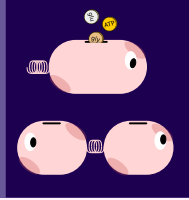
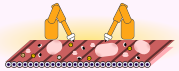
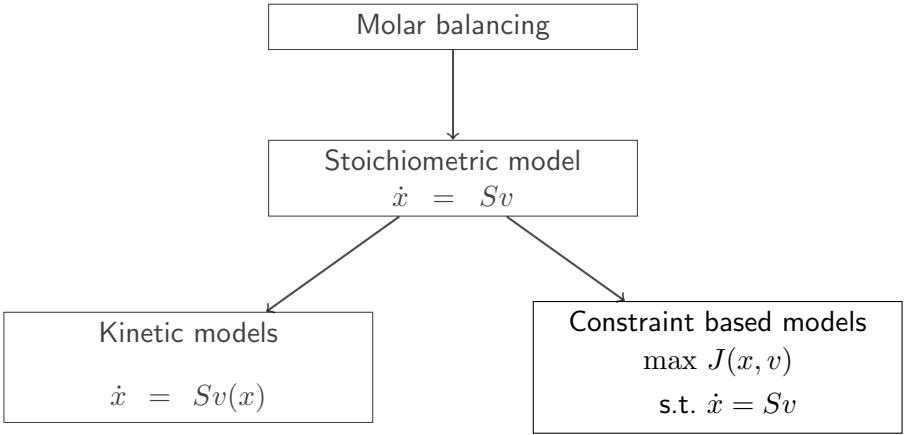


Flux balance analysis

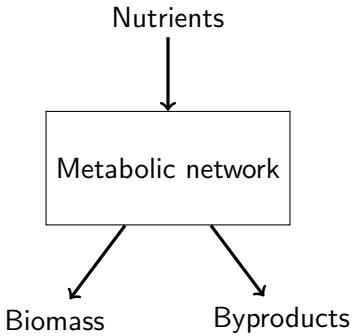
Steffen Waldherr



From metabolic networks to models



A whole-cell perspective on metabolism



Optimality principle instead of kinetics: Maximize growth subject to flux balance and uptake constraints

Restrictions for "basic" FBA:

- Fixed composition of biomass
- (Quasi-)steady state
- Metabolite concentrations fall out of the model
- Limited consideration of thermodynamics

Constraints on fluxes

1. Steady state constraint

$$Sv = 0$$

   - Fluxes constrained to **subspace**

2. Irreversibility constraints on some fluxes (from thermodynamics/heuristics/empirical evidence)

$$v_i \geq 0, \qquad i \text{ irreversible}$$

   - Fluxes constraint to **flux cone**

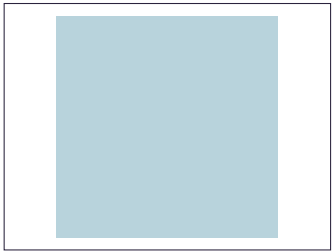
3. Flux bounds from capacity constraints, maintenance, ...

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

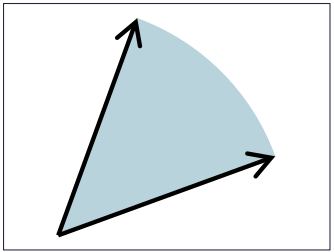
   - Fluxes constraint to **convex polytope**

Geometric illustration

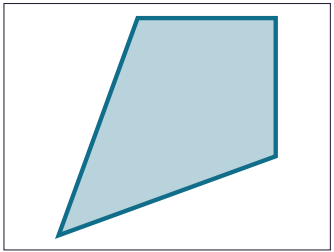
Flux space



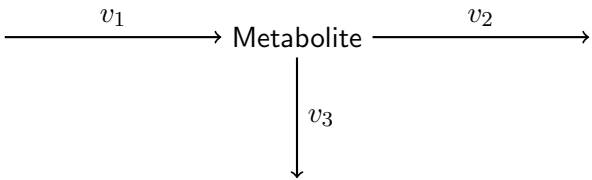
Flux cone



Flux polytope



Flux space $\rightarrow$ cone $\rightarrow$ polytope example



Construct the ...

