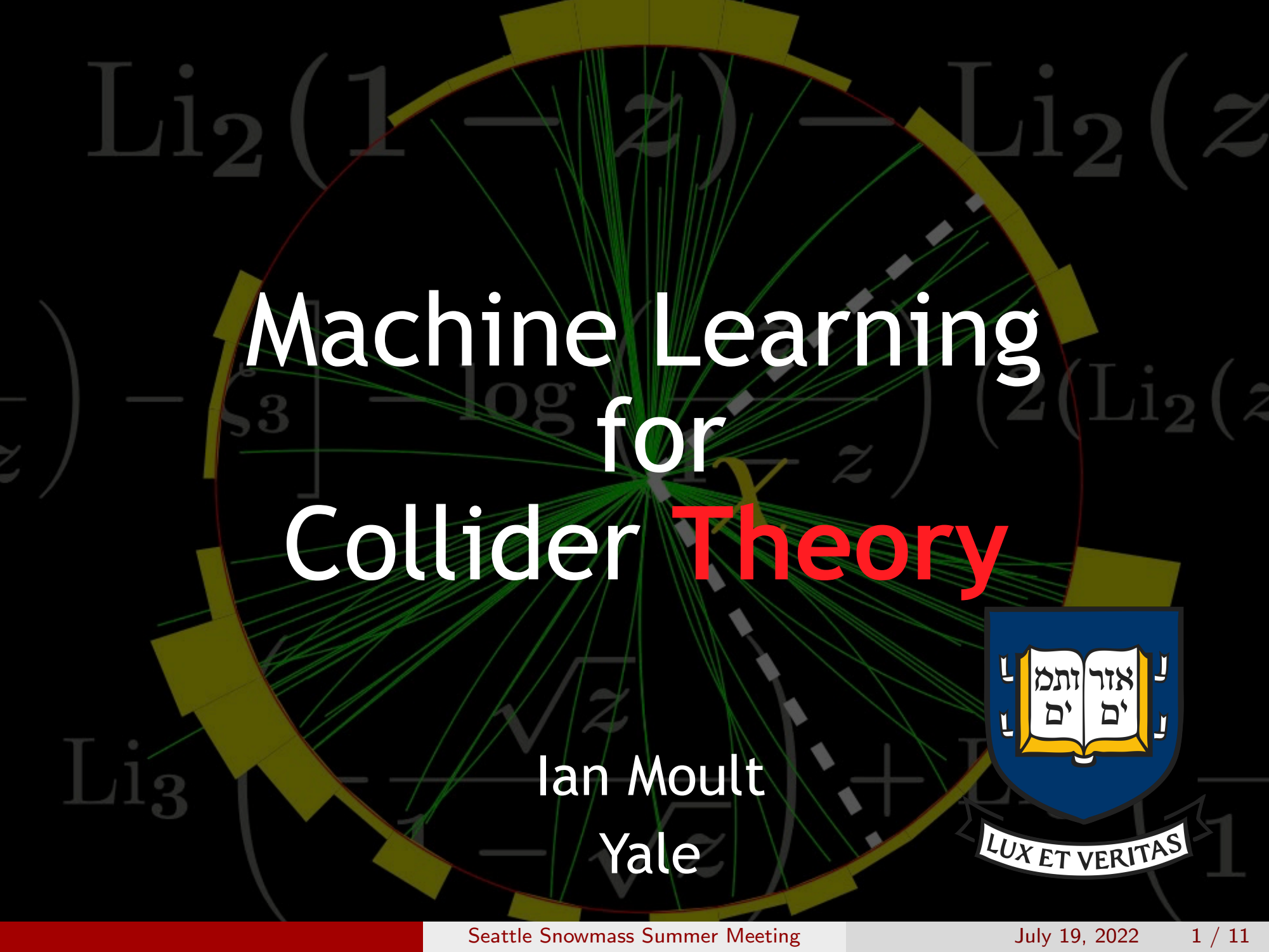


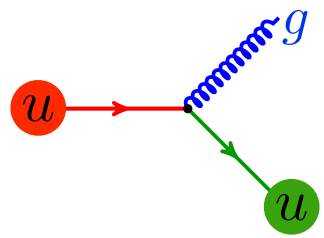
Machine Learning for Collider Theory

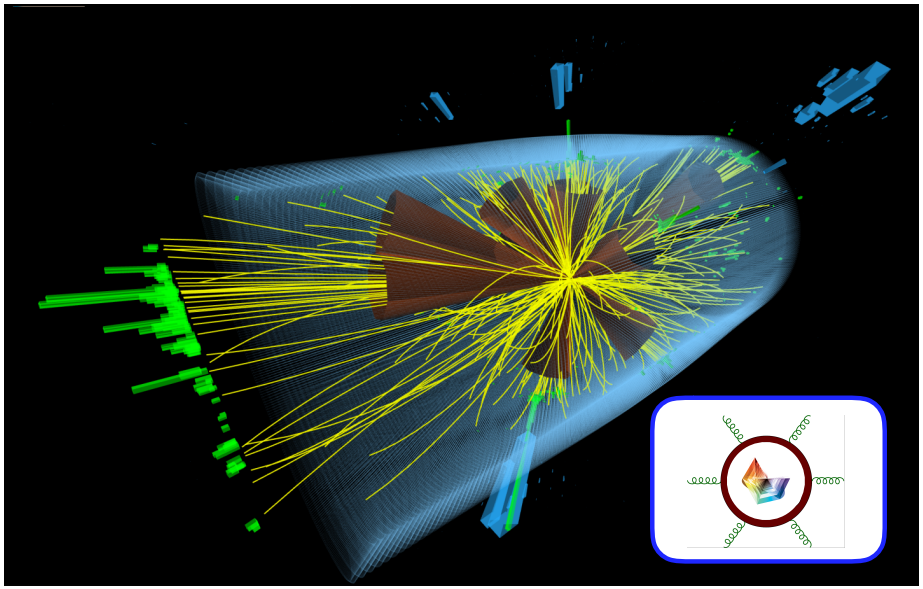
Ian Moult
Yale

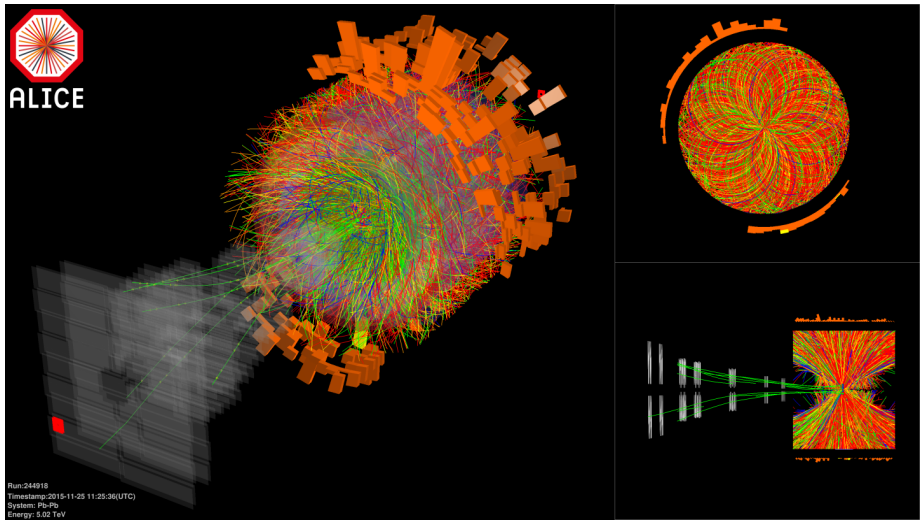
Particle Colliders

- Particle colliders provide one of the most spectacular examples of a simple underlying theory producing remarkably complicated data sets.

$$\mathcal{L}_{\text{QCD}} = -\tfrac{1}{4} G^a_{\mu\nu} G^{\mu\nu a} + \sum_f \bar{q}_f (i \not{D} - m_f) q_f$$



