Multifidelity Approaches to Approximate Bayesian Computation

Thomas P. Prescott
Wolfson Centre for Mathematical Biology
University of Oxford

Banff International Research Station
11th – 16th November 2018


Buses on Diversion
15 17
521 15H
N15 N199

Approximate Bayesian computation

Mathematical Institute

Multifidelity methods

- Multifidelity modelling example: Repressilator

Multifidelity ABC

- Early accept/reject multifidelity ABC

Performance analysis

- Repressilator revisited

Implementation and future extensions

UNIVERSITY OF OXFORD
Mathematical Institute

- Real-world system defined by uncertain parameters $\theta \sim \pi(\cdot)$ belonging to a prior distribution.
  - Summary statistics, y_{obs}, are gathered from observed data.
- A model $p(\cdot \mid \theta)$ is a map from θ to a distribution.
  - Distribution $p(\cdot \mid \theta)$ on an output space containing y_{obs}.
  - Likelihood of observed data is $p(y_{\text{obs}} \mid \theta)$.
  - Simulation of a model is a draw $y \sim p(\cdot \mid \theta)$ from the distribution.
  - Cost of simulation is the time taken to simulate y.
- 'Model' here refers to combination of mathematical model and numerical implementation to generate y.

Bayesian parameter inference

UNIVERSITY OF OXFORD
Mathematical Institute

Posterior distribution for parameters is:

$$p(\theta \mid y_{\mathrm{obs}}) = \frac{p(y_{\mathrm{obs}} \mid \theta)\pi(\theta)}{\int p(y_{\mathrm{obs}} \mid \theta)\pi(\theta)\ \mathrm{d}\theta}.$$

Assume that the likelihood, $p(y_{\mathrm{obs}} \mid \theta)$, is unavailable.

UNIVERSITY OF OXFORD
Mathematical Institute

Attempt 1: Approximate likelihood, $p(y_{\mathrm{obs}} \mid \theta)$, using Monte Carlo simulation:

- Simulate $y \sim p(\cdot \mid \theta)$ and set $\hat{p}(y_{\mathrm{obs}} \mid \theta) = \mathbb{I}(y = y_{\mathrm{obs}})$.
- Expectation $\mathbb{E}(\hat{p}(y_{\mathrm{obs}}) \mid \theta) = p(y_{\mathrm{obs}} \mid \theta)$.

UNIVERSITY OF OXFORD
Mathematical Institute

Attempt 1: Approximate likelihood, $p(y_{\text{obs}} \mid \theta)$, using Monte Carlo simulation:

- Simulate $y \sim p(\cdot \mid \theta)$ and set $\hat{p}(y_{\text{obs}} \mid \theta) = \mathbb{I}(y = y_{\text{obs}})$.
- Expectation $\mathbb{E}(\hat{p}(y_{\text{obs}}) \mid \theta) = p(y_{\text{obs}} \mid \theta)$.
- Impractical in large output space.

UNIVERSITY OF OXFORD
Mathematical Institute

Attempt 2:

Consider neighbourhood $\Omega(\epsilon) = \{y \mid d(y, y_{\text{obs}}) < \epsilon\}$ around y_{obs}.

- Distance $d(y, y_{\text{obs}})$ in output space.
- Threshold distance ϵ.

Attempt 2:
Consider neighbourhood $\Omega(\epsilon) = \{y \mid d(y, y_{\text{obs}}) < \epsilon\}$ around y_{obs}.

- Distance $d(y, y_{\text{obs}})$ in output space.
- Threshold distance ϵ.

ABC approximation to the likelihood: for $y \sim p(\cdot \mid \theta)$,

$$p(y_{\text{obs}} \mid \theta) \approx p(y \in \Omega(\epsilon) \mid \theta) =: p_{\text{ABC}}(y_{\text{obs}} \mid \theta)$$

Requires $\lim_{\epsilon \to 0} p(y \in \Omega(\epsilon) \mid \theta) = p(y_{\text{obs}} \mid \theta)$.

ABC: Monte Carlo rejection sampling

UNIVERSITY OF OXFORD
Mathematical Institute

For $i = 1, \ldots, N$:

- Select $\theta_i \sim \pi(\cdot)$ from prior and simulate $y_i \sim p(\cdot \mid \theta_i)$.
- If $y_i \in \Omega(\epsilon)$ is in the ϵ-neighbourhood of y_{obs} then accept θ_i into sample, else reject.

Given θ_i, the accept/reject decision is a random weight

$$w(\theta_i) = \mathbb{I}(y_i \in \Omega(\epsilon))$$

with expectation $\mathbb{E}(w(\theta_i)) = p_{\text{ABC}}(y_{\text{obs}} \mid \theta_i) \approx p(y_{\text{obs}} \mid \theta_i)$.

