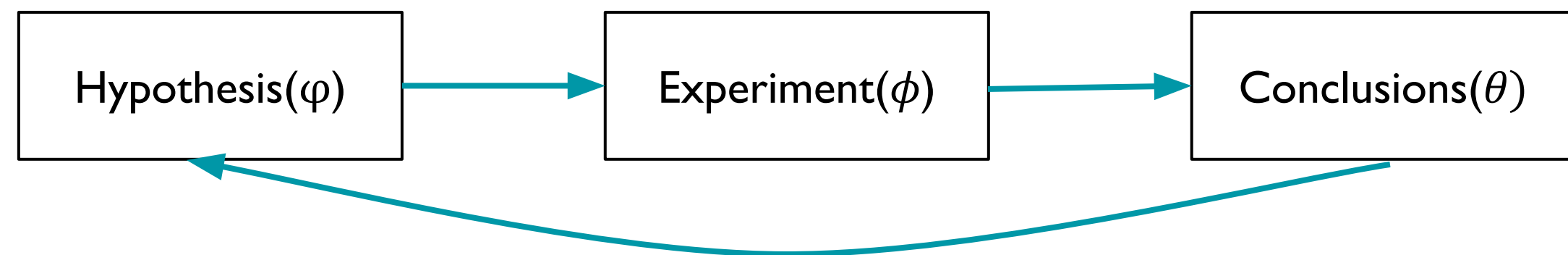


The key to Science



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!