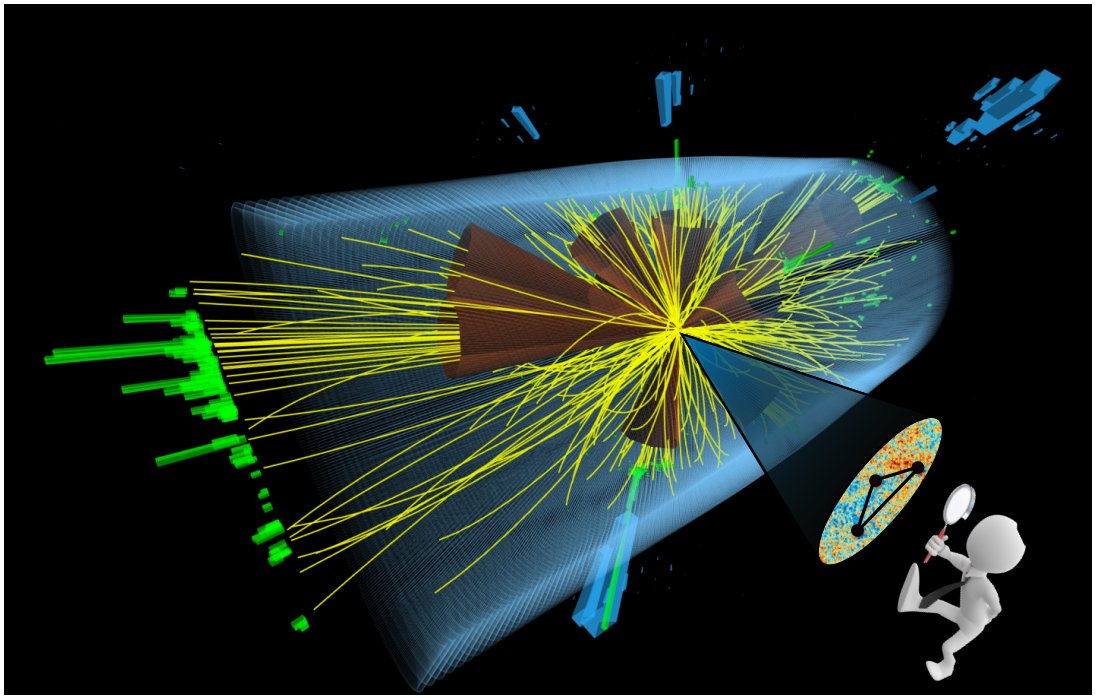




- This has enabled (combined with sustained theory efforts) the development of high fidelity simulations (Pythia, Herwig, Sherpa, ...).

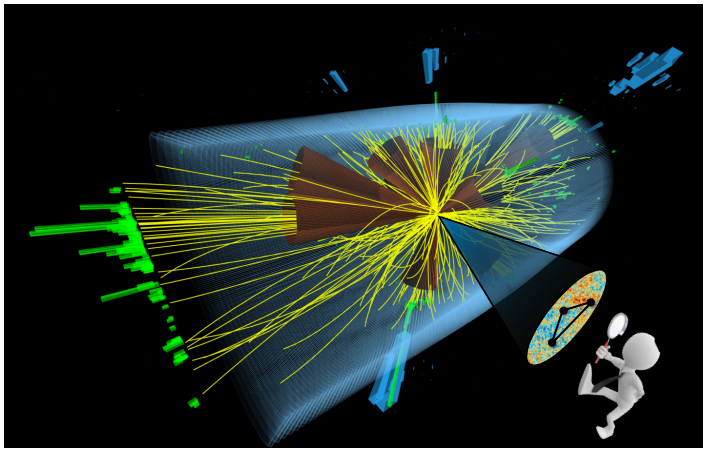
Jet Substructure

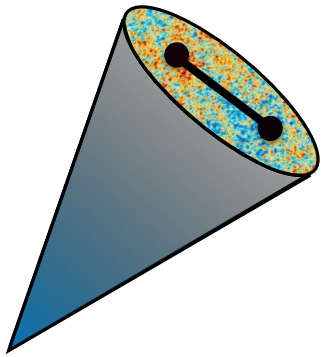
- Many remarkable phenomena of quantum field theory are imprinted in subtle correlations in collider energy flux:
  - Interactions of asymptotically free quarks and gluons.
  - Real time dynamics of confinement.
  - Nature of the QGP.
  - Phase diagram of QCD.
  - ....

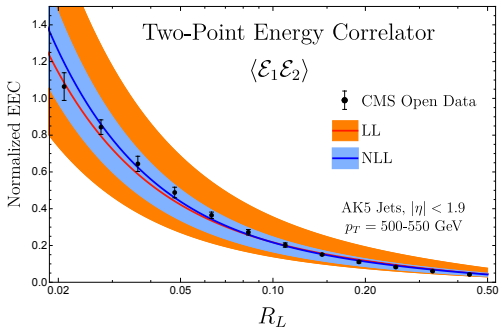


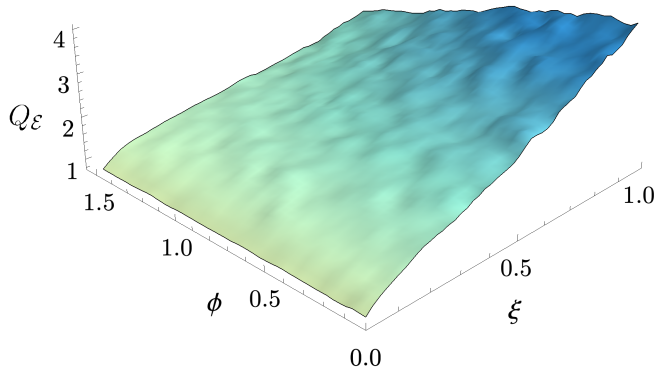
Formal Theory Progress

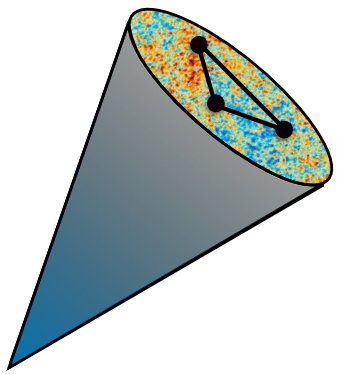
- Formal theory has provided powerful new tools to compute these observables, and relate them to parameters of the underlying theory.





