Contributions in Computational Systems Biology

Eva Sciacca
sciacca@dmi.unict.it
Ph.D. Student at University of Catania

About Me

I belong to the PhD Program in Mathematics for Technology (Coordinator Prof. Giovanni Russo) at the University of Catania, and my work is supervised by:

- Prof. Giuseppe Nicosia;

I spent 16 months at the Department of Mechanical Engineering and Biological Engineering at the Massachusetts Institute of Technology (MIT) in Boston U.S.A. and I was supervised by:

- Prof. C. Forbes Dewey Jr.

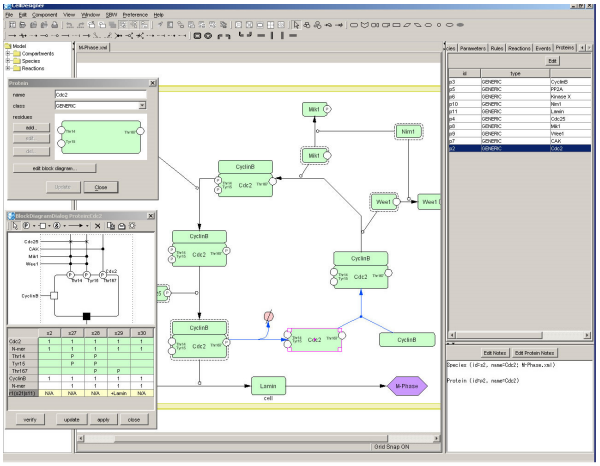
SUMMARY

- Intro: Modeling Biological Systems with ODE
- Robust Parameter Identification for Biological Circuit Calibration; [Nicosia]
- Implementation of A Web Based Tool for Integration of Molecular Pathway Models (CytoSolve); [MIT]
- Reliable Biological Circuit Design Including Uncertain Kinetic Parameters

3

MODELING

Biology is understood and is translated into graphs, equations, reactions.



http://www.celldesigner.org/

4

Main Ingredients

Objects: molecules (cytokines/chemokines/...),
cells (Macrophages, Neutrophils, ...),
organs (lymph node, spleen, .., lung,..)

Actions: trafficking/migration,
interaction (activation/inhibition),
proliferation

Ordinary Differential Equations are about rate of change of quantities

5

Ordinary Differential Equations (ODE)

a – some quantity examples: cell count, receptor expression level, cell damage, ...

We write ODE as
$da/dt = f$ where f may be a complex formula

We interpret this ODE as
$da = f\,dt$

- the change in a during a short time interval dt is equal to f times dt

f may depend on time or may depend on a or both.

6

Basic Examples

Kinetics – how fast does a reaction proceed?

1. da/dt = 0
   - This means da = 0 * dt = 0 → change in a is 0, → a does not change

2. da/dt = 1
   - This means da = dt → a changes by dt

3. da/dt = - a da = - a * dt → a changes by – a * dt
   - This means that a decreases, and the reduction is large when a is large and getting small when a is getting smaller.

7

Modeling Reactions - The Law of Mass Action

The rate of change of products is proportional to the product of reactants concentration

A → 0

The only reagent (left side) is a :

→ rate of change is proportional to a,

ODE da/dt = -k*a (minus sign since we loose a)

A → B :

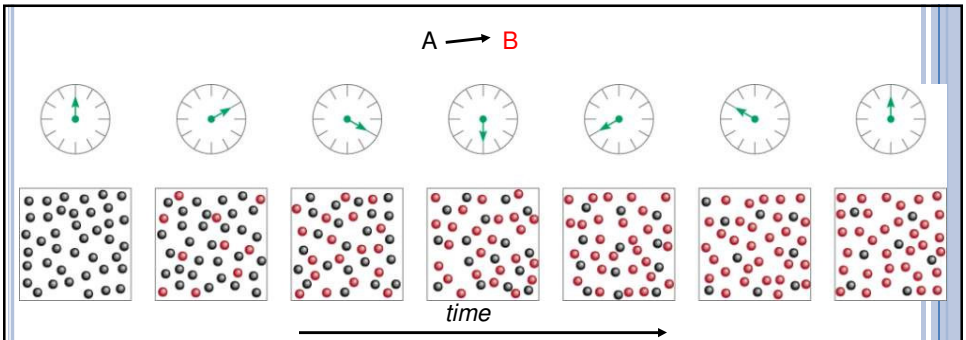
Similar to the previous case but here one B is created per each A that disappear

ODE

da/dt = -k*a as before but we also have

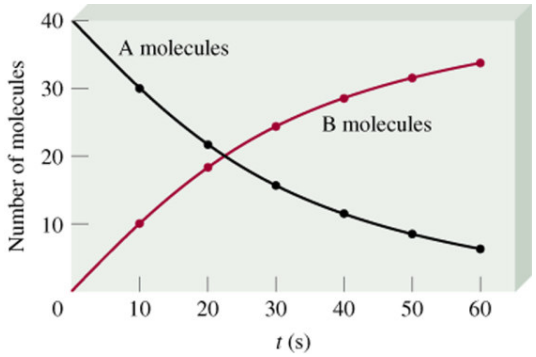
db/dt = k*a; here the sign is +

8



$$\text{rate} = -\frac{\Delta[A]}{\Delta t}$$

$$\text{rate} = \frac{\Delta[B]}{\Delta t}$$



9

Modeling Reactions – cont.

A + B → C;

Here the reactants (left side) is A and B, the product (right side) is C.

dc/dt = k *a*b; C is created at a rate proportional to the product of the concentration of A and B

da/dt = - k*a*b; The rate of change of A is the same as the rate of change of C – per each C that is created one A is lost

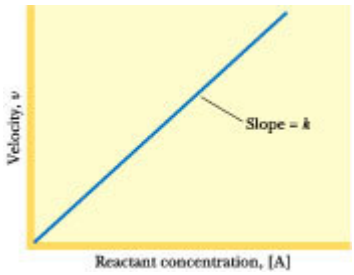
db/dt = - k*a*b, similar to A.

10

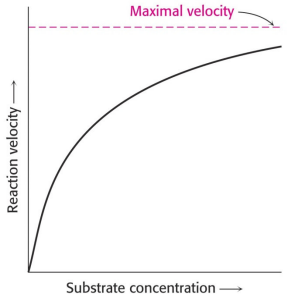
Basic problem of enzyme kinetics

Suppose an enzyme were to react with a substrate, giving a product.

$$S + E \rightleftharpoons P + E$$



If we simply applied the law of mass action to this reaction, the rate of reaction would be a linearly increasing function of [S]. As [S] gets very big, so would the reaction rate.



This doesn't happen. In reality, the reaction rate saturates.

