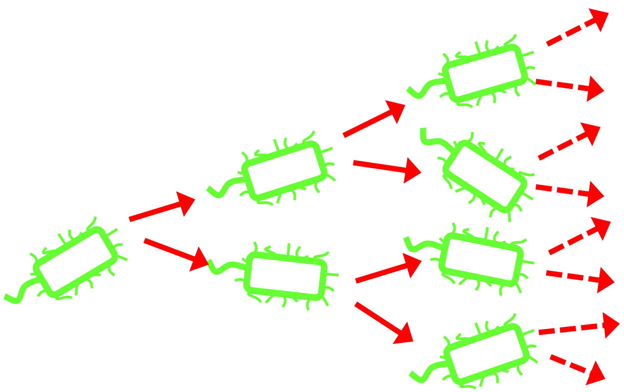
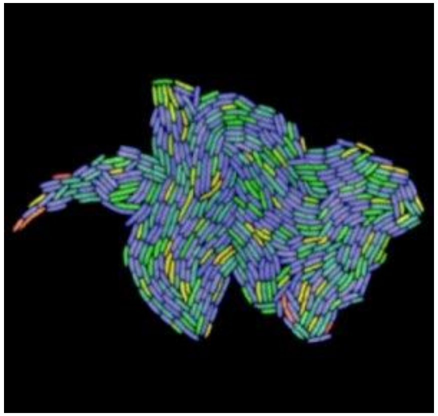
# Dynamic optimization of resource allocation in microorganisms

RESET Workshop, October 2017

**Jean-Luc Gouzé**
INRIA Sophia

# Bacterial growth

- Bacteria are unicellular organisms geared towards **growth**
  - *E. coli* cells have doubling times up to 20 min


Stewart *et al.* (2005), *PLoS Biol.*, 3(2): e45

- **Metabolism** fuels growth by production of energy and building blocks for macromolecules
  - Different **growth rates** depending on medium (carbon, nitrogen, …)

# Growth and macromolecular composition

- Macromolecular composition of cell varies with growth rate
  - Quantity of DNA, RNA, protein, … per cell or per volume

**TABLE 3** Parameters pertaining to the macromolecular synthesis rates in exponentially growing *E. coli* B/r as a function of growth rate at 37°C

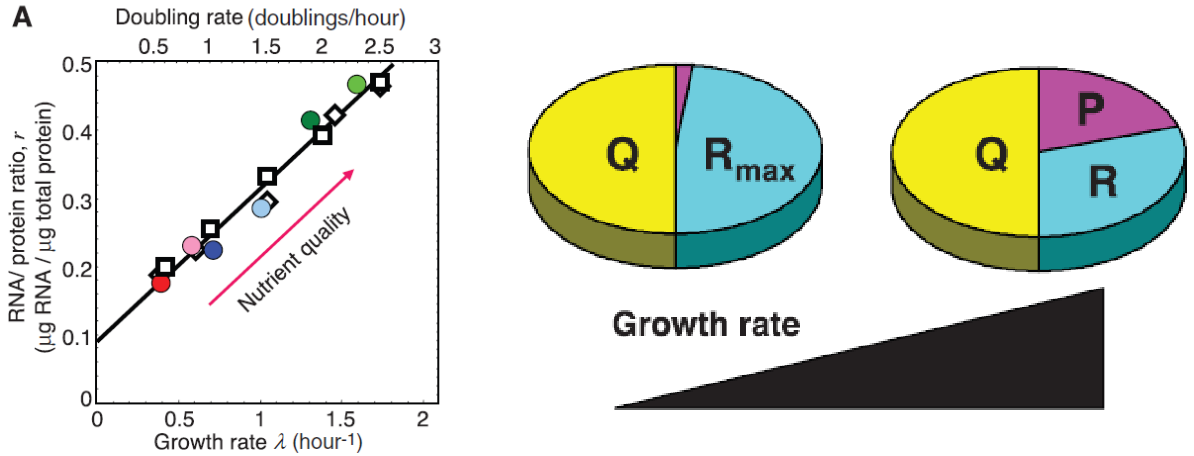
| Parameter | Symbol | Units | At τ (min) and μ (doublings per h): | | | | | Observed parameter(s) | Footnote |
|---|---|---|---|---|---|---|---|---|---|
| | | | τ, 100<br>μ, 0.6 | τ, 60<br>μ, 1.0 | τ, 40<br>μ, 1.5 | τ, 30<br>μ, 2.0 | τ, 24<br>μ, 2.5 | | |
| RNA polymerase protein/total protein | $\alpha_p$ | % | 0.90 | 1.10 | 1.30 | 1.45 | 1.55 | $\alpha_p$ | *a* |
| RNA polymerase molecules/cell | $N_p$ | $10^3$ RNAP/cell | 1.5 | 2.8 | 5.0 | 8.0 | 11.4 | $\alpha_p$, $P_C$ | *b* |
| RNA polymerase activity | $\beta_p$ | % | 17 | 20 | 21 | 24 | 30 | $r_s, r_m, c_s, c_m, N_p$ | *c* |
| Active RNA polymerase per cell | $N_{ap}$ | RNAP/cell | 205 | 503 | 992 | 1,929 | 3,298 | | *c* |
| Stable RNA synthesized per total RNA synthesized | $r_s/r_t$ | % | 41 | 52 | 66 | 78 | 85 | $r_s/r_t$ | *d* |
| Active RNA polymerase synthesizing stable RNA | $\psi_s$ | % | 24 | 36 | 56 | 68 | 79 | $r_s/r_t$<br>$c_s/c_m$ | *e* |
| rRNA chain elongation | $c_s$ | Nucl./s | 85 | 85 | 85 | 85 | 85 | Indirect | *f* |
| mRNA chain elongation | $c_m$ | Nucl./s | 39 | 45 | 50 | 52 | 55 | Indirect | *g* |
| Rate of stable RNA synthesis/cell | $r_s$ | $10^5$ nucl./min/cell | 3.0 | 9.9 | 29.0 | 66.4 | 132.5 | $R_C$ | *h* |
| Rate of mRNA synthesis/cell | $r_m$ | $10^5$ nucl./min/cell | 4.3 | 9.2 | 13.7 | 18.7 | 23.4 | $r_s$, $r_s/r_t$ | *i* |
| ppGpp concentration | ppGpp/$M$ | pmol/$OD_{460}$ | 55 | 38 | 22 | 15 | 10 | ppGpp/$M$ | *j* |
| | ppGpp/$P$ | pmol/$10^{17}$aa | 8.5 | 6.6 | 4.2 | 2.9 | 2.0 | $P_M$ | *j* |
| r-Protein per total protein | $\alpha_r$ | % | 9.0 | 11.4 | 14.8 | 17.5 | 21.1 | $P_M$, $R_M$ | *k* |
| | | | 9 | 11 | 13.5 | 18.0 | 21.6 | $\alpha_r$ | *l* |
| Ribosome activity | $\beta_r$ | % | 80 | 80 | 80 | 80 | 80 | Indirect | *m* |
| Peptide chain elongation | $c_p$ | aa residues/s | 12 | 16 | 18 | 20 | 21 | Indirect | *n* |
| Ribosomes/cell | $N_r$ | $10^3$ ribosomes/cell | 6.8 | 13.5 | 26.3 | 45.1 | 72.0 | $R_C$, $f_s$, $f_t$ | *o* |
| tRNA/cell | $N_t$ | $10^3$ tRNA/cell | 65 | 125 | 244 | 419 | 669 | $N_r$, $f_t$ | *p* |
| *rrn* genes/cell | $N_{rrn}$ | Avg no./cell | 12.4 | 15.1 | 20.0 | 26.9 | 35.9 | $C$, $D$ | *q* |
| *rrn* genes/genome | $N_{rrn}/G$ | Avg no./genome | 7.9 | 8.2 | 8.6 | 9.0 | 9.5 | $C$ | *r* |
| Initiation rate at *rrn* gene | $i_{rrn}$ | Initiations/min/gene | 4 | 10 | 23 | 39 | 58 | $N_r$, $N_{rrn}$ | *s* |
| Distance of ribosomes on mRNA | $R_m/N_r$ | Nucl./ribosome | 79 | 85 | 65 | 52 | 41 | $r_m, c_m, N_r$ | *t* |

