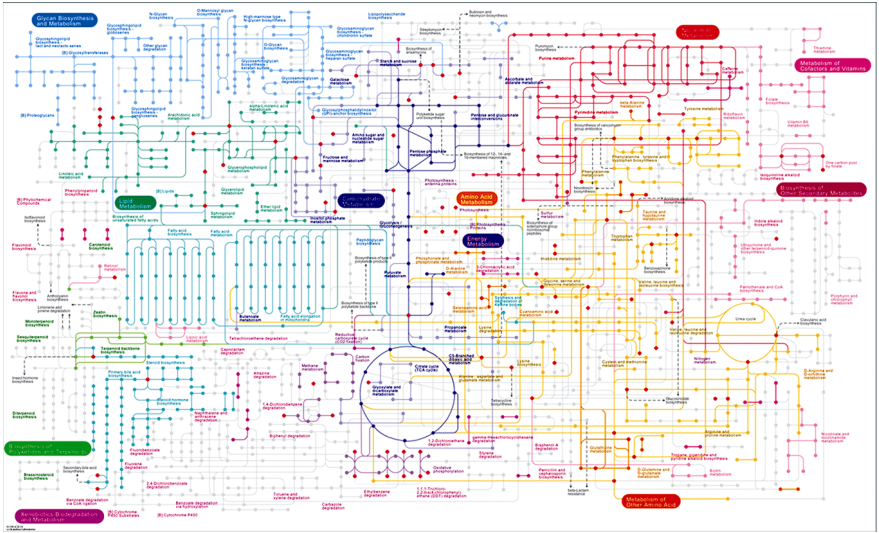
Diversity of metabolic flux distributions

Roberto Mulet

The Problem


Glycan Biosynthesis and Metabolism
Nucleotide Metabolism
Metabolism of Cofactors and Vitamins
Biosynthesis of Other Secondary Metabolites
Lipid Metabolism
Carbohydrate Metabolism
Amino Acid Metabolism
Energy Metabolism
Biosynthesis of Polyketides and Terpenoids
Xenobiotics Biodegradation and Metabolism
Metabolism of Other Amino Acid

The simplest math

$$\frac{dX_i}{dt} = \sum_j S_{ij}\nu_j(\vec{X})$$

- X_i concentration of metabolite $i \in [1, ..M]$
- ν_j velocity of reaction $j \in [1, ..N]$
- S_{ij} Stoichiometric Matrix
- $N > M$