Monte Carlo ABC: Computational bottleneck

UNIVERSITY OF OXFORD
Mathematical Institute

Simulate $y_i \sim p(\cdot \mid \theta_i)$ for $i = 1, \dots, N$ for N large.

- Aim to reduce computational burden.
- Efficiently exploring parameter space can reduce N:
  - Markov chain Monte Carlo (MCMC);
  - Seqential Monte Carlo (SMC).

Monte Carlo ABC: Computational bottleneck

Mathematical Institute

Simulate $y_i \sim p(\cdot \mid \theta_i)$ for $i = 1, \dots, N$ for N large.

- Aim to reduce computational burden.
- Efficiently exploring parameter space can reduce N:
  - Markov chain Monte Carlo (MCMC);
  - Seqential Monte Carlo (SMC).
- Goal: reduce simulation cost within rejection sampling framework.

Approximate Bayesian computation

Multifidelity methods

Multifidelity modelling example: Repressilator

Multifidelity ABC

Early accept/reject multifidelity ABC

Performance analysis

Repressilator revisited

Implementation and future extensions

Alternative models

UNIVERSITY OF OXFORD
Mathematical Institute

Original model $p(\cdot \mid \theta)$: distribution on output space containing y_{obs}.

- Alternative summary statistics $\tilde{y}_{\text{obs}}$ induce new output space.
- Consider new model $\tilde{p}(\cdot \mid \theta)$...
- ...with new cost, $\tilde{c}$, of simulating $\tilde{y} \sim \tilde{p}(\cdot \mid \theta)$.

We assume the new model is cheaper:

$$\mathbb{E}(\tilde{c}) \ll \mathbb{E}(c).$$

Cheaper simulations

Mathematical
Institute

- Early stopping: simulate to $t < T$;
- Coarser discretisations of space/time;
- Reduce model dimension;
- Langevin, mean field, or deterministic approximations;
- Surrogate models and regressions;
- Steady state analysis.

Note: we're assuming that θ well-defines both p and $\tilde{p}$.

UNIVERSITY OF OXFORD
Mathematical Institute

Using alternative model $\tilde{p}$:
Consider neighbourhood $\tilde{\Omega}(\tilde{\epsilon}) = \{\tilde{y} \mid \tilde{d}(\tilde{y}, \tilde{y}_{\text{obs}}) < \tilde{\epsilon}\}$ around $\tilde{y}_{\text{obs}}$.

- Distance $\tilde{d}(\tilde{y}, \tilde{y}_{\text{obs}})$ in output space.
- Threshold distance $\tilde{\epsilon}$.

University of Oxford
Mathematical Institute

Using alternative model $\tilde{p}$:
Consider neighbourhood $\tilde{\Omega}(\tilde{\epsilon}) = \{\tilde{y} \mid \tilde{d}(\tilde{y}, \tilde{y}_{\text{obs}}) < \tilde{\epsilon}\}$ around $\tilde{y}_{\text{obs}}$.

- Distance $\tilde{d}(\tilde{y}, \tilde{y}_{\text{obs}})$ in output space.
- Threshold distance $\tilde{\epsilon}$.

ABC approximation to a different likelihood: for $\tilde{y} \sim \tilde{p}(\cdot \mid \theta)$,

$$\tilde{p}(\tilde{y}_{\text{obs}} \mid \theta) \approx \tilde{p}_{\text{ABC}}(\tilde{y}_{\text{obs}} \mid \theta) = \tilde{p}(\tilde{y} \in \tilde{\Omega}(\tilde{\epsilon}) \mid \theta).$$

Requires $\lim_{\epsilon \to 0} \tilde{p}(\tilde{y} \in \tilde{\Omega}(\tilde{\epsilon}) \mid \theta) = \tilde{p}(\tilde{y}_{\text{obs}} \mid \theta)$.

ABC: Monte Carlo rejection sampling

University of Oxford
Mathematical Institute

For $i = 1, \dots, N$:

- Select $\theta_i \sim \pi(\cdot)$ from prior and simulate $\tilde{y}_i \sim \tilde{p}(\cdot \mid \theta_i)$.
- If $\tilde{y}_i \in \tilde{\Omega}(\tilde{\epsilon})$ is in the $\tilde{\epsilon}$-neighbourhood of $\tilde{y}_{\mathrm{obs}}$ then accept θ_i into sample, else reject.

