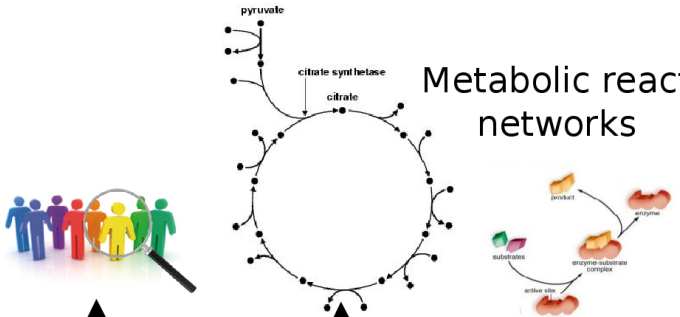
Parameter synthesis for biological systems modelling

Thao Dang, Eric Fanchon, Marcelo Forets, Victor Magron, Alexandre Rocca

VERIMAG/CNRS and TIMC-IMAG, Grenoble Université Grenoble Alpes, France
France

SynCop and PV
14-15 April 2018, Thessaloniki, Greece

Biological Modelling

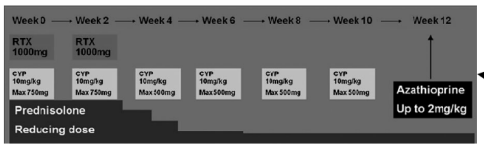


Metabolic reactions networks

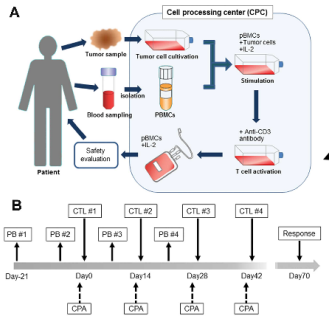
Population studies

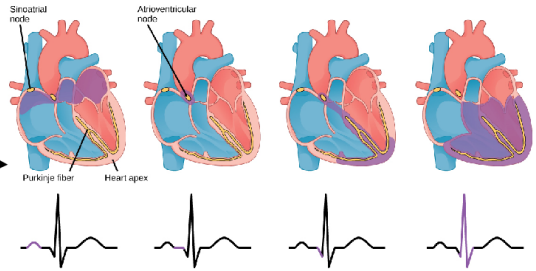
and more . . .