- flux space;
- flux cone assuming $v_2,\ v_3 \geq 0$;
- flux polytope assuming $v_1 \leq 0.5$.

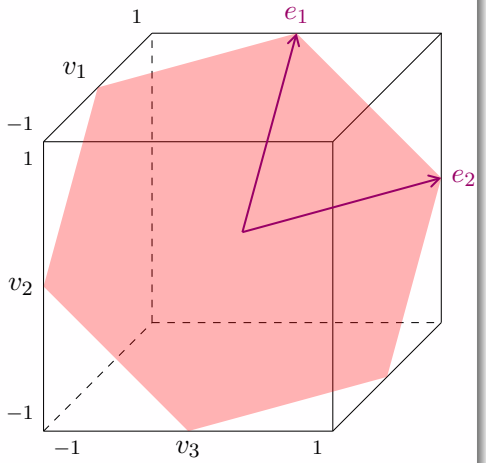
Molar balancing:

$$\dot{x} = \begin{pmatrix} 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

Flux space from $Sv = 0$

- Plane defined by

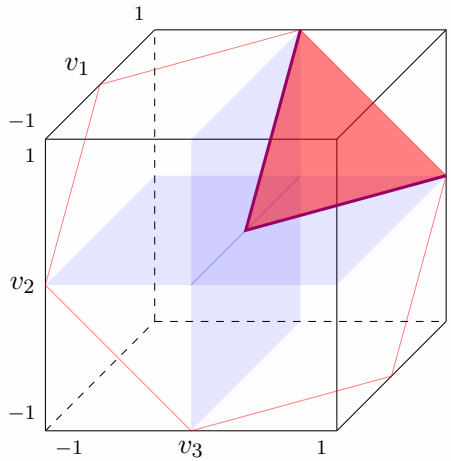
$$v_1 - v_2 - v_3 = 0$$



Flux cone

- Add irreversibility

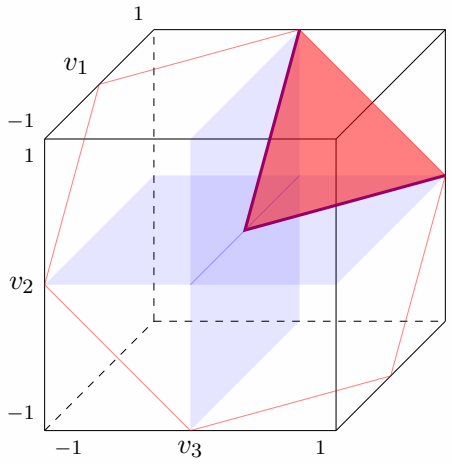
$$v_2,\ v_3 \geq 0$$



Flux cone

- Add irreversibility

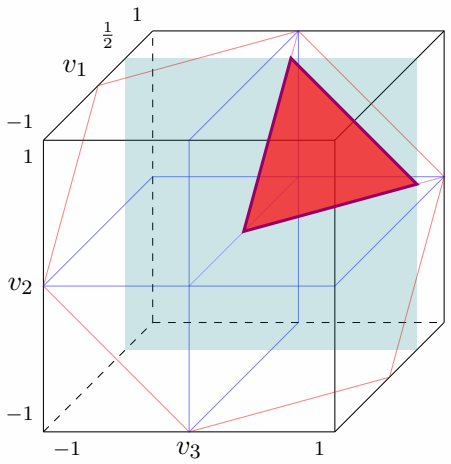
$$v_2,\ v_3 \geq 0$$



Flux polytope

- Add upper bound(s)

$$v_1 \leq 0.5$$



Setting up the constraint based model (CBM)

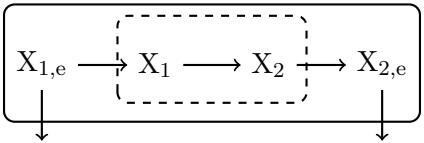
Constraint based model useful if non-trivial steady state fluxes exist

- The steady state equation
$$Sv = 0$$
should have a non-zero solution $v \Rightarrow$ non-trivial steady state flux space
- We need $\operatorname{rank} S < m$; most models have more reactions than metabolites anyway.

