Applying the Hybridization Approach to Biological Models

Thao Dang

VERIMAG, CNRS (France)

Joint work with Colas Leguernic, Oded Maler, Romain Testylier

Modeling and Analysis of biological systems

Goal of this work: investigating the application of formal methods to biological systems

Mitochondria Theory of Aging

Mitochondria

- Generate the majority of the cellular ATP
- Produce reactive oxygen species that damage proteins, membranes and the mitochondrial DNA (mtDNA)

Damages impair ATP production but not replication of mtDNA

How defective mitochondria might accumulate?

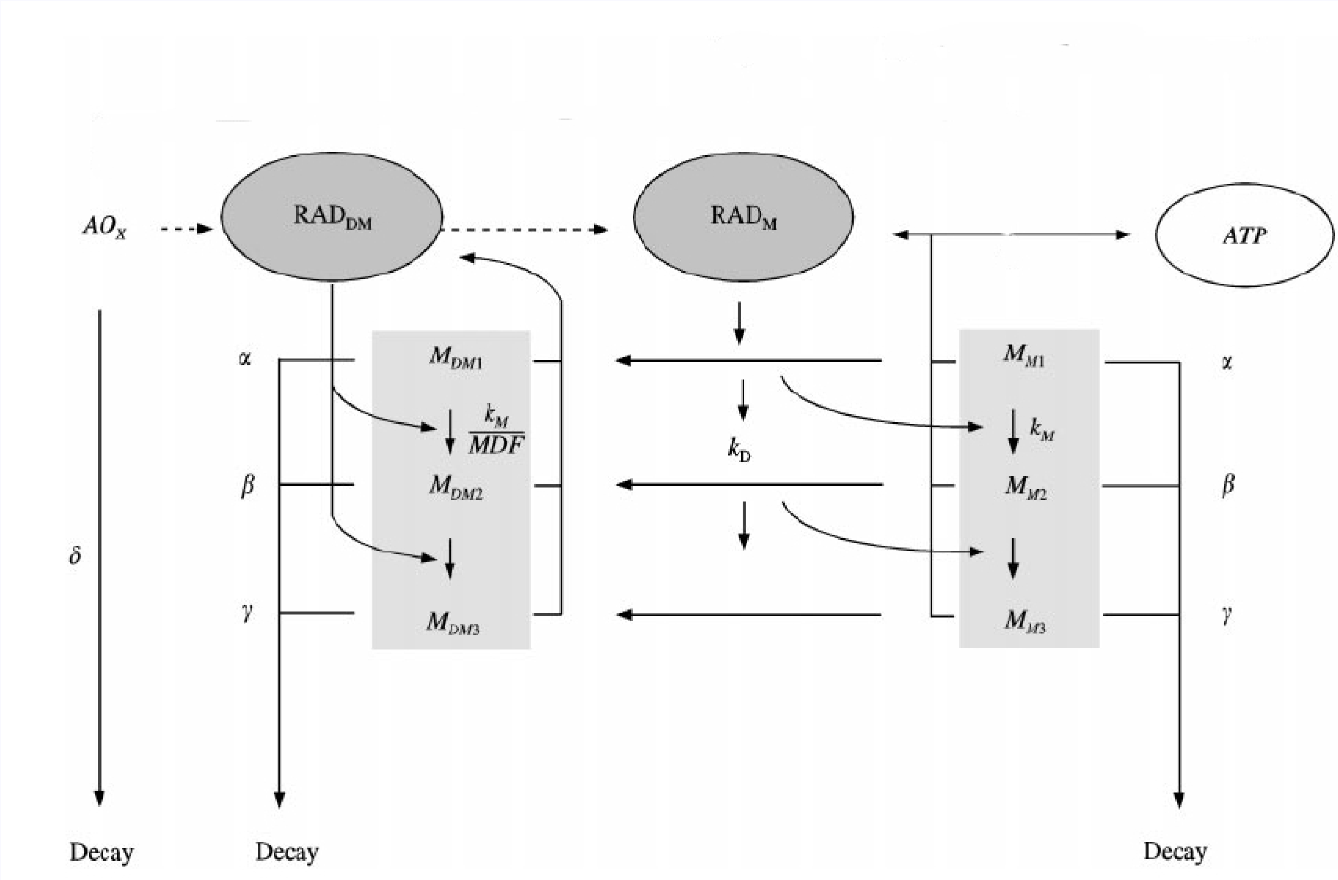
"Survival of the slowest" hypothesis [Grey 1997]:

- Accumulation by lowering degradation rate
- Degradation depends on membrane damage
- Decreased respiratory activity ⇒ Inflict membrane damage at a slower rate

A mathematical model proposed by [Kowald and Kirkwood 2000] to examine this hypothesis

Mathematical Model [Kowald and Kirkwood 2000]

Additional hypothesis: defective mitochondria have a reduced growth rate.