Bremer and Dennis (1996), *Escherichia Coli and Salmonella*, ASM Press, 1553-69

# Growth and macromolecular composition

- Phenomenological **growth laws** capture variation of macromolecular composition with growth rate
  - Distribution of proteins over different categories

Scott *et al.* (2010), *Science*, 330(6007):1099-102




- Explanation of growth laws (implicitly or explicitly) based on **optimization principle**
  - Bacteria have evolved so as to distribute limited resources over cellular processes in order to optimize growth (biomass)

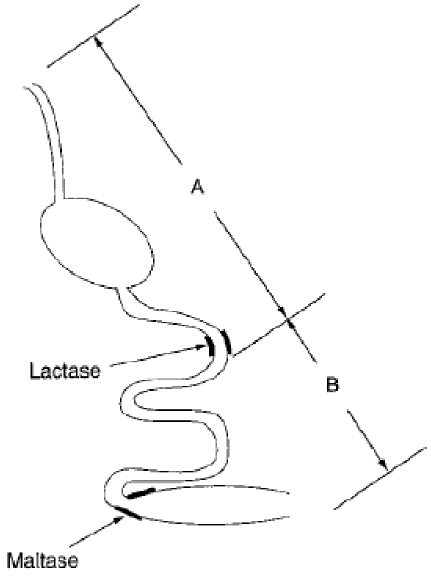
# Growth and optimization

"The aim of the RESET project is to break with these classical approaches and propose a novel strategy for improving product yield and productivity."

- Optimization: models and optimal control
- How does the cell optimize its growth?
- How can we optimize production with external inducer?
- A nice blend of biology, modelling, and mathematics…

# Steady-state and dynamic optimization

- Most growth laws and data concern **steady state** (balanced growth)
  - Well-controlled and reproducible in laboratory
- However, most bacteria evolve in **dynamic** environment
  - Example: *E. coli* in human colon

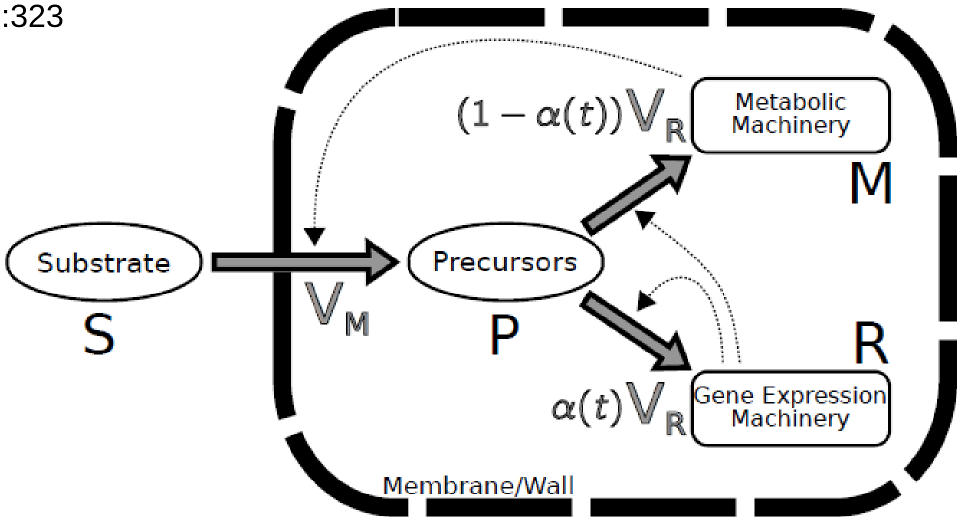



Savageau (1983), *Am. Natural.*, 122(6):732-44

# Towards dynamic growth laws

- **Aim**: study optimal allocation of resources to gene expression machinery and metabolism during growth-phase transitions
  - Which allocation is optimal for sustaining maximal growth (biomass)?
- Simple model of cell: bacteria as **self-replicators**

