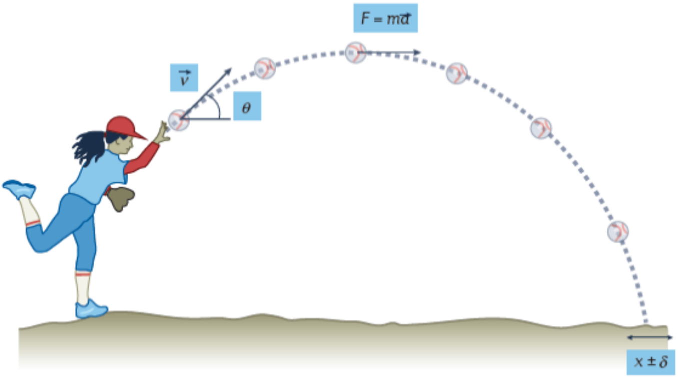
Simulation-Based Inference

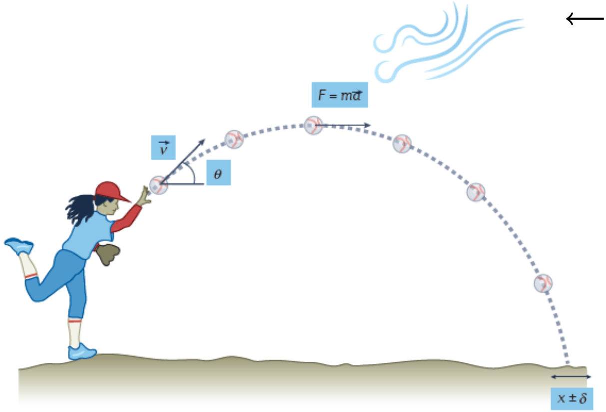
Arnaud Delaunoy



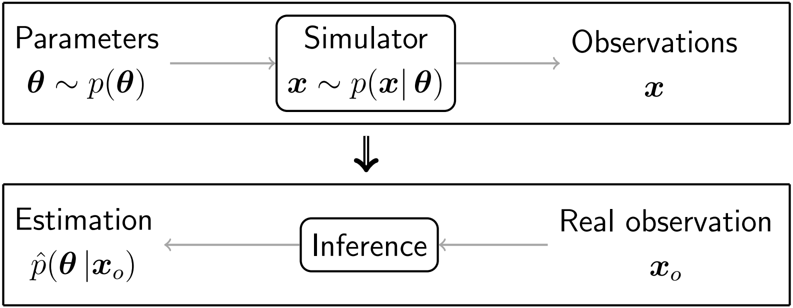
$$X \sim \mathcal{N}(\underbrace{\mu(v,\theta)}_{\text{Newton}}, \quad \underbrace{\Sigma(v,\theta)}_{\text{Measurement error}} \quad)$$

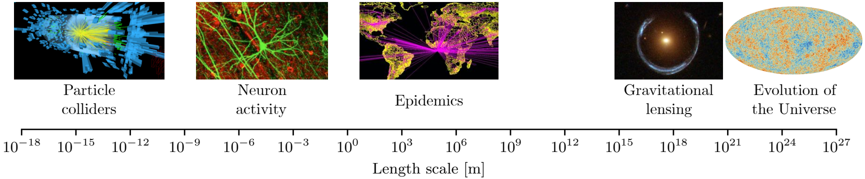
Gal, Yarin, et al. "Bayesian uncertainty quantification for machine-learned models in physics." Nature Reviews Physics 4.9 (2022): 573-577.

⟵ **Intractable likelihood**



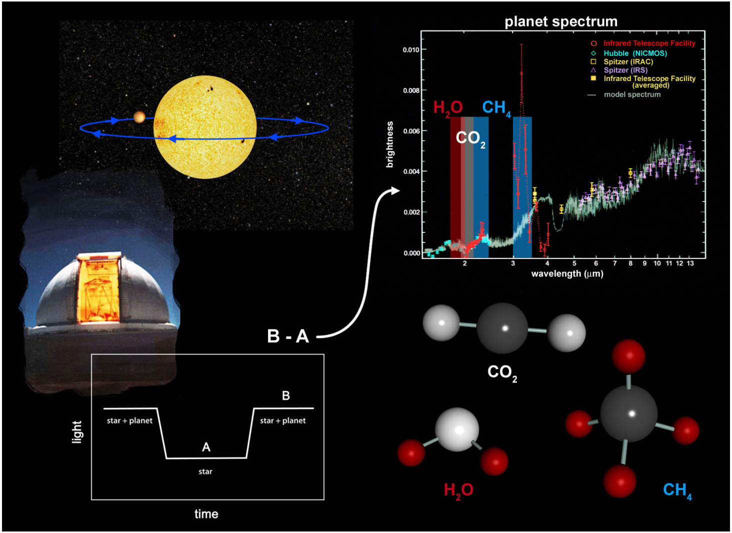
$$\cancel{X \sim \mathcal{N}(\mu(v,\theta), \Sigma(v,\theta))} \quad \rightarrow \quad \begin{aligned} \vec{v}_{t+1} &= \vec{v}_t + \text{Newton}(\vec{v}_t) + \underbrace{\mathcal{N}(0,\Sigma)}_{\text{Wind}} \\ \vec{x}_{t+1} &= \vec{x}_t + \Delta t \; \vec{v}_{t+1} \end{aligned}$$


