HelmholtzZentrum münchen
German Research Center for Environmental Health

MASAMB 2016

Efficient Parameter Estimation for ODE Models from Relative Data Using Hierarchical Optimization

Sabrina Krause, Carolin Loos, Jan Hasenauer

Helmholtz Zentrum München
Institute of Computational Biology

Data-driven Computational Modelling

Cambridge, 02/10/16

HELMHOLTZ ASSOCIATION

Parameter Estimation

Parameter Estimation

ODE model:

$$\frac{dx}{dt} = f(\theta, x(t,\theta)), \quad x(0,\theta) = x_0(\theta) \quad \textit{dynamics}$$

$$y(t) = h(\theta, x(t,\theta)) \quad \textit{observables}$$

Parameter Estimation

ODE model:

$$\frac{dx}{dt} = f(\theta, x(t,\theta)), \quad x(0,\theta) = x_0(\theta) \quad \textit{dynamics}$$

$$y(t) = h(\theta, x(t,\theta)) \quad \textit{observables}$$

Measurements: $\bar{y}_k = h(\theta, x(t_k,\theta)) + \varepsilon_k, \quad \varepsilon_k \sim \mathcal{N}(0,\sigma^2), \quad k = 1,\dots,n_t$

Parameter Estimation

ODE model:

$$\frac{dx}{dt} = f(\theta, x(t,\theta)), \quad x(0,\theta) = x_0(\theta) \quad \textit{dynamics}$$

$$y(t) = h(\theta, x(t,\theta)) \quad \textit{observables}$$

Measurements: $\bar{y}_k = h(\theta, x(t_k,\theta)) + \varepsilon_k$, $\varepsilon_k \sim \mathcal{N}(0,\sigma^2)$, $k = 1,\ldots,n_t$

Maximize the likelihood function:

$$\max_{\theta} \left\{ p(\mathcal{D}|\theta) = \prod_k \frac{1}{\sqrt{2\pi\sigma^2}} \, exp \left\{ -\frac{1}{2} \left(\frac{\bar{y}_k - h(\theta, x(t_k,\theta))}{\sigma} \right)^2 \right\} \right\}$$

Parameter Estimation

ODE model:

$$\frac{dx}{dt} = f(\theta, x(t,\theta)), \quad x(0,\theta) = x_0(\theta) \quad \textit{dynamics}$$

$$y(t) = h(\theta, x(t,\theta)) \quad \textit{observables}$$

Measurements: $\bar{y}_k = h(\theta, x(t_k,\theta)) + \varepsilon_k, \quad \varepsilon_k \sim \mathcal{N}(0,\sigma^2), \quad k = 1,\dots,n_t$

Minimize the negative log likelihood function:

$$\min_{\theta} \left\{ J(\theta) = \frac{1}{2} \sum_k log(2\pi\sigma^2) + \left(\frac{\bar{y}_k - h(\theta, x(t_k,\theta))}{\sigma} \right)^2 \right\}$$

Parameter Estimation

ODE model:

$$\frac{dx}{dt} = f(\theta, x(t,\theta)), \quad x(0,\theta) = x_0(\theta) \quad \textit{dynamics}$$

$$y(t) = h(\theta, x(t,\theta)) \quad \textit{observables}$$

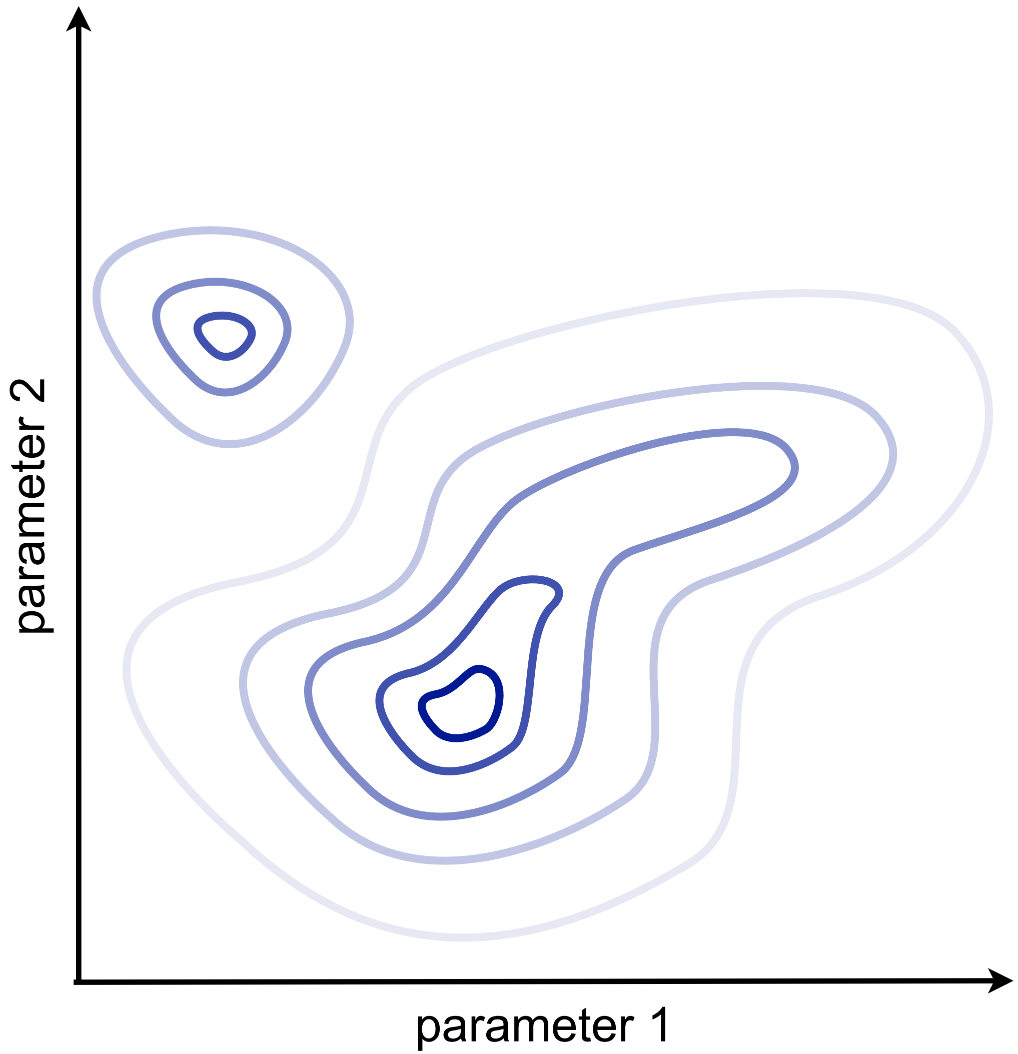
Measurements: $\bar{y}_k = h(\theta, x(t_k,\theta)) + \varepsilon_k$, $\varepsilon_k \sim \mathcal{N}(0,\sigma^2)$, $k = 1,\ldots,n_t$

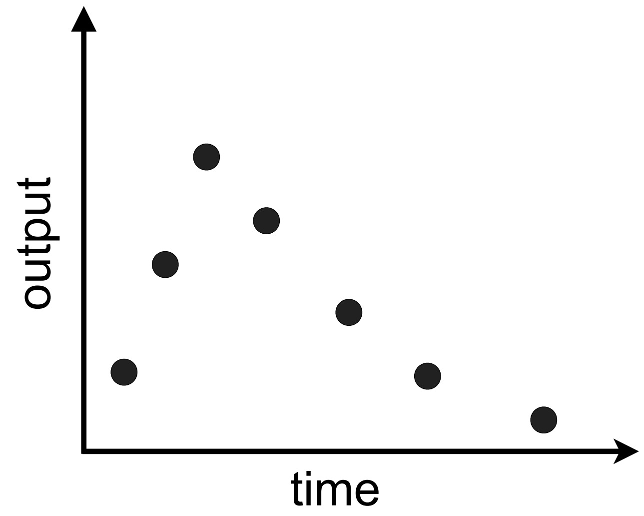
Minimize the negative log likelihood function:

$$\min_{\theta} \left\{ J(\theta) = \frac{1}{2} \sum_k log(2\pi\sigma^2) + \left(\frac{\bar{y}_k - h(\theta, x(t_k,\theta))}{\sigma} \right)^2 \right\}$$

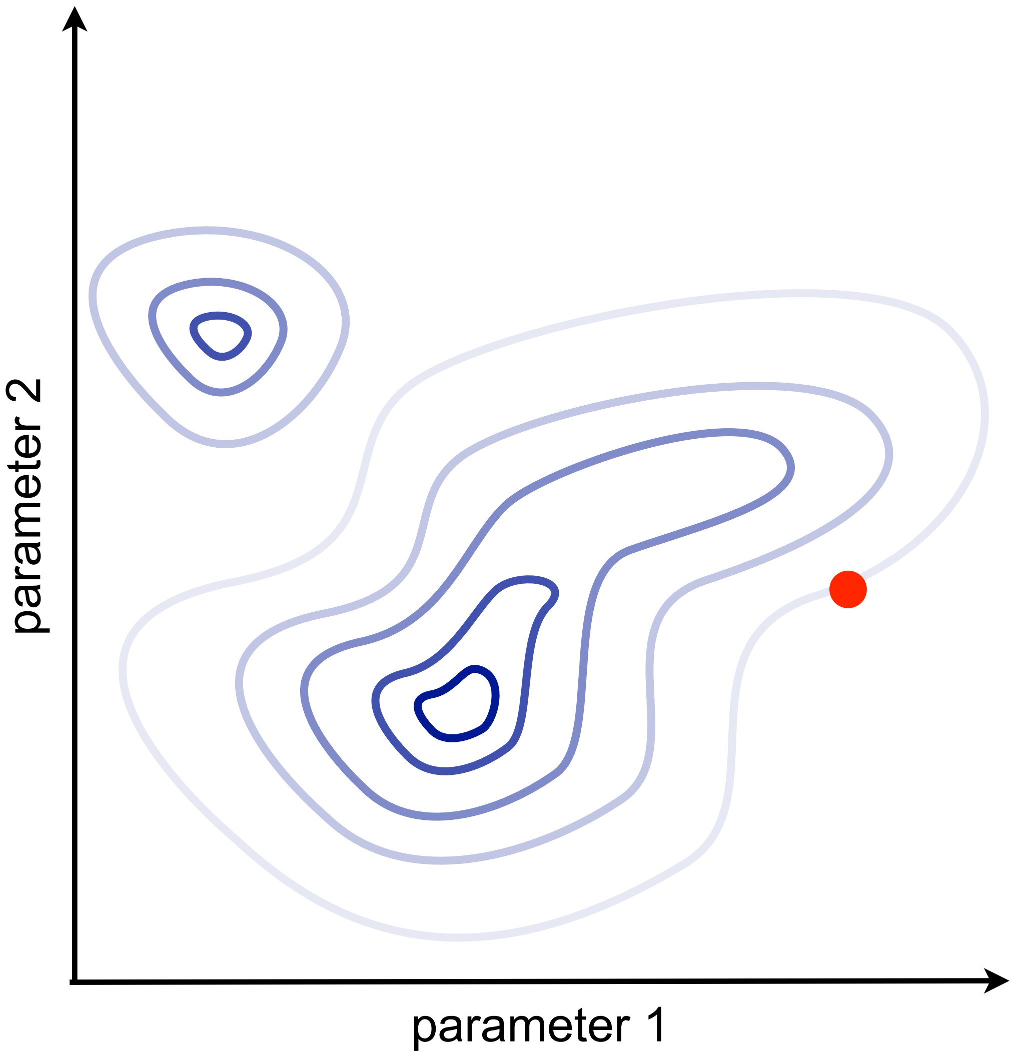
Optimization problem with n_θ parameters