Stationarity

$$\frac{dX_i}{dt} = \sum_j S_{ij}\nu_j(\vec{X}) = 0$$

Constraint modelling

$$\mathcal{S}\vec{\nu} = 0$$

Stationarity

$$\frac{dX_i}{dt} = \sum_j S_{ij}\nu_j(\vec{X}) = 0$$

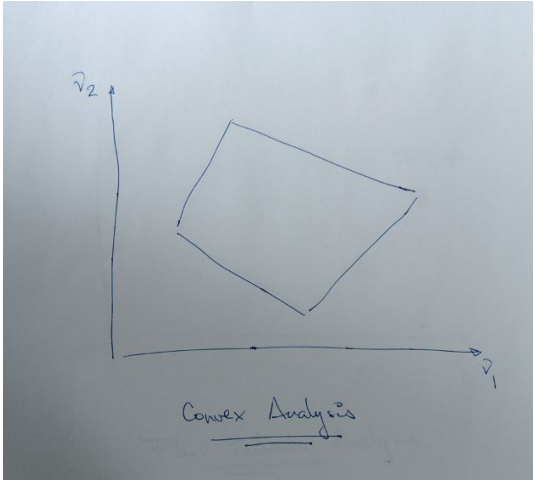
Constraint modelling

$$\mathcal{S}\vec{\nu} = 0$$

Constraint modelling

$$\mathcal{S}\vec{\nu} = \vec{b}$$

Graphical Representation



Additional Assumption

- Maximize: $E = \sum_j h_j \nu_j$

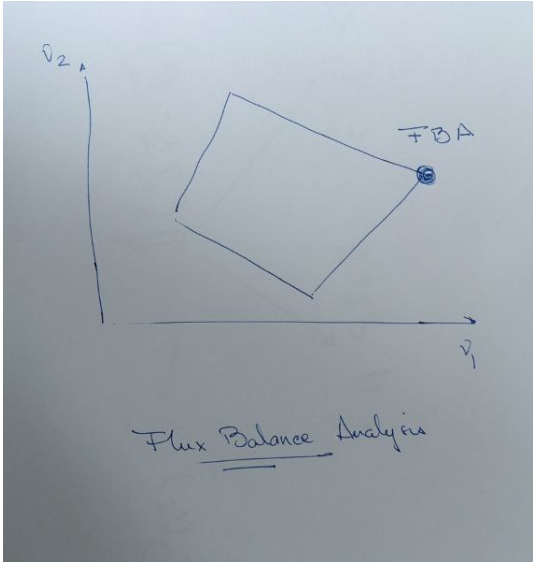
Additional Assumption

- Maximize: $E = \sum_j h_j \nu_j$

Flux Balance Analysis = Linear Programming

$$\mathcal{S}\vec{\nu} = \vec{b}$$
$$\max_{\vec{\nu}} E$$

Graphical Representation



Experimental Support

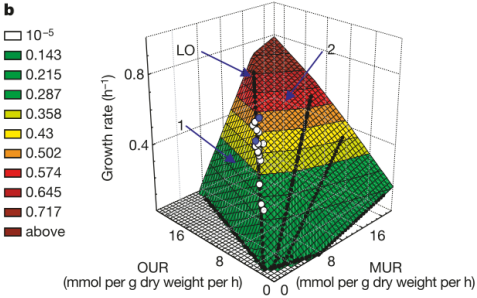


Figure 1 Growth of *E. coli* K-12 on malate. **a**, The malate–oxygen phenotype phase plane (PPP) Phase 1 is characterized by metabolic futile cycles, whereas phase 2 is characterized by acetate overflow metabolism. The line of optimality (LO, in red) separates phases 1 and 2 (ref. 21.) Data points (open circles) represent malate concentrations ranging from 0.25–3 g l^{-1}; and temperatures ranging from 29–37 °C. The two data points in blue represent the starting point (day 0) and endpoint (day 30) of adaptive evolution respectively, at a malate concentration of 2 g l^{-1} and a temperature of 37 °C. These data points represent a span of 500 generations. **b**, Three-dimensional representation of growth rates. The *x* and *y* axes represent the same variables as in **a**. The *z* axis represents the cellular growth rate (h^{-1}). OUR, oxygen uptake rate; MUR, malate uptake rate.

But Life is more complex than that

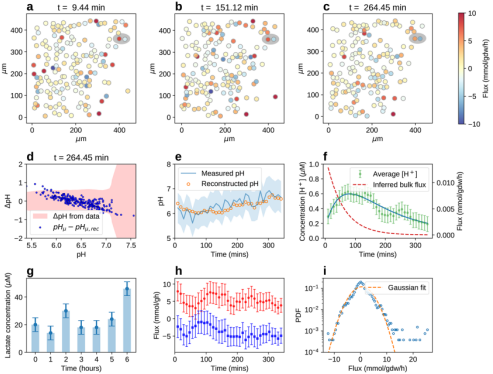
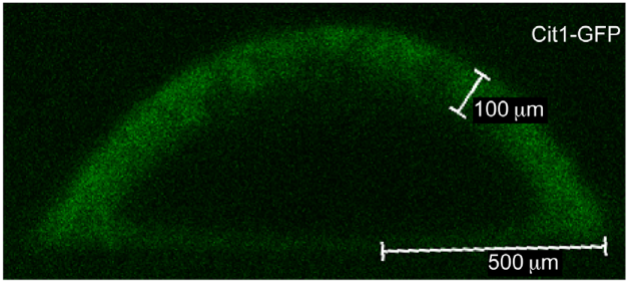