Given θ_i, the accept/reject decision is a random weight

$$\tilde{w}(\theta_i) = \mathbb{I}(\tilde{y}_i \in \tilde{\Omega}(\tilde{\epsilon}))$$

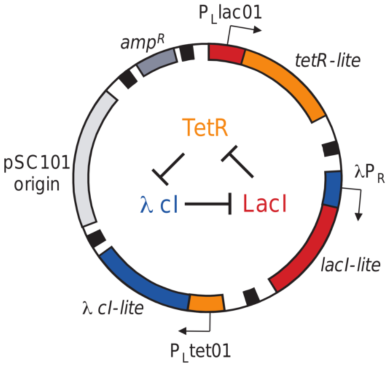
with expectation $\mathbb{E}(\tilde{w}(\theta_i)) = \tilde{p}_{\mathrm{ABC}}(\tilde{y}_{\mathrm{obs}} \mid \theta_i) \approx \tilde{p}(\tilde{y}_{\mathrm{obs}} \mid \theta_i)$.

Multifidelity terminology

UNIVERSITY OF OXFORD
Mathematical Institute

- High-fidelity model $p(\cdot \mid \theta)$:
  - HF simulation $y \sim p(\cdot \mid \theta)$ with cost c;
  - HF accept/reject weight $w(\theta) = \mathbb{I}(y \in \Omega(\epsilon))$.
- Low-fidelity model $\tilde{p}(\cdot \mid \theta)$:
  - LF simulation $\tilde{y} \sim \tilde{p}(\cdot \mid \theta)$ with cost $\tilde{c}$;
  - LF accept/reject weight $\tilde{w}(\theta) = \mathbb{I}(\tilde{y} \in \tilde{\Omega}(\tilde{\epsilon}))$.

Example: Repressilator



Elowitz & Leibler, *Nature*, 2000.

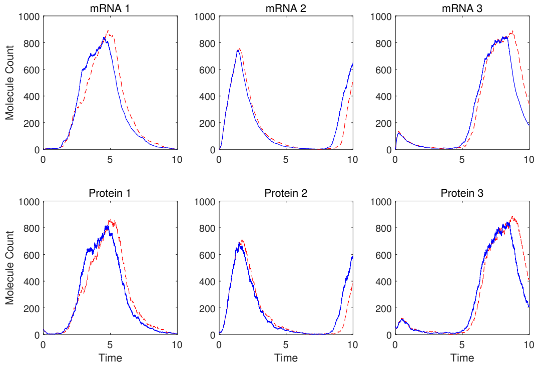
Three genes G_1, G_2, and G_3, transcribed and translated into proteins P_1, P_2, and P_3.

- P_1 represses G_2 transcription;
- P_2 represses G_3 transcription;
- P_3 represses G_1 transcription.

Goal: identify n and K_h in repression $f(p) = K_h^n/(K_h^n + p^n)$.

Multifidelity models

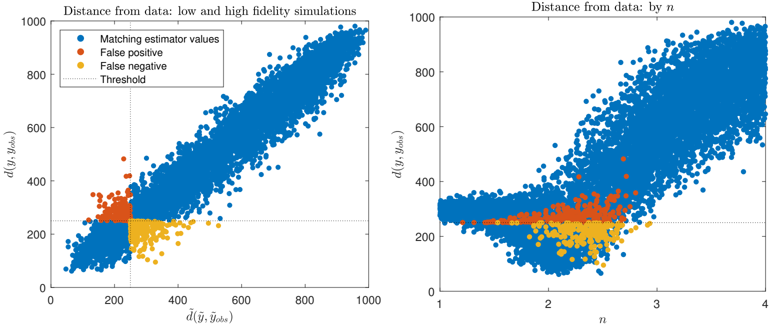
University of Oxford Mathematical Institute



Simulate network with:

- High fidelity — Stochastic Simulation Algorithm (Gillespie).
- Low fidelity — tau-leap discretisation.

Low-fidelity rejection sampling is biased



Distance $d(y, y_{\text{obs}})$ varies with $\tilde{d}(\tilde{y}, \tilde{y}_{\text{obs}})$ (left) and n (right). Each point ($N = 10^4$) corresponds to parameter sampled from uniform prior. Observed data $y_{\text{obs}} = \tilde{y}_{\text{obs}}$ is synthetic data using $n = 2$.

Low-fidelity rejection sampling is biased

High fidelity:

$$\mathbb{E}(w(\theta)) = p(y \in \Omega(\epsilon) \mid \theta) \approx p(y_{\mathrm{obs}} \mid \theta).$$

Low fidelity:

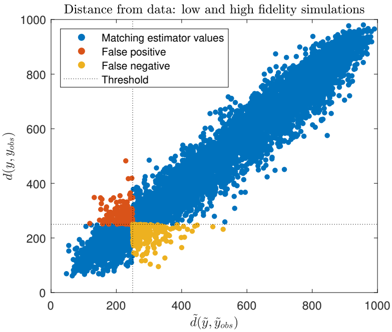
$$\begin{aligned}\mathbb{E}(\tilde{w}(\theta)) &= \tilde{p}(\tilde{y} \in \tilde{\Omega}(\tilde{\epsilon}) \mid \theta) \approx \tilde{p}(\tilde{y}_{\mathrm{obs}} \mid \theta) \\ &\neq p(y_{\mathrm{obs}} \mid \theta).\end{aligned}$$

Can we combine low and high-fidelity models to reduce simulation costs, but without introducing bias?