11

Michaelis and Menten

In 1913, Michaelis and Menten proposed the following mechanism for a saturating reaction rate

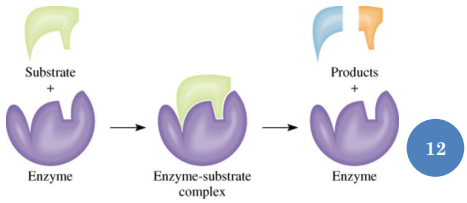
$$S + E \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} P + E$$

Complex.

product

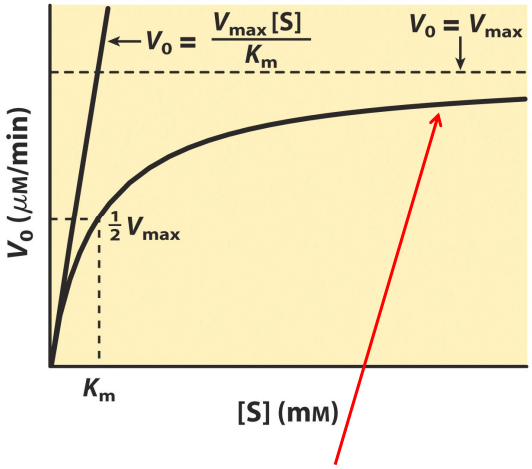
Leonor Michaelis
1875–1949

Maud Menten
1879–1960



12

Michaelis-Menten Kinetics



- When [S] << K_M, the reaction increases linearly with [S]; I.e. $v_o = (V_{max} / K_M)$ [S]
  Very little [ES] is formed
- When [S] = K_M, $v_o = V_{max}/2$ (half maximal velocity); this is a definition of K_M: the concentration of substrate which gives $\frac{1}{2}$ of V_{max}. This means that low values of K_M imply the enzyme achieves maximal catalytic efficiency at low [S].
- When [S] >> K_m, $v_o = V_{max}$

Where activity measurements should be performed: 1. [S] very high
2. all enzyme bound in [ES] complex

13

Michaelis-Menten Kinetics

When the enzyme is saturated with substrate, the reaction is progressing at its maximal velocity, V_{max}.

Combing the steady-state assumption (d[ES]/dt=0) with the conservation condition ([E]T=[E] + [ES]) v_o leads to the Michaelis-Menten Equation of enzyme kinetics:

$$v = \frac{V_{max}[S]}{K_M + [S]}$$

where Km is

$K_M = (k_{-1} + k_2)/k_1$

14

Michaelis-Menten Kinetics

What is V_{max} and K_M ?

- **K_M gives an idea of the range of [S] at which a reaction will occur. The larger the K_M, the WEAKER the binding affinity of enzyme for substrate.**

- **V_{max} gives an idea of how fast the reaction can occur under ideal circumstances.**

15

Analogies between Biological Networks and Electronic Circuits

Biological Domain	Electrical Domain
mass	charge
Mass flux	current
concentration	voltage
stoichiometric conservation	Kirchhoffs voltage law
mass conservation	Kirchhoffs current law

16

Robust Parameter Identification for Biological Circuit Calibration

17

Outline

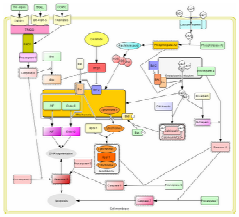
- Motivation
- Problem Definition
- Tested Algorithms
- Case study: P53/MDM2 model
- Results
- Conclusions

18

Motivation

- Accurately modeling and simulating biological networks is a ***challenging problem***, due to the complex interaction between large numbers of interacting pathways, feedback inherent to the system, and the ***stochastic*** nature of biological processes.
- The aim of this work is to give a computational tool to analyze the ***robustness*** (less sensitive to the noise of the experimental data) of the parameters of multivariate, multi-scale, hybrid biological networks.
- We have tested classical methods such as LSQNONLIN of MATLAB, DIRECT, and a Pattern Search Algorithm and two evolutionary algorithms: **C**ovariance **M**atrix **A**daptation **E**volution **S**trategy and **D**ifferential **E**volution.

Problem Definition: Parameter Identification



Biological Model

Described by

Given

S1→S2
S1+S3 →S4
…

Chemical Reactions

Modeled with

$$\frac{d[S_1]}{dt} = K_1 * [S_2] + K_2 * [S_1]$$
$$\frac{d[S_2]}{dt} = K_3 * [S_2] + \frac{K_4 [S_3]}{[S_4]}$$
…

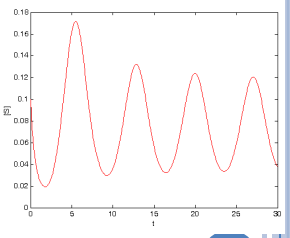
System of ODE

Find

K_1 $[S_1](t_0)$
K_2 or $[S_2](t_0)$
K_3 $[S_3](t_0)$
… …

Initial Conditions
Or
Reaction Rates

such that



Simulated Concentrations
best fit Experimental data

Optimization Problem

Find p to minimize:

Error Function

$$Z = \int_{t_0}^{t_f} (y_{msd}(t) - y(p,t))^T W(t)(y_{msd}(t) - y(p,t))dt$$

$s.t.$

Set of Equality Constraints

$$f(\frac{dx}{dt}, x, y, p, v, t) = 0$$

$$x(t_0) = x_0$$

$$h(x, y, p, v) = 0$$

$$g(x, y, p, v) \le 0$$

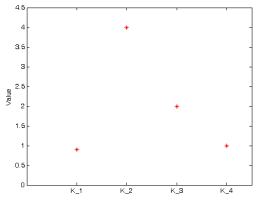
Additional equality/inequality Constraints

$$p^L \le p \le p^U$$

Domain Bounds

21

Optimization Problem



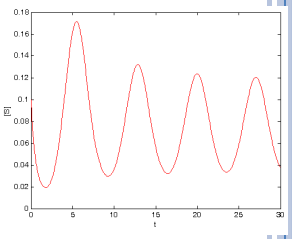
Canonical Parameters

$$\frac{d[S_1]}{dt} = K_1 * [S_2] + K_2 * [S_1]$$

$$\frac{d[S_2]}{dt} = K_3 * [S_2] + \frac{K_4[S_3]}{[S_4]}$$

...

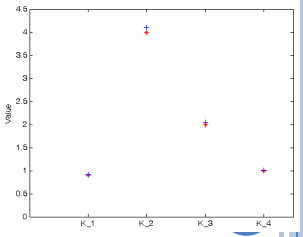
ODE Model



"Measured" Data

Optimization Algorithm (fitting simulated data with "Measured" data)

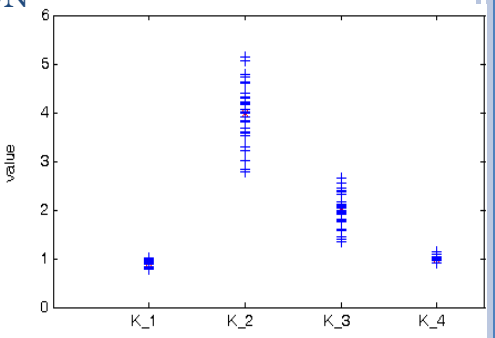
$$Z = \sum_{j=1}^{n} \sum_{i=1}^{m} \frac{[y_{j,pred}(i) - y_{j,\exp}(i)]^2}{[\max(y_{\exp}(i))_j]^2}$$



Estimated Parameters

ROBUST OPTIMIZATION

Monte Carlo Simulation adding error to the "Measured" Data



Ranges of Estimated Parameters

GOAL

Analyze Algorithms that find Smaller ranges of the estimated parameters

23

TESTED ALGORITHMS

- **LSQNonLin**

The function LSQNONLIN of Matlab with the default options of large scale optimization, which uses the subspace trust method based on the Levenberg-Marquardt method to compute the decreasing directions.