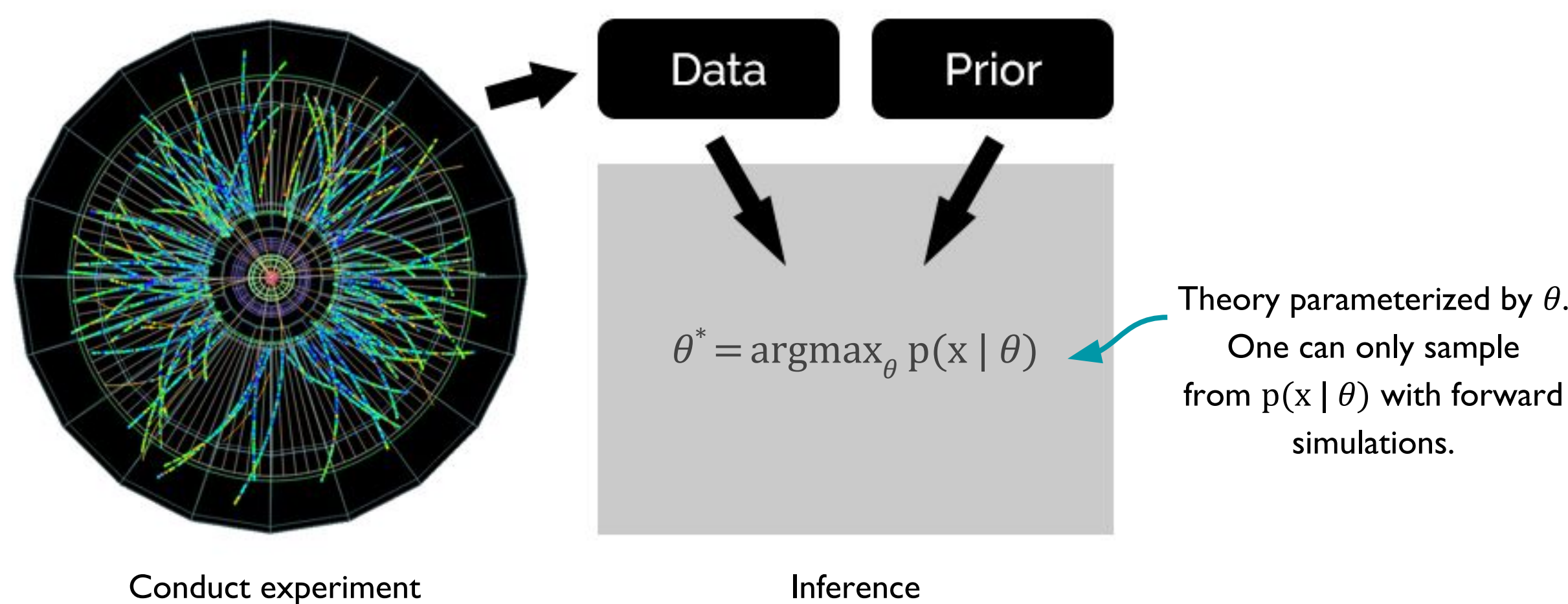
A **guess** is a parameterized theory

$$\begin{aligned} \mathcal{L}_{\text{CWS}} = & \sum_f (\Psi_f (i\gamma^\mu \partial_\mu - m_f) \Psi_f - e Q_f \bar{\Psi}_f \gamma^\mu \Psi_f A_\mu) + \\ & + \frac{g}{\sqrt{2}} \sum_l (\bar{a}_L \gamma^\mu b_L W_\mu^+ + \bar{b}_L \gamma^\mu a_L W_\mu^-) + \frac{g}{2c_w} \sum_f \bar{\Psi}_f \gamma^\mu (I_f^3 - 2s_w^2 Q_f - I_f^3 \gamma_5) \Psi_f Z_\mu + \\ & - \frac{1}{4} [\partial_\mu A_\nu - \partial_\nu A_\mu - ie(W_\mu^+ W_\nu^- - W_\mu^- W_\nu^+) - \frac{1}{2} \partial_\mu W_\nu^+ - \partial_\nu W_\mu^+ + \\ & - ie(W_\mu^+ A_\nu - W_\nu^+ A_\mu) + ig' c_w (W_\mu^+ W_\nu^- - W_\mu^- W_\nu^+)^2 + \\ & - \frac{1}{4} [\partial_\mu Z_\nu - \partial_\nu Z_\mu + ig' c_w (W_\mu^+ W_\nu^+ - W_\mu^- W_\nu^-)^2] + \\ & - \frac{1}{2} M_W^2 \eta^2 + \frac{M_W^2}{2M_W} - \frac{g^2 M_W^2}{2M_W} + (M_W \Psi_\mu^+ + \frac{g}{2} \eta W_\mu^+)^2 + \\ & + \frac{1}{2} [\partial_\mu \eta + i M_Z Z_\mu + \frac{g}{2c_w} \eta \gamma_5]^2 - \sum_f \frac{g}{2} \bar{\Psi}_f \gamma^\mu \Psi_f \eta \end{aligned}$$

