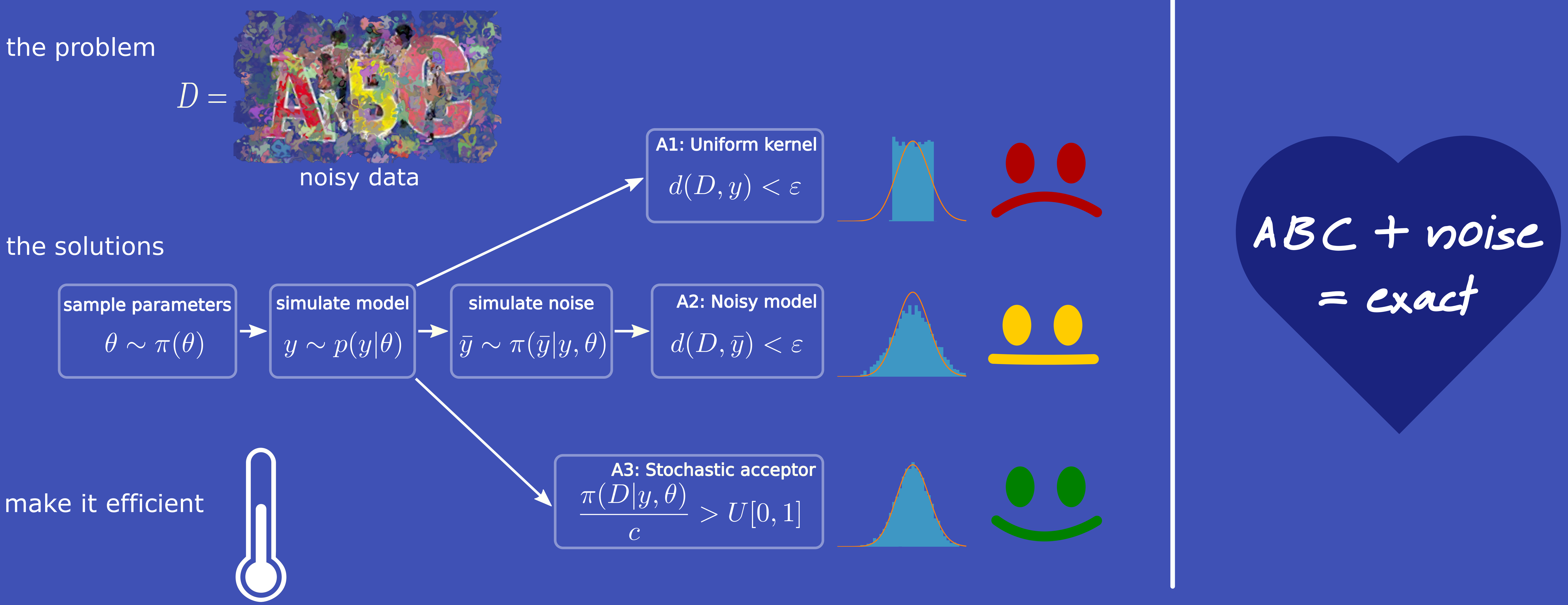
HelmholtzZentrum münchen
German Research Center for Environmental Health

Efficient noise assessment in sequential ABC

Yannik Schaelte[1,2], Jan Hasenauer[1,2,3]

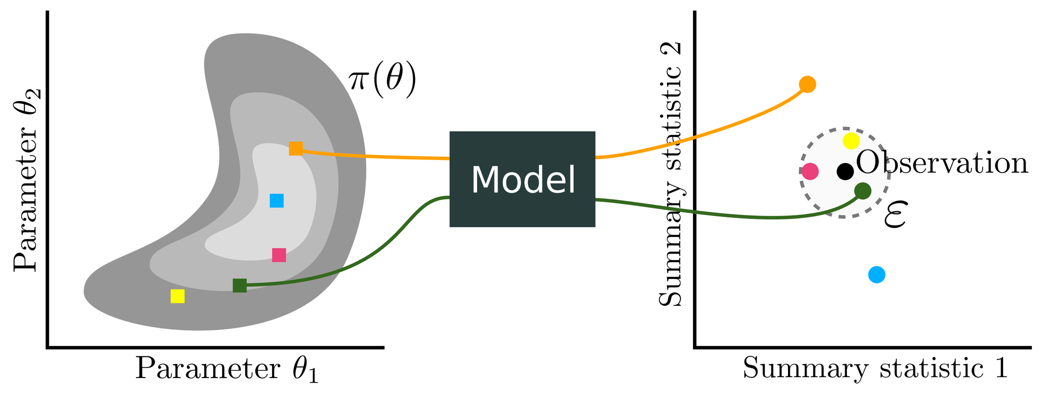
1: Helmholtz Center Munich, Institute of Computational Biology, 2: Technical University Munich, Department of Mathematics, 3: University Bonn, Research Unit Biomathematics

Graphical abstract



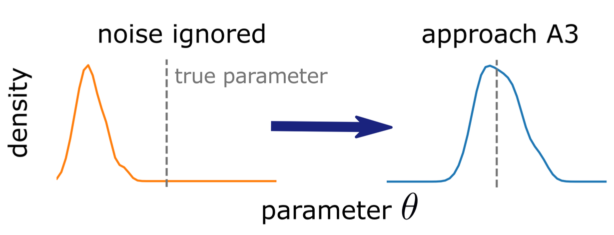
Mini-Intro: Basic Rejection ABC

- Approximate Bayesian Computation (ABC) enables **likelihood-free inference**



Measurement noise in ABC

- experimental data usually noise corrupted
- easy to ignore → **wrong parameter estimates**
- **modified acceptor** (A3) yields **exact likelihood-free inference** from the true posterior



Integrate in sequential ABC

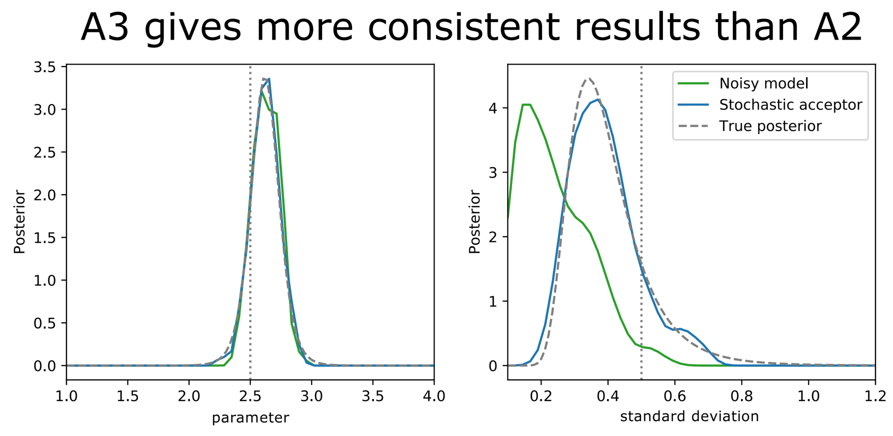
- problem: Rejection ABC **inefficient** → use ABC-SMC (Sequential Monte Carlo)
- instead of ε_t, **temperate** noise likelihoods via $T_t \searrow 1$,

$$\pi_{\text{ABC},t}(\theta|D) \propto \int \pi(D|y,\theta)^{1/T_t} p(y|\theta) dy \cdot \pi(\theta)$$

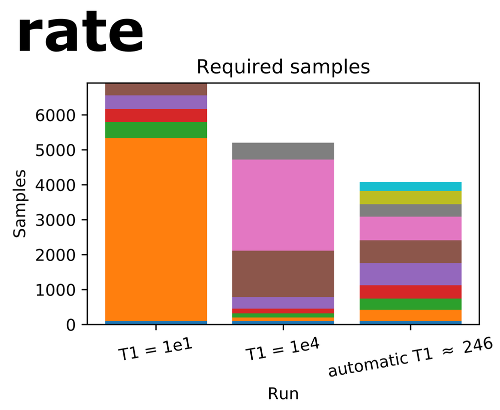
- how to choose T_t? update by geometric progression (based on ideas from parallel tempering)

