Workshop "Identification and Control of Biological Interaction Networks"
INRIA Grenoble, 8-Feb-2011

Global optimization approaches for the identification and control of biological systems

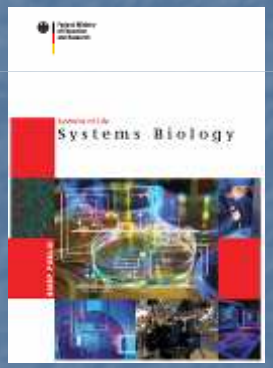
Julio R. Banga

IIM-CSIC, Vigo

julio@iim.csic.es

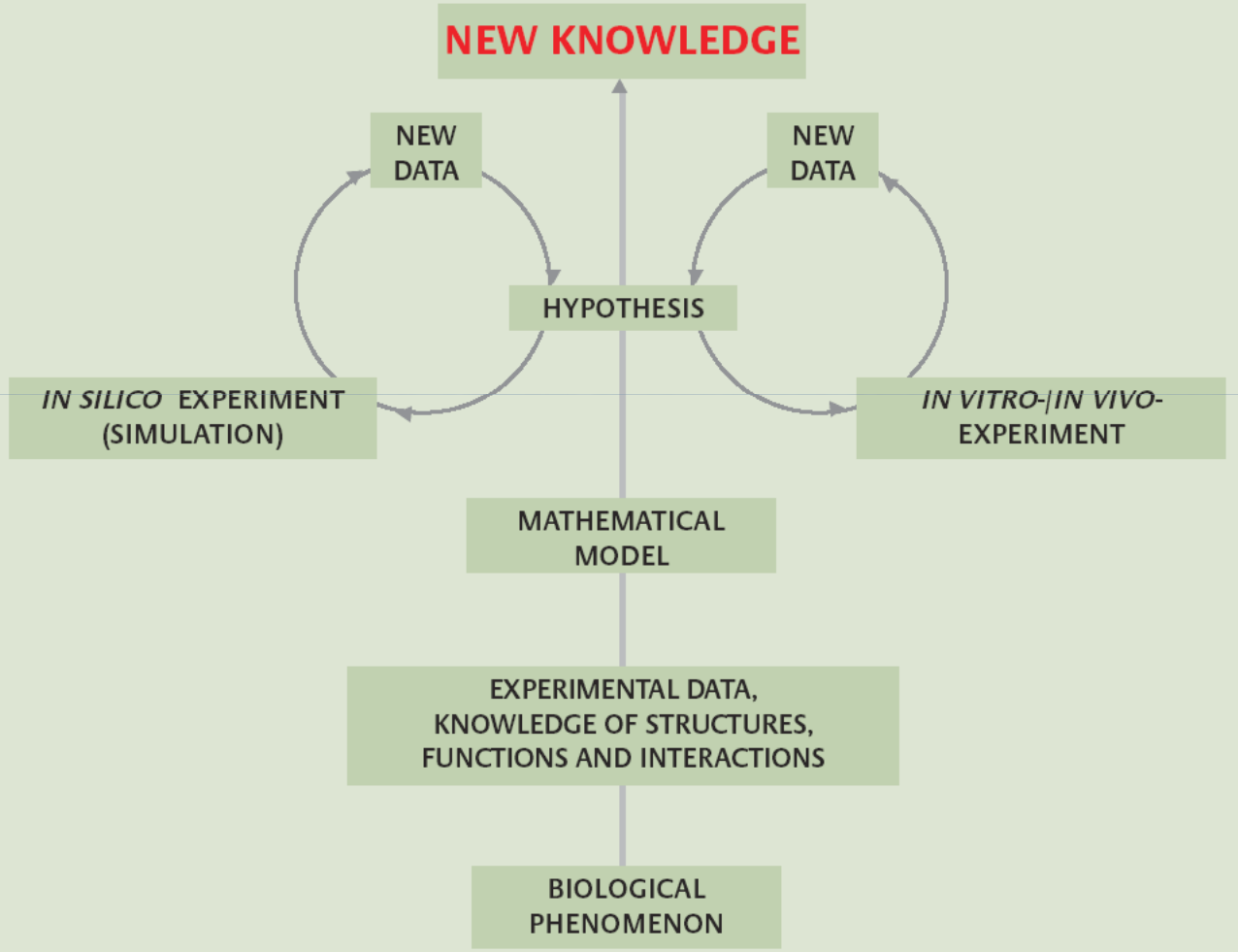
Outline

- Systems biology
- Optimality and/in biology
- Optimization in biochemical pathways
- Some applications
  - Model building / inference
  - Analysis
  - Optimal control
- Conclusions



Federal Ministry
of Education
and Research

The process of knowledge generation in systems biology



CASE WESTERN RESERVE UNIVERSITY EST. 1826

Mike Mesarovic in "*System Theory and Biology*", 1968:

*"The **real advance in the application of systems theory to biology will come about only when the biologists start asking questions which are based on the system-theoretic concepts** rather than using these concepts to represent in still another way the phenomena which are already explained in terms of biophysical or biochemical principles.*

Then we will [...] have [...] a field of Systems Biology "

Key issues in SB

Dynamics, feedback, optimality

Optimization

Optimization

To OPTIMIZE: "*to make as perfect, effective, or functional as possible*" (Webster Dictionary)

Model-based Optimization:
To find the best solution, from the set of all the possible ones, in an efficient and systematic way, using model-based simulation to evaluate candidate solutions

In general, it implies finding the best compromise among several conflicting demands

Structure of an optimization problem

Decision variables

Find ***x***

Objective function

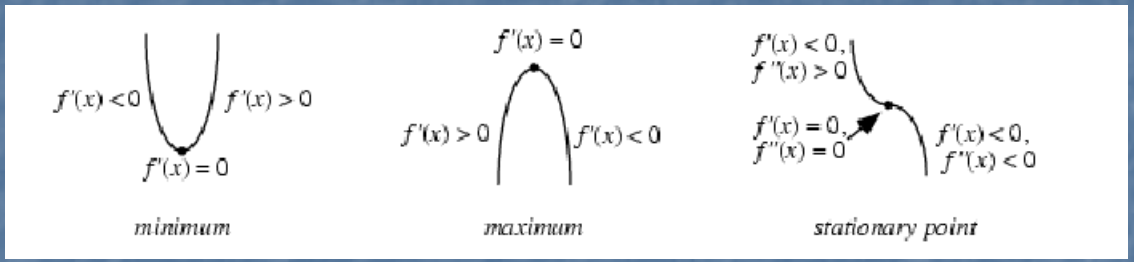
To Minimize (or Maximize) ***f(x)***

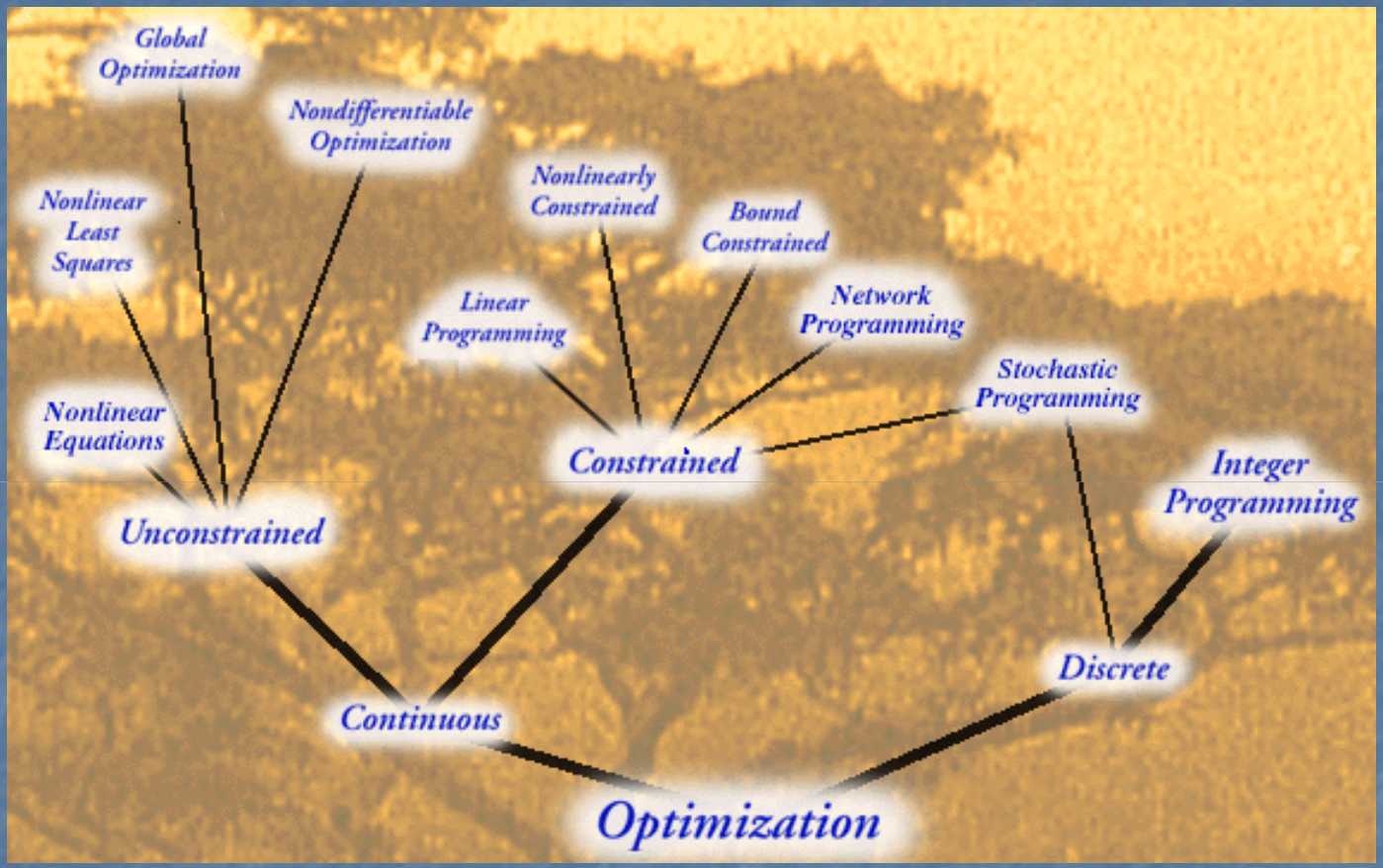
Subject to

Constraints *(requirements)*

- ***Origins: classical theory of calculus***

Fermat, Newton, Euler, Lagrange,...





Modern Optimization Tree

http://www-fp.mcs.anl.gov/otc/Guide/OptWeb/

Optimization in biology

Living organisms have evolved to maximize their chances for survival (Darwin)

The best solution

Optimization: this beguilingly simply idea allows biologists not only to understand current adaptations, but also to predict new designs that may yet evolve.

William J. Sutherland

"If one way be better than another, that you may be sure is nature's way." Aristotle clearly stated the basic premise of optimization in biology, yet it was almost 2,000 years before the power of this idea was appreciated. The essence of optimization is to calculate the most efficient solution to a given problem, and then to test the prediction. The concept has already revolutionized some aspects of biology, but it has the potential for much wider application.

Of course, optimization has long been employed effectively in subjects other than biology. Economists have traditionally calculated the options that result in the greatest profit, and engineers routinely calculate the best design solution, such as the strongest bridge of a given weight.

Darwin's theory of natural selection provided an obvious mechanism for explaining optimization in biology: more efficiently designed individuals will leave more offspring. But it was another century before biologists calculated optimal solutions. David Lack pioneered its use in biology with his concept of the optimal clutch size — the number of eggs that would produce the greatest number of offspring.

The use of optimization has allowed biologists to move from merely describing patterns or mechanisms to being able to predict, from first principles, how organisms should be designed. Optimality models are based on three elements: the optimal. Actually, it is the assumptions of optimality that are tested. The failure to find support for a prediction can be used to determine whether an assumption is wrong. For example, if animals do not select the diet that maximizes energy intake, it may be because they are choosing a diet that optimizes a balance of different components, or that avoids the costs associated with obtaining larger prey. Once such possibilities have been identified, a new theory can be devised and its predictions tested. It has been argued that this process is circular but in practice it is no different from the successive predicting and testing that underlies most science.

A recent example of the insight that optimization can provide concerns the design of mammalian mouths. It is possible to predict, on the basis of efficient food tools. Further calculations give the prediction that the reduction in molars and premolars depends on the cube root of the drop in food toughness. On the basis of these predictions, the changes caused by cooking would have to be vast to match the changes caused by tool use. As predicted, although all teeth have become reduced, the face and incisors have become proportionately smaller. This means the mouth can no longer accommodate the molars, hence the squeezed or missing third molars (wisdom teeth) of many modern humans.

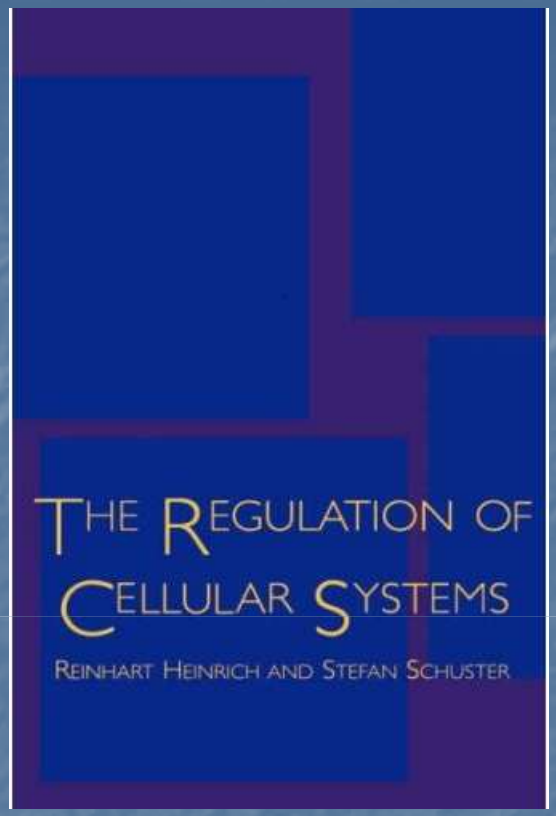
A considerable strength of using optimization is that once we understand why organisms are as they are, then it should be possible to understand how they will respond to new conditions. Optimization can therefore be used to understand behav-

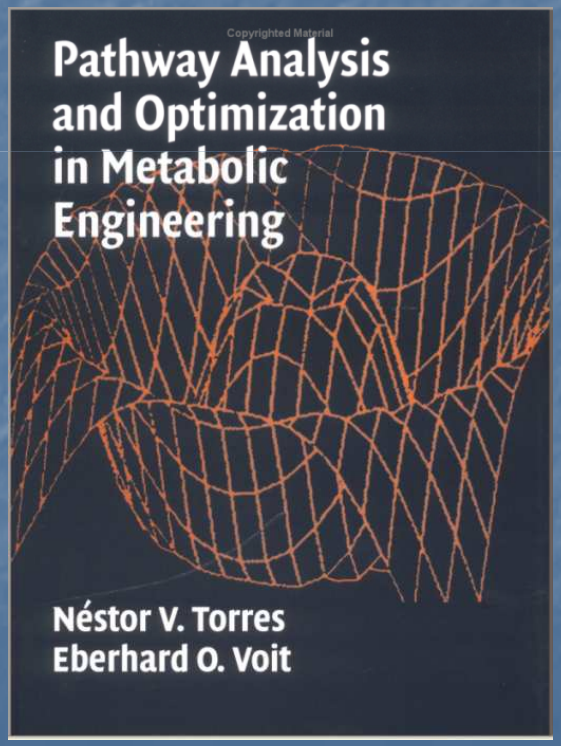
"…in a context of increasing calls for biology to be predictive, optimization is the only approach biology has for making predictions from first principles."

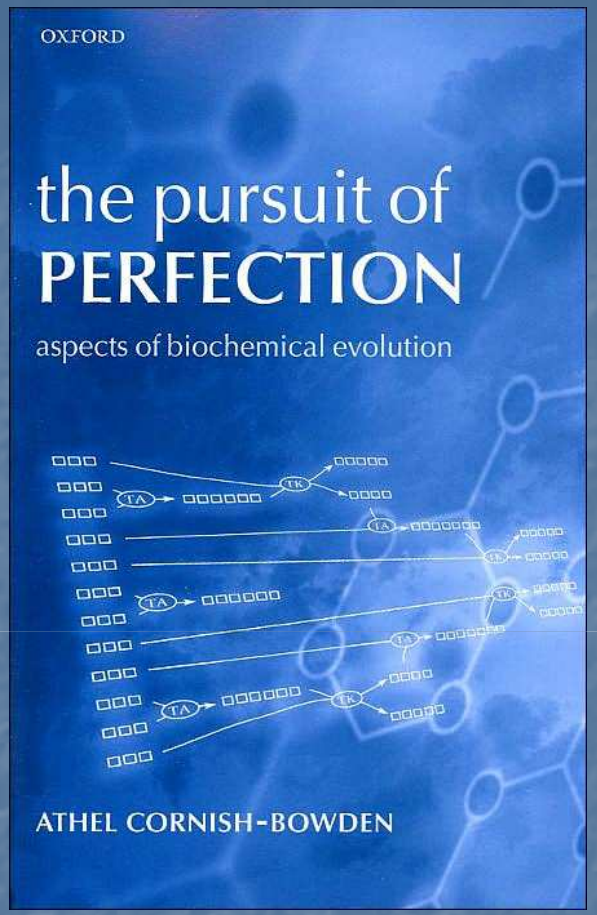
Revealed: optimal-design theory can be used to assess how selective forces have shaped teeth.