Standard Model of Particle Physics

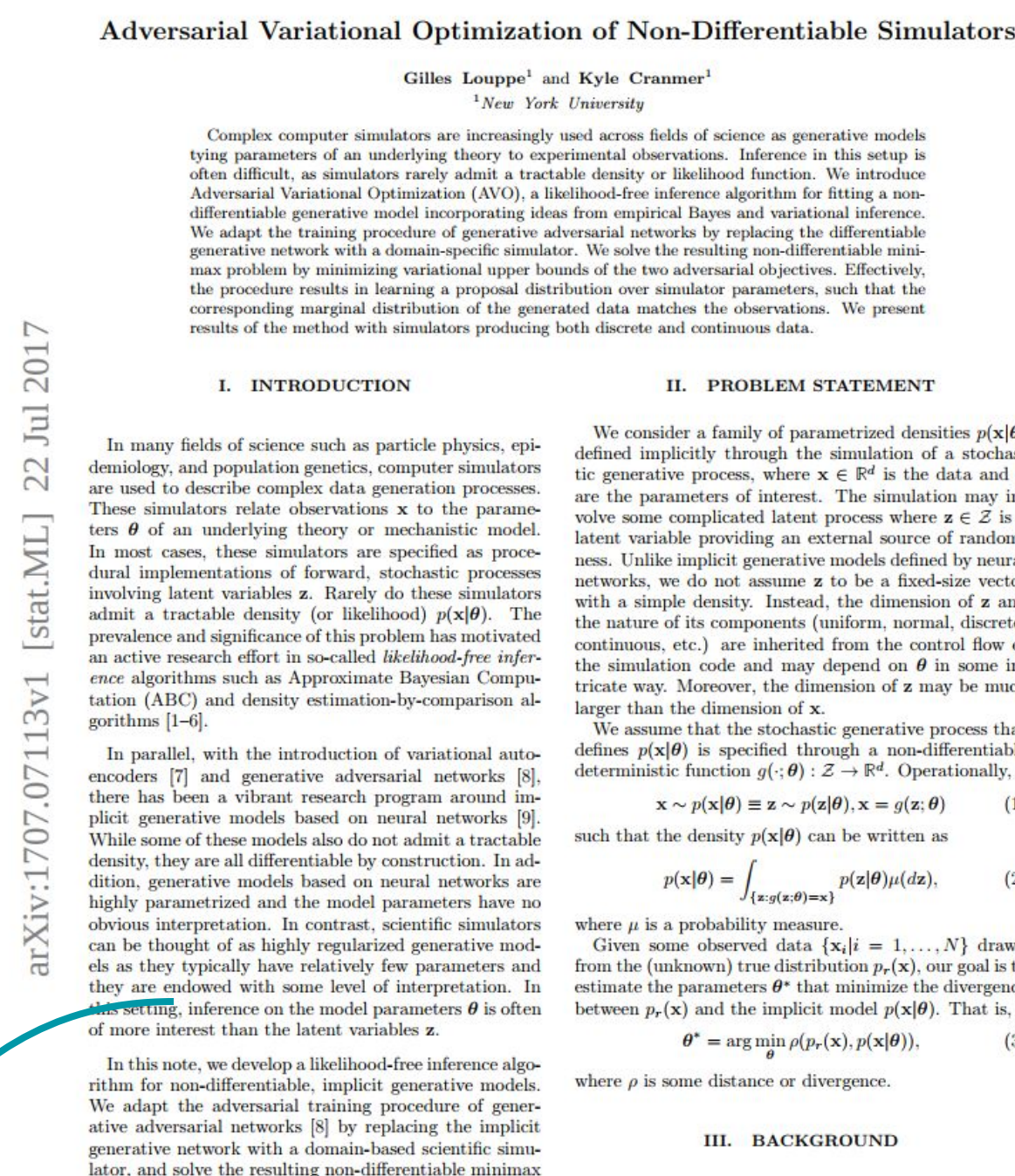
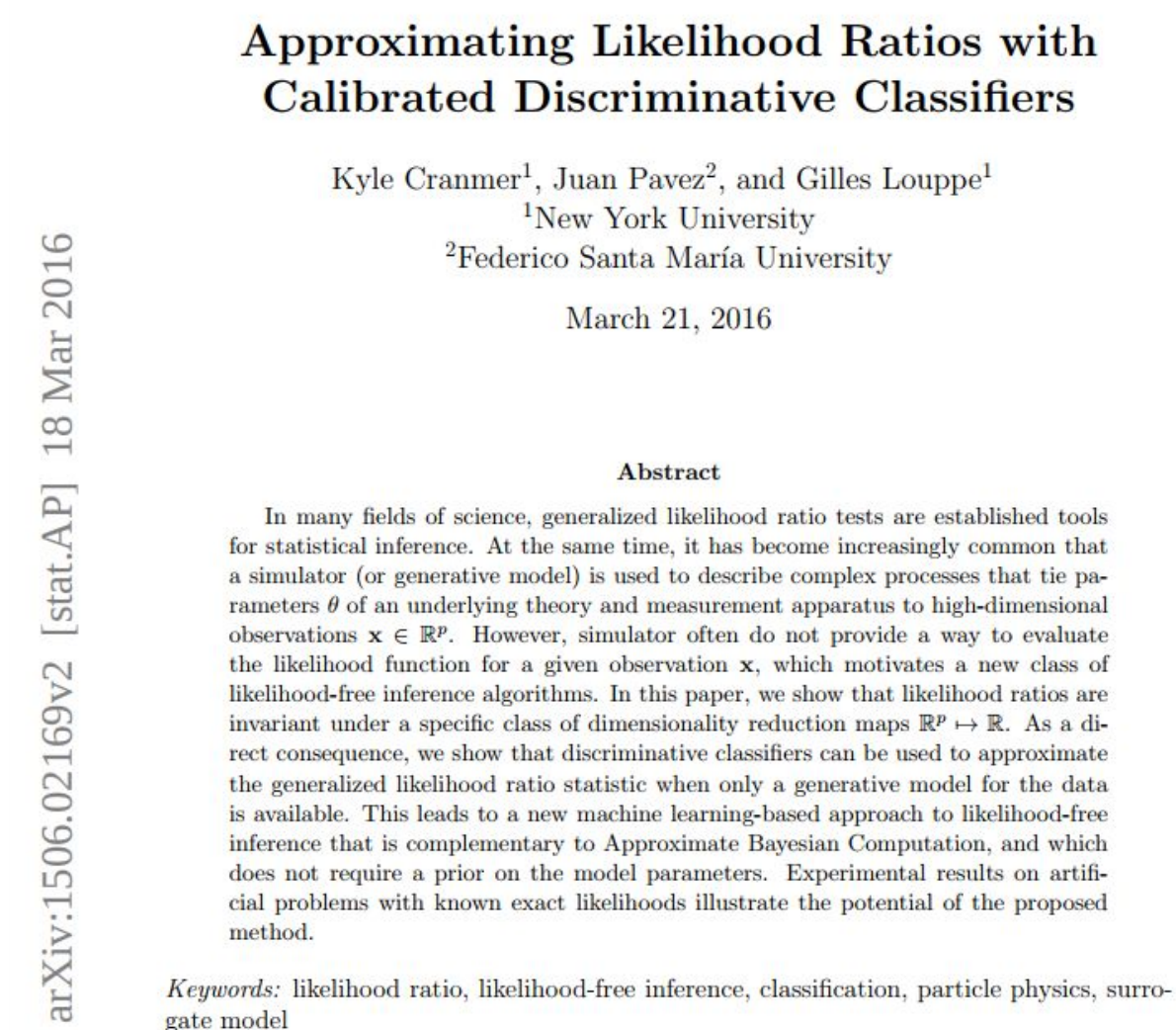
I. Likelihood-free inference