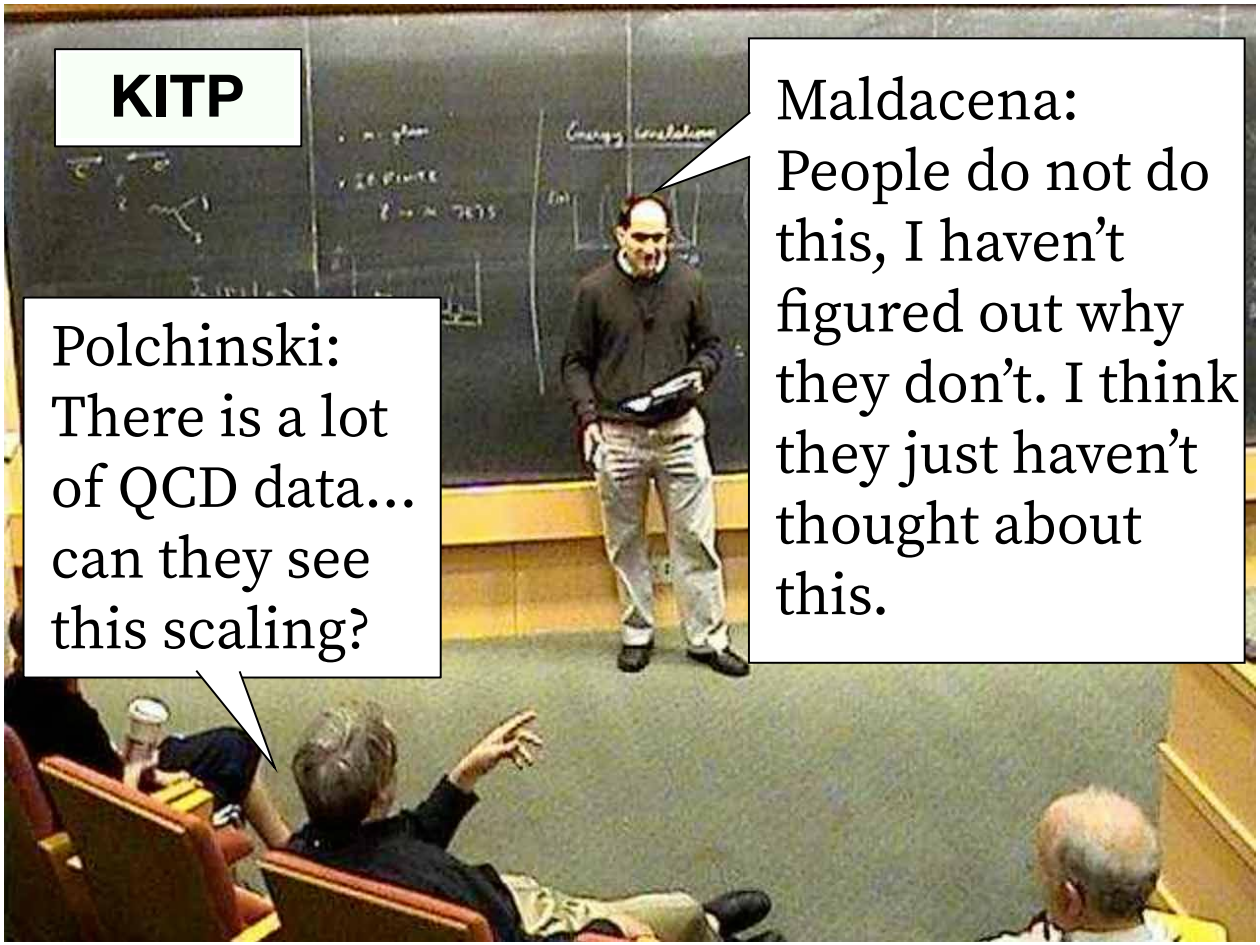






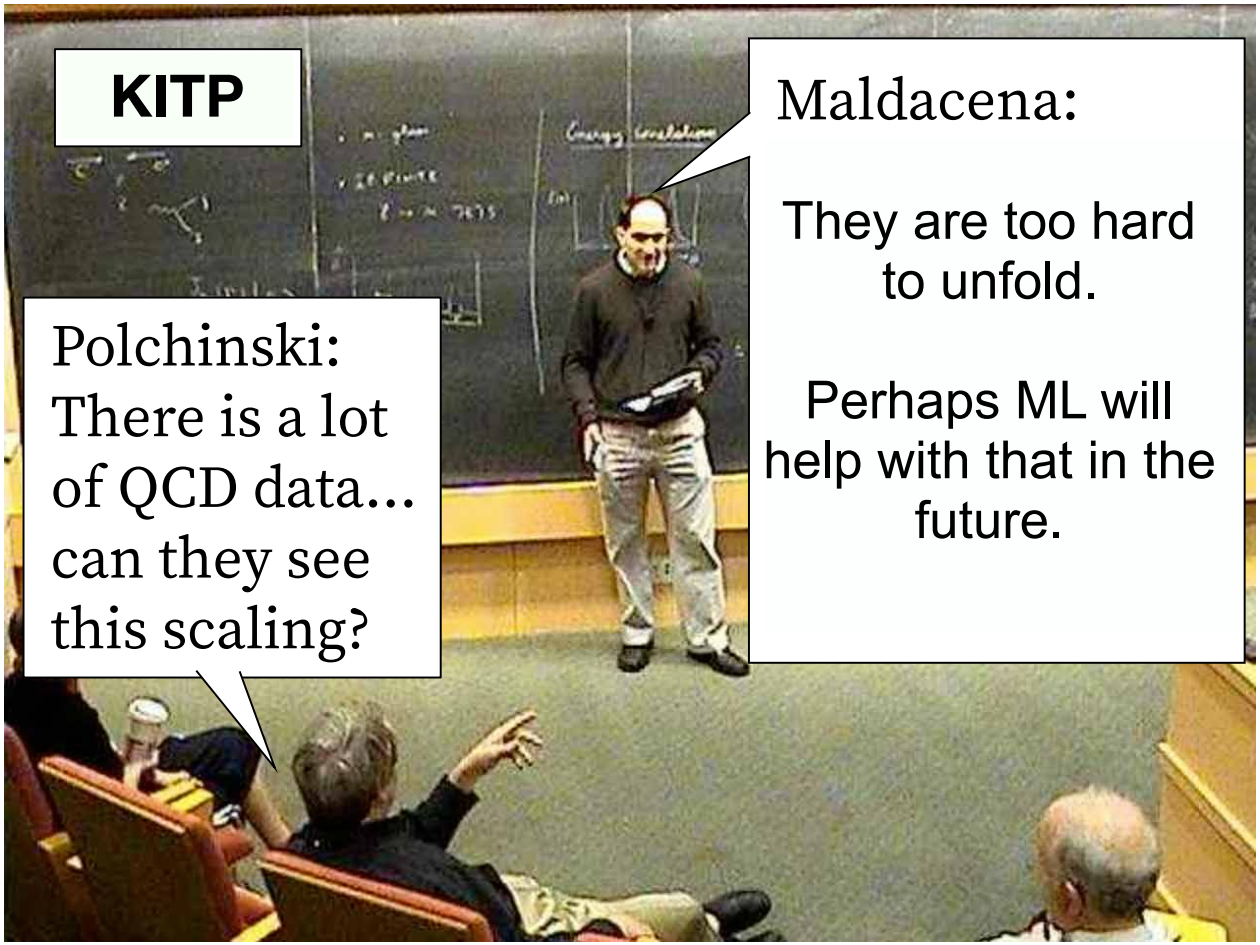
Theory-Experiment Gap

- The extraordinary complexity of the LHC dataset, has produced a gap between what theorists want, and what can be measured.