Figure 3. (a−c) Snapshots at different time points (at t_a = 9 min, t_B = 151 min, and t_c = 264 min after the cell culture is settled, all frames are reported in the Supporting Information Figures S1−S5) of the same square visual field (length L = 500 μm) during a typical experiment. Cells are represented schematically as disks of diameter 10 μm whose color intensity scales with the flux (side bar, blue vs red for importing vs exporting flux). Probes not shown. (d,e) Quality of the reconstructed pH gradient profile. In (d), the error between the pH calculated from the inferred fluxes and the experimentally observed pH is plotted against the latter for each probe (at time t_c = 264 min, all frames are reported in the Supporting Information Figures S6−S10). In (e), the time trace of the pH measured by a given probe is reported alongside the reconstructed trend at that spatial point. Shaded areas represent the experimental error on the pH at the probes. (f) Time trends of the bulk [H^+] concentration (experimental, dots and reconstructed, continuous line, left y scale) and inferred bulk acidic efflux (dashed line, right y scale). (g) Time trend of the experimentally measured bulk lactate concentration in a biological replicate. (h) Single-cell flux intensity (in mmol/gdw/h) as a function of time (in min, sampling every 10 min) of the cells forming the dipole motif highlighted in the upper right corner of the frames in (a−c). (i) Single-cell experimental flux distribution (in mmol/gdw/h, (dots) and its Gaussian approximation (lines) in linear-logarithmic scale. The histogram is built from all single-cell flux values (100−200 cells per frame) and time frames (36 frames resulting from a 6 h experiment sampled every 10 min) tracked in one visual field of one experiment.

V. Onesto el al, Probing Single-Cell Fermentation Fluxes and Exchange Networks via pH-Sensing Hybrid Nanofibers, ACS Nano 2023, 17, 4, 3313–3323

But Life is more complex than that



A. Traven et al, Transcriptional profiling of a yeast colony provides new insight into the heterogeneity of multicellular fungal communities. PLoS One. 2012;7(9):e46243.

But Life is more complex than that

REVIEWS

Physiological heterogeneity in biofilms

Philip S. Stewart[*1] *and Michael J. Franklin*[*§]

Abstract | Biofilms contain bacterial cells that are in a wide range of physiological states. Within a biofilm population, cells with diverse genotypes and phenotypes that express distinct metabolic pathways, stress responses and other specific biological activities are juxtaposed. The mechanisms that contribute to this genetic and physiological heterogeneity include microscale chemical gradients, adaptation to local environmental conditions, stochastic gene expression and the genotypic variation that occurs through mutation and selection. Here, we discuss the processes that generate chemical gradients in

The ISME Journal (2018) 12:1199–1209
https://doi.org/10.1038/s41396-017-0036-2

ARTICLE

The emergence of metabolic heterogeneity and diverse growth responses in isogenic bacterial cells

Emrah Şimşek[1] · Minsu Kim[1,2]

Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Microbiology
ELSEVIER

Metabolic heterogeneity in clonal microbial populations

Vakil Takhaveev and Matthias Heinemann

In the past decades, numerous instances of phenotypic diversity were observed in clonal microbial populations, particularly, on the gene expression level. Much less is,

extreme case of subpopulations having distinctly different activities of metabolic pathways [4–6]. Furthermore, recent discoveries show that even under constant condi-

Building a Probability Measure

$$P(\vec{v}) \sim \mathbb{1}(S\vec{v} - \vec{b})$$

$$1 \text{ if } S\vec{v} = \vec{b} \qquad 0 \text{ if } S\vec{v} \neq \vec{b}$$

Building a Probability Measure

$$\mathbb{P}(\vec{v}) \sim e^{\beta \sum_i h_i v_i} \mathbb{1}(S\vec{v} - \vec{b})$$

if $\beta \to \infty$ $\equiv$ Dominated by largest v_i

if $\beta \to 0$ $\equiv 1$

Building a Probability Measure

$$\boxed{P(\vec{v}) = \frac{1}{Z} e^{\rho \vec{h} \cdot \vec{v}} \mathbb{1}(S\vec{v} - \vec{b})}$$

Maximum Entropy Principle

$S = -\int_{\mathcal{P}} P(\nu) \log P(\nu)$

Among all the probability densities compatible with the data (or knowledge), the one having the largest value of S is the one that best represents our knowledge of the system

Derivation

$$max_{P(\nu)} - [\int_{\mathcal{P}} P(\nu) \log P(\nu)]$$

Derivation

$$max_{P(\nu)} - [\int_{\mathcal{P}} P(\nu) \log P(\nu)]$$

subject to: $\int_{\mathcal{P}} P(\nu) = 1$ and $\int_{\mathcal{P}} f(\nu) P(\nu) =< f >$

Derivation

$$\mathcal{L} = -\int_{\mathcal{P}} P(\nu)\log P(\nu) - \alpha(\int_{\mathcal{P}} P(\nu) - 1) - \beta(\int_{\mathcal{P}} f(\nu)P(\nu) - < f >)$$