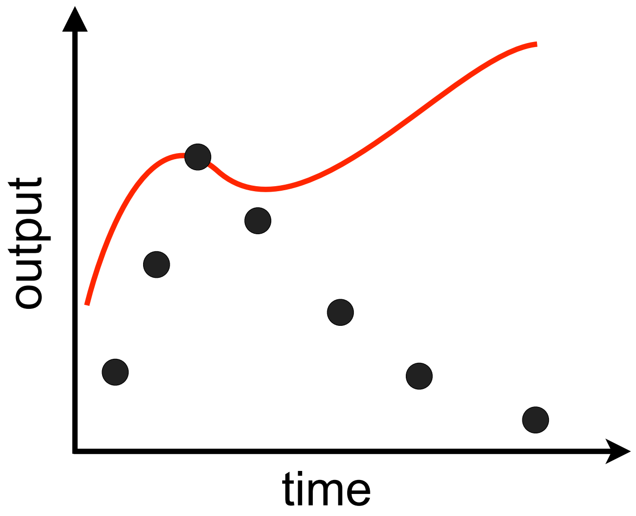
Multi-Start Optimization



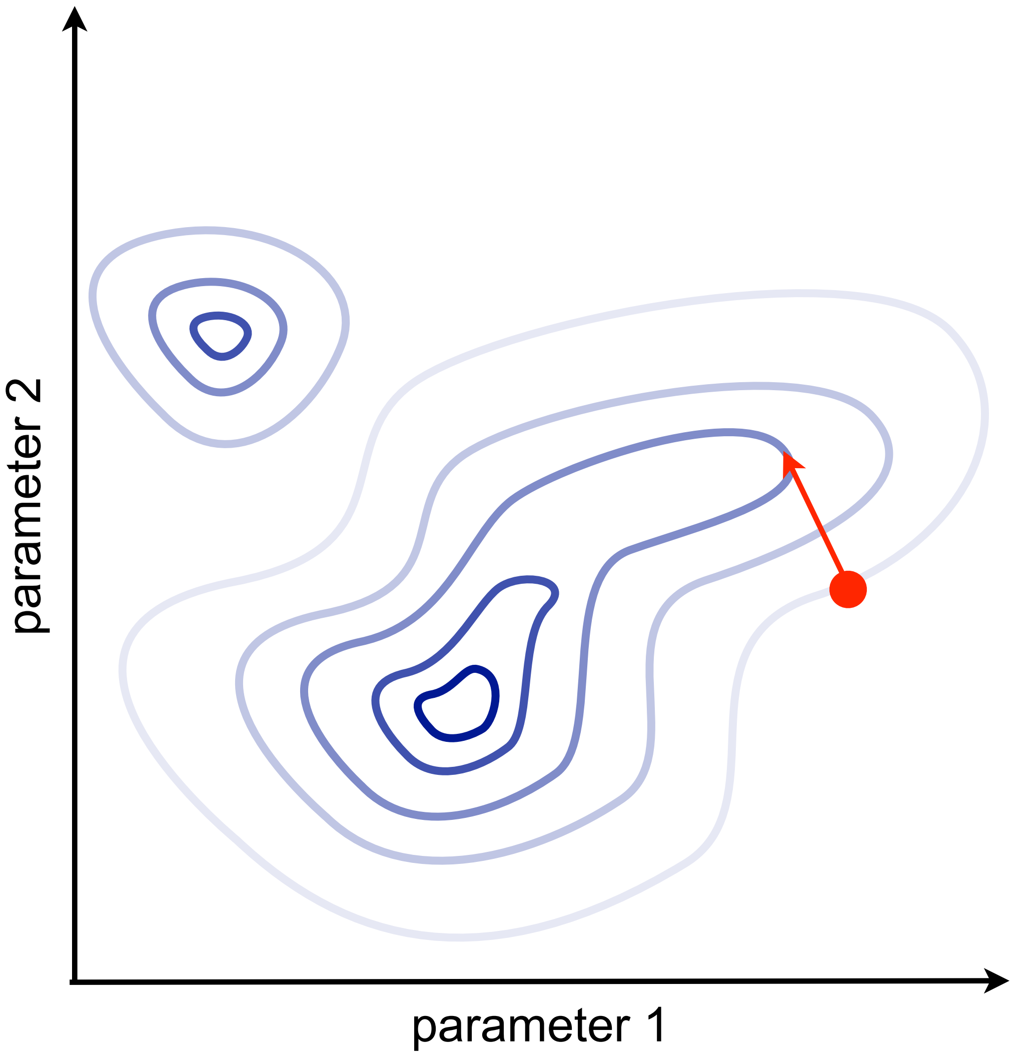


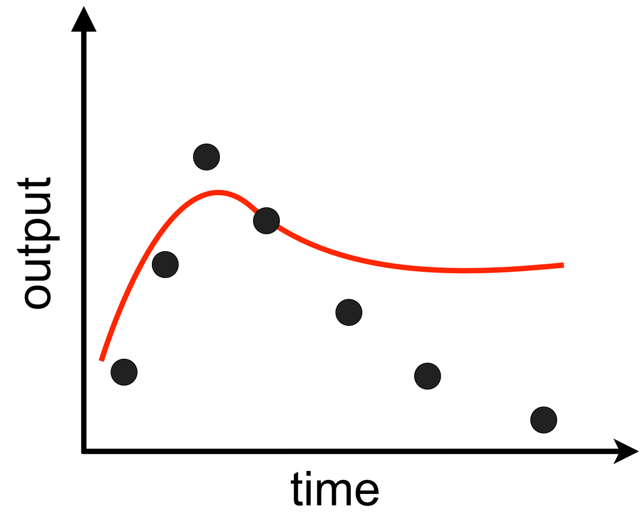
Multi-Start Optimization



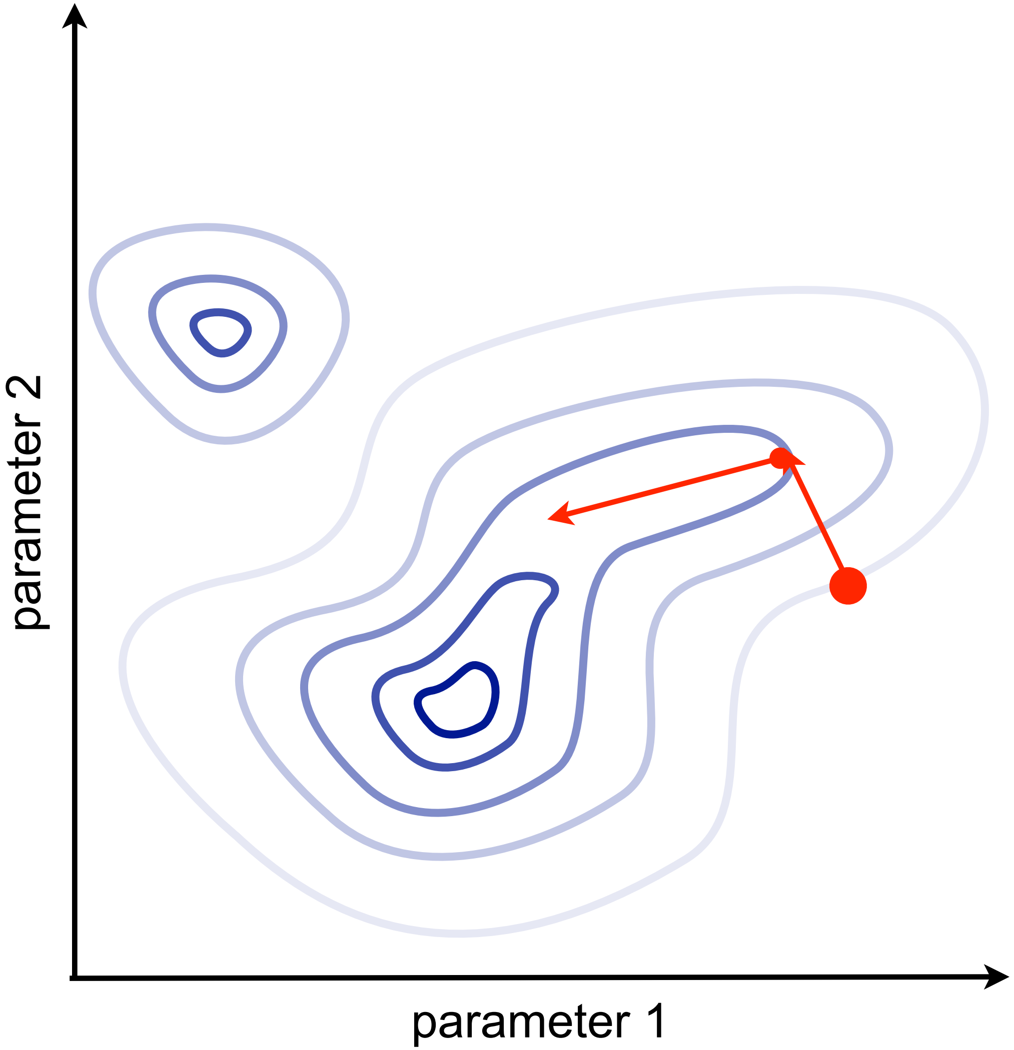


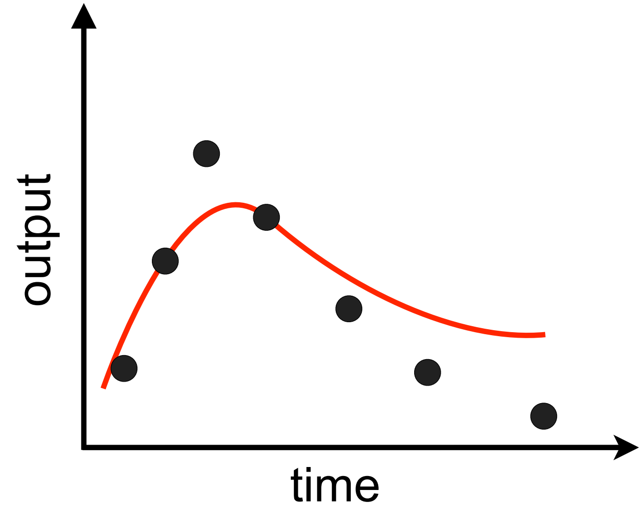
Multi-Start Optimization



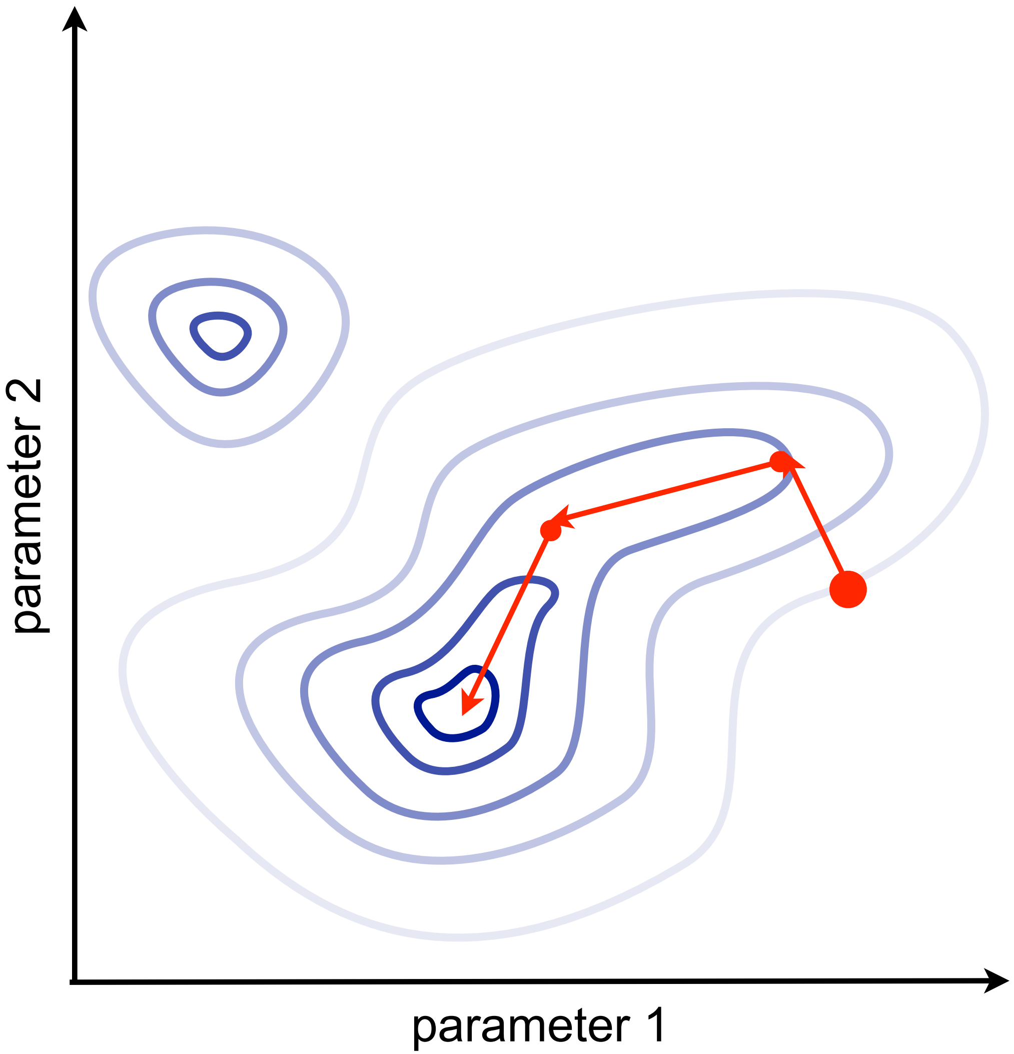


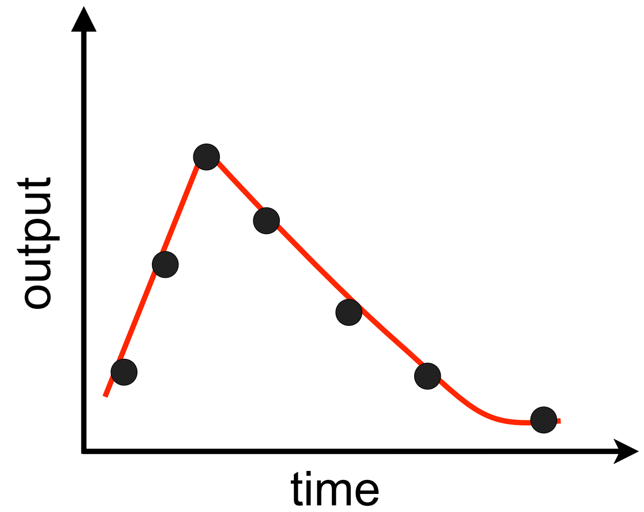
Multi-Start Optimization



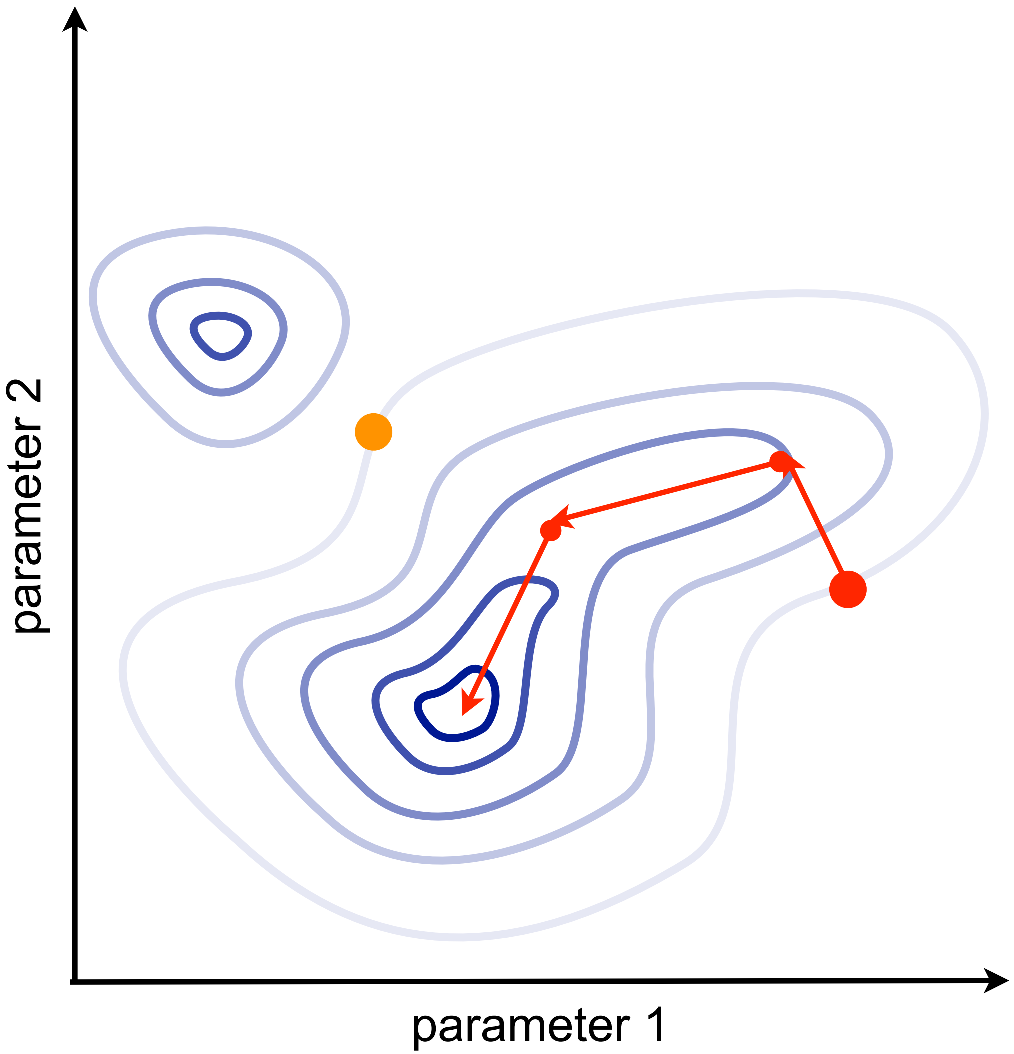


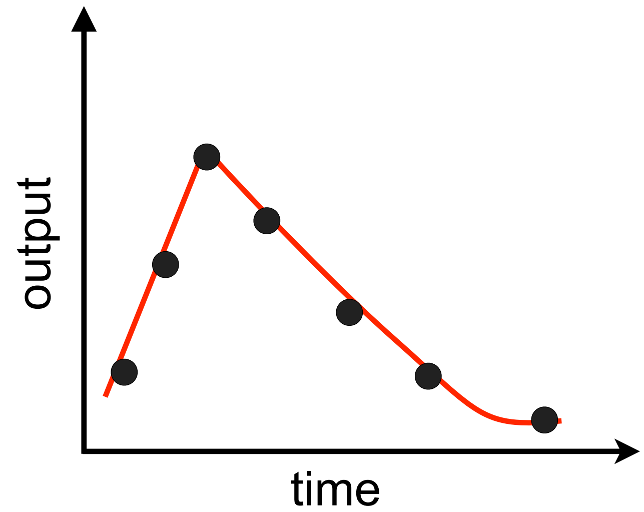
Multi-Start Optimization



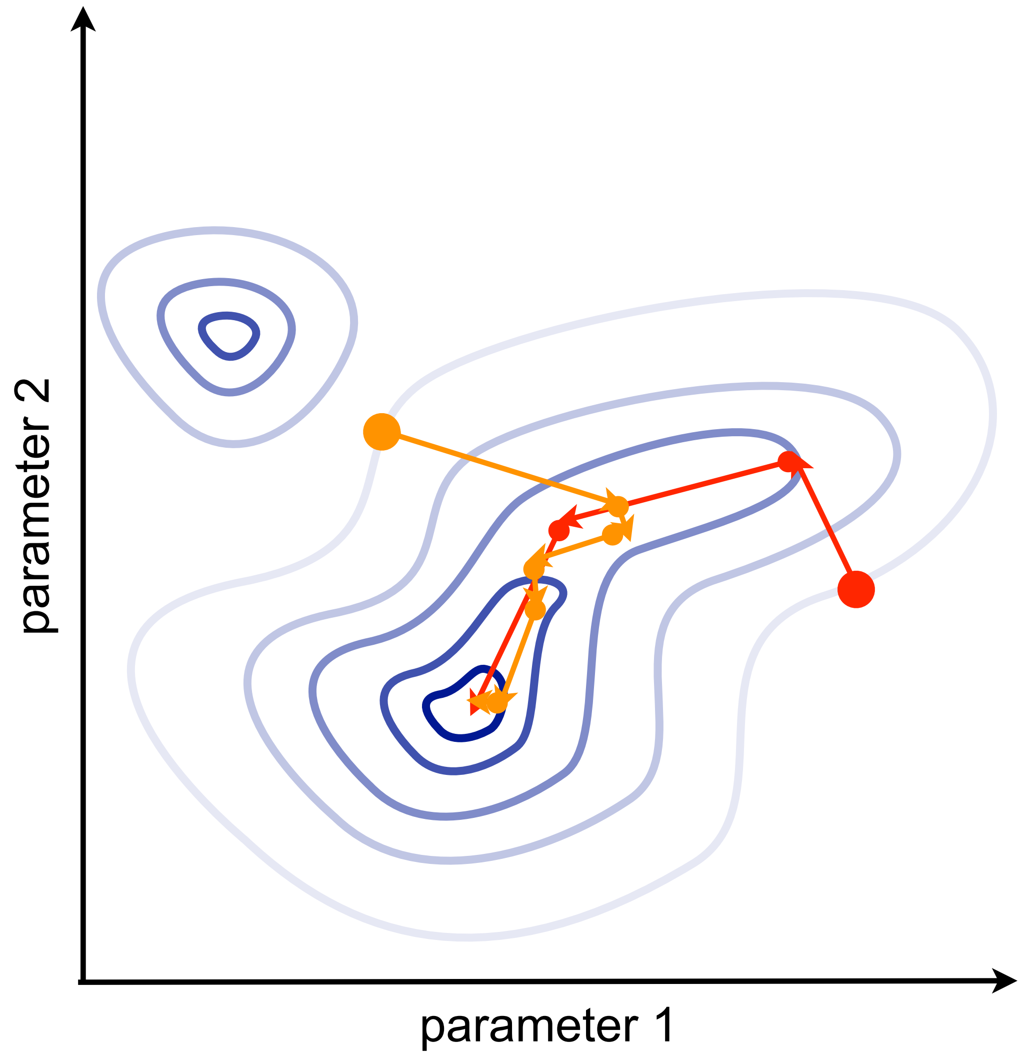


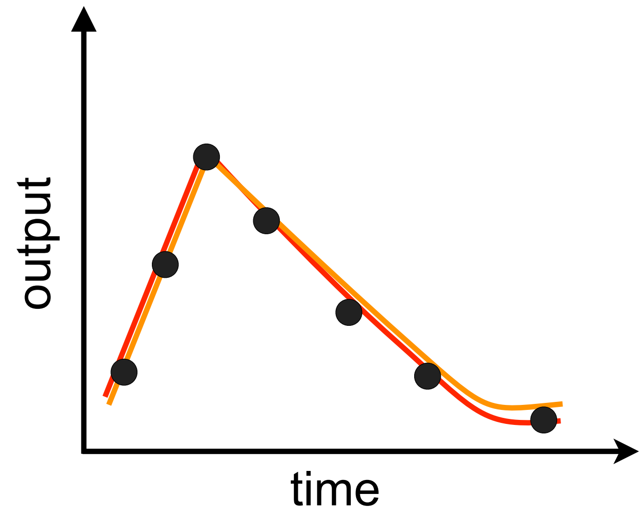
Multi-Start Optimization



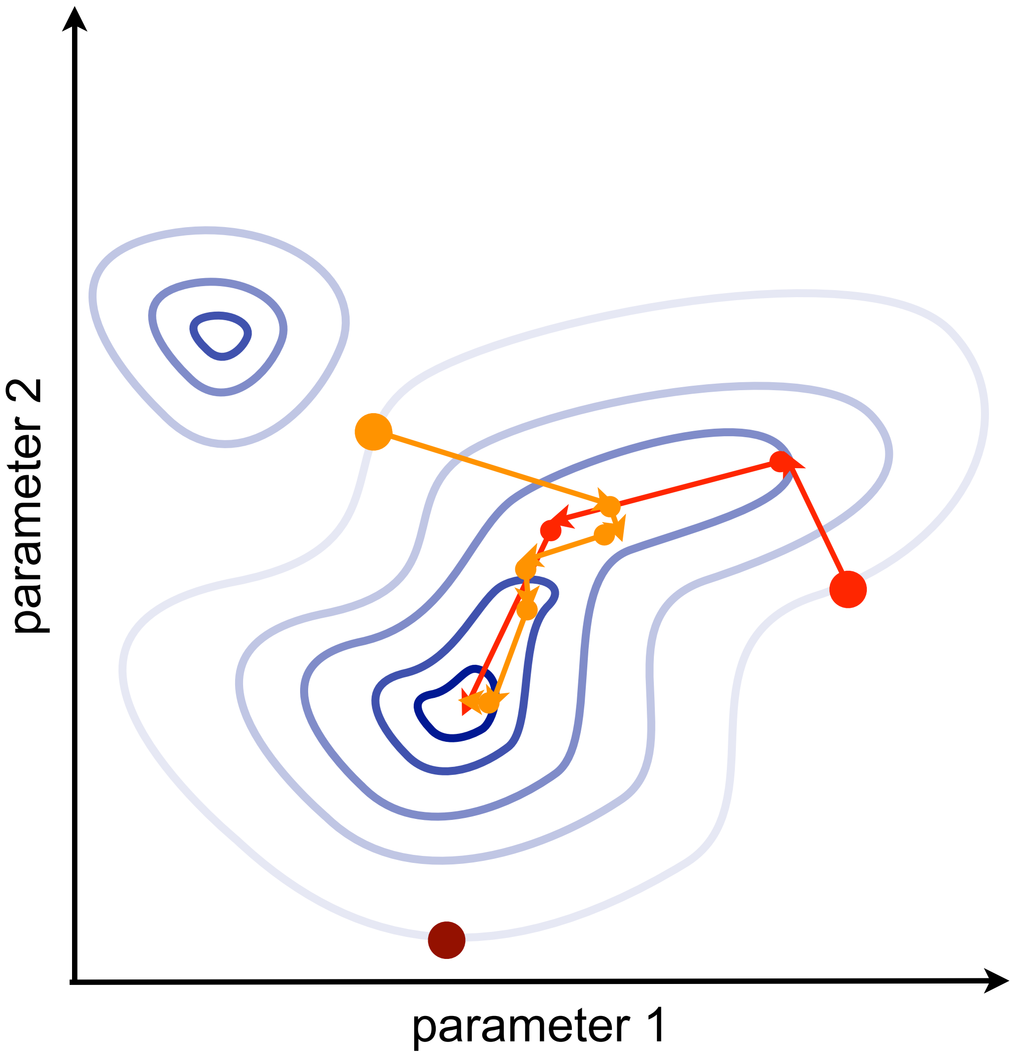


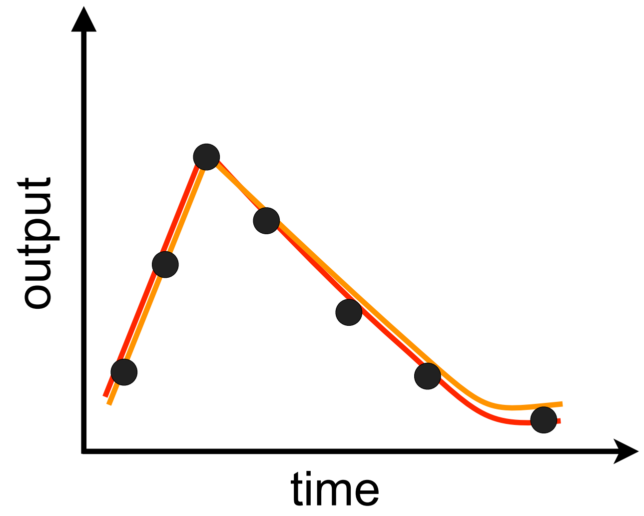
Multi-Start Optimization



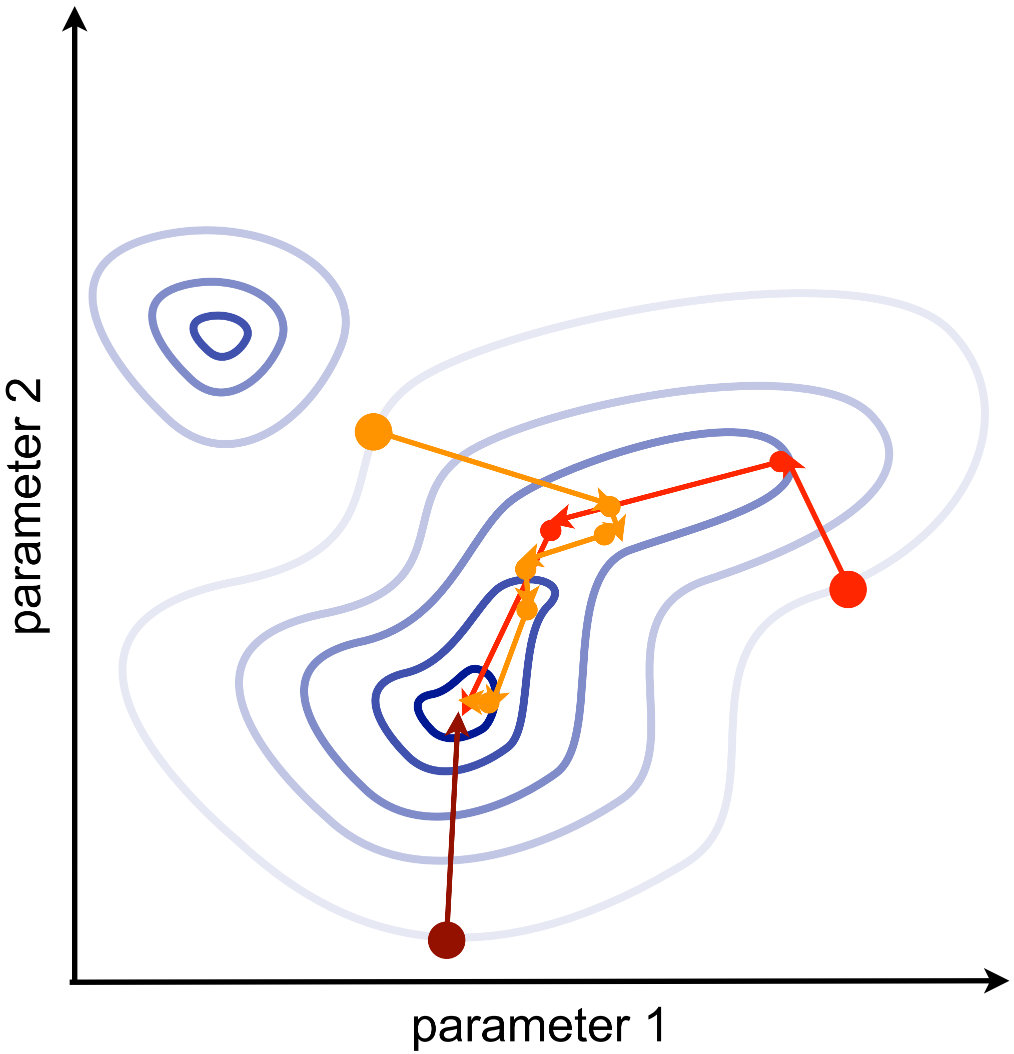


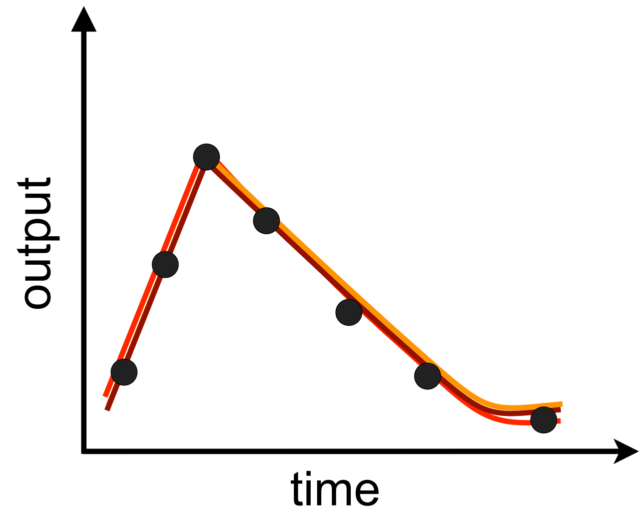
Multi-Start Optimization



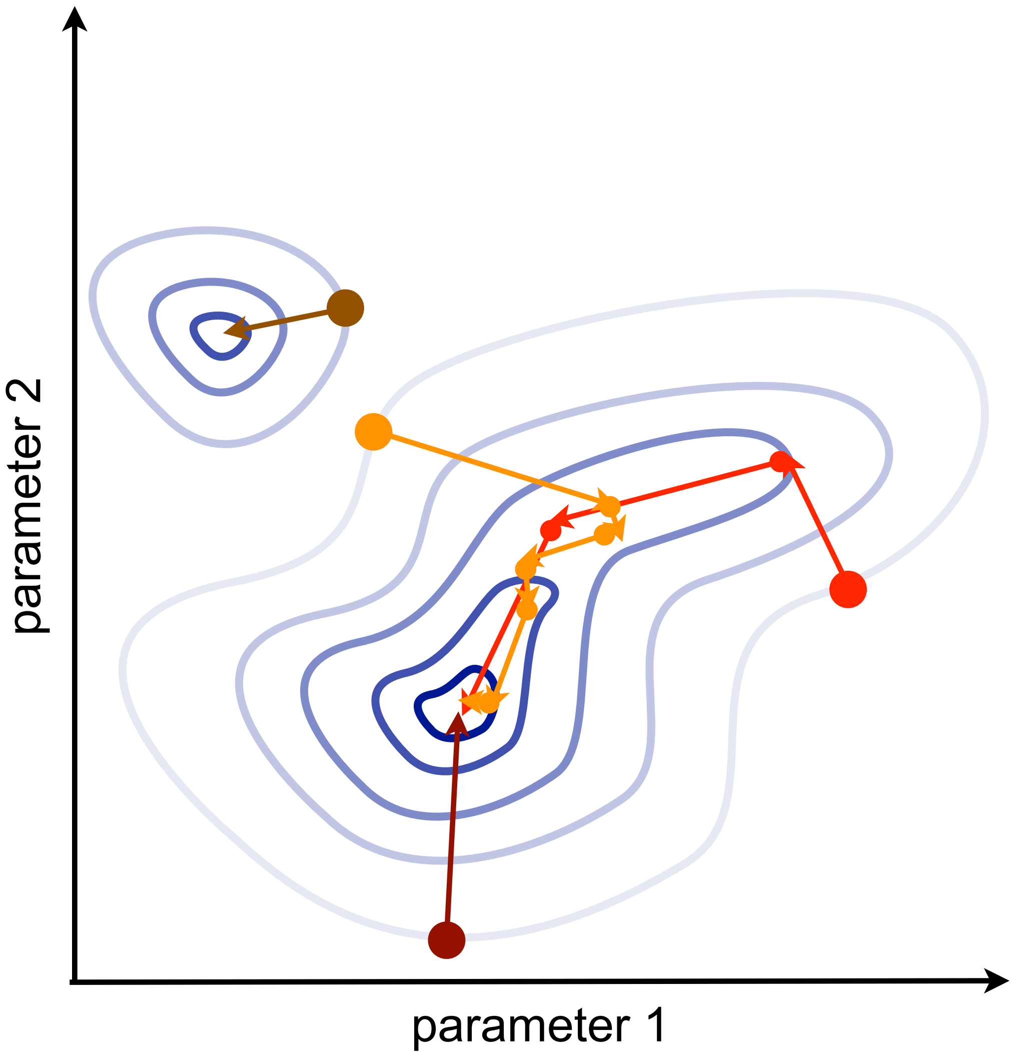


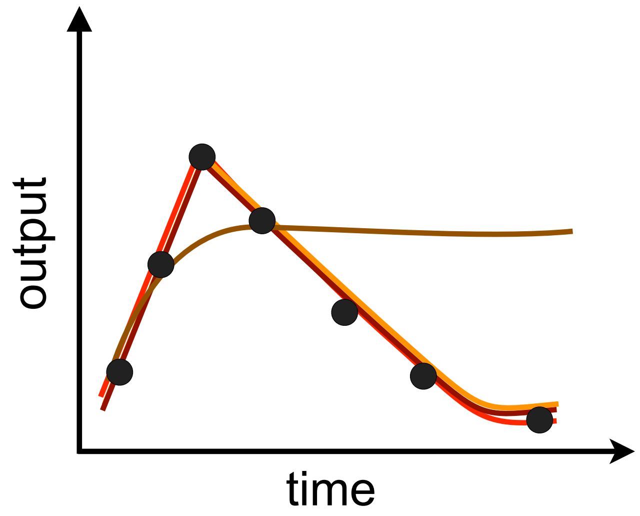
Multi-Start Optimization



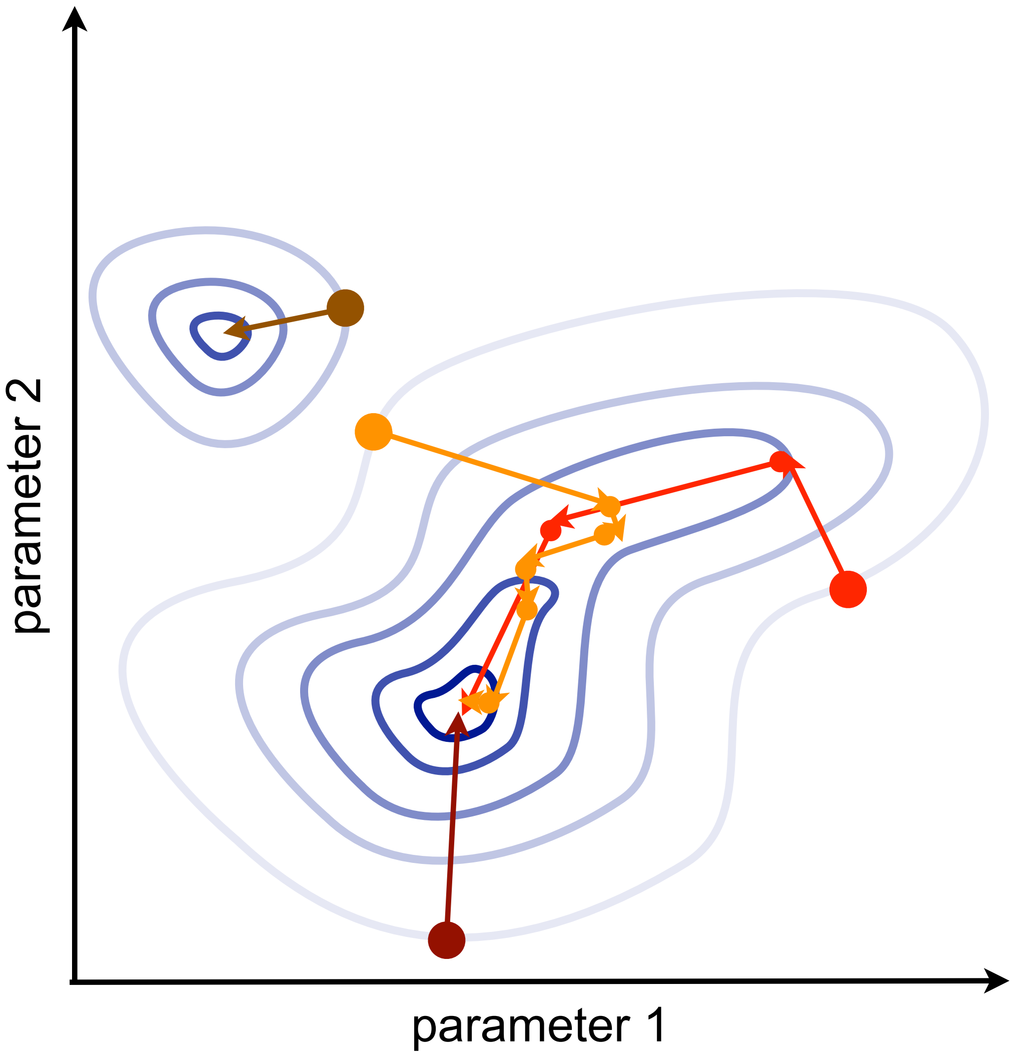


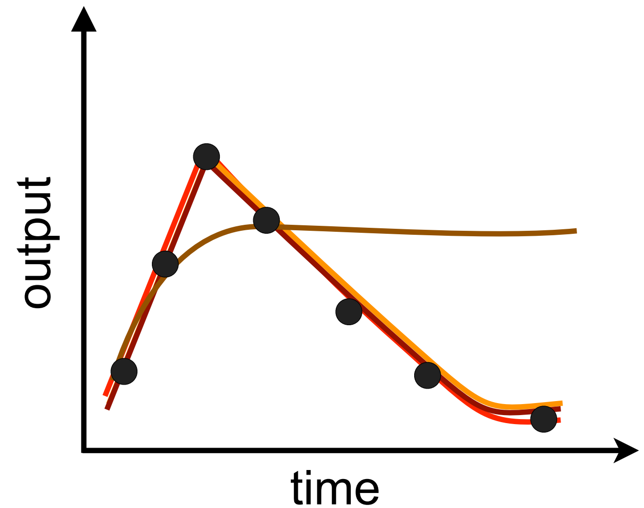
Multi-Start Optimization

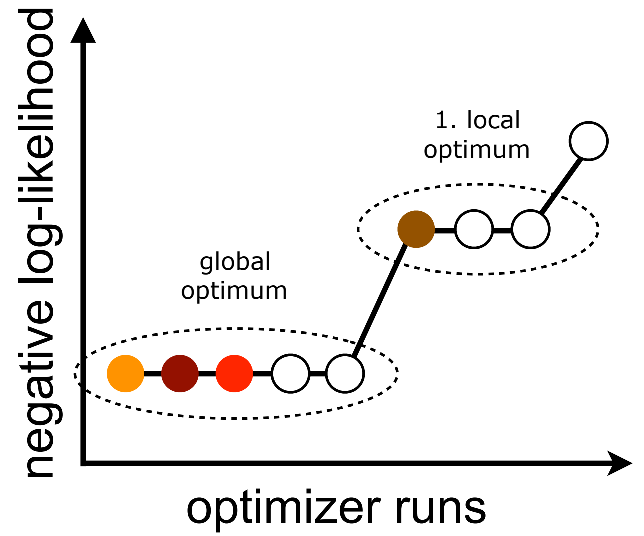




Multi-Start Optimization

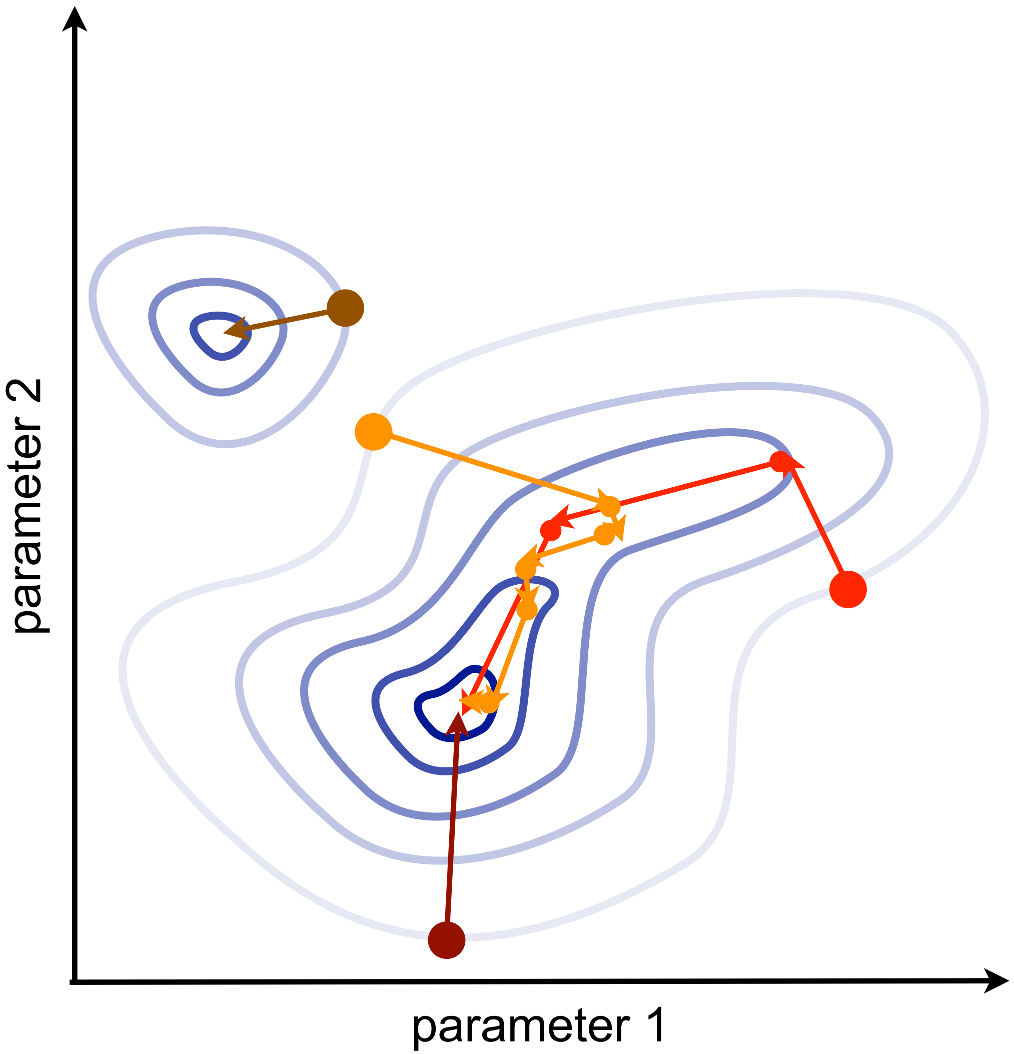