Early rejection

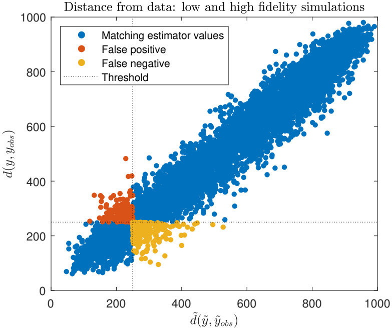
UNIVERSITY OF OXFORD
Mathematical Institute

First idea: Use low-fidelity simulation to decide when it's worth simulating the high fidelity model.



Early rejection

First idea: Use low-fidelity simulation to decide when it's worth simulating the high fidelity model.



- Simulate $\tilde{y}_i \sim \tilde{p}(\cdot \mid \theta_i)$.
- Define continuation probability $\alpha(\tilde{y}_i) \in (0, 1]$.
- w.p. $1 - \alpha$:
  - Reject without simulating y_i
- Else, simulate y_i:
  - If $y_i \in \Omega(\epsilon)$ then accept (with appropriate weight);
  - Else, reject.

UNIVERSITY OF OXFORD
Mathematical Institute

Formulate as early rejection weight:

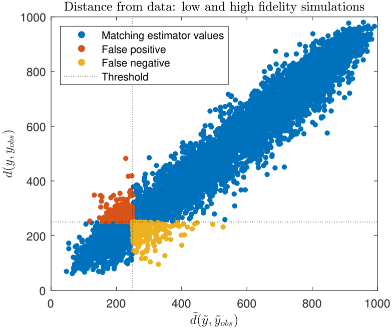
$$w_{\mathrm{er}}(\theta) = \frac{\mathbb{I}(U < \alpha(\tilde{y}))}{\alpha(\tilde{y})} w(\theta)$$

for unit uniform r.v. $U \sim \mathrm{U}(0, 1)$.

$w_{\mathrm{er}}(\theta)$ is an unbiased estimator of $p(y \in \Omega(\epsilon) \mid \theta)$, and we avoid some high-fidelity simulations.

Early rejection

UNIVERSITY OF OXFORD
Mathematical Institute



$$w_{\text{er}}(\theta) = \frac{\mathbb{I}(U < \alpha(\tilde{y}))}{\alpha(\tilde{y})} w(\theta)$$

Choosing $\alpha(\tilde{y})$:

- Make $\alpha(\tilde{y})$ small when $\tilde{y}$ suggests $y \notin \Omega(\epsilon)$ is likely.
- Converse requires larger $\alpha(\tilde{y})$ when $y \in \Omega(\epsilon)$ more likely.

See also: Lazy ABC (Prangle 2016), Delayed Acceptance ABC (Everitt & Rowinska 2017).

Re-examine converse

Mathematical Institute

$$w_{\text{er}}(\theta) = \frac{\mathbb{I}(U < \alpha(\tilde{y}))}{\alpha(\tilde{y})} w(\theta).$$

- Larger $\alpha(\tilde{y})$ when $y \in \Omega(\epsilon)$ more likely.

Re-examine converse

$$w_{\mathrm{er}}(\theta) = \frac{\mathbb{I}(U < \alpha(\tilde{y}))}{\alpha(\tilde{y})} w(\theta).$$

- Larger $\alpha(\tilde{y})$ when $y \in \Omega(\epsilon)$ more likely.
- More likely to simulate high-fidelity model if low-fidelity simulation is close to the data.
- Why not accept early without simulating?

Re-examine converse

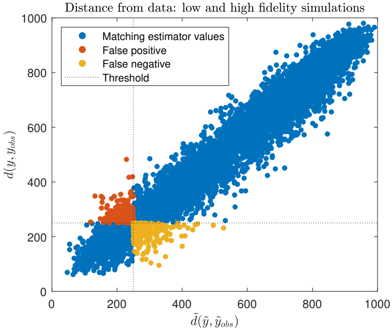
$$w_{\text{er}}(\theta) = \frac{\mathbb{I}(U < \alpha(\tilde{y}))}{\alpha(\tilde{y})} w(\theta).$$

- Larger $\alpha(\tilde{y})$ when $y \in \Omega(\epsilon)$ more likely.
- More likely to simulate high-fidelity model if low-fidelity simulation is close to the data.
- Why not accept early without simulating?
- Early acceptance requires positive weight without high-fidelity simulation: not possible using w_{er}.