Biological modelling

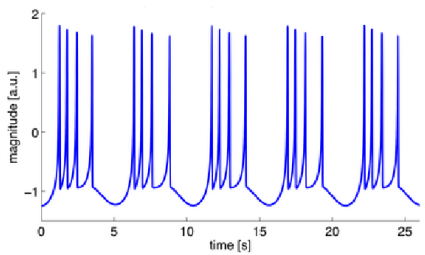


Experimental protocols



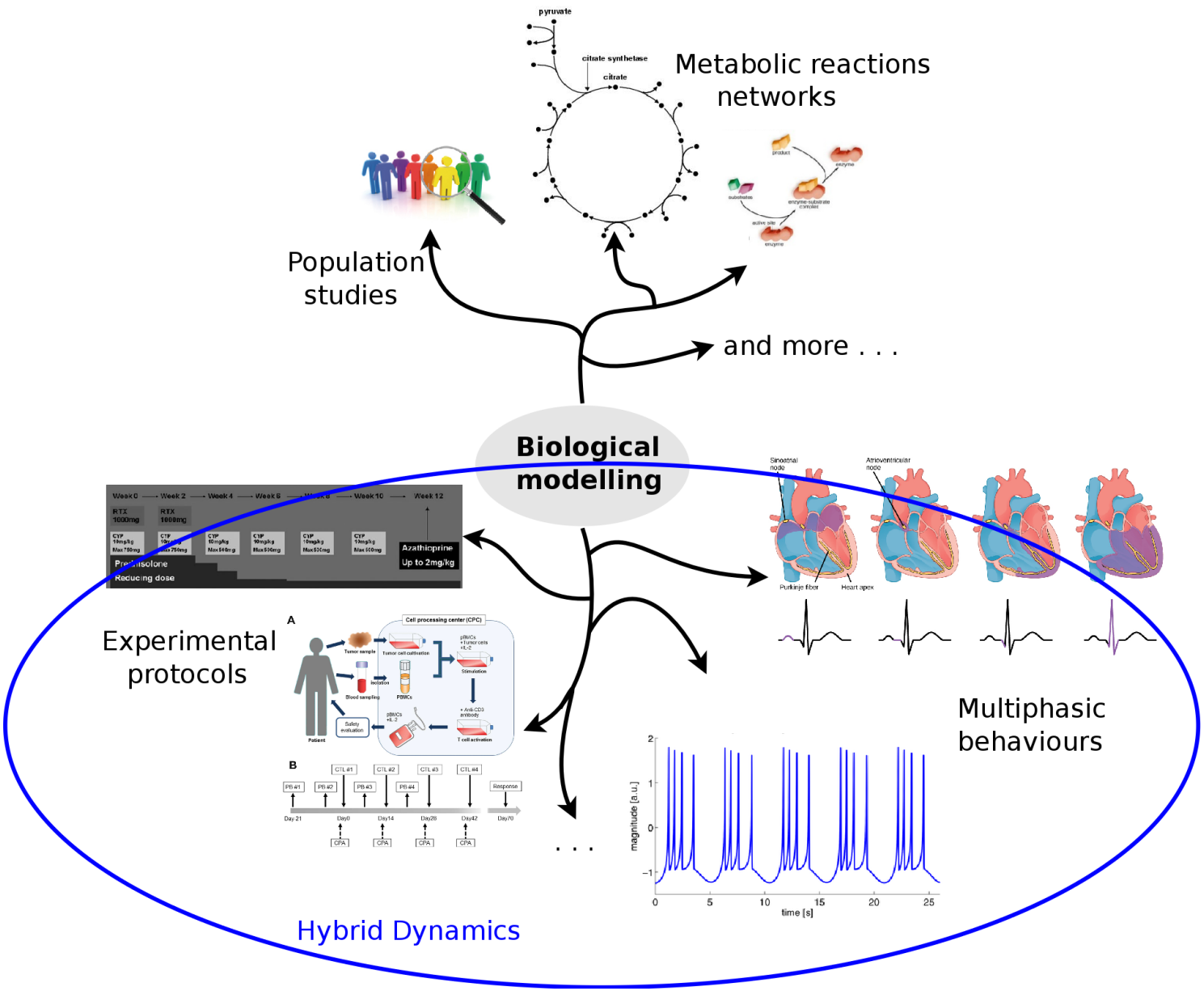


Multiphasic behaviours

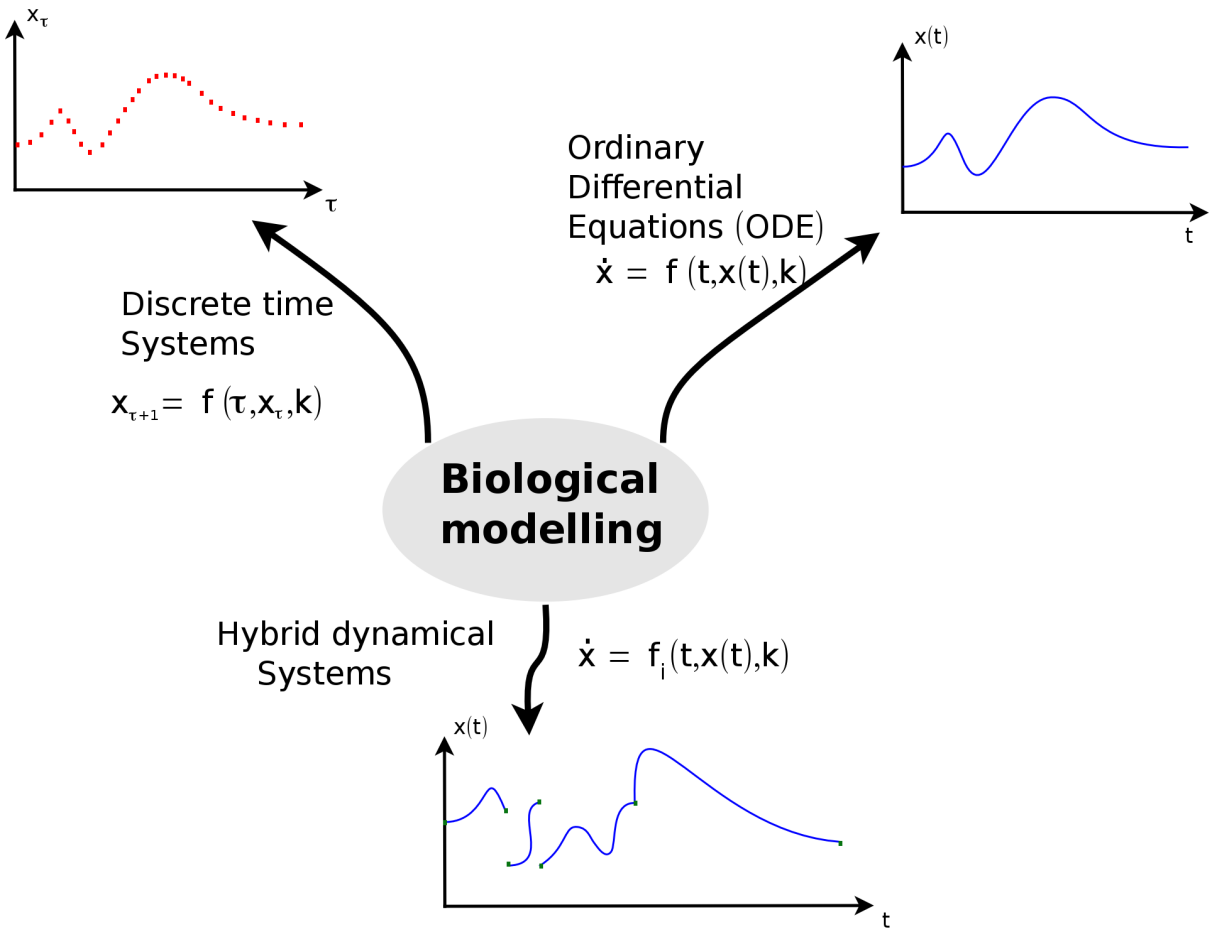


. . .

Biological Modelling



Biological Modelling



Evolution over time of certain quantities (e.g. concentrations of species)
From pathway structure, equations are established using kinetic laws (e.g. mass action or Michaelis-Menten law)

Parameters in Biological Models and Formal Methods

Parameters used to capture uncertainty

- in extrapolation of experimental data
- variety in species

Formal Methods

- efficiently handling **uncertainty**
  - tubes of behaviors rather than single behavior obtained by numerical simulation

Parameter Estimation

- Only some parameters (such as reaction rates, and production and decay coefficients) have phical meanings and can be experimentally measured
- Others come from approximations or reductions, therefore might have no direct biological or biochemical interpretation.
- Fitting : using time series of experimental measurements to determine parameter values that minimize error between these measurements and corresponding model predictions.
- Measurements can be very noisy, and obtained at a limited number of time points.

Model Revision

We have :

- A First ODE model (of Haemoglobin production)
- Data from multi-stage experimental procedures.
- Previous results from simulations on a Monte-Carlo sampling of the parameter space

Model Revision

We have :

- A First ODE model (of Haemoglobin production)
- Data from multi-stage experimental procedures.
- Previous results from simulations on a Monte-Carlo sampling of the parameter space

Model Revision

We have :

- A First ODE model (of Haemoglobin production)
- Data from multi-stage experimental procedures.
- Previous results from simulations on a Monte-Carlo sampling of the parameter space

Model Revision

We have :

- A First ODE model (of Haemoglobin production)
- Data from multi-stage experimental procedures.
- Previous results from simulations on a Monte-Carlo sampling of the parameter space

We want to :

Model Revision

We have :

- A First ODE model (of Haemoglobin production)
- Data from multi-stage experimental procedures.
- Previous results from simulations on a Monte-Carlo sampling of the parameter space

We want to :

- Improve the quality of the fitting

Model Revision

We have :

- A First ODE model (of Haemoglobin production)
- Data from multi-stage experimental procedures.
- Previous results from simulations on a Monte-Carlo sampling of the parameter space

We want to :

- Improve the quality of the fitting

How ?

