Automating Science with Artificial Intelligence

Gilles Louppe, Faculty of Applied Sciences
g.louppe@uliege.be @glouppe

Joint work with Joeri Hermans (ULiège), Kyle Cranmer (NYU), Lukas Heinrich (NYU), David Rousseau (LAL)

Looking for **new** PhD students! (Computer Science, Engineering or Physical Sciences)

The key to Science

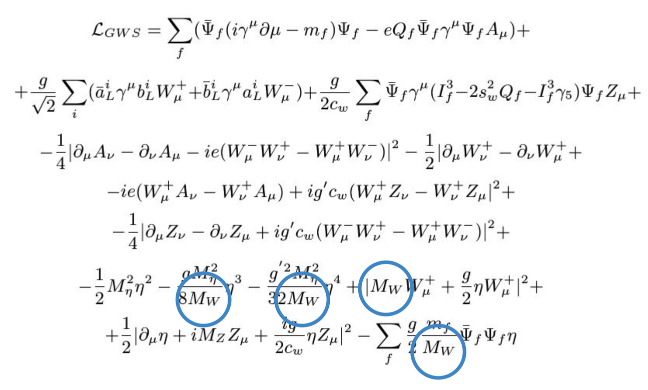
Hypothesis(φ) → Experiment(ϕ) → Conclusions(θ)

How do we look for new scientific laws?

1. Make a guess
2. Compute the consequences of the guess
3. Compare the results to Nature

Long-term research goal: automate the Scientific method through modern Artificial Intelligence and Deep Learning methods!