Derivation

$$\mathcal{L} = -\int_{\mathcal{P}} P(\nu)\log P(\nu) - \alpha(\int_{\mathcal{P}} P(\nu) - 1) - \beta(\int_{\mathcal{P}} f(\nu)P(\nu) - < f >)$$

$$P(\nu) \sim e^{\beta f(\nu)}$$

Summarizing

Probability densities over the flux polytope 267

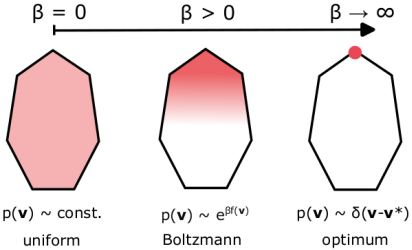
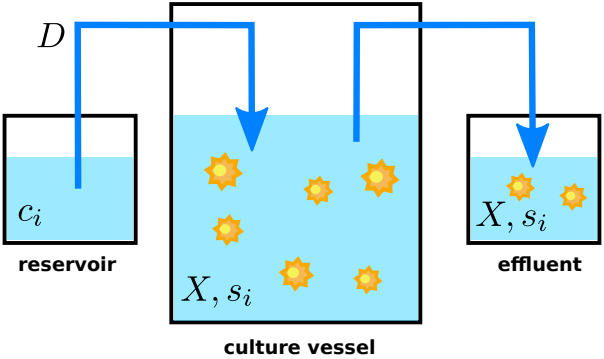


Figure 19.3: Boltzmann distribution on the flux polytope. The Boltzmann distribution, Eqn (19.10), morphs from a uniform probability density to a δ-distribution concentrated on the flux vector that maximizes the function f as β varies from 0 to $+\infty$.

Problem One



Mathematical framework

$$\frac{dX}{dt} = (\mu - \sigma - D)X$$

$$\mu = \mu(u, r) \qquad \sigma = \sigma(s)$$

$$\frac{ds_i}{dt} = -u_i X - (s_i - c_i)D$$

$$lb_k \leq r_q \leq ub_k$$

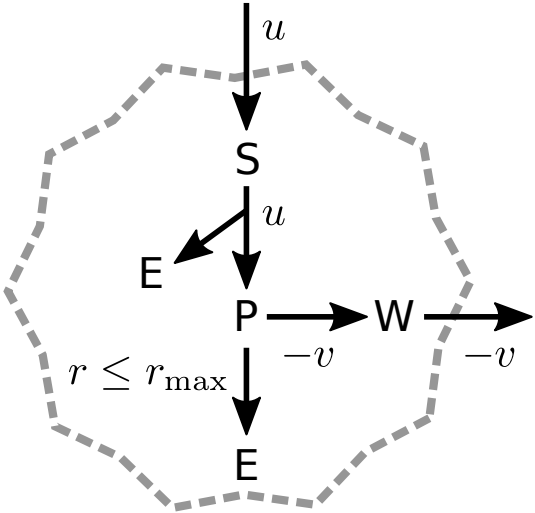
$$-L_i \leq u_i \leq min\{V_i, c_i \frac{D}{X}\}$$