Metabolite / flux units

- In CBMs, metabolites are usually considered in **molar amounts per dry biomass**: $\mathrm{mmol/g}$
- Fluxes are then in $\mathrm{mmol/(gh)}$

System and cellular exchange fluxes



$$S = \begin{pmatrix} -1 & 0 & 0 & -1 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{pmatrix}$$

- Steady state on system level: $Sv = 0$
- "Outer" exchange fluxes for nutrients / secretions
- Convention: Outer exchange fluxes go outwards of system
- Consumption: negative flux — Production: positive flux

Biomass composition

E. coli biomass composition

Compound	Proportion [% g/g DW]
Protein	72
DNA	4
RNA	10
Lipids	9
Polysaccharides	2.5
Mureine	2.5

Chassagnole *et al.* 2002, via bionumbers.hms.harvard.edu, ID 108705
Varies depending on environmental conditions (nutrients, aerobic/anaerobic, growth rate, ...)

Break down to metabolites

- 20 proteinogenic amino acids
- 8 D/R nucleotides
- phospholipids
- cofactors / vitamins
- ATP hydrolysis required for biomass assembly ("**growth associated maintenance**" GAM)

Biomass reaction

- Biomass reaction formalizes consumption of metabolites to generate biomass

$$v_{bio}: \qquad \sum_{i=1}^{n} c_i X_i \rightarrow 1 \text{ g dry biomass}$$

- Based on pre-determined constant biomass composition
- Coefficients c_i commonly in mmol / g dry biomass
- Unit of v_{bio} becomes $1/\mathrm{h}$: interpretable as dry biomass growth rate μ!

Maintenance

- "**Non-growth associated maintenance**" (NGAM):
  - membrane voltage gradients and osmolarity (ion pumps)
  - movement (flagella)
  - macromolecule (RNA/protein/carbohydrates) turnover
- Energy demand is commonly represented by a single ATP hydrolysis reaction

$$v_{maint}: \qquad \mathrm{ATP} + \mathrm{H_2O} \rightarrow \mathrm{ADP} + \mathrm{Pi} + \mathrm{H^+}$$

- Put as constraint into constraint based model
  - $v_{maint} \geq \alpha$ [mmol / (h · g biomass)]
  - NGAM rate estimates: *E. coli* 8.4 mmol/g/h; *S. cerevisiae* 1.0 mmol/g/h

Optimization principle

Constraint based model

$$Sv = 0$$

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

- **Underdetermined** system of equalities / inequalities: flux polytope
- How do we determine fluxes v that we expect to occur in nature?

Add an optimization objective

- **Hypothesis:** Cells regulate fluxes within constraints to achieve an "optimal" configuration from an evolutionary perspective.

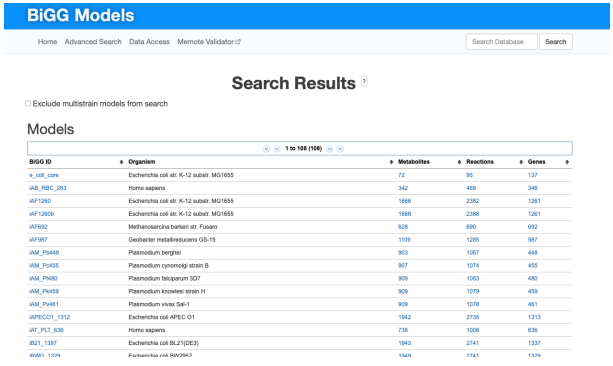
$$\begin{aligned} \max\ & J(v) \\ \text{s.t.}\ \ & Sv = 0 \\ & v_{i,min} \leq v_i \leq v_{i,max} \end{aligned}$$

Useful objective functions

Type	Objective $J(v)$	Principle
Biomass yield	$\max v_{bio}$	Biomass flux at fixed max. substrate uptake
ATP yield	$\max v_{ATP}$	ATP flux at fixed max. substrate uptake
Minimal flux	$\min \|v\|^2$	Minimization of overall flux ($\sim$ enzyme usage)
Biomass flux yield	$\max v_{bio}/\|v\|^2$	Biomass yield per overall flux unit

Empirical evaluation of objective functions: Schuetz, R., Kuepfer, L., & Sauer, U. (2007). Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*. Molecular Systems Biology, 3, 119.