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 $p(x | \theta)$ is in most cases **intractable**.

Solutions?

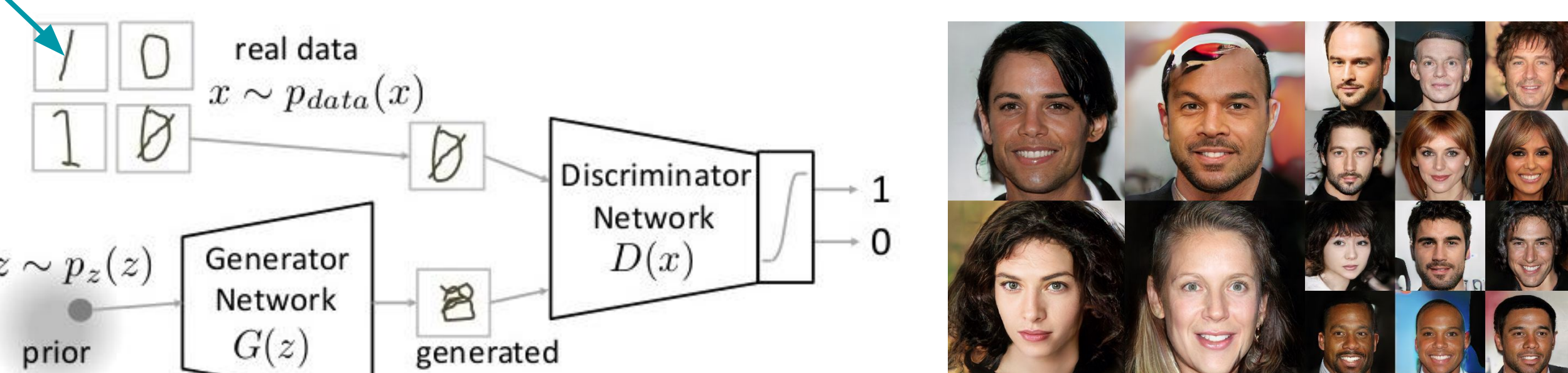


Adversarial Variational Optimization (Louppe, Cranmer, 2017):