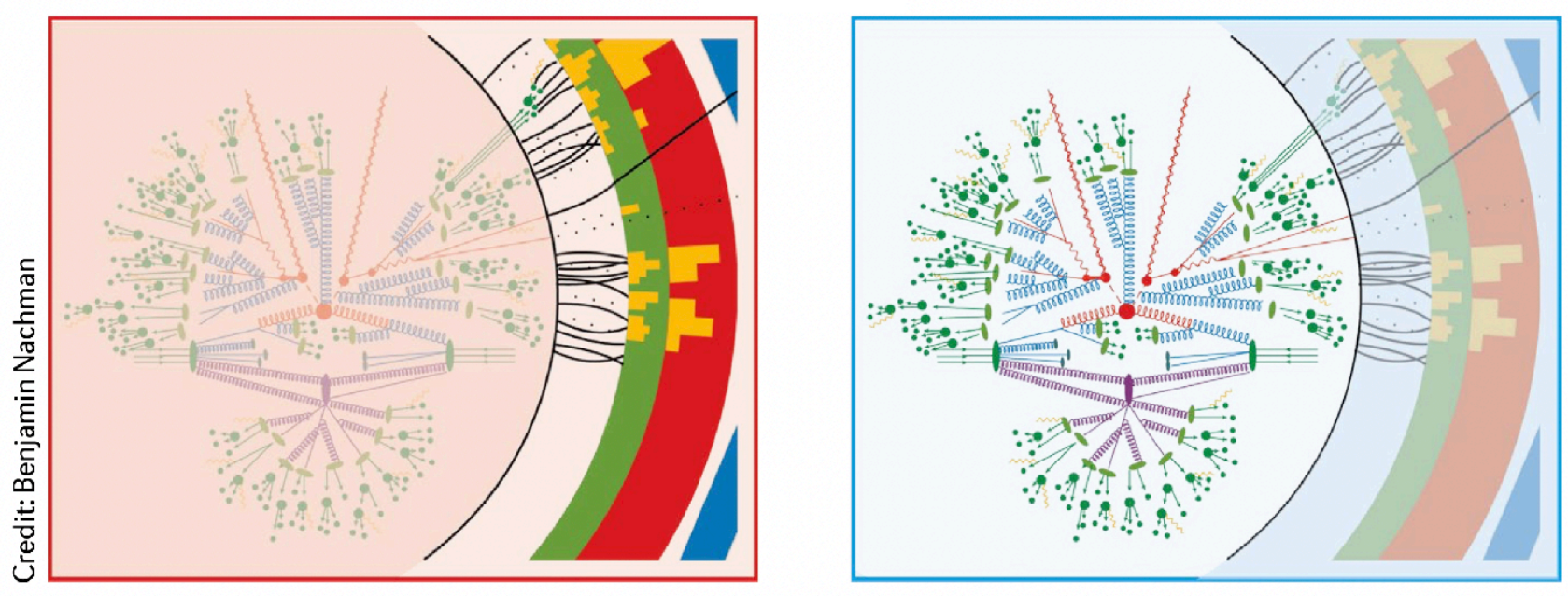
Theory-Experiment Gap

- The extraordinary complexity of the LHC dataset, has produced a gap between what theorists want, and what can be measured.



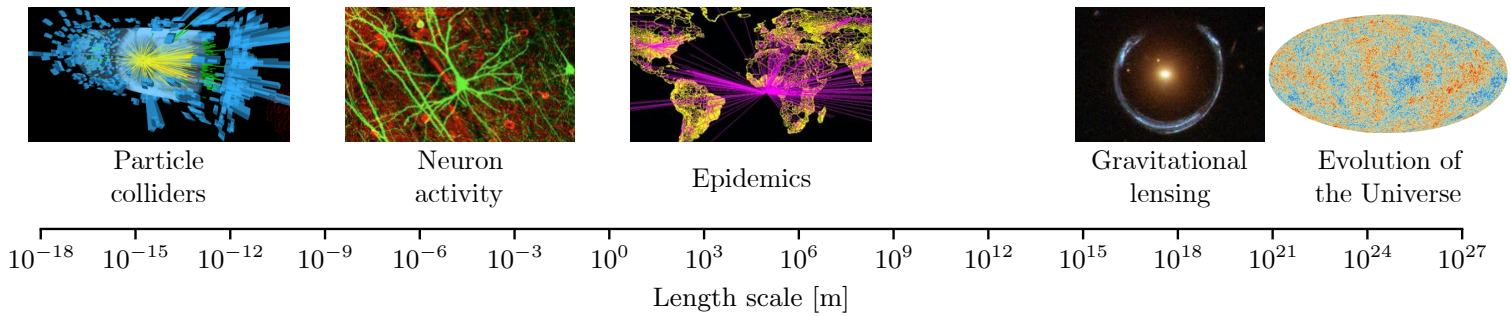
From Detector Data to Theory Comparisons

- To interpret theoretically, measurements require "inverting" the effects of the detector: Unfolding.
- "Simple" if one projects to low dimensional features such as jets.
- To measure statistical properties of energy flux requires unfolding the full particle phase space.



Likelihood Free Inference

- Traditional approaches to unfolding that explicitly determine the likelihood fail:
  - High dimensional input space - full phase space of detector effects
  - High dimensional output space - space of energy correlations
- This is a common feature of many modern data sets in the physical sciences, for which we have high fidelity simulations.