- **Direct**

Global search method that applies to Lipschitz continuous function and, after an initial implicit estimate of the Lipschitz constant chooses the potentially optimal rectangles and resamples them across their axis. Subsequently it divides these rectangles and proceed sampling and dividing until a stopping criteria is met.

- **GPS**

Pattern Search algorithm known as Search and Poll algorithms. In the *search* step, any finite set of mesh points can be evaluated. When the search step fails the algorithm calls the *poll* procedure that consists in evaluating the objective function at the neighboring mesh points to see if a lower function value can be found.

24

TESTED ALGORITHMS

- **CMA-ES**

The **C**ovariance **M**atrix **A**daptation **E**volution **S**trategy (CMA-ES) is an evolutionary algorithm for difficult **non-linear non-convex optimization** problems in continuous domain. The CMA-ES is a second order approach and estimates a covariance matrix within an iterative procedure. Adaptation of the covariance matrix amounts is similar to the approximation of the inverse Hessian matrix. Restarts with increasing population size improve the global search performance.

- **DE**

Differential Evolution (DE) was introduced by Storn and Price. DE works as follow: after a random initialization, the objective function is evaluated and the following steps are repeated until a termination condition is satisfied. The crucial idea behind DE is this new scheme for generating trial parameter vectors. DE generates new parameter vectors by adding the weighted difference vector between two population members to a third member. We used the classical version of DE *DE/rand/1*.

25

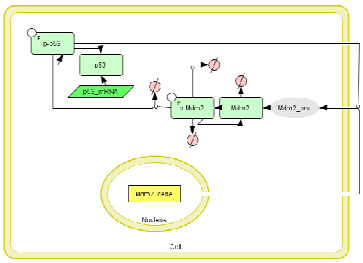
TEST-CASE: P53-MDM2 MODEL

- The negative feedback loop between the tumor suppressor **p53** and the oncogene **Mdm2** is one of the best-studied protein circuits in human cells.
- In the p53 system, ***p53*** transcriptionally activates ***Mdm2***. Mdm2, in turn, negatively regulates p53 by both inhibiting its activity as a transcription factor and by enhancing its degradation rate.
- For different parameters of the feedback loop, the dynamics can show either a monotonic response, damped oscillations, or undamped (sustained) oscillations in which each peak has the same amplitude as the previous peak. The stronger the interactions between the proteins, the more oscillatory the dynamics. Other parameters, such as high basal degradation rates of the proteins, tend to damp out the oscillations.

N. Geva-Zatorsky, N. Rosenfeld, S. Itzkovitz, R. Milo, A. Sigal, E. Dekel, T. Yarnitzky, Y. Liron, P. Polak, G. Lahav and U. Alon. Oscillations and variability in the p53 system, *Molecular Systems Biology* (2006)

26

TEST CASE



$$\frac{d[p53]}{dt} = -\frac{V1\cdot[p53]}{[p53]+kp1} + kdp1\cdot[p-p53] + kf1 + \\ + \frac{kd1\cdot[p53]\cdot[p-Mdm2]}{[p53]+kdeg1}$$

$$\frac{d[p-p53]}{dt} = \frac{V1\cdot[p53]}{[p53]+kp1} - kdp1\cdot[p-p53] + \\ - \frac{kd2\cdot[p-p53]\cdot p-Mdm2}{[p-p53]+kdeg2}$$

$$\frac{d[Mdm2]}{dt} = -\frac{V2\cdot[Mdm2]}{[Mdm2]+kp2} + kdp2\cdot[p-Mdm2] + \\ + kf3\cdot[Mdm2\_pre] - kd3\cdot[Mdm2]$$

$$\frac{d[p-Mdm2]}{dt} = -kd4\cdot[p-Mdm2] + \frac{V2\cdot[Mdm2]}{[Mdm2]+kp2} + \\ - kdp2\cdot[p-Mdm2]$$

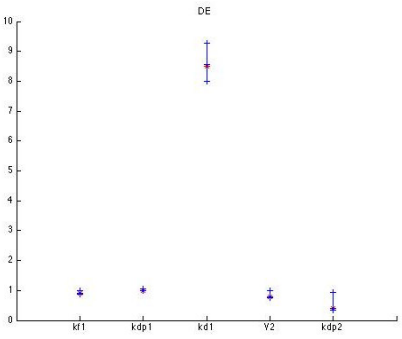
$$\frac{d[Mdm2\_pre]}{dt} = kf2\cdot[p-p53] - kf3\cdot[Mdm2\_pre]$$

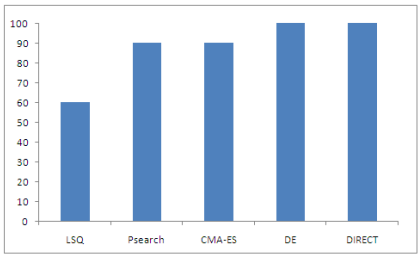
Parameters	Nominal Value
kf1	0. 9
V1	4
Kp1	2
kdp1	1
kd1	8. 5
Kdeg1	0. 1
kd2	0. 85
Kdeg2	0. 01
kf2	1. 1
kf3	0. 8
V2	0. 8
Kp2	0. 2
kdp2	0. 4
kd3	0. 08
kd4	0. 8

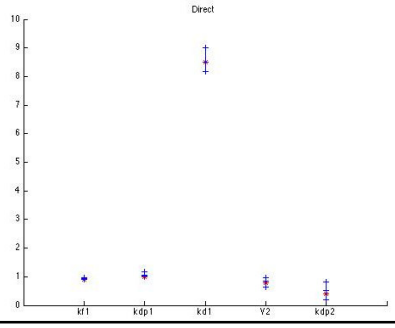
COMPUTATIONAL TOOL

1. Perform a first **global search** on the complete parameter set.
   - We used *LSQNonlin* algorithm which requires less computational effort in terms of CPU time and number of function evaluations but using a multi-start strategy using the method repeatedly, starting from a number of different initial points to avoid to stuck in a local minimum.
2. After defining the value of the whole parameters of the set, **identify** the parameters which are more **critical** for matching the known experimental data.
   - E.g. through *Latin Hypercube* sampling and calculation of *correlation coefficients*.
3. Perform the **robustness** of that parameters throw **Montecarlo** simulation adding noise to measured data.
   - Results demonstrated that **DE** identified the most robust set of parameters and was the most effective in terms of the data fitting.

