HELMHOLTZ MUNICH
TUM
UNIVERSITÄT BONN

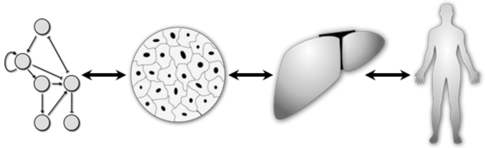
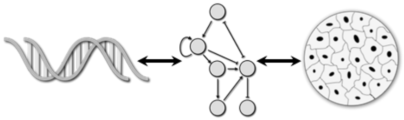
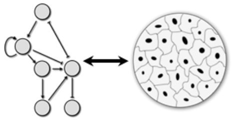
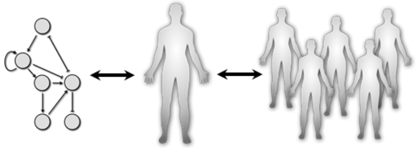
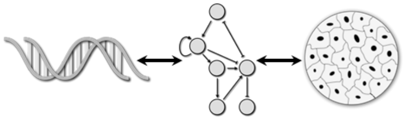
pyABC: Efficient and robust easy-to-use approximate Bayesian computation

Yannik Schälte
Inverse Problems Summer School, HCM Bonn 2022-08-26

@yannik_schaelte
yannikschaelte

WHY?

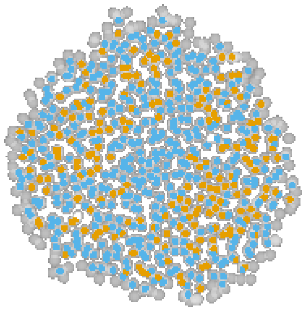
MODEL TYPES

		Boolean / Petri net models	Constraint-based models	Markov chain models	SDE models	ODE models	PDE models	Agent-based models	
whole-heart (tissue and fluid mechanics, signaling)		■				■	■		Hunter and Borg, *Nature Reviews Molecular Cell Biology*, 4(3):237-243, 2003
cancer growth (signaling, cell division and death, angiogenesis, tissue remodeling)					■		■	■	Anderson and Quaranta, *Nature Reviews Cancer*, 8(3):227-234. 2008
liver lobule (cell division and cell death, tissue mechanics)				■	■		■	■	Hoehme et al., *PNAS*, 107(23):10371-10376, 2010
glucose-insulin-glucagon regulation (blood and interstitial flow, organ uptake, signaling)						■	■	■	Schaller et al., *CPT: Pharmacometrics and Systems Pharmacology*, 2:e65, 2013
whole-cell (transcription, translation, DNA replication, metabolism, replication)		■	■	■		■		■	Karr et al., *Cell*, 150(2):389-401, 2012

based on Hasenauer et al., J. Coup. Sys. and Mult. Dyn., 2015

EXAMPLE: TUMOR GROWTH MULTI-SCALE MODEL

based on Jagiella et al., Cell Systems 2017



- cells modeled as interacting stochastic agents, dynamics of extracellular substances by PDEs
- simulate up to 10^6 cells
- 10s - 1h for one forward simulation

WHAT WE TRIED

- multi-start local methods
  - deterministic gradient descent
    - Levenberg-Marquardt
    - interior-point
    - trust-region
  - stochastic gradient descent
  - Bayesian optimization
- global methods
  - simulated annealing
  - > 20 particle methods
  - enhanced scatter search

FAILED	WORKED

- multi-start local methods
  - deterministic gradient descent
    - Levenberg-Marquardt
    - interior-point
    - trust-region
  - stochastic gradient descent
  - Bayesian optimization
- global methods
  - simulated annealing
  - > 20 particle methods
  - enhanced scatter search

WHAT?

ABC

likelihood-free approximate Bayesian computation

posterior

$$\pi(\theta|\bar{y}_{\text{obs}}) \propto \cancel{\pi(\bar{y}_{\text{obs}}|\theta)}\pi(\theta)$$