Parameters
$\boldsymbol{\theta} \sim p(\boldsymbol{\theta})$
Simulator
$\boldsymbol{x} \sim p(\boldsymbol{x}|\,\boldsymbol{\theta})$
Observations
$\boldsymbol{x}$
Estimation
$\hat{p}(\boldsymbol{\theta}\,|\boldsymbol{x}_o)$
Inference
Real observation
$\boldsymbol{x}_o$



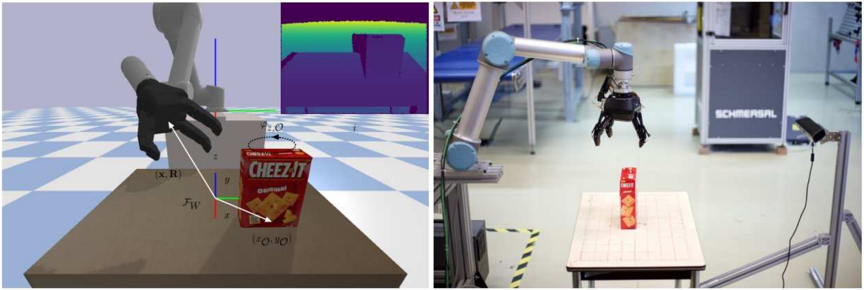
Cranmer, Kyle, Johann Brehmer, and Gilles Louppe. "The frontier of simulation-based inference." Proceedings of the National Academy of Sciences 117.48 (2020): 30055-30062.

Exoplanet characterization



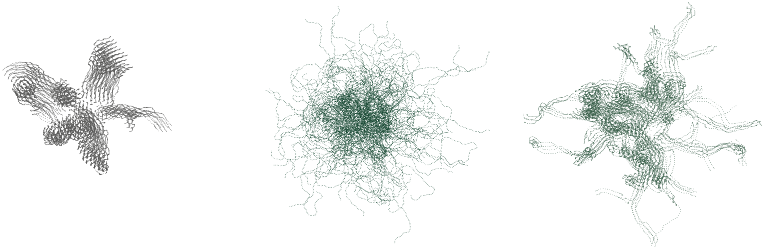
Vasist, Malavika, et al. "Neural posterior estimation for exoplanetary atmospheric retrieval." (2023)

Robotics



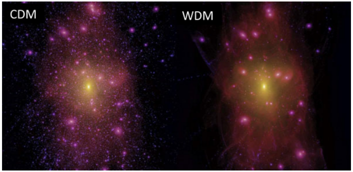
Marlier, Norman, Olivier Brüls, and Gilles Louppe. "Simulation-based Bayesian inference for robotic grasping." (2023)

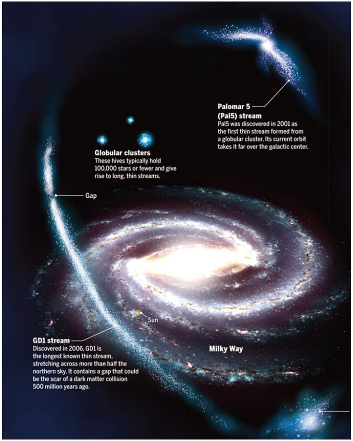
Cell migration



Stillman, Namid R., et al. "Graph-informed simulation-based inference for models of active matter." (2023)

Dark matter





Hermans, Joeri, et al. "Towards constraining warm dark matter with stellar streams through neural simulation-based inference." Monthly Notices of the Royal Astronomical Society 507.2 (2021): 1999-2011.

Approximate Bayesian Computation (ABC)

1. Draw proposal parameters $\boldsymbol{\theta}_i$ from the prior $p(\boldsymbol{\theta})$.
2. Simulate synthetic observations for each parameter $\boldsymbol{x}_i \sim p(\boldsymbol{x}|\,\boldsymbol{\theta} = \boldsymbol{\theta}_i)$.
3. Accept parameters $\boldsymbol{\theta}_i$ if $d(\boldsymbol{x}_i, \boldsymbol{x}_o) < \epsilon$ for some distance d.
4. For $\epsilon \to 0$, accepted parameters are sampled from the posterior $p(\boldsymbol{\theta}\,|\boldsymbol{x}_o)$.

