Bayesian Estimation and Analysis of Pathway Models using Kernel-enhanced Particle Filters

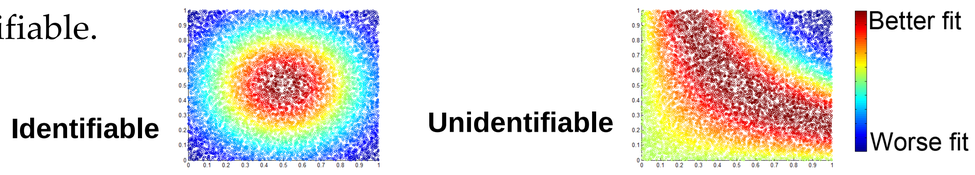
Benjamin Gyori[1,2] David Hsu[2]

1 NUS Graduate School for Integrative Sciences & Engineering (NGS), National University of Singapore;
2 School of Computing, National University of Singapore.

INTRODUCTION

Motivation

Pathway model parameters are commonly under-constrained and unidentifiable.



Large number of parameters fit measurement data so a Bayesian posterior estimate is more adequate than a single best estimate (ML/MAP) for making predictions.

Bayesian framework

Recover posterior over parameters with respect to measurement data

$$(1) \qquad p(\theta \mid Y) = \frac{p(Y \mid \theta)\, p(\theta)}{\int p(Y \mid \theta)\, p(\theta) d\theta}$$

Make predictions based on full posterior (model averaging)

$$(2) \qquad p(x \mid Y) = \int p(x \mid \theta)\, p(\theta \mid Y) d\theta \approx \sum_{i=1}^{N} w^{(i)}\, p(x \mid \theta^{(i)})$$

Non-linear system model and high-dimensional parameter space makes approximating the posterior computationally challenging.

Particle filtering

Produces a sample based, **sequential approximation** of the posterior using **weighted samples,** which can be directly used to **make predictions** (2). However, **degenerate estimates** can result from limited sample size.

Previously proposed solutions

- Use very high number of particles (~10^8)
- Add random noise to the particles at each step

Our approach

- Use an MCMC **kernel to diversify** the sample set at each step of the particle filter while **preserving the posterior**

METHOD

- State extension: joint state and parameter space
- Samples obtained from $p(x, \theta \mid Y_{1:t})$ at each time step
- Propagate and update samples until all data is used

Kernel enhanced particle filter:

Sample prior $(x_0^{(i)}, \theta_0^{(i)}) \sim p_0(x, \theta)$

Loop over all measurement time points

1) **Propose** next state and **update** likelihood weights

$$x_t^{(i)} \sim p(x_t \mid x_{t-1}^{(i)}, \theta_{t-1}^{(i)}) \quad w_t^{(i)} \sim w_{t-1}^{(i)} L(y_t \mid x_t^{(i)})$$

2) **Resample** if weights degrade

$$(x_t^{(i)}, \theta_t^{(i)}) \sim Mult(\{w_t^{(j)}\}_{j=1:N})$$

3) **Diversify** resampled particles

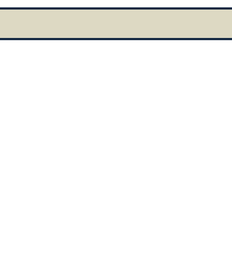
3.1) **Choose directions** $d_1 \ldots d_n$ based on particle spread

3.2) **Propose** new parameter along chosen directions

$$\theta_t'^{(i)} \sim T_{d_j}(\theta_t'^{(i)} \mid \theta_t^{(i)})$$

3.3) **Accept** according to the Metropolis-Hasting kernel

$$\mathrm{P(accept)} = \min\left(1, \frac{T_{d_j}(\theta_t^{(i)} \mid \theta_t'^{(i)})\, p(Y_{1:t} \mid \theta_t'^{(i)})}{T_{d_j}(\theta_t'^{(i)} \mid \theta_t^{(i)})\, p(Y_{1:t} \mid \theta_t^{(i)})}\right)$$



3.4) **Adapt** T based on acceptance rates.

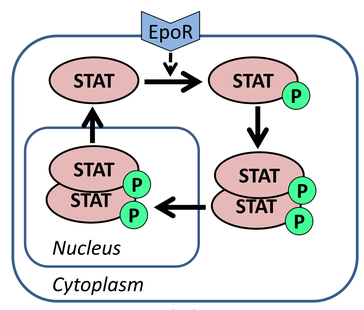
MCMC kernel ensures that particles are diversified while still being distributed according to the the posterior $p(x, \theta \mid Y_{1:t})$ at current time step.