Mathematical Model [Kowald and Kirkwood 2000]

γ, β, α such that $\gamma > \beta > \alpha$ are the decay rates

9 differential equations: 6 equations are for the various types of mitochondria, 1 for the level of antioxidant, 1 for the ATP level and one for the amount of radicals.

Variables M_{Mi} and M_{DMi}: populations of intact and damaged mitochondria. $i \in \{1, 2, 3\}$ level of membrane damage.

Variables Rad_M and Rad_{DM}: radical concentrations in intact and damaged mitochondria, related by RDF (radical difference factor).

Rate k_M of moving to a higher membrane damage class, a rate k_D of converting intact into defective mitochondria.

The model also contains a generic antioxidant species (AOx) that destroys radicals.

Synthesis Rate

Synthesis rate controlled by the cellular energy level, modeled using "artificial promoter"

$$\frac{k_1}{1 + (ATP/ATP_c)^n}$$

(constant n modelling how sensitive the promoter to deviation from control parameter $ATPc$)

Synthesis rate has an upper limit k_1

Synthesis requires energy $\Rightarrow$ Synthesis rate depends on ATP concentration

$$\frac{k_1}{1 + (ATP/ATP_c)^n} \frac{ATP}{ATP + ATP_c}$$

Growth disadvantages are different for each class of defective mitochondria

$$\frac{dM_{M1}}{dt} = S\, M_{M1} + \frac{2S}{GDF+1} M_{M2} - (\alpha + (k_M + k_D)\, Rad_M)\, M_{M1} \tag{8}$$

$$\frac{dM_{M2}}{dt} = -\frac{2S}{GDF+1} M_{M2} + \frac{2S}{GDF} M_{M3} + k_M Rad_M\, M_{M1} - (\beta + (k_M + k_D)\, Rad_M)\, M_{M2} \tag{9}$$

$$\frac{dM_{M3}}{dt} = -\frac{2S}{GDF} M_{M3} + k_M\, Rad_M\, M_{M2} - (\gamma + k_D\, Rad_M)\, M_{M3} \tag{10}$$

$$\frac{dM_{DM1}}{dt} = \frac{S}{GDF} (M_{DM1} + M_{DM2}) + k_D\, Rad_M\, M_{M1} - \left(\alpha + k_M \frac{RDF}{MDF}\, Rad_M\right) M_{DM1} \tag{11}$$

$$\frac{dM_{DM2}}{dt} = -\frac{S}{GDF} M_{DM2} + \frac{2S}{GDF} M_{DM3} + k_D Rad_M\, M_{M2} + k_M \frac{RDF}{MDF}\, Rad_M\, M_{DM1} - \left(\beta + k_M \frac{RDF}{MDF}\, Rad_M\, M_{DM2}\right) \tag{12}$$

$$\frac{dM_{DM3}}{dt} = -\frac{2S}{GDF} M_{DM3} + k_D\, Rad_M\, M_{M3} + k_M \frac{RDF}{MDF}\, Rad_M\, M_{DM2} - \gamma\, M_{DM3} \tag{13}$$

$$\frac{dAOx}{dt} = \frac{ATP}{ATP + ATP_c} \frac{k_2}{1 + \left[\frac{PAOx}{Rad_M (M_{M1} + M_{M2} + M_{M3}) + RDF\, Rad_M\, (M_{DM1} + M_{DM2} + M_{DM3})}\right]^3} - \delta\, AO_x \tag{14}$$

$$\frac{dRad_M}{dt} = k_R - k_3 \frac{AO_x\, Rad_M}{M_{M1} + M_{M2} + M_{M3} + M_{DM1} + M_{DM2} + M_{DM3}} \tag{15}$$

$$\frac{dATP}{dt} = k_{ATP}\, M_{M1} + \tfrac{1}{2}\, k_{ATP}\, M_{M2} - \frac{ATP}{ATP + ATP_c} \left[\frac{k_{EM}\, k_1}{1 + \left[\frac{ATP}{ATP_c}\right]^3} + k_{EC} + \frac{k_{EP}\, k_2}{1 + \left[\frac{PAOx}{Rad_M\, (M_{M1} + M_{M2} + M_{M3}) + RDF\, Rad_M\, (M_{DM1} + M_{DM2} + M_{DM3})}\right]^3}\right] \tag{16}$$

$$S = \frac{ATP}{ATP + ATP_c} \frac{k_1}{1 + \left[\frac{ATP}{ATP_c}\right]^3} \frac{1}{M_{M1} + \frac{2}{GDF+1}\, M_{M2} + \frac{1}{GDF}\, (M_{M3} + M_{DM1} + M_{DM2} + M_{DM3})} \tag{17}$$