Neural methods

$$p(\boldsymbol{\theta} \,|\boldsymbol{x}) = \frac{p(\boldsymbol{x}|\,\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\boldsymbol{x})}$$

- Neural Likelihood Estimation (NLE): estimate $p(\boldsymbol{x}|\,\boldsymbol{\theta})$
- Neural Posterior Estimation (NPE): estimate $p(\boldsymbol{\theta} \,|\boldsymbol{x})$
- Neural Ratio Estimation (NRE): estimate $\frac{p(\boldsymbol{x}|\,\boldsymbol{\theta})}{p(\boldsymbol{x})}$

Neural Likelihood Estimation (NLE)

- Sample a dataset $(\boldsymbol{\theta}, \boldsymbol{x}) \sim p(\boldsymbol{x}|\,\boldsymbol{\theta})p(\boldsymbol{\theta})$
- Train a neural density estimator on $\boldsymbol{x}$ conditioned on $\boldsymbol{\theta}$.
- At the optimum the density estimator models $p(\boldsymbol{x}|\,\boldsymbol{\theta})$
- Use Markov Chain Monte-Carlo or Variational Inference to approximate $p(\boldsymbol{\theta}\,|\boldsymbol{x} = \boldsymbol{x}_o)$.

Neural density estimators are typically normalizing flows.

Neural Posterior Estimation (NPE)

- Sample a dataset $(\boldsymbol{\theta}, \boldsymbol{x}) \sim p(\boldsymbol{x}|\,\boldsymbol{\theta})p(\boldsymbol{\theta})$
- Train a neural density estimator on $\boldsymbol{\theta}$ conditioned on $\boldsymbol{x}$.
- At the optimum the density estimator models $p(\boldsymbol{\theta}\,|\boldsymbol{x})$

Neural Ratio Estimation (NRE)

Sample dataset and train a classifier

$(\boldsymbol{x}, \boldsymbol{\theta}) \sim p(\boldsymbol{x}, \boldsymbol{\theta})$	$y = 1$
$(\boldsymbol{x}, \boldsymbol{\theta}) \sim p(\boldsymbol{x})p(\boldsymbol{\theta})$	$y = 0$

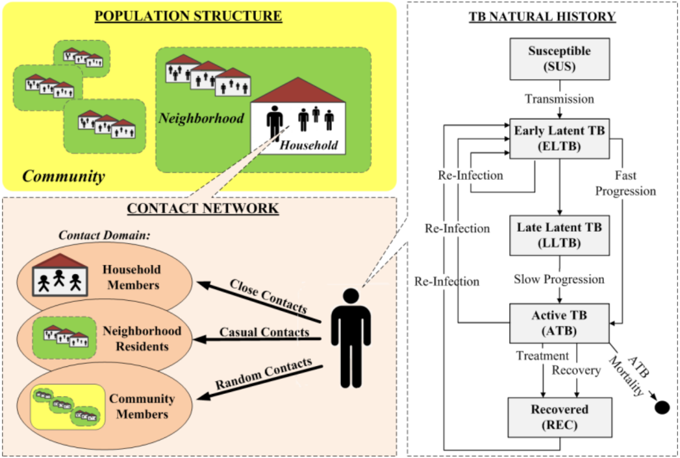
The Bayes optimal classifier d^* can be expressed

$$d^*(\boldsymbol{\theta}, \boldsymbol{x}) = \frac{p(\boldsymbol{\theta}, \boldsymbol{x})}{p(\boldsymbol{\theta}, \boldsymbol{x}) + p(\boldsymbol{\theta})p(\boldsymbol{x})} \Leftrightarrow \frac{p(\boldsymbol{x}|\, \boldsymbol{\theta})}{p(\boldsymbol{x})} = \frac{d^*(\boldsymbol{\theta}, \boldsymbol{x})}{1 - d^*(\boldsymbol{\theta}, \boldsymbol{x})}.$$