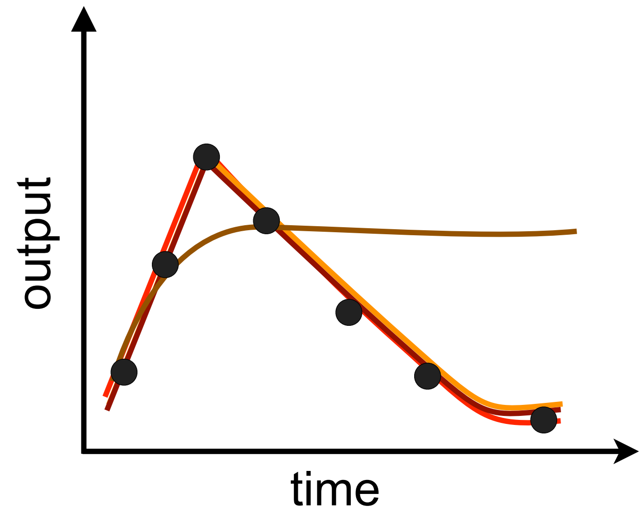


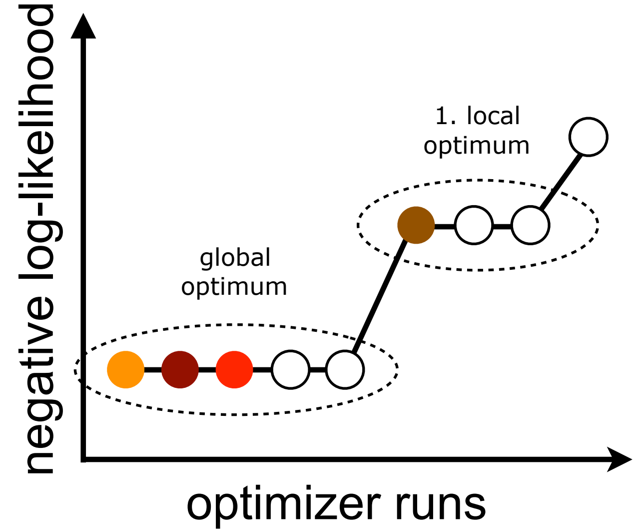


Multi-Start Optimization

https://github.com/ICB-DCM/PESTO
https://github.com/ICB-DCM/AMICI







Parameter Estimation

ODE model:

$$\frac{dx}{dt} = f(\theta, x(t,\theta)), \quad x(0,\theta) = x_0(\theta) \quad \textit{dynamics}$$

$$y(t) = h(\theta, x(t,\theta)) \quad \textit{observables}$$

Measurements: $\bar{y}_k = h(\theta, x(t_k,\theta)) + \varepsilon_k, \quad \varepsilon_k \sim \mathcal{N}(0,\sigma^2), \quad k = 1,\dots,n_t$

Minimize the negative log likelihood function:

$$\min_{\theta} \left\{ J(\theta) = \frac{1}{2} \sum_k log(2\pi\sigma^2) + \left(\frac{\bar{y}_k - h(\theta, x(t_k,\theta))}{\sigma} \right)^2 \right\}$$

Optimization problem with n_θ parameters

Problem Statement

ODE model:

$$\frac{dx}{dt} = f(\theta, x(t,\theta)), \quad x(0,\theta) = x_0(\theta) \quad \textit{dynamics}$$

$$y(t) = c \cdot h(\theta, x(t,\theta)) \quad \textit{observables}$$

Measurements that provide relative data:

$$\bar{y}_k = c \cdot h(\theta, x(t_k,\theta)) + \varepsilon_k, \quad \varepsilon_k \sim \mathcal{N}(0, \sigma^2), \quad k = 1, \ldots, n_t$$

with unknown variance σ^2 of the measurement noise