Hypothesis Validation

Analysis Problem: studying the influence of the turnover rate and initial situations on the stability of the system.

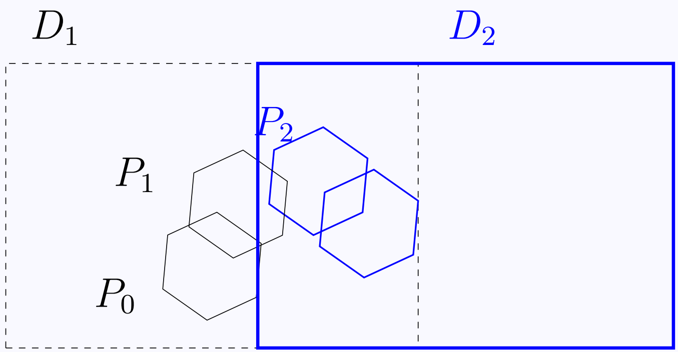
- Numerical solution can only approximate single solutions
- Reachability computation can characterize sets of all possible solutions. Systems with non-linear dynamics remain a challenging problem.

Hybridization Approach

$$\dot{x} = f(x), \quad x \in \mathcal{X}, f \text{ is Lipschitz}$$

Principle

- Complex system (difficult to analyse) $\longrightarrow$ piecewise less complex system (easier to analyse)
- In this work, we use different affine dynamics in different approximation domains
- Control of dynamics approximation error $\longrightarrow$ Accuracy of trajectory approximation



Approximation Domain Construction

- Given a desired error bound ϵ, compute a domain D such that
  - $|f(x) - l(x)| \leq \epsilon, x \in D$
  - Large domains $\longrightarrow$ less frequent domain construction

- The accuracy of dynamics approximation is important (in hybrid systems, the problem of spurious trajectories can be aggravated by discrete transitions)

Interpolation over Simplicial Approximation Domains

$$\dot{x}(t) = f(x(t)), x(t) \in \mathbb{R}^n$$

Approximate Dynamics

- $\dot{x}(t) = l(x(t)) + u(t)$
  - Affine function: $l(x(t)) = Ax(t) + b$
  - Input: $||u(\cdot)|| \leq \mu$ such that $\forall x \in \Delta$ $||f(x) - l(x)|| \leq \mu$

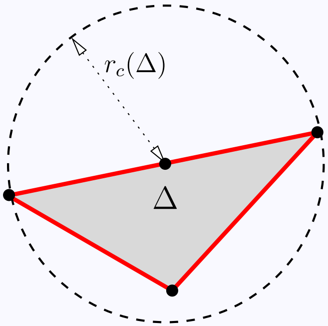
Interpolation Over a Simplex Δ

- Interpolation: $l(v_i) = f(v_i)$, for all vertices v_i of Δ

Interpolation Error

For all $x \in \Delta$, $||f(x) - l(x)|| \leq \delta_\Delta \frac{r_c^2(\Delta)}{2}$.

- δ_Δ is the maximal curvature of f in Δ
- $r_c(\Delta)$ is the radius of the smallest ball containing the simplex Δ.



min-containment circle

- By exploiting the curvature of $f(x)$ we can compute a larger simplex that guarantee the same error bound