RESULTS







Results: μ(P) and σ(P)

Algorithms/ Parameters	Kf1	Kdp1	Kd1	V2	Kdp2
Direct	0.93 0.015	1.08 0.054	8.56 0.25	0.81 0.072	0.46 0.148
LSQ	0.67 0.251	1.34 0.471	6.46 2.11	0.73 0.159	0.54 0.181
GPS	0.86 0.142	1.06 0.11	8.04 0.965	0.87 0.124	0.55 0.178
CMA-ES	0.88 0.061	1.04 0.15	8.18 0.86	0.82 0.122	0.5 0.198
DE	**0.9 0.019**	**1 0.01**	**8.52 0.25**	**0.8 0.031**	**0.41 0.074**

CONCLUSIONS

- We presented a methodology inspired by the electronic circuit design to study the most robust optimization algorithms (less sensitive to the noise of the experimental data) for parameter identification that are critical for matching the known experimental data.
- Considering the complete problem of identification of the whole set of parameters LSQnonLin showed the best results in terms of reached object function value and the number of function evaluations even if it is more dependent to the chosen initial search point.
- Using a **Montecarlo** simulation, the evolutionary strategy **DE** and the deterministic method **Direct** are the most robust in the sense that they are less sensitive to the noise of the experimental data. Both Direct and DE showed 100% of success in the identification of the parameters that characterize the curves of the variation of the concentrations over the time accurately with respect to the experimental data.

32

QUESTIONS?

33

A Web Based Tool for Integration of Molecular Pathway Models

34

Outline

- Motivation
- CytoSolve
  - Remote Local Solver
  - Controller
- Tools
- CytoSolve Web Portal
  - Use Case
- Test Case: EGFR Model
- Conclusions and Future Works

35

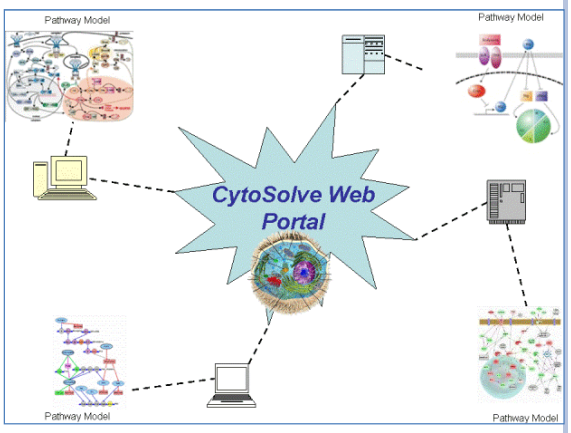
MOTIVATION

- A grand challenge of Systems Biology is to model the whole cell by capturing the quantitative kinetics of interactions between organelles at the molecular mechanistic level to derive descriptions of higher level cellular functions.
- Modeling a cell or a cellular function requires a computational architecture that *integrates* multiple biological pathway models in a *scalable* way, ensuring minimal effort to maintain the resulting integrated model. *Distributed control* allows the maintenance of each model at the local level, not at a central level.

36

GOALS

The goal of this work was to develop a *Web Application* in order to allow users to *simulate* models of biochemical reaction networks as a *distributed* ensemble of biological pathway models and *integrate* computed solutions by building a web GUI to CytoSolve*.



The computational architecture used by Cytosolve is based on the *dynamic messaging approach*. *The dynamic messaging* approach implies that the models remain independently executing programs that interact only by exchanging data via message passing during execution.

37

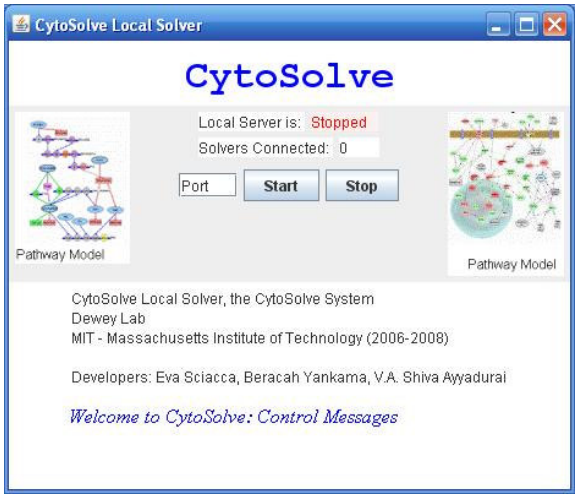
*Ayyadurai, V.A.S., Dewey, C.F., Jr. (2008) Cytosolve: Scalable method for dynamic integration of a distributed ensemble of biological pathway models, *BMC Bioinformatics*, in review

CytoSolve: Local Solver

- The CytoSolve Local Solver is designed for remote SBML model simulations. It can support multiple models on multiple computers distributed across the web.
- The Local Solver can be instructed by a central controller:
  - To simulate a local model over a single time step. After simulation, the service sends back new concentration values calculated by the model.
  - To insert new species concentration values into the model simulation based on the combined results of many models.
- These two instructions allow external control of the global simulation.

38

CytoSolve Local Solver: GUI



39

CytoSolve: Controller

A centralized *Controller* couples multiple SBML models together, sending instructions to the Local Solver.

The Controller of Cytosolve consists of three main components:

- the Monitor serves to track the progress of each pathway's solution time;
- the Communications Manager mediates communication across all pathway models;
- the Mass Balance algorithm provides computational steering by ensuring mass conservation across all integrated models for each time step.

Mass Balance Algorithm

The Mass Balance serves to provide the calculation of species concentration for each time step n across the ensemble of m models. Each model i was treated as a black box with the input and $S_n^{j,i}$ output $S_{n+1}^{j,i}$

The *jth* species concentration of the integrated model in the global vector is calculated as:

$$S_{n+1}^{j,G} = S_n^{j,G} + \sum_{i=1}^{m} (S_{n+1}^{j,i} - S_n^{j,i})$$

41

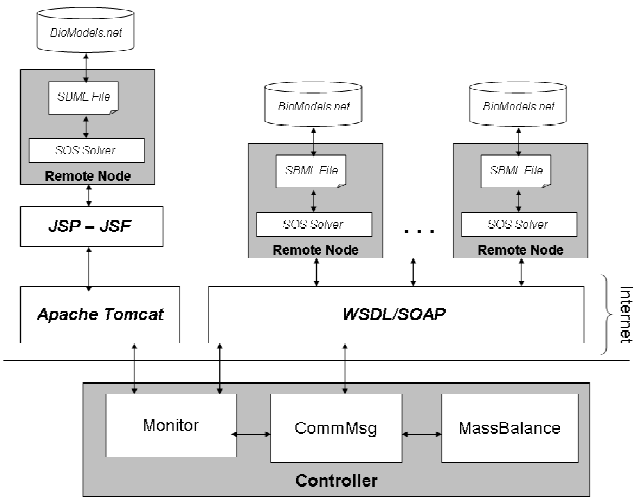
TOOLS

- CytoSolve Local Solver
  - **Interface**: Java
  - **SBML ODE Solver***: C based on CVODE** library
- CytoSolve Web Application
  - **CytoSolve Controller:** Java
  - **Web Application Framework:** Java Server Pages
  - **Integrated Development Environment:** Java Studio Creator/ NetBeans
  - **Application Server:** GlassFish

