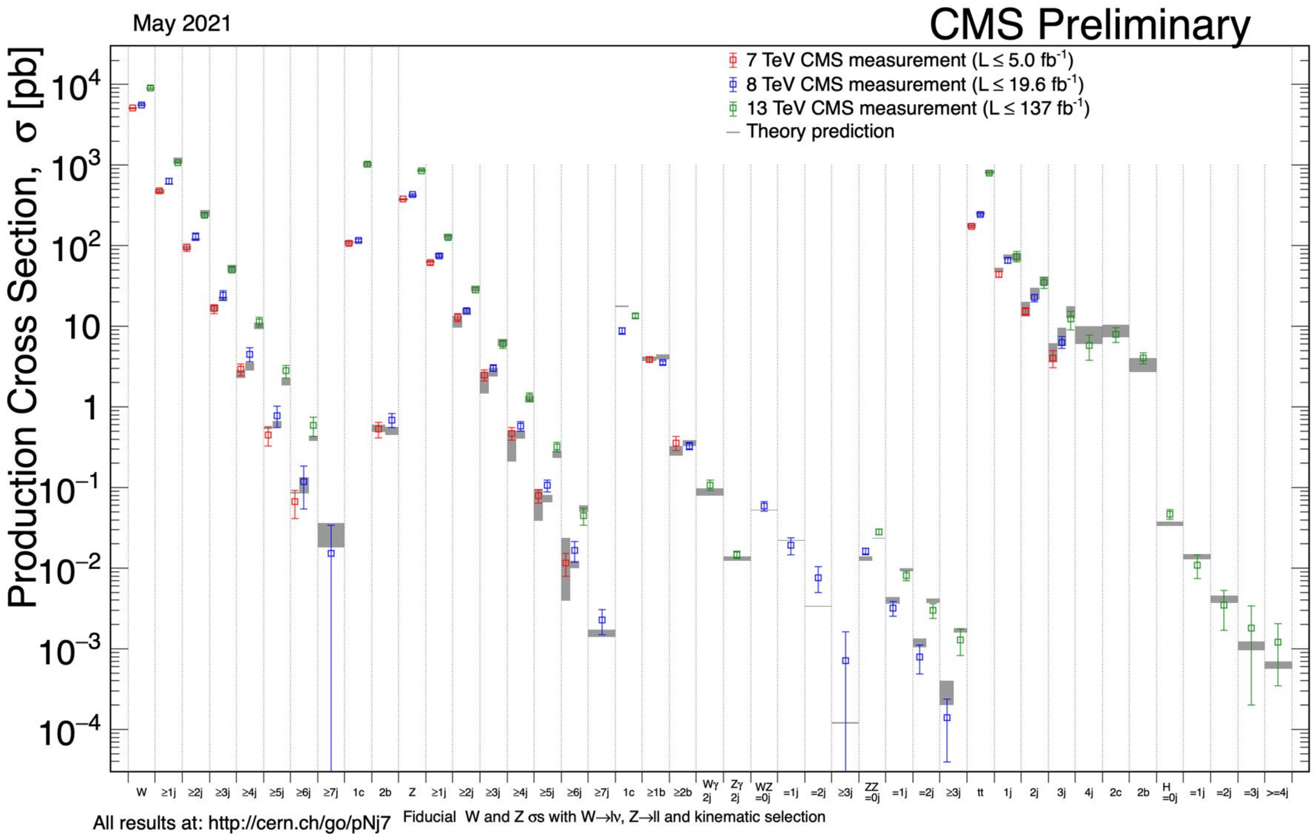
Anomaly detection for new physics searches