likelihood

prior

$$\bar{y} \sim \pi(\bar{y}|\theta)$$ ✓

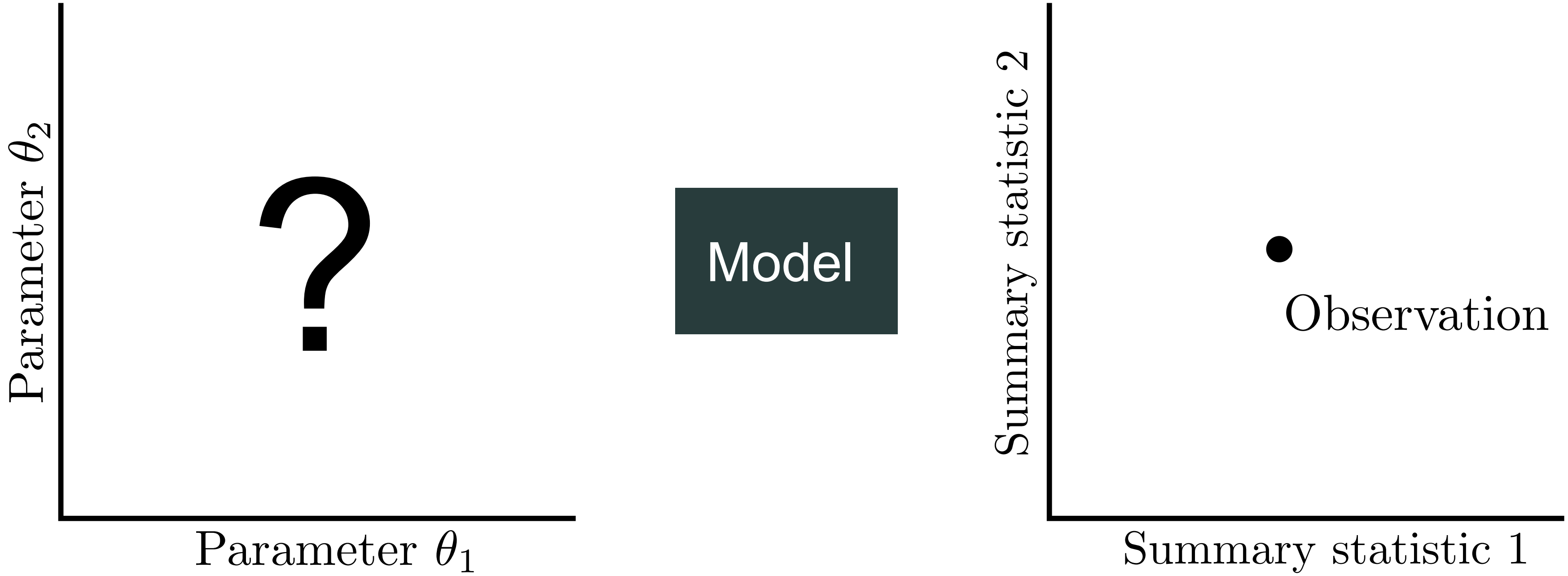
ABC

likelihood-free approximate Bayesian computation

Model

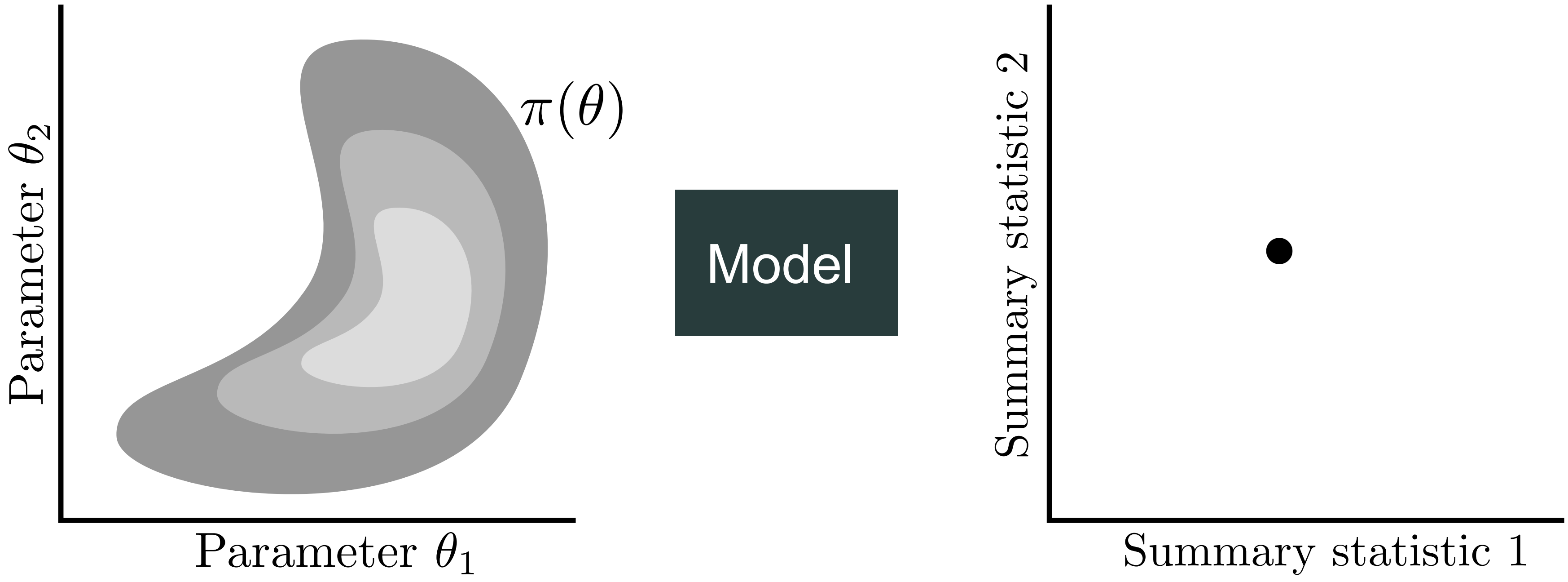
ABC

likelihood-free approximate Bayesian computation



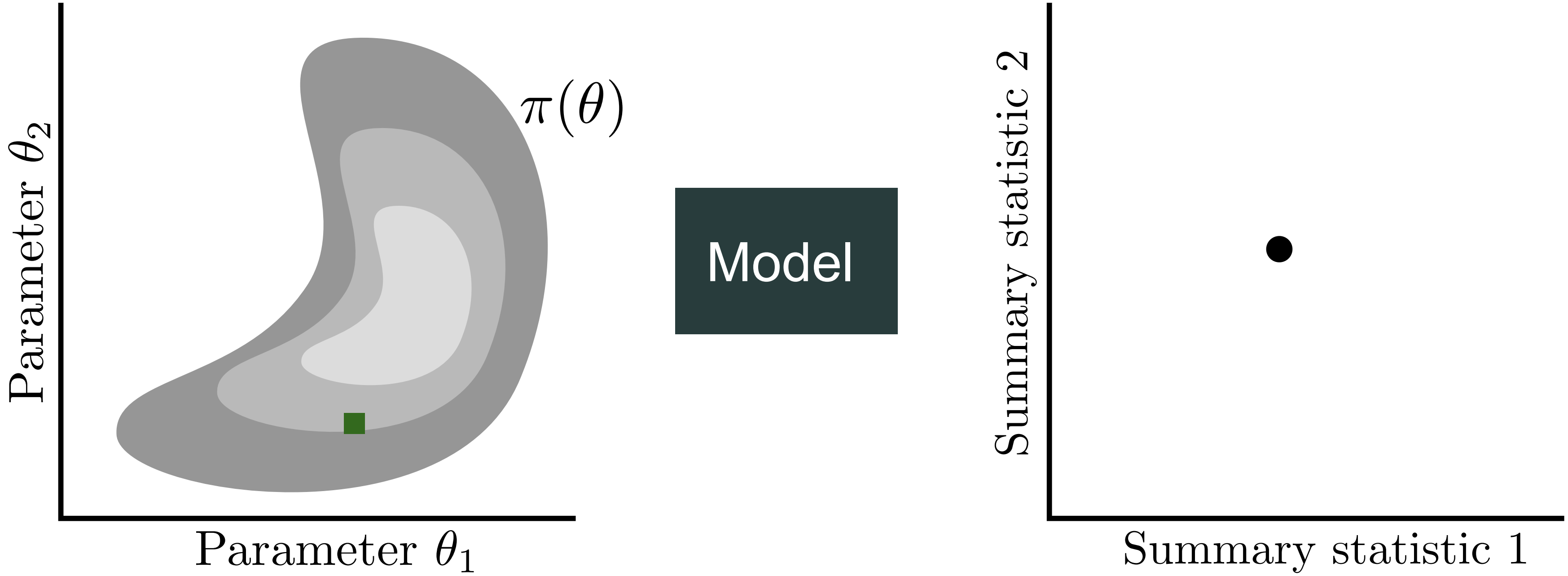
ABC

likelihood-free approximate Bayesian computation



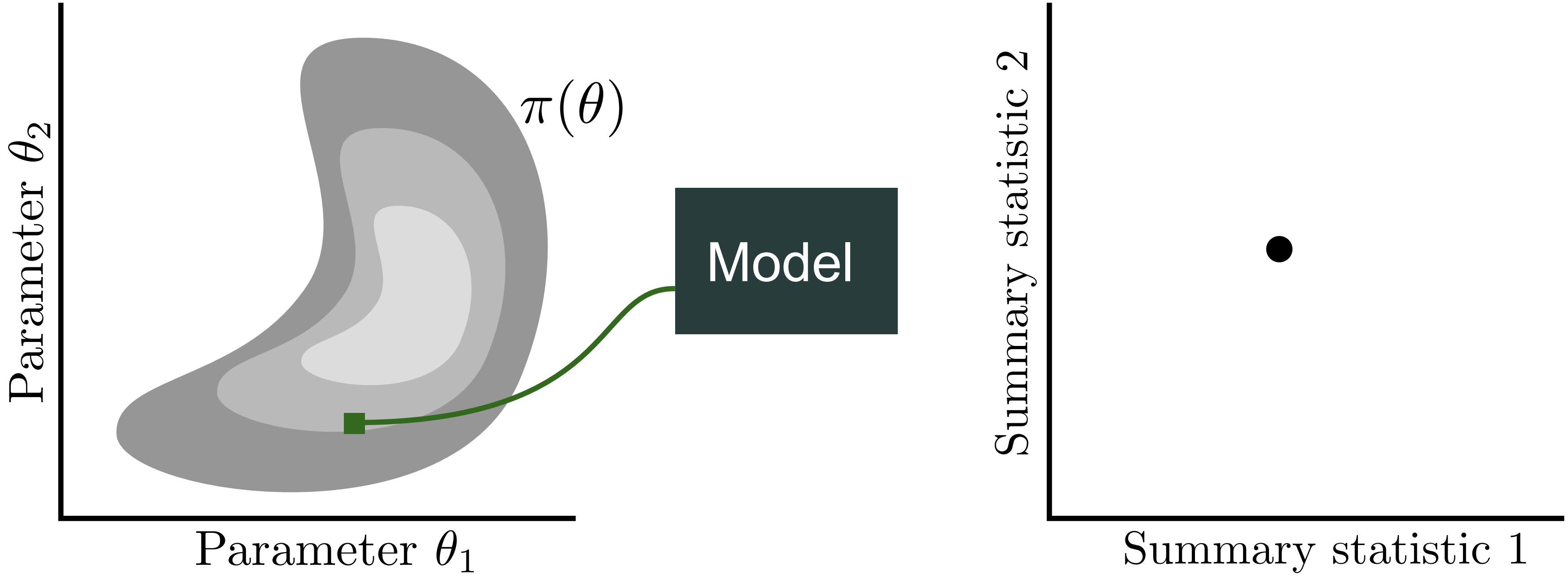
ABC

likelihood-free approximate Bayesian computation



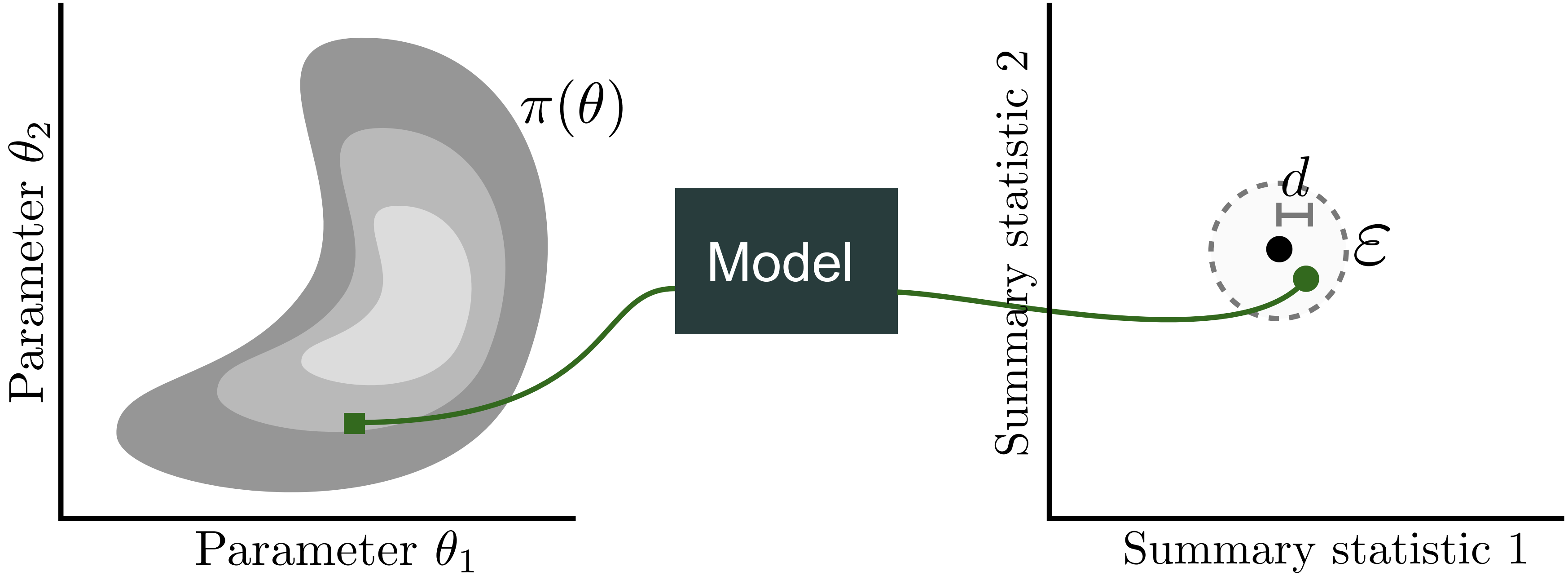
ABC

likelihood-free approximate Bayesian computation



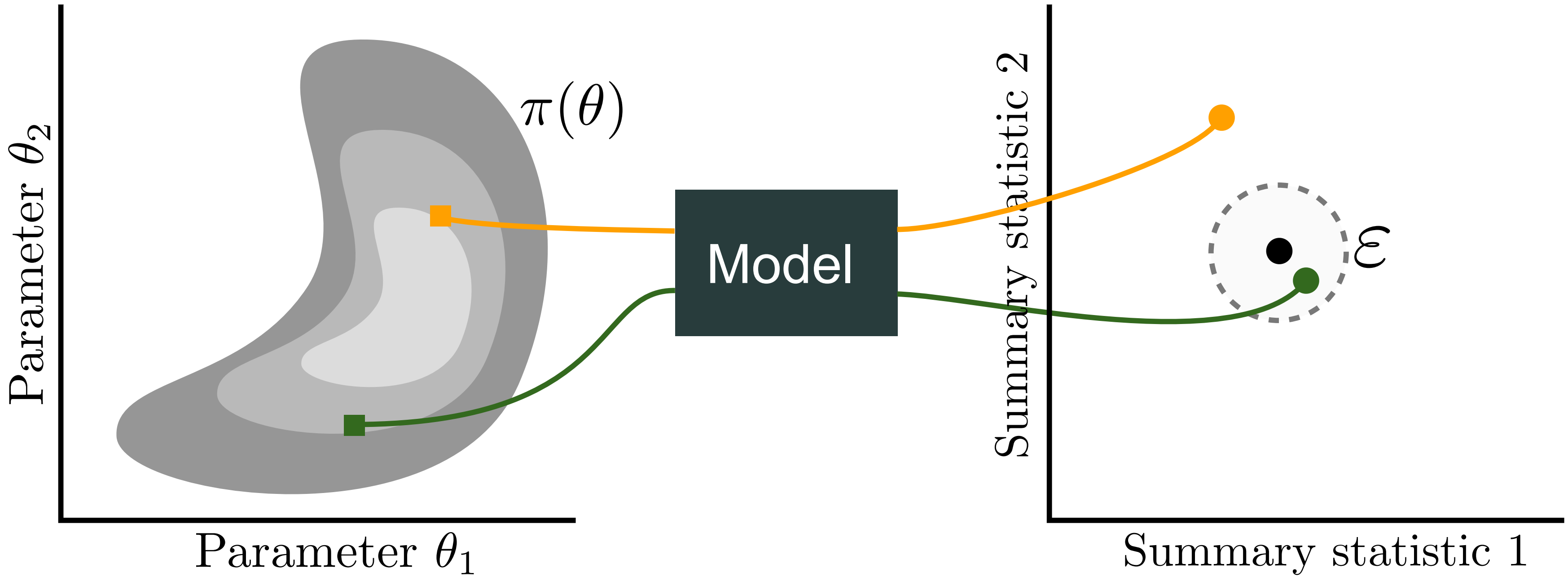
ABC

likelihood-free approximate Bayesian computation



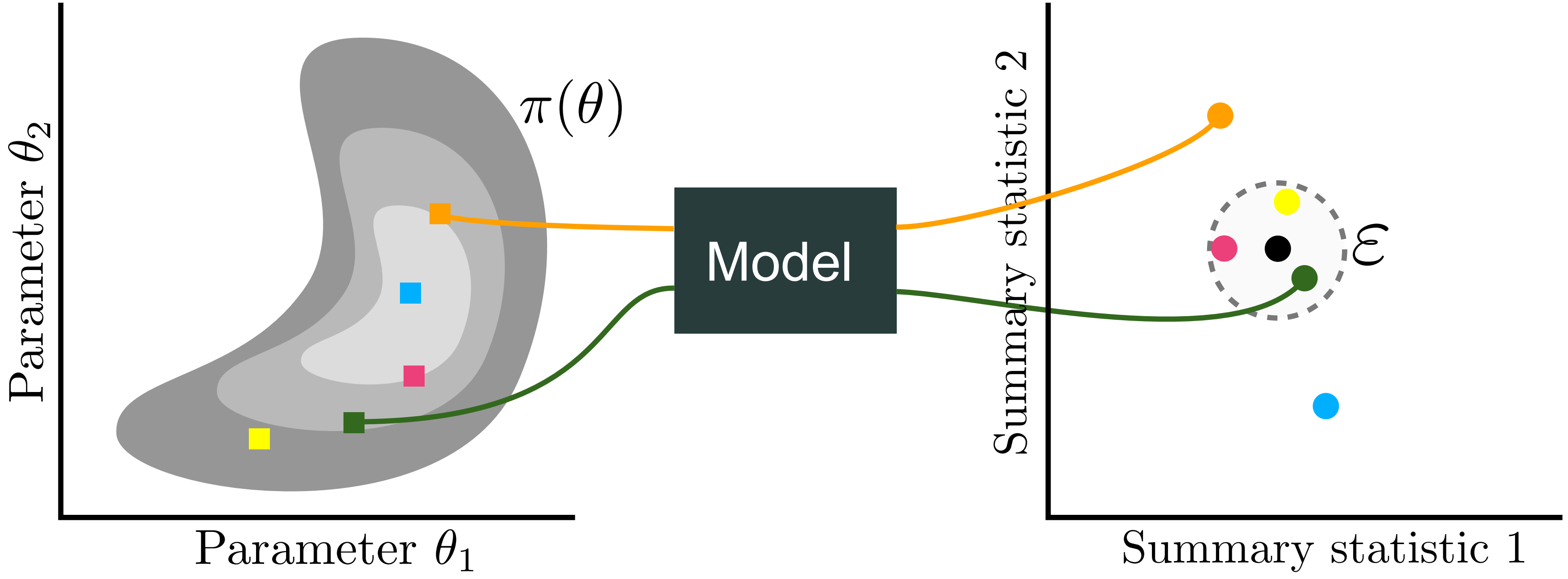
ABC

likelihood-free approximate Bayesian computation



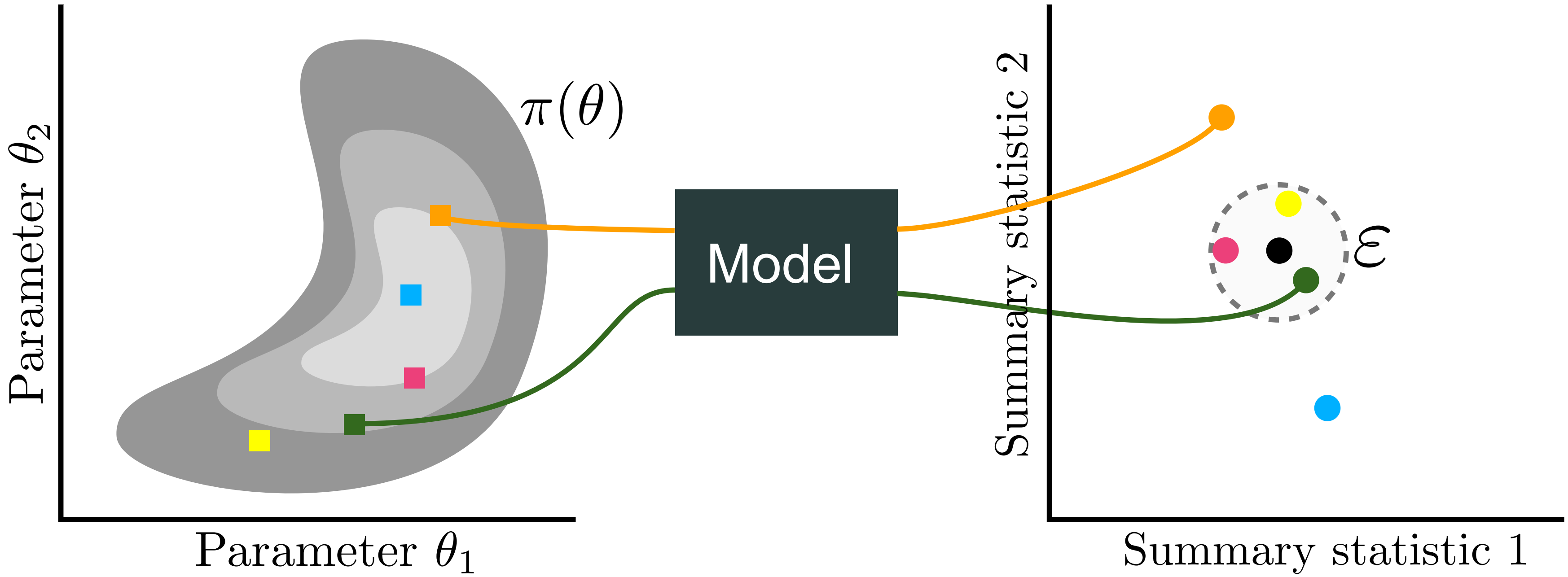
ABC

likelihood-free approximate Bayesian computation



ABC

likelihood-free approximate Bayesian computation

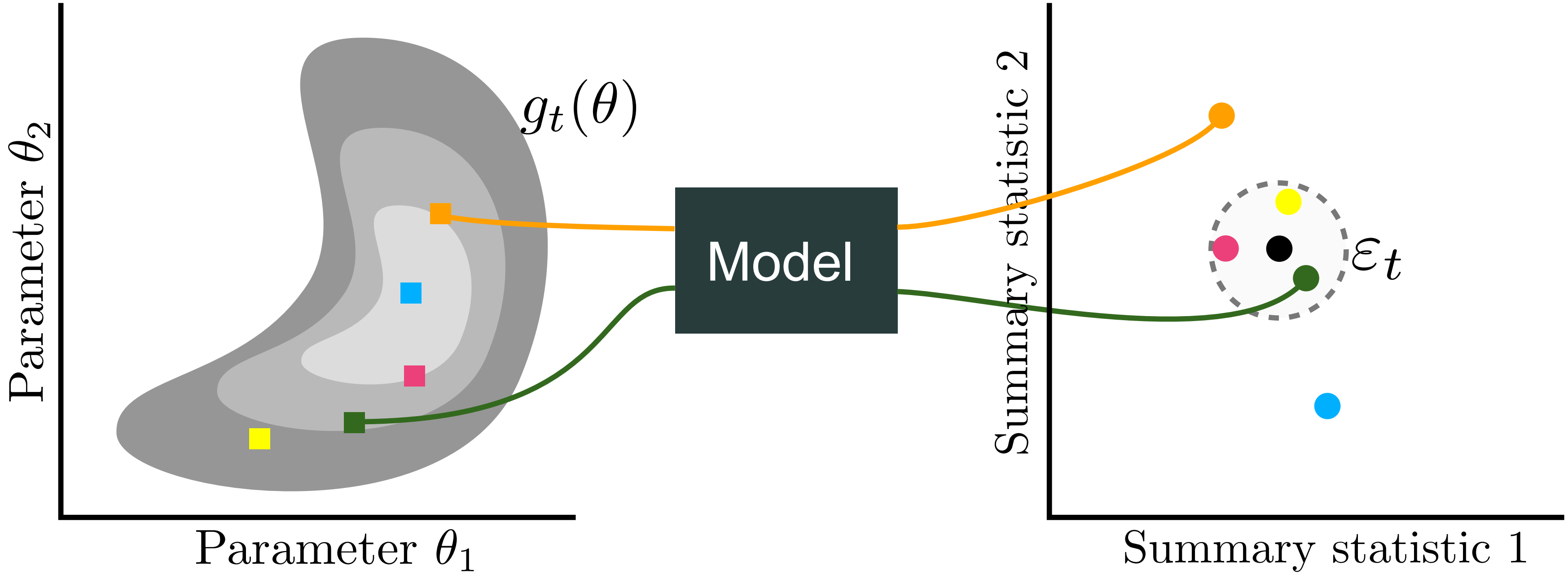


conflicting goals:

- reduce approximation error ε
- keep high acceptance rates

ABC-SMC

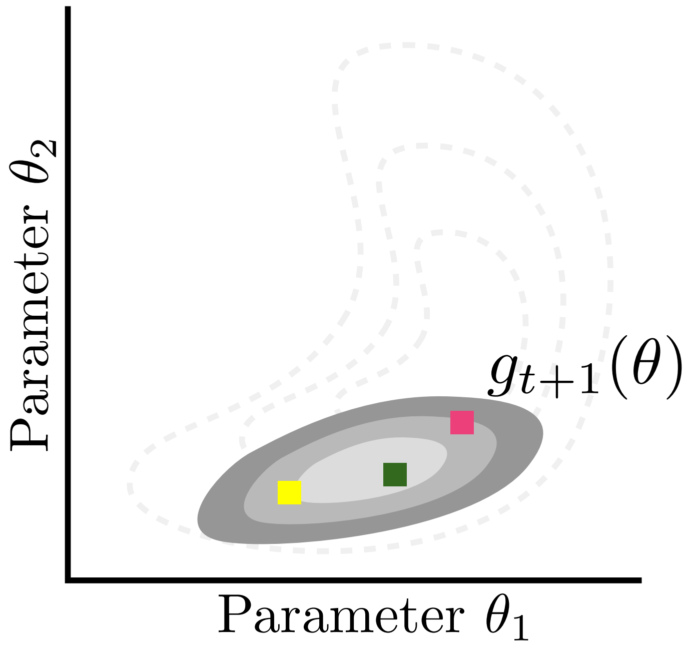
combine with a sequential Monte-Carlo scheme



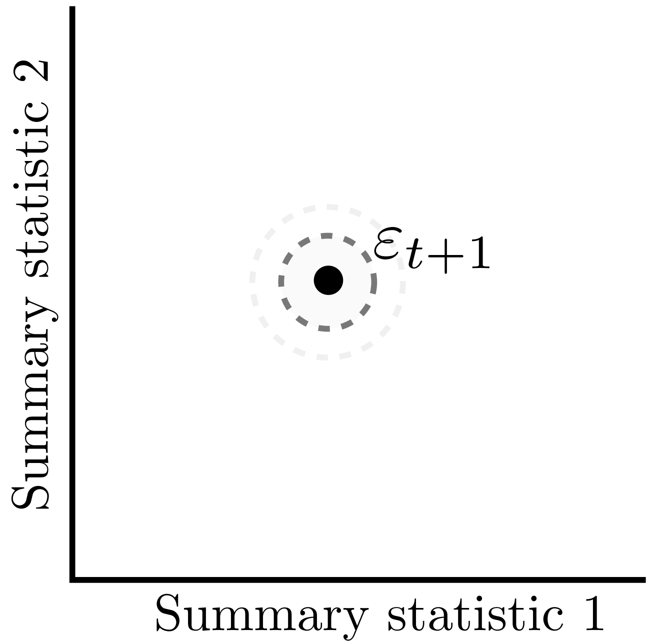
similar to Toni et al., JRS 2009

ABC-SMC

combine with a sequential Monte-Carlo scheme



$$w_t(\theta) \propto \frac{\pi(\theta)}{g_t(\theta)}$$



similar to Toni et al., JRS 2009

ABC

With distance d, threshold $\varepsilon > 0$, summary statistics s:

until N acceptances:
1. sample $\theta \sim g(\theta)$
2. simulate data $y \sim \pi(y|\theta)$
3. accept θ if $d(s(y), s(y_{\text{obs}})) \leq \varepsilon$

A "DERIVATION"

REJECTION SAMPLING

Background: Want to sample from f, but can only sample from $g \gg f$.

until N acceptances:
1. sample $\theta \sim g(\theta)$
2. accept θ with probability $\propto \frac{f(\theta)}{g(\theta)}$

Accepted θ are independent samples from $f(\theta)$.