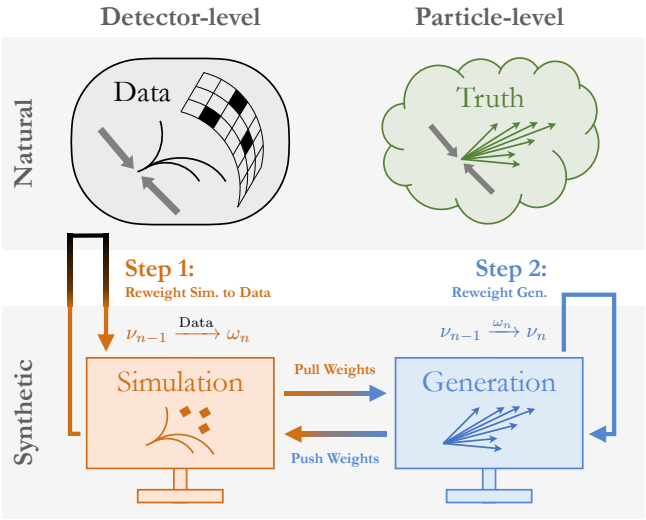


[Cranmer, Brehmer, Louppe]
[Karagiori, Kasieczka, Kravitz, Nachman, Shih]

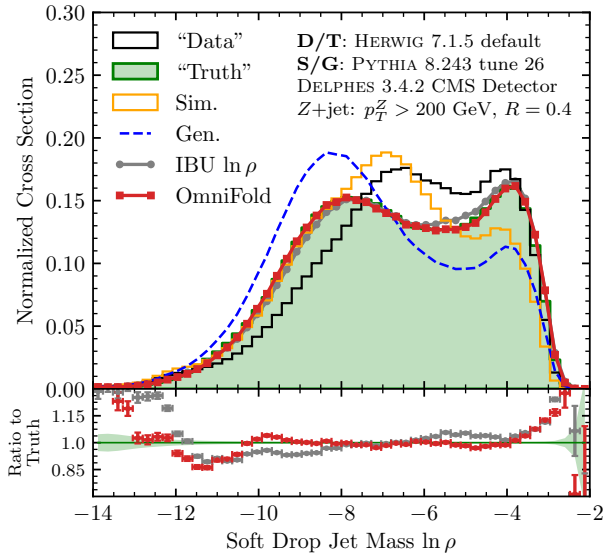
- Significant progress in "Likelihood free inference" using ML based techniques.

Omnifold

- Seminal advance in unfolding for collider physics: Omnifold



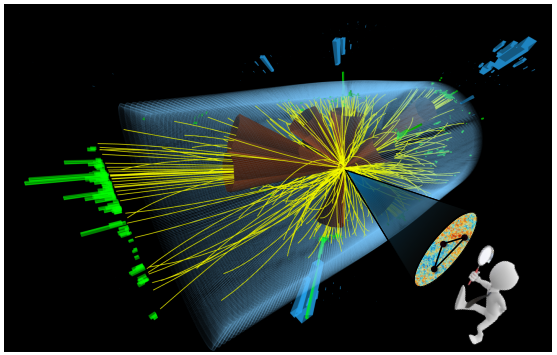
[Andreassen, Nachman]
[Andreassen, Komiske, Metodiev, Nachman, Thaler]



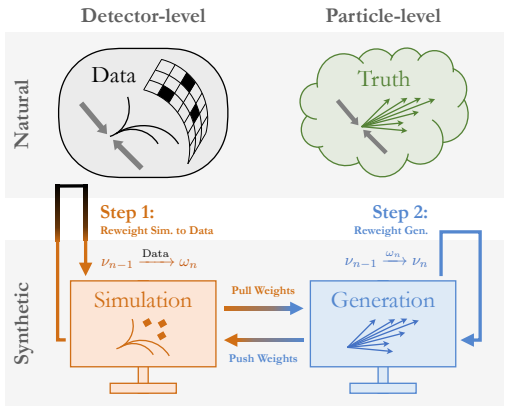
- Rigorously proven to reduce to Iterative Bayesian Unfolding.
- Explicit expression for likelihood intractable in high dimension $\implies$ circumvented by classification task.
- Unfolding of qualitatively new observables (Energy Correlators), combined with theory progress $\implies$ transformative progress in QCD.