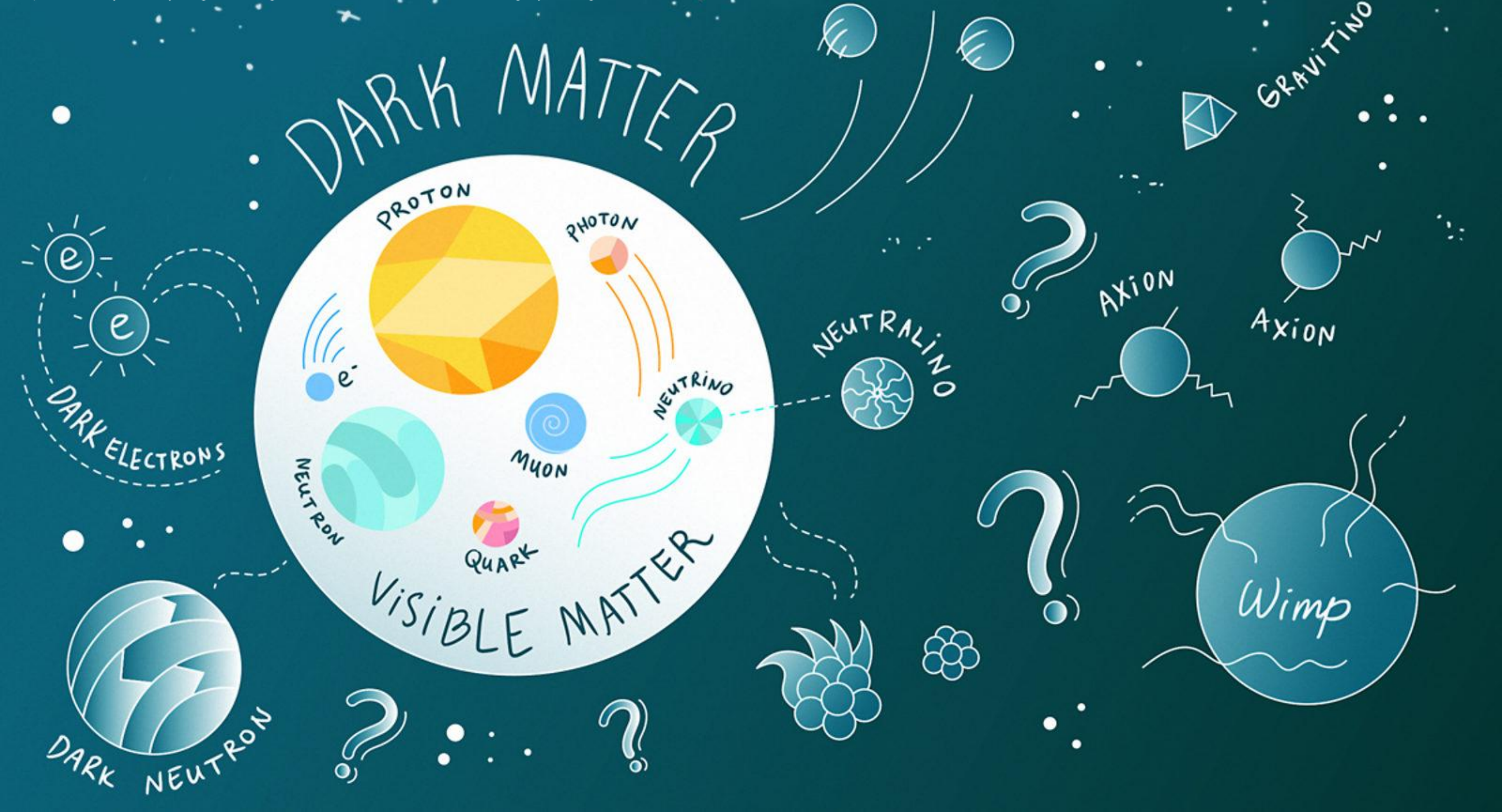
Vinicius M. Mikuni

Introduction



- The **Standard Model** of particle physics is one of the most successful theories of all time
- Several processes with several orders of magnitude are well described by the model
- **Is that everything?**

https://www.symmetrymagazine.org/article/december-2013/four-things-you-might-not-know-about-dark-matter



Introduction

Model dependent searches:

- Take a **well-motivated** new physics scenario with well-defined phenomenology
- **Maximize** the search sensitivity based on the signal properties
- **Most sensitive** strategy to probe that particular new physics scenario, but unlikely to be useful for other searches

Model independent searches:

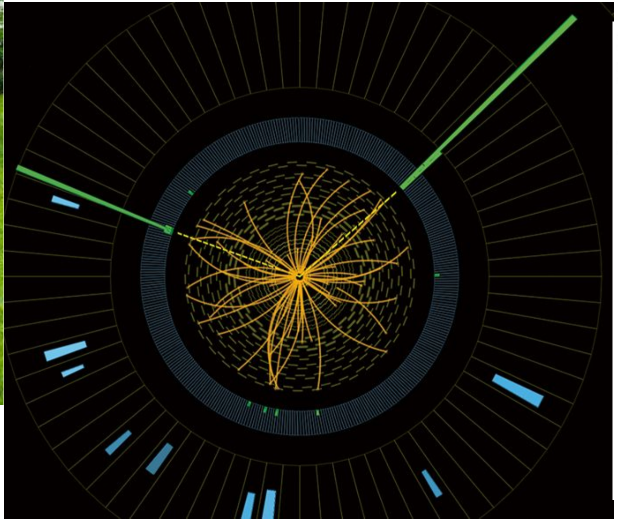
- **Minimal** set of assumptions for signal properties
- Look for **deviations** of the background only hypothesis
- Not optimal for a particular new physics scenario, but likely to be **sensitive to multiple scenarios** satisfying the minimal assumptions

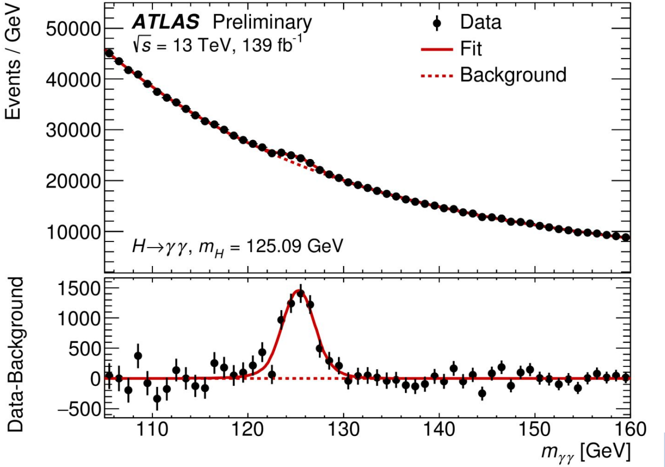
Will focus on this one today

What is an anomaly anyway?

- **Anomaly detection** is often associated to **outlier detection**
- Our application is a bit different: a **single particle collision is not very informative**, only an ensemble of events are!

- There are also examples of outlier detection in HEP such as **detector quality monitoring**