Sutherland, W.J. (2005) The best solution. *Nature* 435:569

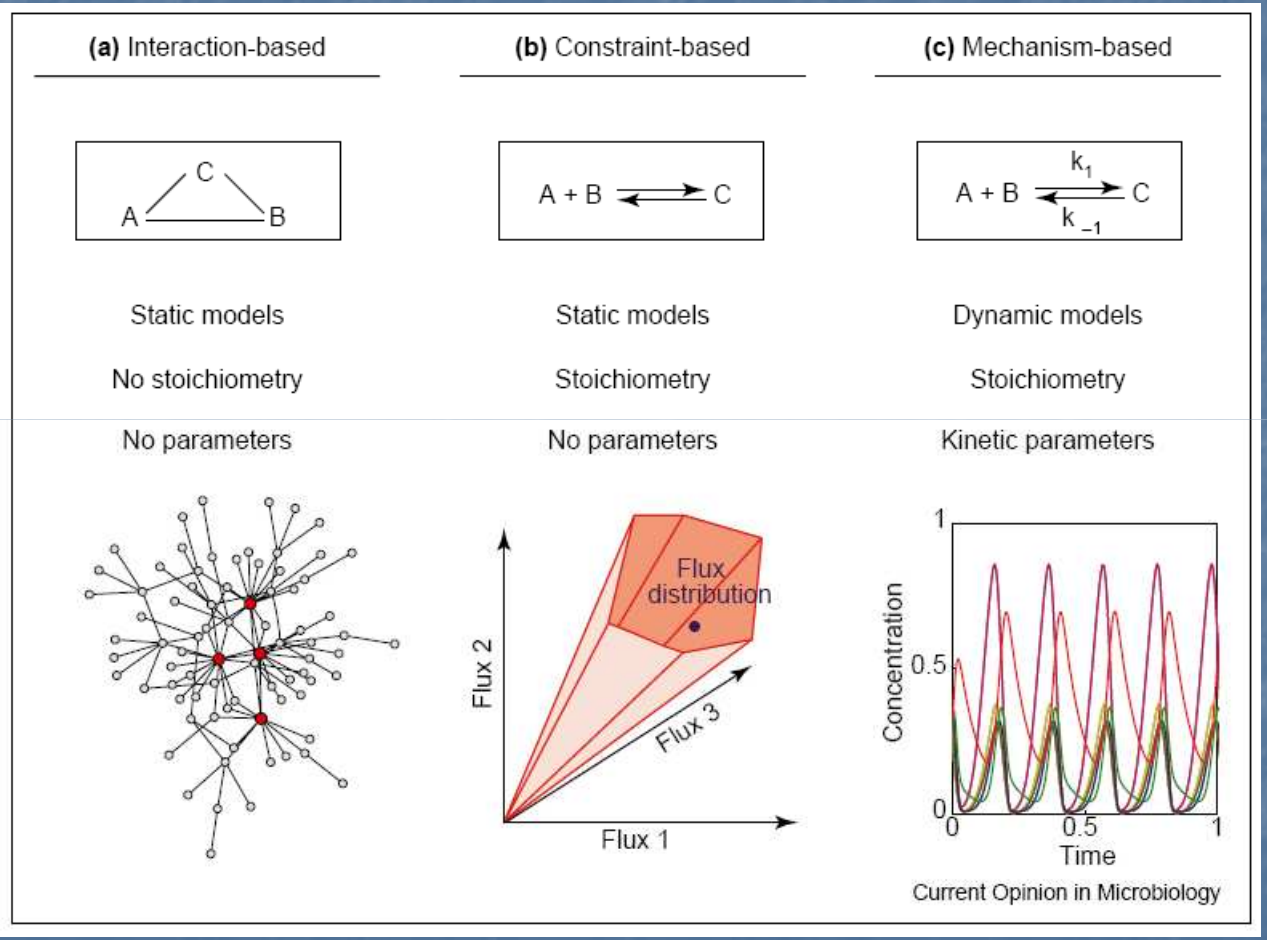
Optimization in biochemical pathways


THE REGULATION OF
CELLULAR SYSTEMS
REINHART HEINRICH AND STEFAN SCHUSTER


Copyrighted Material
Pathway Analysis
and Optimization
in Metabolic
Engineering
Néstor V. Torres
Eberhard O. Voit


OXFORD
the pursuit of
PERFECTION
aspects of biochemical evolution
ATHEL CORNISH-BOWDEN

Models in Systems Biology



Stelling, J. (2004) *Current Opinion in Microbiology, 7*(5), 513-518.

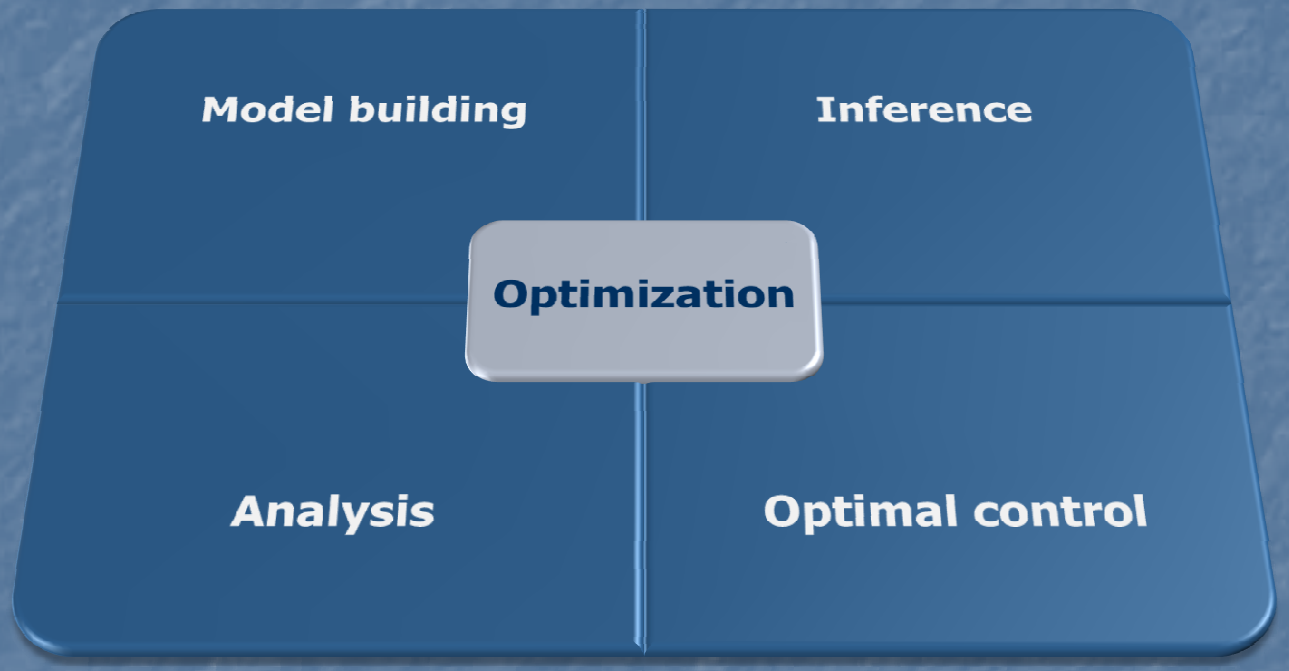
Examples of applications of optimization in systems biology

Problem type or application	Description	Examples with references
Linear programming (LP)	linear objective and constraints	maximal possible yield of a fermentation [83]; metabolic flux balancing [18,83]; review of flux balance analysis in [30]; use of LP with genome scale models reviewed in [27]; inference of regulatory networks [40,42]
Nonlinear programming (NLP)	some of the constraints or the objective function are nonlinear	applications to metabolic engineering and parameter estimation in pathways [69]; substrate metabolism in cardiomyocytes using ^{13}C data [84]; analysis of energy metabolism [85]
Semidefinite programming (SDP)	problems over symmetric positive semidefinite matrix variables with linear cost function and linear constraints	partitioning the parameter space of a model into feasible and infeasible regions [86]
Bilevel optimization (BLO)	objective subject to constraints which arise from solving an inner optimization problem	framework for identifying gene knockout strategies [87]; optimization of metabolic pathways under stability considerations [88]; optimal profiles of genetic alterations in metabolic engineering [89]
Mixed integer linear programming (MILP)	linear problem with both discrete and continuous decision variables	finding all alternate optima in metabolic networks [90,91]; optimal intervention strategies for designing strains with enhanced capabilities [91]; framework for finding biological network topologies [47]; inferring gene regulatory networks [41]
Mixed integer nonlinear programming (MINLP)	nonlinear problem with both discrete and continuous decision variables	analysis and design of metabolic reaction networks and their regulatory architecture [92,93]; inference of regulatory interactions using time-course DNA microarray expression data [45]
Parameter estimation	model calibration minimizing differences between predicted and experimental values	tutorial focused in systems biology [53]; parameter estimation using global and hybrid methods [52,54,55,59,70]; parameter estimation in stochastic models [58]
Dynamic optimization (DO)	Optimization with differential equations as constraints (and possible time-dependent decision variables)	discovery of biological network design strategies [94]; dynamic flux balance analysis [29]; optimal control for modification of self-organized dynamics [95]; optimal experimental design [66]
Mixed-integer dynamic optimization (MIDO)	Optimization with differential equations as constraints and both discrete and continuous decision variables (possibly time-dependent)	computational design of genetic circuits [76]

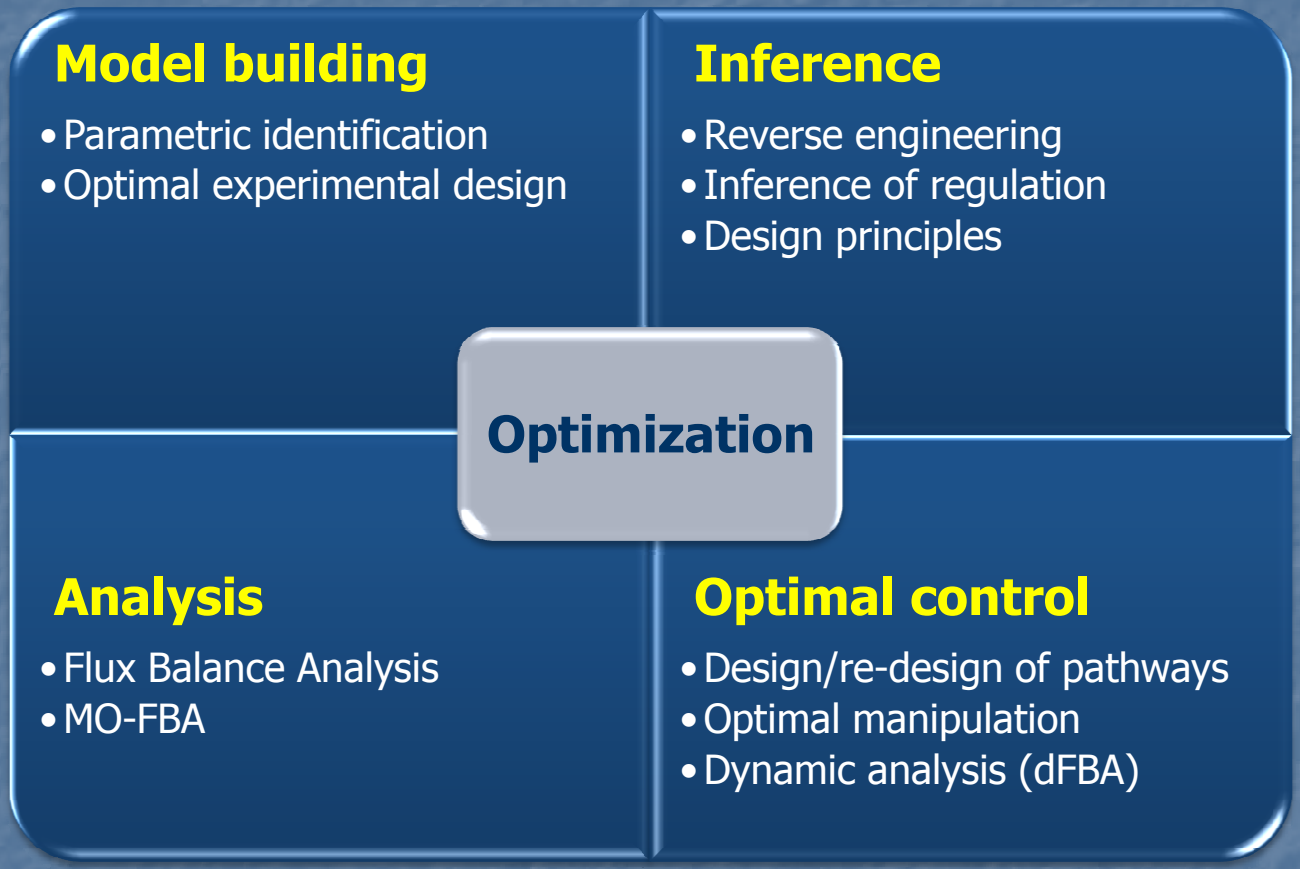
Banga, J.R. (2008). Optimization in computational systems biology. *BMC Systems Biology 2:47*.

- ***Linear Programming (LP)***
- ***Non-linear Programming (NLP)***
- ***Mixed-integer NLP (MINLP)***
- ***Dynamic Optimization (optimal control)***
- ***Mixed Integer Dynamic Optimization***

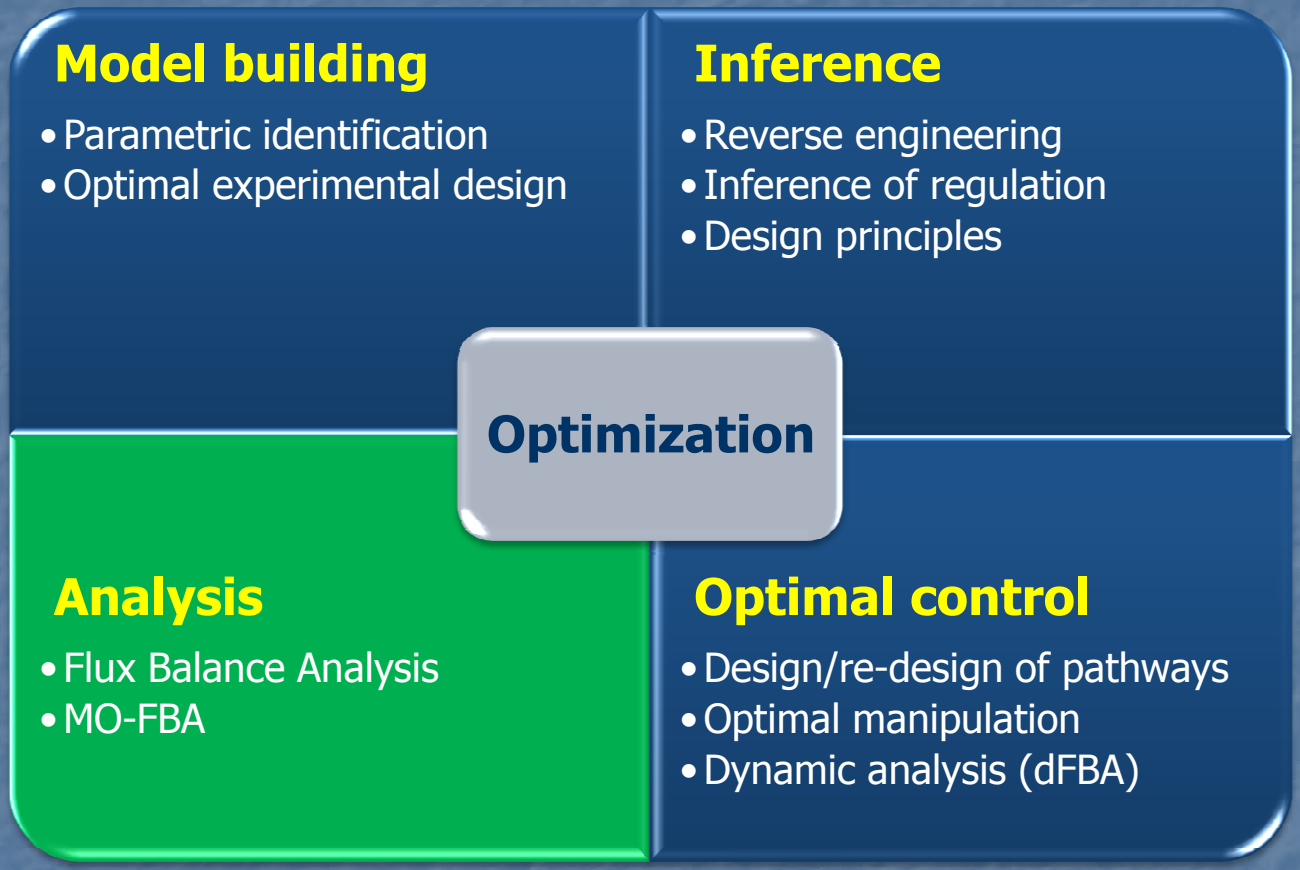
Optimization in systems biology: types of applications