42

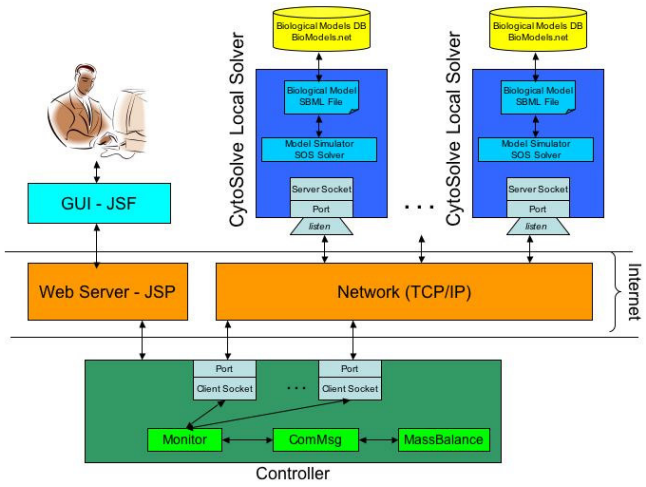
*Machn´e,R., Finney,A., Muller,S., Lu, J., Widder, S. and Flamm,C. (2006) The SBML ODE Solver Library: a native API for symbolic and fast numerical analysis of reaction networks. *Bioinformatics* 22(11),1406-1407
Scott D. Cohen, Alan, C. Hindmarsh **CVODE, a stiff/nonstiff ODE solver (1996) *Computers in Physics*

CYTOSOLVE: PREVIOUS ARCHITECTURE



43

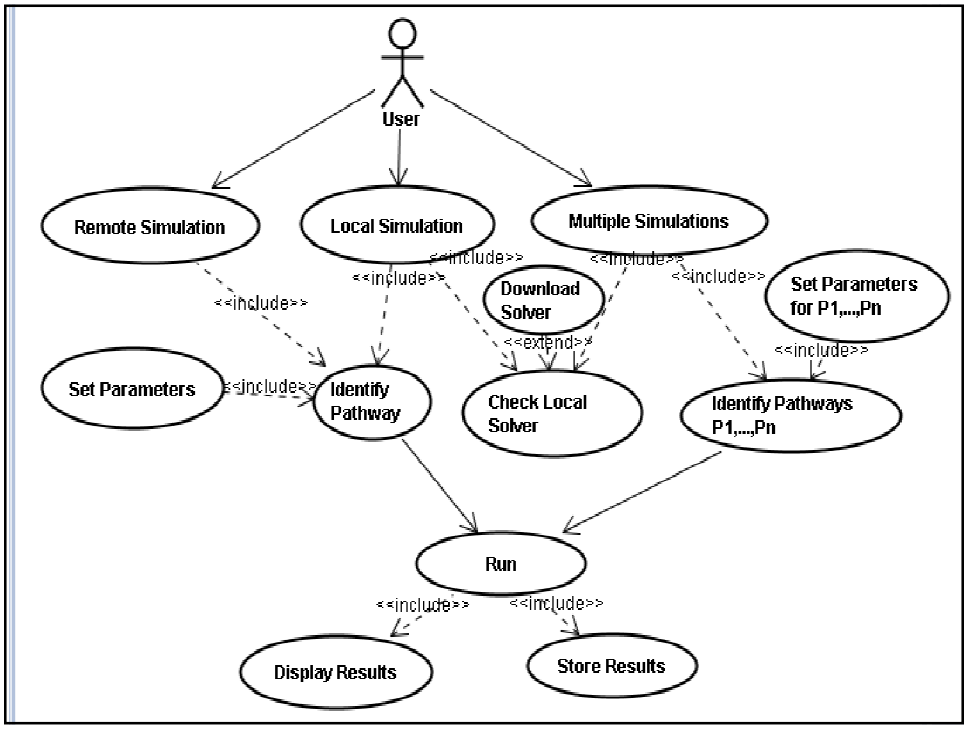
CytoSolve: Current Architecture



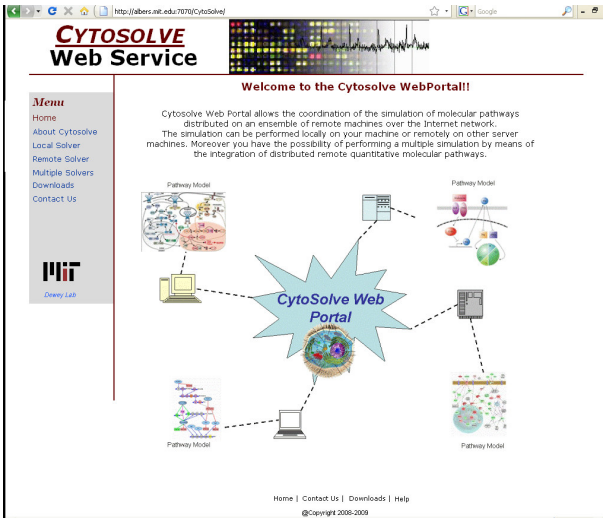
44

CytoSolve Web Portal

1. Simulate a single model on the *local* machine;
2. Simulate a single model on a *remote* machine;
3. Simulate an ensemble of models *remotely* integrating their simulation on the *remote* nodes machines on the network;
4. Simulate a model *locally* integrating this simulation with the other *remote* nodes machines on the network.

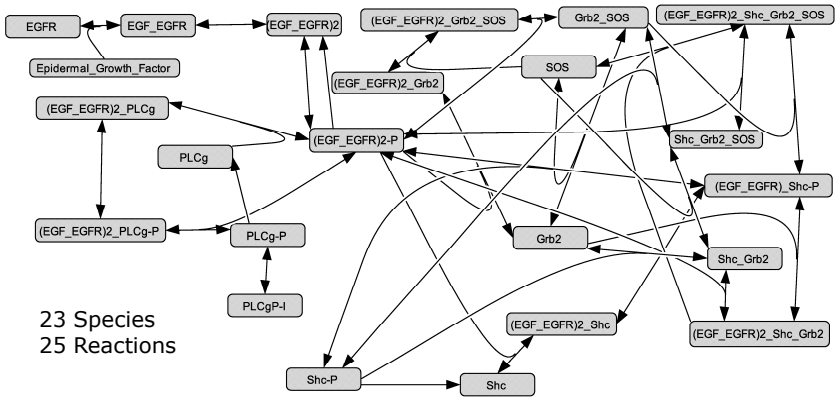


HOME PAGE: HTTP://CYTOSOLVE.MIT.EDU



47

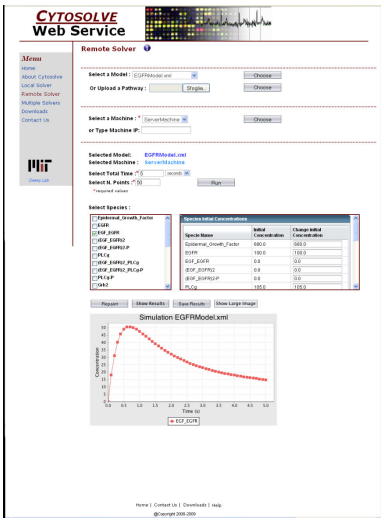
Test Case: EGFR Model*



48

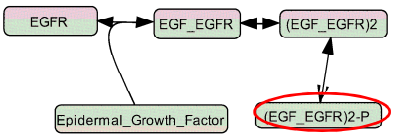
*Kholodenko, B.N., Demin, O.V., Moehren, G. and Hoek, J.B. (1999) Quantification of short term signaling by the epidermal growth factor receptor, *J Biol Chem*