General strategies

https://iml-wg.github.io/HEPML-LivingReview/

- Anomaly detection.
  - Learning New Physics from a Machine [DOI]
  - Anomaly Detection for Resonant New Physics with Machine Learning [DOI]
  - Extending the search for new resonances with machine learning [DOI]
  - Learning Multivariate New Physics [DOI]
  - Searching for New Physics with Deep Autoencoders [DOI]
  - QCD or What? [DOI]
  - A robust anomaly finder based on autoencoder
  - Variational Autoencoders for New Physics Mining at the Large Hadron Collider [DOI]
  - Adversarially-trained autoencoders for robust unsupervised new physics searches [DOI]
  - Novelty Detection Meets Collider Physics [DOI]
  - Guiding New Physics Searches with Unsupervised Learning [DOI]
  - Does SUSY have friends? A new approach for LHC event analysis [DOI]
  - Nonparametric semisupervised classification for signal detection in high energy physics
  - Uncovering latent jet substructure [DOI]
  - Simulation Assisted Likelihood-free Anomaly Detection [DOI]
  - Anomaly Detection with Density Estimation [DOI]
  - A generic anti-QCD jet tagger [DOI]
  - Transferability of Deep Learning Models in Searches for New Physics at Colliders [DOI]
  - Use of a Generalized Energy Mover's Distance in the Search for Rare Phenomena at Colliders [DOI]
  - Adversarially Learned Anomaly Detection on CMS Open Data: re-discovering the top quark [DOI]
  - Dijet resonance search with weak supervision using 13 TeV pp collisions in the ATLAS detector [DOI]
  - Learning the latent structure of collider events [DOI]
  - Finding New Physics without learning about it: Anomaly Detection as a tool for Searches at Colliders [DOI]
  - Tag N' Train: A Technique to Train Improved Classifiers on Unlabeled Data [DOI]
  - Variational Autoencoders for Anomalous Jet Tagging
  - Anomaly Awareness
  - Unsupervised Outlier Detection in Heavy-Ion Collisions
  - Decoding Dark Matter Substructure without Supervision
  - Mass Unspecific Supervised Tagging (MUST) for boosted jets [DOI]
  - Simulation-Assisted Decorrelation for Resonant Anomaly Detection
  - Anomaly Detection With Conditional Variational Autoencoders
  - Unsupervised clustering for collider physics
  - Combining outlier analysis algorithms to identify new physics at the LHC
  - Quasi Anomalous Knowledge: Searching for new physics with embedded knowledge
  - Uncovering hidden patterns in collider events with Bayesian probabilistic models
  - Unsupervised in-distribution anomaly detection of new physics through conditional density estimation
  - The LHC Olympics 2020: A Community Challenge for Anomaly Detection in High Energy Physics
  - Model-Independent Detection of New Physics Signals Using Interpretable Semi-Supervised Classifier Tests
  - Topological Obstructions to Autoencoding
  - Unsupervised Event Classification with Graphs on Classical and Photonic Quantum Computers
  - Bump Hunting in Latent Space
  - Comparing Weak- and Unsupervised Methods for Resonant Anomaly Detection
  - Better Latent Spaces for Better Autoencoders
  - Autoencoders for unsupervised anomaly detection in high energy physics
  - Via Machinae: Searching for Stellar Streams using Unsupervised Machine Learning
  - Anomaly detection with Convolutional Graph Neural Networks
  - Anomalous Jet Identification via Sequence Modeling
  - The Dark Machines Anomaly Score Challenge: Benchmark Data and Model Independent Event Classification for the Large Hadron Collider
  - RanBox: Anomaly Detection in the Copula Space
  - Rare and Different: Anomaly Scores from a combination of likelihood and out-of-distribution models to detect new physics at the LHC
  - LHC physics dataset for unsupervised New Physics detection at 40 MHz
  - New Methods and Datasets for Group Anomaly Detection From Fundamental Physics
  - Autoencoders on FPGAs for real-time, unsupervised new physics detection at 40 MHz at the Large Hadron Collider
  - Classifying Anomalies THrough Outer Density Estimation (CATHODE)
  - Deep Set Auto Encoders for Anomaly Detection in Particle Physics
  - Challenges for Unsupervised Anomaly Detection in Particle Physics
  - Improving Variational Autoencoders for New Physics Detection at the LHC with Normalizing Flows
  - Signal-agnostic dark matter searches in direct detection data with machine learning

- What is the **best** anomaly detection method?

General strategies

https://iml-wg.github.io/HEPML-LivingReview/

- Anomaly detection.
  - Learning New Physics from a Machine [DOI]
  - Anomaly Detection for Resonant New Physics with Machine Learning [DOI]
  - Extendir
  - Learning
  - Searchi
  - QCD or
  - A robus
  - Variation
  - Adversa
  - Novelty
  - Guiding
  - Does SU
  - Nonpara
  - Uncover
  - Simulati
  - Anomaly
  - A gener
  - Transfe
  - Use of
  - Advers
  - Dijet re
  - Learnin
  - Finding
  - Tag N'
  - Variational Autoencoders for Anomalous Jet Tagging
  - Anomaly Awareness
  - Unsupervised Outlier Detection in Heavy-Ion Collisions
  - Decoding Dark Matter Substructure without Supervision
  - Mass Unspecific Supervised Tagging (MUST) for boosted jets [DOI]
  - Simulation-Assisted Decorrelation for Resonant Anomaly Detection
  - Anomaly Detection With Conditional Variational Autoencoders
  - Unsupervised clustering for collider physics
  - Combining outlier analysis algorithms to identify new physics at the LHC
  - Quasi Anomalous Knowledge: Searching for new physics with embedded knowledge
  - Challenges for Unsupervised Anomaly Detection in Particle Physics
  - Improving Variational Autoencoders for New Physics Detection at the LHC with Normalizing Flows
  - Signal-agnostic dark matter searches in direct detection data with machine learning

What is the **best**

ection

No clear winner! Different methods use different assumptions and are often complementary!