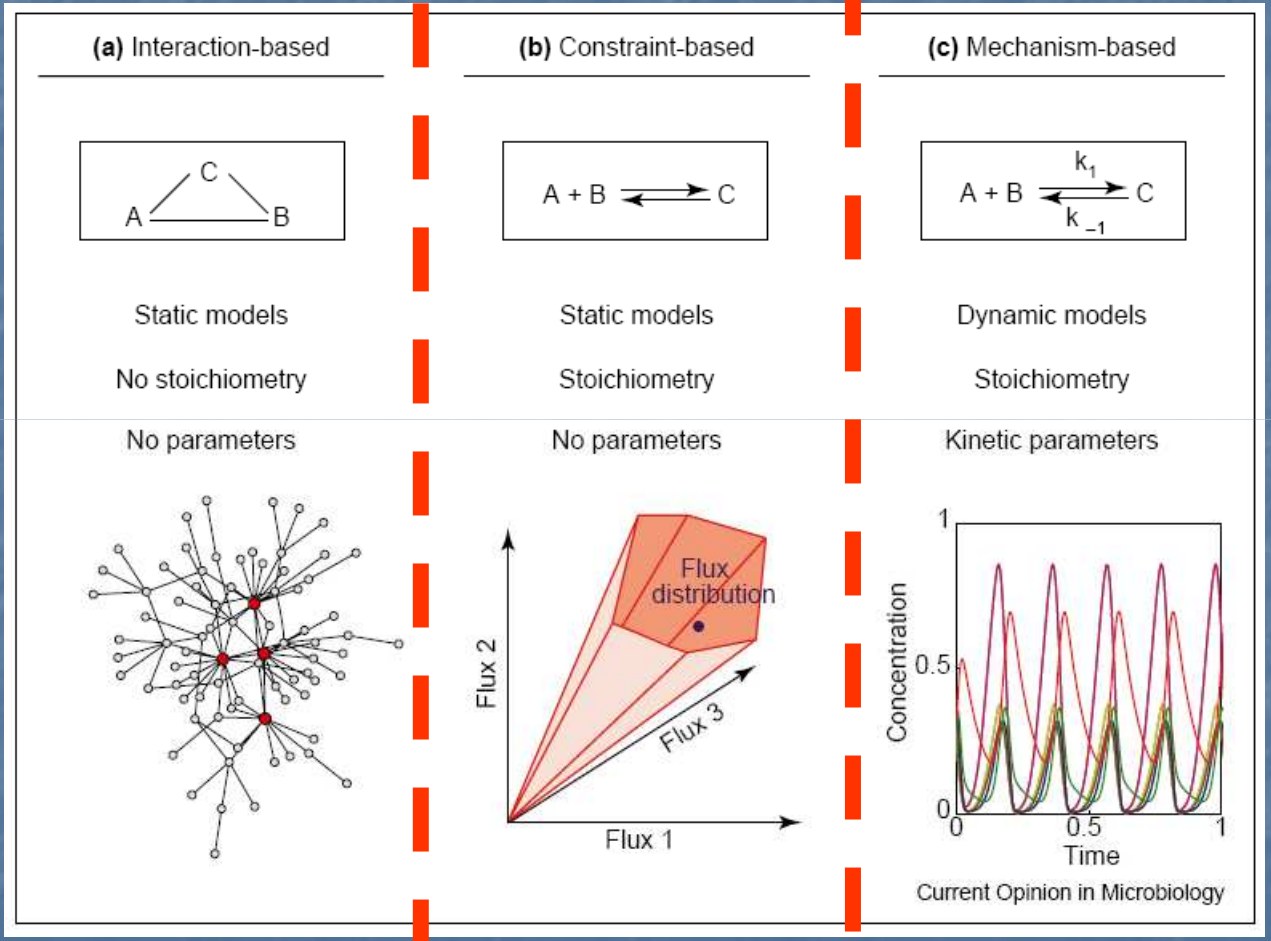
Optimization in systems biology: examples of applications



Optimization in systems biology: examples of applications



Models in Systems Biology



Stelling, J. (2004). Mathematical models in microbial systems biology. *Current Opinion in Microbiology, 7*(5), 513-518.

Convex Optimization

- ***Linear Programming (LP)***

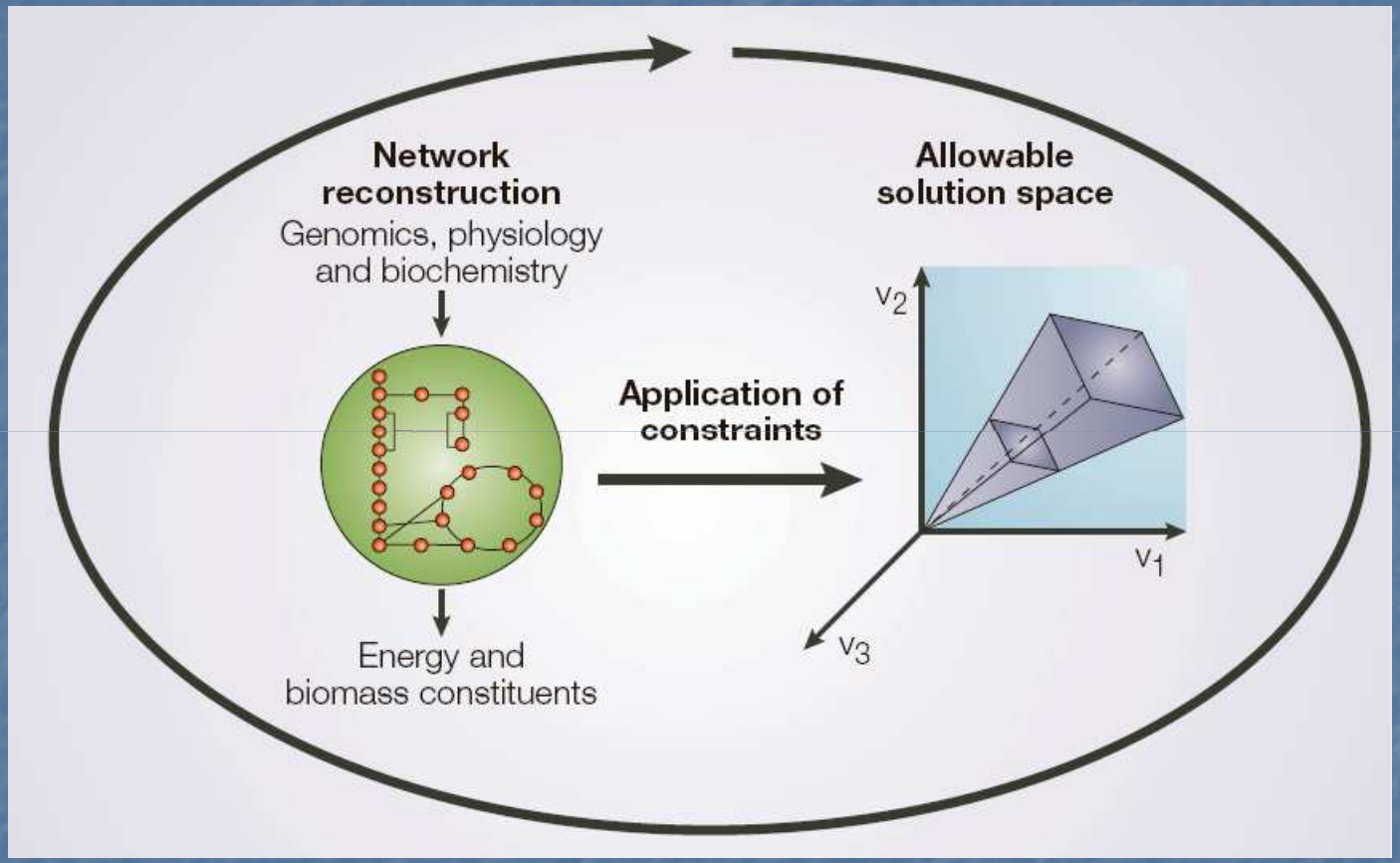
Find x to

minimize $f(x)=c*x$

subject to $Ax<=b$

- ***where x are decision variables,***
- ***f is the objective function***
- ***LP: linear f(x) & linear constraints***

Flux Balance Analysis



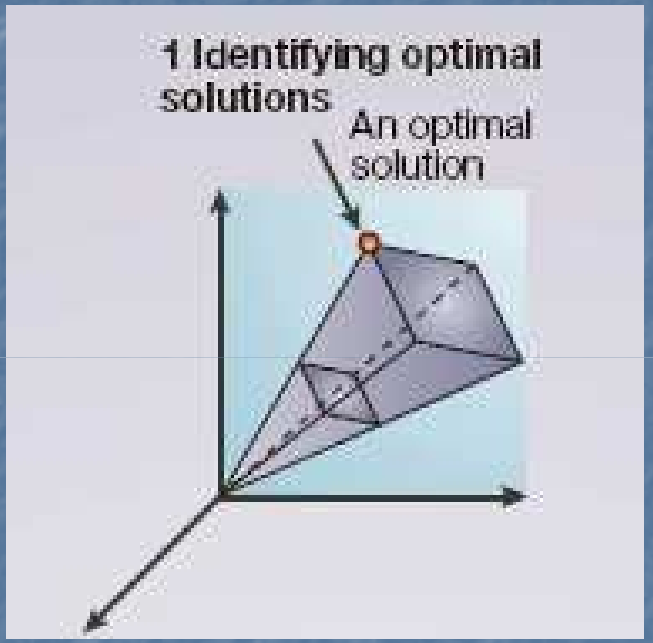
Price, N.D., Reed, J.L. and Palsson, B.O., Genome-scale models of microbial cells:evaluating the consequences of constraints, *Nature Reviews Microbiology*, 2:886-897(2004).

Flux Balance Analysis

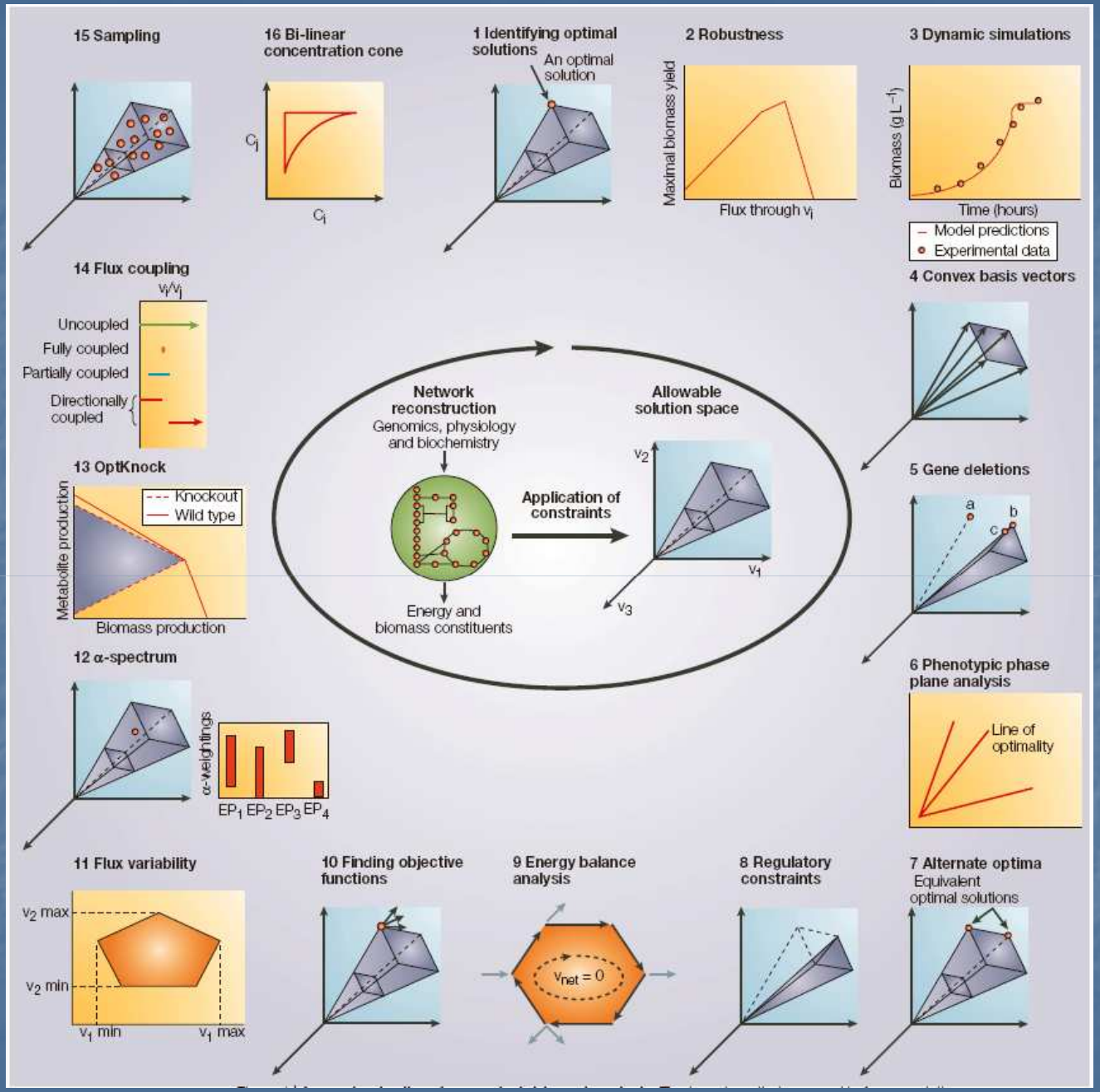
If we assume the organism is optimizing a certain objective function (e.g., maximizing biomass), we get an optimal solution by solving a

LINEAR PROGRAMMING (LP) problem

And we can do this for genome-scale models!



Price, N.D., Reed, J.L. and Palsson, B.O., Genome-scale models of microbial cells:evaluating the consequences of constraints, *Nature Reviews Microbiology*, 2:886-897(2004).



Price, N.D., Reed, J.L. and Palsson, B.O., Genome-scale models of microbial cells:evaluating the consequences of constraints, *Nature Reviews Microbiology*, 2:886-897(2004).