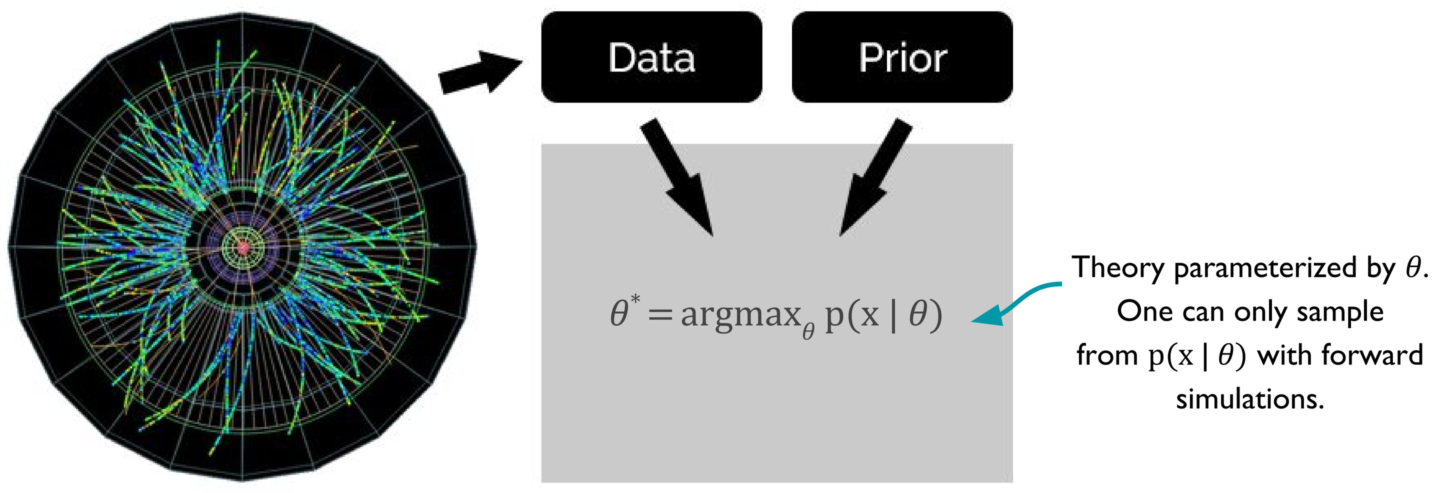
A guess is a parameterized theory



Standard Model of Particle Physics

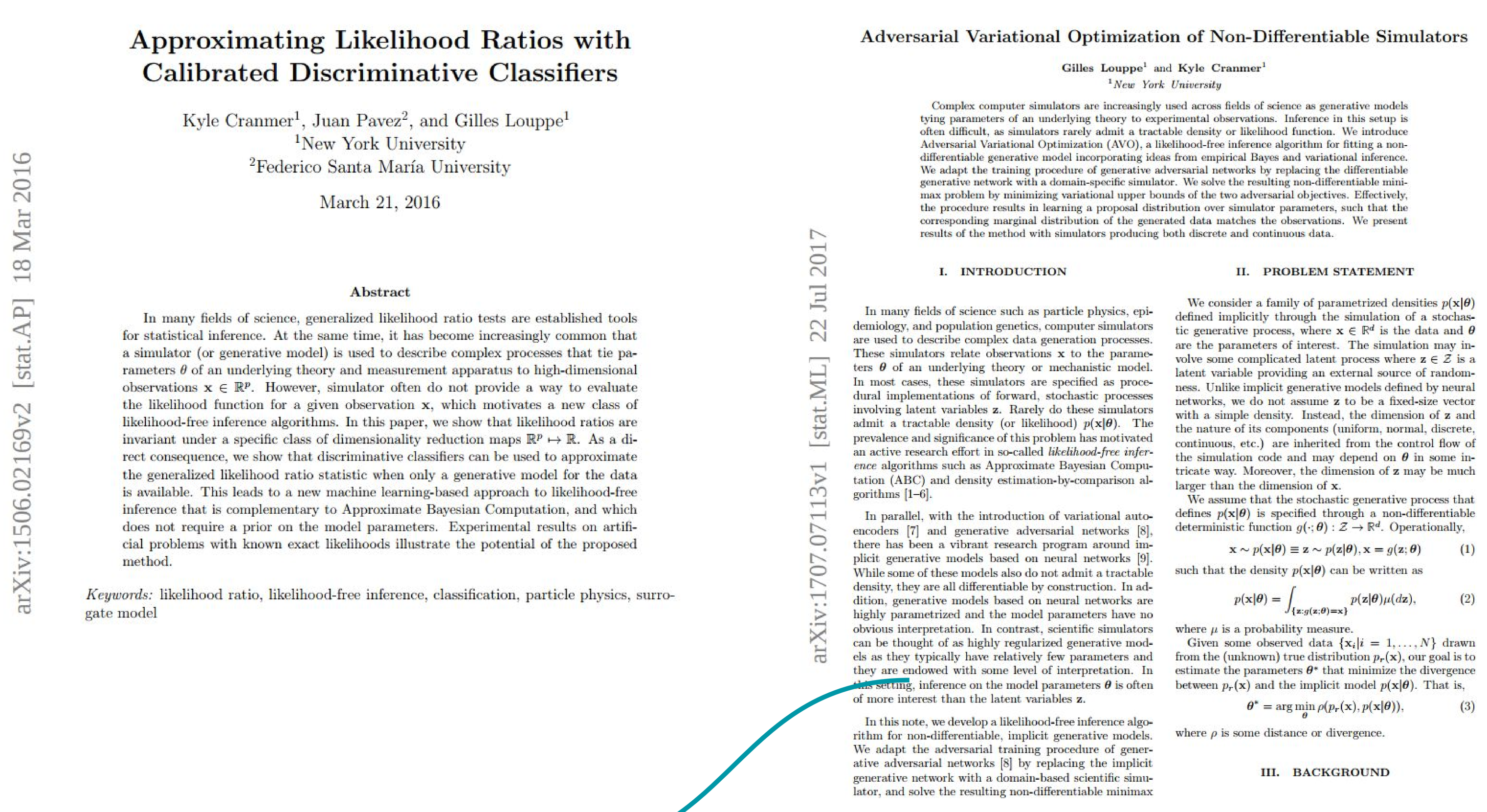
I. Likelihood-free inference

To validate a guess and compare it to Nature, we need to compute its consequences. The expected consequences are usually modelled through a computer simulator.



Issue: computation of $p(x \mid \theta)$ is in most cases intractable.

Solutions?