Model Revision

We have :

- A First ODE model (of Haemoglobin production)
- Data from multi-stage experimental procedures.
- Previous results from simulations on a Monte-Carlo sampling of the parameter space

We want to :

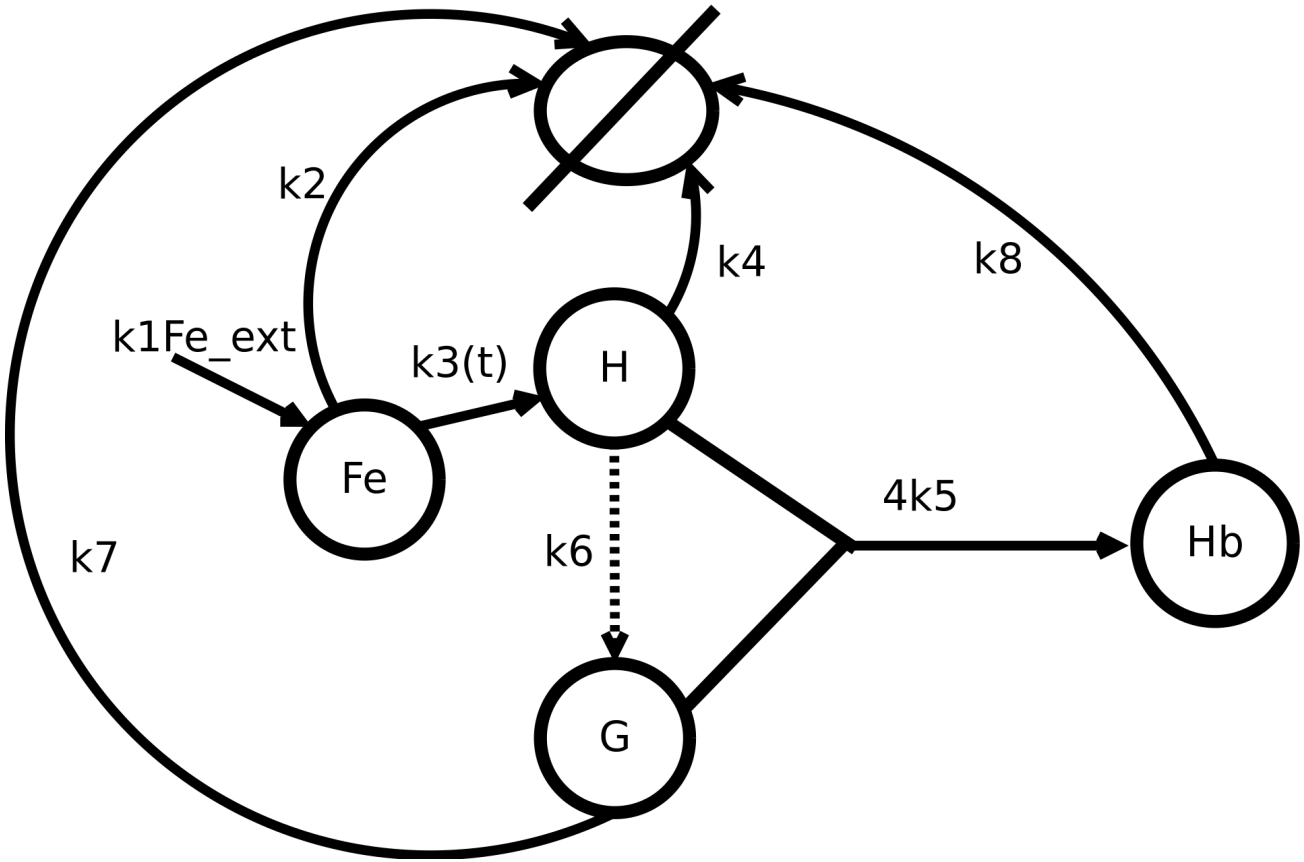
- Improve the quality of the fitting

One solution

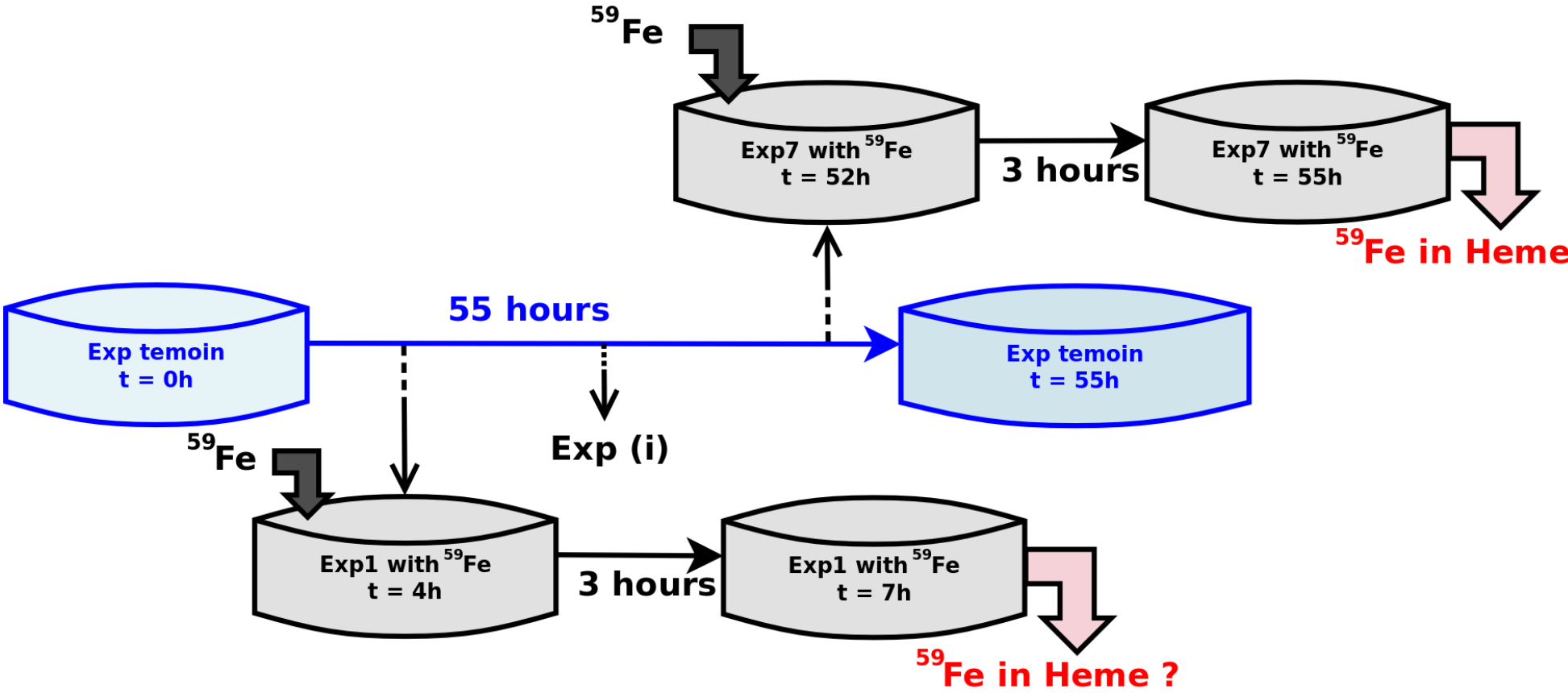
- Transform a parameters $\mathbf{k}$ into a time varying function $\mathbf{k}(t)$