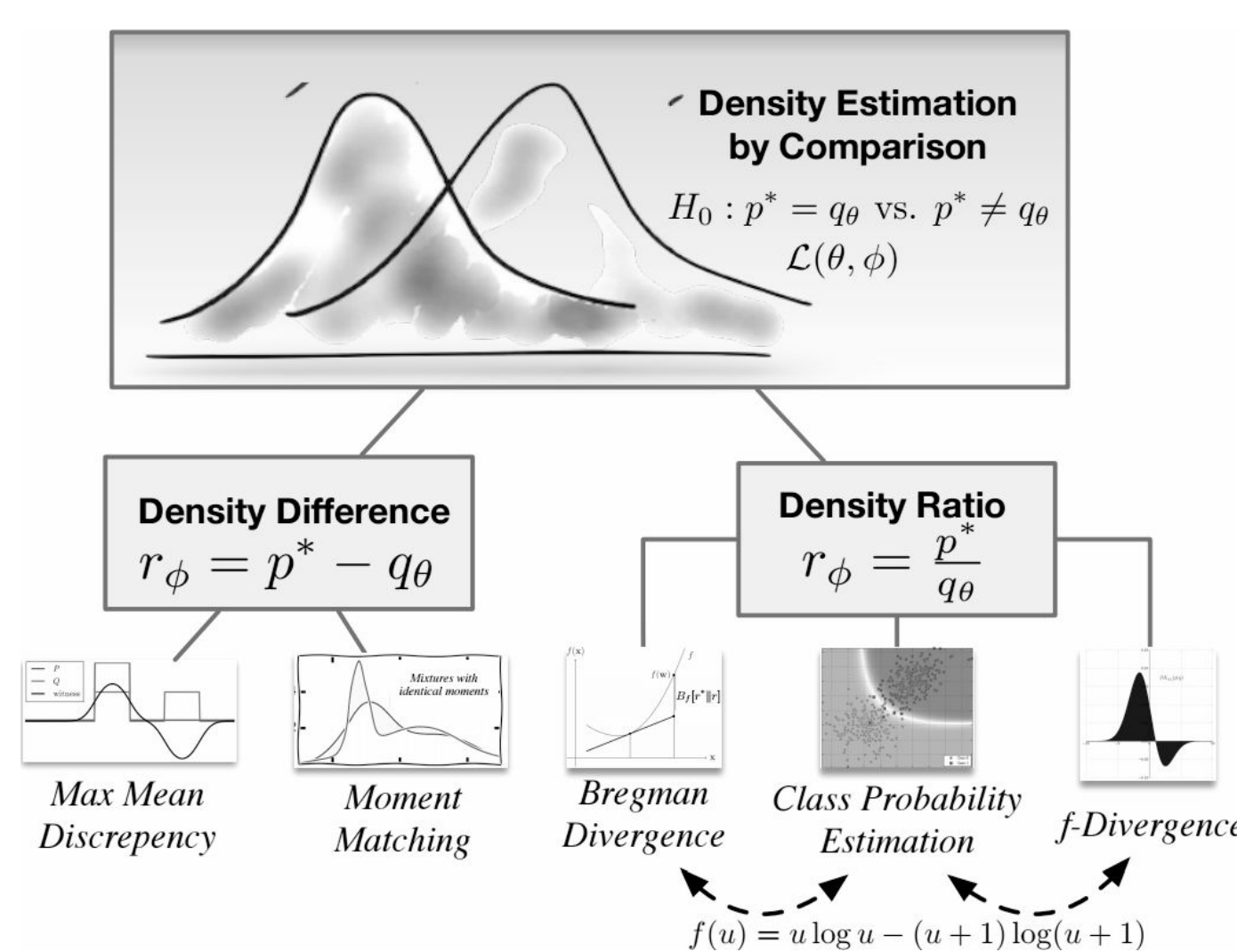
- Same as generative adversarial networks (see below).
- Replace generator network with domain simulator.
- Cope with non-differentiability using variational optimization.
- Require state-of-the-art classifiers for optimal performance.

Deep learning for physical sciences strongly needed here!

Generative adversarial networks:



Which of those are real faces?