and unknown proportionality factor c

Standard Approach

ODE model:

$$\frac{dx}{dt} = f(\theta, x(t,\theta)), \quad x(0,\theta) = x_0(\theta) \quad \textit{dynamics}$$

$$y(t) = c \cdot h(\theta, x(t,\theta)) \quad \textit{observables}$$

Measurements that provide relative data:

$$\bar{y}_k = c \cdot h(\theta, x(t_k,\theta)) + \varepsilon_k, \quad \varepsilon_k \sim \mathcal{N}(0, \sigma^2), \quad k = 1, \ldots, n_t$$

with unknown variance σ^2 of the measurement noise

and unknown proportionality factor c

Minimize the negative log likelihood function:

$$\min_{\theta, c, \sigma^2} \left\{ J(\theta, c, \sigma^2) = \frac{1}{2} \sum_k log(2\pi\sigma^2) + \left(\frac{\bar{y}_k - c \cdot h(\theta, x(t_k,\theta))}{\sigma} \right)^2 \right\}$$

Standard Approach

ODE model:

$$\frac{dx}{dt} = f(\theta, x(t,\theta)), \quad x(0,\theta) = x_0(\theta) \quad \textit{dynamics}$$

$$y(t) = c \cdot h(\theta, x(t,\theta)) \quad \textit{observables}$$

Measurements that provide relative data:

$$\bar{y}_k = c \cdot h(\theta, x(t_k,\theta)) + \varepsilon_k, \quad \varepsilon_k \sim \mathcal{N}(0, \sigma^2), \quad k = 1, \ldots, n_t$$

with unknown variance σ^2 of the measurement noise