Let $f = \pi(\theta | y_{\text{obs}}), g = \pi(\theta) \Rightarrow \frac{\pi(\theta | y_{\text{obs}})}{\pi(\theta)} \propto \pi(y_{\text{obs}} | \theta)$

- not available
- idea: **circumvent likelihood evaluation** by **simulating data** and matching them to the observed data

LIKELIHOOD-FREE REJECTION SAMPLING

until N acceptances:
1. sample $\theta \sim \pi(\theta)$
2. simulate data $y \sim \pi(y|\theta)$
3. accept θ if $y = y_{\text{obs}}$

- Acceptance probability: $\mathbb{P}[y_{\text{obs}}]$
- can be small in particular for continuous data
- idea: accept simulations that are **similar** to y_{obs}

ABC-REJECTION

With distance d, threshold $\varepsilon > 0$:

until N acceptances:
1. sample $\theta \sim \pi(\theta)$
2. simulate data $y \sim \pi(y|\theta)$
3. accept θ if $d(y, y_{\text{obs}}) \leq \varepsilon$

- **curse of dimensionality:** if the data are too high-dimensional, the probability of simulating similar data sets is small
- idea: create an informative lower-dimensional representation via **summary statistics**
- ideally minimal sufficient statistics

APPROXIMATE BAYESIAN POSTERIOR

We want:

$$\pi(\theta|y_{\text{obs}}) \propto p(y_{\text{obs}}|\theta)\pi(\theta)$$

APPROXIMATE BAYESIAN POSTERIOR

We want:

$$\pi(\theta \mid y_{\text{obs}}) \propto p(y_{\text{obs}} \mid \theta)\pi(\theta)$$

We get:

$$\pi_{ABC}(\theta \mid s(y_{\text{obs}})) \propto \int I(\{d(s(y), s(y_{\text{obs}})) \leq \varepsilon\})p(y \mid \theta)\mathrm{d}y\pi(\theta) \approx \frac{1}{N}\sum_{i=1}^{N} \delta_{\theta^{(i)}}(\theta)$$

with distance d, threshold $\varepsilon > 0$, and summary statistics s

SOURCES OF APPROXIMATION ERRORS IN ABC

- model error (as for every model of reality)
- Monte-Carlo error (as for sampling in general)
- summary statistics
- epsilon threshold

SOURCES OF APPROXIMATION ERRORS IN ABC

- model error (as for every model of reality)
- Monte-Carlo error (as for sampling in general)
- summary statistics
- epsilon threshold

> *Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise.*

John Tukey 1962

HOW?

πABC

github.com/icb-dcm/pyabc

Klinger et al., Bioinformatics 2018 and Schälte et al., JOSS 2022

flexible

modular
configure
extend

user-friendly

documented
self-tuning
robust

scalable

1 to 1,000s cores
efficient
tested

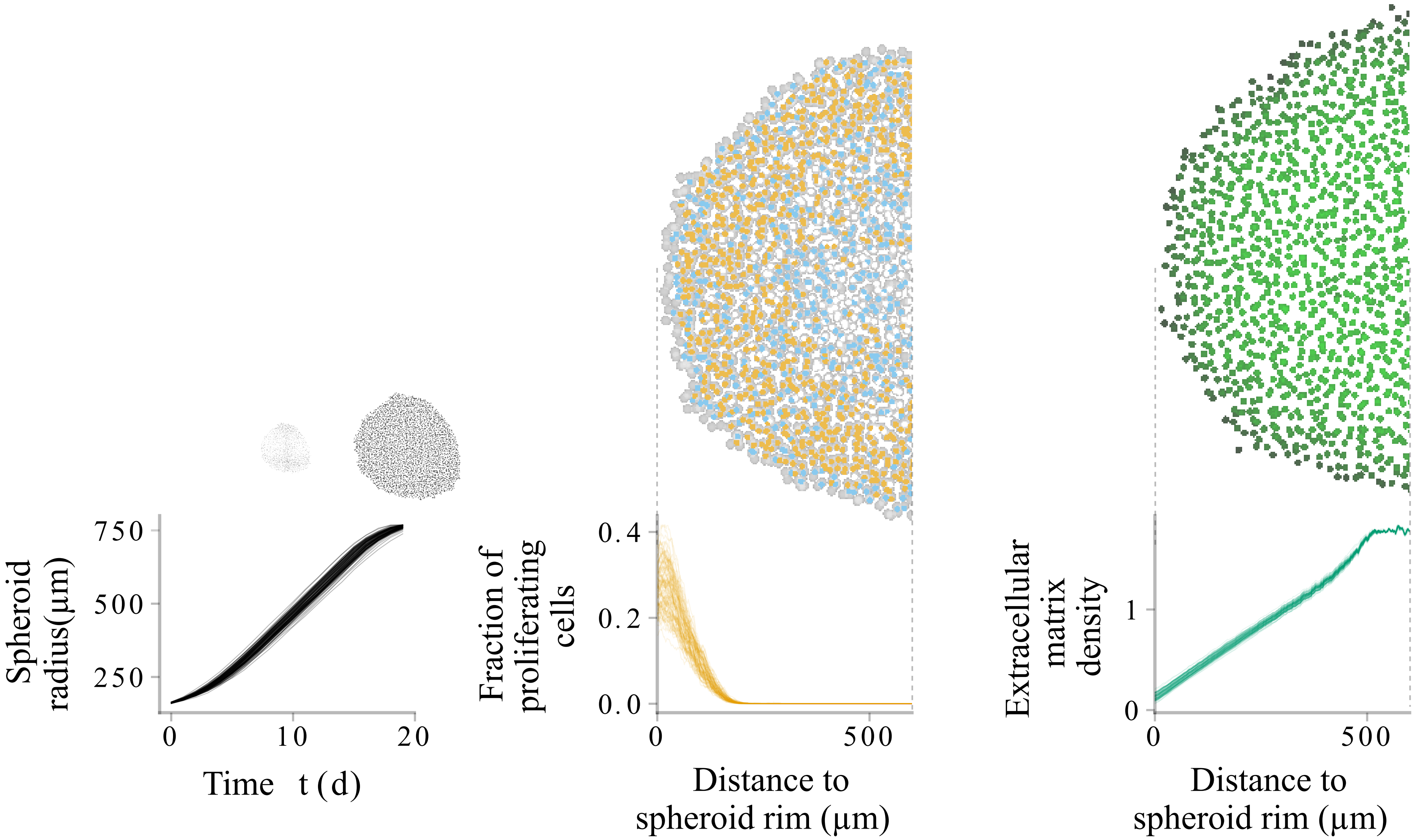
EASY TO USE

```
# define problem
abc = pyabc.ABCSMC(model, prior, distance)

# pass data
abc.new(database, observation)

# run it
abc.run()
```

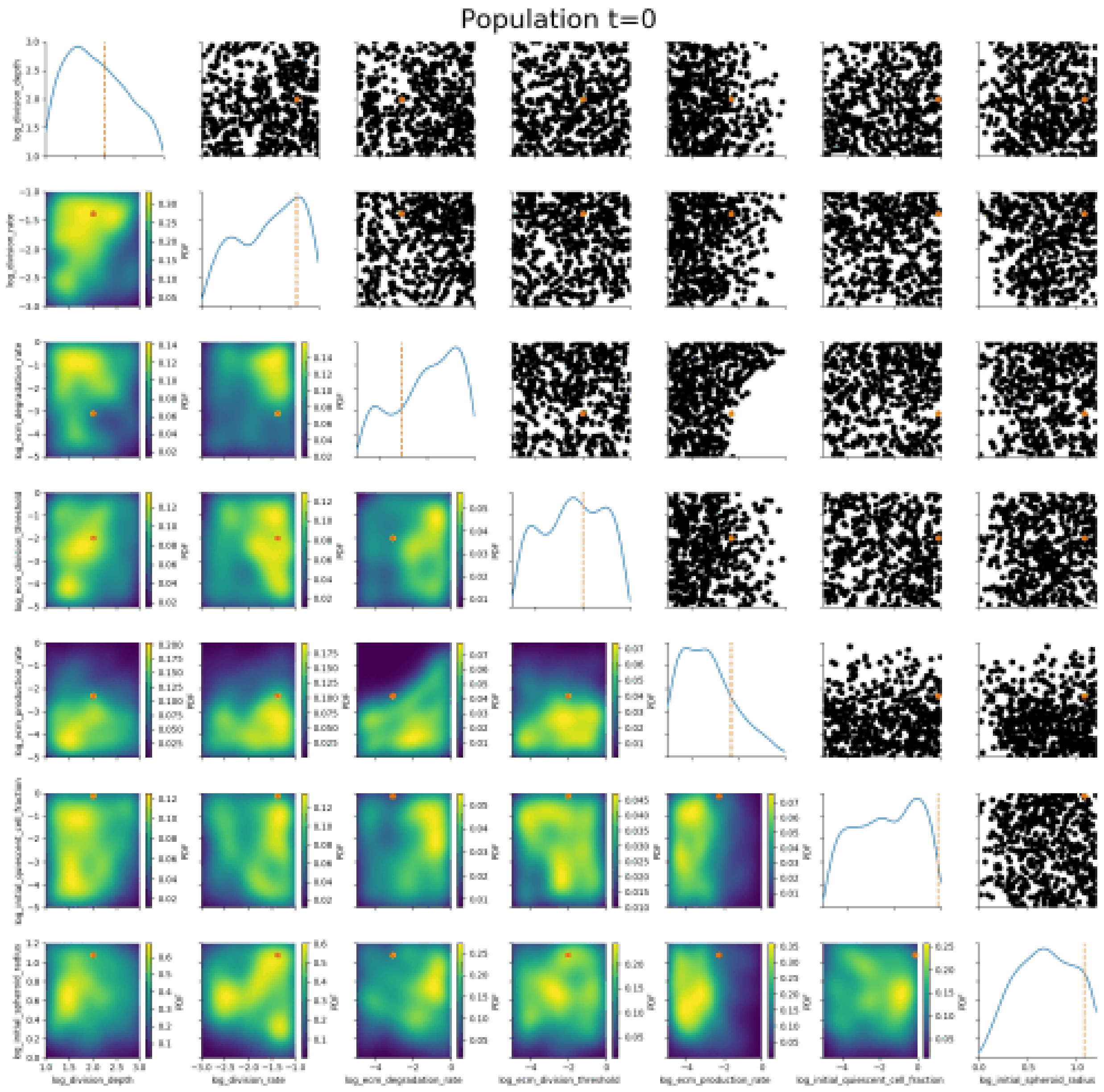
EXAMPLE: TUMOR GROWTH MODEL

based on Jagiella et al., Cell Systems 2017

DEFINE SUMMARY STATISTICS

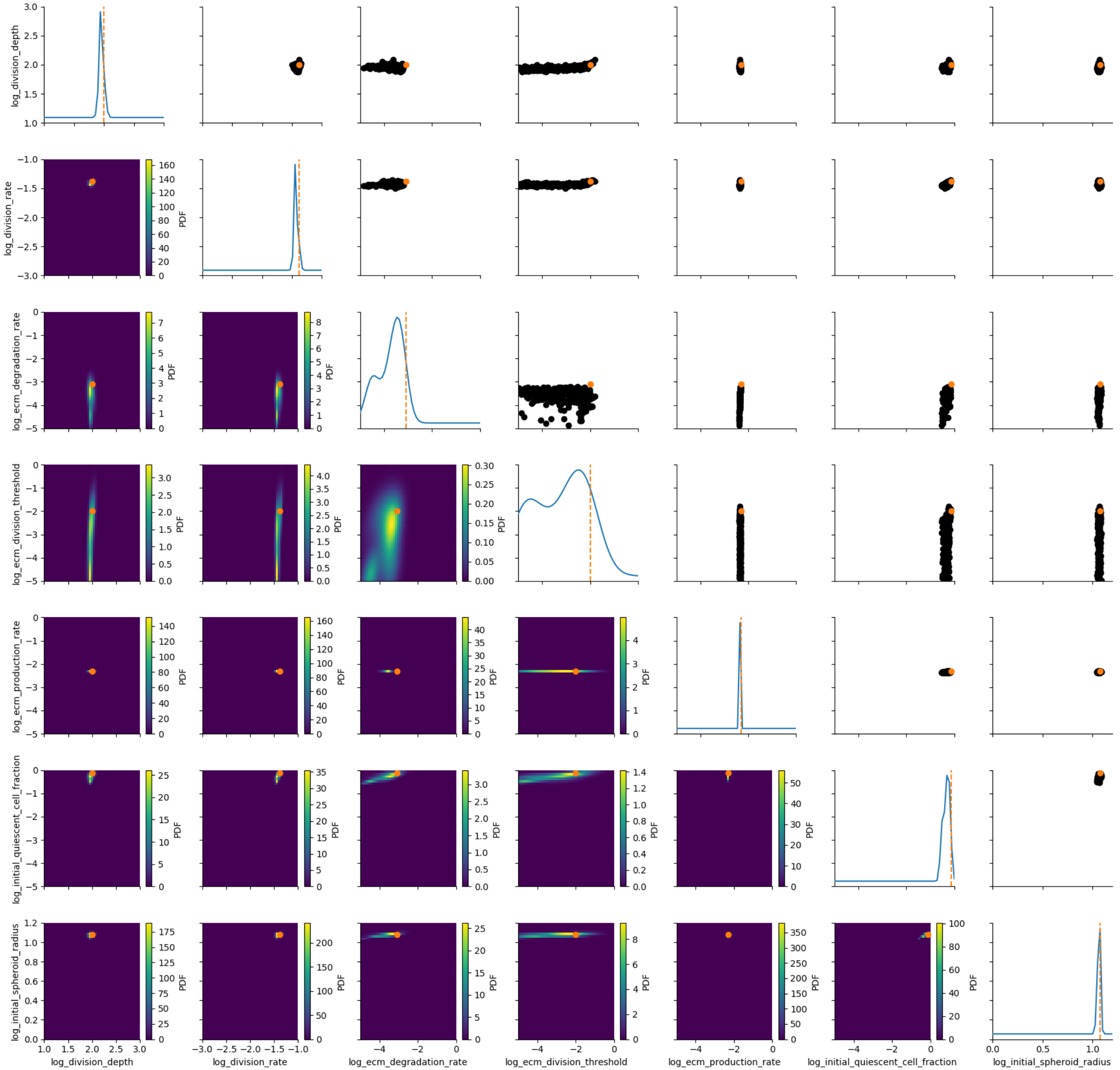


- 400 cores
- 2 days
- 1.8e6 simulations

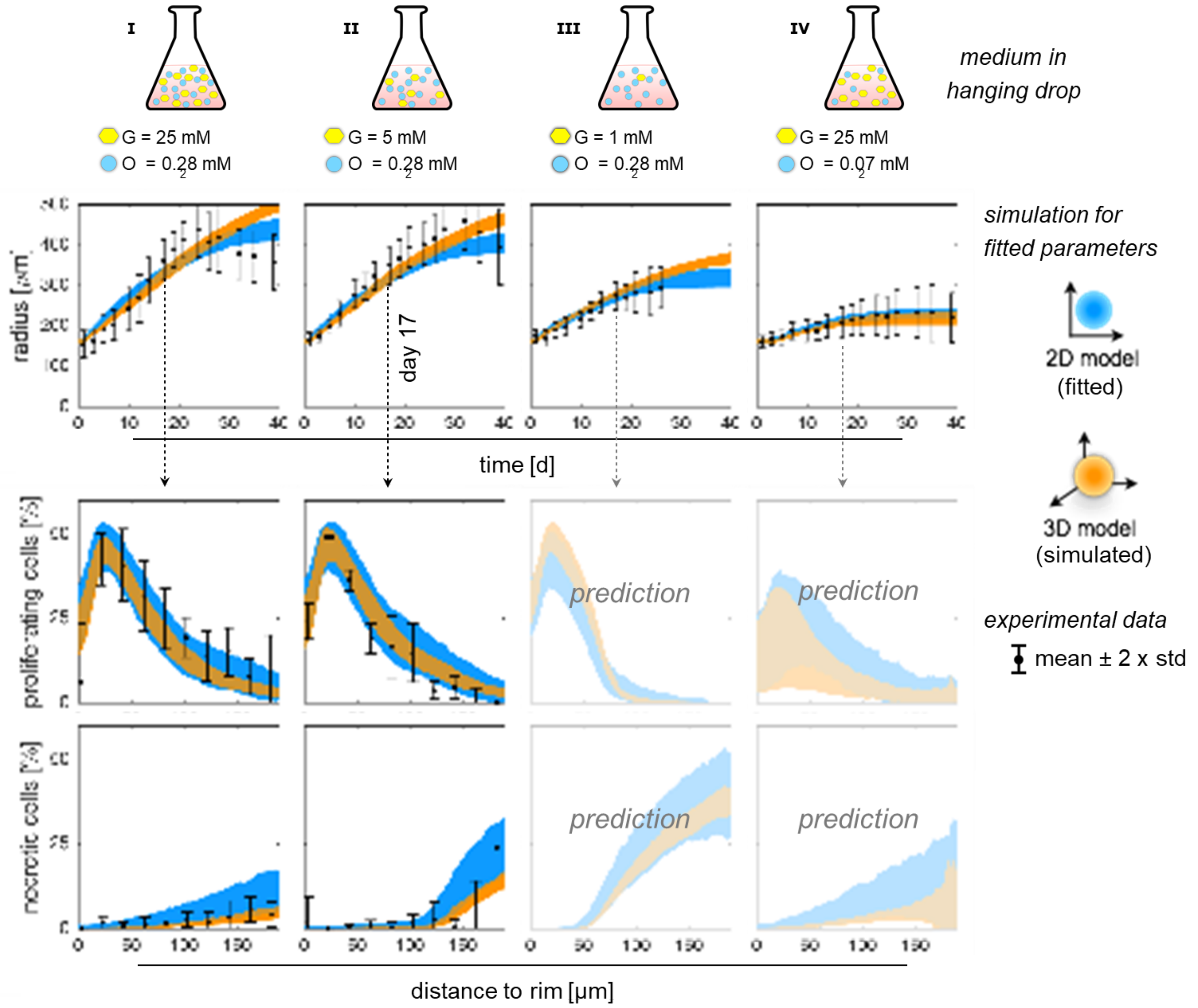


ABC worked where many other methods had failed.