Remote Simulation

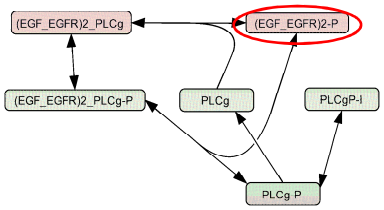


49

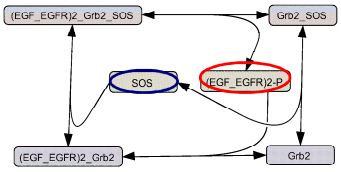
Test Case: EGFR - 4 Models



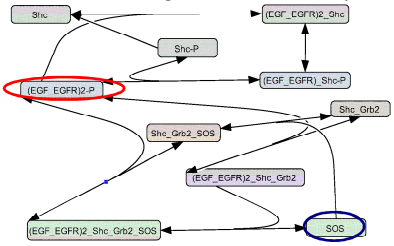
EGFR Dimerization Pathway



PLCg Production Pathway



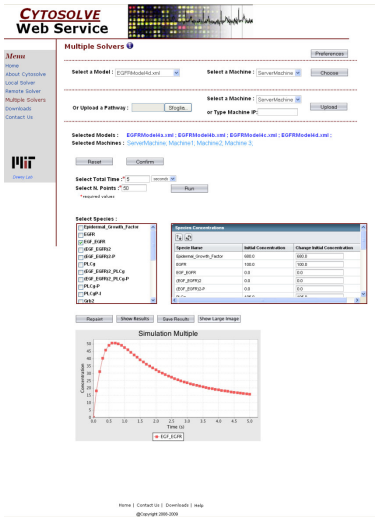
SOS Production Pathway



Shc Production Pathway

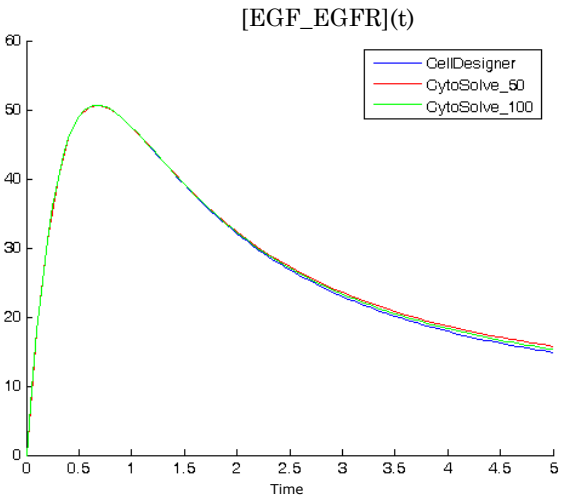
50

Integrating Multiple Simulations



51

RESULTS



err = 0.0176
err = 0.0090

52

CONCLUSIONS

- ***Ease of access:*** The architecture is able to integrate new pathway models.
- ***The multiple platform availability for the simulations:*** The computational system allows models developed on different hardware and computing environments to be integrated.
- ***Open accessibility:*** The computational system supports integration of models across geographical boundaries. The architecture and the implementation of the web application support communication with models anywhere without regard to geographical location.
- ***Ensures protection of proprietary models*** (models where the source code is inaccessible). The black-box simulations of the models allow researchers to integrate Public models with existing Proprietary models to learn some new aspect of science, without violating confidentiality issues.

53

Future Works

- Employ *Ontologies* to manage semantics and species identification across all individual biological pathway models
  - Unique Identifiers for uniquely name biological resources. E.g. *LSID* Life Science Identifiers.
- Analyzing the size of the time step.
  - Adaptive time stepping at the Controller level to observe the time scales of different models and invoke them only when necessary.
- Testing with N other models. (N ~ 10^3?)
- Support CellML and other descriptions.

54

Questions?

55

Reliable Biological Circuit Design Including Uncertain Kinetic Parameters

56

Motivation

Generally, the design frameworks that support Research & Development technologies must face with *uncertainty* modeling. In biological field the uncertainty arise from different sources related to:

- behavioral models, that connect uncertain model parameters to observed evolution of system state;
- equivalent models, that connect uncertain model parameters to system feedbacks;
- approximated models, that approximate various aspects of a system in a computational tractable manner.

57

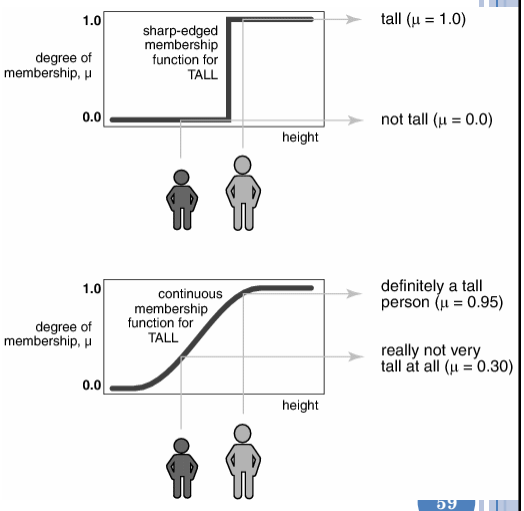
Introduction

- The biological systems is investigated in terms of *performances* or *specifications* of the system.
- The analysis of the biological system is based on the idea that the set of parameters involved in the model can be classified into two different typologies: the *uncertain kinetic parameters* and the *control design parameters*.
- To take into account the uncertainties arising from the estimations of the kinetic parameters, the function representing the feedback of the system is *fuzzified* and a *measure of failure* of the designed biological circuit is minimized to reach the required performance.
- The methodology set up the design parameter values to balance the uncertainty of the kinetic parameters.

58

Fuzzy Sets

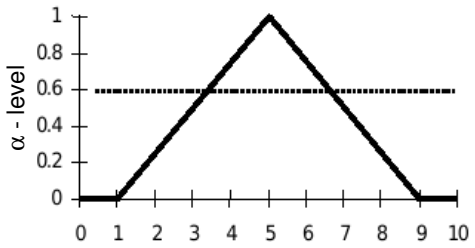
- *Fuzzy sets* have been introduced by [Zadeh, 1965] as an extension of the classical notion of set. In classical set theory, the membership of elements in a set is assessed in binary terms. Fuzzy set theory permits the gradual assessment (in the real unit set [0;1]) of the *membership* of elements



59

Fuzzy Number

- A *fuzzy number* is a convex, normalized fuzzy set.