and unknown proportionality factor c

Minimize the negative log likelihood function:

$$\min_{\theta, c, \sigma^2} \left\{ J(\theta, c, \sigma^2) = \frac{1}{2} \sum_k log(2\pi\sigma^2) + \left(\frac{\bar{y}_k - c \cdot h(\theta, x(t_k,\theta))}{\sigma} \right)^2 \right\}$$

Number of parameters:
n_θ + number of proportionality factors + number of variances

Hierarchical Approach

Hierarchical optimization problem:

$$\min_{\theta} \left\{ \min_{c,\sigma^2} \left\{ J(\theta, c, \sigma^2) = \frac{1}{2} \sum_k log(2\pi\sigma^2) + \left(\frac{\bar{y}_k - c \cdot h(\theta, x(t_k, \theta))}{\sigma} \right)^2 \right\} \right\}$$

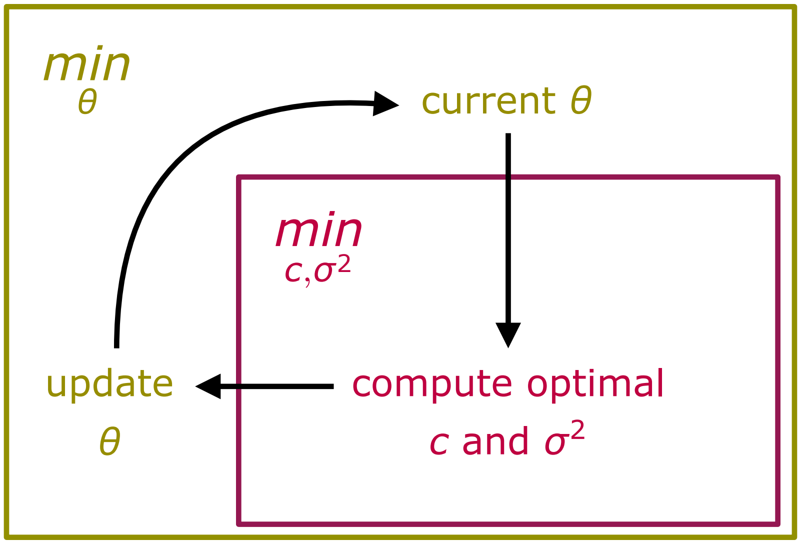
Hierarchical Approach

Hierarchical optimization problem:

$$\min_{\theta}\left\{\min_{c,\sigma^2}\left\{J(\theta,c,\sigma^2)=\frac{1}{2}\sum_k log(2\pi\sigma^2)+\left(\frac{\bar{y}_k-c\cdot h(\theta,x(t_k,\theta))}{\sigma}\right)^2\right\}\right\}$$

In each step of the optimization:

1. Calculate optimal proportionality factors and variances analytically for a given θ
2. Use analytical results to do the update step in the outer optimization to estimate the remaining dynamical parameters