Model of Haemoglobin Production

- Model of Haemoglobin production, from ANR Cadmidia Project
- 52 first hours of an eryhtroblast, before leaving the bone marrow



Experimental Workflow



Model of Hemoglobin Production

Mechanism

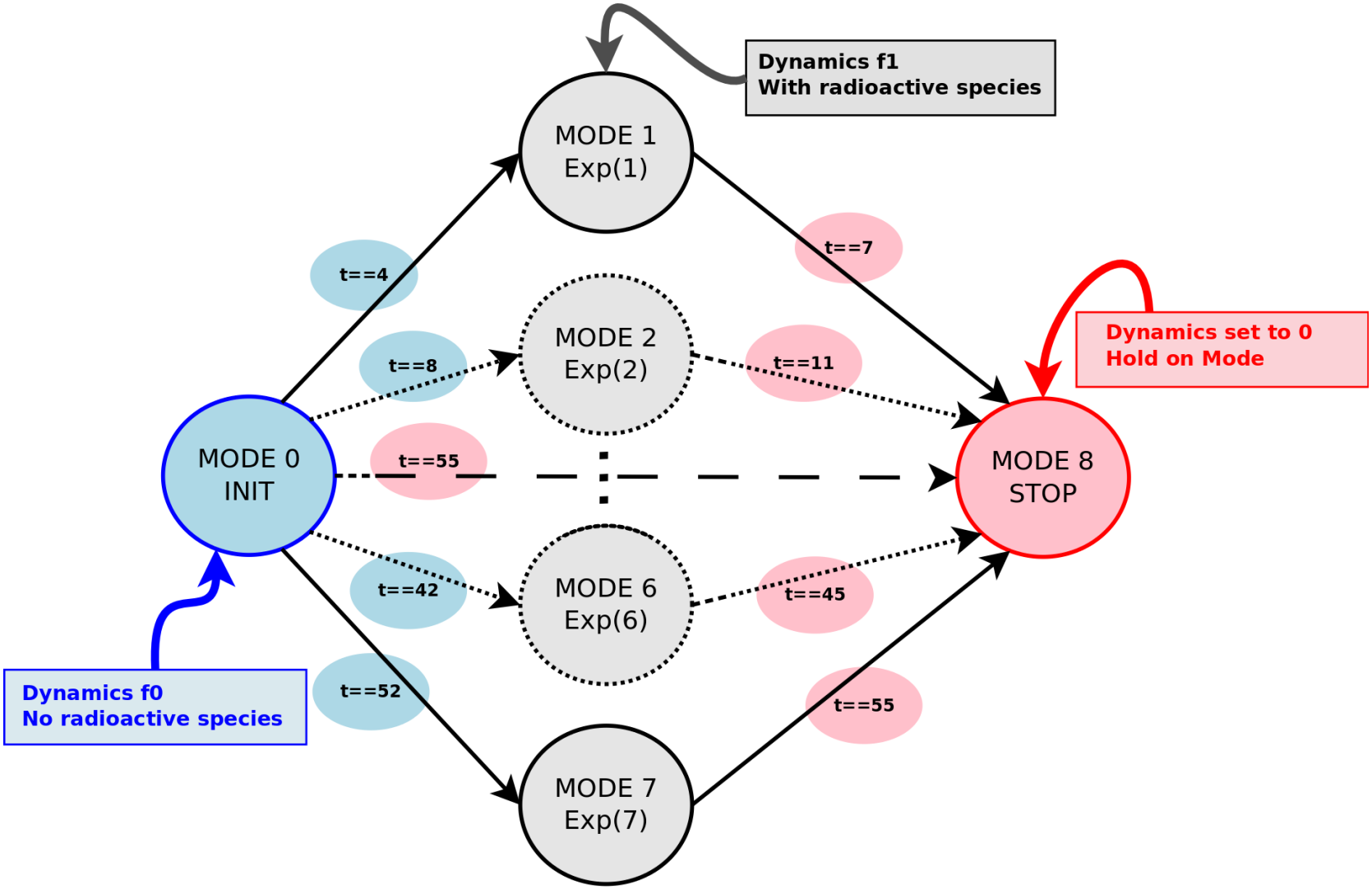
- $\mathcal{F}_{norm}$: System without radioactive species.

$$\begin{cases} \frac{dFe}{dt} &= k_1\mathrm{Fe}_{\mathrm{ext}} - k_2\mathrm{Fe} - k_3(t)\mathrm{Fe} \\ \frac{dH}{dt} &= k_3(t)\mathrm{Fe} - k_4\mathrm{H} - 4k_5\mathrm{HG} \\ \frac{dG}{dt} &= k_6\mathrm{H} - 4k_5\mathrm{HG} - k_7\mathrm{G} \\ \frac{dHb}{dt} &= k_5\mathrm{HG} - k_8\mathrm{Hb} \end{cases}$$

- $\mathcal{F}_{rad}$: System with (additional) radioactive species.

