

- Goal: Determine metabolic fluxes maximizing a cellular objective over time (based on biomass concentration $P(t)$).
- Network model: Metabolism & dynamic

$$\dot{y}(t) = -S_y^y v_y(t)\, P(t)$$
$$\dot{x}(t) = S_y^x v_y(t)\, P(t) + S_x^x v_x(t)\, P(t) - S_{bm}^x v_{bm}(t)\, P(t)$$
$$\dot{P}(t) = v_{bm}(t)\, P(t)$$

- Constraints: Bounds on fluxes & flux changes

$$v_{i,min}(y,P) \leq v_i(t) \leq v_{i,max}(y,P), \qquad |\dot{v}_i(t)| \leq \dot{v}_{i,max}$$

- Optimisation: Non-linear dynamic

$$\max_{v(t)} P(t_{end})$$

- Goal: Determine cell composition (protein concentrations p) and metabolic fluxes v maximizing the growth rate μ.
- Network model: Metabolism+enzyme production & steady state

$$0 = S_y^x v_y + S_x^x v_x - \alpha\epsilon S_p^x v_p$$
$$0 = \varepsilon S_p^p v_p - \mu p$$

- Constraints: Enzyme capacity & cellular composition

$$\sum_{j\in\mathcal{V}_i} |v_j(t)/k_j| \leq p_i, \qquad \sum_i c_i p_i \leq 1$$

- Optimization: Iteratively solving LPs

$$\max_{v,p} \mu$$

- Goal: Determine the dynamic cell composition and metabolic fluxes to maximize a cellular objective over a time interval
- Network model: Metabolism+enzyme production & dynamic

$$\dot{Y} = -S_y^y V_y, \qquad \dot{P} = \varepsilon S_p^p V_p,$$
$$0 = S_y^x V_y + S_x^x V_x - \alpha \varepsilon S_p^x V_p$$

- Constraints: Enzyme capacity & cellular composition

$$\sum_{j \in \mathcal{V}_i} |V_j(t)/k_j| \leq P_i(t), \qquad \sum_i c_i P_i(t) \leq 1$$

- Optimization: Linear dynamic

$$\max_{V,Y,P} \int_0^{t_{end}} c^{\mathrm{T}} P(t) dt$$

Dynamic optimization problem [with $z = (Y, P)$]

$$\max_{\mathcal{V}(\mathcal{Z}, z_0)} \int_0^{t_{end}} \Phi(z(t), v(t))dt \; + \; \Psi(z(t_{end}))$$

s.t. $\dot{Y} = -S_y^y V_y, \quad \dot{P} = \varepsilon S_p^p V_p,$

$S_y^x V_y + S_x^x V_x - \alpha \varepsilon S_p^x V_p = 0,$ Quasi steady-state

$z(0) = z_0, z(t) \geq 0,$

$v_{min} \leq v(t) \leq v_{max},$

$H_C v(t) \leq H_E P(t),$ Enzyme capacity

$H_B P(t) \leq h_B$ Cellular composition

- Components: Nutrient Y, metabolite X, generic enzyme P
- Reactions

$$V_y : Y \to X \ \text{(uptake)}, \quad V_p : \alpha X \to P \ \text{(biomass)}$$

- Enzymatic constraint

$$\frac{V_y}{k_y} + \frac{\varepsilon V_p}{k_p} \leq P$$

- Quasi steady-state approximation

$$\dot{Y} = -V_y, \quad \dot{P} = \varepsilon V_p, \qquad V_y = \alpha \varepsilon V_p$$