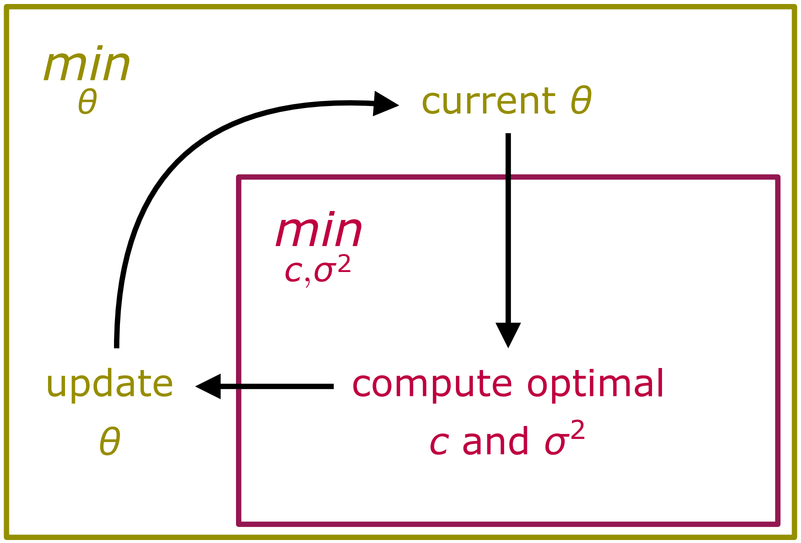
Hierarchical Approach

Hierarchical optimization problem:

$$\min_{\theta} \left\{ \min_{c,\sigma^2} \left\{ J(\theta, c, \sigma^2) = \frac{1}{2} \sum_k \log(2\pi\sigma^2) + \left(\frac{\bar{y}_k - c \cdot h(\theta, x(t_k, \theta))}{\sigma} \right)^2 \right\} \right\}$$

In each step of the optimization:

1. Calculate optimal proportionality factors and variances analytically for a given θ
2. Use analytical results to do the update step in the outer optimization to estimate the remaining dynamical parameters



Advantage: Outer optimization problem has n_θ parameters

Analytical Derivation of the Proportionality Factors and Variances

Necessary first order optimality condition:

Is $(\hat{\theta}, \hat{c}, \hat{\sigma}^2)^T$ a local minimum of J, with J continuously differentiable, then

$$\nabla J(\hat{\theta}, \hat{c}, \hat{\sigma}^2) = 0.$$

Analytical Derivation of the Proportionality Factors and Variances

Necessary first order optimality condition:

Is $(\hat{\theta}, \hat{c}, \hat{\sigma}^2)^T$ a local minimum of J, with J continuously differentiable, then

$$\nabla J(\hat{\theta}, \hat{c}, \hat{\sigma}^2) = 0.$$

$$J(\theta, c, \sigma^2) = \frac{1}{2} \sum_k log(2\pi\sigma^2) + \left(\frac{\bar{y}_k - c \cdot h(\theta, x(t_k, \theta))}{\sigma} \right)^2$$

Analytical Derivation of the Proportionality Factors and Variances

Necessary first order optimality condition:

Is $(\hat{\theta}, \hat{c}, \hat{\sigma}^2)^T$ a local minimum of J, with J continuously differentiable, then

$$\nabla J(\hat{\theta}, \hat{c}, \hat{\sigma}^2) = 0.$$

$$J(\theta, c, \sigma^2) = \frac{1}{2} \sum_k log(2\pi\sigma^2) + \left(\frac{\bar{y}_k - c \cdot h(\theta, x(t_k, \theta))}{\sigma} \right)^2$$

$$\left. \frac{\partial J}{\partial c} \right|_{(\theta, \hat{c}, \hat{\sigma}^2)} \overset{!}{=} 0$$

$$-\frac{1}{\hat{\sigma}^2} \sum_k \bar{y}_k h(\theta, x(t_k, \theta)) - \hat{c} \cdot h(\theta, x(t_k, \theta))^2 = 0$$

$$\sum_k \bar{y}_k h(\theta, x(t_k, \theta)) = \hat{c} \sum_k h(\theta, x(t_k, \theta))^2$$

Analytical Derivation of the Proportionality Factors and Variances

Necessary first order optimality condition:

Is $(\hat{\theta}, \hat{c}, \hat{\sigma}^2)^T$ a local minimum of J, with J continuously differentiable, then

$$\nabla J(\hat{\theta}, \hat{c}, \hat{\sigma}^2) = 0.$$

$$J(\theta, c, \sigma^2) = \frac{1}{2} \sum_k log(2\pi\sigma^2) + \left(\frac{\bar{y}_k - c \cdot h(\theta, x(t_k, \theta))}{\sigma} \right)^2$$

$$\left. \frac{\partial J}{\partial c} \right|_{(\theta, \hat{c}, \hat{\sigma}^2)} \overset{!}{=} 0$$

$$-\frac{1}{\hat{\sigma}^2} \sum_k \bar{y}_k h(\theta, x(t_k, \theta)) - \hat{c} \cdot h(\theta, x(t_k, \theta))^2 = 0 \longrightarrow \hat{c}(\theta) = \frac{\sum_k \bar{y}_k h(\theta, x(t_k, \theta))}{\sum_k h(\theta, x(t_k, \theta))^2}$$

$$\sum_k \bar{y}_k h(\theta, x(t_k, \theta)) = \hat{c} \sum_k h(\theta, x(t_k, \theta))^2$$

Analytical Derivation of the Proportionality Factors and Variances

Necessary first order optimality condition:

Is $(\hat{\theta}, \hat{c}, \hat{\sigma}^2)^T$ a local minimum of J, with J continuously differentiable, then

$$\nabla J(\hat{\theta}, \hat{c}, \hat{\sigma}^2) = 0.$$

$$J(\theta, c, \sigma^2) = \frac{1}{2} \sum_k log(2\pi\sigma^2) + \left(\frac{\bar{y}_k - c \cdot h(\theta, x(t_k, \theta))}{\sigma} \right)^2$$

Analytical Derivation of the Proportionality Factors and Variances

Necessary first order optimality condition:

Is $(\hat{\theta}, \hat{c}, \hat{\sigma}^2)^T$ a local minimum of J, with J continuously differentiable, then

$$\nabla J(\hat{\theta}, \hat{c}, \hat{\sigma}^2) = 0.$$

$$J(\theta, c, \sigma^2) = \frac{1}{2} \sum_k log(2\pi\sigma^2) + \left(\frac{\bar{y}_k - c \cdot h(\theta, x(t_k, \theta))}{\sigma}\right)^2$$

$$\left.\frac{\partial J}{\partial \sigma^2}\right|_{(\theta, \hat{c}, \hat{\sigma}^2)} \stackrel{!}{=} 0$$

$$\frac{1}{2\hat{\sigma}^2} \sum_k 1 - \frac{\left(\bar{y}_k - \hat{c} \cdot h(\theta, x(t_k, \theta))\right)^2}{\hat{\sigma}^2} = 0$$

$$\sum_k 1 = \frac{1}{\hat{\sigma}^2} \sum_k \left(\bar{y}_k - \hat{c} \cdot h(\theta, x(t_k, \theta))\right)^2$$

Analytical Derivation of the Proportionality Factors and Variances

Necessary first order optimality condition:

Is $(\hat{\theta}, \hat{c}, \hat{\sigma}^2)^T$ a local minimum of J, with J continuously differentiable, then

$$\nabla J(\hat{\theta}, \hat{c}, \hat{\sigma}^2) = 0.$$

$$J(\theta, c, \sigma^2) = \frac{1}{2} \sum_k log(2\pi\sigma^2) + \left(\frac{\bar{y}_k - c \cdot h(\theta, x(t_k, \theta))}{\sigma} \right)^2$$

$$\left. \frac{\partial J}{\partial \sigma^2} \right|_{(\theta, \hat{c}, \hat{\sigma}^2)} \stackrel{!}{=} 0$$

$$\frac{1}{2\hat{\sigma}^2} \sum_k 1 - \frac{\left(\bar{y}_k - \hat{c} \cdot h(\theta, x(t_k, \theta))\right)^2}{\hat{\sigma}^2} = 0 \longrightarrow \hat{\sigma}^2(\theta) = \frac{1}{n_t} \sum_k \left(\bar{y}_k - \hat{c} \cdot h(\theta, x(t_k, \theta))\right)^2$$

$$\sum_k 1 = \frac{1}{\hat{\sigma}^2} \sum_k \left(\bar{y}_k - \hat{c} \cdot h(\theta, x(t_k, \theta))\right)^2$$

Several experiments, observables and replicates

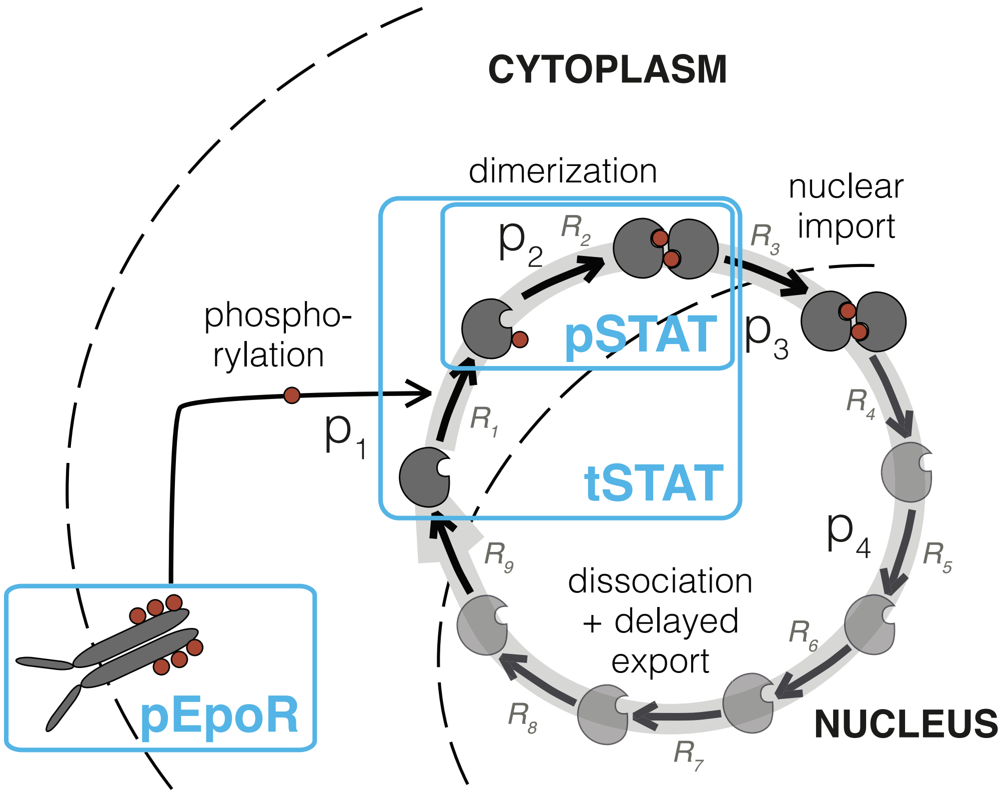
Proportionality factors c_{il} and variances σ_{il}^2 for each observable and replicate, $i = 1, \ldots, n_y$, $l = 1, \ldots, n_r$

$$J(\theta, c, \sigma^2) = \frac{1}{2} \sum_{j=1}^{n_e} \sum_{i \in I_j} \sum_{l=1}^{n_{r_{ji}}} \sum_{k=1}^{n_{t_{jil}}} \left[log(2\pi\sigma_{il}^2) + \frac{\left(\bar{y}_{jilk} - c_{il} \cdot h_{ji}(\theta, x(t_k, \theta))\right)^2}{\sigma_{il}^2} \right]$$

Analytical solutions for the proportionality factors and the variances:

$$\hat{c}_{il}(\theta) = \frac{\sum_{j \in \mathcal{E}_i} \sum_{k=1}^{n_{t_{jil}}} \bar{y}_{jilk} h_{ji}(\theta, x(t_k, \theta))}{\sum_{j \in \mathcal{E}_i} \sum_{k=1}^{n_{t_{jil}}} h_{ji}(\theta, x(t_k, \theta))^2} \qquad \hat{\sigma}_{il}^2(\theta) = \frac{\sum_{j \in \mathcal{E}_i} \sum_{k=1}^{n_{t_{jil}}} \left(\bar{y}_{jilk} - \hat{c}_{il}(\theta) h_{ji}(\theta, x(t_k, \theta))\right)^2}{\sum_{j \in \mathcal{E}_i} n_{t_{jil}}}$$