Summary

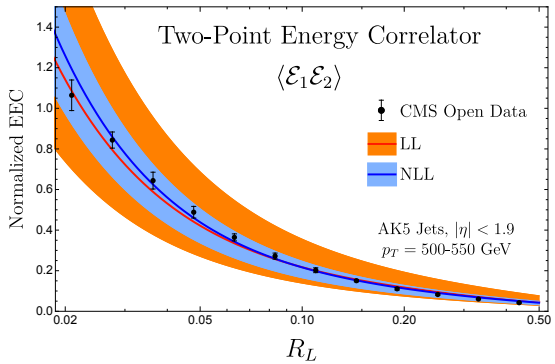
- Collider physics inextricably ties Data Science and Quantum Field Theory.

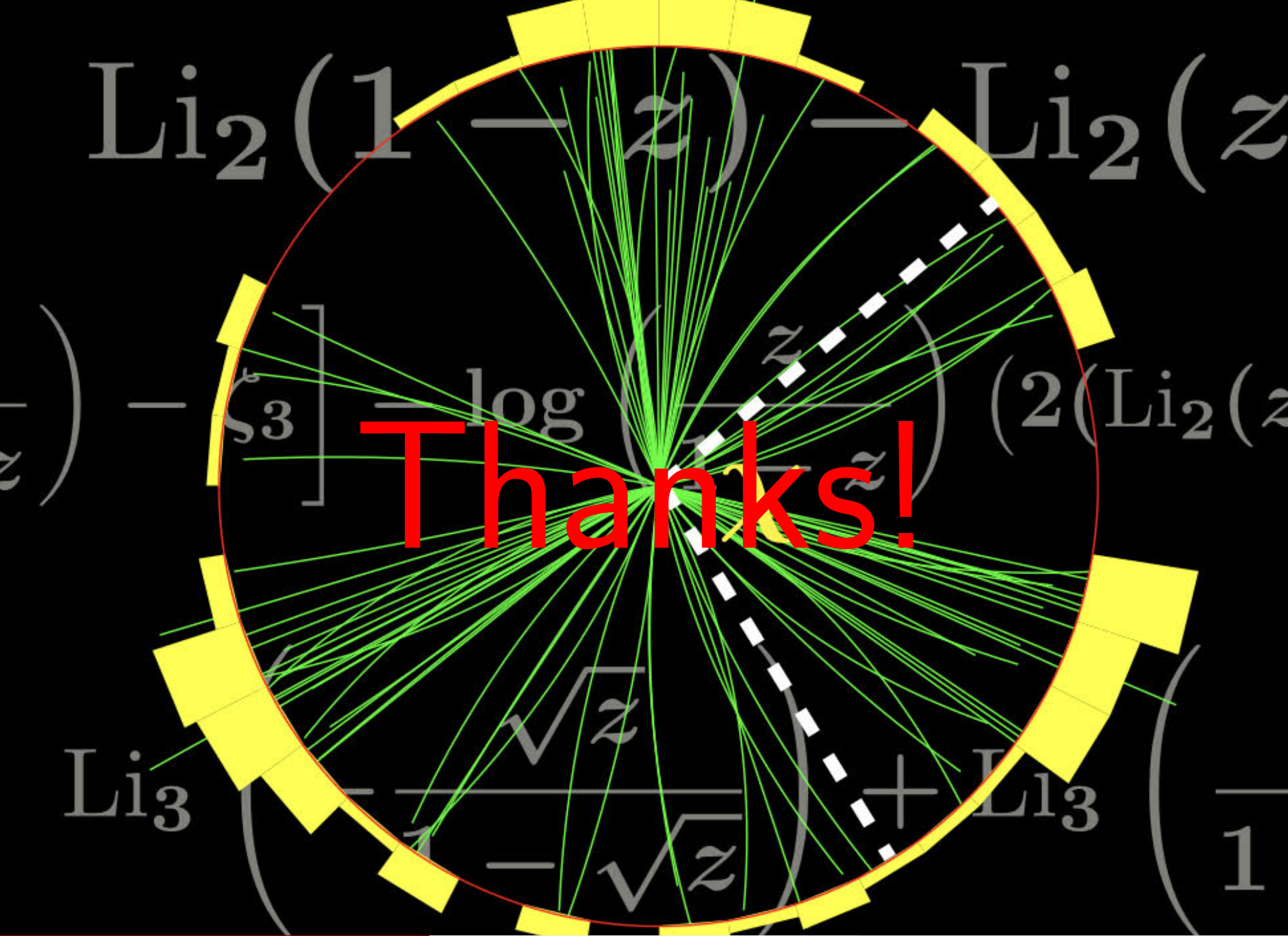


- Progress in Likelihood Free Inference will enable the measurement of qualitatively new classes of Jet Substructure observables.



- Combined with theory progress, this opens the door to a precision physics program using jet substructure!




Thanks!