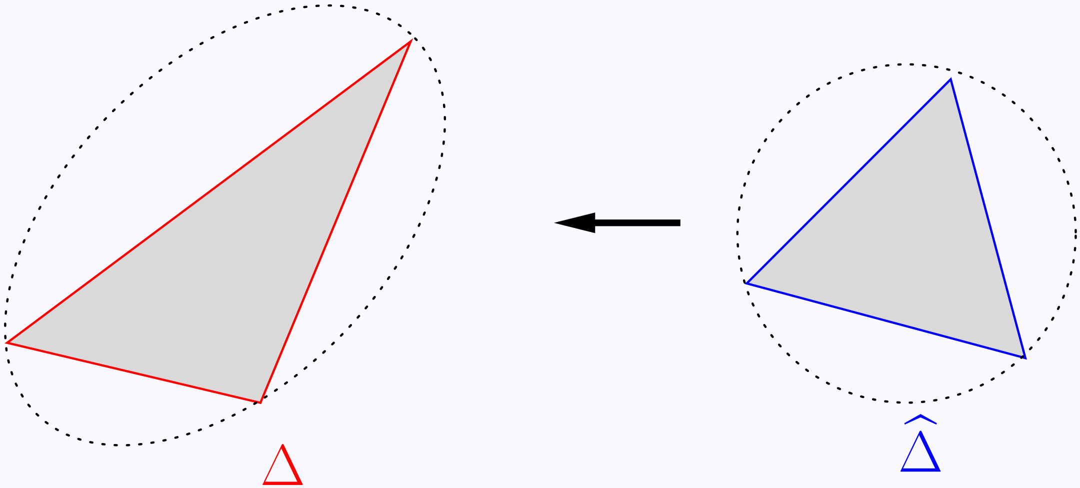
Curvature

$f_i(x)$ is the i^{th} component of the vector of functions f
Curvature

- Hessian matrix $H^i(x)$ associated with each f_i is a matrix whose element $H^i_{jk}(x) = \frac{\partial^2 f_i}{\partial x_j x_k}$

- For a unit vector v, the curvature of f_i along the direction v is $v^T H^i(x) v$.

Exploiting Curvature



Observation

$$||f(x) - l(x)|| \leq \delta_\Delta \frac{r_c^2(\widehat{\Delta})}{2}$$

When the largest curvature in one direction is much greater than the largest curvature in another $\Rightarrow$ **Shrink** along the directions with **small curvatures**

“Isotropic” Mapping

Curvature Bound Matrix C
$\forall i \in \{1, \ldots, n\}\ \forall x \in \Delta; \forall v \in \mathbb{R}^n$

$$||v|| = 1 \wedge |v^T H^i(x) v| \leq v^T C v.$$

Transformation Ω matrix formed by eigenvectors of C, ξ_i eigenvalues of C

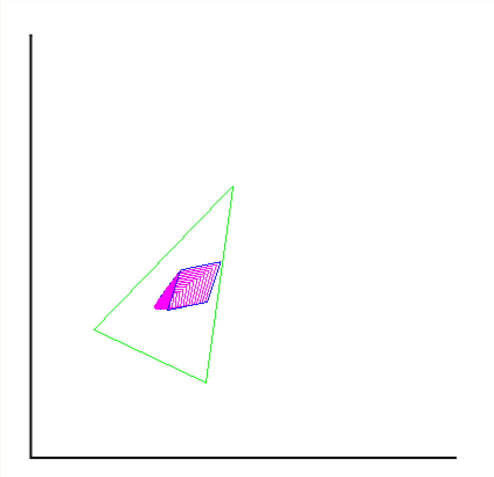
$$T = \Omega \begin{pmatrix} \sqrt{\xi_1/\xi_{max}} & 0 & \ldots & 0 \\ 0 & \sqrt{\xi_2/\xi_{max}} & \ldots & 0 \\ & & \ldots & \\ 0 & & \ldots & \sqrt{\xi_n/\xi_{max}} \end{pmatrix} \Omega^T.$$

Summary of Recent Results

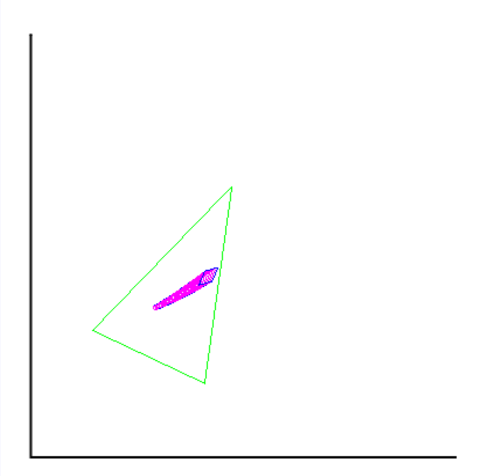
Methods for computing C for systems with non-constant Hessian matrices