Collections of constraint based models

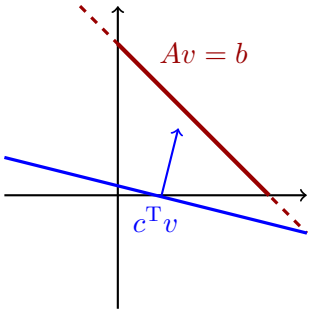


BiGG models database: `http://bigg.ucsd.edu/models`

Linear programs

A linear program in standard form:

$$\begin{aligned} \max\ & c^{\mathrm{T}} v \\ \text{s.t. } & Av = b \\ & v \geq 0 \end{aligned}$$



Objective

$c^{\mathrm{T}} v$

Equality constraint

$Av = b$

Inequality constraint

(Cone constraint) $v \geq 0$

Example

$$\max_{v_1, v_2} v_2$$

$$\text{s.t. } v_1 + v_2 = 1$$

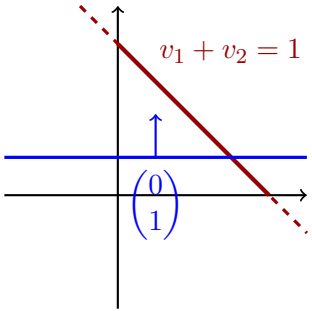
$$v \geq 0$$

Thus:

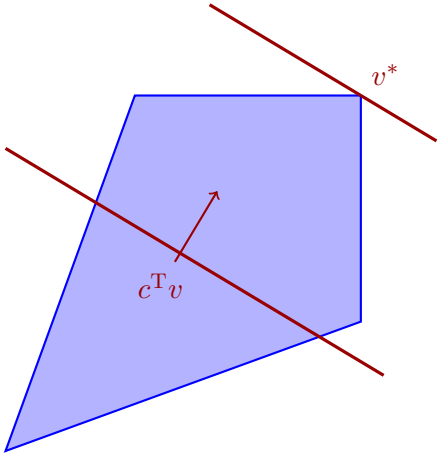
$$c^{\mathrm{T}} = \begin{pmatrix} 0 & 1 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 1 \end{pmatrix}$$

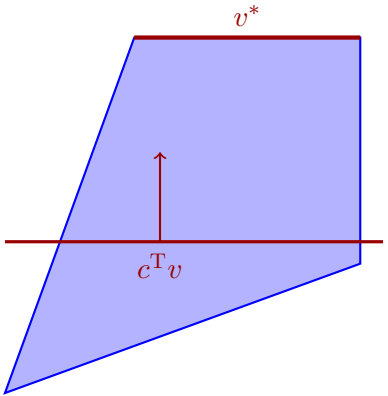
$$b = 1$$



Generalized geometrical interpretation

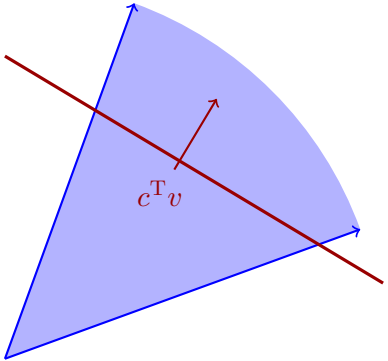


Non-uniqueness of optimal solutions



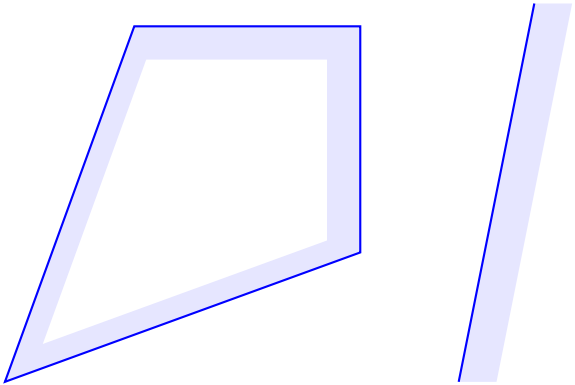
Set of optimal solutions is a **face of the polytope**

Unboundedness



Unboundedness: $\max c^{\mathrm{T}}v = \infty$

Infeasibility: Constraint set is empty



Example

$$v_1 + v_2 \leq -1$$
$$v_1, v_2 \geq 0$$

Flux balance analysis (FBA)