Early decision

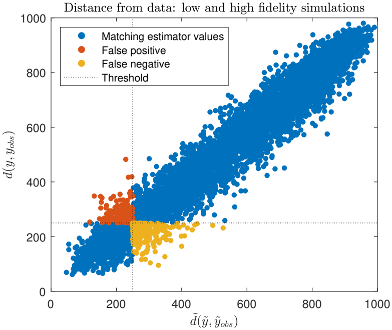
UNIVERSITY OF OXFORD
Mathematical Institute

Second idea: Sometimes make an early accept/reject decision based on the low-fidelity simulation alone.



Mathematical Institute

Second idea: Sometimes make an early accept/reject decision based on the low-fidelity simulation alone.



- Simulate $\tilde{y}_i \sim \tilde{p}(\cdot \mid \theta_i)$.
- Make an early decision $\tilde{w}$.
- Define probability $\eta \in (0, 1]$.
- w.p. $1 - \eta$:
  - Use early decision.
- Else, simulate y_i:
  - Correct early decision, based on w.
  - Correction appropriately weighted.

Early decision

UNIVERSITY OF OXFORD
Mathematical Institute

Formulate as early decision weight:

$$w_{\text{ed}}(\theta) = \tilde{w}(\theta) + \frac{\mathbb{I}(U < \eta)}{\eta}\left[w(\theta) - \tilde{w}(\theta)\right]$$

for unit uniform r.v. $U \sim \mathrm{U}(0,1)$.

$w_{\text{ed}}(\theta)$ is an unbiased estimator of $p(y \in \Omega(\epsilon) \mid \theta)$, and we avoid some high-fidelity simulations.

Combining existing approaches

UNIVERSITY OF OXFORD
Mathematical Institute

Early rejection — use $\tilde{y}$ to determine when to generate y:

$$w_{\mathrm{er}} = \frac{\mathbb{I}(U < \alpha(\tilde{y}))}{\alpha(\tilde{y})} \mathbb{I}(y \in \Omega(\epsilon)).$$

Early decision — use $\tilde{y}$ to determine the early decision:

$$w_{\mathrm{ed}} = \mathbb{I}(\tilde{y} \in \tilde{\Omega}(\tilde{\epsilon})) + \frac{\mathbb{I}(U < \eta)}{\eta} \left(\mathbb{I}(y \in \Omega(\epsilon)) - \mathbb{I}(\tilde{y} \in \tilde{\Omega}(\tilde{\epsilon})) \right).$$

Special cases of multifidelity ABC.

Early accept/reject multifidelity ABC

UNIVERSITY OF OXFORD
Mathematical Institute

Combine approaches into a multifidelity weight:

$$w_{\mathrm{mf}}(\theta) = \mathbb{I}(\tilde{y} \in \tilde{\Omega}(\tilde{\epsilon})) + \frac{\mathbb{I}(U < \eta(\tilde{y}))}{\eta(\tilde{y})} \Big(\mathbb{I}(y \in \Omega(\epsilon)) - \mathbb{I}(\tilde{y} \in \tilde{\Omega}(\tilde{\epsilon})) \Big)$$

Early accept/reject multifidelity ABC

Mathematical Institute

Combine approaches into a multifidelity weight:

$$w_{\mathrm{mf}}(\theta) = \mathbb{I}(\tilde{y} \in \tilde{\Omega}(\tilde{\epsilon})) + \frac{\mathbb{I}(U < \eta(\tilde{y}))}{\eta(\tilde{y})} \left(\mathbb{I}(y \in \Omega(\epsilon)) - \mathbb{I}(\tilde{y} \in \tilde{\Omega}(\tilde{\epsilon})) \right)$$

Natural form of $\eta(\tilde{y})$ to give early accept/reject multifidelity ABC:

$$\eta(\tilde{y}) = \eta_1 \mathbb{I}(\tilde{y} \in \tilde{\Omega}(\tilde{\epsilon})) + \eta_2 \mathbb{I}(\tilde{y} \notin \tilde{\Omega}(\tilde{\epsilon}))$$

for $\eta_1, \eta_2 \in (0, 1]$.

- $\mathbb{E}(w_{\mathrm{mf}}(\theta)) = p(y \in \Omega(\epsilon) \mid \theta)$.
- $1 - \eta_1$ and $1 - \eta_2$ are early accept/reject probabilities.