Popular strategies

Signal is an over density for some feature:

- Even though new physics is rare, assume there is at least one feature where **$p_s(x)/p_b(x)$ is high**: often assumed to be some invariant mass combination
- Requires an estimate of $p_b(x)$ and prior knowledge of the resonant feature to use

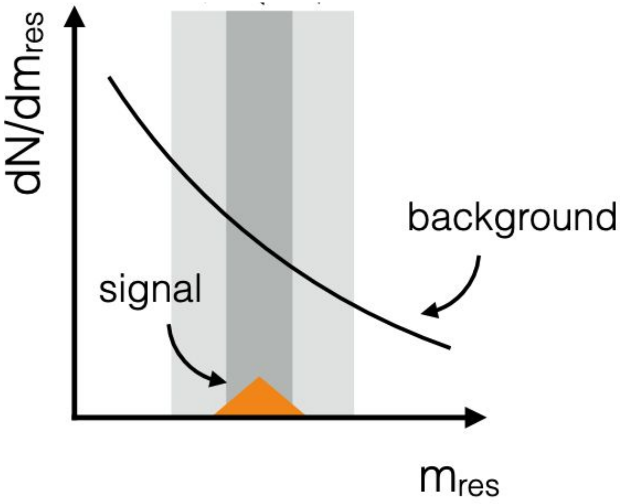
Signal is located where the background density is low

- Assume that "**tails**" of distributions are informative
- Can be **trained using data** directly, as long as possible signal contamination is low
- **Low $p_b(x)$ is not universal** and the choice of x determines the quality of the algorithm

Resonant anomaly detection

Signal is an over density for some feature:

- Even though new physics is rare, assume there is at least one feature where **$p_s(x)/p_b(x)$ is high**: often assumed to be some invariant mass combination
- Requires a control region with only $p_b(x)$ and prior knowledge of the resonant feature to use



Resources for resonant anomaly detection:

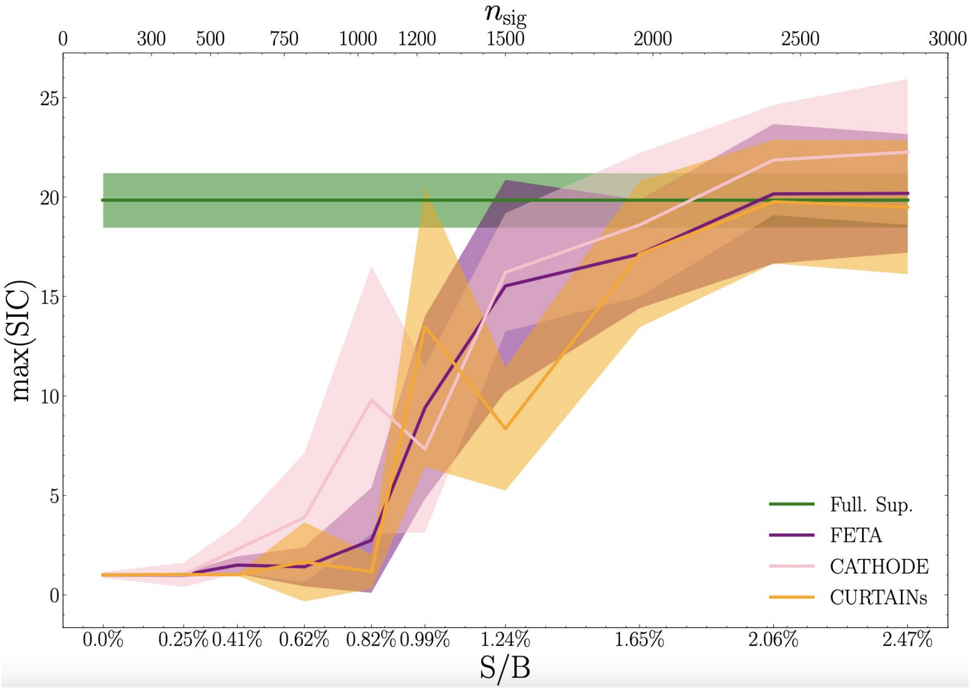
- PRL 121 (2018) 241803, 1805.02664
- PRL 125 (2020) 131801, 2005.02983
- PRD 101 (2020) 095004
- Hallin et al., 2109.00546
- PRD 101 (2020) 9, 095004
- PRD 101 (2020) 075042
- Raine et al., 2203.09470
- Golling, Tobias, et al., 2212.11285