- Maximization of terminal biomass

$$J_1 = P(t_{end})$$

- Maximization of discounted biomass integral

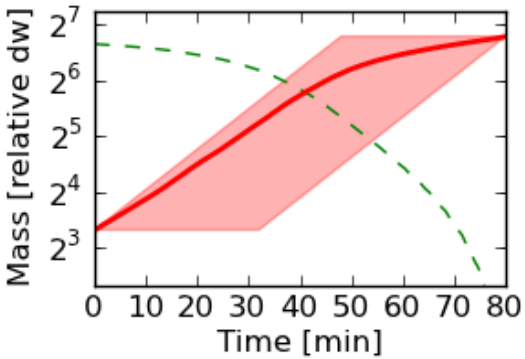
$$J_2 = \int_0^{t_{end}} P(\tau) e^{-\mu\tau} d\tau$$

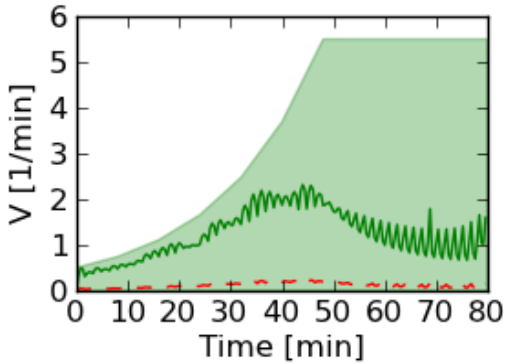
- Minimization of time to consume nutrients

$$J_3 = -\int_0^{t_{end}} d\tau = -t_{end}$$

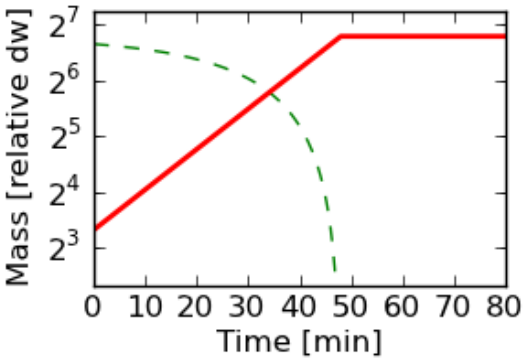
with $Y(t_{end}) = 0$.

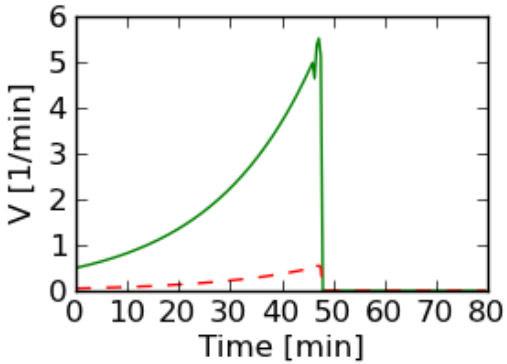
Terminal biomass



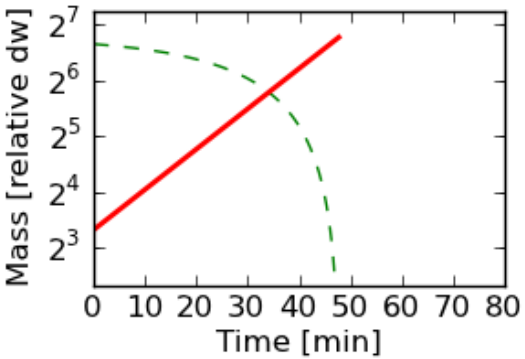


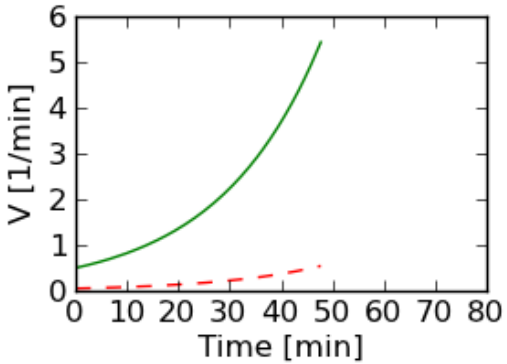
Discounted biomass





Minimal Time





Analytical proof of existence and uniqueness for J_2 and J_3

$\rightsquigarrow$ mathematical optimum biologically meaningful?