Early accept/reject multifidelity ABC

University of Oxford
Mathematical Institute

There are 6 cases (at most 4 values) of $w_{\mathrm{mf}}(\theta)$:

$$w_{\mathrm{mf}}(\theta) = \begin{cases} 1 & \tilde{y} \in \tilde{\Omega}(\tilde{\epsilon}) \ \& \ U \geq \eta_1 \\ 0 & \tilde{y} \notin \tilde{\Omega}(\tilde{\epsilon}) \ \& \ U \geq \eta_2 \\ 1 & \tilde{y} \in \tilde{\Omega}(\tilde{\epsilon}) \ \& \ y \in \Omega(\epsilon) \ \& \ U < \eta_1 \\ 0 & \tilde{y} \notin \tilde{\Omega}(\tilde{\epsilon}) \ \& \ y \notin \Omega(\epsilon) \ \& \ U < \eta_2 \\ 1 - 1/\eta_1 & \tilde{y} \in \tilde{\Omega}(\tilde{\epsilon}) \ \& \ y \notin \Omega(\epsilon) \ \& \ U < \eta_1 \\ 0 + 1/\eta_2 & \tilde{y} \notin \tilde{\Omega}(\tilde{\epsilon}) \ \& \ y \in \Omega(\epsilon) \ \& \ U < \eta_2 \end{cases}$$

Early accept/reject multifidelity ABC

University of Oxford
Mathematical Institute

There are 6 cases (at most 4 values) of $w_{\mathrm{mf}}(\theta)$:

$$w_{\mathrm{mf}}(\theta) = \begin{cases} 1 & \text{Predicted positive (unchecked)} \\ 0 & \text{Predicted negative (unchecked)} \\ 1 & \text{True positive (checked)} \\ 0 & \text{True negative (checked)} \\ 1 - 1/\eta_1 & \text{False positive (checked)} \\ 0 + 1/\eta_2 & \text{False negative (checked)} \end{cases}$$

UNIVERSITY OF OXFORD
Mathematical Institute

Approximate Bayesian computation

Multifidelity methods
- Multifidelity modelling example: Repressilator

Multifidelity ABC
- Early accept/reject multifidelity ABC

Performance analysis
- Repressilator revisited

Implementation and future extensions

Early accept/reject multifidelity ABC algorithm

University of Oxford
Mathematical Institute

For $i = 1, \dots, N$:

- Generate $\theta_i \sim \pi(\cdot)$.
- Simulate (low-fidelity) $\tilde{y}_i \sim \tilde{p}(\cdot \mid \theta_i)$.
- Calculate $\tilde{w} = \mathbb{I}(\tilde{y}_i \in \tilde{\Omega}(\tilde{\epsilon}))$.
- Set $w_i = \tilde{w}$ and $\eta = \eta_1 \tilde{w} + \eta_2 (1 - \tilde{w})$.
- If $U < \eta$:
  - Simulate (high fidelity) $y_i \sim p(\cdot \mid \theta)$.
  - Calculate $w = \mathbb{I}(y_i \in \Omega(\epsilon))$.
  - Update $w_i = w_i + (w - \tilde{w})/\eta$.

Return $\{w_i, \theta_i\}_{i=1}^N$.

Weighted sample

UNIVERSITY OF OXFORD
Mathematical Institute

Input to algorithm is continuation probabilities η_1, η_2.

Output of algorithm is weighted sample $\{w_i, \theta_i\}_{i=1}^N$.

Each weight w_i incurs simulation cost T_i.

Use sample in ABC estimator for a function, $F(\theta)$, of uncertain parameters:

$$\mu_{\mathrm{ABC}}(F) = \frac{\sum_i w_i F(\theta_i)}{\sum_j w_j} \approx \int F(\theta) p_{\mathrm{ABC}}(y_{\mathrm{obs}} \mid \theta) \pi(\theta) \, \mathrm{d}\theta.$$

Quality of a sample

Monte Carlo sample $\{w_i, \theta_i\}_{i=1}^N$:

- Effective Sample Size:

$$\mathrm{ESS} = \frac{(\sum_i w_i)^2}{\sum_i w_i^2} = N \frac{(\sum_i w_i/N)^2}{\sum_i w_i^2/N}$$

- Simulation time:

$$T_{\mathrm{tot}} = \sum T_i$$

Quality of a sample

UNIVERSITY OF OXFORD
Mathematical Institute

Efficiency is ratio of ESS and simulation time:

$$\frac{\mathrm{ESS}}{T_{\mathrm{tot}}} = \frac{\left(\sum_i w_i/N\right)^2}{\left(\sum_i w_i^2/N\right)\left(\sum_i T_i/N\right)} \approx \frac{\mathbb{E}(w_i)^2}{\mathbb{E}(w_i^2)\mathbb{E}(T_i)}$$

Approximation is in the limit as $N \to \infty$.