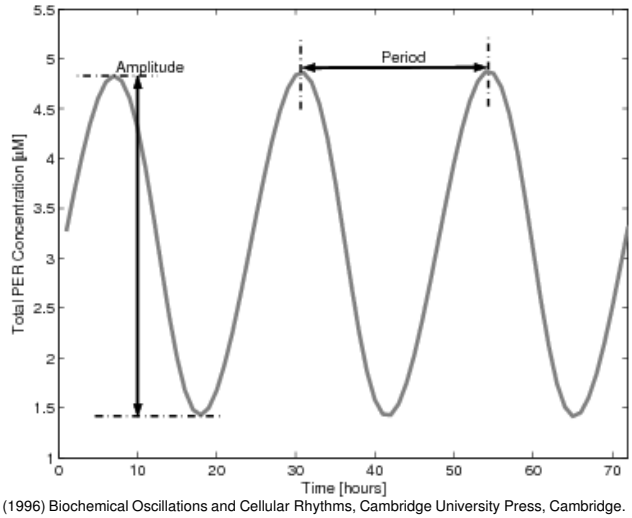
60

Parameters

- **Uncertain kinetic parameters K_U**, that are known in terms of confidence intervals. Those parameters are for example the Michaelis constants or the constant rates for the kinases and the phosphatases.
- **Control design parameters K_{CD}**, that can be defined in a reliable way and determine the behavior of the biological system. Those are for example the maximum rate of degradation of an enzyme or the first order transportation rate parameters.

61

PERFORMANCES



Goldbeter, A. (1996) Biochemical Oscillations and Cellular Rhythms, Cambridge University Press, Cambridge.

62

PERFORMANCES

Performances can be defined in functional terms as:

$$P = y(K_{CD}, K_U,)$$

The uncertainty concerning the uncertain kinetic parameters compel to consider these parameters as *fuzzy numbers* and to interpret the previous formulation as:

$$\tilde{P} = \tilde{y}(K_{CD}, \tilde{K}_U)$$

Which is the fuzzy representation of the performances and only the controllable parameters are considered crisp.

63

APPROXIMATION

- Designing for uncertainty is computationally expensive.
- Approximation methods reduce the high computational cost associated with designing for uncertainty by using approximations.
- Linearization:

$$\overline{P} = a + B_U K_U$$

- Response Surface:

$$\overline{P} = a + \sum_{i=1}^{n} \sum_{j=1}^{h_i} (b_{ij} K_i^j) + \sum_{q=1}^{m} c_q \prod_{i=1}^{n} K_i^{p_{iq}}$$

64

FUZZYFICATION

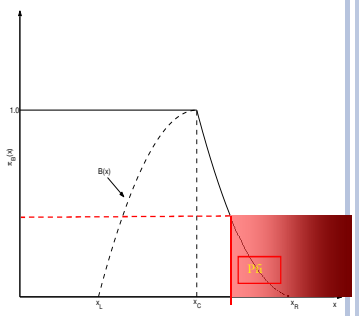
- The fuzzy representation of the performance is constructed *enveloping* the fitted data by intervals.
- The fuzzy map is built by α-level considering the minimum *median interval* which envelopes a fraction (1- α) of the performance values.

65

POSSIBILITY MEASURE OF FAILURE

In order to deal with design specifications it is necessary to *compare* the *fuzzy numbers* representing the performances with crisp numbers representing the *design constraints* and give a measure of satisfaction of these constraints.

For this purpose the *possibility measure of failure* with respect to the specification constraints can give useful information to improve the design.

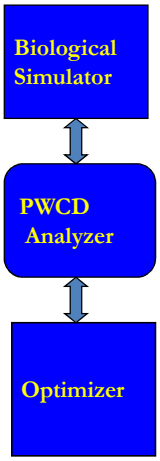


The possibility measure of failure P with respect to the specification constraints P_f ($P \geq P_f$):

$$\Pi_P([P_f,+\infty]) = \sup_{y \geq P_f}(P)(y)$$

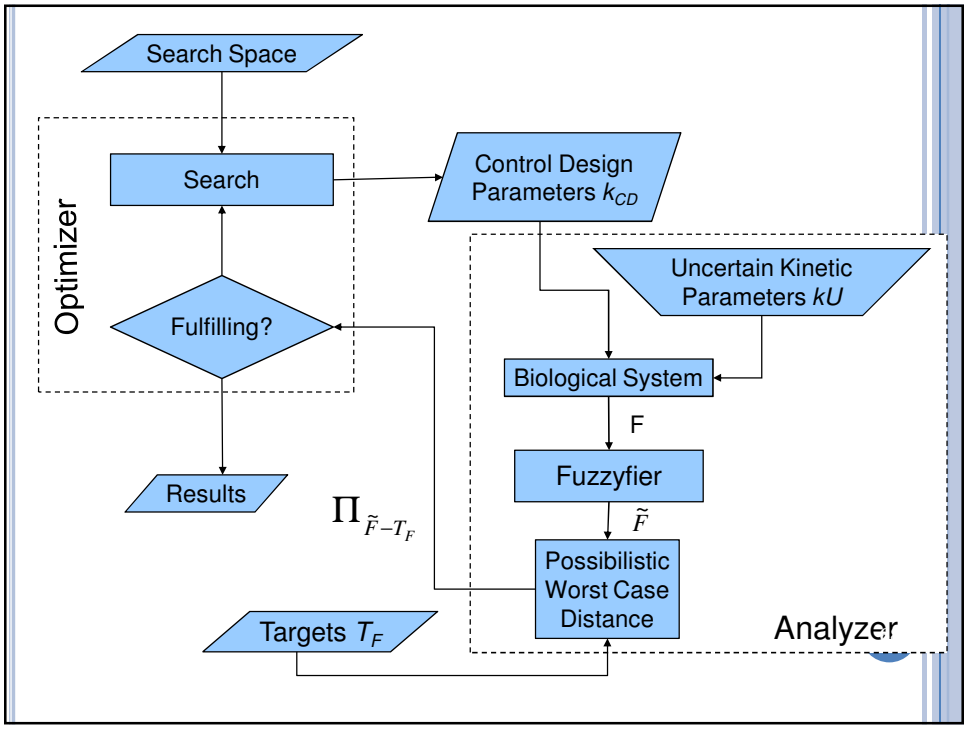
Dubois, D. and Prade, H. (1988) Possibility theory: An approach to computerized processing of uncertainty. New York: Plenum Press.