Molecular Systems Biology 3; Article number 119; doi:10.1038/msb4100162
Citation: *Molecular Systems Biology* 3:119
© 2007 EMBO and Nature Publishing Group All rights reserved 1744-4292/07
www.molecularsystemsbiology.com

molecular systems biology

Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*

Robert Schuetz[1,2,3], Lars Kuepfer[1,3,4] and Uwe Sauer[1,2,*]

[1] Institute of Molecular Systems Biology, ETH Zurich, Switzerland and [2] Life Science Zurich PhD program on Systems Physiology and Metabolic Diseases, Zurich, Switzerland
[3] These authors contributed equally to this work
[4] Present address: Bayer Technology Services, Leverkusen 51368, Germany
* Corresponding author. Institute of Molecular Systems Biology, ETH Zurich, Wolfgang-Pauli-Strasse 16, Zurich 8093, Switzerland. Tel.: +41 44 633 3672; Fax: +41 44 633 1051; E-mail: sauer@imsb.biol.ethz.ch

Received 10.1.07; accepted 2.5.07

To which extent can optimality principles describe the operation of metabolic networks? By explicitly considering experimental errors and *in silico* alternate optima in flux balance analysis, we systematically evaluate the capacity of 11 objective functions combined with eight adjustable constraints to predict ^{13}C-determined *in vivo* fluxes in *Escherichia coli* under six environmental conditions. While no single objective describes the flux states under all conditions, we identified two sets of objectives for biologically meaningful predictions without the need for further, potentially artificial constraints. Unlimited growth on glucose in oxygen or nitrate respiring batch cultures is best described by nonlinear maximization of the ATP yield per flux unit. Under nutrient scarcity in continuous cultures, in contrast, linear maximization of the overall ATP or biomass yields achieved the highest predictive accuracy. Since these particular objectives predict the system behavior without preconditioning of the network structure, the identified optimality principles reflect, to some extent, the evolutionary selection of metabolic network regulation that realizes the various flux states.
Molecular Systems Biology 10 July 2007; doi:10.1038/msb4100162
Subject Categories: metabolic and regulatory networks; simulation and data analysis
Keywords: ^{13}C-flux; evolution; flux balance analysis; metabolic network; network optimality

Multi-Objective Flux Balance Analysis (MOFBA)

Motivation for MOFBA

- **Flux Balance Analysis (FBA):** determination of flux distribution by optimizing a single (usually **linear)** objective function (typically, growth or biomass yield).
- Other objectives (**linear as well as non-linear**) are possible, often conflicting each other, e.g. maximization of ATP, minimization of overall intracellular flux, etc.
- **Main Idea:** fluxes are distributed to optimize **simultaneously** two or more biological objective functions, i.e. **Multi-Objective FBA**
- **Objective:** to find a set of optimal trade-offs between different criteria (**Pareto Optimal-Set**)

Multi-Objective Flux Balance Analysis (MOFBA)

The aim of MOFBA is to find the values of the fluxes **v** which maximize/minimize simultaneously a set of *p* objective functions subject to a number of equality and inequality constraints (mass-balance equations, bounds on fluxes, …).

$$\min/\max_{\mathbf{v}} \left[Z_1(\mathbf{v}), Z_2(\mathbf{v}), ..., Z_p(\mathbf{v}) \right]^{\mathrm{T}}$$

subject to:

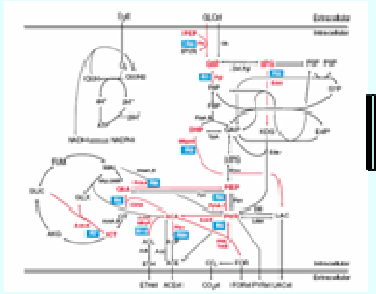
$$\mathbf{S}\mathbf{v} = \mathbf{0}$$

$$\mathbf{v}^L \leq \mathbf{v} \leq \mathbf{v}^U$$

Z_i = objective function i
S = stoichiometric matrix (m x n)
m = number of metabolites
n = number of reactions (fluxes)
$\mathbf{v}^L$, $\mathbf{v}^U$ = lower and upper bounds

Additional constraints can be imposed to reflect biological knowledge and other physiological restrictions

Multi-Objective Flux Balance Analysis (MOFBA)



- **Stoichiometric Matrix S**
- Define **objectives** and **constraints**

Optimize each objective separately ⇨Z_1*; Z_2*; … (similar to traditional FBA)

Transform the original MOP into standard single-objective optimization problems

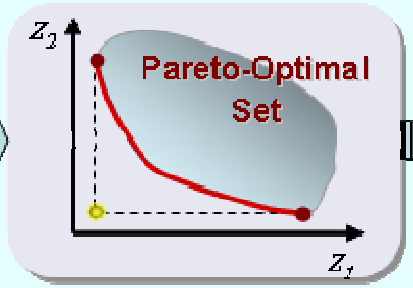
Solve the resulting set of LPs, NLPs, MILPs or MINLPs

Optimization Solvers (single-objective)

- **glpk** (LPs, MILPs)
- **GLOBALm** (NLPs)
- **MITS** (MINLPs)

Multi-Objective Methods

- Normal Boundary Intersection (NBI)
- Weighted Sum Approach
- ε-Constraint
- Min-Max Formulation
- …



Analysis of solutions

Multi-Objective Optimization of Biological Networks for Prediction of Intracellular Fluxes

José-Oscar H. Sendín, Antonio A. Alonso, and Julio R. Banga

Process Engineering Group
Instituto de Investigaciones Marinas (CSIC)
C/ Eduardo Cabello 6, 36208, Vigo, Spain
osendin@iim.csic.es, antonio@iim.csic.es, julio@iim.csic.es

Abstract. In this contribution, we face the problem of predicting intracellular fluxes using a multi-criteria optimization approach, i.e. the simultaneous optimization of two or more cellular functions. Based on Flux Balance Analysis, we calculate the Pareto set of optimal flux distributions in *E. coli* for three objectives: maximization of biomass and ATP, and minimization of intracellular fluxes. These solutions are able to predict flux distributions for different environmental conditions without requiring specific constraints, and improve previous published results. We thus illustrate the usefulness of multi-objective optimization for a better understanding of complex biological networks.

Keywords: Multi-objective optimization, Pareto front, Flux Balance Analysis.

1 Introduction

Intracellular fluxes in biochemical networks can be calculated *in silico* under the assumption that cellular systems operate in an optimal way with respect to a certain biological objective. Network capabilities and flux distributions have thus been predicted by using, for example, Metabolic Flux Balance Analysis (FBA), the fundamentals of which can be found in e.g. (Varma and Palsson 1994). FBA only requires the stoichiometric model of the network, but since the linear system of mass balance equations at steady-state is generally under-determined, appropriate cellular functions (objectives) must be defined, as well as other possible additional constraints, to find a unique solution. Successful applications of FBA include the prediction of *E. coli* metabolic capabilities (Edwards et al. 2001) and the genome-scale reconstruction of the metabolic network in *S. cerevisiae* (Forster et al. 2003).

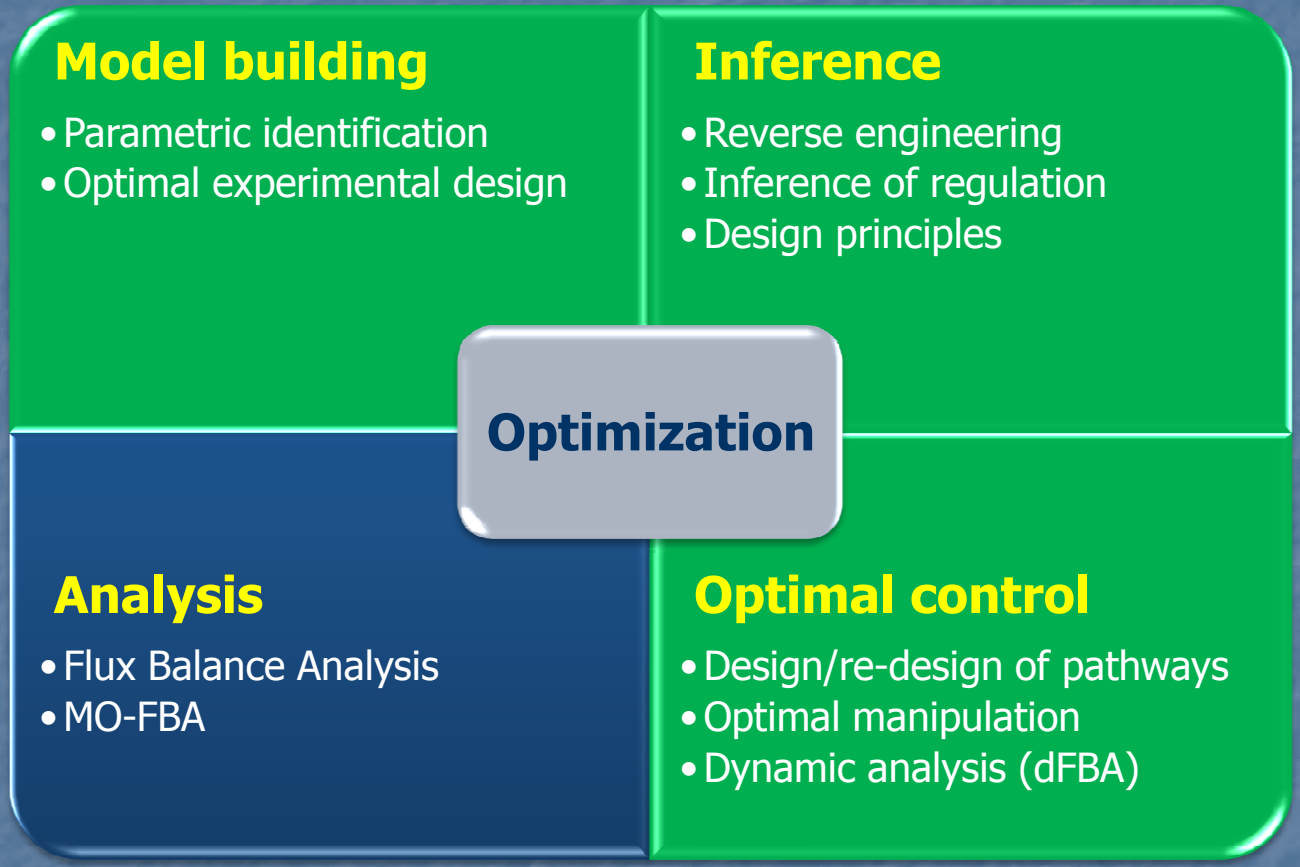
In this context, a particularly interesting question which have been addressed recently in detail (Schuetz et al. 2007; Nielsen 2007) concerns the principles behind the optimal biochemical network operation, i.e.: "which are the criteria being optimized in these systems?" By far, the most common objective considered is the maximization of growth (or biomass yield), although other criteria, such as maximization of ATP yield (van Gulik and Heijnen 1995) or minimization of the overall intracellular flux (Bonarios et al. 1996), have been proposed for different systems and conditions.

Since neither we nor nature have a single goal, a more desirable and realistic approach is to consider the simultaneous optimization of two or more criteria, often conflicting. As a consequence, the solution will not be unique but instead this strategy

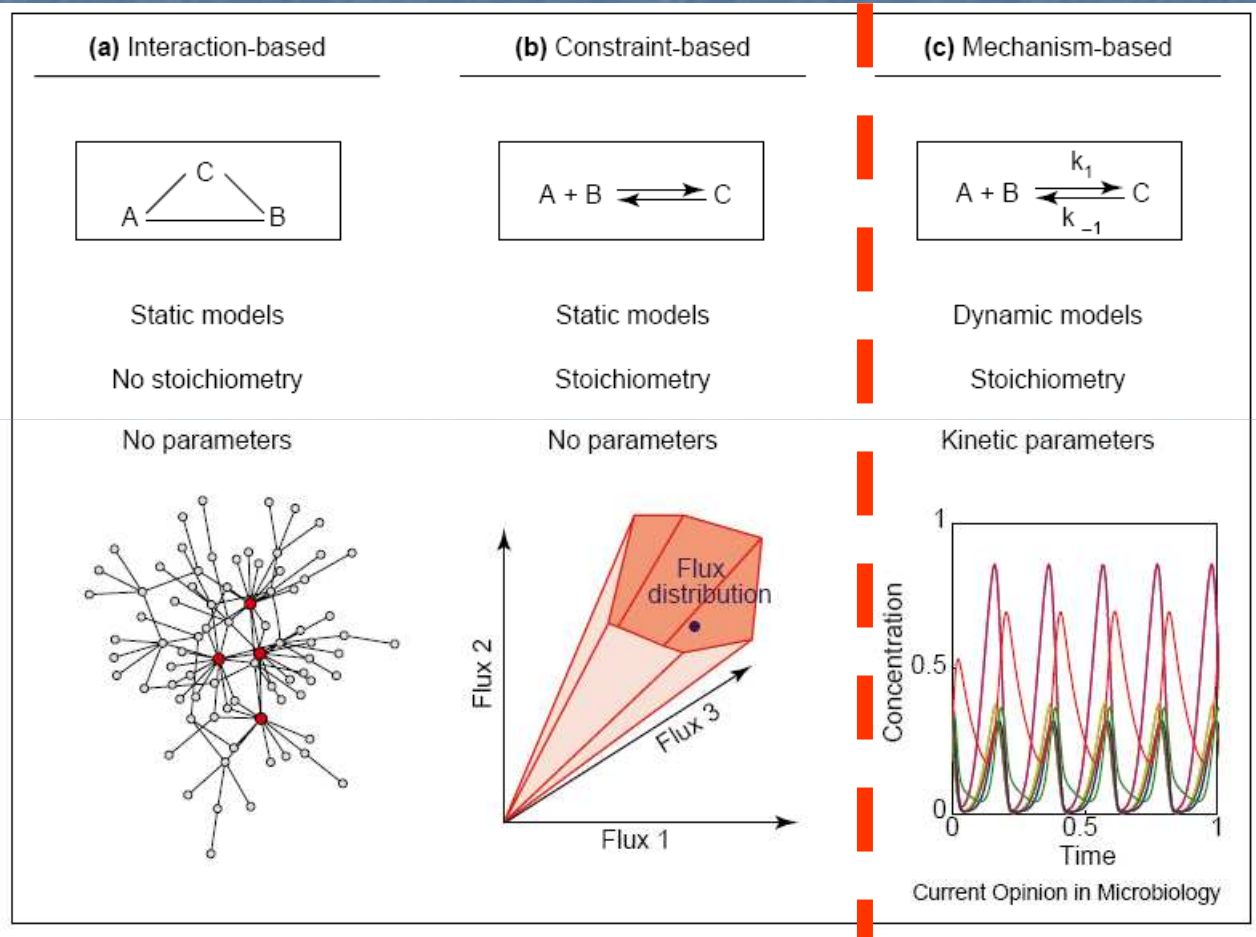
J.M. Corchado (Eds.): IWPACBB 2008, ASC 49, pp. 197–205, 2009.
springerlink.com © Springer-Verlag Berlin Heidelberg 2009

Sendín, J.O.H., Alonso, A. A., & Banga, J. R. (2009). Multi-objective optimization of biological networks for prediction of intracellular fluxes. Advances in Soft Computing 49, pp. 197-205