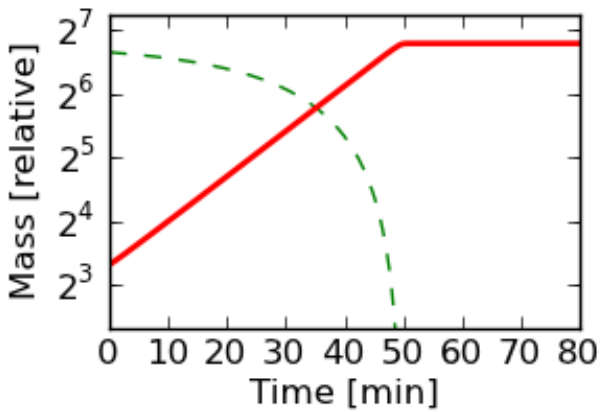
- Assume Michaelis-Menten kinetics

$$V_y = \frac{V_{m,y} P Y}{K_y \vartheta_e + Y}, \qquad \varepsilon V_p = \frac{V_{m,p} P X}{K_p + X}.$$

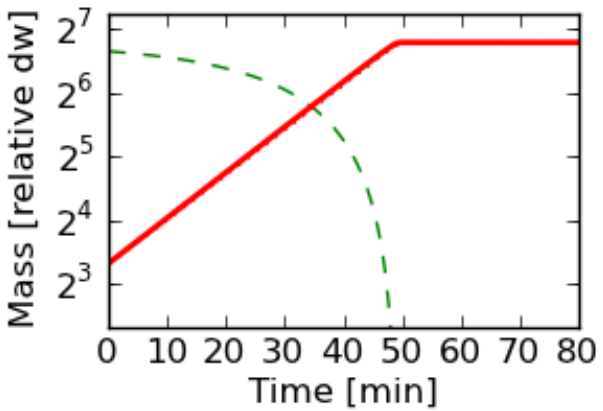
- Choose consistent parameters

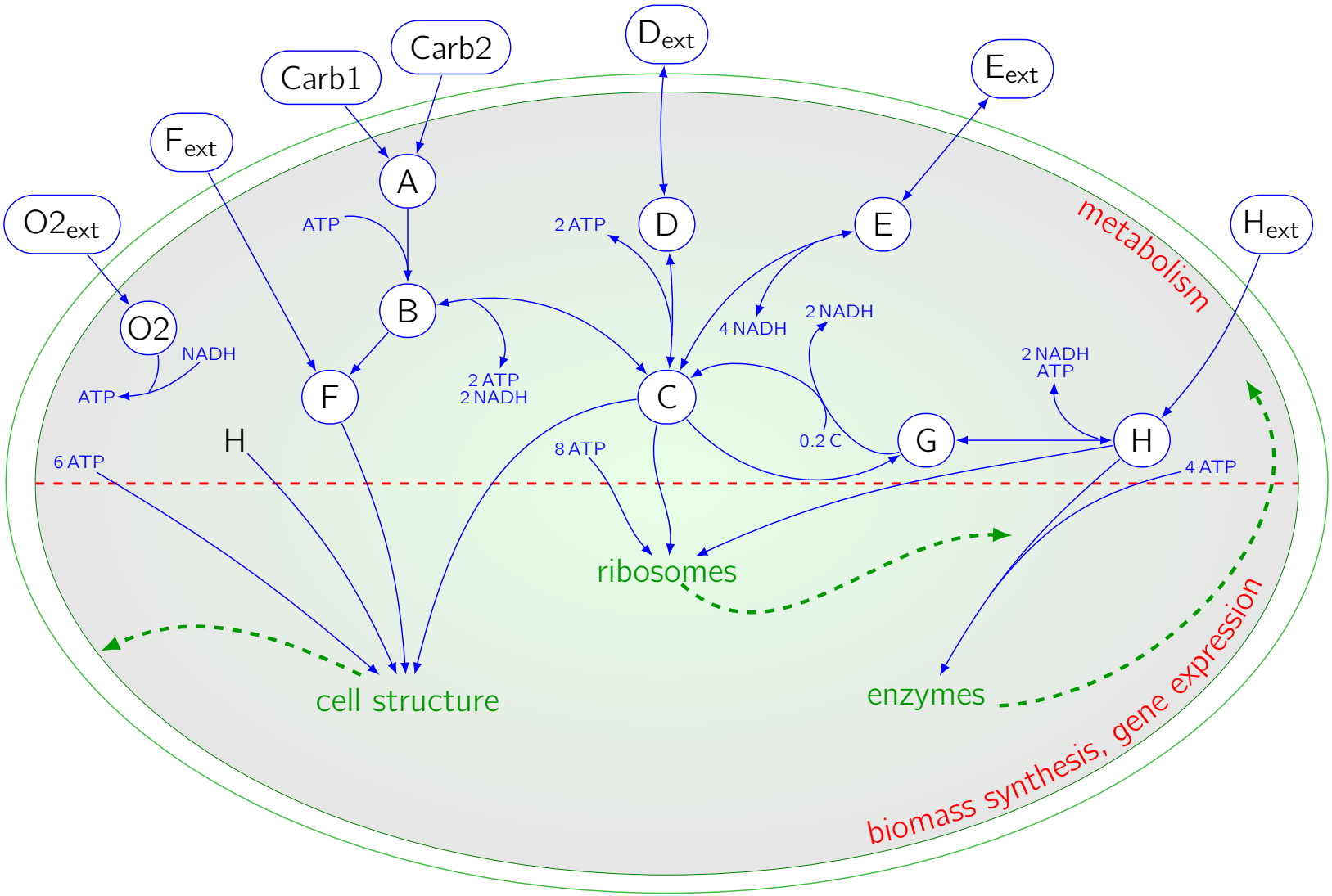
$$V_{m,y} = \left(\frac{1}{k_y} + \frac{1}{\alpha k_p}\right)^{-1}, \qquad V_{m,p} = \left(\frac{\alpha}{k_y} + \frac{1}{k_p}\right)^{-1}.$$

Simulation
(assuming kinetics)



Dynamic optimization
(no kinetics assumed)





Covert et al. 2001 [for the metabolic part]

Reaction	Enz	k_{cat}
Exchange reactions		
$Carb1 \rightarrow A$	T_{C1}	$3000\ \frac{1}{\mathrm{min}}$
$Carb2 \rightarrow A$	T_{C2}	$2000\ \frac{1}{\mathrm{min}}$
$F_{ext} \rightarrow F$	T_F	$3000\ \frac{1}{\mathrm{min}}$