POSSIBILISTIC WORST CASE DISTANCE

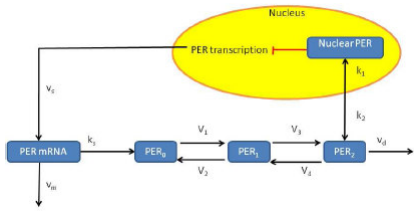


- In order to model the uncertainties arising from the circuit simulation of the performance, the *fuzzy set theory* has been used.
- A *response surface* of the biological circuit performances as function of kinetic parameters has been fitted as suitable *approximation* in a finite range.
- By means of the approximation the performances were fuzzyfied and it was computed *the possibility measures of performances failure*
- The function to *optimize* was the sum of possibilities measure of the performances failure.

$$\sum_{i \in I} \Pi_{(P_i - P_{fi})}([0,+\infty))$$



DROSOPHILA PERIOD PROTEIN CASE STUDY



$$\frac{dM}{dt} = v_s \frac{K_I^n}{K_I^n + P_N^n} - v_m \frac{M}{K_m + M}$$

$$\frac{dP_0}{dt} = k_s M - V_1 \frac{P_0}{K_1 + P_0} + V_2 \frac{P_1}{K_2 + P_1}$$

$$\frac{dP_1}{dt} = V_1 \frac{P_0}{K_1 + P_0} - V_2 \frac{P_1}{K_2 + P_1} - V_3 \frac{P_1}{K_3 + P_1} + V_4 \frac{P_2}{K_4 + P_2}$$

$$\frac{dP_2}{dt} = V_3 \frac{P_1}{K_3 + P_1} - V_4 \frac{P_2}{K_4 + P_2} - k_1 P_2 + k_2 P_N - v_d \frac{P_2}{K_d + P_2}$$

$$\frac{dP_N}{dt} = k_1 P_2 - k_2 P_N$$

v_s is the rate of mRNA production,

v_m is the rate of mRNA degradation,

K_I is mRNA repression constant,

K_m is the Michaelis constant for the mRNA, degradation,

k_s is the rate of PER production,

V_i is the rate for the kinase,

K_i is the rate for the phosphatase,

v_d is the rate of degradation of the bisphosphorylated PER form,

k_1 is the transportation rate of the bisphosphorylated PER form in the nucleus,

k_2 is the transportation rate of bisphosphorylated nuclear PER form in the cytosol,

K_d is the Michaelis constant for the degradation of bisphosphorylated PER form

69

Goldbeter, A. (1995) A model for circadian oscillations in the Drosophila period protein (PER) *Proc. R. Soc. Lond. B, 261 , pp. 319-324.*

PERFORMANCES

In this case study, the target performances of the required design problem are:

- *the period (measured in hours), and*
- *the amplitude (measured in μmol)*

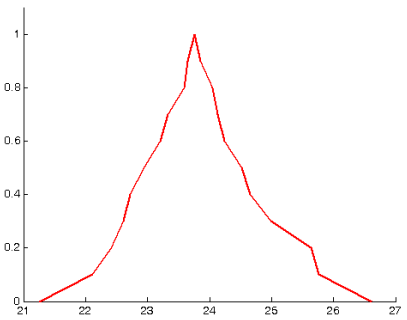
of the concentration of the total quantity of the PER protein (PT) which is given by:

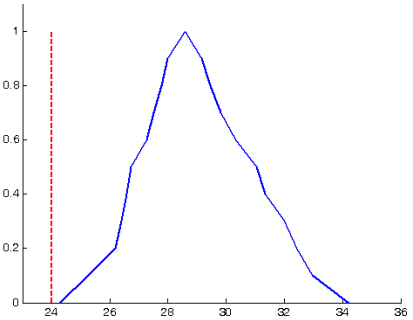
$$PT = P0+P1+P2+PN$$

These performances are optimized by the methodology and they are expressed in terms of possibility of failure. In this particular test case, the minimum threshold for the period of the PER protein oscillations is fixed to 24 hours while the minimum threshold of the amplitude is fixed to a rather large value in order to guarantee significant oscillations.

70

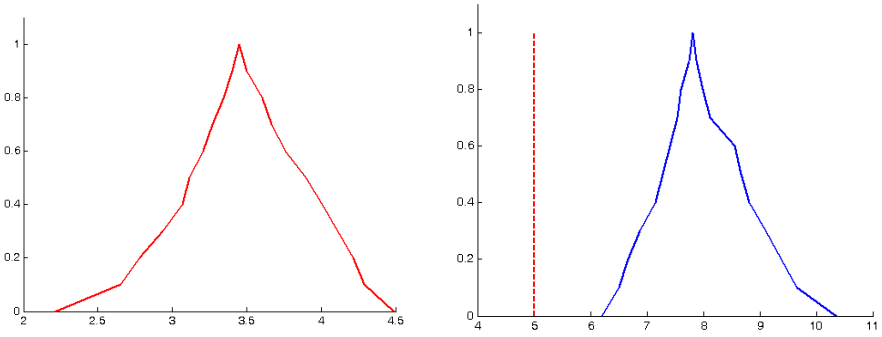
RESULTS: PERIOD





71

RESULTS: AMPLITUDE



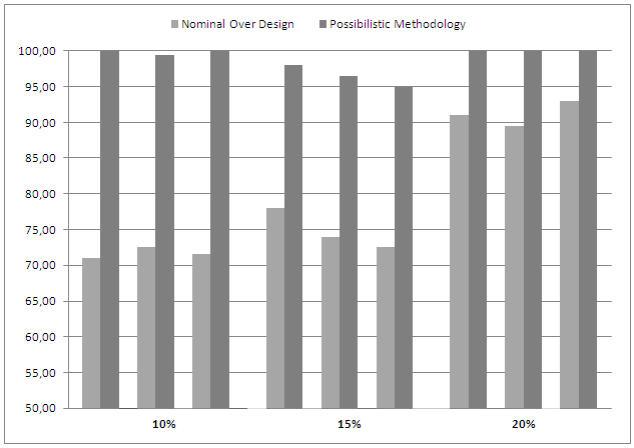
72

COMPARISON BETWEEN STANDARD DESIGN METHODOLOGY AND PWCD

- We made the comparison between Possibilistic Worst-Case Distance methodology and the most common design methodology named "Nominal Over-Design".
- The *Nominal Over-Design* methodology sets every objective to a secure value with respect to the nominal specifications. In this test case the objectives were increased of a 10% with regard to the minimum thresholds.

73

COMPARISON



74

QUESTIONS?

75

PUBLICATIONS

- NICOSIA G, SCIACCA E. (2008). **Robust Parameter Identification for Biological Circuit Calibration**. In: IEEE BIBE 2008 Proceedings, p. 1-6
- SCIACCA E., V. A. SHIVA AYYADURAI, C. FORBES DEWEY (2008). **A Web Based Tool for Integration of Molecular Pathway Models**. In: IEEE BIBE 2008, p. 1-6
- SCIACCA E., V. AYYADURAI, C. F. DEWEY (2008). **A Web Based Application for the Integration of Quantitative Molecular Pathway Simulations**. In: BMES 2008. St. Louis USA, 1-4 Ottobre, p. 1
- SCIACCA E., SPINELLA S. (2009). **Reliable Biological Circuit Design Including Uncertain Kinetic Parameters.** In: Fuzzy Optimization: Recent Developments and Applications. Springer. In press.