Goal: Tune inputs η_1, η_2 to maximise efficiency (in the limit).

Maximising efficiency

UNIVERSITY OF OXFORD
Mathematical Institute

$$\frac{\mathrm{ESS}}{T_{\mathrm{tot}}} \approx \frac{\mathbb{E}(w_{\mathrm{mf}})^2}{\mathbb{E}(w_{\mathrm{mf}}^2)\mathbb{E}(T)} \quad \text{over } \theta \sim \pi(\cdot)$$

- Numerator: $\mathbb{E}(w_{\mathrm{mf}}) = p(y \in \Omega(\epsilon))$ is independent of η_1, η_2.

Maximising efficiency

UNIVERSITY OF OXFORD
Mathematical Institute

$$\frac{\text{ESS}}{T_{\text{tot}}} \approx \frac{\mathbb{E}(w_{\text{mf}})^2}{\mathbb{E}(w_{\text{mf}}^2)\mathbb{E}(T)} \quad \text{over } \theta \sim \pi(\cdot)$$

- Numerator: $\mathbb{E}(w_{\text{mf}}) = p(y \in \Omega(\epsilon))$ is independent of η_1, η_2.
- Maximising efficiency over $\eta_1, \eta_2 \in (0, 1]$ equivalent to minimising product $\mathbb{E}(w_{\text{mf}}^2)\mathbb{E}(T_i)$.

UNIVERSITY OF OXFORD
Mathematical Institute

Simulation times:

- Low-fidelity simulation $\tilde{y} \sim \tilde{p}(\cdot \mid \theta)$ costs $\tilde{c}$.
- High-fidelity simulation $y \sim p(\cdot \mid \theta)$ costs c.
- w_{mf} costs $T = \tilde{c} + c\mathbb{I}(U < \eta(\tilde{y}))$.

Expected time to construct $w_{\mathrm{mf}}(\theta)$ is therefore

$$\mathbb{E}(T) = \mathbb{E}(\tilde{c}) + \eta_1 \mathbb{E}[c\mathbb{I}(\tilde{y} \in \tilde{\Omega}(\tilde{\epsilon}))] + \eta_2 \mathbb{E}[c\mathbb{I}(\tilde{y} \notin \tilde{\Omega}(\tilde{\epsilon}))].$$

Depends on simulation cost c relative to $\tilde{c}$.

UNIVERSITY OF OXFORD
Mathematical Institute

Second moment increases as η_i decreases:

$$\mathbb{E}(w_{\mathrm{mf}}^2) = \mathbb{P}(y \in \Omega(\epsilon)) + \left(\frac{1}{\eta_1} - 1\right)\mathbb{P}(\tilde{y} \in \tilde{\Omega}(\tilde{\epsilon}) \cap y \notin \Omega(\epsilon)) + \left(\frac{1}{\eta_2} - 1\right)\mathbb{P}(\tilde{y} \notin \tilde{\Omega}(\tilde{\epsilon}) \cap y \in \Omega(\epsilon))$$

Depends on false positive and false negative rates.

Optimal early accept/reject continuation probabilities

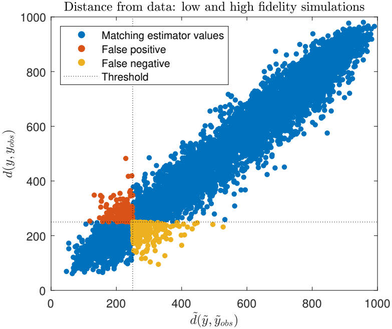
Analytical result: efficiency maximised at

$$(\eta_1^\star, \eta_2^\star) = \frac{1}{\lambda}\left(\sqrt{\frac{\mathbb{P}(y \notin \Omega(\epsilon) \mid \tilde{y} \in \tilde{\Omega}(\tilde{\epsilon}))}{\mathbb{E}(c \mid \tilde{y} \in \tilde{\Omega}(\epsilon))/\mathbb{E}(\tilde{c})}}, \sqrt{\frac{\mathbb{P}(y \in \Omega(\epsilon) \mid \tilde{y} \notin \tilde{\Omega}(\tilde{\epsilon}))}{\mathbb{E}(c \mid \tilde{y} \notin \tilde{\Omega}(\tilde{\epsilon}))/\mathbb{E}(\tilde{c})}} \right)$$

- If probability $\mathbb{P}(y \notin \Omega(\epsilon) \mid \tilde{y} \in \tilde{\Omega}(\tilde{\epsilon}))$ of needing to correct a positive prediction is small, then check less often.
- If the cost $\mathbb{E}(c \mid \tilde{y} \in \tilde{\Omega}(\epsilon))$ of checking a positive prediction is large relative to low-fidelity simulation cost $\mathbb{E}(\tilde{c})$, then check less often.
- Similar for negative predictions.
- Less effective when $\lambda = \sqrt{\mathbb{P}(y \in \Omega(\epsilon)) - \mathbb{P}(w \neq \tilde{w})}$ is small.