Optimization in systems biology: examples of applications

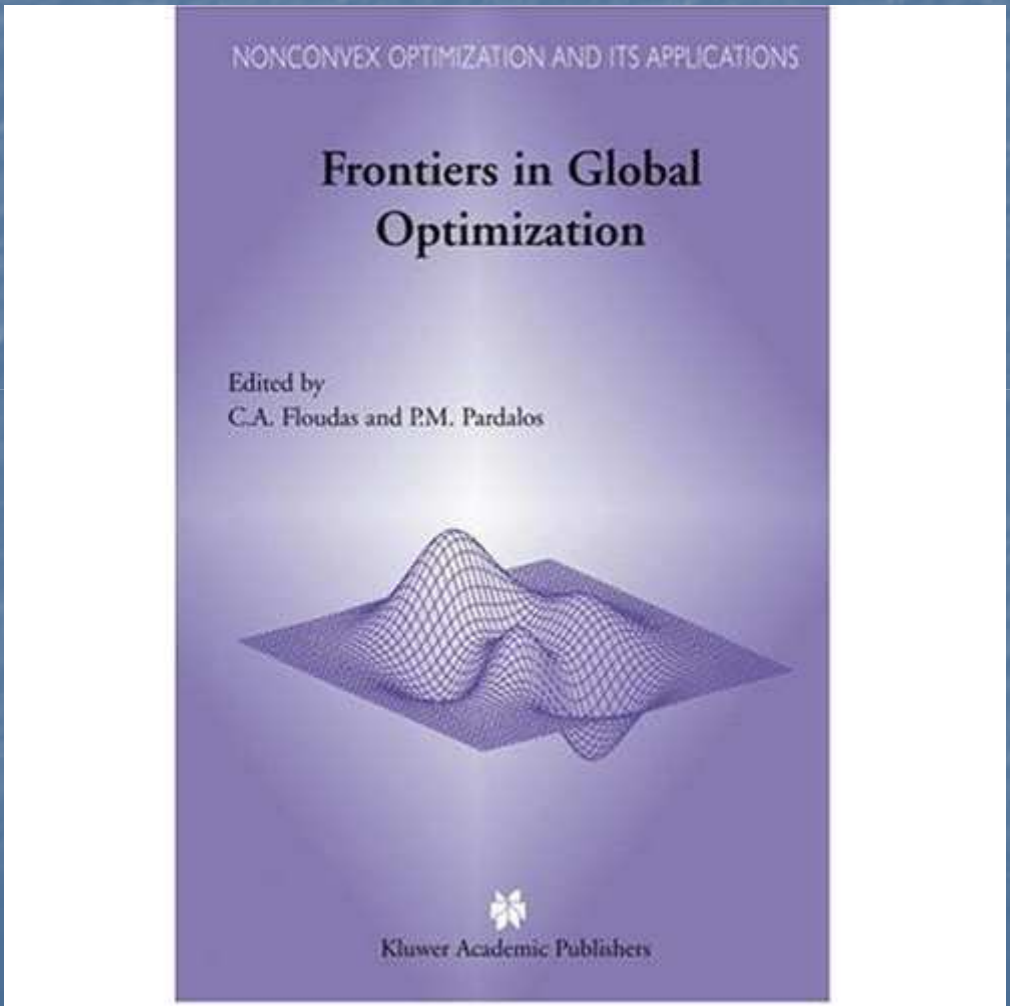


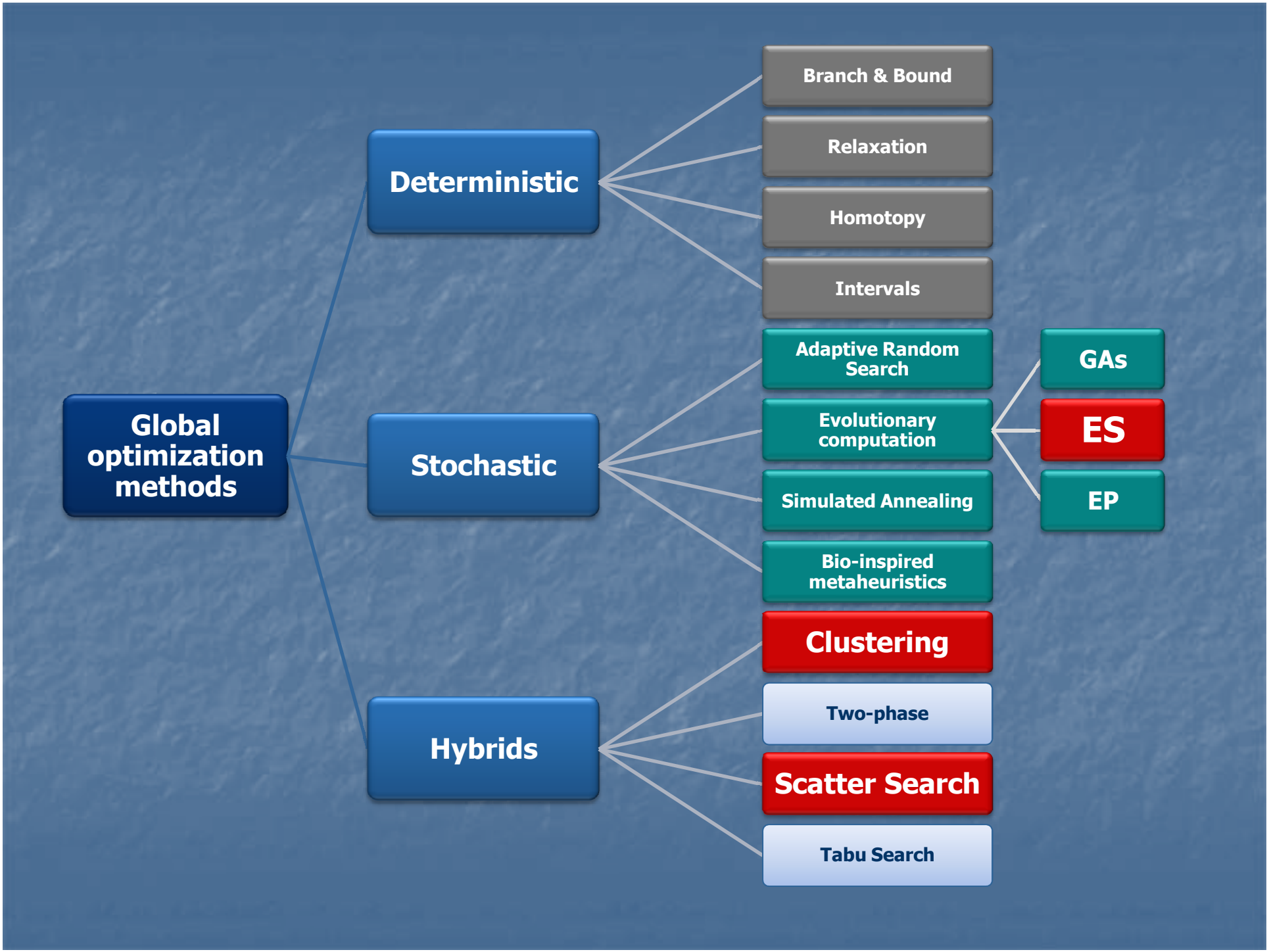
Models in Systems Biology



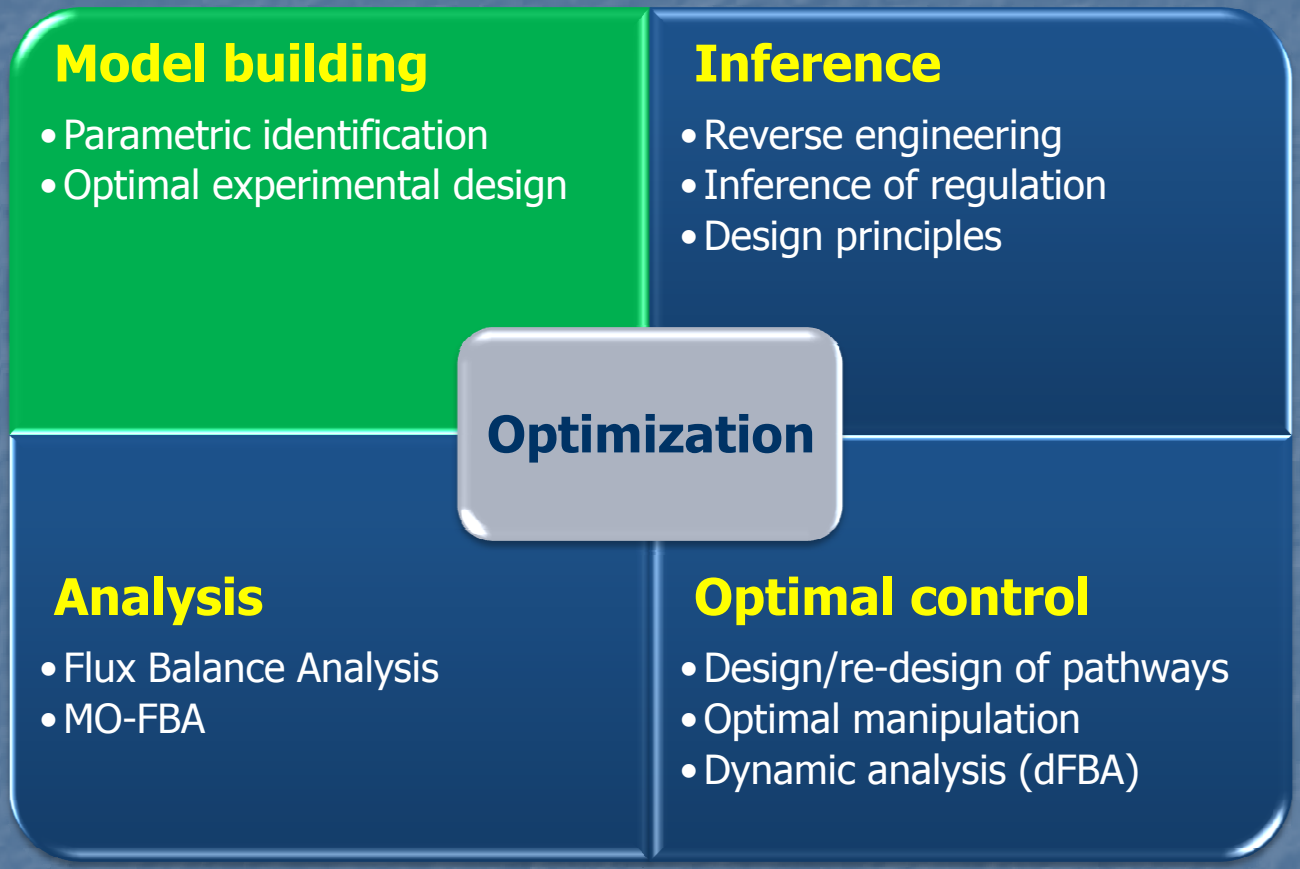
Stelling, J. (2004). Mathematical models in microbial systems biology. *Current Opinion in Microbiology, 7*(5), 513-518.

Non-convex Optimization



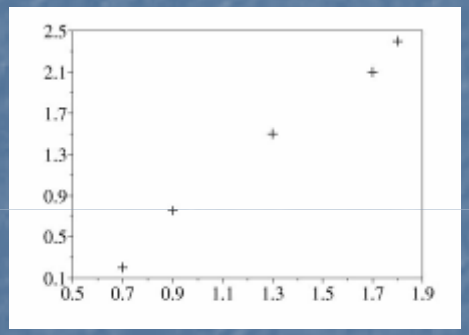

Global optimization methods
Deterministic
Stochastic
Hybrids
Branch & Bound
Relaxation
Homotopy
Intervals
Adaptive Random Search
Evolutionary computation
Simulated Annealing
Bio-inspired metaheuristics
GAs
ES
EP
Clustering
Two-phase
Scatter Search
Tabu Search

Optimization in systems biology: examples of applications



Model building

Experiment



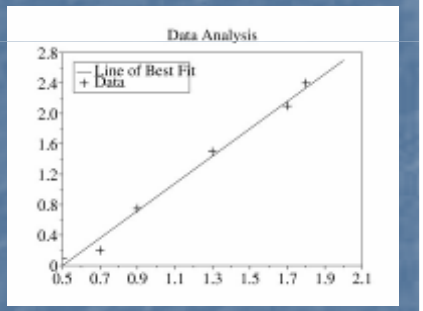
Data

$$\dot{x}(p,t) = f[x(p,t), u(t), p],\ x(0) = x_0,$$

$$y(p,t) = g[x(p,t), u(p,t), p]$$

Model

Solver



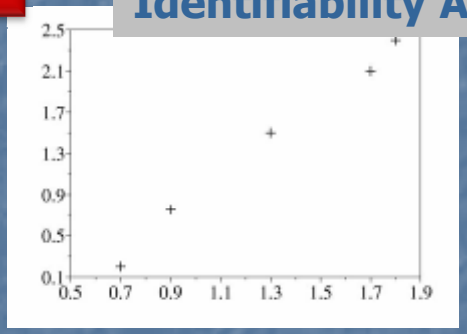
Fitted Model

Model building

Optimal Experimental Design

Experiment

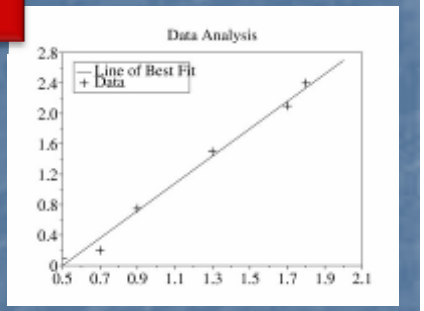
Identifiability Analysis



Data

Parameter Estimation

Solver

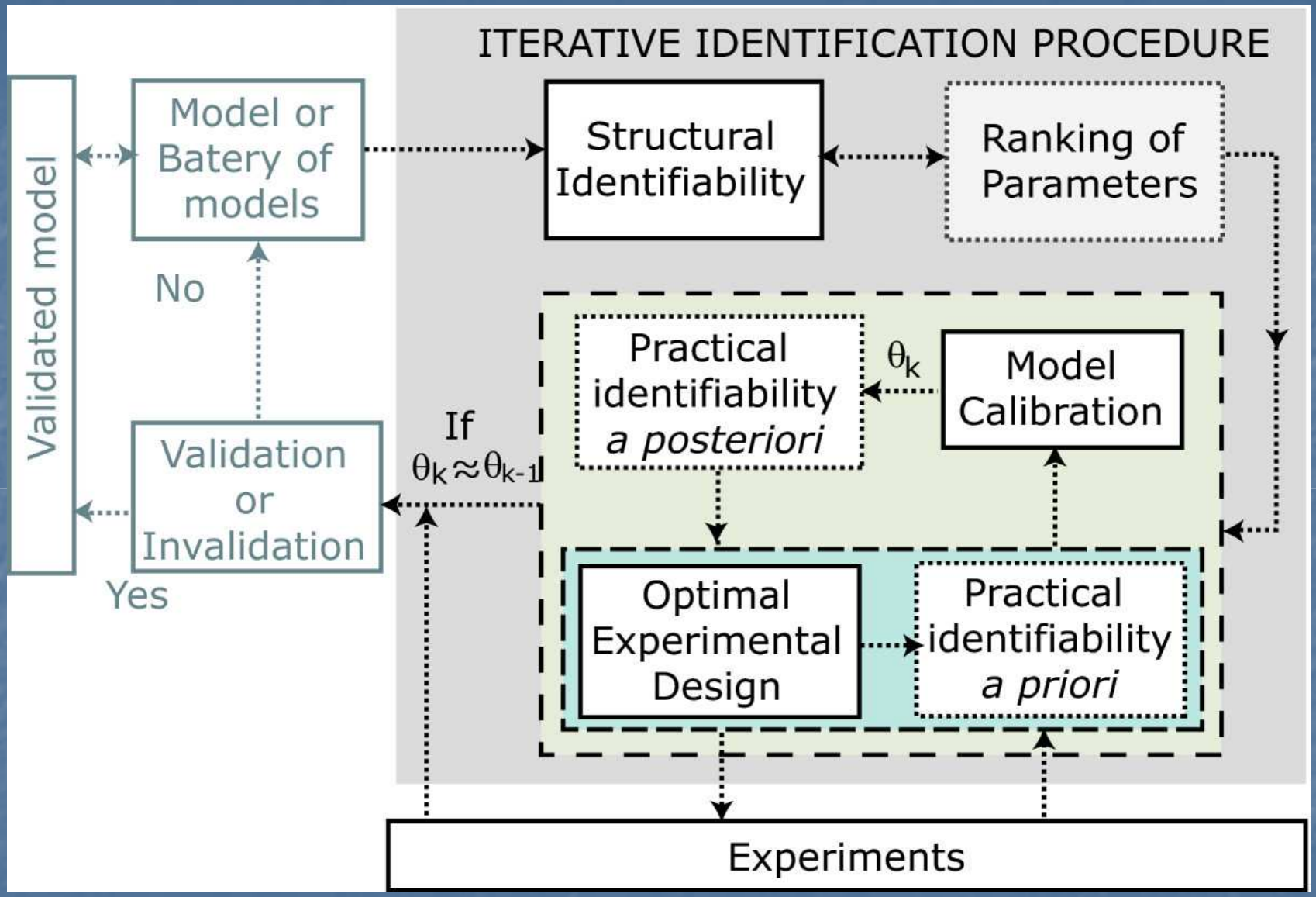


Fitted Model

Identifiability Analysis

$$y(p,t) = g[x(p,t), u(p,t), p]$$

Model

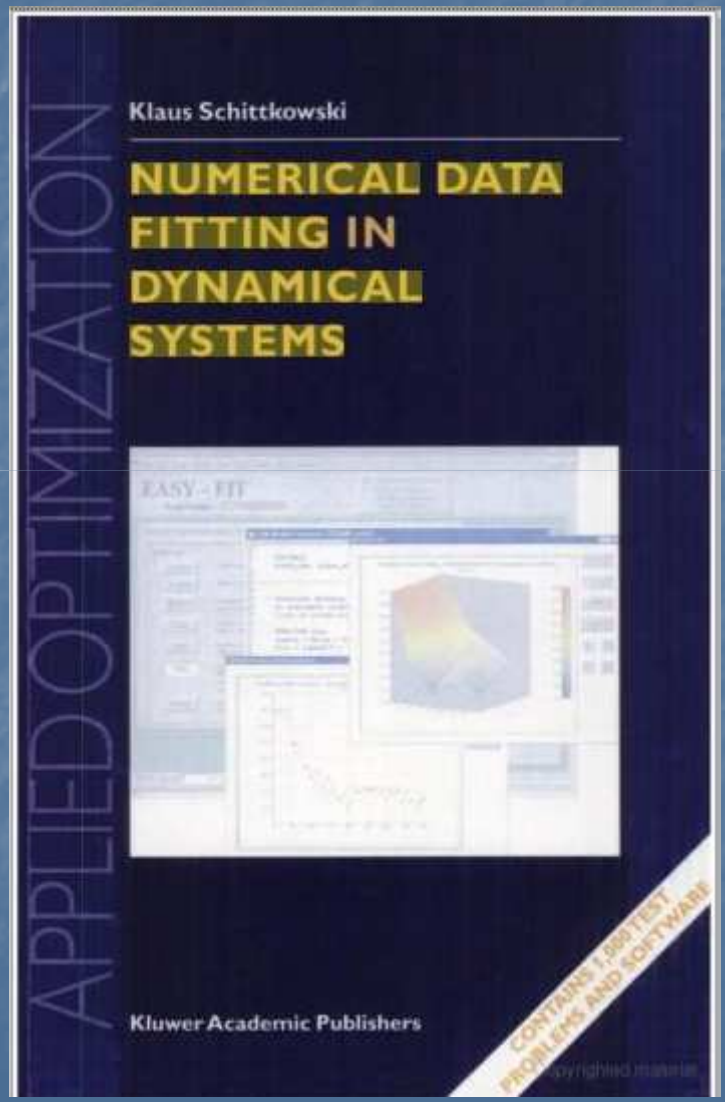


Balsa-Canto, E., Alonso, A. A., & Banga, J. R. (2010) ***BMC Systems Biology*** 4:11

Parameter estimation

Possible pitfalls and difficulties *(Schittkowski, 2002)*

- *Convergence to a **local solutions** (with standard local methods)*
- *Narrow curved valleys*
- *Very flat objective function in the neighbourhood of a solution*
- *Over-determined models, leading to many solution vectors*
- *Bad starting values for parameters*
- *Badly scaled model functions*
- *Non-differentiable model functions*



Parameter estimation

Local methods (L-M, G-N, etc.) may converge to local optima...