Population t=37



ABC worked where many other methods had failed.



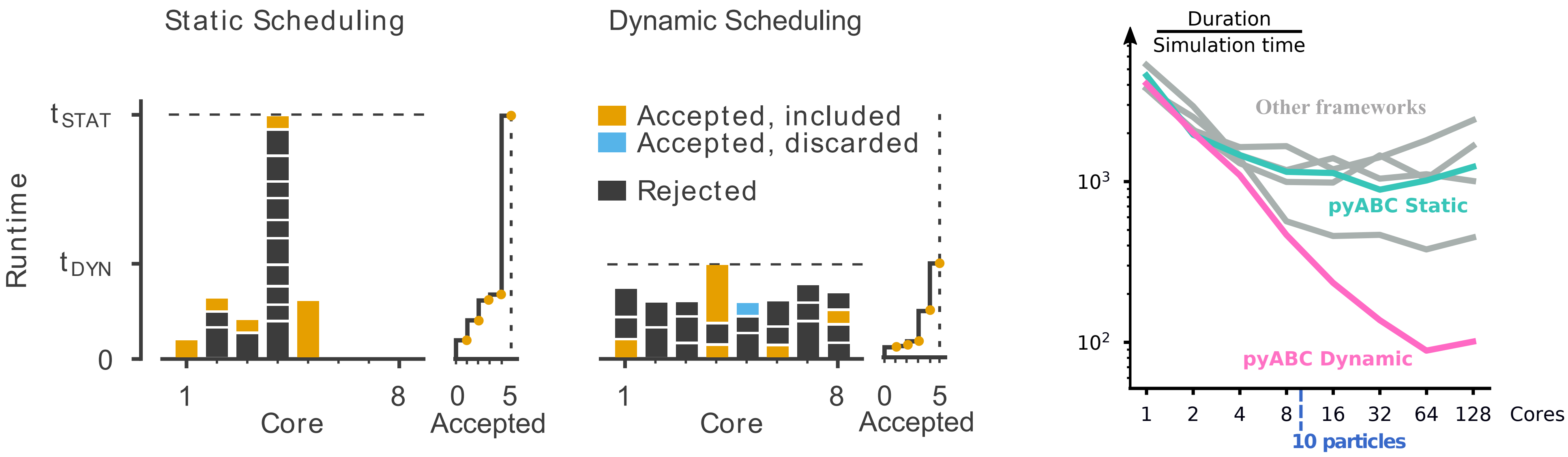
Uncertainty-aware predictions, easy data integration.

ALGORITHMIC DETAILS

PARALLEL BACKENDS: 1 TO 1,000S CORES

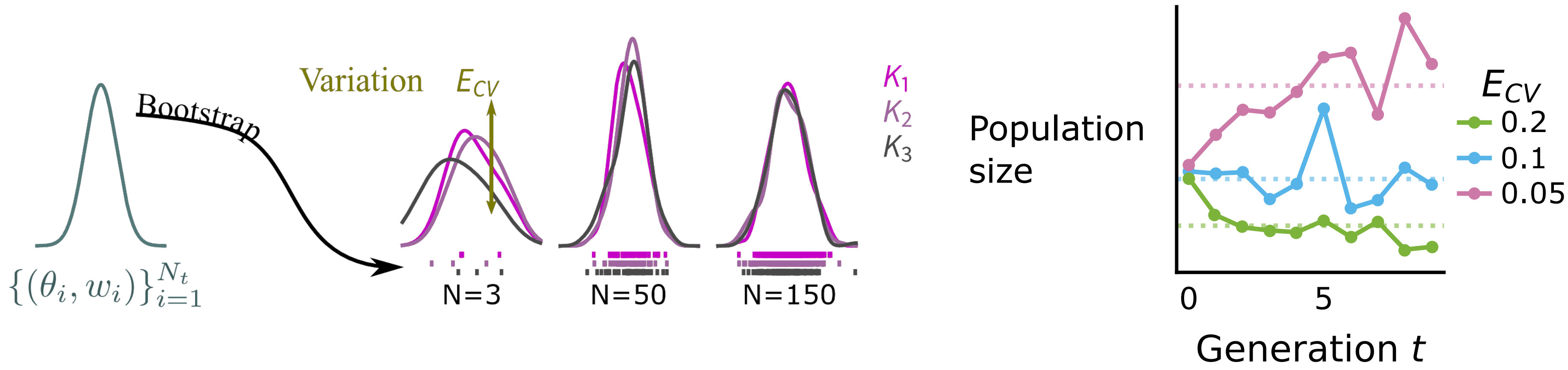
PARALLELIZATION STRATEGIES

Klinger et al., CMSB Proceedings 2017



ADAPTIVE POPULATION SIZES

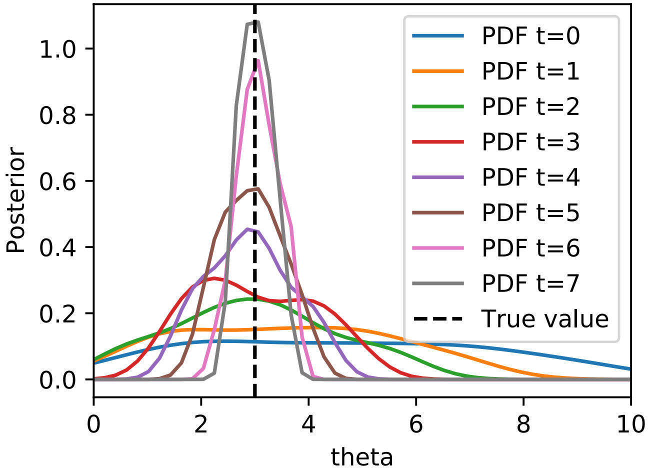
Klinger et al., CMSB Proceedings 2017



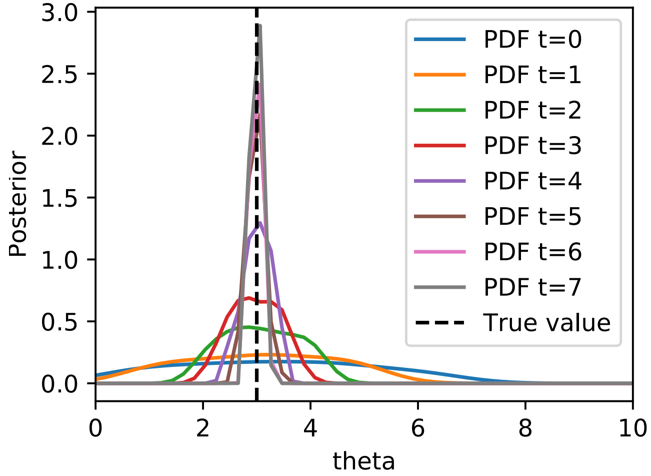
idea: adapt population size trying to match a target accuracy

SELF-TUNING DISTANCE FUNCTIONS

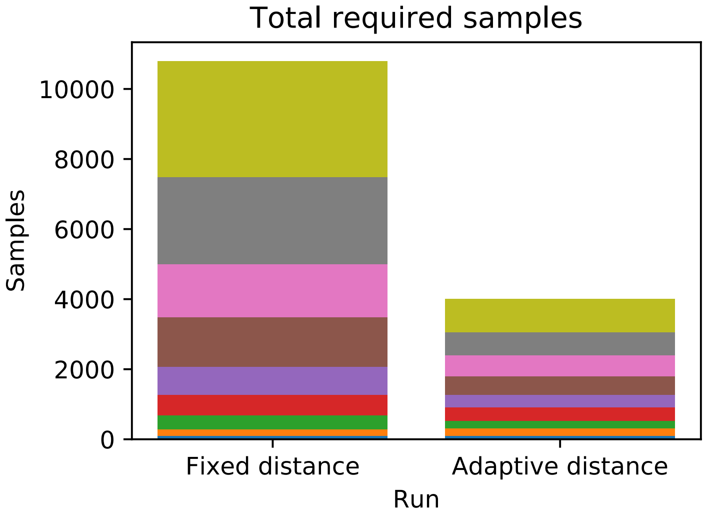
based on Prangle, Bayesian Analysis 2015



standard



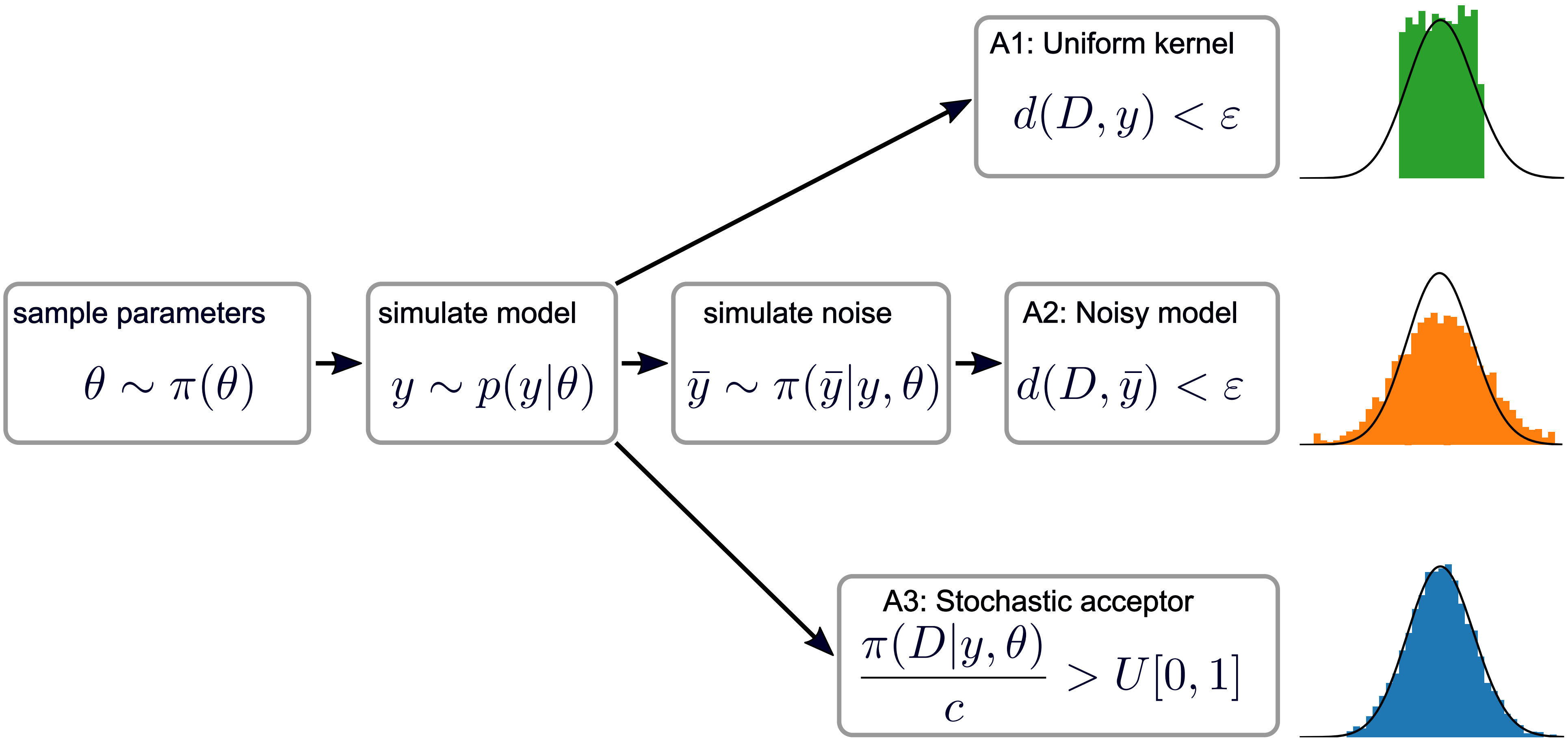
adaptive



MEASUREMENT NOISE

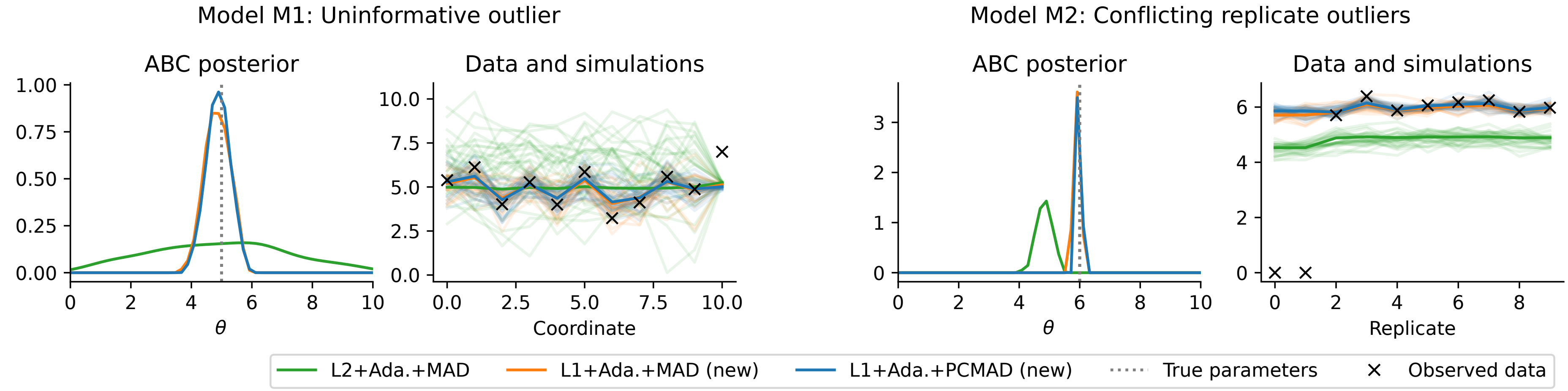
Schälte et al., Bioinformatics 2020

How to efficiently account for measurement noise in ABC?