Molenaar *et al.* (2009), *Mol. Syst. Biol.*, 5:323

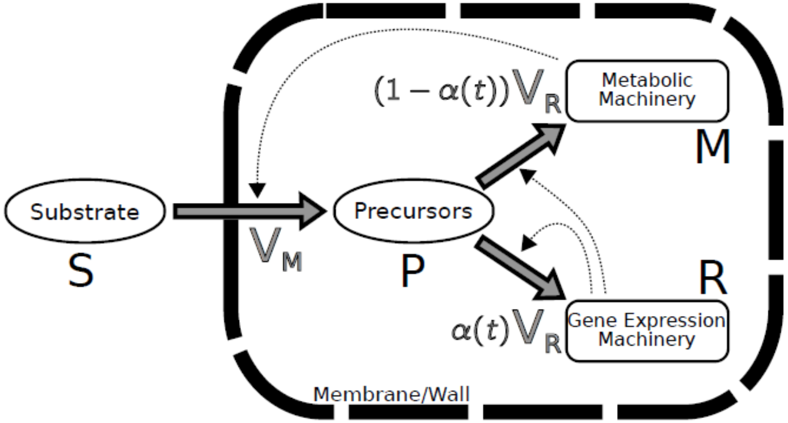



- Tools from **optimal control** theory

# Self-replicator model of cell

- Reaction scheme:

$$S \xrightarrow{V_M} P$$
$$nP \xrightarrow{V_R} \alpha R + (1-\alpha)M$$




- Stochiometry model with **extensive** variables:

$$\frac{d}{dt}\begin{bmatrix} P \\ M \\ R \end{bmatrix} = \begin{bmatrix} 1 & -n \\ 0 & 1-\alpha \\ 0 & \alpha \end{bmatrix} \cdot \begin{bmatrix} V_M \\ V_R \end{bmatrix} = N \cdot V$$

- Volume and growth rate:

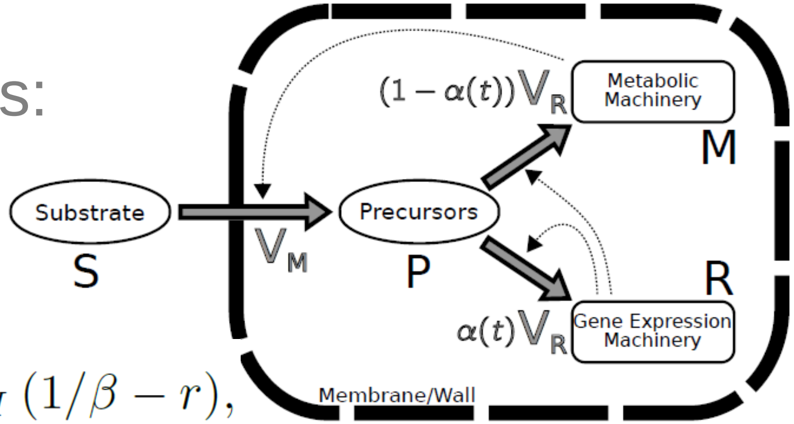
$$Vol = \beta(M+R)$$

$$\mu = \frac{1}{Vol}\frac{d\,Vol}{dt} = \frac{1}{M+R}\frac{d(M+R)}{dt}$$

# Reformulated self-replicator model of cell

- Definition of intensive variables:

$$p = \frac{P}{Vol}, \; m = \frac{M}{Vol}, \text{ and } r = \frac{R}{Vol}$$




- Kinetics:

$$v_M = k_M \frac{s}{K_M + s} m = e_M \left(1/\beta - r\right),$$

$$v_R = k_R \frac{p}{K_R + p} r,$$

$$\mu = \frac{V_R}{R + M} = \beta v_R.$$

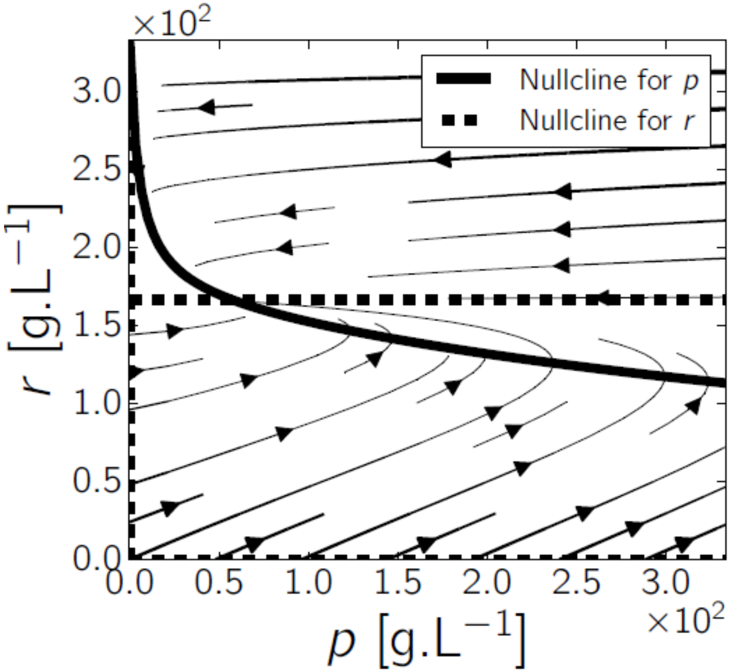
- Model with **intensive** (dimensionless) variables:

$$\begin{cases} \frac{dp}{dt} = E_M \cdot (1 - r) - (1 + p)\frac{p}{K+p} r, \\ \frac{dr}{dt} = (\alpha - r)\frac{p}{K+p} r. \end{cases}$$

# Steady-state analysis of model

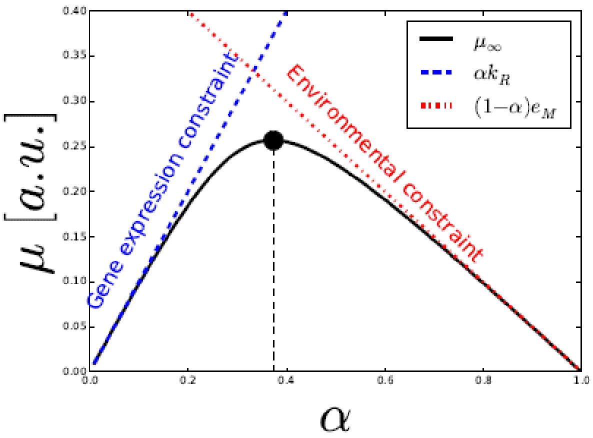
- Control parameter $\alpha$ determines fractional distribution of resources over metabolic and gene expression subsystems
- **Result**: for constant $\alpha$, the system has a single steady state with growth rate $\mu^*(p^*, r^*, \alpha)$