JAK-STAT Signaling Pathway



R_1 : STAT $\xrightarrow{p_1}$ pSTAT

R_2 : 2pSTAT $\xrightarrow{p_2}$ pSTAT:pSTAT

R_3 : pSTAT:pSTAT $\xrightarrow{p_2}$ npSTAT:npSTAT

R_4 : npSTAT:npSTAT $\xrightarrow{p_4}$ 2nSTAT1

R_5 : nSTAT1 $\xrightarrow{p_4}$ nSTAT2

R_6 : nSTAT2 $\xrightarrow{p_4}$ nSTAT3

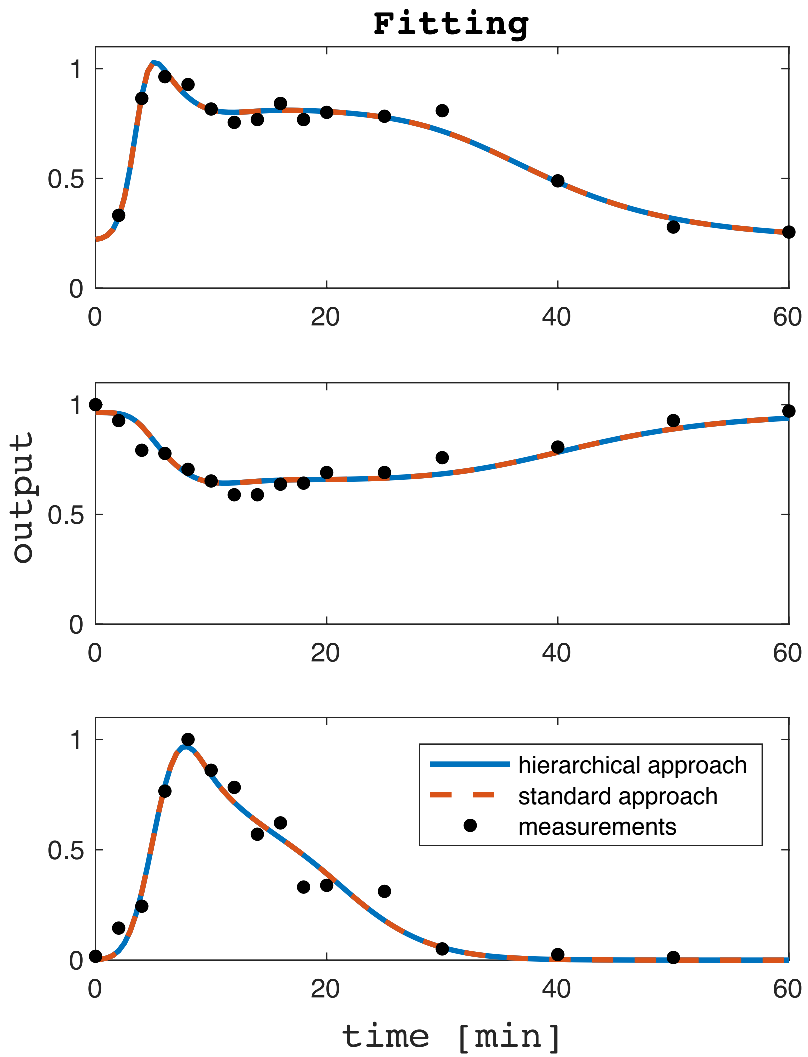
R_7 : nSTAT3 $\xrightarrow{p_4}$ nSTAT4

R_8 : nSTAT4 $\xrightarrow{p_4}$ nSTAT5

R_9 : nSTAT5 $\xrightarrow{p_4}$ STAT

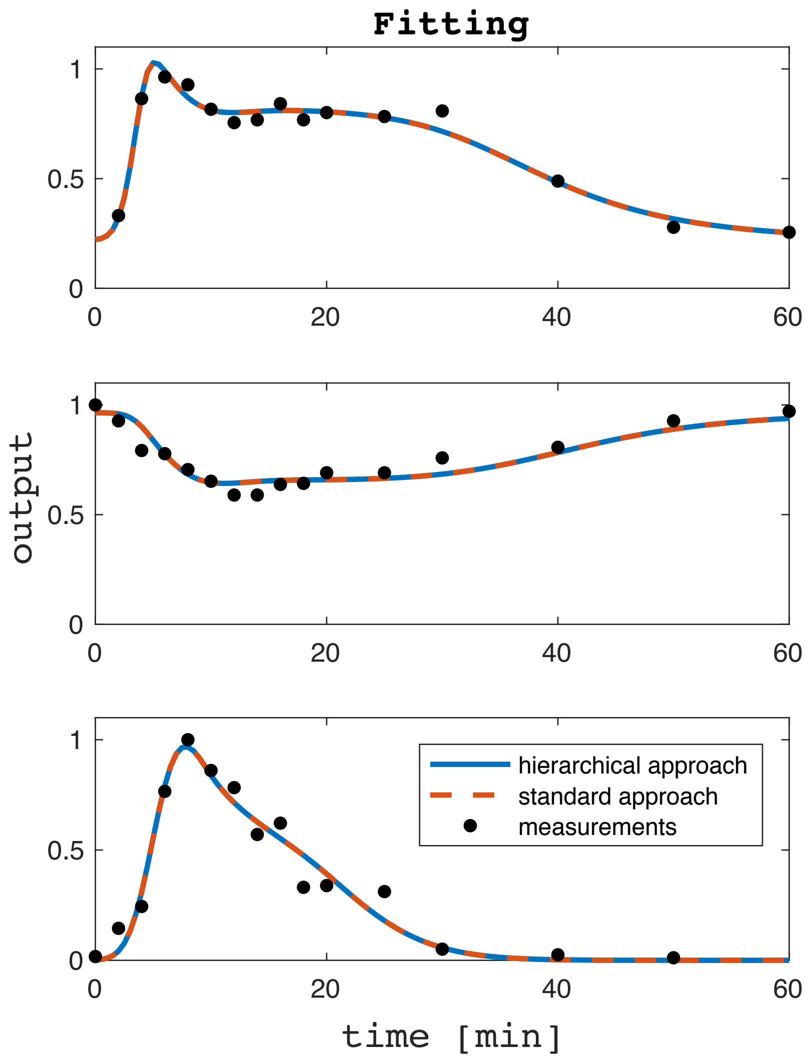
Fröhlich F et al. (2016) PLoS Comput Biol 12(7)

Fitting and Convergence



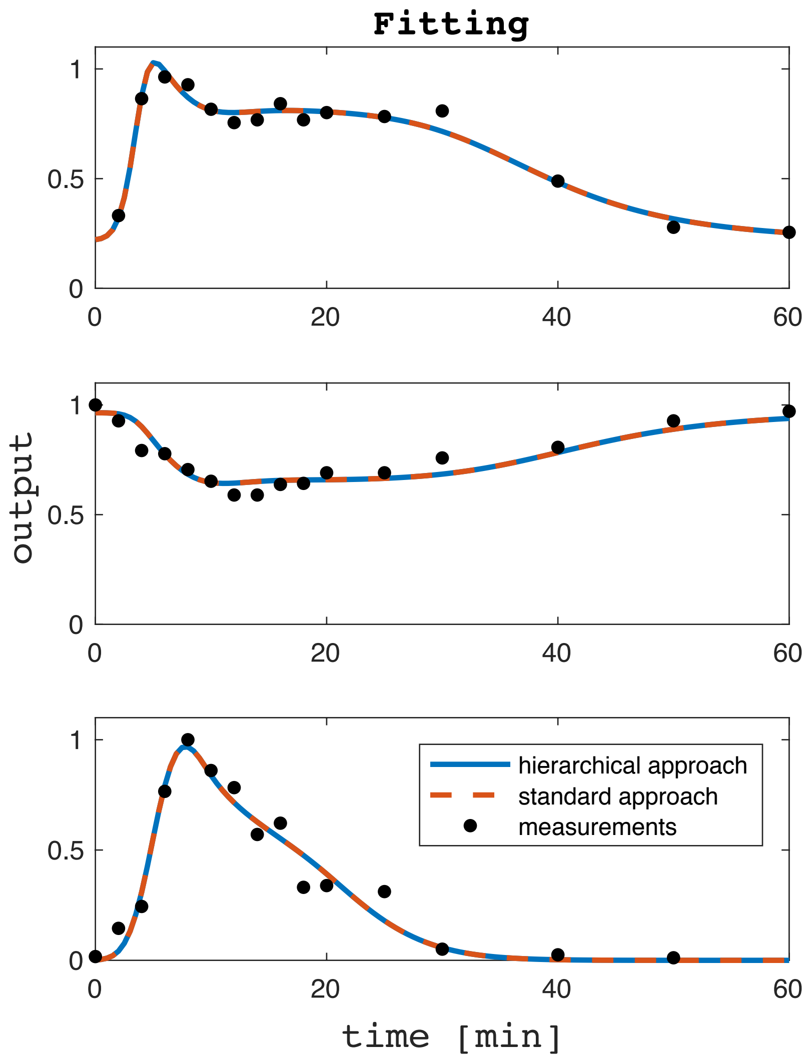
Data from:
Swaneye *et al.* (2003) *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.0237333100