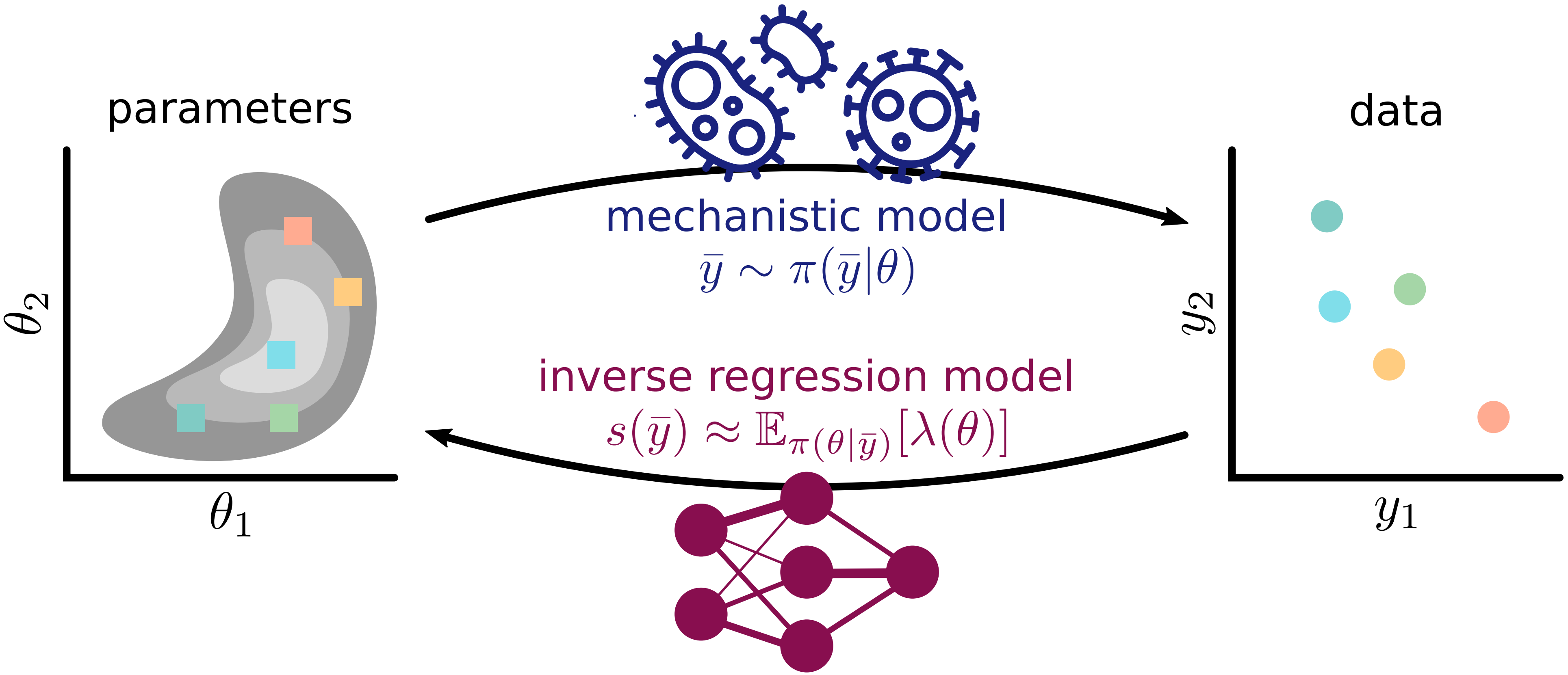
ROBUSTNESS TO OUTLIERS

Schälte et al., bioRxiv 2021



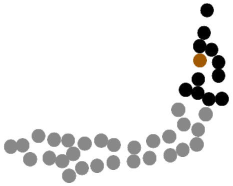
ASSESS DATA INFORMATIVENESS

based on Fearnhead and Prangle, RSS B 2012
Schälte and Hasenauer, bioRxiv 2022

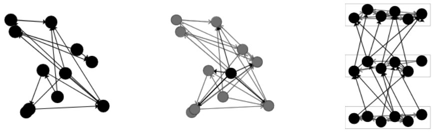


AND ...

adaptive proposal kernels

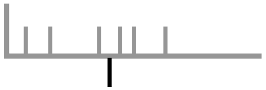


model selection



speaks python, R and julia

adaptive epsilon schedules

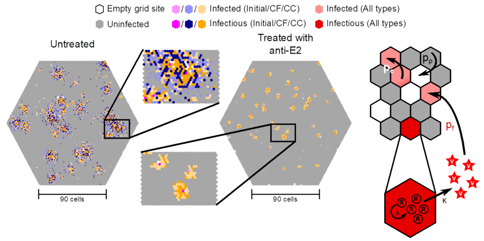


ε

early rejection

$$\sum_i d_i < \varepsilon$$

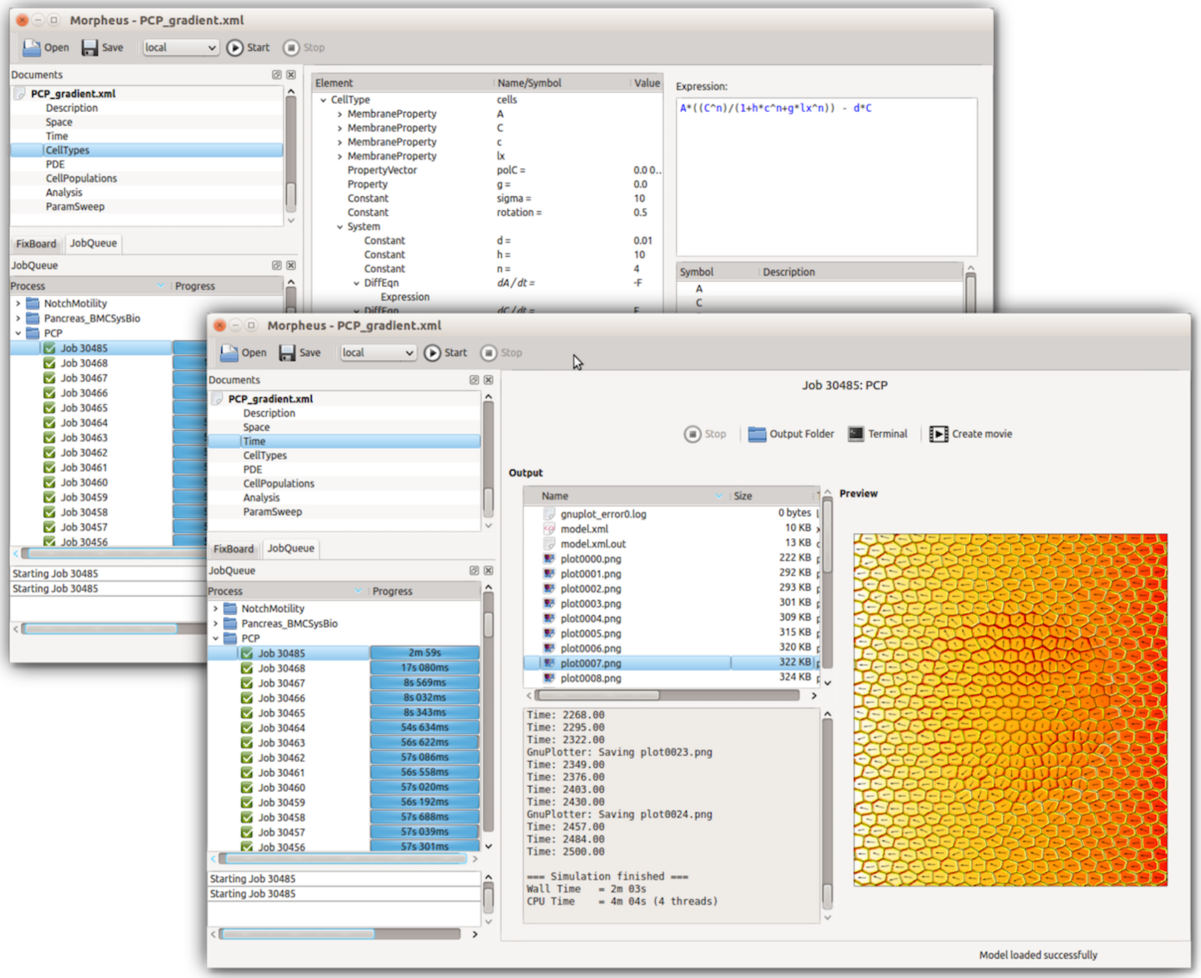
applications



...

FitMultiCell

Joint initiative to perform inference for multi-cellular models



Morpheus toolbox: Staruß et al., Bioinformatics 2014

EFFICIENT EXACT ABC WITH NOISE

THE PROBLEM: (BIOLOGICAL) DATA ARE NOISY

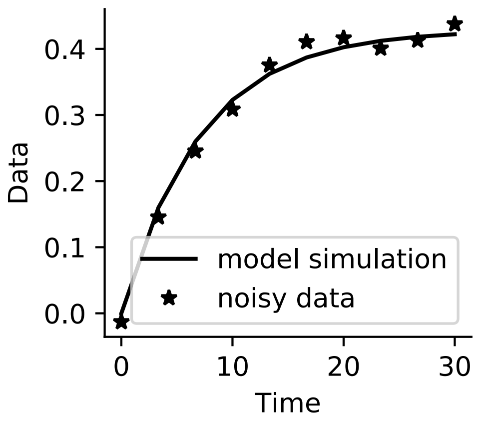
$\bar{y}_{\mathrm{obs}} =$ 

WHAT HAPPENS WHEN IGNORING NOISE IN ABC?

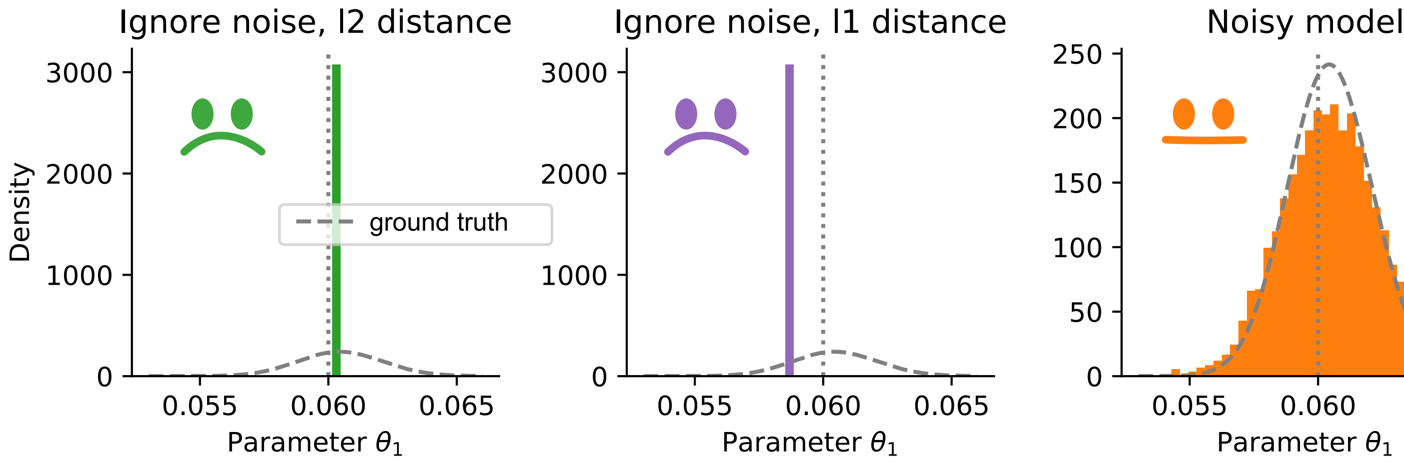
Assume: Model $y \sim \pi(y|\theta)$ does not account for noise.

But: Measurements are noisy, $\bar{y}_{\text{obs}} \sim \pi(\bar{y}|y,\theta)$.

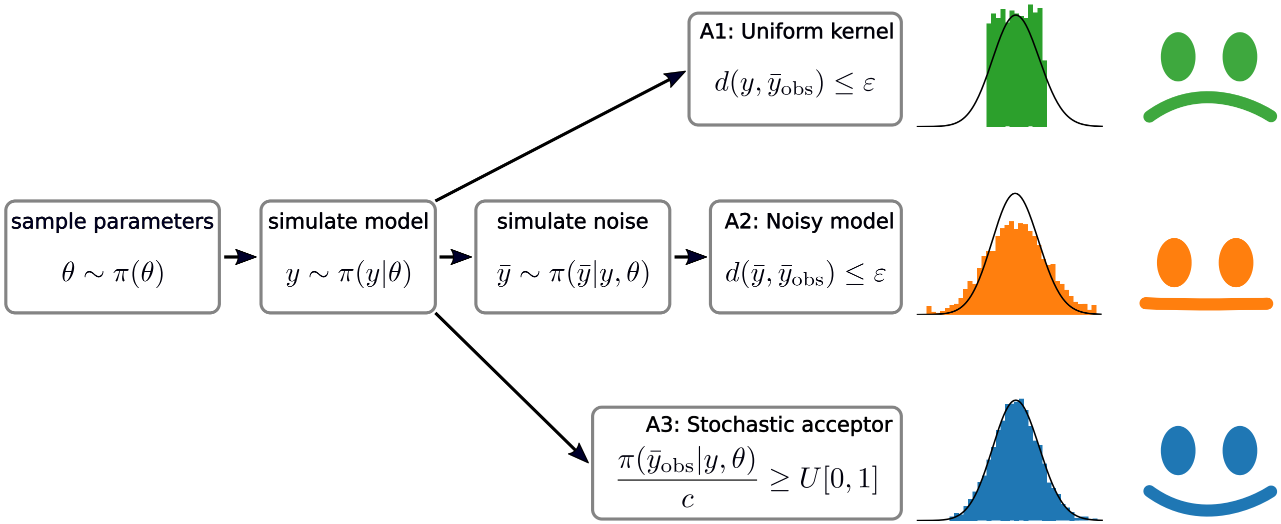
data and simulations



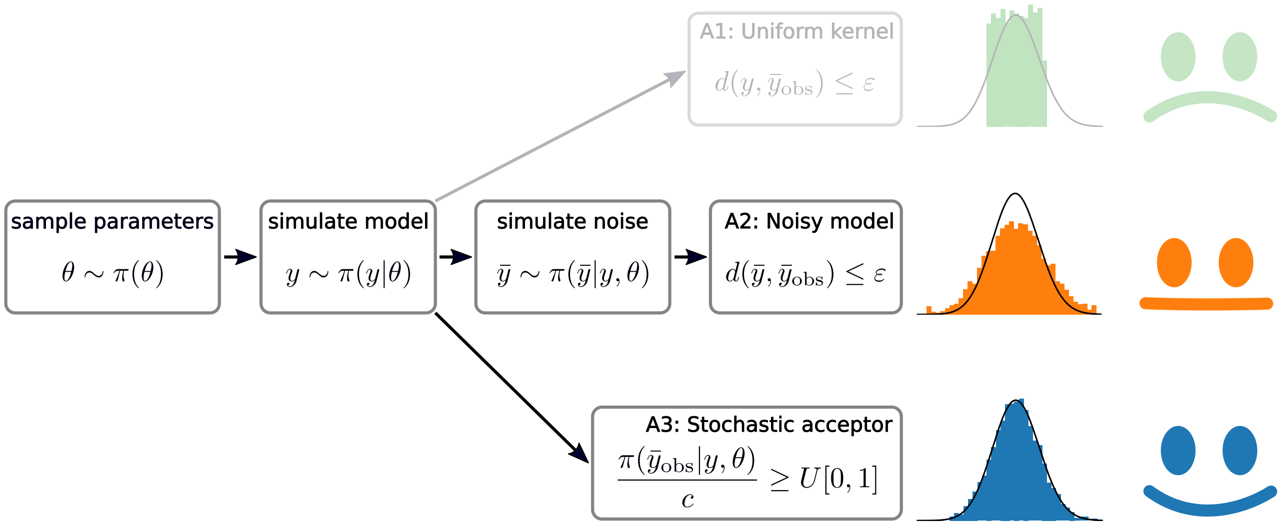
different ABC posteriors



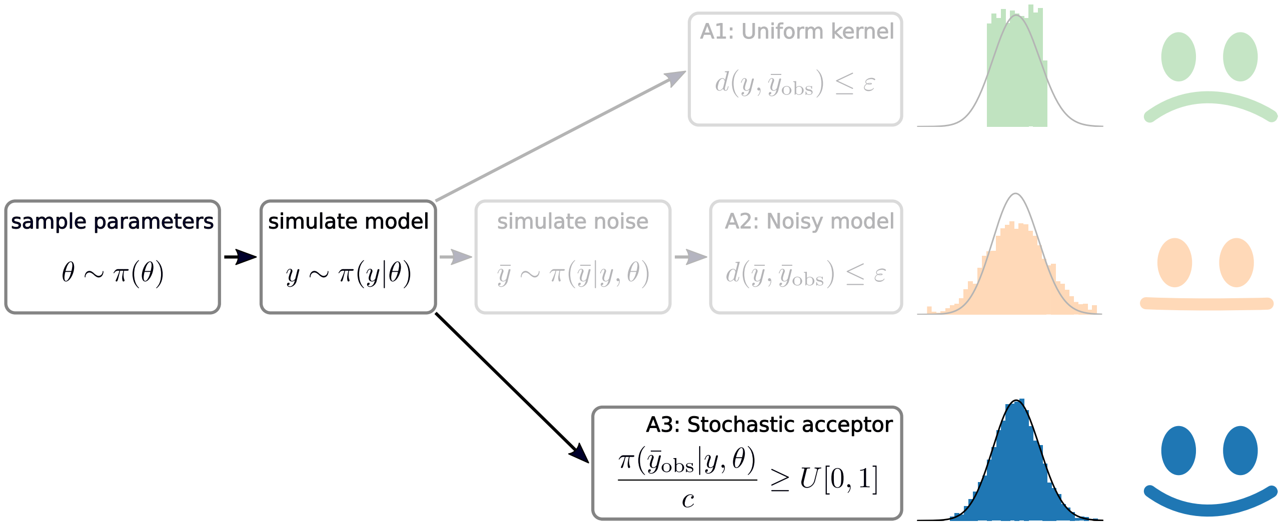
HOW TO ACCOUNT FOR NOISE?