If one gets a bad fit, there is a Key Question:

Is my model wrong, or has my solver failed?

- **Need of Global Optimization methods**

GENOME RESEARCH

Downloaded from www.genome.org on April 5, 2006 - Published by Cold Spring Harbor Laboratory Press

Methods

Parameter Estimation in Biochemical Pathways: A Comparison of Global Optimization Methods

Carmen G. Moles,[1] Pedro Mendes,[2] and Julio R. Banga[1,3]

[1]Process Engineering Group, Instituto de Investigaciones Marinas (CSIC), 36208 Vigo, Spain; [2]Virginia Bioinformatics Institute, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061, USA

Here we address the problem of parameter estimation (inverse problem) of nonlinear dynamic biochemical pathways. This problem is stated as a nonlinear programming (NLP) problem subject to nonlinear differential-algebraic constraints. These problems are known to be frequently ill-conditioned and multimodal. Thus, traditional (gradient-based) local optimization methods fail to arrive at satisfactory solutions. To surmount this limitation, the use of several state-of-the-art deterministic and stochastic global optimization methods is explored. A case study considering the estimation of 36 parameters of a nonlinear biochemical dynamic model is taken as a benchmark. Only a certain type of stochastic algorithm, evolution strategies (ES), is able to solve this problem successfully. Although these stochastic methods cannot guarantee global optimality with certainty, their robustness, plus the fact that in inverse problems they have a known lower bound for the cost function, make them the best available candidates.

Mathematical optimization can be used as a computational engine to arrive at the best solution for a given problem in a systematic and efficient way. In the context of biochemical systems, coupling optimization with suitable simulation modules opens a whole new avenue of possibilities. Mendes and Kell (1998) highlight two types of important applications:

1. Design problems: How to rationally design improved metabolic pathways to maximize the flux of interesting products and minimize the production of undesired by-products (metabolic engineering and biochemical evolution studies);
2. Parameter estimation: Given a set of experimental data, calibrate the model so as to reproduce the experimental results in the best possible way.

This contribution considers the latter case, that is, the so-called inverse problem. The correct solution of inverse problems plays a key role in the development of dynamic models, which, in turn, can promote functional understanding at the systems level, as shown by, for example, Swameye et al. (2003) and Cho et al. (2003) for signaling pathways.

The paper is structured as follows: In the next section, we state the mathematical problem, highlighting its main characteristics, and very especially its challenging nature for traditional local optimization methods. Next, global optimization (GO) is presented as an alternative to surmount these difficulties. A brief review of GO methods is given, and a selection of the presently most promising alternatives is presented. The following section outlines a case study considering the estimation of 36 parameters of a three-step pathway, which will be used as a benchmark to compare the different GO methods selected. A Results and Discussion section follows, ending with a set of Conclusions.

METHODS

Statement of the Inverse Problem

Parameter estimation problems of nonlinear dynamic systems are stated as minimizing a cost function that measures the goodness of the fit of the model with respect to a given experimental data set, subject to the dynamics of the system (acting as a set of differential equality constraints) plus possibly other algebraic constraints. Mathematically, the formulation is that of a nonlinear programming problem (NLP) with differential-algebraic constraints:

Find $\mathbf{p}$ to minimize

$$J = \int_0^{t_f} (\mathbf{y}_{meas}(t) - \mathbf{y}(\mathbf{p},t))^T \mathbf{W}(t) (\mathbf{y}_{meas}(t) - \mathbf{y}(\mathbf{p},t)) dt \quad (1)$$

subject to

$$\mathbf{f}\left(\frac{d\mathbf{x}}{dt}, \mathbf{x}, \mathbf{y}, \mathbf{p}, \mathbf{v}, t\right) = 0 \quad (2)$$

$$\mathbf{x}(0) = \mathbf{x}_0 \quad (3)$$

$$\mathbf{h}(\mathbf{x}, \mathbf{y}, \mathbf{p}, \mathbf{v}) = 0 \quad (4)$$

$$\mathbf{g}(\mathbf{x}, \mathbf{y}, \mathbf{p}, \mathbf{v}) \le 0 \quad (5)$$

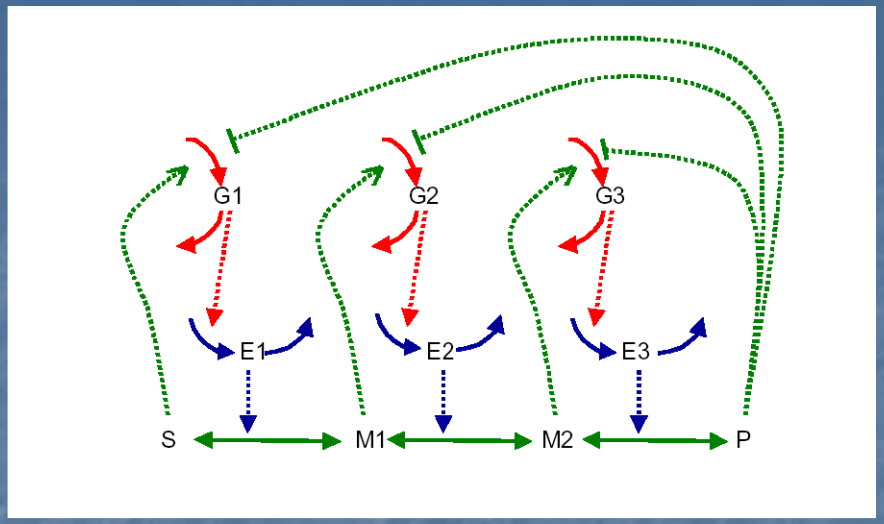
$$\mathbf{p}^L \le \mathbf{p} \le \mathbf{p}^U \quad (6)$$

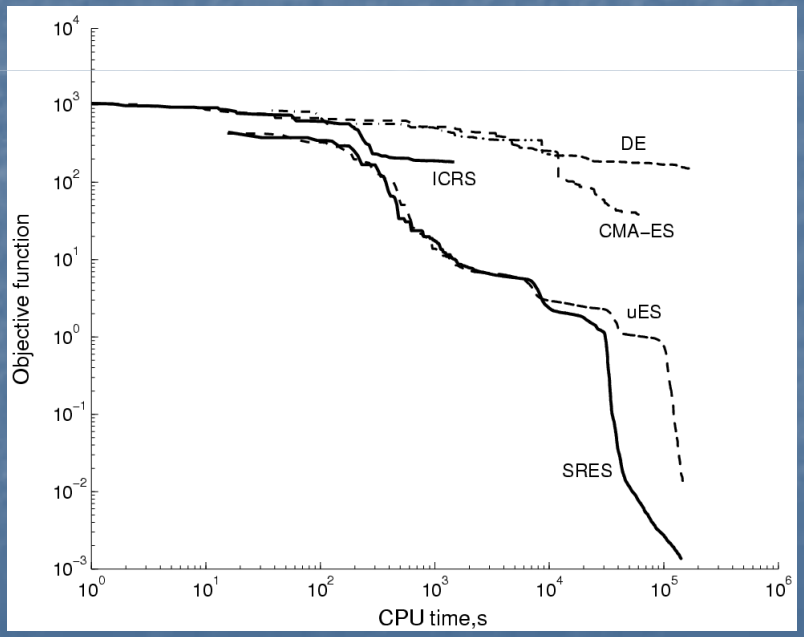
where J is the cost function to be minimized, $\mathbf{p}$ is the vector of decision variables of the optimization problem, the set of parameters to be estimated, $\mathbf{y}_{meas}$ is the experimental measure of a subset of the (so-called) output state variables, $\mathbf{y}(\mathbf{p}, t)$ is the model prediction for those outputs, $\mathbf{W}(t)$ is a weighting (or scaling) matrix, $\mathbf{x}$ is the differential state variables, $\mathbf{v}$ is a vector of other (usually time-invariant) parameters that are not estimated, $\mathbf{f}$ is the set of differential and algebraic equality constraints describing the system dynamics (i.e., the nonlinear process model), and $\mathbf{h}$ and $\mathbf{g}$ are the possible equality and inequality path and point constraints that express additional requirements for the system performance. Finally, $\mathbf{p}$ is subject to upper and lower bounds acting as inequality constraints.

The formulation above is that of a nonlinear programming problem (NLP) with differential-algebraic (DAE) constraints. Because of the nonlinear and constrained nature of the system dynamics, these problems are very often multimodal (nonconvex). Therefore, if this NLP-DAE is solved via standard local methods, such as the standard Levenberg-Marquardt method, it is very likely that the solution found will be of local nature, as discussed by Mendes and Kell (1998), for example.

The earliest and simplest attempt to surmount the nonconvexity of many optimization problems was based on the idea of using a local method repeatedly, starting from a number of dif-

[3]Corresponding author.
E-MAIL jrbanga@iim.csic.es; FAX 34-986292762.
Article and publication are at http://www.genome.org/cgi/doi/10.1101/gr.1262503. Article published online before print in October 2003.





Moles, C. G., P. Mendes and J. R. Banga (2003) ***Genome Research***, 13(11):2467-2474

New metaheuristics

BMC Bioinformatics

BioMed Central

Methodology article

Open Access

Novel metaheuristic for parameter estimation in nonlinear dynamic biological systems

Maria Rodriguez-Fernandez, Jose A Egea and Julio R Banga*

Address: Process Engineering Group, Instituto de Investigaciones Marinas (C.S.I.C.), Spanish Council for Scientific Research, C/Eduardo Cabello, 6. 36208 Vigo, Spain

Email: Maria Rodriguez-Fernandez - mrodriguez@iim.csic.es; Jose A Egea - jegea@iim.csic.es; Julio R Banga* - julio@iim.csic.es

* Corresponding author

Published: 02 November 2006

BMC Bioinformatics 2006, 7:483 doi:10.1186/1471-2105-7-483

Received: 29 May 2006
Accepted: 02 November 2006

This article is available from: http://www.biomedcentral.com/1471-2105/7/483

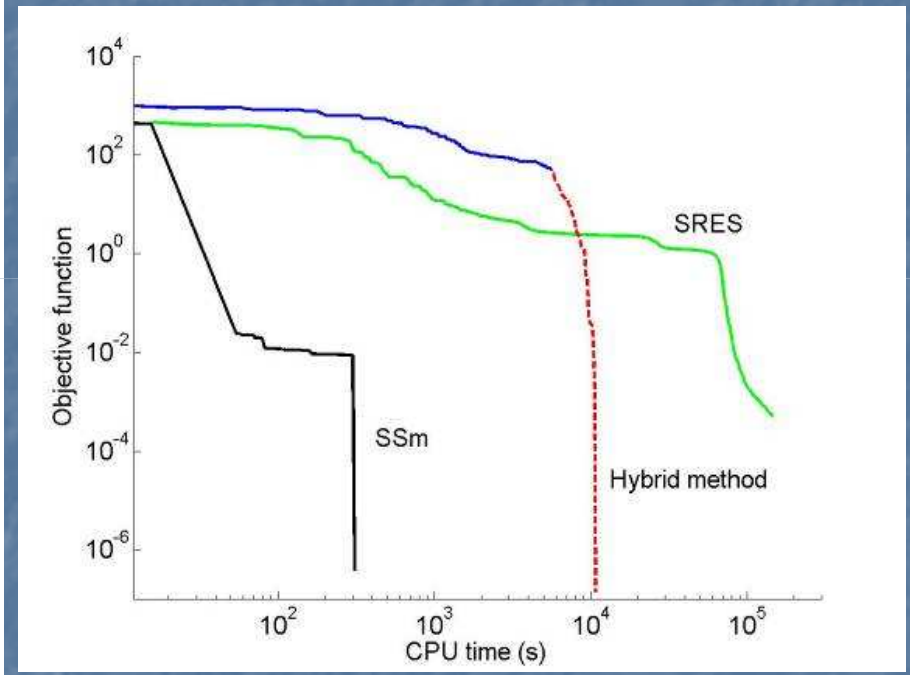
© 2006 Rodriguez-Fernandez et al; licensee BioMed Central Ltd.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: We consider the problem of parameter estimation (model calibration) in nonlinear dynamic models of biological systems. Due to the frequent ill-conditioning and multi-modality of many of these problems, traditional local methods usually fail (unless initialized with very good guesses of the parameter vector). In order to surmount these difficulties, global optimization (GO) methods have been suggested as robust alternatives. Currently, deterministic GO methods can not solve problems of realistic size within this class in reasonable computation times. In contrast, certain types of stochastic GO methods have shown promising results, although the computational cost remains large. Rodriguez-Fernandez and coworkers have presented hybrid stochastic-deterministic GO methods which could reduce computation time by one order of magnitude while guaranteeing robustness. Our goal here was to further reduce the computational effort without loosing robustness.

Results: We have developed a new procedure based on the scatter search methodology for nonlinear optimization of dynamic models of arbitrary (or even unknown) structure (i.e. black-box models). In this contribution, we describe and apply this novel metaheuristic, inspired by recent developments in the field of operations research, to a set of complex identification problems and we make a critical comparison with respect to the previous (above mentioned) successful methods.

Conclusion: Robust and efficient methods for parameter estimation are of key importance in systems biology and related areas. The new metaheuristic presented in this paper aims to ensure the proper solution of these problems by adopting a global optimization approach, while keeping the computational effort under reasonable values. This new metaheuristic was applied to a set of three challenging parameter estimation problems of nonlinear dynamic biological systems, outperforming very significantly all the methods previously used for these benchmark problems.



Rodriguez-Fernandez, M., J. A. Egea and J. R. Banga (2006) ***BMC Bioinformatics*** 7:483

Optimal dynamic experimental design

- *To design optimal dynamic experiments which maximize information content*

- *Measurements:*
  - *What?*
  - *When? (sampling times)*
- *Stimuli / controls:*
  - *Which? How? (dynamic profile)*

- *Number and type of experiments (sequential, parallel)*

- *Mixed-integer dynamic optimization (MIDO) problem formulation*

- *Cost function based on INFORMATION CRITERIA (Fisher, Akaike)*

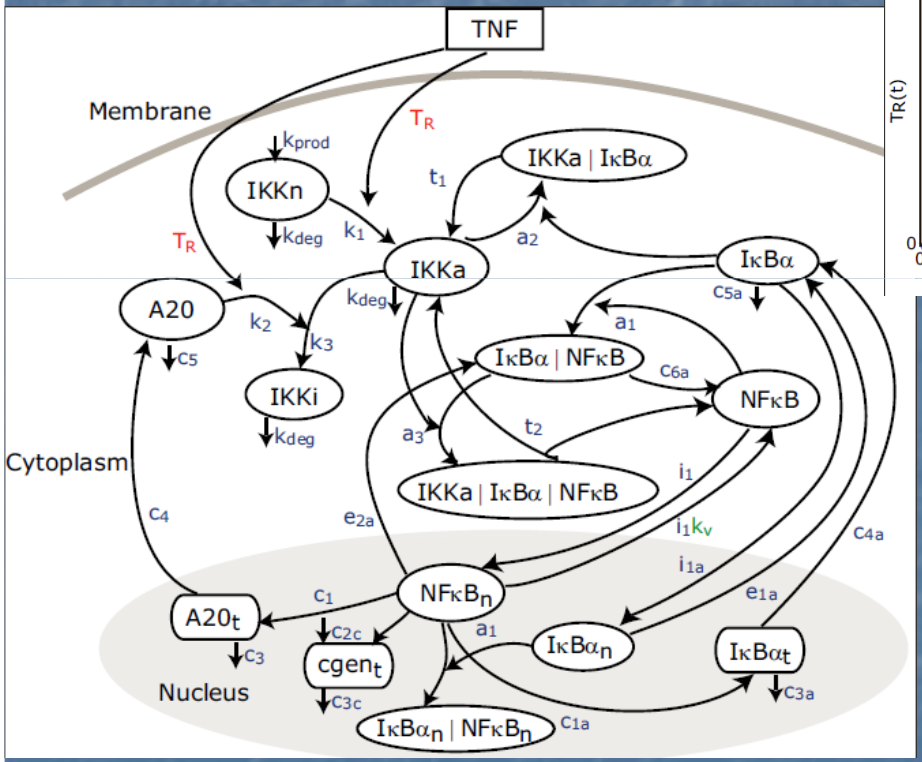
- *Model calibration, selection and discrimination*