Optimal Domains for a Class of Quadratic Systems

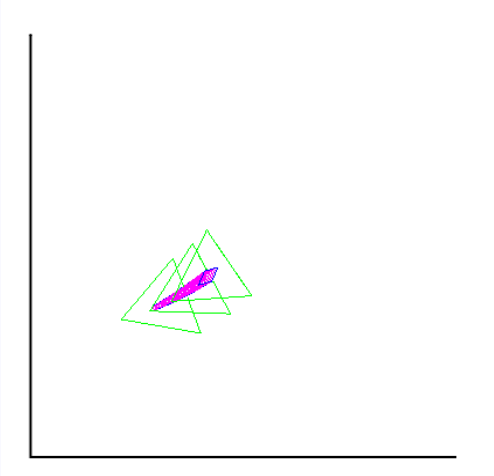
Large Error Bound



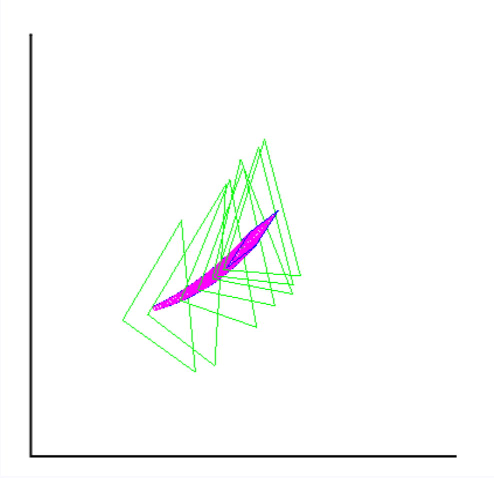
Small Error Bound



Regular Simplex

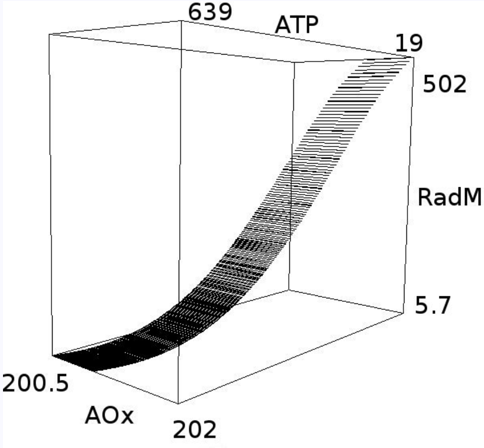


Isotropic Transformation



Aging Model

Reachability computation results are consistent with the simulation results: With (normalized) turnover rate too small (≤ 0.6) or too high (> 11) the system is unstable



The computation time for 1000 iterations is 23.3 minutes (for standard turnover rate).

Lac Operon

Lac Operon: biochemical feedback mechanism through which the bacterium *E. Coli* adapts to the lack of Glucose in its environment by switching to a Lactose diet.

$$
\begin{aligned}
\dot{R}_a &= \tau - \mu * R_a - k_2 R_a O_f + k_{-2}(\chi - O_f) - k_3 R_a I_i^2 + k_8 R_i G^2 \\
\dot{O}_f &= -k_2 r_a O_f + k_{-2}(\chi - O_f) \\
\dot{E} &= \nu k_4 O_f - k_7 E \\
\dot{M} &= \nu k_4 O_f - k_6 M \\
\dot{I}_i &= -2k_3 R_a I_i^2 + 2k_{-3} F_1 + k_5 I_r M - k_{-5} I_i M - k_9 I_i E \\
\dot{G} &= -2k_8 R_i G^2 + 2k_{-8} R_a + k_9 I_i E
\end{aligned}
$$

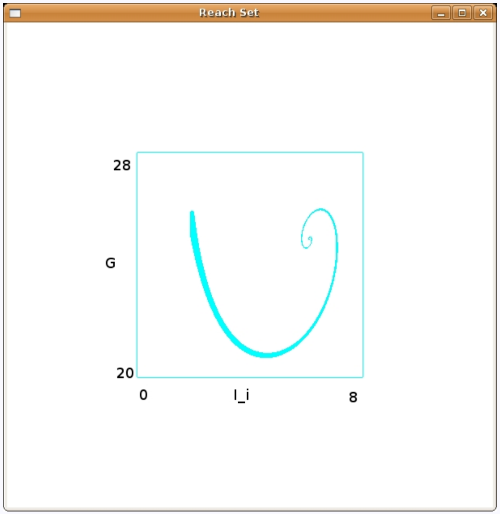
Variables denote the concentrations of different reactants, such as R_a (active repressor) O_f (free operator), E (enzyme), M (mRNA), I_i (internal inducer), and G (glucose).