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

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

II. Optimal experimental design

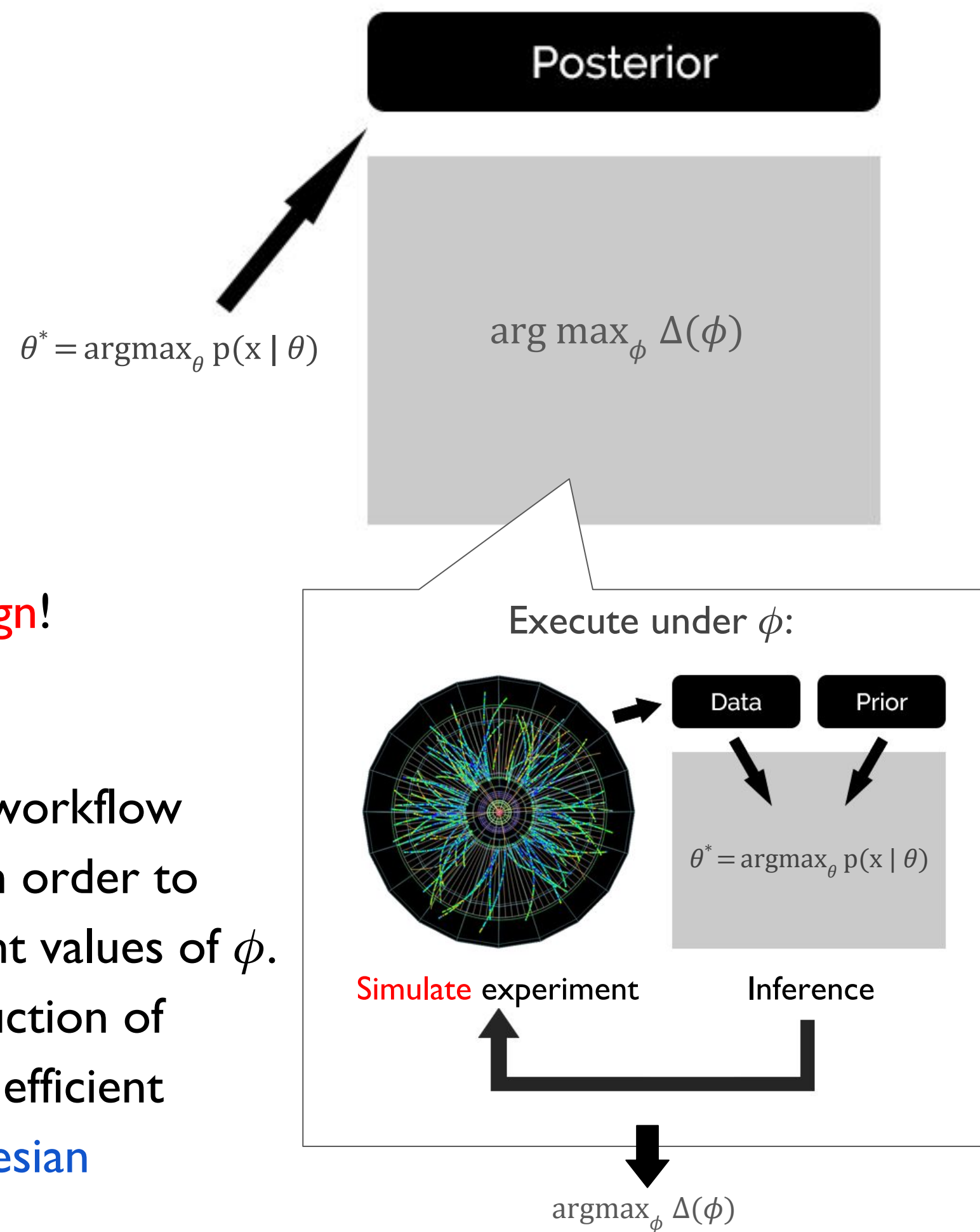
How do we design the best scientific experiment?

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 $H[\theta]$.

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**).