Example: Construct realisations of early accept/reject samples

UNIVERSITY OF OXFORD
Mathematical Institute

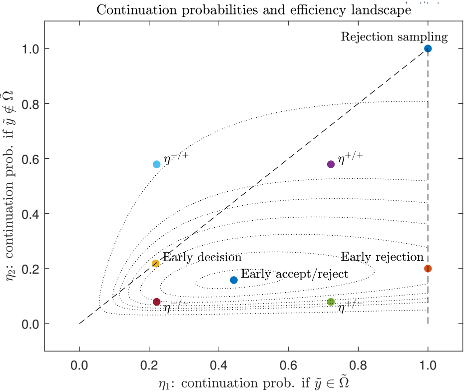


- Calculate $w_{\text{mf}}(\theta_i)$ for $i = 1, \dots, 10^4$.
- $\mathbb{I}(U < \eta(\tilde{y}_i))$ is the random decision.
- Evaluate $\text{ESS}/T_{\text{tot}}$ for 500 realisations.

Optimal $(\eta_1^\star, \eta_2^\star)$ taken from baseline dataset: compare to other continuation probabilities.

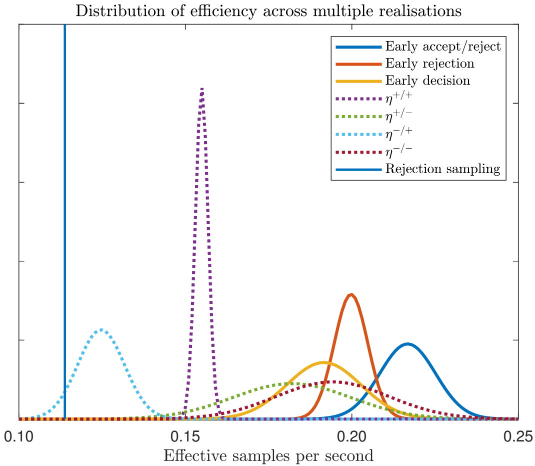
Example: Efficiency of different continuation probabilities

- Early decision:
  $\eta_1 = \eta_2$.
- Early rejection:
  $\eta_1 = 1$.
- Rejection sampling:
  $\eta_1 = \eta_2 = 1$.



Level sets are 99%, 95%, 90%, 85%, 80%, 75%, 60% of maximum efficiency.

Example: Distribution of observed efficiencies



Mathematical Institute

Approximate Bayesian computation

Multifidelity methods

Multifidelity modelling example: Repressilator

Multifidelity ABC

Early accept/reject multifidelity ABC

Performance analysis

Repressilator revisited

Implementation and future extensions

UNIVERSITY OF OXFORD
Mathematical Institute

Dealt with in recent work:

- No *a priori* knowledge of location of $(\eta_1^\star, \eta_2^\star)$:
  - Simulation times and false positive/negative probabilities.
  - Use burn-in approach first, then estimate.
- Methods to reduce false positive/negative probabilities.
  - Coupling low and high fidelity simulations.
- Optimising efficiency in terms of estimating specific functions

$$\mathbb{E}_{\mathrm{ABC}}(F(\theta)) \approx \sum_i w_i F(\theta_i) / \sum_i w_i$$

  - Minimize $\mathbb{E}(w_{\mathrm{mf}}^2 (F - \mathbb{E}_{\mathrm{ABC}}(F))^2) \mathbb{E}(T)$ instead.

Outlook: future work

UNIVERSITY OF OXFORD
Mathematical Institute

- Applicability to MCMC/SMC methods.
  - Deal with $w_{\mathrm{mf}} < 0$.
- Multiple low-fidelity models.
  - Local hierarchies of accuracy/speed-up.
  - Parameter/estimator-specific low-fidelity model.
- Alternative forms for $\eta(\tilde{y})$ and $\tilde{w}$.
  - Parameter-dependent η.
- Analytic/computational bounds on error:
  - $\tilde{y}$ versus y;
  - $\tilde{w}$ versus w.

Acknowledgements

UNIVERSITY OF OXFORD
Mathematical Institute

BBSRC
bioscience for the future

UK Research and Innovation

Project: *Next generation approaches to connect models and quantitative data.*
PIs: Ruth Baker & Michael Stumpf.