HOW TO ACCOUNT FOR NOISE?



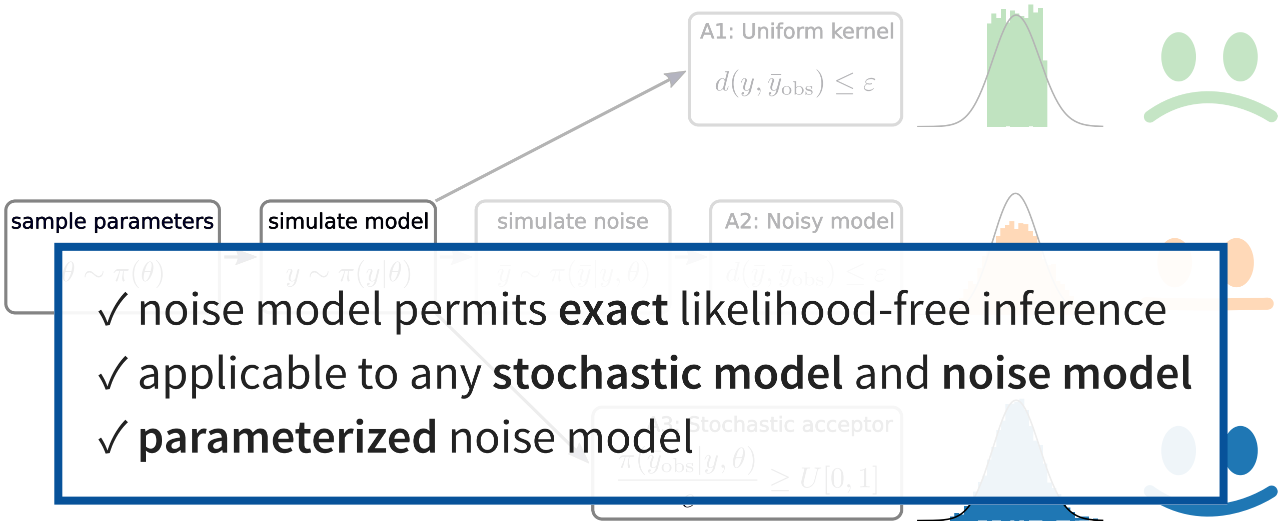
HOW TO ACCOUNT FOR NOISE?



"ABC gives exact inference for the wrong model"

Richard Wilkinson, Stat. App. Gen. Mol. Bio. 2013

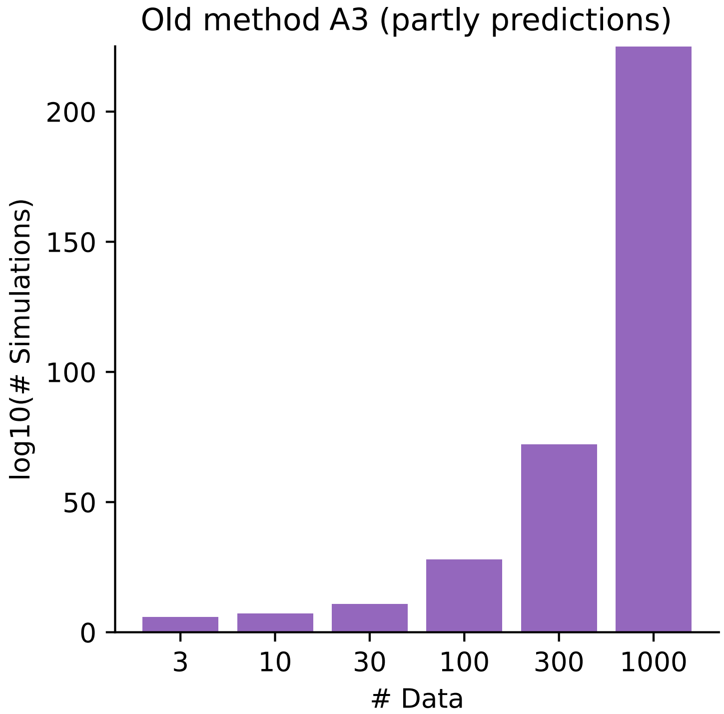
HOW TO ACCOUNT FOR NOISE?



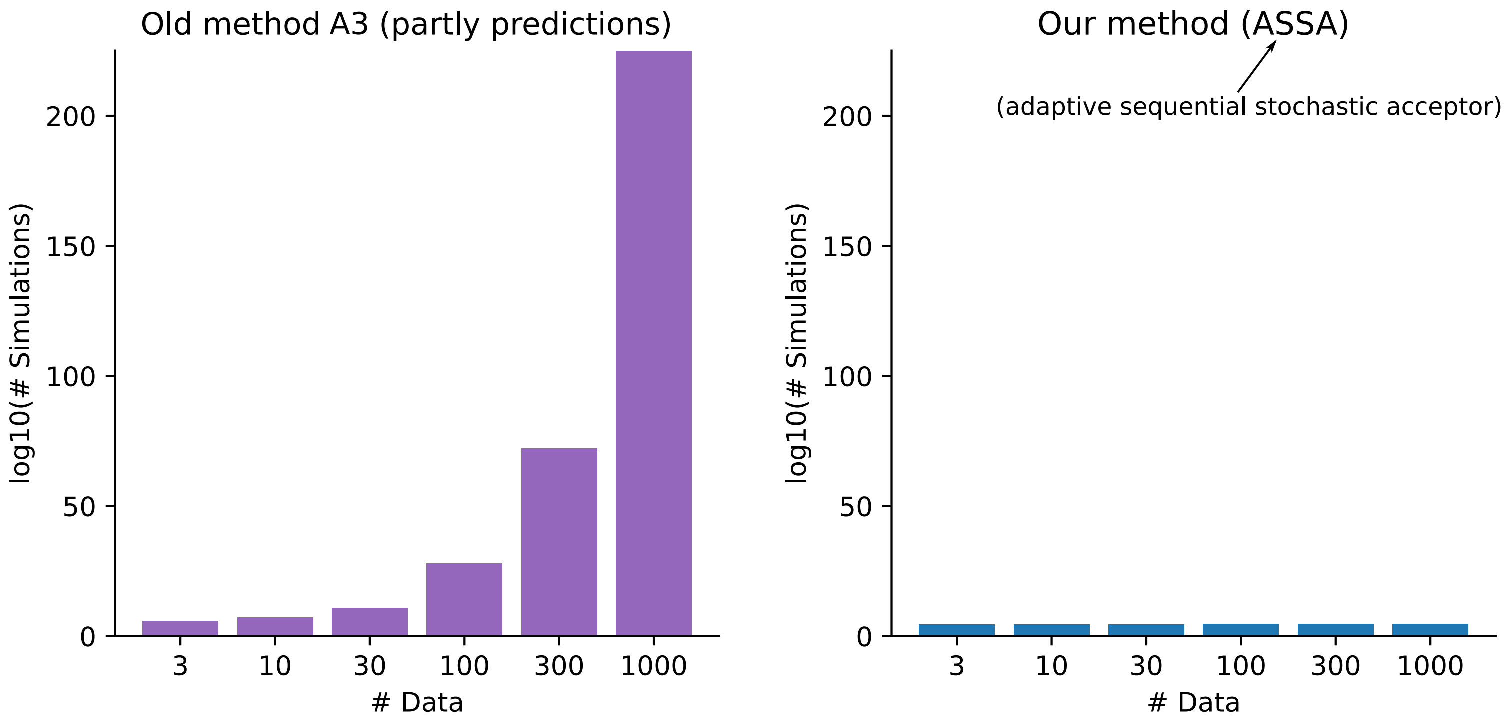
- ✓ noise model permits **exact** likelihood-free inference
- ✓ applicable to any **stochastic model** and **noise model**
- ✓ **parameterized** noise model

> *"ABC gives exact inference for the wrong model"*
>
> *Richard Wilkinson, Stat. App. Gen. Mol. Bio. 2013*

PROBLEM: EXISTING METHODS DO NOT SCALE IN PRACTICE

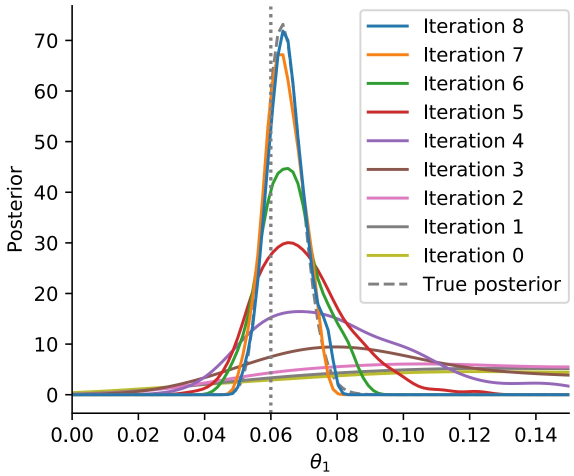


PROBLEM: EXISTING METHODS DO NOT SCALE IN PRACTICE



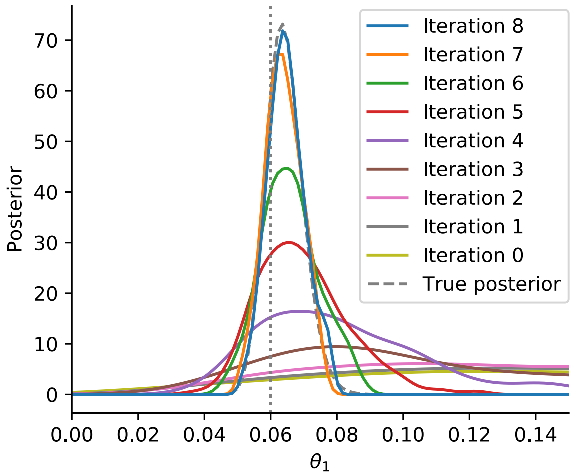
CAN WE MAKE IT MORE EFFICIENT?

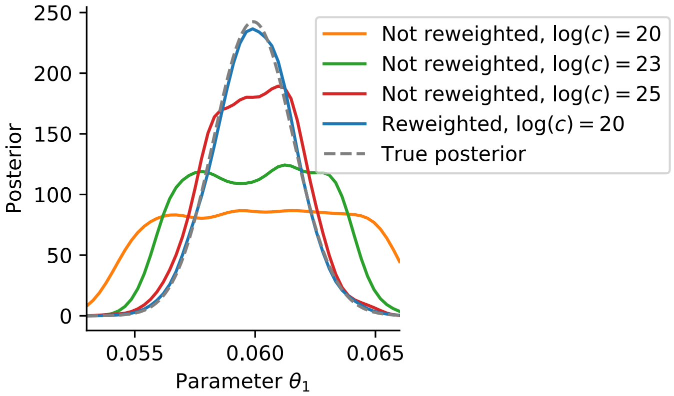
CAN WE MAKE IT MORE EFFICIENT?



- How to *propose parameters*?
  - ⇝ integrate in **SMC** via **tempering,** $\pi(\bar{y}_{\text{obs}} \,|\, y, \theta)^{1/T_t}$.

CAN WE MAKE IT MORE EFFICIENT?



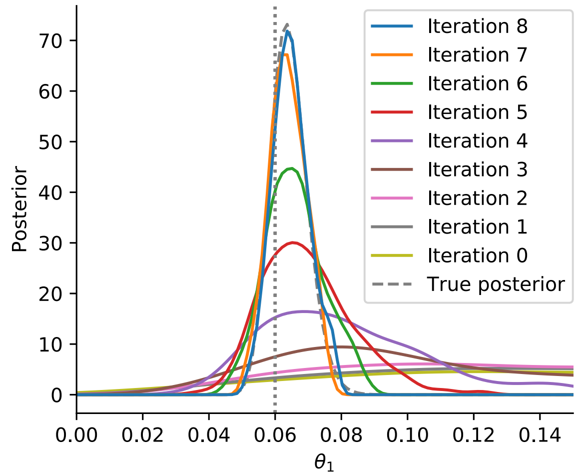


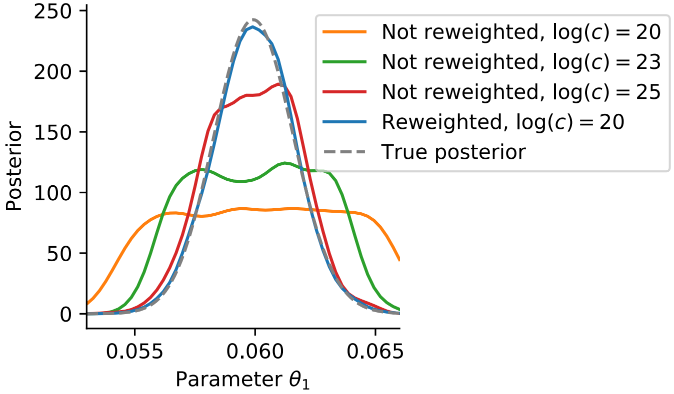
- How to *propose parameters*?
  - ⇝ integrate in **SMC** via **tempering,** $\pi(\bar{y}_{\text{obs}}\,|\,y,\,\theta)^{1/T_t}$.

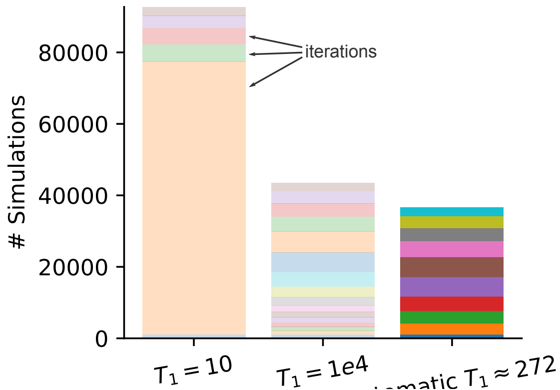
- How to choose the *normalization* c?
  - ⇝ based on **previous samples**, and avoid decapitation via **reweighting**

$$\tilde{w} := \frac{\left(\frac{\pi(\bar{y}_{\text{obs}}|y,\theta)}{c_t}\right)^{1/T_t}}{\min\left[\frac{\pi(\bar{y}_{\text{obs}}|y,\theta)}{c_t},1\right]^{1/T_t}} \cdot \frac{\pi(\theta)}{g_t(\theta)}$$

CAN WE MAKE IT MORE EFFICIENT?







automati

- How to *propose parameters*?

  ⇝ integrate in **SMC** via **tempering,** $\pi(\bar{y}_{\text{obs}}\,|\,y,\theta)^{1/T_t}$.

- How to choose the *normalization* c?

  ⇝ based on **previous samples**, and avoid decapitation via **reweighting**

$$\tilde{w} := \frac{\left(\frac{\pi(\bar{y}_{\text{obs}}|y,\theta)}{c_t}\right)^{1/T_t}}{\min\left[\frac{\pi(\bar{y}_{\text{obs}}|y,\theta)}{c_t}, 1\right]^{1/T_t}} \cdot \frac{\pi(\theta)}{g_t(\theta)}$$

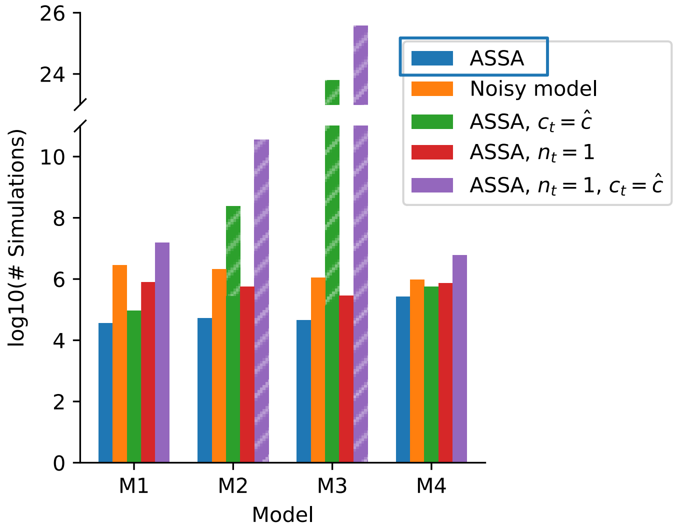
- How to choose the *temperatures* T_t, $t = 1, \ldots, n_t$?

