Machine Learning

Some direction for LH BSM 2023

Gregor Kasiecka (gregor.kasieczka@uni-hamburg.de) - Universtität Hamburg
Andrea Wulzer (andrea.wulzer@cern.ch) - ICREA

ICREA @ IFAE
Institut de Física d'Altes Energies

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

CLUSTER OF EXCELLENCE
QUANTUM UNIVERSE

Agnostic searches for new physics

We search for a needle in the haystack, knowing
(more or less) how the haystack looks like (the SM),
but **not knowing** how the needle (BSM) looks like.

Lots of work done, also by many people here.

Two main threads:

Anomaly detection:

— define some notion of "rarity" (e.g. autoencoders) or "over-density" (e.g. weak supervision)

— select anomalous events and search for BSM in there

Goodness of fit:

— inspect all events searching for statistically significant departures from SM

— can make stronger or weaker assumption on BSM nature and SM knowledge

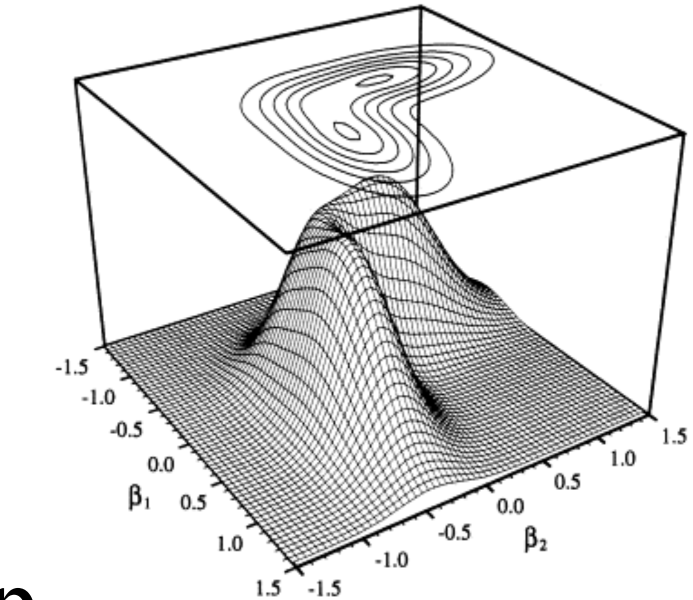
What we could discuss:

Mapping the space of agnostic searches: Define Benchmarks**:** having in mind possible real applications; Strategies for assessing coverage beyond individual benchmarks; A unifying view of AD & GOF; Sharing of results

Use cases Beyond BSM: e.g., spotting out mistakes/differences in generators; validating generative models; data quality monitoring. Comparison with traditional methods and/or cutting edge advances in other fields; other

Likelihood Learning

Aiming at “optimal” BSM sensitivity by precision.
The SM EFT at the LHC and HL-LHC as an ideal use case for these ideas.



Also on this topic, much activity by this group

Several proposals, based on the “likelihood-ratio trick”.

Towards automation, for extensive deployment at LHC

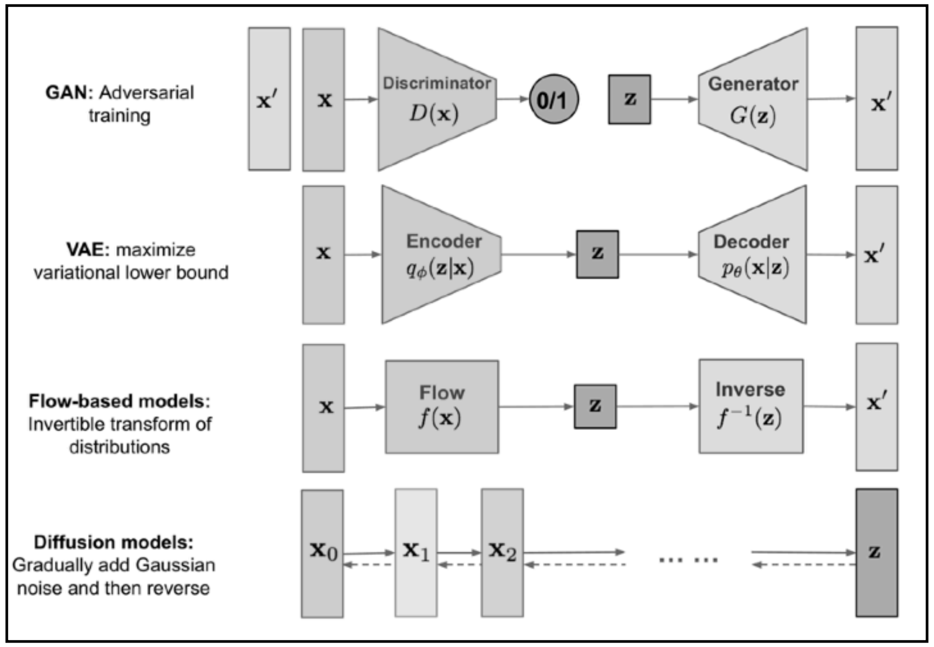
What we could discuss:

Comparing performances: assessing also validation tools, uncertainties, robustness, simplicity, perspective for automation of different proposals.

Towards real analyses: one CSM result using these ideas. What is (or should be) coming next? What is missing from the theory side?

Generative Models

Active use of generative models across for e.g. detector simulation and event generation



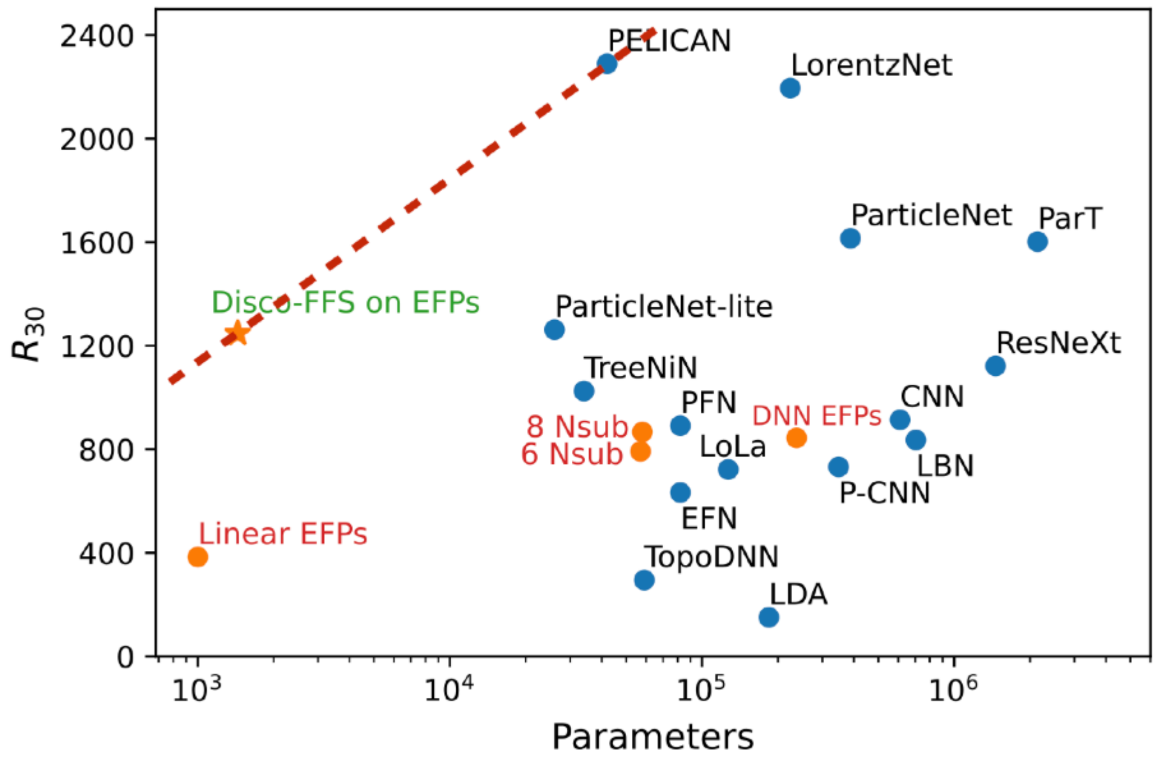
What we could discuss:

In-situ generative models: Are there other uses of learning background distributions (either as likelihood or as sample-able generativer model) from data (also related to anomalies)

Specifically: What can be done with diffusion models?

Physical Symmetries

For tagging algorithms, see benefit from including physical symmetries (primarily Lorentz group)



What we could discuss:

Other types of physics information to include in classifiers: decay chains, IRC safety, …

Adding symmetries to generative models