$O2_{ext} \rightarrow O2$	S	$1000\ \frac{1}{\mathrm{min}}$
$D \leftrightarrow D_{ext}$	S	$1000\ \frac{1}{\mathrm{min}}$
$E \leftrightarrow E_{ext}$	S	$1000\ \frac{1}{\mathrm{min}}$
$H_{ext} \rightarrow A$	T_H	$3000\ \frac{1}{\mathrm{min}}$
Metabolic reactions		
$A + ATP \rightarrow B$	E_B	$1800\ \frac{1}{\mathrm{min}}$
$B \rightarrow C + 2\,ATP + 2\,NADH$	E_C	$1800\ \frac{1}{\mathrm{min}}$
$B \rightarrow F$	E_F	$1800\ \frac{1}{\mathrm{min}}$
$C \rightarrow G$	E_G	$1800\ \frac{1}{\mathrm{min}}$
$G \rightarrow 0.8\,C + 2\,NADH$	E_N	$1800\ \frac{1}{\mathrm{min}}$
$C \leftrightarrow 2\,ATP + 3\,D$	E_D	$1800\ \frac{1}{\mathrm{min}}$
$C + 4\,NADH \leftrightarrow 3\,E$	E_E	$1800\ \frac{1}{\mathrm{min}}$
$G + ATP + 2\,NADH \leftrightarrow H$	E_H	$1800\ \frac{1}{\mathrm{min}}$
$NADH + O \rightarrow ATP$	E_T	$1800\ \frac{1}{\mathrm{min}}$