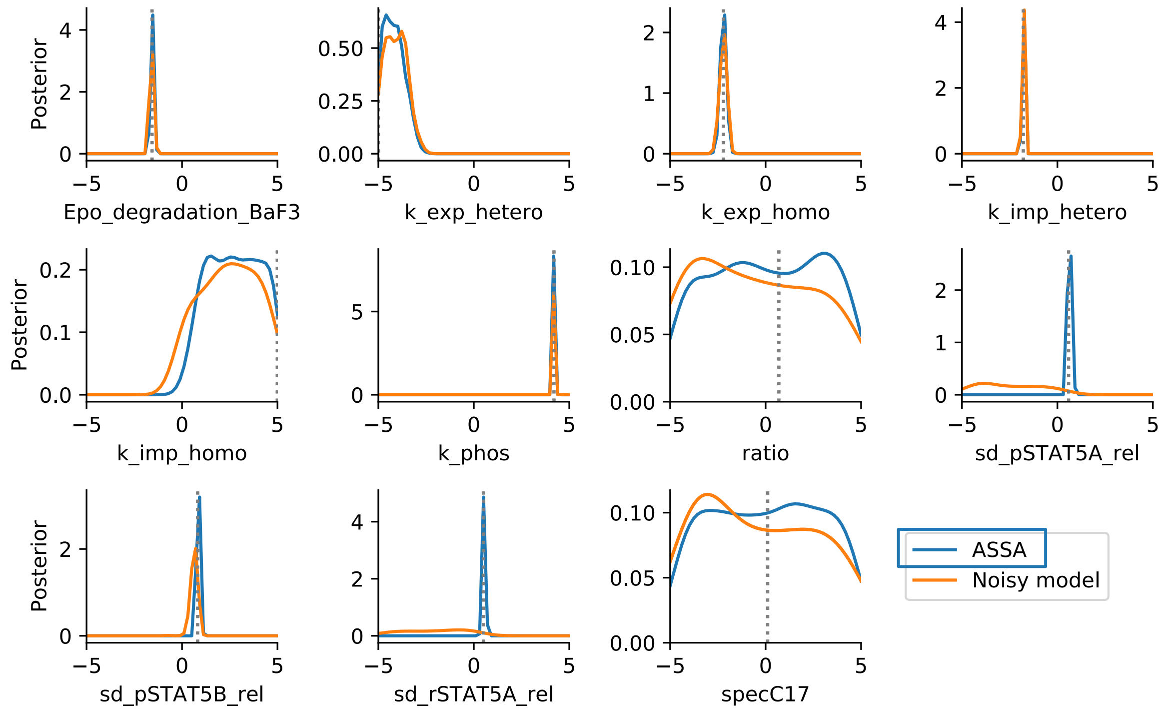
  ⇝ **predict the acceptance rate**

$$\gamma = \int\left(\int \min\left[\left(\frac{\pi(\bar{y}_{\text{obs}}|y,\theta)}{c_t}\right)^{1/T}, 1\right]\pi(y\,|\,\theta)dy\right)g_t(\theta)d\theta$$

(esp. allows choosing T_1)

EVALUATION



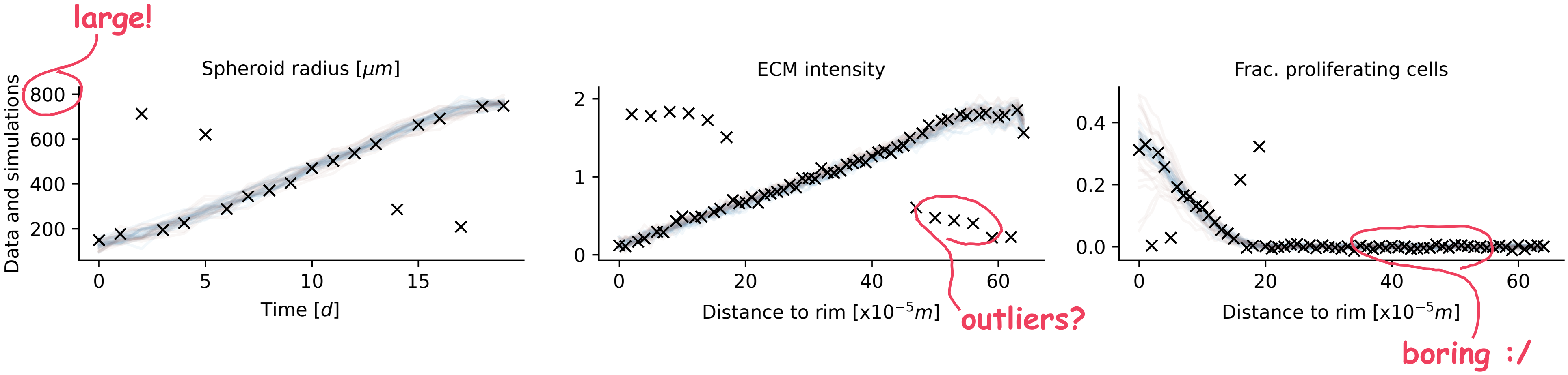


✓ Applicable to **various model** and **noise model** types
✓ orders of magnitude **speed-up**
✓ **scales** to challenging application problems

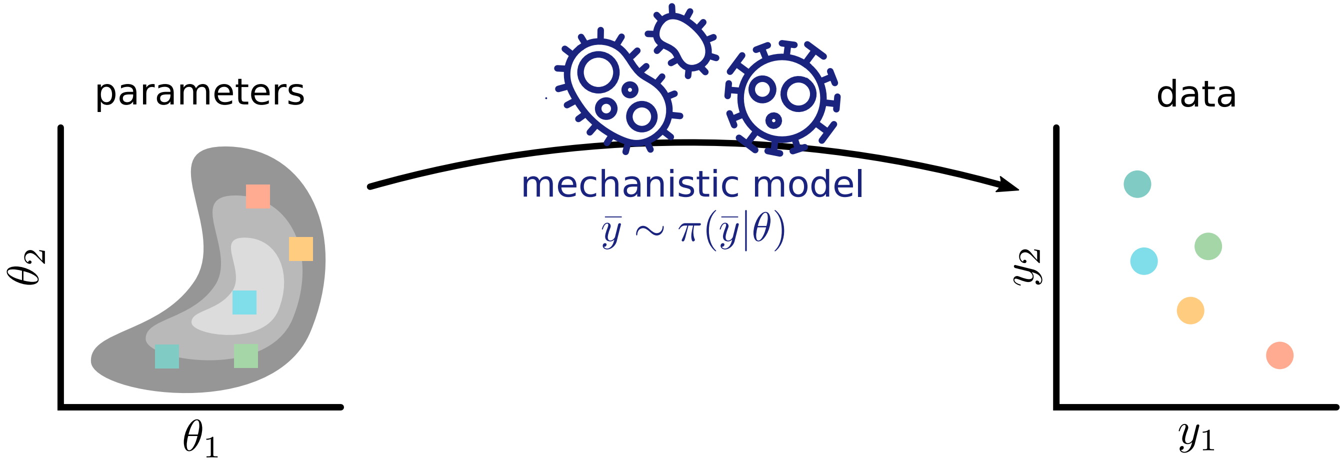
ROBUST AND EFFICIENT ABC VIA INVERSE MACHINE LEARNING MODELS

THE PROBLEM:

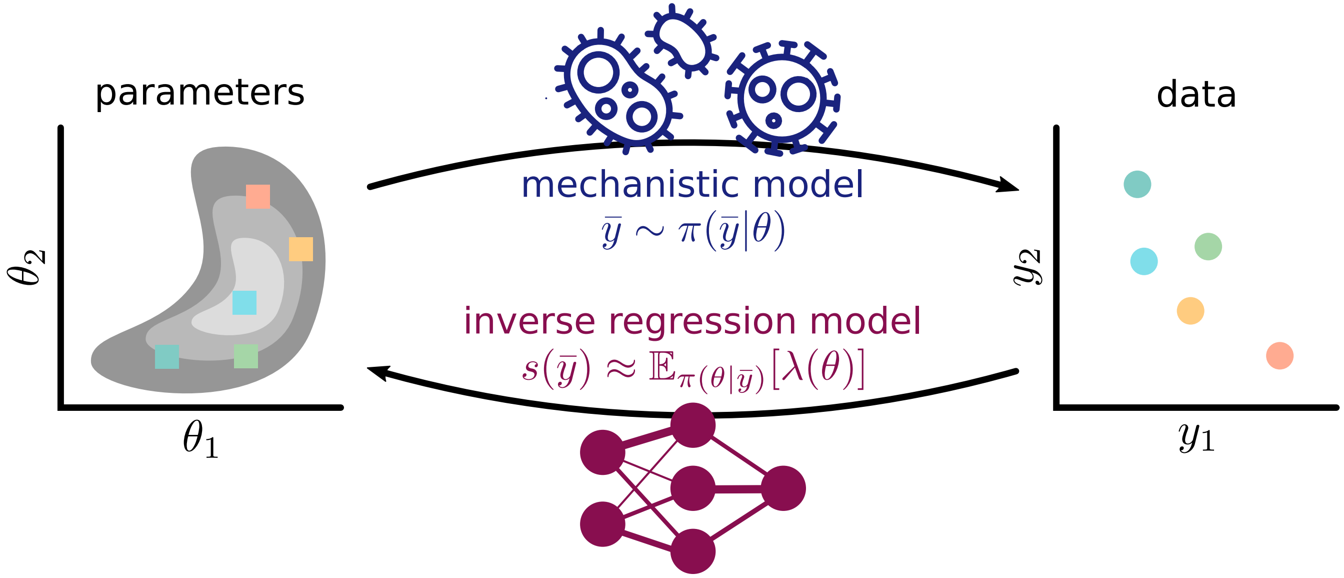
FITTING HETEROGENEOUS DATA



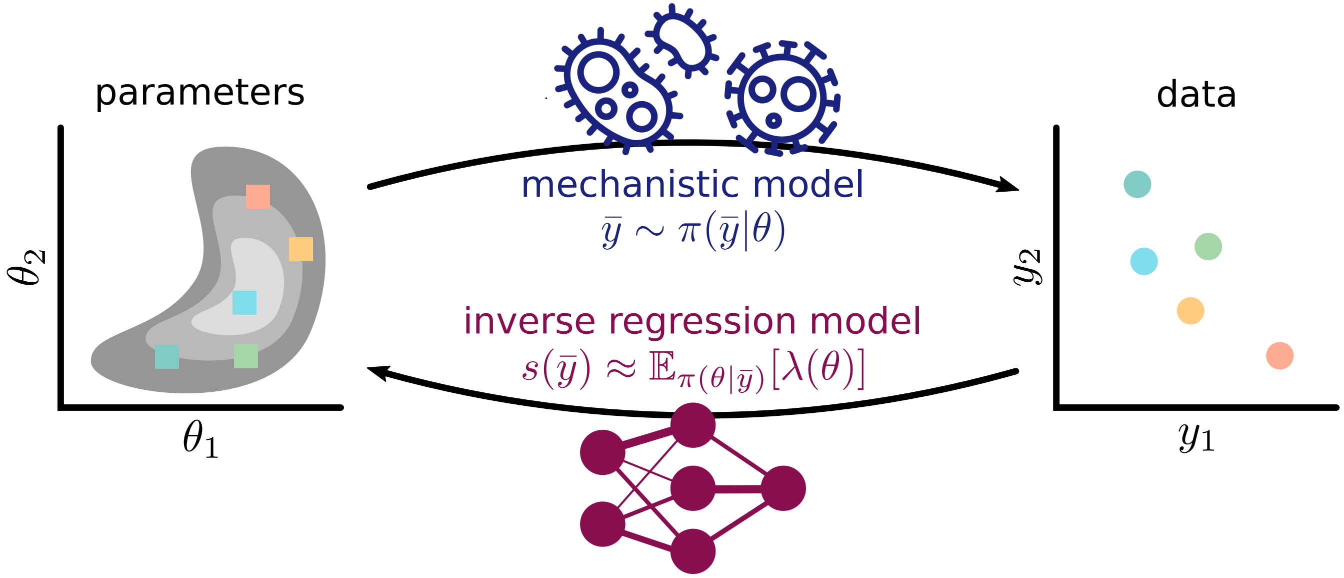
HOW TO ACCOUNT FOR DATA INFORMATIVENESS?



HOW TO ACCOUNT FOR DATA INFORMATIVENESS?

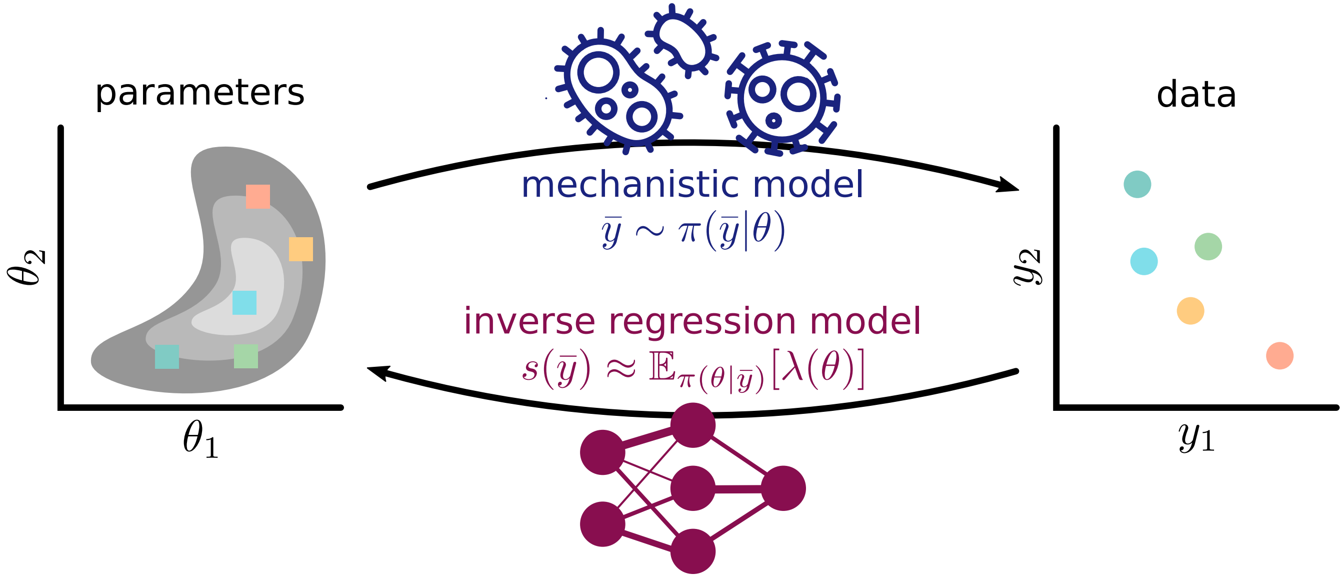


HOW TO ACCOUNT FOR DATA INFORMATIVENESS?



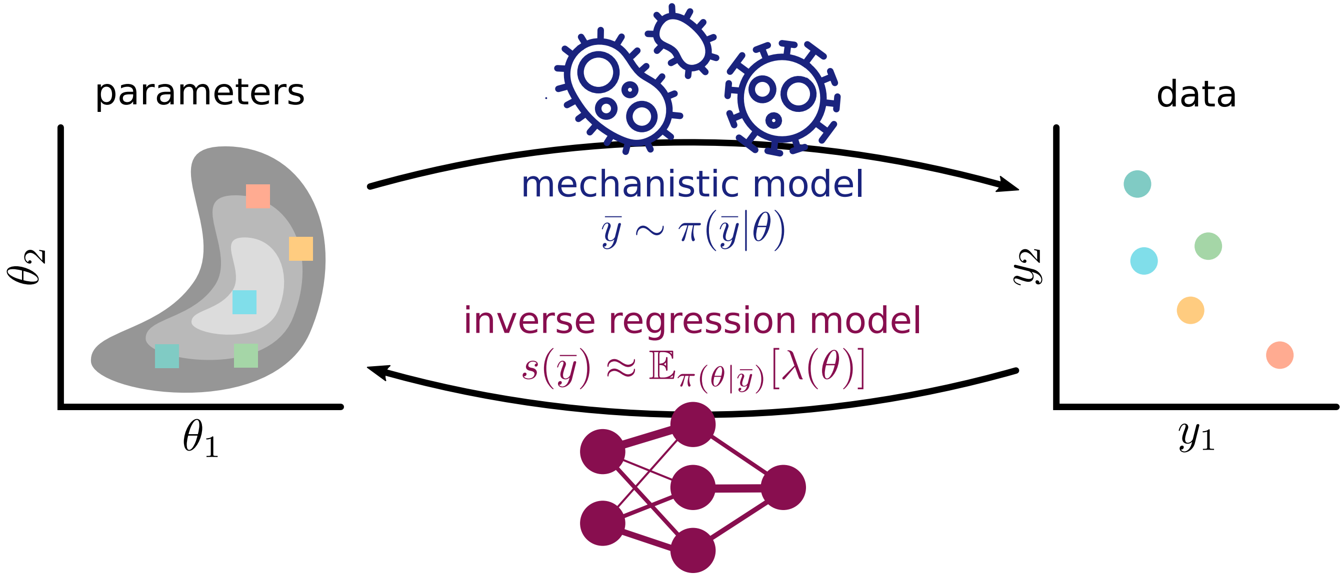
- construct low-dimensional **summary statistics** (see Fearnhead & Prangle, JRSS 2012)

HOW TO ACCOUNT FOR DATA INFORMATIVENESS?