$$\sum_k r_k < K$$

$$\sum_k S_{ik} r_k - e_i - y_i \mu + u_i = 0$$

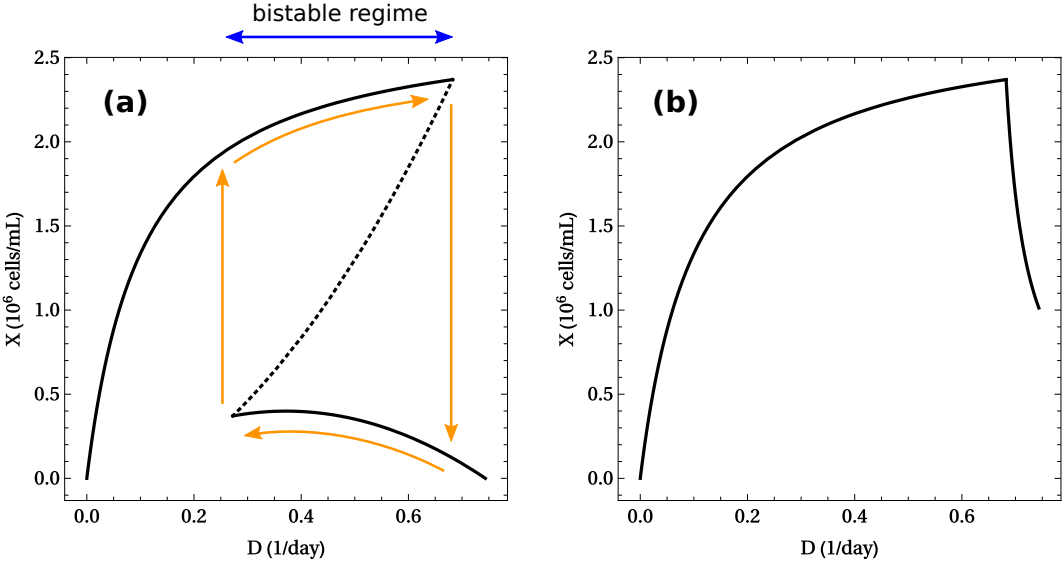
The cell maximizes biomass production μ

Small Network

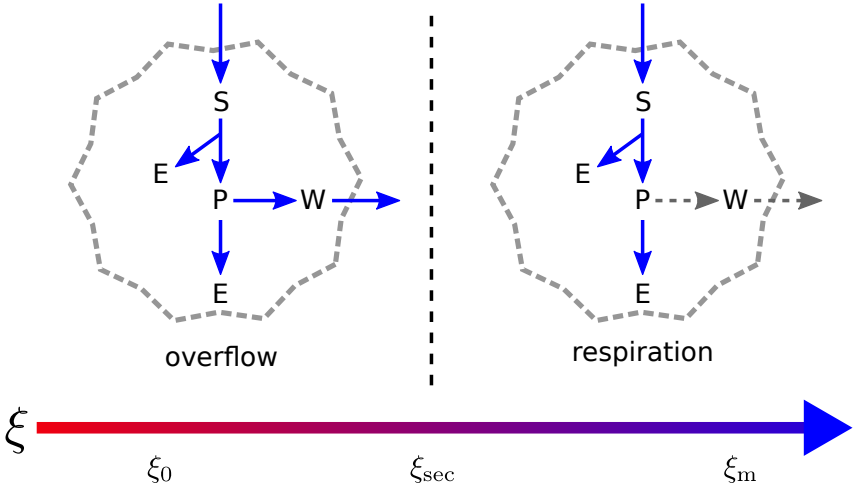


Vazquez et al.. Macromolecular crowding explains overflow metabolism in cells. Scientific Reports 6, 31007 (2016)

Toxicity is the key point



General Picture



(a) Overflow. At high enough nutrient uptake the respiratory flux hit s the upper bound $r_{\max}$ and the remaining nutrients are exported as W. **(b)** Respiration. The nutrient is completely oxidized with a large energy yield. **(c)** Threshold values of ξ. ξ_0 delimits the nutrient excess regime ($\xi < \xi_0$) from the competition regime ($\xi > \xi_0$). ξ_{sec} delimits the transition between overflow metabolism ($\xi < \xi_{\mathrm{sec}}$ and respiration ($\xi > \xi_{\mathrm{sec}}$).

Homogeneous Chemostat

$$\frac{dX}{dt} = (\mu - \sigma - D)X = 0$$

Homogeneous Chemostat

$$\frac{dX}{dt} = (\mu - \sigma - D)X = 0$$

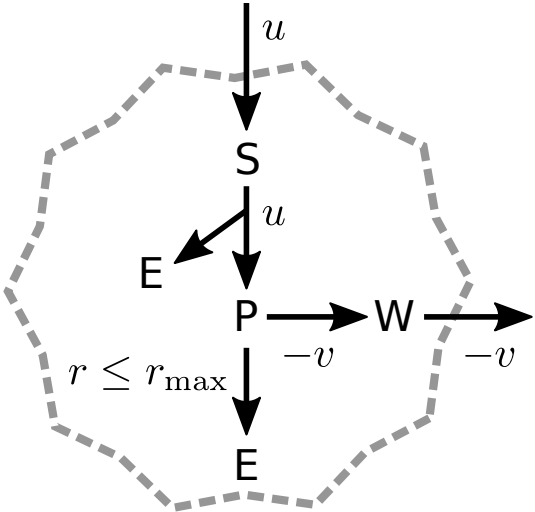
If stationary $\mu(u, r) - \sigma(s^*) = D$