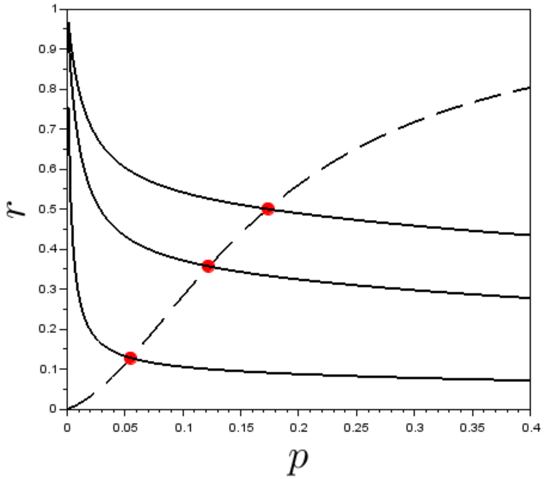
$$\begin{cases} \frac{dp}{dt} = E_M \cdot (1-r) - (1+p)\frac{p}{K+p}r, \\ \frac{dr}{dt} = (\alpha - r)\frac{p}{K+p}r. \end{cases}$$



# Steady-state analysis of model

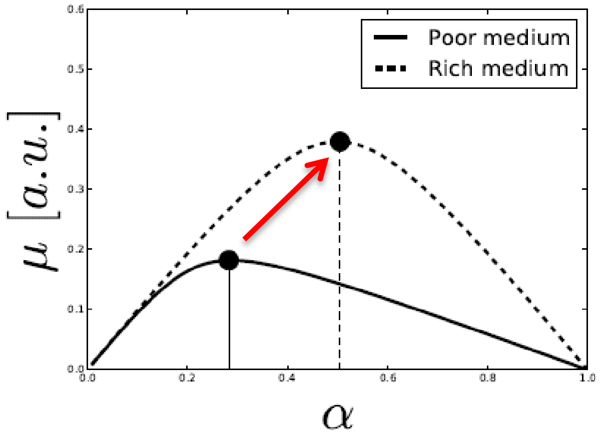
- **Result**: system admits single maximum growth rate for value $\alpha = \alpha_{opt} \in [0, 1]$
  - Maximum varies with medium quality, represented by parameter $e_M$



# Dynamic optimal control problem

- Bacterial cell has to reallocate resources after change in environment to reach optimal growth rate (change $\alpha$)




- What is the **best dynamic resource allocation strategy**?
- **Optimal control problem** for biomass

$$\max_{\alpha \in \mathcal{U}} J(\alpha) := \int_0^{+\infty} \mu(p, r, \alpha, t) dt$$

with set of admissible controls: $\mathcal{U} = \{\alpha : \mathbb{R}^+ \to [0, 1]\}$

# Pontryagin Maximum Principe

$$H := \lambda_p E_M(1-r) - \frac{p}{K+p} r \left[\lambda_p(1+p) + \lambda_r r + \lambda_0\right] + \alpha\lambda_r \frac{p}{K+p} r.$$

$$\dot{\lambda}_p = \frac{K}{(K+p)^2} r \left[\lambda_p(1+p) + \lambda_r(r-\alpha) + \lambda_0\right] + \frac{p}{K+p} r \lambda_p,$$
$$\dot{\lambda}_r = \lambda_p E_M + \frac{p}{K+p} \left[\lambda_p(1+p) + \lambda_r(2r-\alpha) + \lambda_0\right].$$

The maximization condition is given by:

$$\alpha(t) \in \operatorname{argmax}_{v\in[0,1]} H(x(t), \lambda(t), \lambda_0, v),$$
$$\text{a.e. } t \in [0, +\infty).$$

The switching function:

$$\phi := \lambda_r \frac{p}{K+p} r$$

$$\begin{cases} \alpha = 1 \iff \phi > 0, \\ \alpha = 0 \iff \phi < 0. \end{cases}$$

# Characterization of singular arcs

- The singular arc corresponds to the optimal steady-state:

$$\phi(t) = \dot{\phi}(t) = 0, \forall t \in [t_1, t_2] \quad \Longrightarrow \quad (\hat{p}(t), \hat{r}(t)) = (\hat{p}^\star_{opt}, \hat{r}^\star_{opt})$$

- Kelley condition:

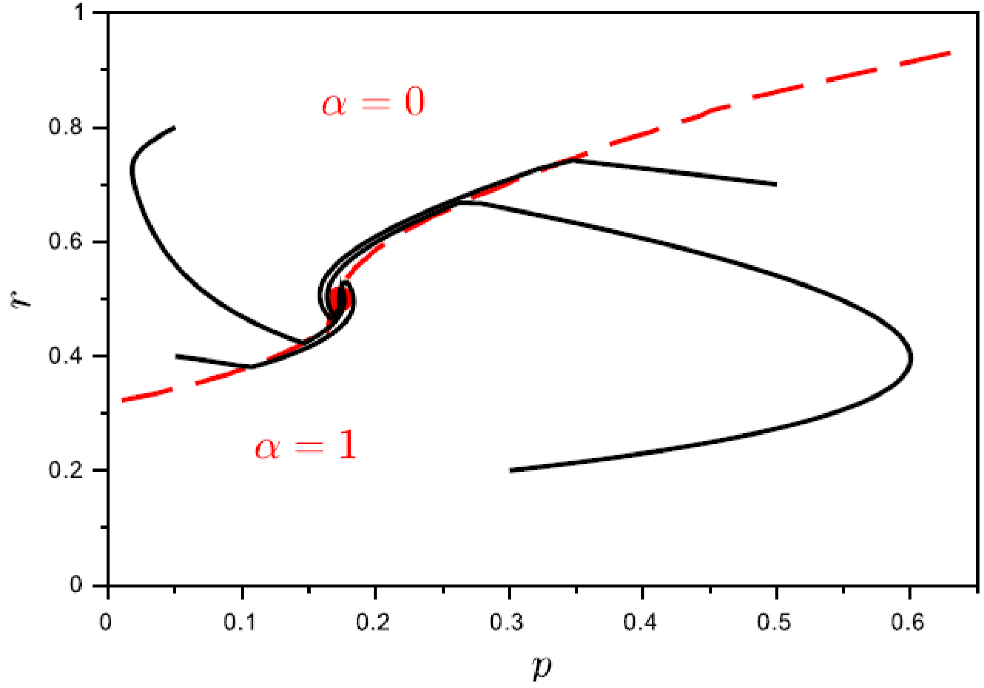
$$(-1)^q \frac{\partial}{\partial \alpha} \frac{d^{2q}}{dt^{2q}} \phi(t) < 0 \quad \text{for } q = 2 \quad \Longrightarrow$$