Reaction	Enz	k_{cat}
Biomass reactions		
$400\,H + 1600\,ATP \rightarrow T_{C1}$	R	$2.5\ \frac{1}{\mathrm{min}}$
$1500\,H + 6000\,ATP \rightarrow T_{C2}$	R	$0.67\ \frac{1}{\mathrm{min}}$
$400\,H + 1600\,ATP \rightarrow T_F$	R	$2.5\ \frac{1}{\mathrm{min}}$
$400\,H + 1600\,ATP \rightarrow T_O$	R	$2.5\ \frac{1}{\mathrm{min}}$
$400\,H + 1600\,ATP \rightarrow T_D$	R	$2.5\ \frac{1}{\mathrm{min}}$
$400\,H + 1600\,ATP \rightarrow T_E$	R	$2.5\ \frac{1}{\mathrm{min}}$
$400\,H + 1600\,ATP \rightarrow T_H$	R	$2.5\ \frac{1}{\mathrm{min}}$
$500\,H + 2000\,ATP \rightarrow E_B$	R	$2\ \frac{1}{\mathrm{min}}$
$500\,H + 2000\,ATP \rightarrow E_C$	R	$2\ \frac{1}{\mathrm{min}}$
$1000\,H + 4000\,ATP \rightarrow E_D$	R	$1\ \frac{1}{\mathrm{min}}$
$1000\,H + 4000\,ATP \rightarrow E_E$	R	$1\ \frac{1}{\mathrm{min}}$
$1500\,H + 6000\,ATP \rightarrow E_F$	R	$0.67\ \frac{1}{\mathrm{min}}$
$500\,H + 2000\,ATP \rightarrow E_G$	R	$2\ \frac{1}{\mathrm{min}}$
$2500\,H + 10000\,ATP \rightarrow E_H$	R	$0.4\ \frac{1}{\mathrm{min}}$
$500\,H + 2000\,ATP \rightarrow E_N$	R	$2\ \frac{1}{\mathrm{min}}$
$500\,H + 2000\,ATP \rightarrow E_T$	R	$2\ \frac{1}{\mathrm{min}}$
$2000\,H + 4000\,C + 16000\,ATP \rightarrow R$	R	$0.2\ \frac{1}{\mathrm{min}}$
$250\,H + 250\,C + 250\,F + 1500\,ATP \rightarrow S$	R	$3\ \frac{1}{\mathrm{min}}$

$\rightsquigarrow$ preferred carbon source Carb1

- Carbon switch
  - Low amount of preferred carbon source C_1
  - High amount of non-preferred carbon source C_2
  - Ample oxygen supply

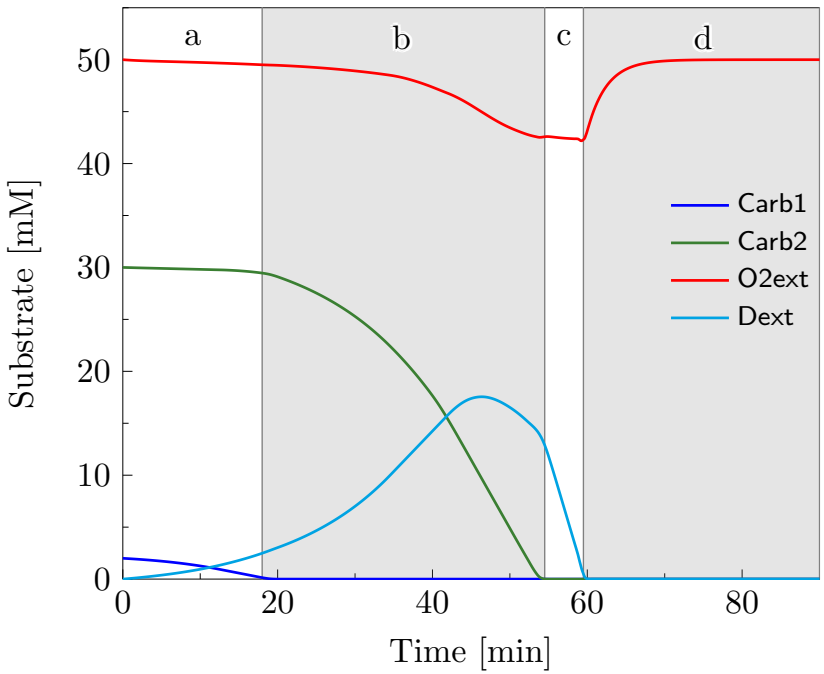
  $\rightsquigarrow$ first consume C_1, then C_2.