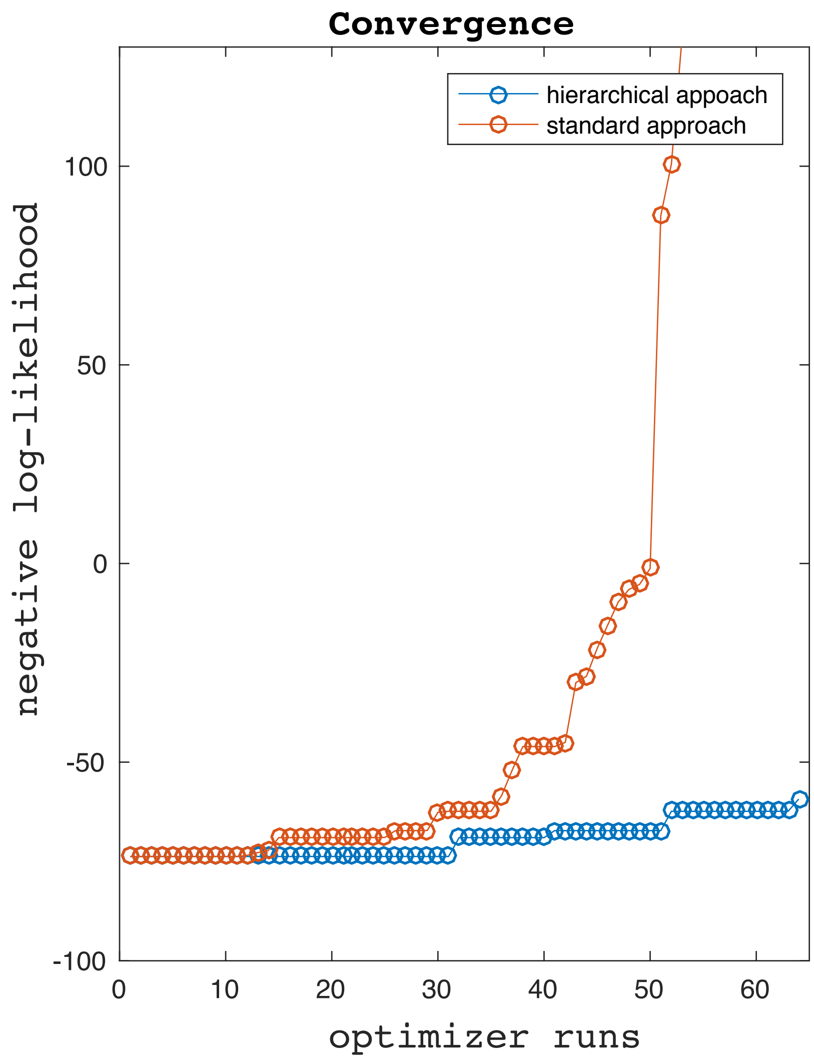
Fitting and Convergence



Fits of same quality

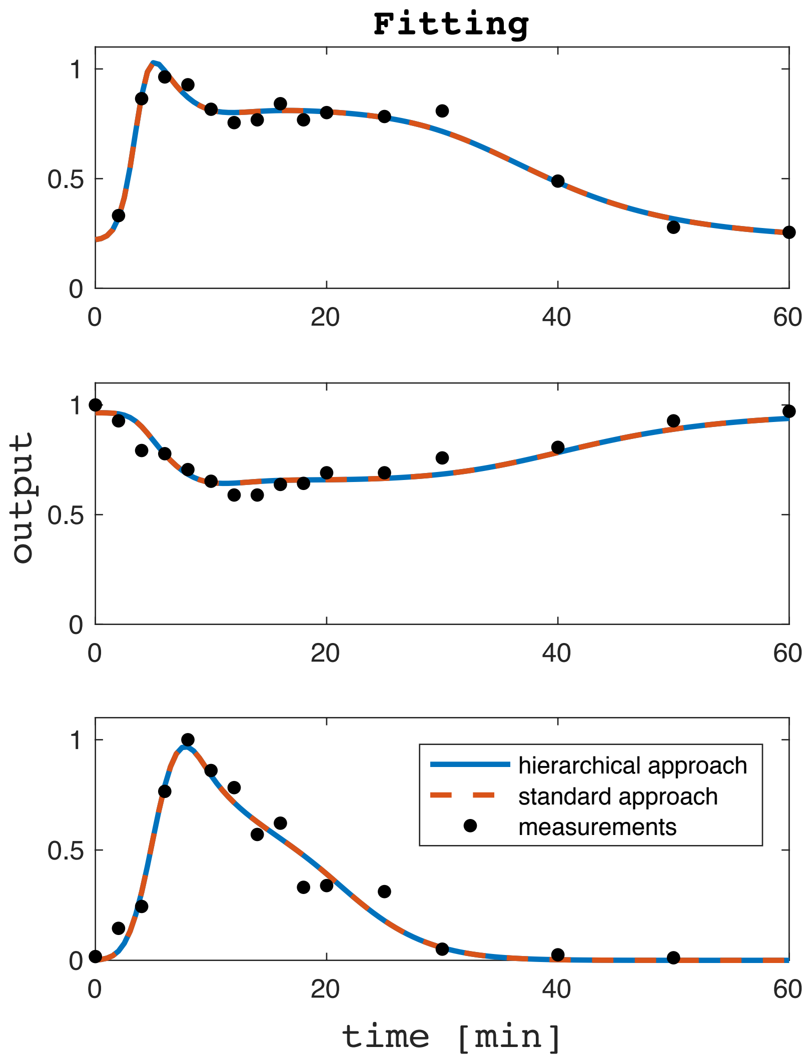
Fitting and Convergence

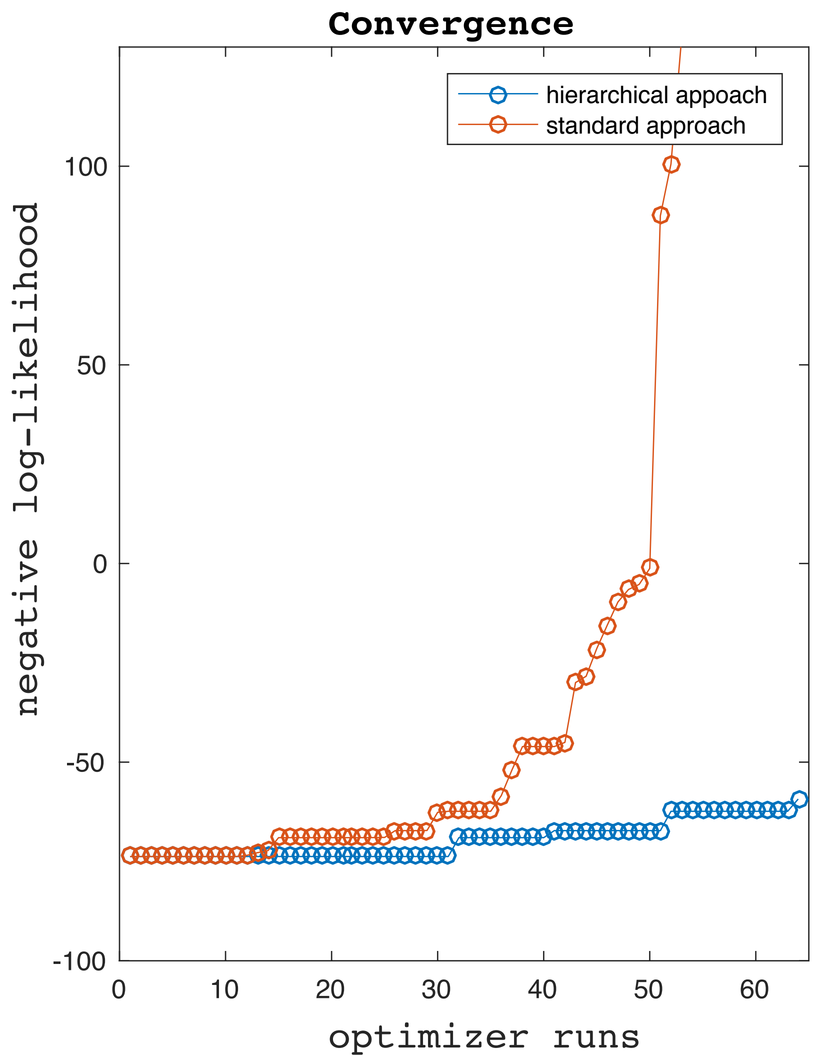




Fits of same quality

Fitting and Convergence

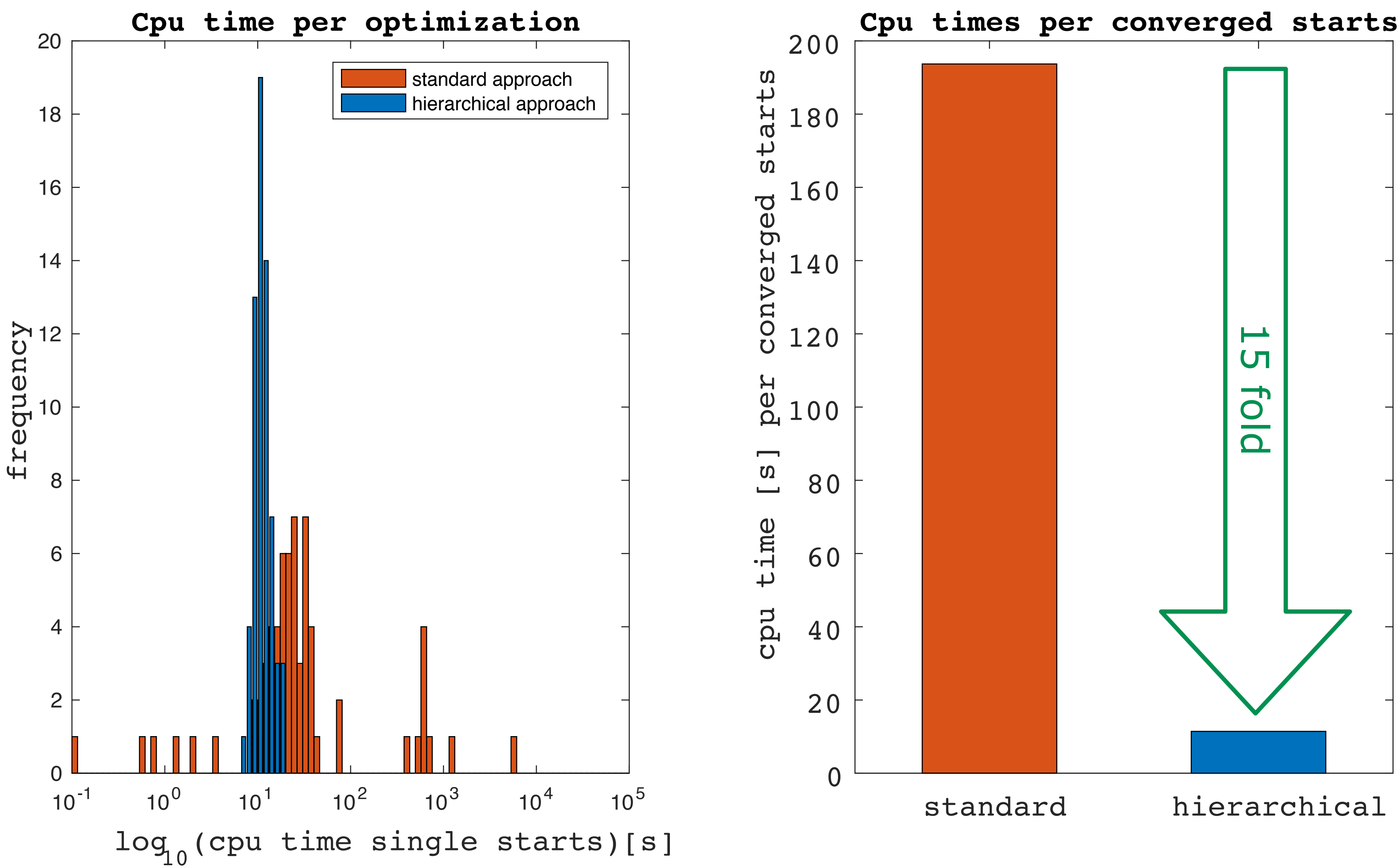




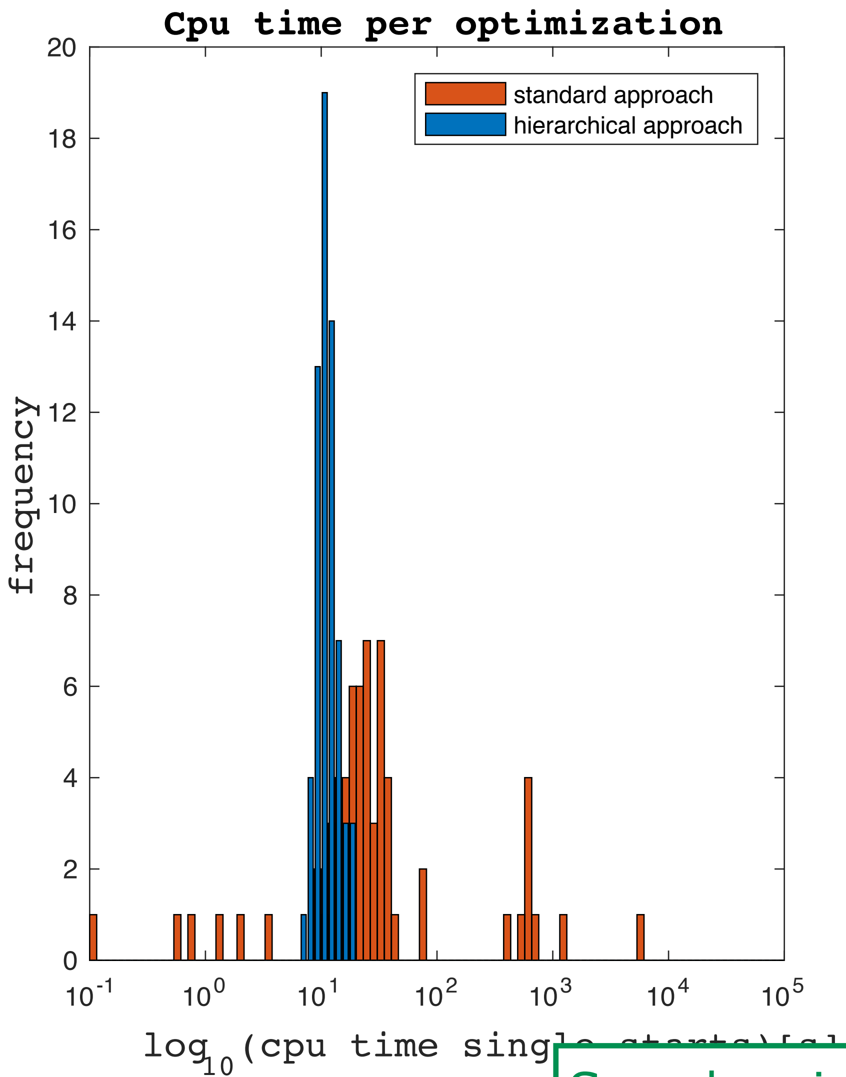
Fits of same quality

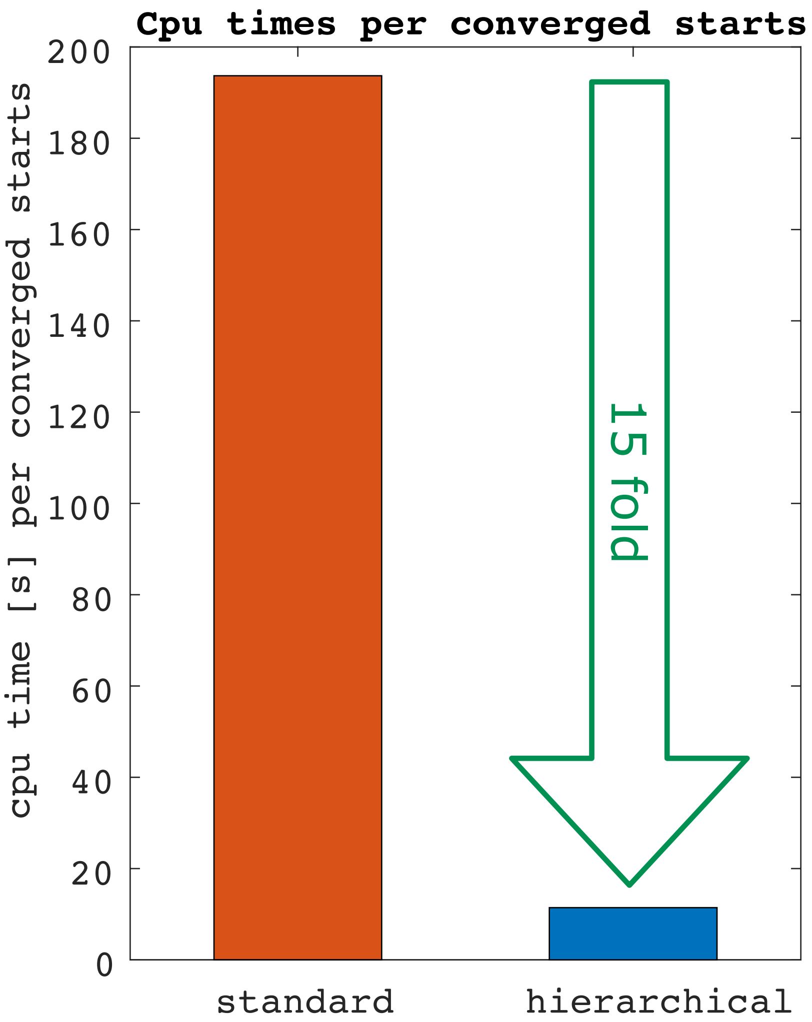
Better convergence

Comparison of Computation Times



Comparison of Computation Times





Speed up in computation time

Summary

- Development of an hierarchical approach to parameter estimation for models with relative data
- Analytical derivation of equations for proportionality factors and variances
- Implementation of the method
- Evaluation of the method for JAK-STAT signaling pathway with better convergence results and a substantial speed up in computation time

Acknowledgements

Institute of Computational Biology
Jan Hasenauer
Carolin Loos

Data-driven Computational Modelling

CanPathPro

European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding through the European Union's Horizon 2020 research and innovation programme under grant agreement no. 686282