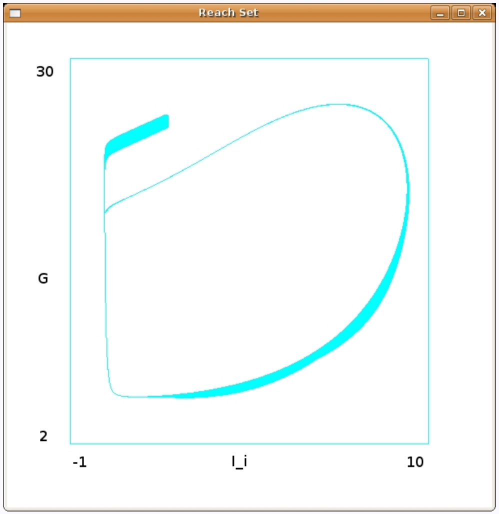
Lac Operon

We studied the behavior of this 6-dimensional system around a quasi-steady state for the first 4 variables.

Initial states in $I_i \in [1.9, 2.0]$ and $G \in [25.9, 26]$.

When $k_{-1} = 2.0$ the system exhibits a stable focus and when $k_{-1} = 0.008$ the system exhibits a limit cycle. Computation times are 3 and 5 minutes, respectively.

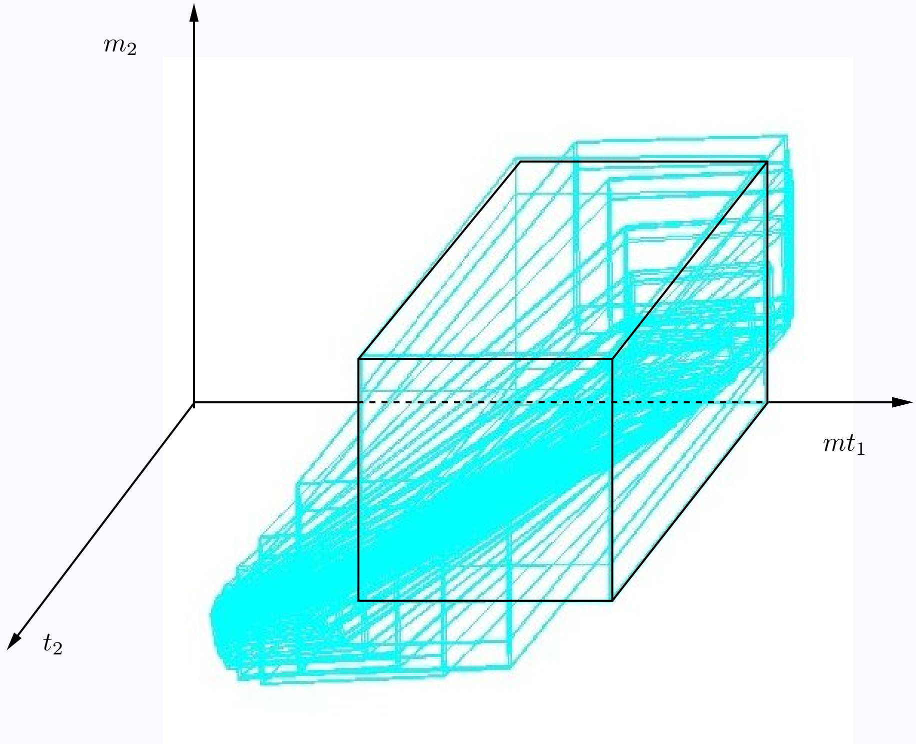




Blood vessels

A biochemical network [Karagiannis, Popel 2002] modelling the loosening of the extra-cellular matrix around blood vessels.

System of **quadratic differential equations with** 12 **variables**



Thank You

Reachability analysis methods

Direct methods

- Track the evolution of the reachable set under the flow of the system. Various set representations: *e.g.* polyhedra, zonotopes, ellipsoids, level sets
- Exact results, or accurate approximations with error bounds. Using symbolic or numerical computations

Indirect methods

- Abstraction methods: reducing to a simpler system that preserves the property (*e.g.* [Tiwari & Khanna 02; Alur et al. 02; Clarke et al. 03])
- Achieve a proof of the property without computing the reachable set: *e.g.* Barrier certificates [Prajna & Jadbabaie04], polynomial invariants [Tiwari & Khanna 04].

⋆ **Scalability** is still challenging (complexity and size of real-life systems)