$$\begin{cases} \dots &= \dots \\ \frac{d^{59}\mathrm{Fe}}{dt} &= k_1{}^{59}\mathrm{Fe}_{\mathrm{ext}} - k_2{}^{59}\mathrm{Fe} - k_3(t){}^{59}\mathrm{Fe} \\ \frac{d^{59}\mathrm{H}}{dt} &= k_3(t){}^{59}\mathrm{H} - k_4{}^{59}\mathrm{H} - 4k_5{}^{59}\mathrm{HG}_{\mathrm{rad}} \\ \frac{d\mathrm{G}_{\mathrm{rad}}}{dt} &= k_6({}^{59}\mathrm{H} + \mathrm{H}) - 4k_5({}^{59}\mathrm{H} + \mathrm{H})\mathrm{G}_{\mathrm{rad}} - k_7\mathrm{G}_{\mathrm{rad}} \\ \frac{d^{59}\mathrm{Hb}}{dt} &= k_5{}^{59}\mathrm{HG}_{\mathrm{rad}} - k_8{}^{59}\mathrm{Hb} \end{cases}$$

Modelling as an Hybrid Automaton



Parameter Synthesis by Optimal Control

Approach

- Formulating reachable sets by occupation measures ([Henrion et 2013, Shia 2014, Zhao 2017]
- Formulating data fitting error as objective function
- Solving optimal control problem for hybrid automata to find inputs allows best data fitting

Occupation Measures : Basic Definitions

- **Measure** μ defined on a set $X \subset \mathbb{R}^n$ s.t. $\mu(\emptyset) = 0$ and $\mu(\bigcup_i^\infty X_i) = \bigcup_1^\infty \mu(X_i)$ (pairwise disjoint collection of sets $X_i \subset X$)
- **Dynamical system** $\dot{x}(t) = f(t, x, u)$ where t $in[0, T]$, $x \in \mathcal{X}$, $u \in \mathcal{U}$
  - Given an initial state x_0, and a trajectory $x(t|x_0)$ under an input $u(t|x_0) \in \mathcal{U}$, **occupation measure** $\mu(\cdot|x_0)$ for $\Delta \times X \times U \subset [0, T] \times \mathcal{X} \times \mathcal{U}$

$$\mu(\Delta \times X \times U|x_0) := \int_0^T \mathbb{I}_{\Delta \times X \times U}(t, x(t|x_0), u(t|x_0))dt$$

    $\mathbb{I}$ is the indicator function
  - This measure defines **the time the system spends** in $\times X \times U$.
- **Property** : for any measurable function $g(t, x, u)$

$$\int_0^T g(t, x(t|x_0), u(t|x_0))dt = \int_0^T \mathbb{I}_{[0,T] \times \mathcal{X} \times \mathcal{U}}\, g(t, x, u)d\mu(t, x, u|x_0)$$

Occupation Measures

- **Linear operator** for continuously differentiable functions $v \in C^1([0,T],\mathcal{X})$, $v \to \mathcal{L}v := \frac{\partial v}{\partial t} + \sum_{i=1}^{n} \frac{\partial v}{\partial x_i} f_i(t,x,u)$
- **Adjoint operator** $\langle \mathcal{L}'v, \mu \rangle := \langle \mu, \mathcal{L}v \rangle = \int_{[0,T]\times\mathcal{X}\times\mathcal{U}} \mathcal{L}v \, d\mu(t,x,u|x_0)$
- Then, for a test function $v \in C^1([0,T],\mathcal{X})$

$$v(T, x(T|x_0)) = v(0,x_0) + \langle \mathcal{L}'\mu(\cdot|x_0), v \rangle$$

Occupation Measures for Reachability Analysis

- Now we consider initial measure μ_0 for the state space X. Initial state $x_0 \in spt(\mu_0)$
- Average occupational measure for a set $\Delta \times X \times U \subset [0, T] \times \mathcal{X} \times \mathcal{U}$,

$$\mu(\Delta \times X \times U) = \int_{\mathcal{X}} \mu(\Delta \times X \times U \,|\, x_0) d\mu_0(x_0)$$

- For final time T, $\mu_T(X) = \int_{\mathcal{X}} \mathbb{I}_X(x(T \,|\, x_0)) d\mu_0(x_0)$

Liouville equation (conservation of the information along the trajectory for any test function v)

$$\int_{X_T} v(T, x) d\mu_T(x) = \int_X v(0, x) d\mu_0(x) + \int_{[0,T]\times X \times U} \mathcal{L}v(t, x, u)\, d\mu(t, x, u)$$

More concisely,

$$\langle \mu_T, v \rangle - \langle \mu_0, v \rangle = \langle \mu, \mathcal{L}v \rangle$$

Occupation Measures for Hybrid Systems

Consider a transition from location j to location i, with guard $G_{j,i}$, $\mathcal{X}_i$ is the state space of location i.

- Final occupation at location i : occupation at location i by continuous dynamics and occupation by taking transitions to i

$$\mu_T^i = \mu_T(\mathcal{X}_i) + \sum_{(j,i)} \mu(G_{j,i}$$

- If transition (j, i) has a reset map $R_{j,i}$, then $\langle \mu(R_{j,i}(G_{j,i})), v \rangle = \langle R_{j,i}^* \, \mu(G_{j,i}), v \rangle$
- For initial measure, $\mu_0^i = \mu_0(\mathcal{X}_i) + R_{j,i}^* \, \mu(G_{j,i})$
- Liouville's equation

$$\mathcal{L}'\mu^i = \mu_T(\mathcal{X}_i) + \sum_{(j,i)} \mu(G(j,i)) - \mu_0(\mathcal{X}_i) - R_{j,i}^* \, \mu(G_{j,i})$$

Optimal Control Problem

- Terminal time cost

$$\inf_{(x,u)} \int_0^T h\left(t, x(t), u(t)\right) dt + H\left(x(T)\right).$$

- Additional costs at intermediate time points T_j

$$\inf_{(x,u)} \int_0^T h(t, x(t), u(t)) dt + \sum_{0 \leq j \leq n_{exp}} H\left(x(T_j)\right)$$

Optimal Control Problem Formulation in Occupation Measure

$$\inf \sum_{i \in I} \mu^i h_i + \sum_{i \in I} \mu^i_T H_i$$

such that

$$\mathcal{L}' \mu^i = \mu_T(\mathcal{X}_i) + \sum_{(j,i)} \mu(G(j,i)) - \mu_0(\mathcal{X}_i) - R^*_{j,i}\, \mu(G_{j,i})$$

and constraints of positivity of measure variables

Finite Dimensional Semidefinite Program Formulation

Given a multi-index $\alpha \in \mathbb{N}^n$, moments of a measure μ is $[y_\mu]_\alpha := \int x^\alpha d\mu(x)$