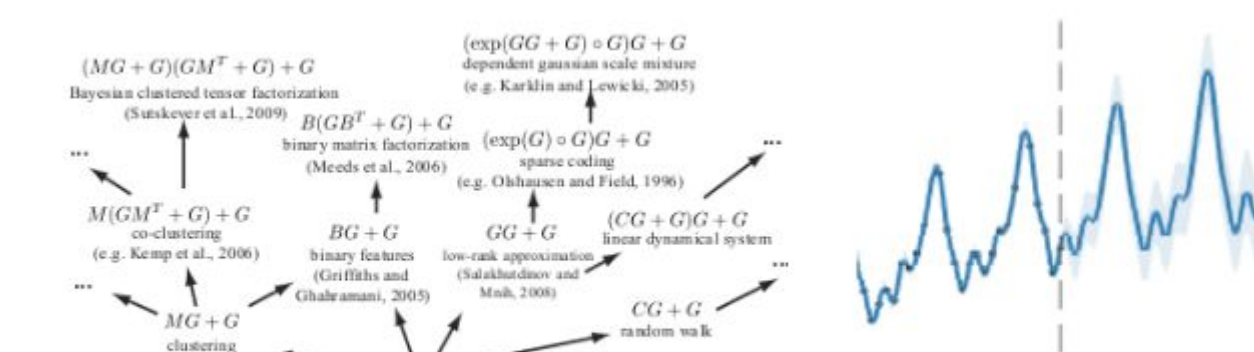


III. Exploring the theory space

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



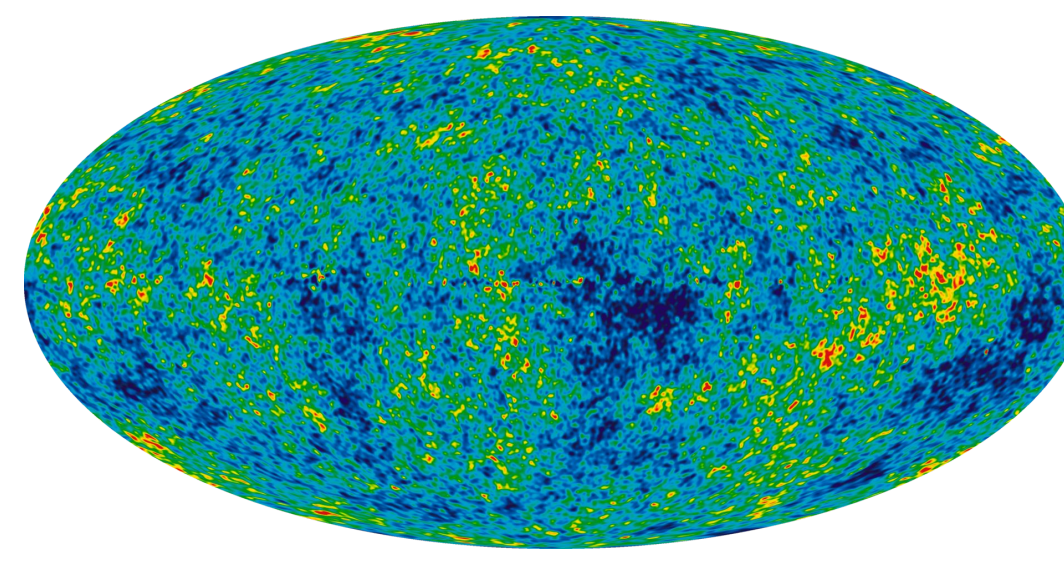
IV. Use cases (ongoing work)

Particle physics (collaborations with CERN and Oxford)

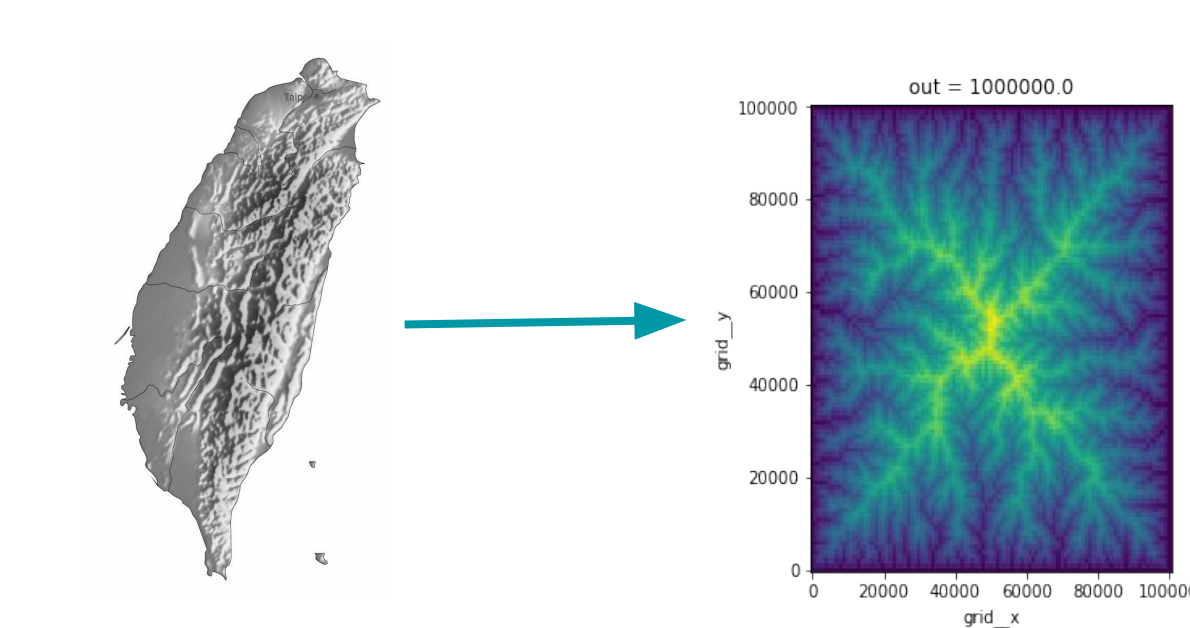
- Exploration of Beyond the Standard Model theories
- Automated tuning of particle detectors