Hetergeneous Chemostat

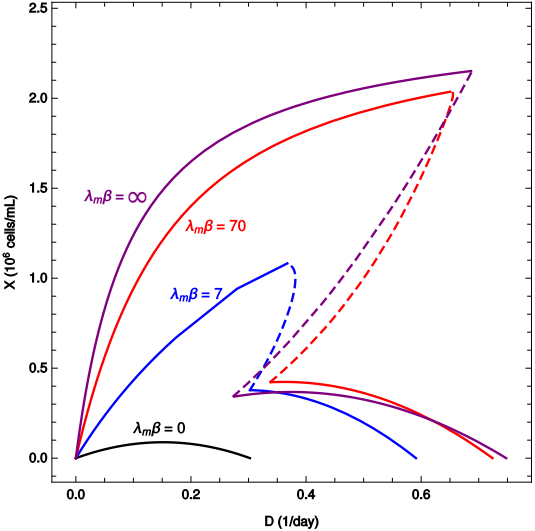
$$D = \frac{X}{\xi} = \int_{\Pi} (\mu(\nu) - \sigma(s^*)) P_{s^*}(u, r) d(u, r)$$

$$s_i^* = c_i - \xi \int_{\Pi} u_i(\nu) P_{s^*} d(u, r)$$

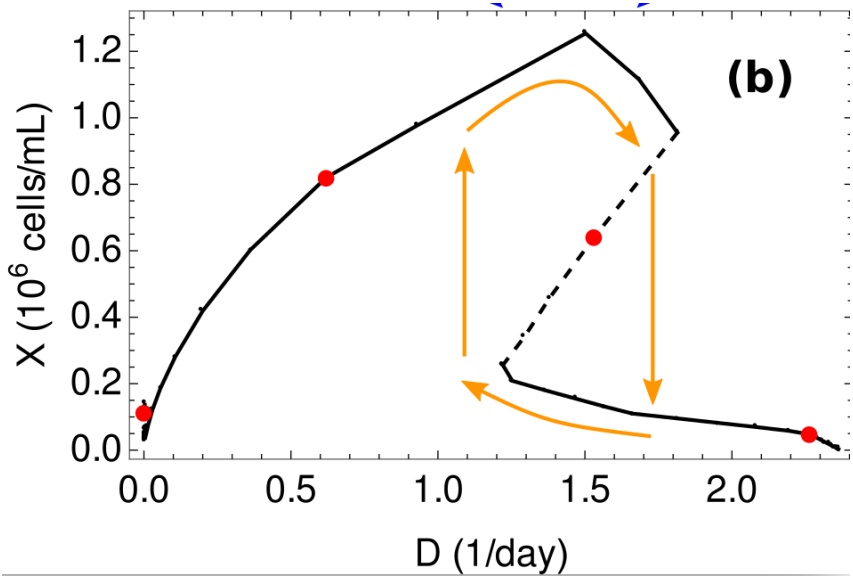
Small Network again



Effect of the heterogeneity

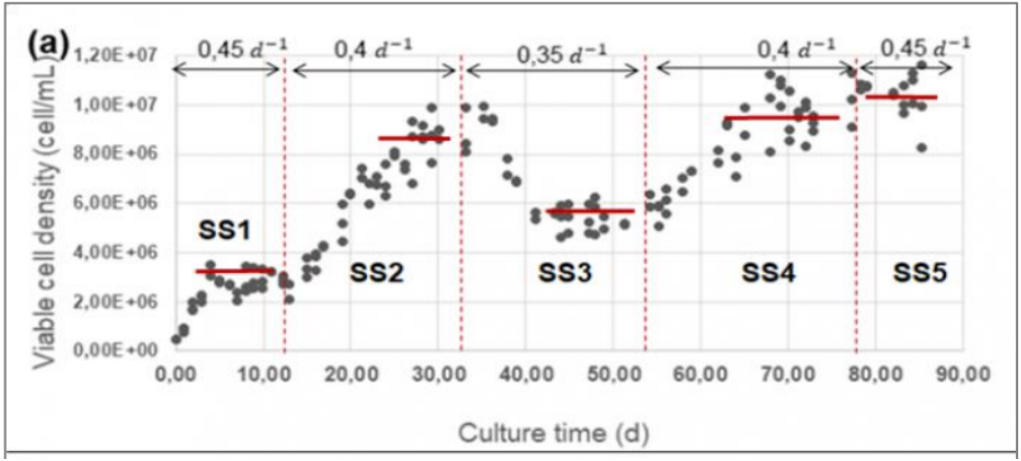


Genome Scale Metabolic Network

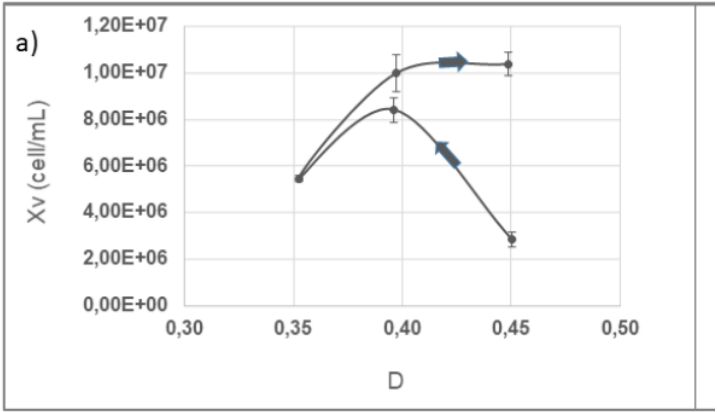


J. Fernandez-de-Cossio

Experimental results



Experimental results



L. Calzadilla-Rosado, E. Hernández, J. Dustet, J. Fernández-de-Cossio-Díaz, M. Pietzke, A. Vazquez, K. León, **R. M.** *and T. Boggiano, Multiple steady states and metabolic switches in continuous cultures of HEK293: Experimental evidences and metabolomics analysis, Biochemical Engineering Journal* **198**, *109010 (2023)*

Problem Two

Can we estimate the fluxes of the cultures in a chemostat?

Standard approach

Given some constraints: $\mathcal{S}\vec{\nu} = \vec{b}$

$$< v_i >_{exp} \in \text{polytope}$$