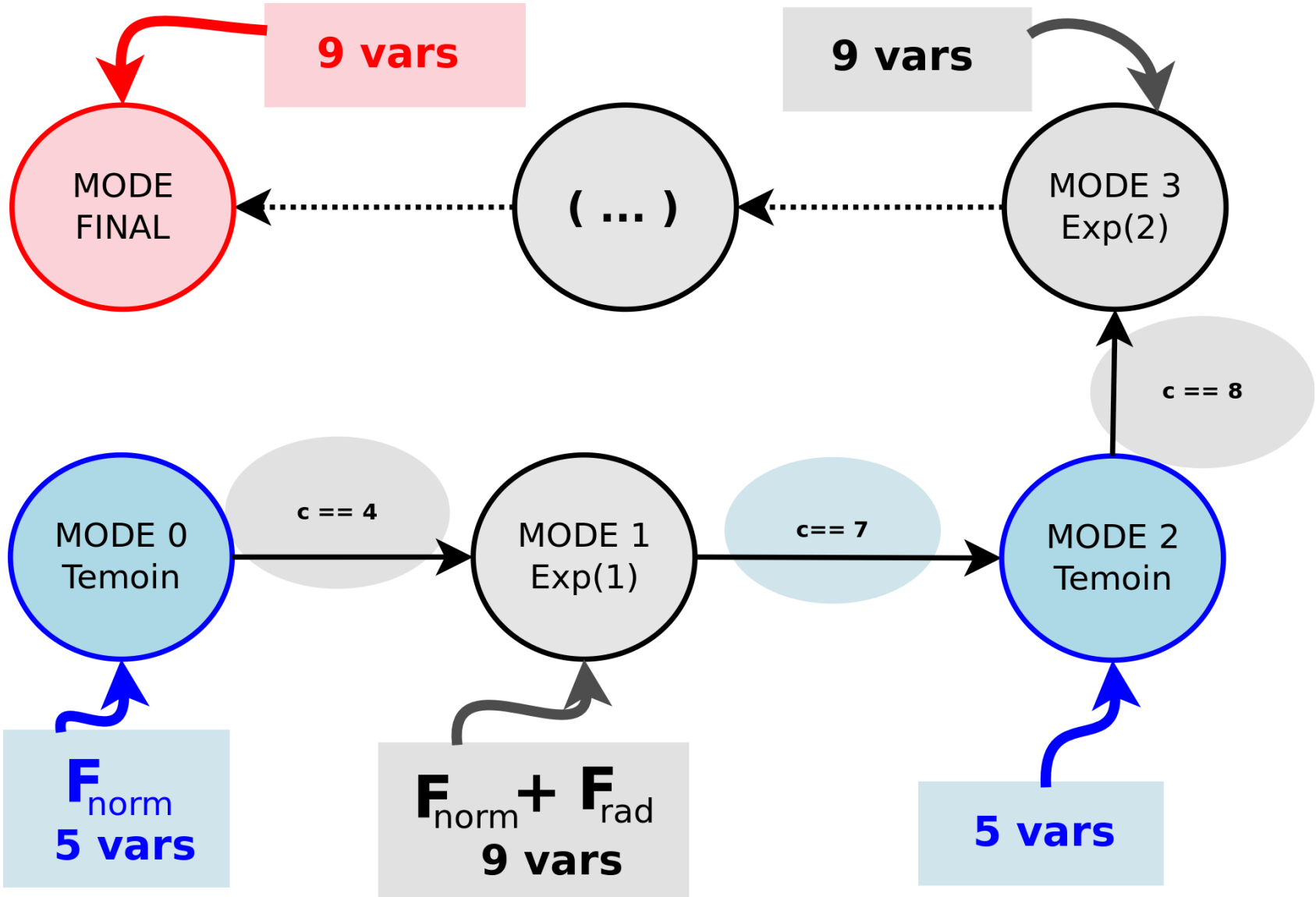
Polynomial $p \in \mathbb{R}_r[x]$, let $L_{y_\mu}(p) := \langle \mu, p \rangle = \int (\sum_{|\alpha| \leq r} p_\alpha x^\alpha) d\mu$.

Moment matrix $[M_r(y_\mu)]_{(\alpha,\beta)} := [y_\mu]_{(\alpha+\beta)}$, where $|\alpha + \beta| \leq 2r$.

Given a polynomial g of degree $l < r$, localizing matrix $[L_r(g, y_\mu)]_{(\alpha,\beta)} := \sum_{|\gamma| \leq l} g_\gamma [y_\mu]_{(\alpha+\beta)}$

$\Rightarrow$ Rewrite the positivity constraints as semidefinite constraints on moments and localizing matrices, and then truncating the degree of the moments to $2r \Rightarrow$ finite dimensional semidefinite program [Nesterov, Lasserre, Parrilo (2000-)]

Hybrid automaton of Hemoglobin Production



Model Revision

- Model revision problem $\Rightarrow$ intermediate costs :
  - n_{exp} experimental data : $(\mathbf{z}_j, T_j)$, $1 \leq j \leq n_{exp}$.
  - At intermediate time T_j, $H(T_j) = ||\mathbf{m}(\mathbf{x}(T_j)) - \mathbf{z}_j||_2^2$.

$$
\begin{aligned}
J^* := \inf_{(\lambda,\mathbf{x}),\mathbf{u}} \quad & \int_0^T h(t, \mathbf{x}(t), \mathbf{u}(t))dt + \sum_{0 \leq j \leq n_{exp}} H(\mathbf{x}(T_j)) \\
\text{s.t.} \quad & (\lambda(t), \mathbf{x}(t)), \forall t \in [0, T] \text{ a trajectory of } \mathcal{H}, \\
& (\lambda(t), \mathbf{x}(t)) \in I \times X_{\lambda(t)}, \forall t \in [0, T], \\
& (\lambda(0), \mathbf{x}(0)) = (i_0, \mathbf{x}_0) \in I \times X_{i_0}, \\
& (\lambda(T_j), \mathbf{x}(T_j)) \in \mathcal{X}_{T_j} \forall 1 \leq j \leq n_{exp}, \\
& \mathbf{u} : T \to U \text{ continuous functions},
\end{aligned}
\tag{1}
$$

Greedy Approach

➢ Handling each each data point $\mathbf{z}_j$ iteratively on $[T_{j-1}, T_j]$ we solve :

$$J_j^* := \inf_{(\mathbf{x}^{(j)}, \tilde{\mathbf{u}}^{(j)})} J_j(t, \mathbf{x}^{(j)}(t), \tilde{\mathbf{u}}^{(j)}(t))$$

s.t.