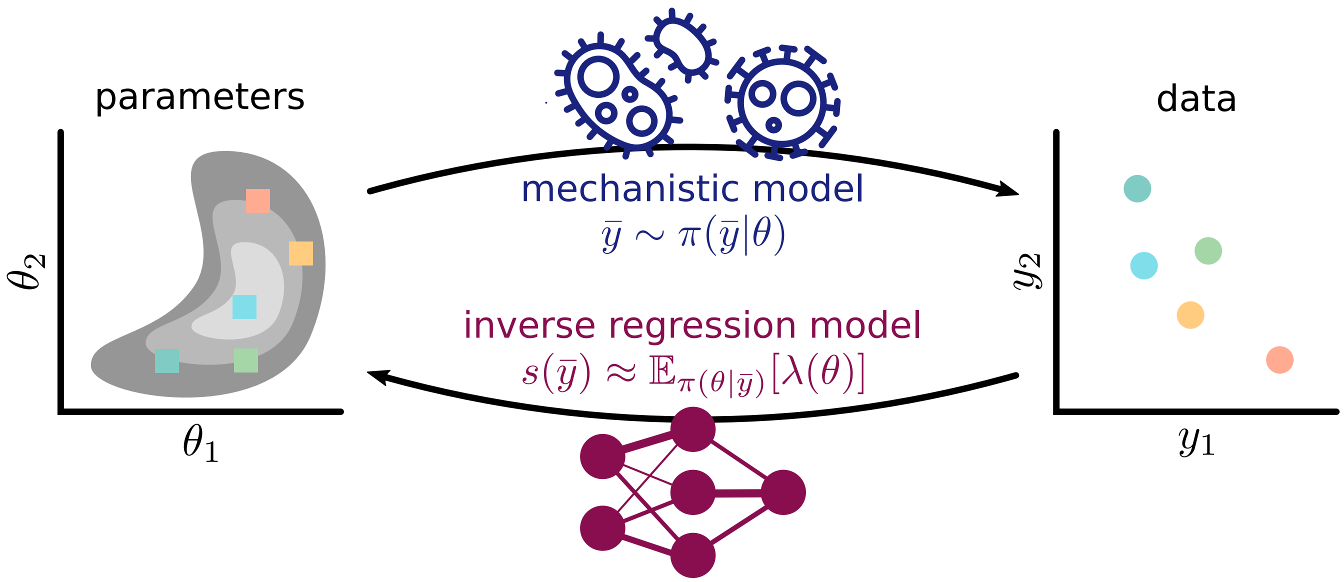
- construct low-dimensional **summary statistics** (see Fearnhead & Prangle, JRSS 2012)
- or: define **sensitivity weights** via the sensitivity matrix $S = \nabla_{\bar{y}} s(\bar{y}_{\text{obs}})$, $q_{i_y} = \sum_{i_\theta=1}^{n_\theta} \frac{\left|S_{i_y i_\theta}\right|}{\sum_{j_y=1}^{n_y} \left|S_{j_y i_\theta}\right|}$

HOW TO ACCOUNT FOR DATA INFORMATIVENESS?



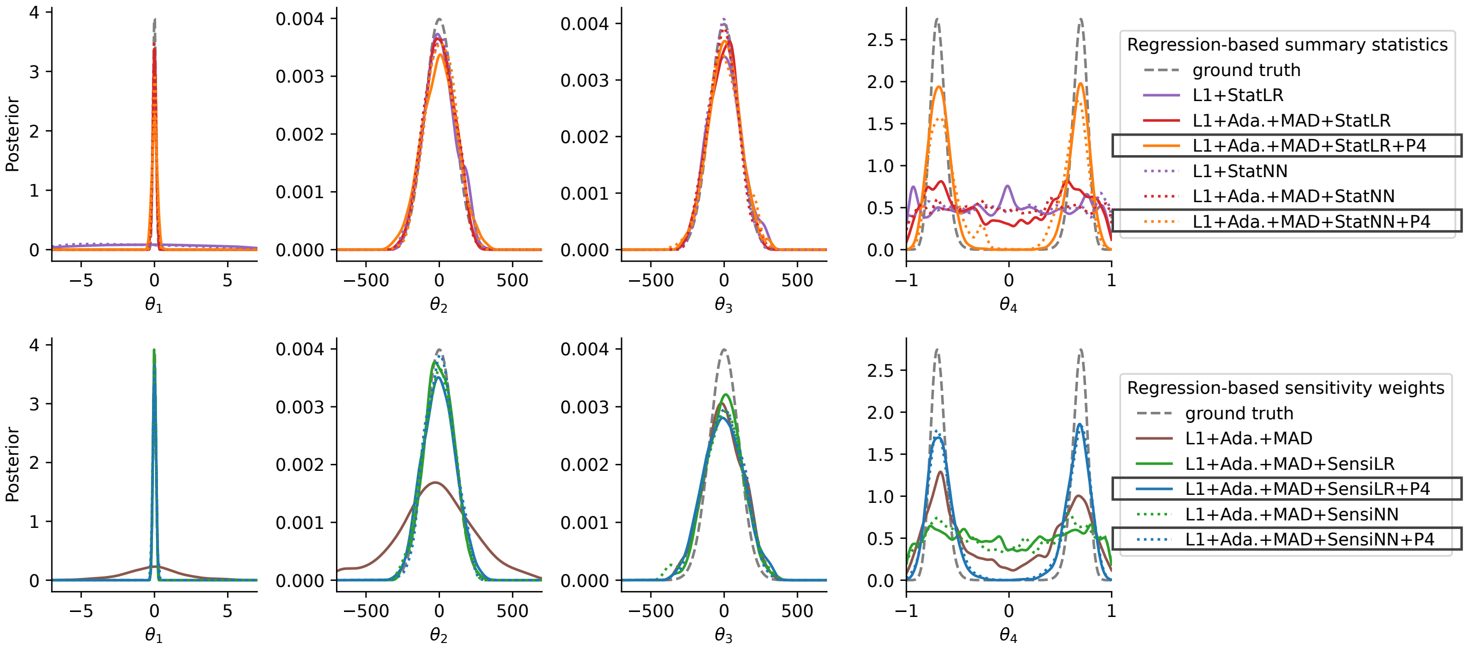
- construct low-dimensional **summary statistics** (see Fearnhead & Prangle, JRSS 2012)
- or: define **sensitivity weights** via the sensitivity matrix $S = \nabla_{\bar{y}} s(\bar{y}_{\text{obs}}),\ q_{i_y} = \sum_{i_\theta=1}^{n_\theta} \frac{\left|S_{i_y i_\theta}\right|}{\sum_{j_y=1}^{n_y} \left|S_{j_y i_\theta}\right|}$
- **combine** with scale normalization and outlier correction via **adaptive weighting** in an **SMC** framework

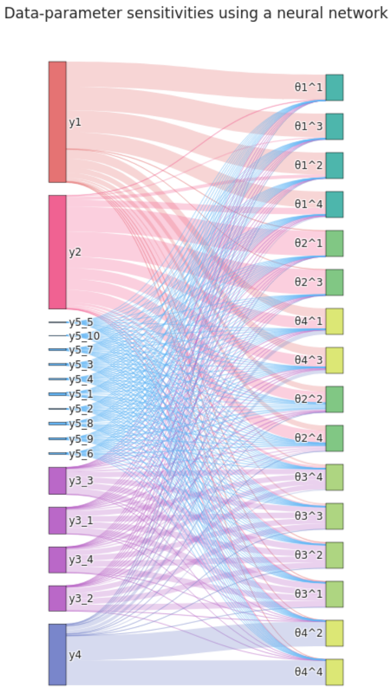
HOW TO ACCOUNT FOR DATA INFORMATIVENESS?



- construct low-dimensional **summary statistics** (see Fearnhead & Prangle, JRSS 2012)
- or: define **sensitivity weights** via the sensitivity matrix $S = \nabla_{\bar{y}} s(\bar{y}_{\text{obs}})$, $q_{i_y} = \sum_{i_\theta=1}^{n_\theta} \frac{\left|S_{i_y i_\theta}\right|}{\sum_{j_y=1}^{n_y} \left|S_{j_y i_\theta}\right|}$
- **combine** with scale normalization and outlier correction via **adaptive weighting** in an **SMC** framework
- learn functions of parameters $\lambda(\theta)$ to capture **higher-order moments**

EVALUATION: SIMPLE TEST MODEL

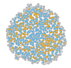


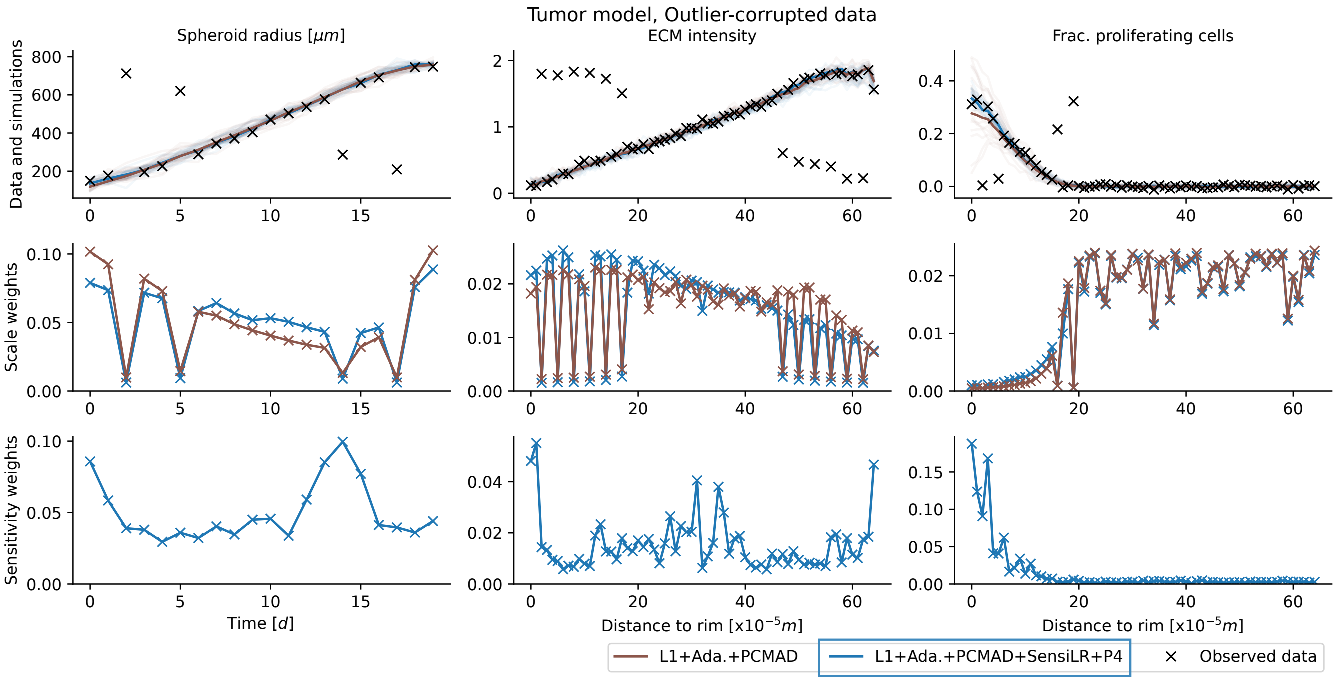


✓ only **combination** of *scale normalization*, *informativeness assessment*, and *regression target augmentation* permits accurate inference

✓ sensitivity weights give further **insights**

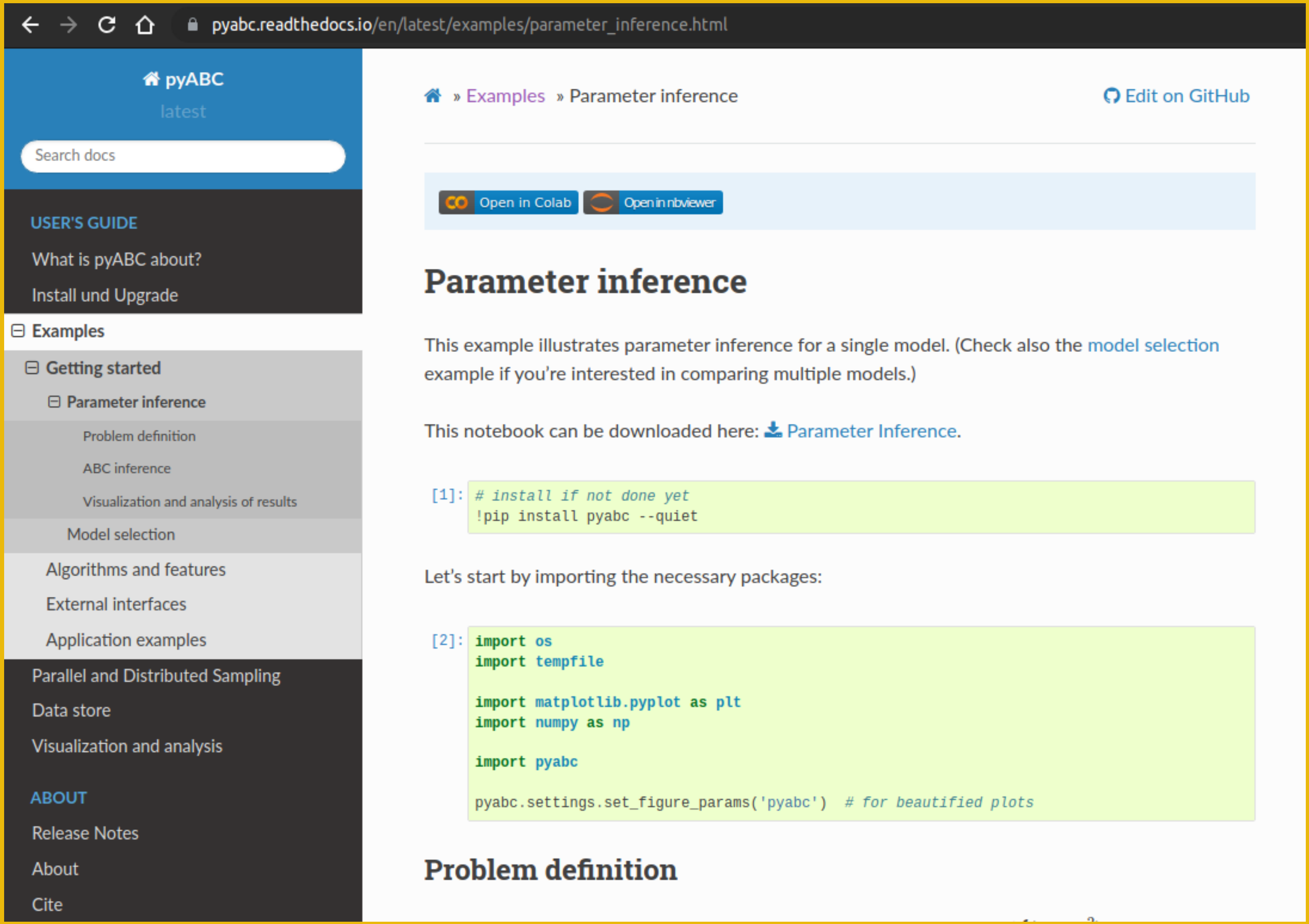
EVALUATION: AGENT-BASED TUMOR SPHEROID MODEL





✓ can via sensitivity weighting in **complex application** simultaneously account for **informativeness** and **outliers**

LIVE DEMO



pyabc.readthedocs.io/en/latest/examples/parameter_inference.html
pyABC
latest
Search docs
USER'S GUIDE
What is pyABC about?
Install und Upgrade
Examples
Getting started
Parameter inference
Problem definition
ABC inference
Visualization and analysis of results
Model selection
Algorithms and features
External interfaces
Application examples
Parallel and Distributed Sampling
Data store
Visualization and analysis
ABOUT
Release Notes
About
Cite
» Examples » Parameter inference
Edit on GitHub
Open in Colab
Open in nbviewer
Parameter inference
This example illustrates parameter inference for a single model. (Check also the model selection example if you're interested in comparing multiple models.)
This notebook can be downloaded here: Parameter Inference.
[1]: # install if not done yet
!pip install pyabc --quiet
Let's start by importing the necessary packages:
[2]: import os
import tempfile
import matplotlib.pyplot as plt
import numpy as np
import pyabc
pyabc.settings.set_figure_params('pyabc') # for beautified plots
Problem definition

SUMMARY

SUMMARY

- parameter estimation when we cannot evaluate the likelihood is challenging
- ABC allows for reliable statistical inference
- pyABC provides an easy-to-use framework

Not everything is a nail.

THANKS! QUESTIONS?