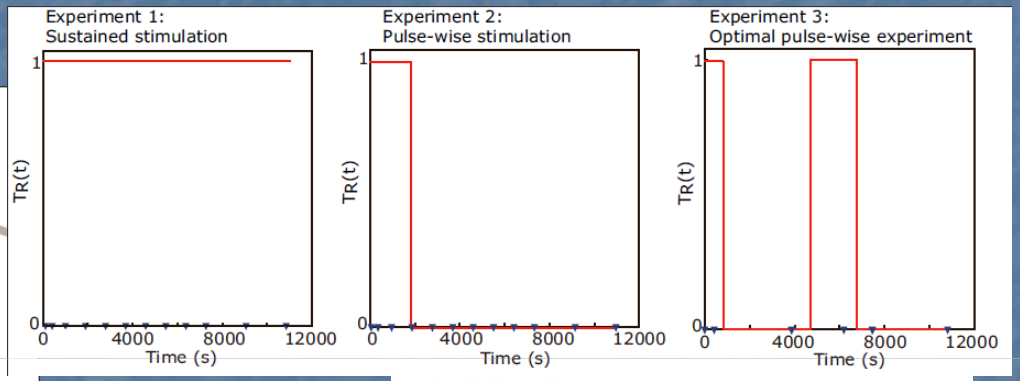
Balsa-Canto, E., Alonso, A. A., & Banga, J. R. (2010) ***BMC Systems Biology*** 4:11

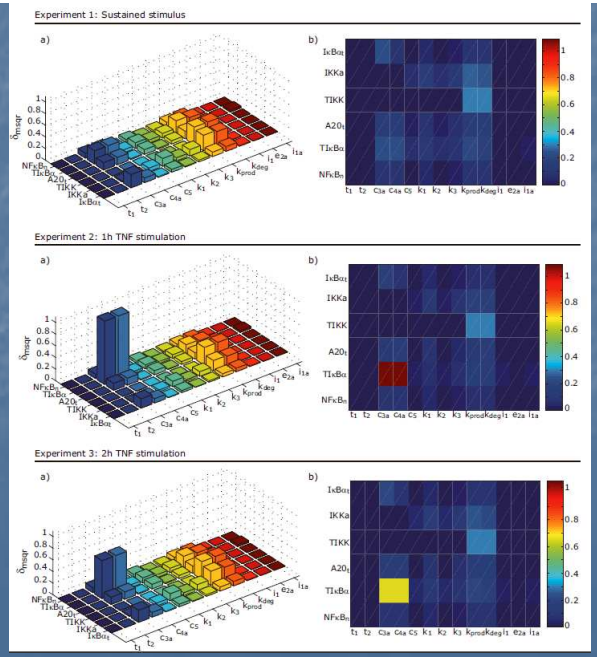
Example: optimal dynamic experimental design in cell signalling

System: ***NF-κB module in cell signalling***

- ***Optimal experiements pulse-wise TNF stimuli***

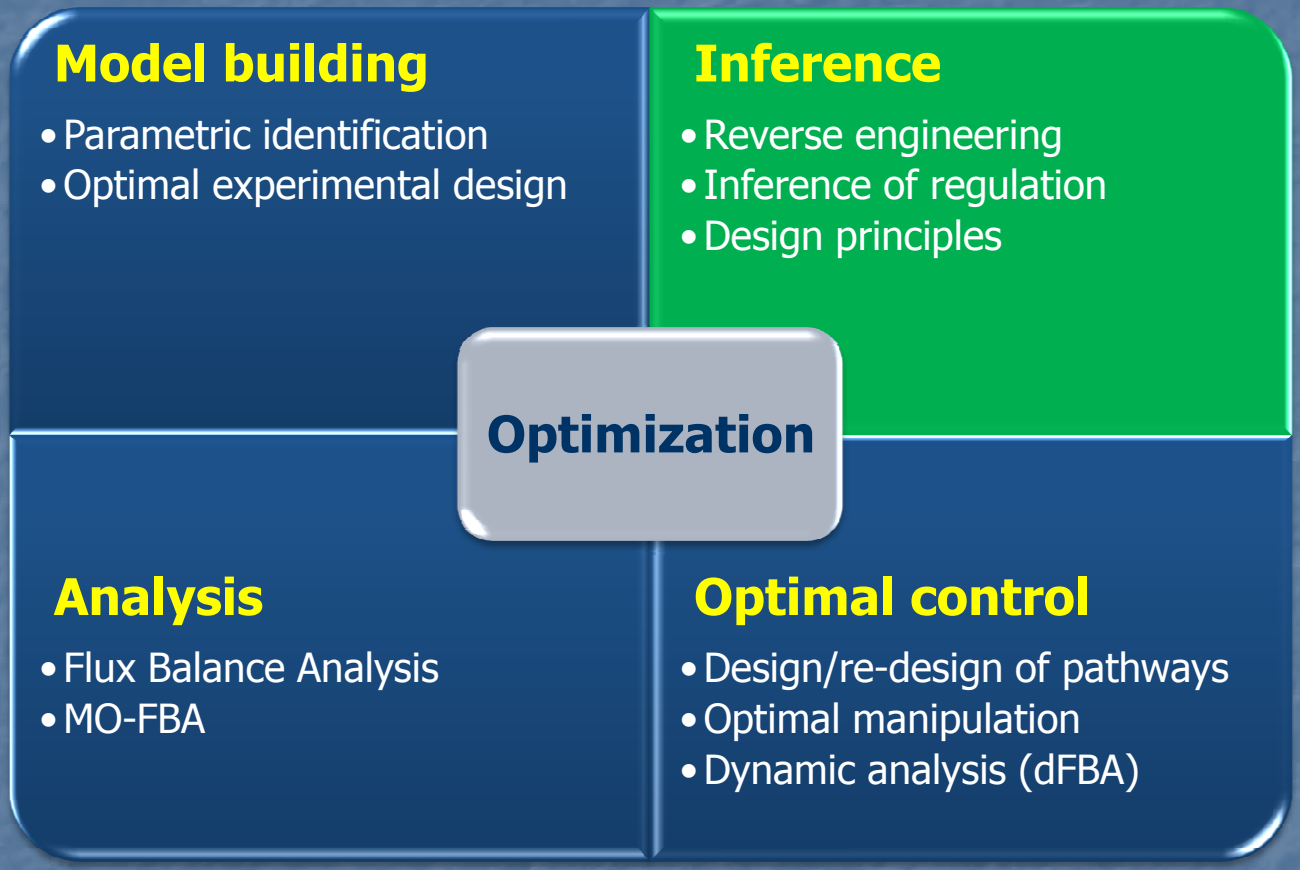






Balsa-Canto, E., Alonso, A. A., & Banga, J. R. (2010) ***BMC Systems Biology*** 4:11

Optimization in systems biology: examples of applications



Design principles, re-design

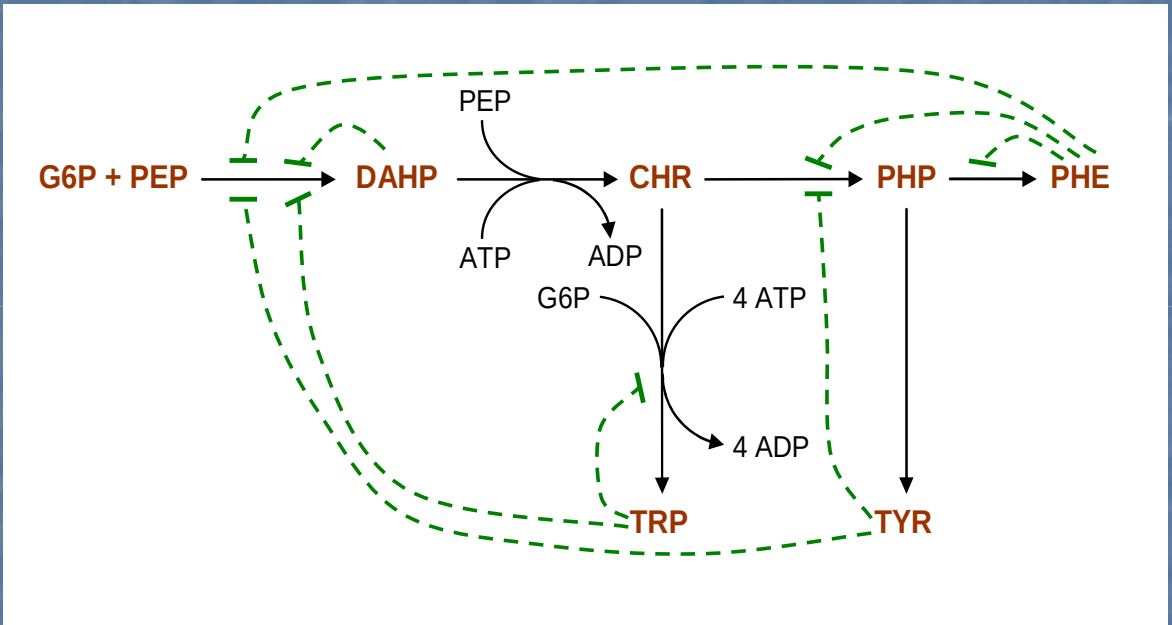
- **Optimality principles in biological systems**

  (objective function not always clear)

- ***Use model-based optimization to:***
  - ***Explain design principles of a biosystem***
  - ***"Re-design" biological systems***
  - ***Synthesis***

Example: Optimal "design" problem (metabolic engineering)

Modification of existing regulatory and activity structure of the aromatic amino acid biosynthesis pathway



The aim of the problem is to determine which of the regulatory loops (**integer variables**) should be retained, and what should be the changes in the enzyme expression levels (**continuous variables**) to optimize a certain objective function

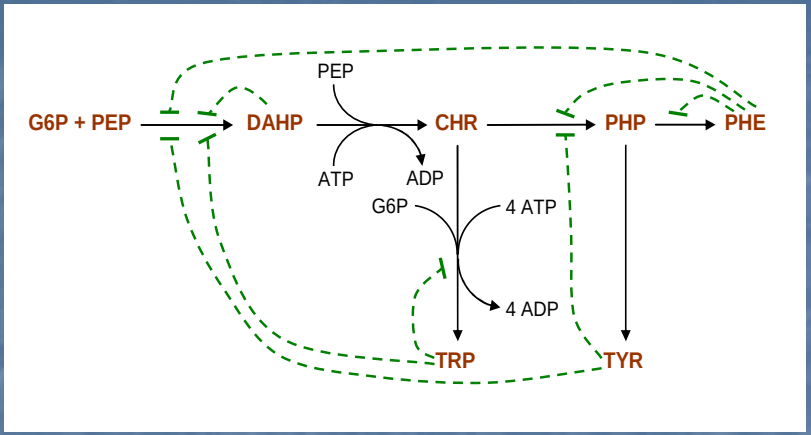
Hatzimanikatis, V., Floudas, C.A. and Bailey, J.E. (1996). *AIChe Journal*, 42(5), 1277-1292.

Example: Optimal "design" problem

Design Problem (MO-MINLP)

Objective Functions

Maximization of Phe Selectivity subject to physiological constraints



Other criteria (usually imposed as '*artificial' constraints*):

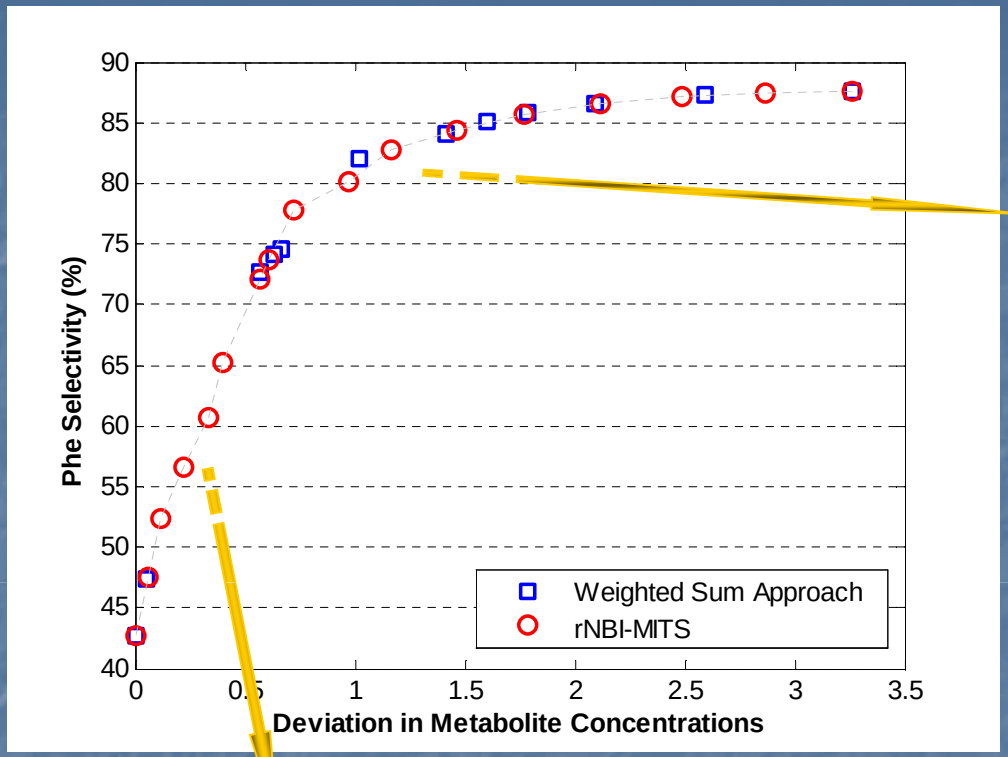
- ☞ Simultaneous minimization of the deviation in metabolite concentrations
- ☞ Simultaneous minimization of total enzyme activity

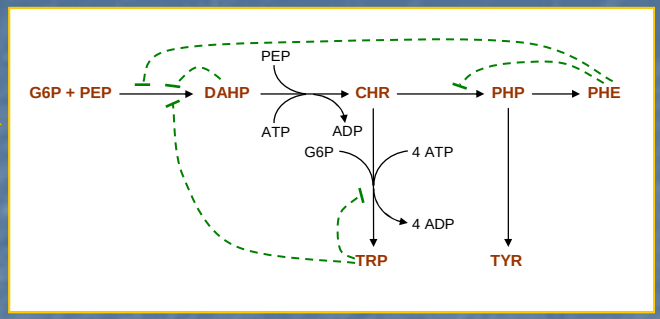
Optimization Method

- New reformulation of the NBI method
- Extension to handle MO-MINLP
- Global Solver for MINLP problems: **MITS** (Exler, Alonso and Banga, 2007)

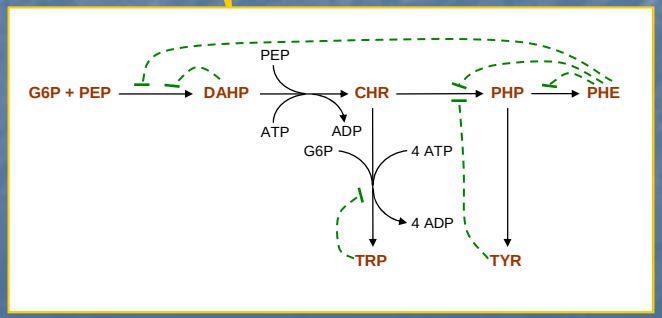
Sendin, J.O.H., O. Exler & J.R. Banga (2010) ***IET Systems Biology*** 4(3):236-248.