- Objective: Discounted biomass

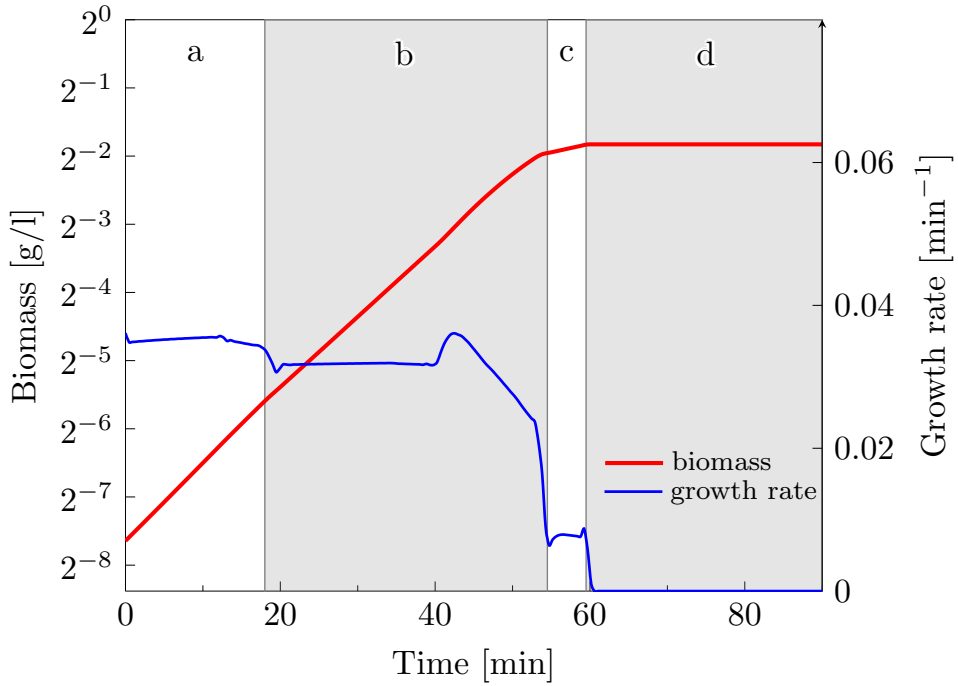
$$J = \int_0^{t_{end}} c_{bm}^{\mathrm{T}} P(t)\, e^{-\mu t}\, dt$$

- Information in the model: Stoichiometry, biomass composition, kcats.
- No information on enzyme activities