Find $\underline{v}$ such that some function $f(\underline{v})$ is maximum

Alternatively: Maximum Entropy

Given some constraints: $\mathcal{S}\vec{\nu} = \vec{b}$ and:

$$< v_i >_{exp} = \int \underline{v} P(\underline{v}) dv \in \text{polytope}$$

$$\underset{P(\underline{v})}{\text{argMaximize}} S = -\int P(\underline{v}) \log P(\underline{v}) dv$$

Toy model

$$\frac{dX}{dt} = (\mu - D)X$$

$$\mu = \mu(u, r)$$

$$\frac{ds}{dt} = -uX + (c_R - s)$$

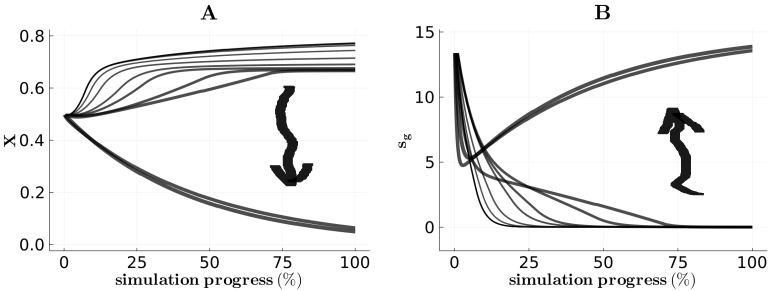
Toy model

$$
\begin{aligned}
\frac{dX(\mu, u_g, u_o)}{dt} &= \mu\, X(\mu, u_g, u_o) - D\, X(\mu, u_g, u_o) \\
\frac{ds_g}{dt} &= -\int_V u_g, X(\mu, u_g, u_o)\, d\mu\, du_g\, du_o + (c_g - s_g)\, D
\end{aligned}
$$

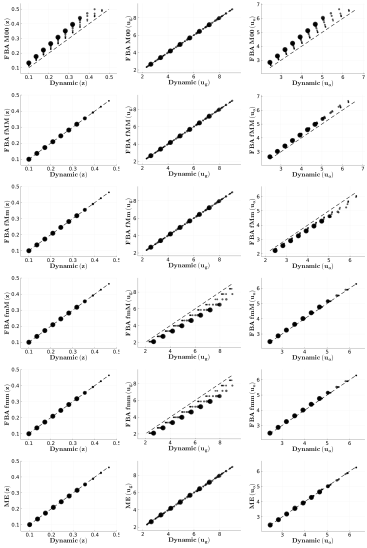
Toy model

$$
\begin{aligned}
\frac{dX(\mu, u_g, u_o)}{dt} &= (1-\epsilon)\,\mu\, X(\mu, u_g, u_o) - D\, X(\mu, u_g, u_o) + \\
&+ \frac{\epsilon}{V} \int_V \mu'\, X(\mu', u'_g, u'_o)\, d\mu', du'_g du'_o \\
\frac{ds_g}{dt} &= -\int_V u_g, X(\mu, u_g, u_o)\, d\mu\, du_g\, du_o + (c_g - s_g)\, D
\end{aligned}
$$