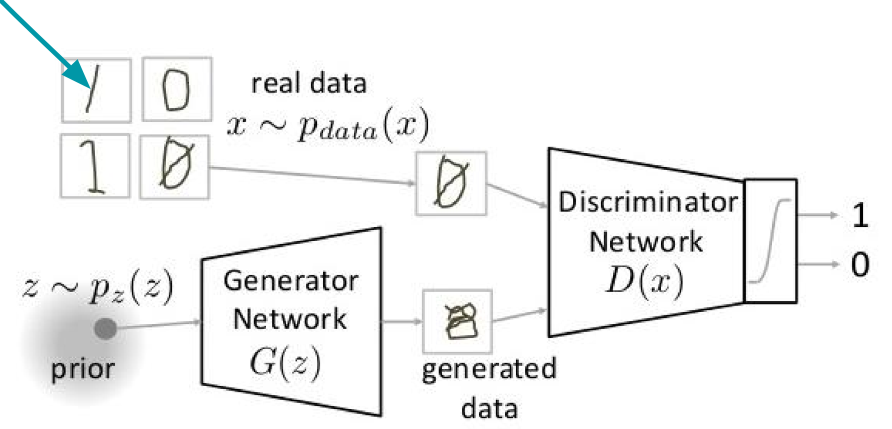


Adversarial Variational Optimization (Louppe, Cranmer, 2017):

- Same as generative adversarial networks (see below).
- Replace generator network with domain simulator.
- Cope with non-differentiability using variational optimization.
- Require state-of-the-art classifiers for optimal performance.

Deep learning for physical sciences strongly needed here!

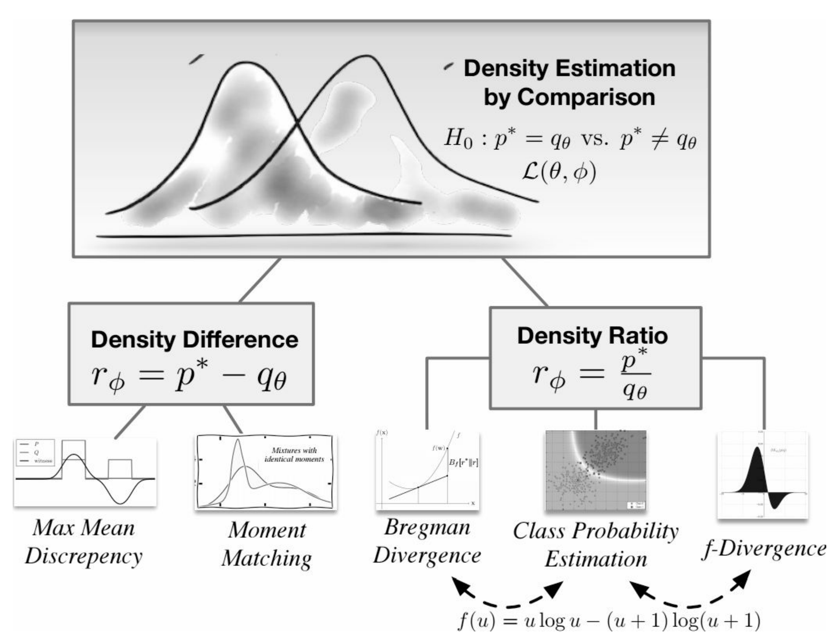
Generative adversarial networks:



Which of those are real faces?

Implicit models: Likehood-free inference relates to a large body of statistical methods developed within the machine learning community:

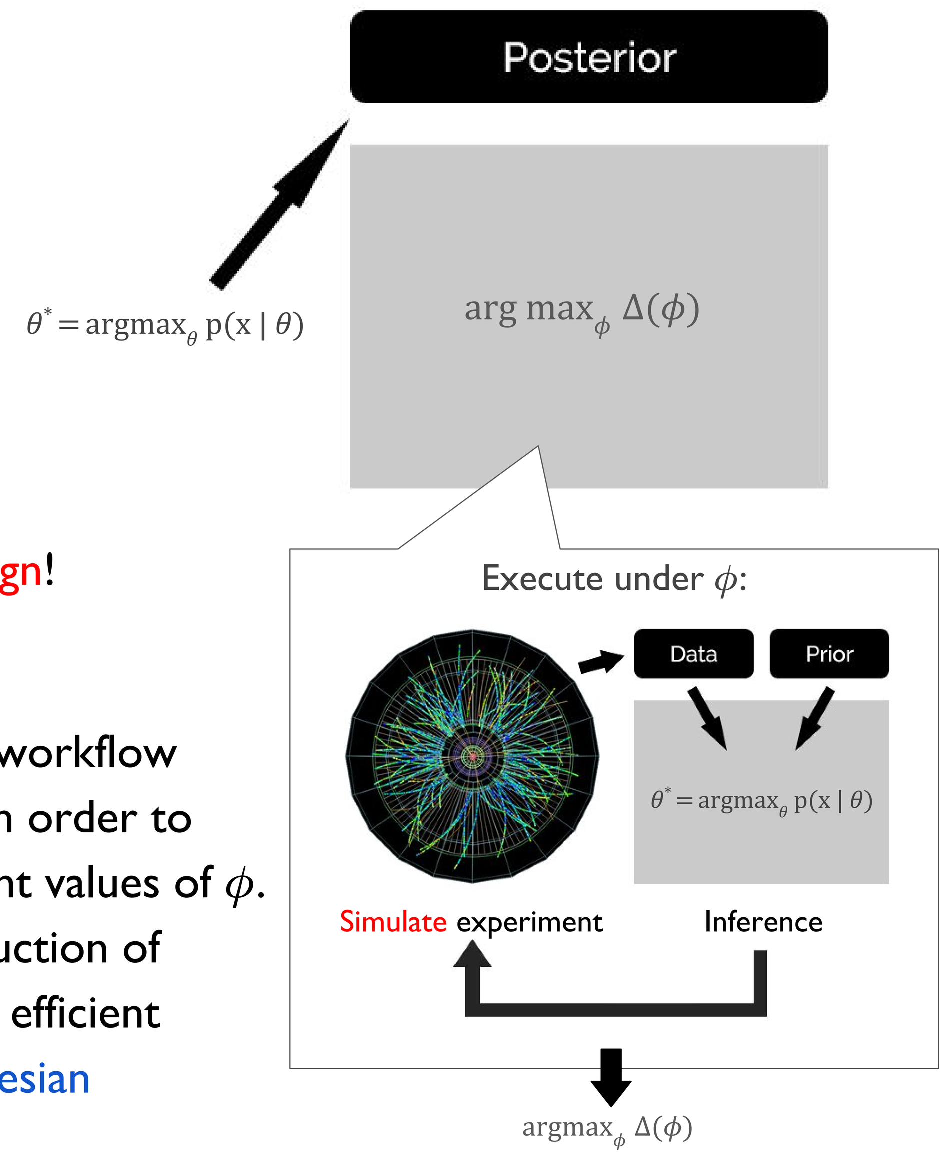
- Approximate Bayesian Computations
- Density estimation-by-comparison
- Varitional inference



II. Optimal experimental design

How do we design the best scientific experiment?

If an experiment is configured by ϕ, then we want to find the experiment that maximizes the reduction in entropy $\Delta(\phi)$, therefore obtaining the least amount of uncertainty $\mathrm{H}[\theta]$.



Solution: optimal experimental design!

Challenges:

- infrastructure: full reproducible workflow needs to be run multiple times in order to make the estimations for different values of ϕ.
- optimization: estimating the reduction of uncertainty is slow. We need an efficient optimization algorithm (e.g., Bayesian optimization).