Substrates

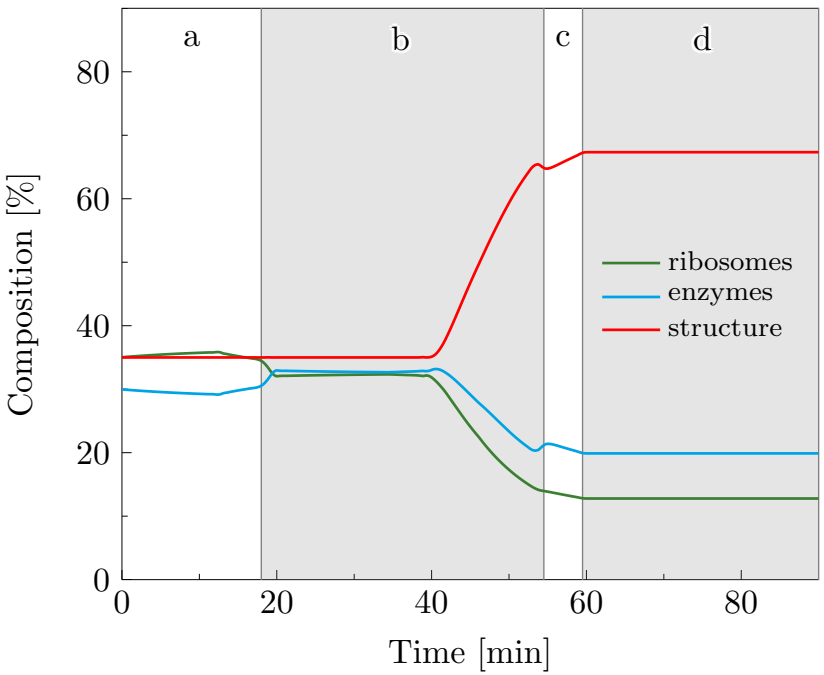


Biomass & Growth

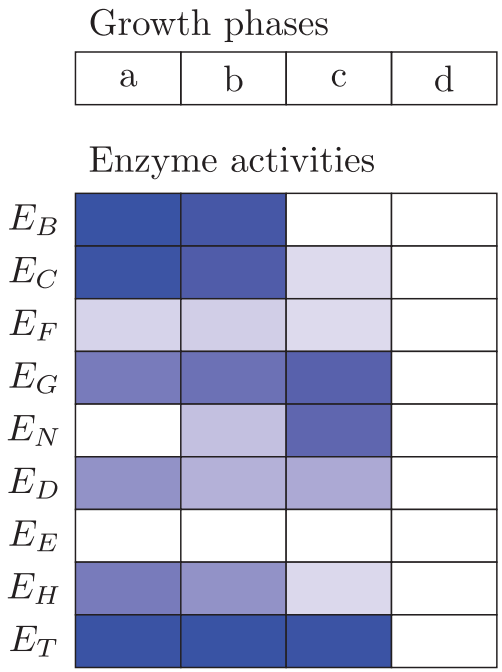


Four cellular growth phases

Cell composition



Enzyme activities



Cellular reorganisation at the end of phase (b)

Modeling metabolism including enzyme costs

- Mass balance ODE model for cellular metabolism and biomass production.
- Time-scale separation yields dynamic model with quasi steady-state metabolic constraints.
- Dynamic optimisation framework: deFBA

Related work

- Model of core phototrophic metabolism of cyanobacteria under day/night conditions (with M. Rügen, R. Steuer)
- Model of genome-scale reconstruction of *Synechococcus elongatus 7942* (with A.-M. Reimers, R. Steuer)
- ERASysApp project ROBUSTYEAST (with S. Waldherr, F. Bruggeman, V. Hatzimanikatis)

- Orth JD, Thiele I, Palsson BO. What is flux balance analysis? Nat Biotechnol., 28(3):245-8, 2010.
- Lewis NE, Nagarajan H, Palsson BO. Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silico methods. Nat Rev Microbiol., 10(4):291-305, 2012.
- Waldherr S, Oyarzún DA, Bockmayr A. Dynamic optimization of metabolic networks coupled with gene expression. J Theor Biol, 365, 469-485, Jan 2015
- Rügen M, Bockmayr A, Steuer R: Elucidating temporal resource allocation and diurnal dynamics in phototrophic metabolism using conditional FBA. Scientific Reports, 5:15247, Oct 2015