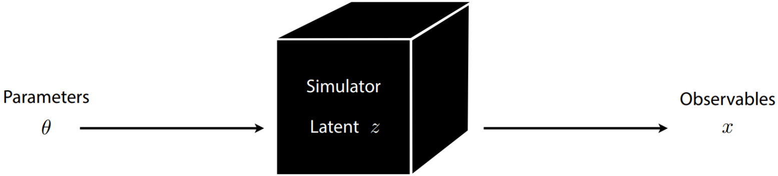
Recover approximate posterior

$$(\boldsymbol{x}, \boldsymbol{\theta}) \longrightarrow \boxed{\text{Classifier}} \rightarrow \hat{d}(\boldsymbol{\theta}, \boldsymbol{x}) \rightarrow \hat{p}(\boldsymbol{\theta}\,|\boldsymbol{x}) = \frac{\hat{d}(\boldsymbol{\theta}, \boldsymbol{x})}{1 - \hat{d}(\boldsymbol{\theta}, \boldsymbol{x})} p(\boldsymbol{\theta})$$

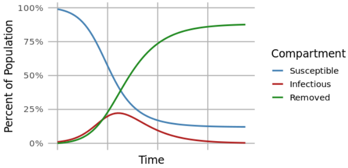
Epidemiology



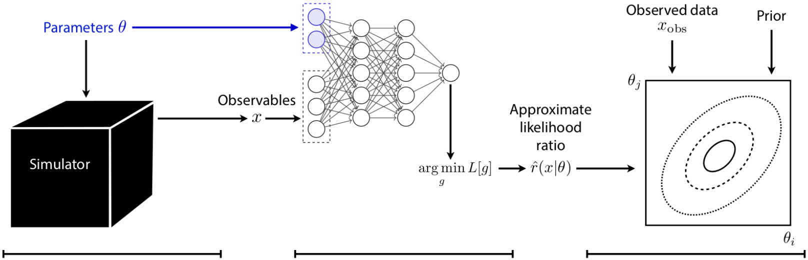
Kasaie, Parastu, David W. Dowdy, and W. David Kelton. "An agent-based simulation of a tuberculosis epidemic: understanding the timing of transmission." 2013 Winter Simulations Conference (WSC). IEEE, 2013.



$$\begin{bmatrix} 0.006 \\ 0.018 \\ 0.34 \\ 0.12 \\ 0.45 \end{bmatrix}$$



Credits: Johann Brehmer
https://www.rgare.com/knowledge-center/article/covid-19-brief-epidemiological-models-explained



1. Simulation

Run simulator and save data

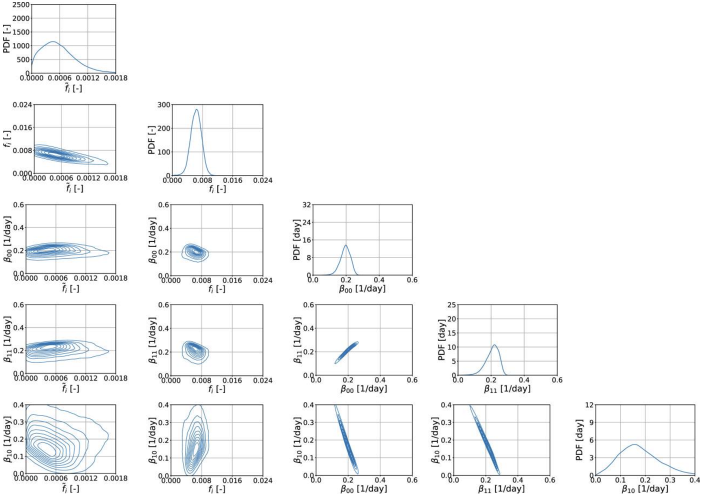
2. Machine Learning

Train NN classifier, interpret as likelihood ratio estimator

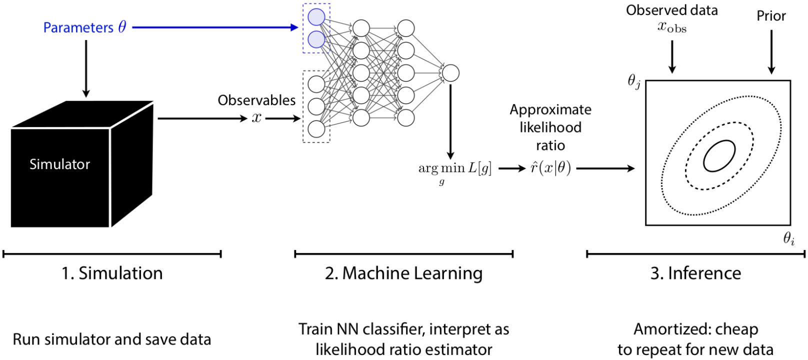
3. Inference

Amortized: cheap to repeat for new data

Credits: Johann Brehmer



Arnst, Maarten, et al. "A hybrid stochastic model and its Bayesian identification for infectious disease screening in a university campus with application to massive COVID-19 screening at the University of Liège." Mathematical Biosciences 347 (2022): 108805.



Credits: Johann Brehmer