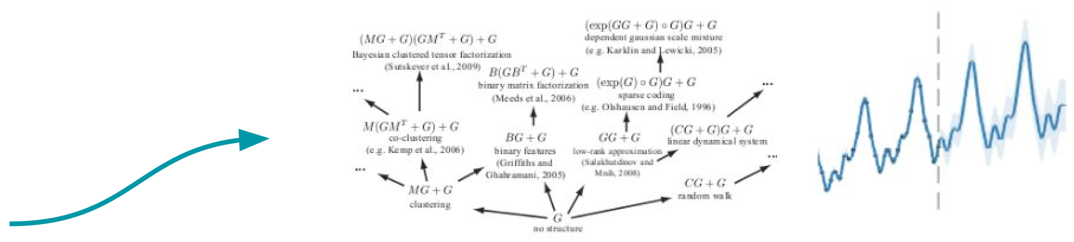
III. Exploring the theory space

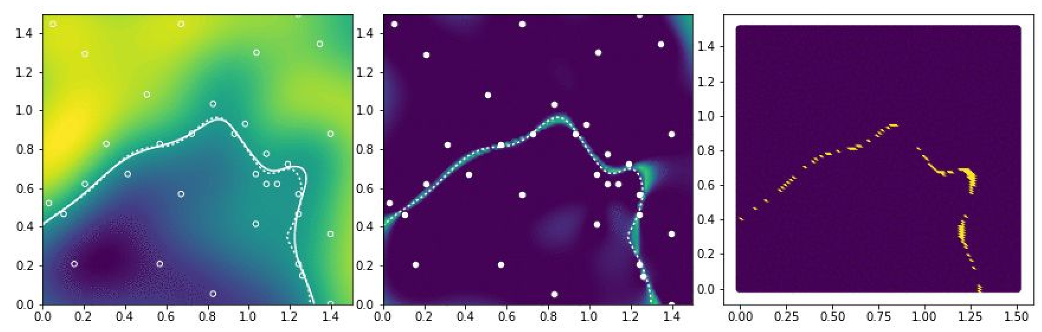
Can we automatically explore the space of theories and find the envelope that agrees with the data?

Solutions?

- Symbolic exploration
- Finding exclusion contours, while minimizing the number of calls (expensive):

$$\{\psi \,|\, \rho(p_r(\mathbf{x}|\phi), p(\mathbf{x}|\psi, \phi, \theta^*)) < \epsilon\}$$

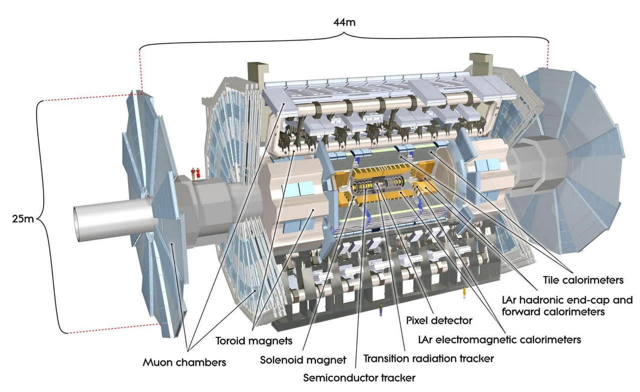




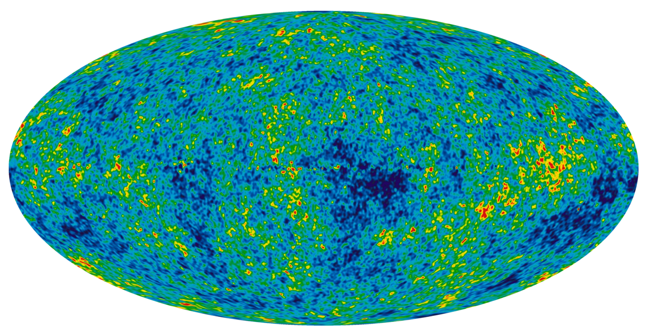
IV. Use cases (ongoing work)

Particle physics (collaborations with CERN and Oxford)

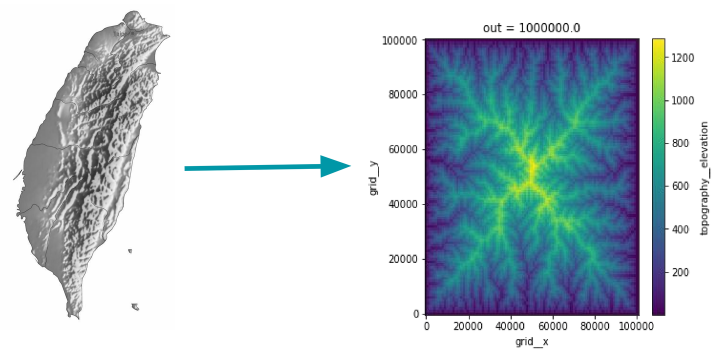
- Exploration of Beyond the Standard Model theories
- Automated tuning of particle detectors



Cosmology (collaboration with astrophysicists)



Geomorphology (collaboration with Earth scientists)



V. Deep learning for physical sciences

This interdisciplinary research usually requires domain-specific deep learning models, which results in contributions of their own.