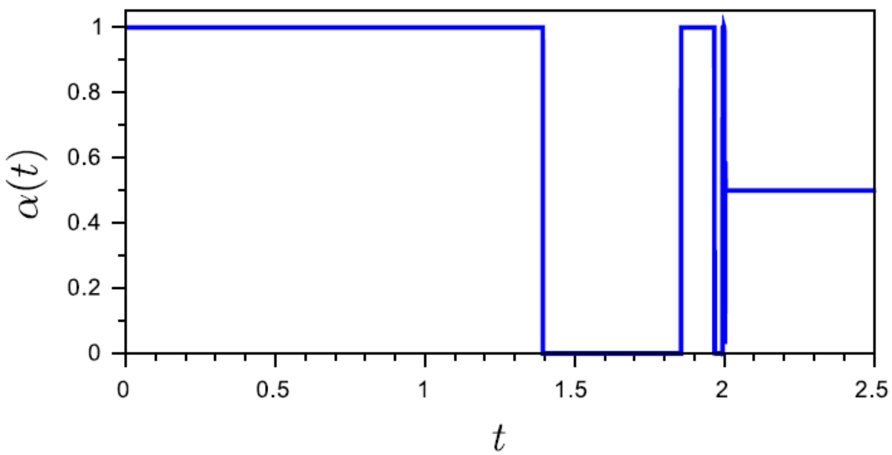
Chattering arc
(Fuller's phenomena)

- Optimal strategy: Turnpike?

# Optimal strategy

- Numerical solutions by a direct method (bocop)