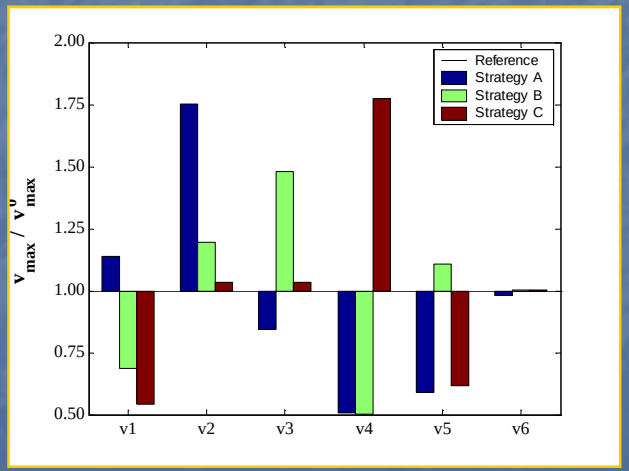
Example: Optimal "design" problem





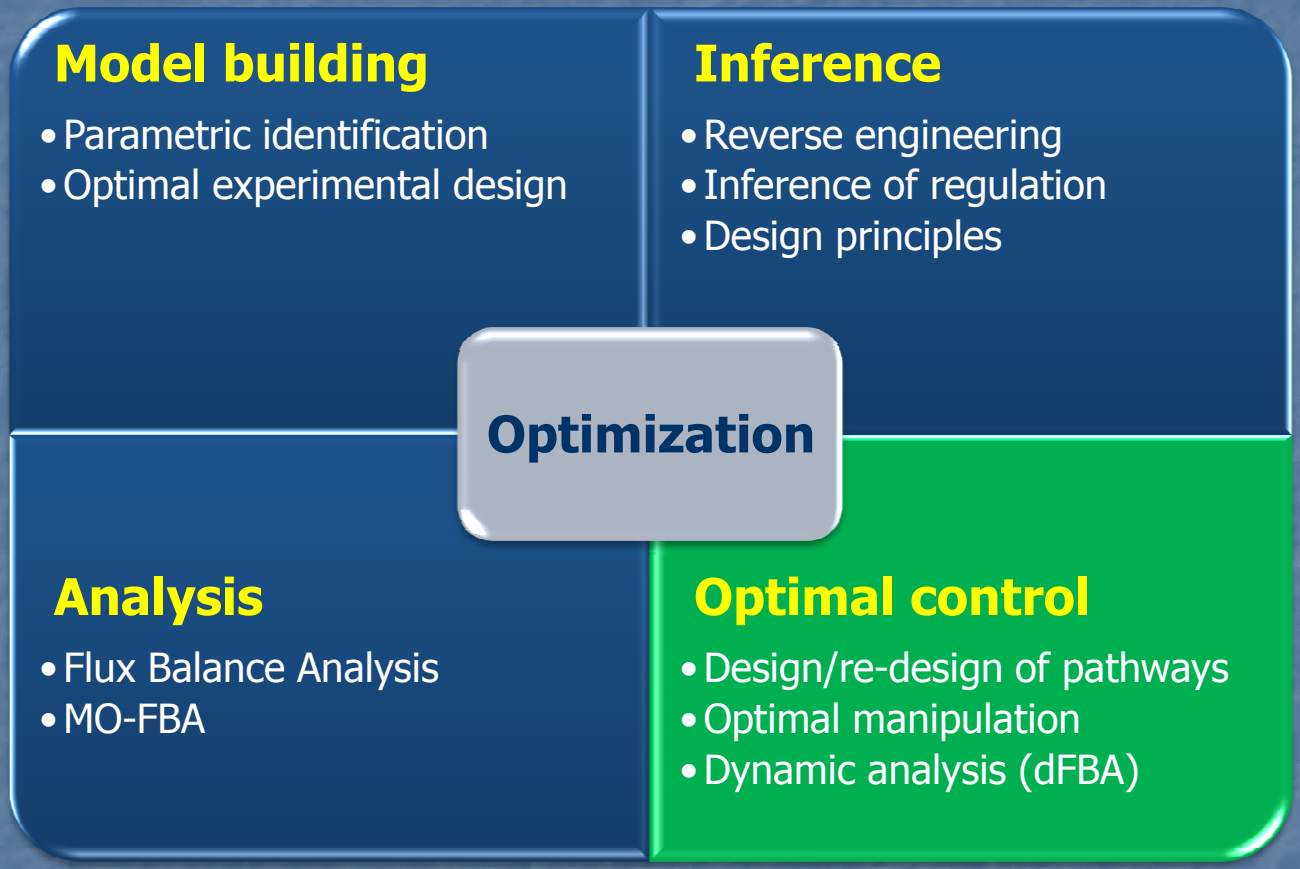
- Different optimal strategies for manipulation
- No « artificial » constraints are needed
- More accurate selection of final solution





Sendin, J.O.H., O. Exler & J.R. Banga (2010) ***IET Systems Biology*** 4(3):236-248.

Optimization in systems biology: examples of applications



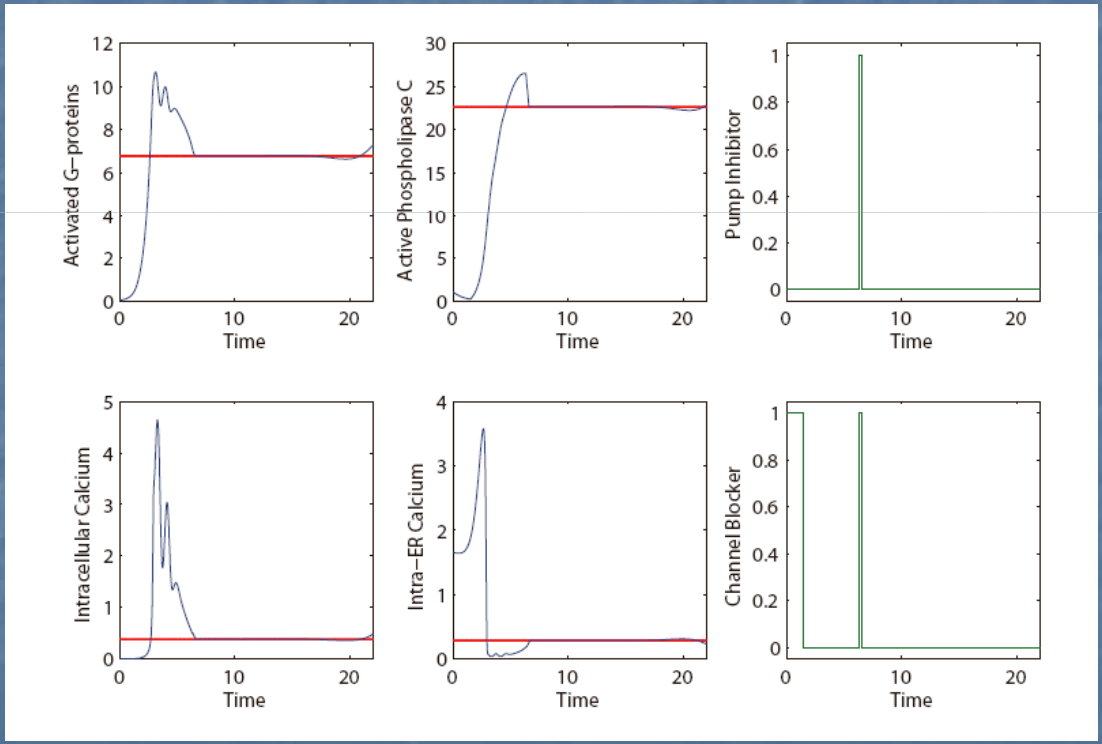
Dynamic optimization (optimal control)

- Optimal « intervention » strategies
- Examples
  - Control of dynamical features (e.g. oscillations) of biosystems
  - Control of spatial structures (patterns)
  - Dynamic analysis (e.g. dFBA)
  - Computer aided design of biological units (synthetic biology)
  - Optimal drug scheduling

Example: phase resetting of calcium oscillator problem

System: *intracellular calcium spiking in hepatocytes induced by an extracellular increase in ATP concentration*
(Lebiedz,D. et al. (2005) Phys. Rev. Lett., 95, 108303)

- *To **minimize the intracellular oscillations**, driving the system to a given desired state by the use of two inhibitors.*
- ***Mixed-integer dynamic optimization** (MIDO) problem formulation*



Hirmajer, T., E. Balsa-Canto and J. R. Banga (2009) **BMC Bioinformatics** 10:199.

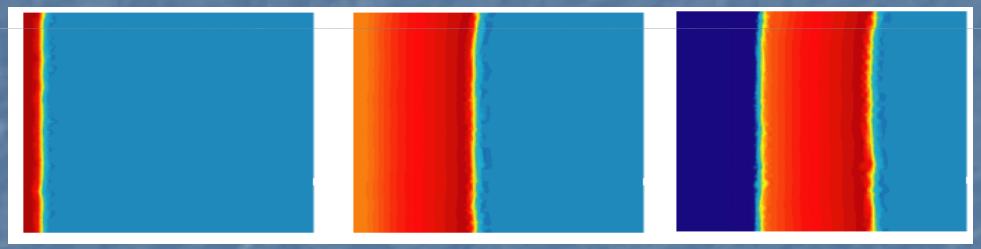
Example: robust control of diffusion-reaction phenomena

The FitzHugh-Nagumo (FHN) Model

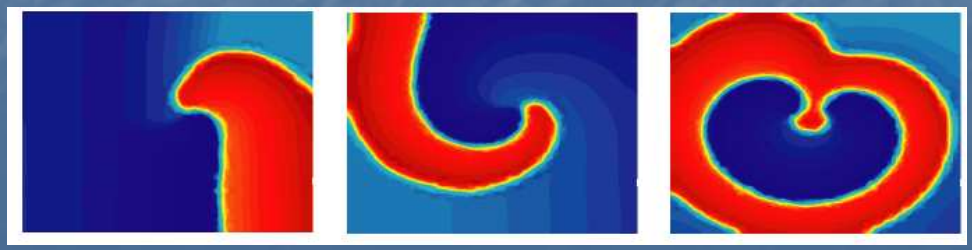
Modelling neuron firing in the brain, the heartbeat, cellular organization activities, etc.

$$\frac{\partial v}{\partial t} = \left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right) + f(v) - w + p;\ f(v) = (\alpha - v)(v-1)v,$$

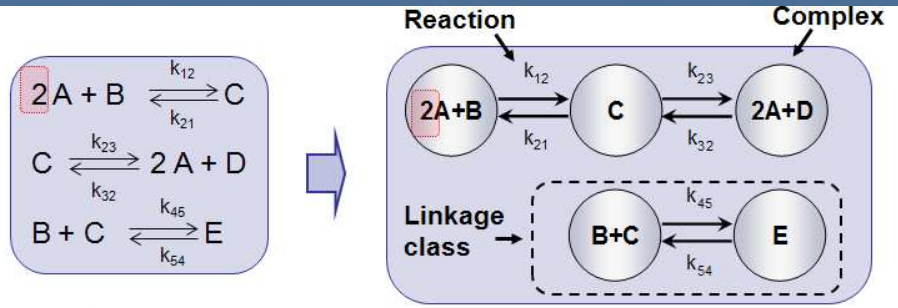
$$\frac{\partial w}{\partial t} = g(w) - \varepsilon\beta v;\ g(w) = \varepsilon(\gamma w - \delta),$$



Vilas, C., M.R. Garcia, J.R. Banga, A.A. Alonso (2008) **Physica D** 237(18):2353-2364
Vilas, C., M. R. Garcia, J. R. Banga and A. A. Alonso (2006) **Journal of Theoretical Biology** 241 (2): 295-306



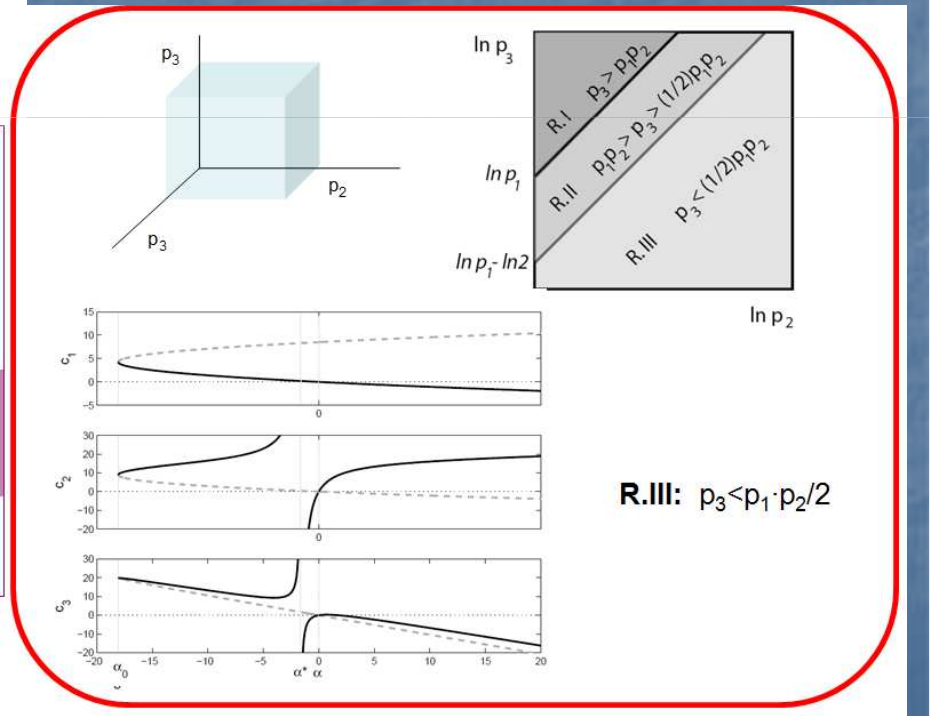
Example: *Multiplicity Conditions in Biochemical Reaction Networks*



$$\dot{c} = \underbrace{\begin{bmatrix} 2 & 0 & 2 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}}_{Y(m\times n)} \underbrace{\begin{bmatrix} -k_{12} & k_{21} & 0 & 0 & 0 \\ k_{12} & -k_{21}-k_{23} & k_{32} & 0 & 0 \\ 0 & k_{23} & -k_{32} & 0 & 0 \\ 0 & 0 & 0 & -k_{45} & k_{54} \\ 0 & 0 & 0 & k_{45} & -k_{54} \end{bmatrix}}_{A_k(n\times n)} \underbrace{\begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \\ \psi_5 \end{bmatrix}}_{\psi(c)}$$

$$\dot{c} = Y \cdot A_k \cdot \psi(c)$$

$$\begin{aligned} \text{Find} \quad & \min_{x\in\mathbb{R}^{n_k+1}} J(x) \qquad (k \in \mathbb{R}^{n_k}, \alpha \in \mathbb{R}) \\ \text{subject to :} \quad & \mathcal{H}_c(k,\alpha) = 0 \\ & p(k) \in R_j \\ & c_i(k,\alpha) > 0, \quad i = 1,\ldots,r \\ & x_l \le x \le x_u \quad (x_l, x_u \in \mathbb{R}^{n_{k+1}}) \end{aligned}$$



Otero, I, J. R. Banga, A. A. Alonso (2009)
Biotechnology Progress 25(3):619-631

Beyond optimization

- standard optimization sometimes insufficient (e.g. co-evolution of biological systems)
- while evolving towards optimal properties, the environment may change or organisms may even change their own environment, which in turn alters the optimum
- in an evolutionary system, continuing development is needed so as to maintain its fitness relative to the systems it is co-evolving with
- Game-theoretic approaches may provide a better framework studying the evolution of biochemical systems.

Red Queen effect

“…it takes all the running you can do, to keep in the same place.”

In an evolutionary system, continuing development is needed just in order to maintain its fitness relative to the systems it is co-evolving with.

Van Valen L. (1973): "A New Evolutionary Law", *Evolutionary Theory* 1, p. 1-30.

Lewis Carroll. 1960 (reprinted). The Annotated Alice: Alice's Adventures in Wonderland and Through the Looking-Glass, illustrated by J. Tenniel, with an Introduction and Notes by M. Gardner. The New American Library, New York, 345 pp

Conclusions

Conclusions

- Systems biology: dynamics, feedback, optimality
- Optimization in systems biology: smart simplification, but need of global optimization in many cases
- Optimization: key element for model building, analysis, inference, design and control
- Scaling-up: novel methods + HPC
- Beyond optimization: game-theoretic approaches

PROCESS ENGINEERING
www.iim.csic.es/~gingproc/software.html
IIM Webmail
:: Julio R. Banga Ho...
Intranet CSIC
Win7
Gmail
iGoogle
Netvibes
Scopus
Google Académico
ISI WoK
Lifehacker, tips and ...
Otros marcadores
< home >
:: welcome
:: members
:: research
:: software
< contact >
Process Engineering Group
IIM-CSIC
C/Eduardo Cabello 6
36208 Vigo
SPAIN
e-mail: julio@iim.csic.es
CSIC
software
by PROCESS ENGINEERING GROUP, IIM-CSIC
The following software packages are freely available for ACADEMIC users under the licenses referenced below. Commercial use of these packages is currently not allowed (contact us for more info).
SSm GO
SSm GO toolbox: Scatter Search for Global Optimization in Matlab
Info / Download / Documentation / References
DOTcvpSB toolbox
Matlab Toolbox for Dynamic Optimization in Systems Biology
Info / Download / Documentation / References
AMIGO
AMIGO: a Matlab toolbox for Advanced Modelling and Identification using Global Optimization
Info / Download / Documentation / References
SensSB
SensSB: a Matlab toolbox for Sensitivity Analysis in Systems Biology models
Info / Download / Documentation / References
GLOBALm
Software available upon request; email us if interested
The method is described in:

Acknowledgements

- IIM-CSIC: Oscar H. Sendín, José Egea, Eva Balsa-Canto, María Rodríguez, Oliver Exler, Martin Schlüter, Tomas Hirmajer, Alex F. Villaverde, Irene Otero, Antonio A. Alonso
- Olaf Wolkenhauer, Julio Vera (COSBICS)
- Jens Timmer, M. Peifer, C. Fleck (COSBICS)
- Wolfram Liebermeister, Edda Klipp (BASYSBIO)
- Joerg Stelling, Markus Uhr, Uwe Sauer (BASYSBIO)
- Ian Booth (Aberdeen), Andreas Kremling (MPI), K. Jung (Munich)

Funding: