HelmholtzZentrum münchen
German Research Center for Environmental Health

github.com/yannikschaelte
yannik.schaelte@helmholtz-muenchen.de
@yannik_schaelte

Noise Up Your ABC

e:Med Bonn, 2020/11/24+25

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

the problem

$D =$

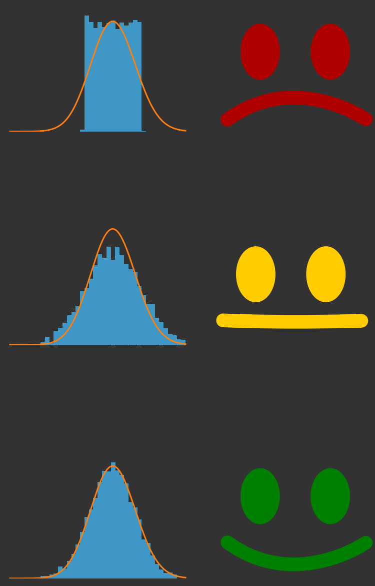
noisy data

the solutions

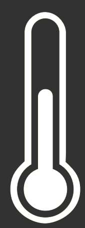
sample parameters $\theta \sim \pi(\theta)$ → simulate model $y \sim p(y|\theta)$ → simulate noise $\bar{y} \sim \pi(\bar{y}|y,\theta)$ → A2: Noisy model $d(D,\bar{y}) < \varepsilon$

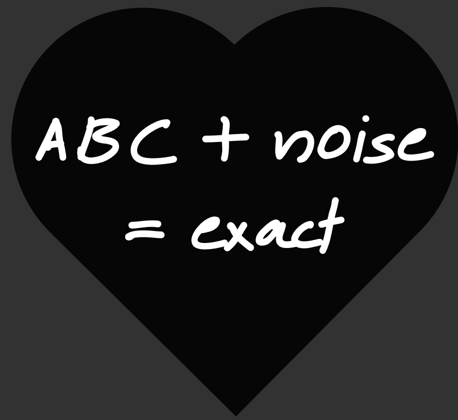
A1: Uniform kernel $d(D,y) < \varepsilon$

A3: Stochastic acceptor $\frac{\pi(D|y,\theta)}{c} > U[0,1]$



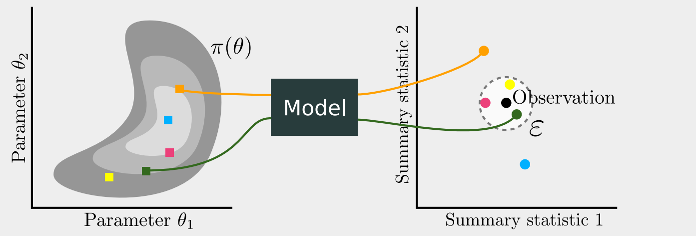
make it efficient





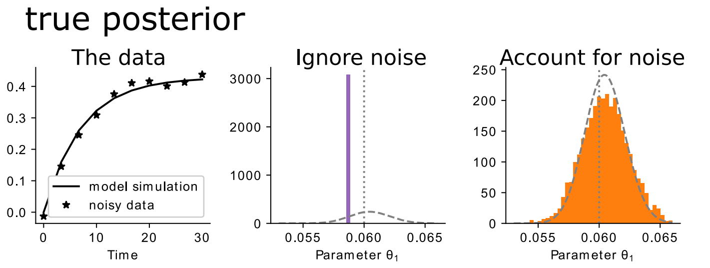
Mini-Intro: Rejection ABC

- Approximate Bayesian Computation enables **likelihood-free inference**
- only requires **forward simulations**



Measurement noise in ABC

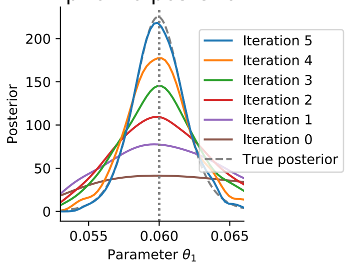
- experimental data usually noise corrupted
- easy to ignore → **wrong estimates**
- **modified acceptor** (A3) yields **exact likelihood-free inference** from the true posterior



Integrate in sequential ABC

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

smooth transition from prior to posterior



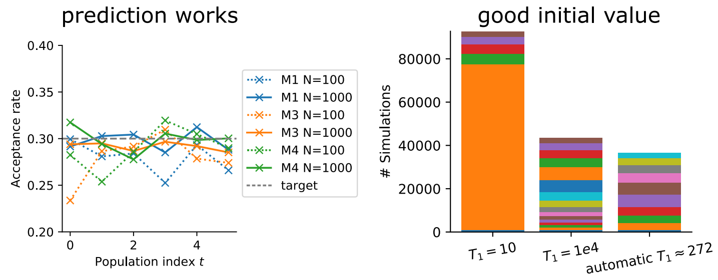
$$\pi_{\mathrm{ABC},t}(\theta|D) \propto \int \pi(D|y,\theta)^{1/T_t} p(y|\theta)dy \cdot \pi(\theta)$$

Choose temperatures

- problem: **sensitive** to initial temperature
- idea: **predict** the next population's **acceptance rate**
- choose T_t (esp. T_1) to match a **target rate**

prediction works

good initial value



Automatically find a good c

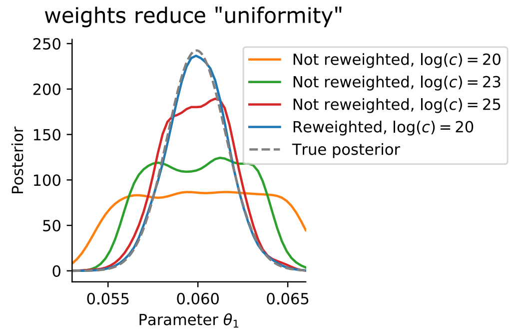
- problem: normalization constant c **affects acceptance rate**
- idea: **update** c after each iteration
- **avoid bias**: accept with probability

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

and **re-weight** by

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

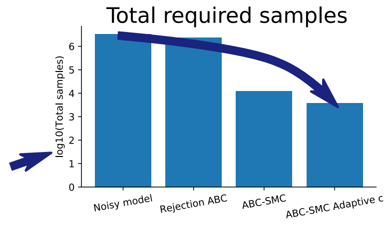
weights reduce "uniformity"



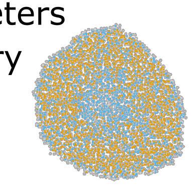
(More) Results

- applied to various model types (ODEs, SDEs, MJPs, ABMs)
- with Gaussian, Laplace, Binomial noise
- orders of magnitude speedups, enabling exact inference for the first time

Total required samples



- can estimate noise parameters
- hardly any tuning necessary
- scales to expensive multi-scale models



- ▸ scalable, modular implementation in the toolbox pyABC

pyABC

https://github.com/icb-dcm/pyabc

- ▸ manuscript:

https://doi.org/10.1093/bioinformatics/btaa397

- ▸ further info:

https://yannikschaelte.github.io/pres_ismb2020

Future

- model selection via thermodynamic integration
- directly target distribution scale parameters
- optimize temperature scheme
- applied in FitMultiCell projects

ICB Institute of Computational Biology
Federal Ministry of Education and Research
e:Med SYSTEMS MEDICINE
FitMultiCell
TUM
UNIVERSITÄT BONN