$(\lambda^{(j)}, \mathbf{x}^{(j)})$ a trajectory of $\mathcal{H}$ on $\mathcal{T}_j$,

$\tilde{\mathbf{u}}^{(j)}(t) \in U\,, \quad \forall t \in \mathcal{T}_j\,,$

$(\lambda^{(j)}(t), \mathbf{x}^{(j)}(t)) \in \mathcal{X}\,, \quad \forall t \in \mathcal{T}_j\,,$

- if $j = 1\,,$

  $(\lambda^{(1)}(0), \mathbf{x}^{(1)}(0)) = (i_0, \mathbf{x}_0) \in I \times X_{i_0}\,,$

- if $j \geq 2.$

  $(\lambda^{(j)}(T_{j-1}), \mathbf{x}^{(j)}(T_{j-1})) = (\lambda^{(j-1)}(T_{j-1}), \mathbf{x}^{(j-1)}(T_{j-1}))\,,$

$(\lambda^{(j)}(T_j), \mathbf{x}^{(j)}(T_j)) \in X_{T_j}.$

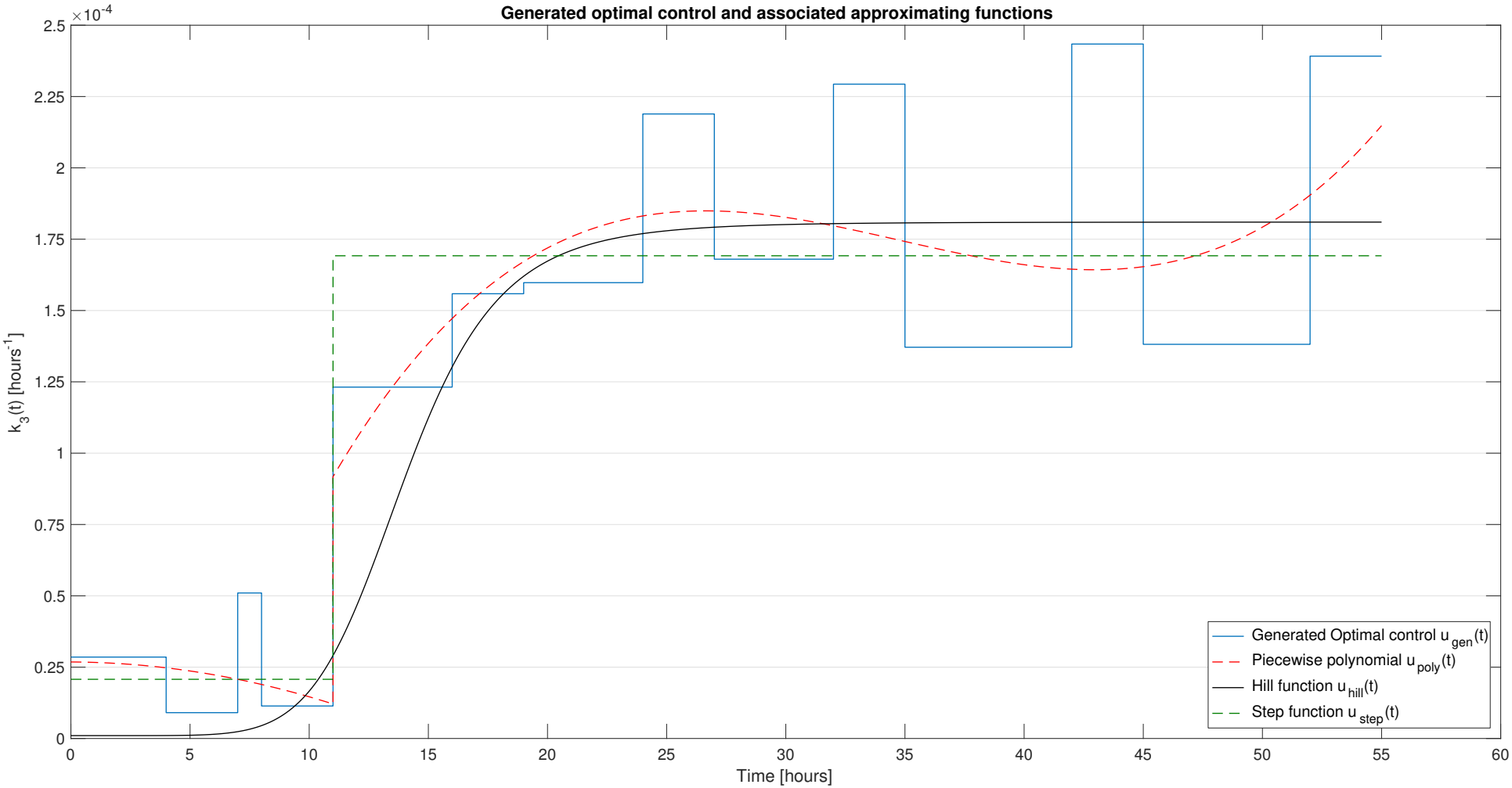
Results on Haemoglobin Case Study

- 11 continuous variables
- 7 data points $\Rightarrow$ 7 iterations
- At each iteration : 2 modes ($\mathcal{F}_{norm}$ and $\mathcal{F}_{rad}$)
- Control searched as piece-wise constant
- Total computation time : 35 minutes