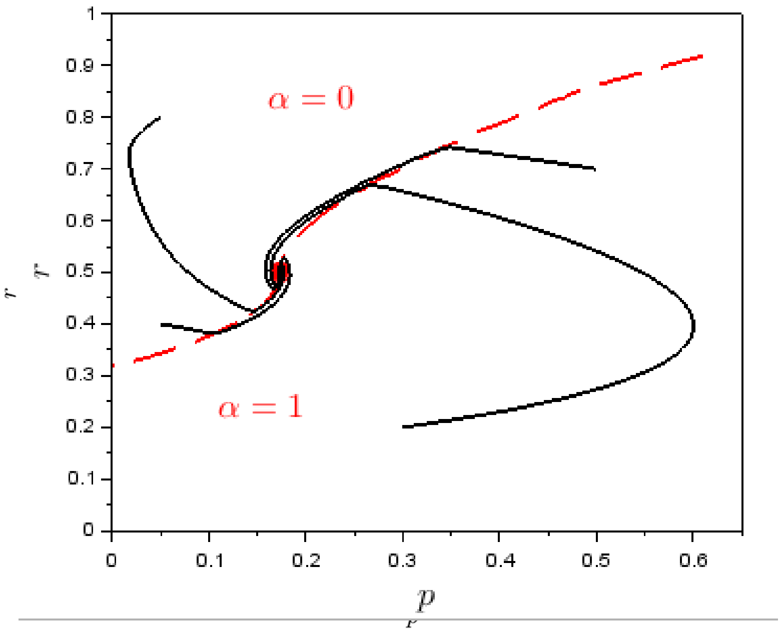


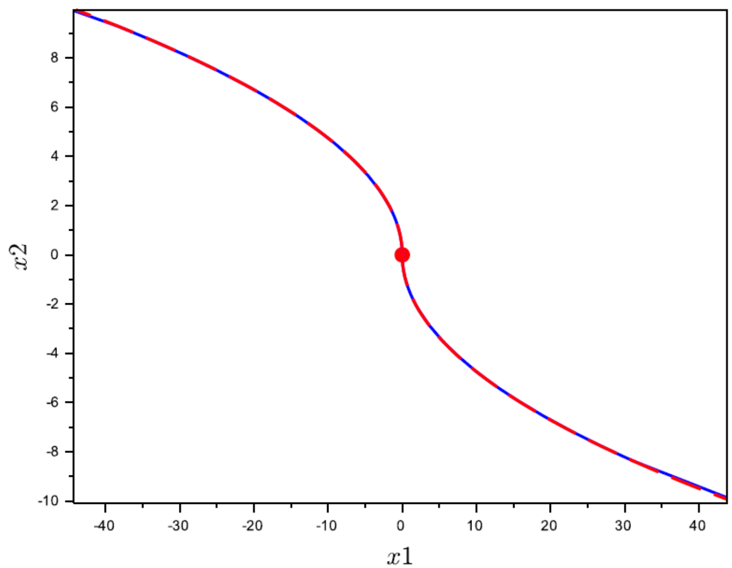



# How to compute the switching curve ?

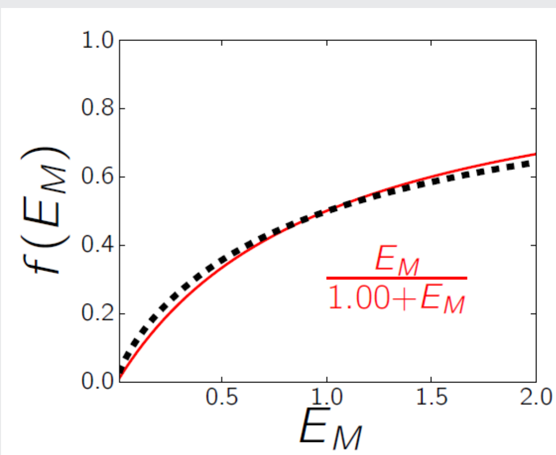
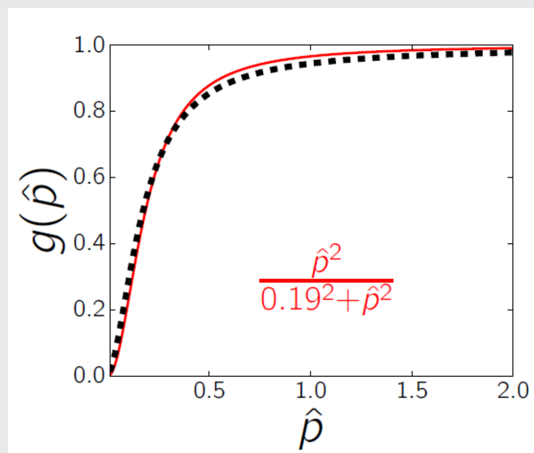
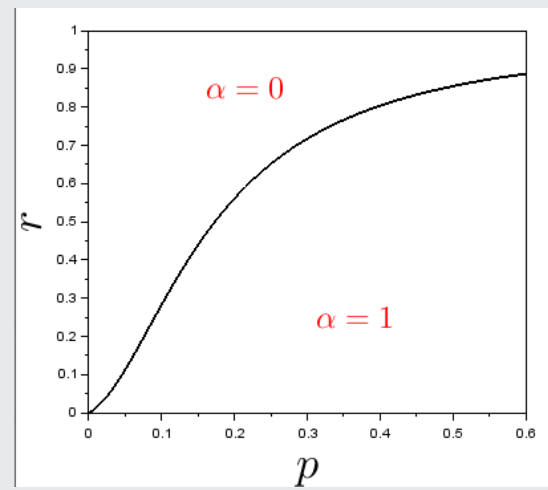
- The tangent of the switching curve at ($p_{opt}$,$r_{opt}$) is vertical.
- Backward integration starting from ($p_{opt}$,$r_{opt}$+ε)
- Validation on Fuller's problem:

$$\text{minimize} \quad \int_0^T x_1^2(t)\,dt \qquad \text{given} \quad \begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = u, \end{cases}$$

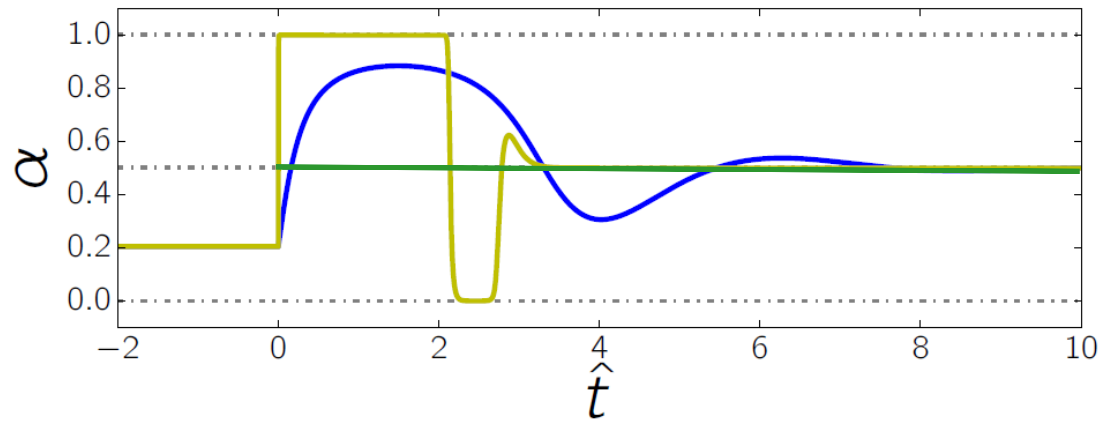


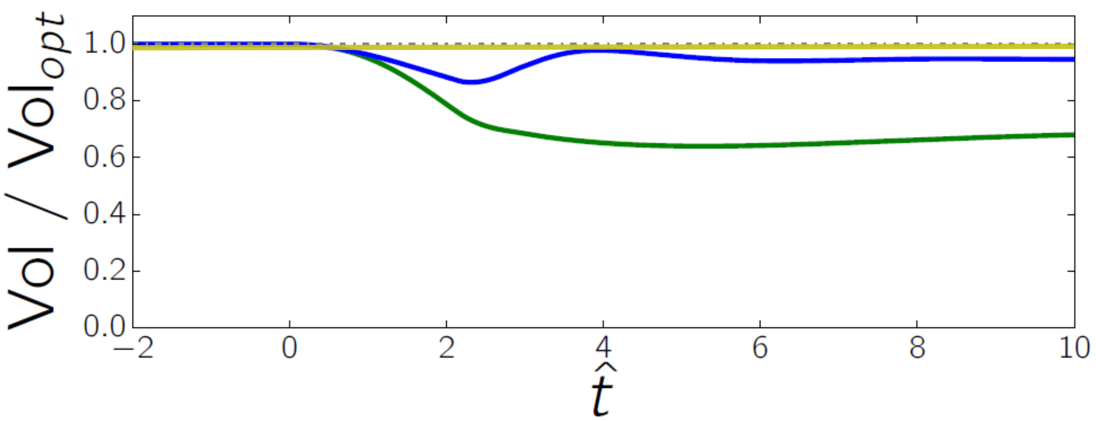



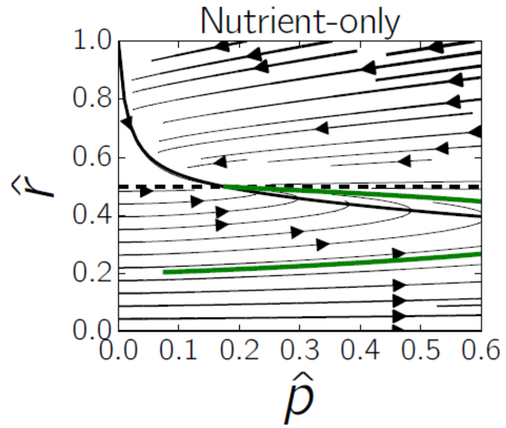
# Simple feedback control strategies

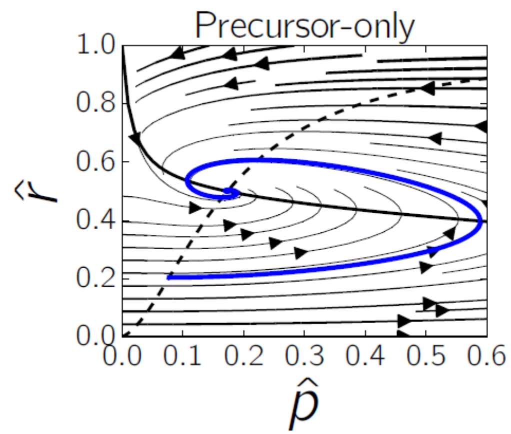
| Substrate | Precursor | Switch |
|---|---|---|
| $\alpha = f(E_M)$ | $\alpha = g(\hat{p})$ | $\alpha = \begin{cases} 0, & \text{if } \hat{r} > g(\hat{p}) \\ 1, & \text{if } \hat{r} < g(\hat{p}) \end{cases}$ |
|  |  |  |