FBA to maximize biomass yield as LP

$$
\begin{aligned}
J^* = \max\ & v_{bio} \\
\text{s.t. } & Sv = 0 \\
& v_{i,min} \leq v_i \leq v_{i,max}
\end{aligned}
$$

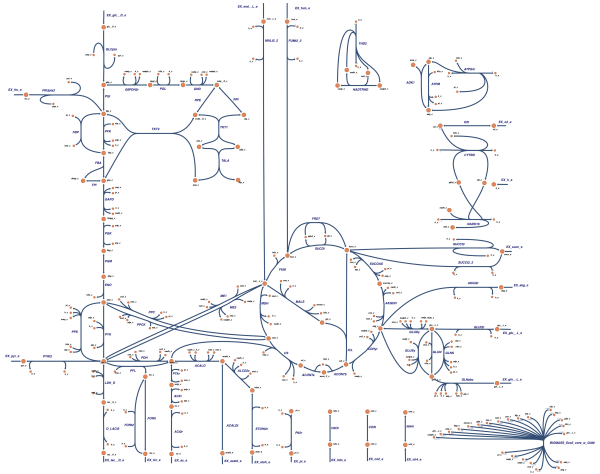
- Typical relevant constraint is glucose / oxygen uptake rate

$$-v_{e,gluc,max} \leq v_{e,gluc} \leq 0$$

- For practical reasons $v_{i,max} = M$ (10^6 mole/h/g) even if no capacity constraint
- Typically no unique optimal flux distribution v^*

FBA example: *E. coli* core

- Core carbon network from BiGG database: 72 metabolites, 95 reactions
- Network visualization from `https://escher.github.io/`



Flux variability analysis (FVA)

What are maximum / minimum fluxes achievable within an "optimal" flux distribution?

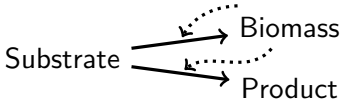
1. Compute objective for optimal flux distribution

$$J^* = \arg\max\ c^{\mathrm{T}} v$$
$$\text{s.t. } \ldots$$

2. Put objective as constraint and optimize individual fluxes

$$\bar{v}_j = \arg\max / \min\ v_j$$
$$\text{s.t. } Sv = 0$$
$$v_{i,min} \leq v_i \leq v_{i,max}$$
$$c^{\mathrm{T}} v \geq (1-\varepsilon) J^*$$

Product yield vs. biomass yield



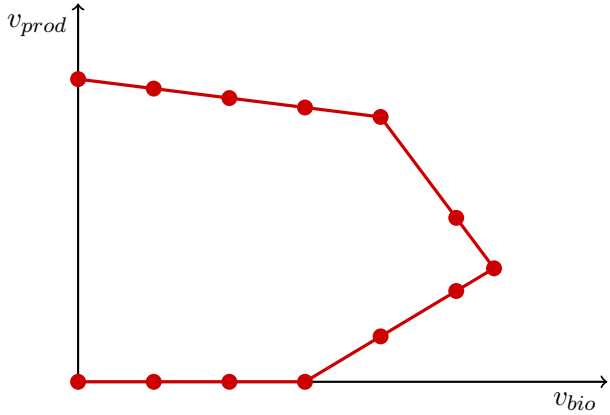
- Extracellular biochemical product $\mathrm{P_e}$, with exchange reaction

$$v_{prod} : \mathrm{P_e} \longrightarrow$$

Minimize / maximize v_{prod} at fixed biomass flux $\bar{v}_{bio}$

$$\begin{aligned} \max \quad & \pm v_{prod} \\ \text{s.t.} \quad & Sv = 0 \\ & v_{i,min} \leq v_i \leq v_{i,max} \\ & v_{bio} = \bar{v}_{bio} \end{aligned}$$

Production envelopes



1. Fix growth rate (v_{bio}) at different values
2. Compute maximum / minimum for product flux for each growth rate by FVA

Outlook: extensions of FBA

- Dynamic FBA
- Thermodynamic FBA
- Resource allocation models:
  - ME models
  - Resource balance analysis
  - Dynamic enzyme-cost FBA

Exercise this afternoon

Run FBA on the carbon core model (Jupyter notebook → Google Colaboratory)