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.

Learning to Pivot with Adversarial Networks	Neural Message Passing for Jet Physics	QCD-Aware Recursive Neural Networks for Jet Physics
Gilles Louppe New York University g.louppe@nyu.edu Michael Kagan SLAC National Accelerator Laboratory mkagan@slac.stanford.edu Kyle Cranmer New York University kyle.cranmer@nyu.edu	Isaac Hninson, Johann Brehmer, Jean Bruna, Kyunghyun Cho, Kyle Cranmer Department of Computer Science New York University New York, NY 10012 {hninson, johann.brehmer, bruna, kyunghyun.cho, kyle.cranmer}@nyu.edu Gilles Louppe Department of Computer Science University of Liège Belgium g.louppe@uliege.ac.be Gasper Rochette Ecole Normale Supérieure Paris, France gaspar.rochette@ens.fr	Gilles Louppe,¹ Kyunghyun Cho,² Cyril Bortol,³ and Kyle Cranmer¹ ¹New York University
Abstract Several techniques for domain adaptation have been proposed to account for differences in the distribution of the data used for training and testing. The majority of this work focuses on binary domain labels. Similar problems occur in a scientific context where there may be a continuous family of plausible data generation processes associated to the presence of systematic uncertainties. Robust inference is possible if it is based on a pivot—a quantity whose distribution does not depend on the unknown values of the nuisance parameters that parameterize the family of data generation processes. In this work, we introduce and derive theoretical results for a training procedure based on adversarial networks for selecting the pivot property (or, equivalently, features with respect to common attributes) on a predictive model. The method includes a hyperparameter to control the trade-off between accuracy and robustness. We demonstrate the effectiveness of this approach with a toy example and examples from particle physics.	Abstract Supervised learning has incredible potential for particle physics, and one application that has received a great deal of attention involves collimated sprays of particles called jets. Recent progress for jet physics has leveraged machine learning techniques based on computer vision and natural language processing. In this work, we consider message passing on a graph where the nodes are the particles in a jet. We design variants of a message-passing neural network (MPNN), (1) with a learnable adjacency matrix, (2) with a learnable symmetric adjacency matrix, and (3) with a self-attention aggregated hidden state and MPNN with a learnable adjacency matrix. We compare these against the previously proposed recursive neural network with a fixed tree structure and show that the MPNN with a learnable adjacency matrix and two message-passing iterations outperforms all the others.	Recent progress in applying machine learning for jet physics has been built upon an analogy between calorimeters and images. In this work, we present a novel class of recursive neural networks built instead upon an analogy between QCD and natural languages. In the analogy, four-momenta are like words, and the jet-based tree structure of sequential reconstruction jet algorithms is like the parsing of a sentence. Our approach works directly with the four-momenta of a variable-length set of particles, and the jet-based tree structure serves as an event-level event basis. Our experiments highlight the flexibility of our method for building task-specific jet embeddings and show that recursive architectures are significantly more accurate and data efficient than previous image-based networks. We extend the analogy from individual jets (sentences) to full events (paragraphs), and show for the first time an event-level classifier operating on all the stable particles produced in an LHC event.