Formulation allows estimating noise parameters

A3 gives more consistent results than A2



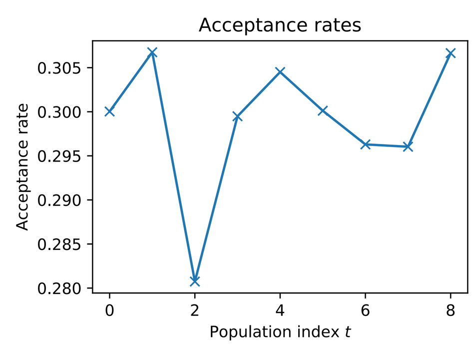
Choose T_1 by predicting acceptance rate



- problem: SMC **sensitive** to initial temperature
- idea: based on existing particles, **predict** the next population's **acceptance rate**

$$\gamma \approx \sum_i \frac{g_t(\theta_i^{(t-1)})}{g_{t-1}(\theta_i^{(t-1)})} \min\left[\left(\frac{\pi(D|y_i^{(t-1)},\theta_i^{(t-1)})}{c_t}\right)^{1/T_t}, 1\right]$$

- choose T_t (esp. T_1) to match a **target rate** γ_{target}



Automatically find a good c

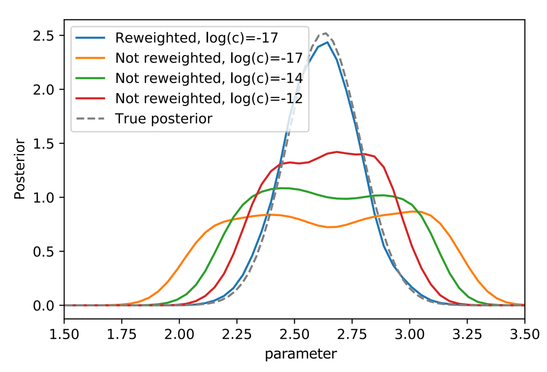
- problem: normalization constant c directly **affects acceptance rate**, difficult to predict
- idea: **update** c after each iteration to so-far maximum
- **avoid bias** (based on ideas from rejection control importance sampling): accept w. prob.

$$\min\left[\frac{\pi(D|y,\theta)}{c_t}, 1\right]^{1/T_t}$$

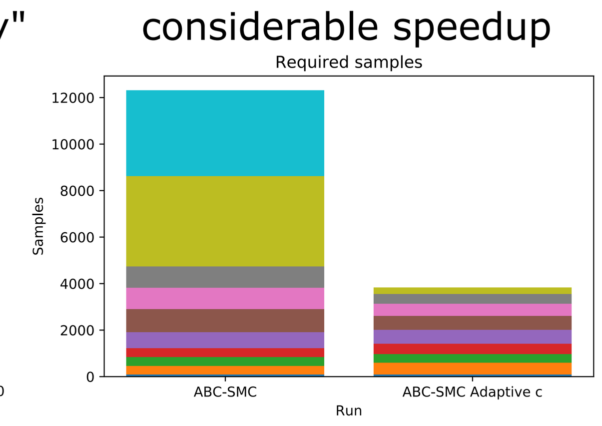
and **re-weight** by

$$w(y,\theta) \propto \frac{\pi(D|y,\theta)^{1/T_t}}{\min\left[\left(\frac{\pi(D|y,\theta)}{c_t}\right)^{1/T_t}, 1\right]} \cdot \frac{\pi(\theta)}{g_t(\theta)}$$

weights reduce "uniformity"