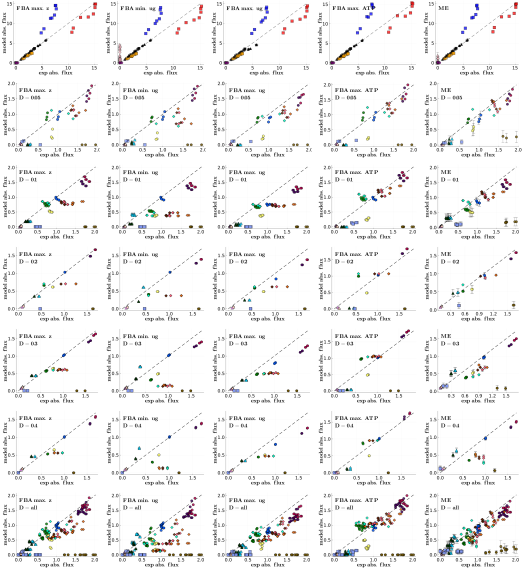
Toy model


A
0.8
0.6
0.4
0.2
0.0
X
0
25
50
75
100
simulation progress (%)
B
15
10
5
0
s_g
0
25
50
75
100
simulation progress (%)

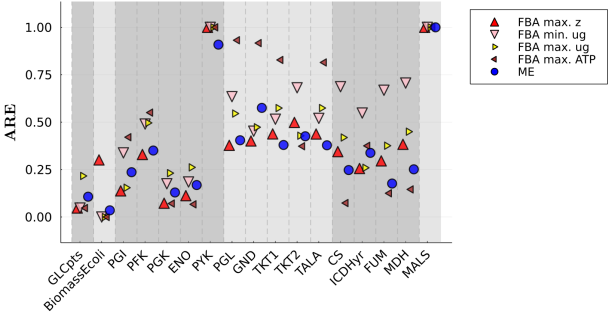
Inference in heterogeneous metabolism



Results for the Genome Scale Ecoli model



Results for the Genome Scale Ecoli model



J.A. Pereiro-Morejón, J. Fernández-de-Cossio Díaz and **R.M**, *Inferring metabolic fluxes in nutrient-limited continuous cultures: A Maximum Entropy Approach with minimum information, iScience* **25**, *105450 (2022)*

Results for the Genome Scale Ecoli model

4

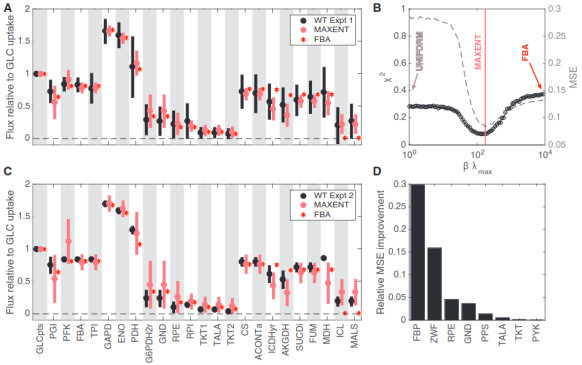


FIG. 2. **Maximum entropy model outperforms FBA flux predictions in *Escherichia coli* during steady state growth. (A, C)** Comparison of measured fluxes (black, mean ± SD over biological replicates; normalized to glucose uptake) with predictions of FBA (red stars) and of the maximum entropy model (pink, mean ± SD from the predicted joint distribution over fluxes). Data for (A) are a collection of 35 experiments from Ref [25]; data for (C) are three replicates from Ref [26] . Wild type *E. coli* was grown in glucose-limited medium with low dilution (growth rates (below 0.4 h^{-1}, no acetate excretion)

D. De Martino et al, Statistical mechanics for metabolic networks during steady state growth, Nat. Comm. **9**, 2988 (2018)

Conclusions

- Every *direct* problem can be studied in a probabilistic setting
- This probabilistic setting can be used to solve every *inverse* problem
- There is a lack of experimental results