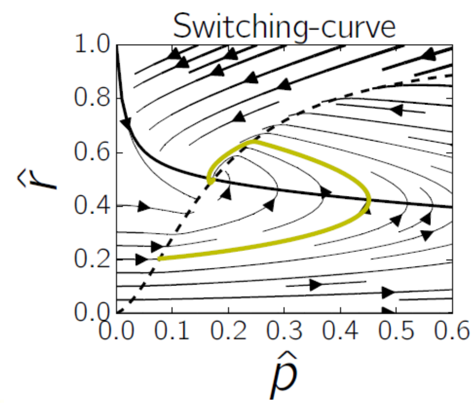
# Comparison of control strategies



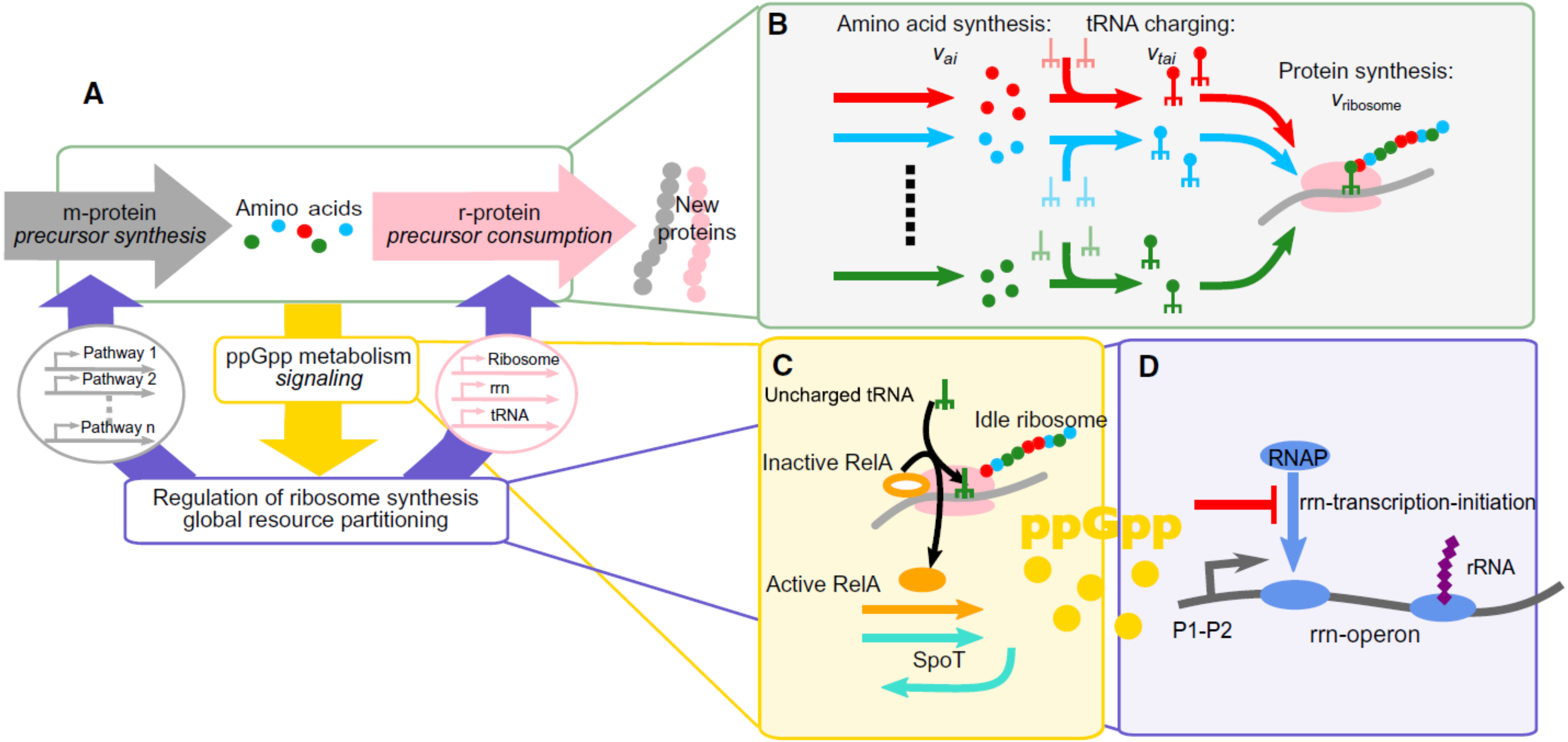



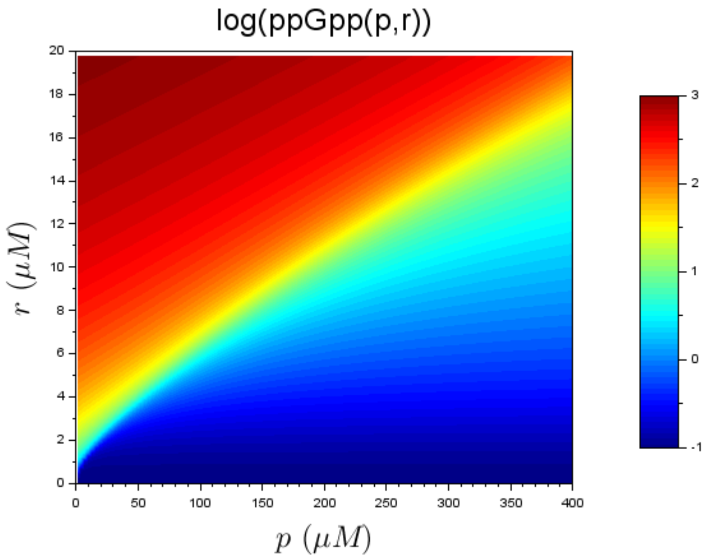
# Biological implementation?