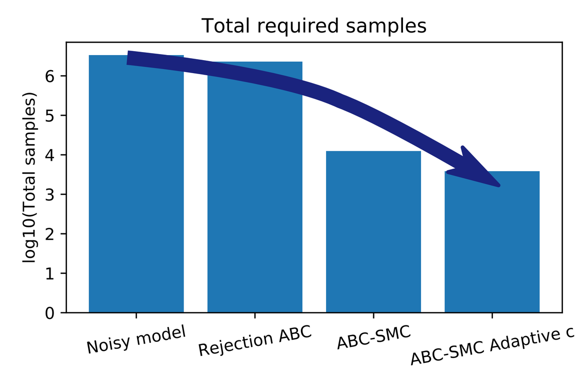


considerable speedup



Results

- applied to ODEs, SDEs, Markov jump processes
- with Gaussian, Laplace, Binomial noise
- A3 more efficient than A2, if applicable
- overall orders of magnitude speedups, enabling inference in some cases for the first time

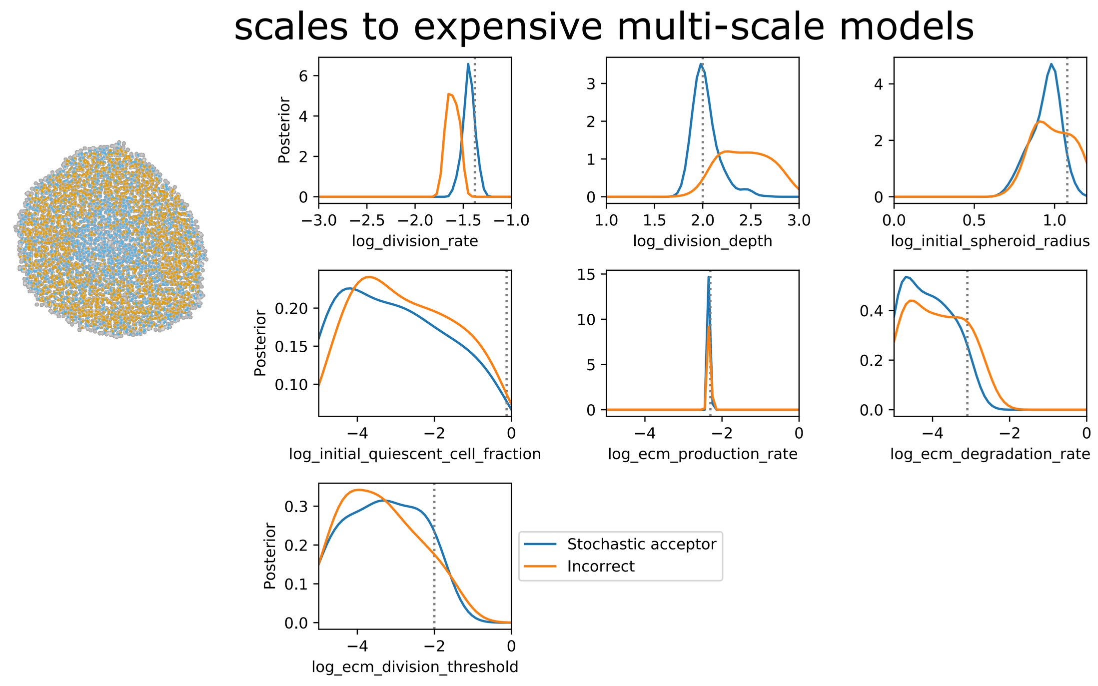


- can estimate noise parameters
- hardly any tuning necessary
- scalable, modular implementation in the toolbox pyabc

πABC

Application example: Multi-scale tumor model

scales to expensive multi-scale models



Future

- model selection via thermodynamic integration
- directly target distribution scale parameters
- optimize temperature scheme
- apply to real data

github.com/icb-dcm/pyabc
yannik.schaelte@helmholtz-muenchen.de

ICB Institute of Computational Biology

UNIVERSITÄT BONN