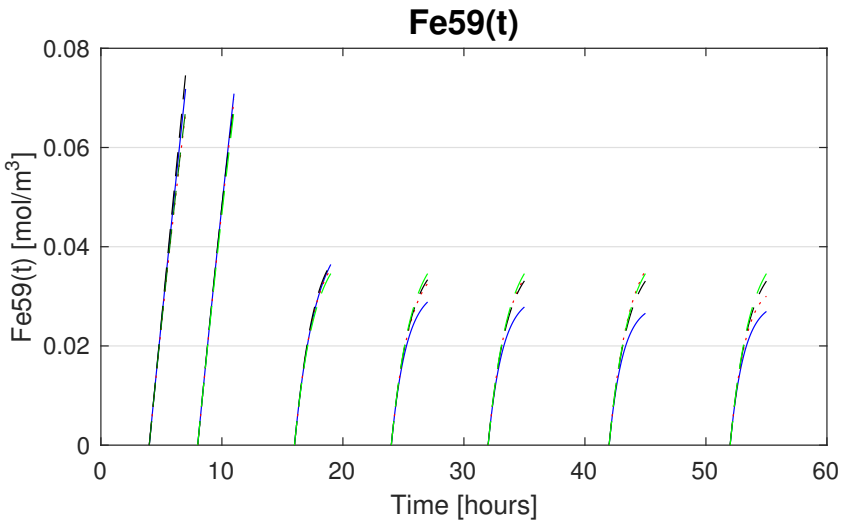
Time (h)	7	11	19	27	35	45	55
Measure ($\frac{cpm}{1e^{-7}\mathrm{L}\cdot h^{-1}}$)	16	85	348	391	399	481	395

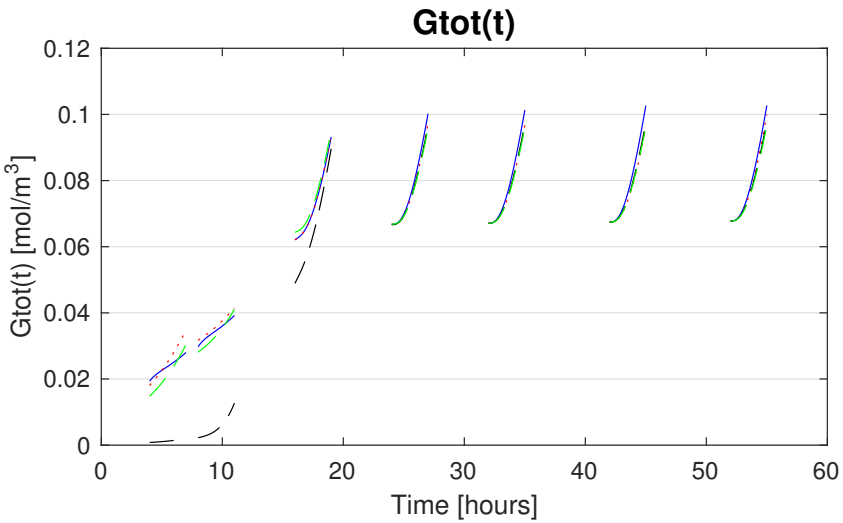
TABLE – Experimental data points (T_j, z_j) used as references (measurement z is function of state x.

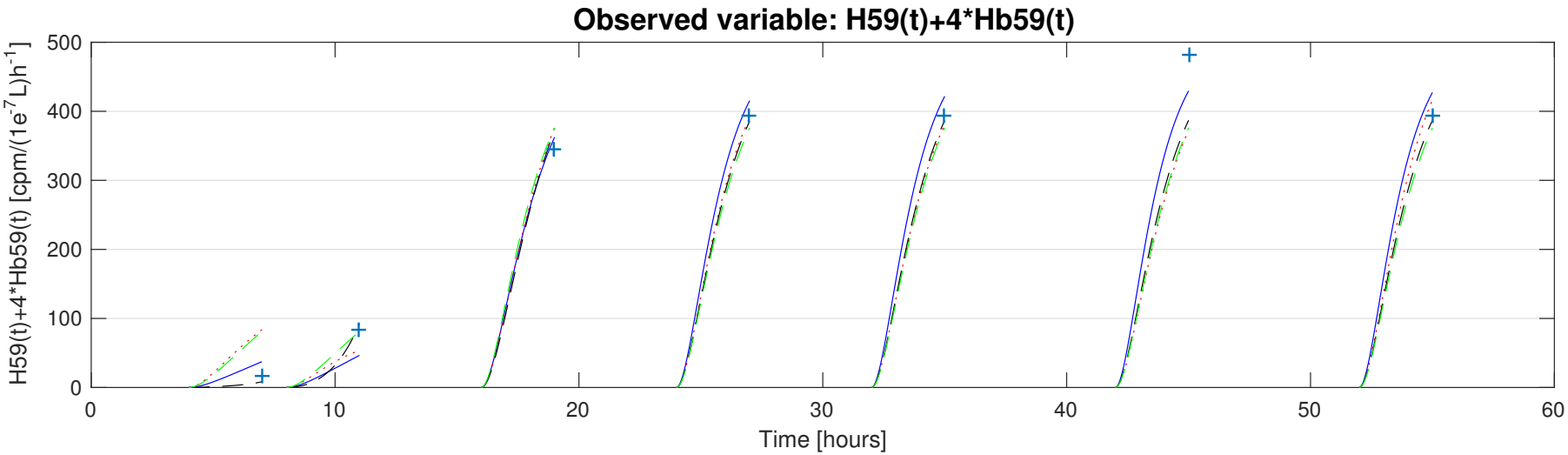
Results on Haemoglobin Case Study



Results on Haemoglobin case study







Results on Haemoglobin Case Study

Control Type	ε_{total}
Previous fit	0.23
Control generated	0.096
Step function fit	0.12
Piecewise Polynomial fit	0.13
Hill function fit	0.075

TABLE – Total error ε_{total} associated to each proposed input.

$$\varepsilon_{total} = \sum_{1 \leq j \leq n_{exp}} \frac{\sqrt{H(\mathbf{x}(T_j))}}{\sum_{1 \leq j \leq n_{exp}} z_j} .$$

Conclusions

- Can produce 'successive' admissible control inputs
- Piecewise polynomial control.
- Large class of hybrid systems.
- Not necessarily optimal...
- ... but good trade-off accuracy/performance.
- Need few a priori knowledge.
- Eases the biological interpretation.

Perspectives

- ☞ Iterative backward reachable set computation using occupation measures for parameter synthesis ?
  - ➢ Intersection of backward reachable set for each target data points.
  - ➢ More expensive than greedy application of optimal control.