RESULTS

JAK-STAT signaling pathway

- Dimerization, nuclear transport of STAT protein upon EPO stimulation (1A)
- ODE model: 7 unknown parameters (1B)
- Experimental data: semi-quantitative, sum of individual species (1C)

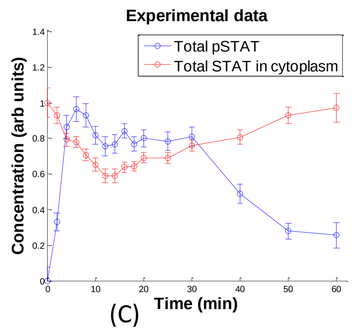
(1)



(A)

$$[STAT]' = -p_1 x_1 [EpoR] + 2 p_4 [npSTAT_2]_r$$
$$[pSTAT]' = +p_1 x_1 [EpoR] - p_2 [pSTAT]^2$$
$$[pSTAT_2]' = \frac{1}{2} p_2 [pSTAT]^2 - p_3 [pSTAT_2]$$
$$[npSTAT_2]' = p_3 [pSTAT_2] - p_4 [npSTAT_2]_r$$

(B)



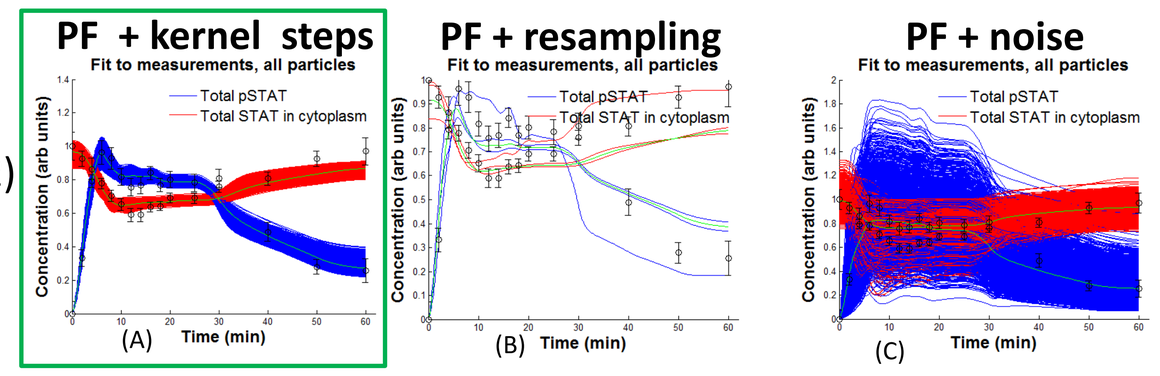
(C)

Particle filter

- Uniform priors on logarithmic scale: 10^{-5}-10^2
- Number of particles: 10^4

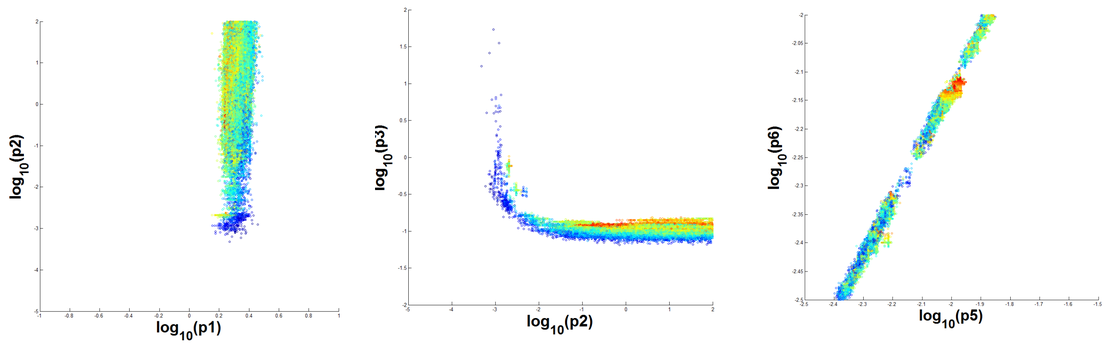
We compare our particle filter enhanced by an MCMC kernel to previously proposed solutions (2). Resampling only [1] results in collapsed particles, while adding noise [2] diversifies particles but gives an inaccurate approximation to the posterior. Our particle filter produces a non-degenerate, yet accurate fit to measurement data.

(2)



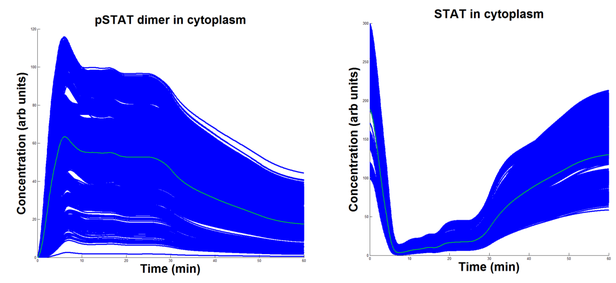
Large areas of parameter space contain acceptable models with respect to the measurement data (3). A single best parameter estimate would not provide an appropriate description.

(3)



While fit to measurement data is tight, unobserved species can have highly varying behavior (4). Using out method, these uncertainties can be quantified.

(4)



DISCUSSION

Our method allows the efficient approximation of a Bayesian posterior over pathway model parameters.

Main advantages:

- Sequential sampling approximates a series of simpler distributions.
- Sample diversification with MCMC kernel results in accurate, complete posterior.
- Kernel design ensures that acceptance rates are high at each step.

By obtaining an accurate parameter posterior, applications such as Bayesian model selection and experimental design become more effective.

[1] Koh CH. et al. (2010) DA 1.0: parameter estimation of biological pathways using data assimilation approach. Bioinformatics, 26. 1794-1796
[2] Nakamura K. et al. (2009) Parameter estimation of in silico biological pathways with particle filtering towards a petascale computing. Pac Symp Biocomput 14: 227–238.