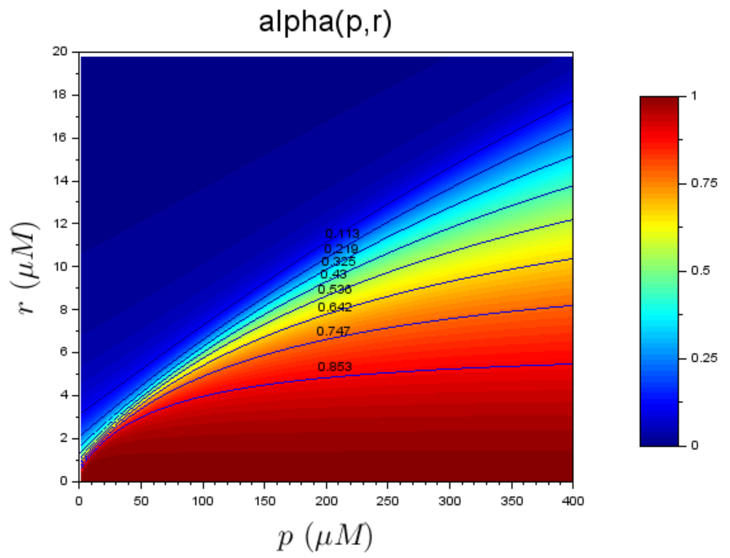
- Regulation of resource allocation via ppGpp (Bosdriesz et al., 2015)



# Biological implementation?

- Regulation of resource allocation via ppGpp (Bosdriesz et al., 2015)
- Quasi steady-state approximation:
  - Fast variables: ppGpp, tRNA



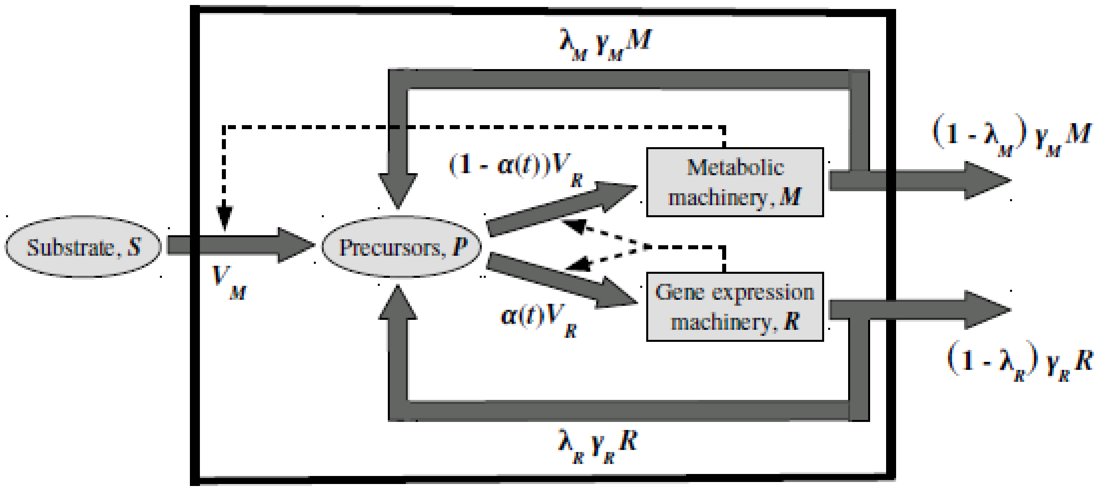



# Conclusions and perspectives

- Study of resource allocation in bacteria from first principles
  - Self-replicator model derived from two macro-reactions and some common assumptions on reaction kinetics
- Optimal strategy: turnpike with chattering
- Near-optimal strategy: switch depending on the imbalance between precursors and ribosomes
- Implementation via ppGpp
- Experimental test of control strategy using fluorescent reporters

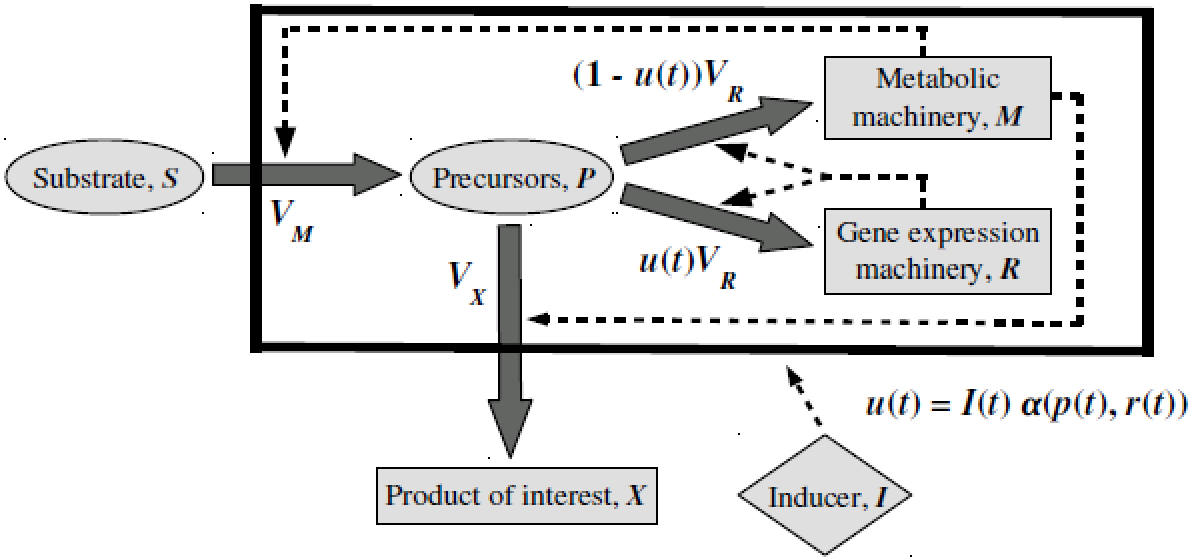
# Conclusions and perspectives

- Some generalizations: degradation and recycling
- Similar results (paper submitted to OCAM)



# Conclusions and perspectives

- Some generalizations: maximum of production with an inducer (Reset)
- Growth or production?
- Limited amount of substrate? ANR Maximic, new…



# Collaborators (RESET)

N. Giordano

H. de Jong

H. Geiselmann

F. Mairet

J.-L. Gouzé

I. Egorov