- Similar to standard **bump hunt**
- Use the **side-bands** to learn the **background** distribution in the signal region: either through **morphing** or **likelihood learning**
- Compare predicted background events with data in the signal region: often a **classifier** is trained to separate data from predicted background

Resonant anomaly detection

Signal is an over density for some feature:

- Even though new physics is rare, assume there is at least one feature where **$p_s(x)/p_b(x)$ is high**: often assumed to be some invariant mass combination
- Requires a control region with only $p_b(x)$ and prior knowledge of the resonant feature to use

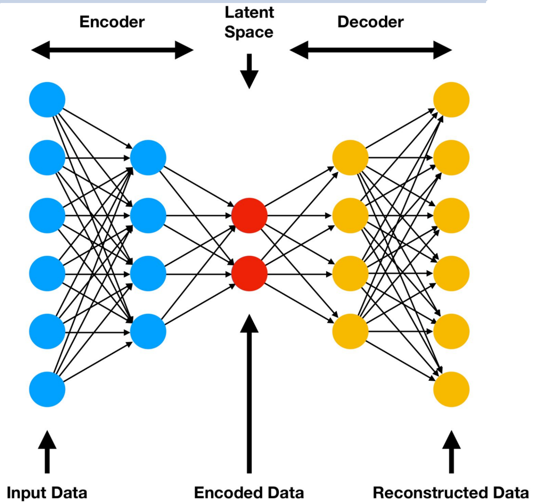


Similar performance between different strategies, almost as good as a **fully supervised classifier**

Non-resonant anomaly detection

Signal is located where the background density is low

- Assume that "**tails**" of distributions are informative
- Can be **trained using data** directly, as long as possible signal contamination is low
- **Low $p_b(x)$ is not universal** and the choice of x determines the quality of the algorithm



- Majority of popular strategies use **autoencoders**
- Train the model with **background-enriched data**
- **Encode** the inputs to a low dimensional representation and try to **decode** it back to the input set
- **Anomalous events** are often **poorly reconstructed**, since there are not many examples seeing during training
- **NFs** and **GANs** have also been used in a similar context

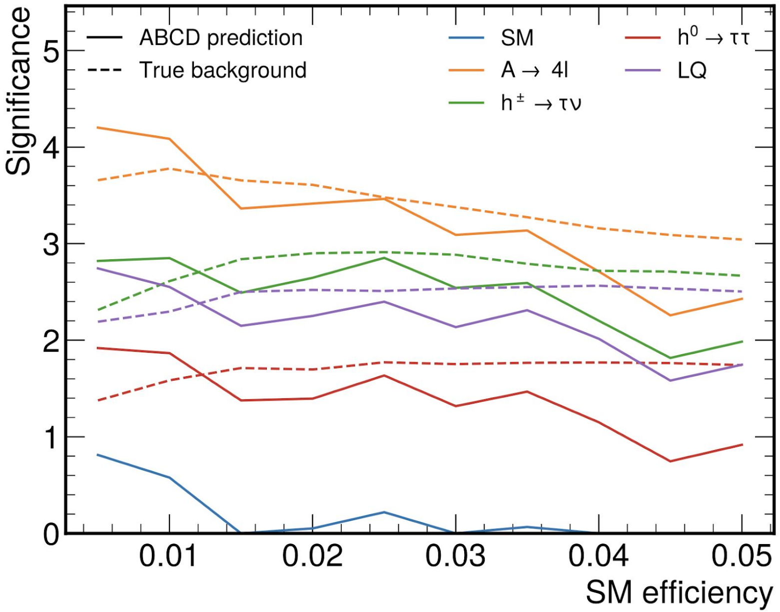
Resources for non-resonant anomaly detection:

- Farina, Marco, Yuichiro Nakai, and David Shih., Physical Review D 101.7 (2020): 075021.
- Finke, Thorben, et al. Journal of High Energy Physics 2021.6 (2021): 1-32.
- Dillon, Barry M., et al. SciPost Physics 11.3 (2021): 061.
- **Mikuni, Vinicius**, and Florencia Canelli. Physical Review D 103.9 (2021): 092007.
- Pol, Adrian Alan, et al. 2019 18th IEEE ICMLA. IEEE, 2019.
- Cheng, Taoli, et al. Physical Review D 107.1 (2023): 016002.
- Blance, Andrew, Michael Spannowsky, and Philip Waite. JHEP 2019.10 (2019).
- Cerri, Olmo, et al. Journal of High Energy Physics 2019.5 (2019): 1-29.
- Roy, Tuhin S., and Aravind H. Vijay. 1903.02032 (2019).
- Ostdiek, Bryan. SciPost Physics 12.1 (2022): 045.

Non-resonant anomaly detection

Signal is located where the background density is low

- Assume that “**tails**” of distributions are informative
- Can be **trained using data** directly, as long as possible signal contamination is low
- **Low $p_b(x)$ is not universal** and the choice of x determines the quality of the algorithm



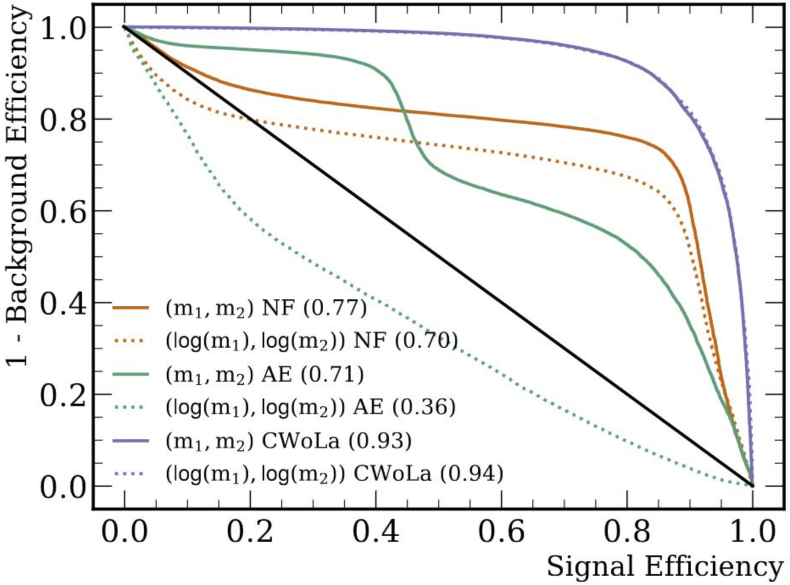
Mikuni, Vinicius, Benjamin Nachman, and David Shih. *Physical Review D* 105.5 (2022): 055006.

- One of the difficulties of using autoencoders is to **determine the background** distribution in the region containing possible anomalies
- Multiple **decorrelated** autoencoders can be used, with background distribution determined by the **ABCD method**

Non-resonant anomaly detection

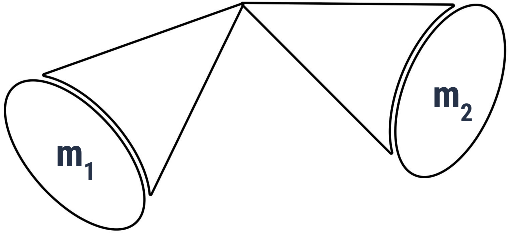
Signal is located where the background density is low

- Assume that "**tails**" of distributions are informative
- Can be **trained using data** directly, as long as possible signal contamination is low
- **Low $p_b(x)$ is not universal** and the choice of x determines the quality of the algorithm

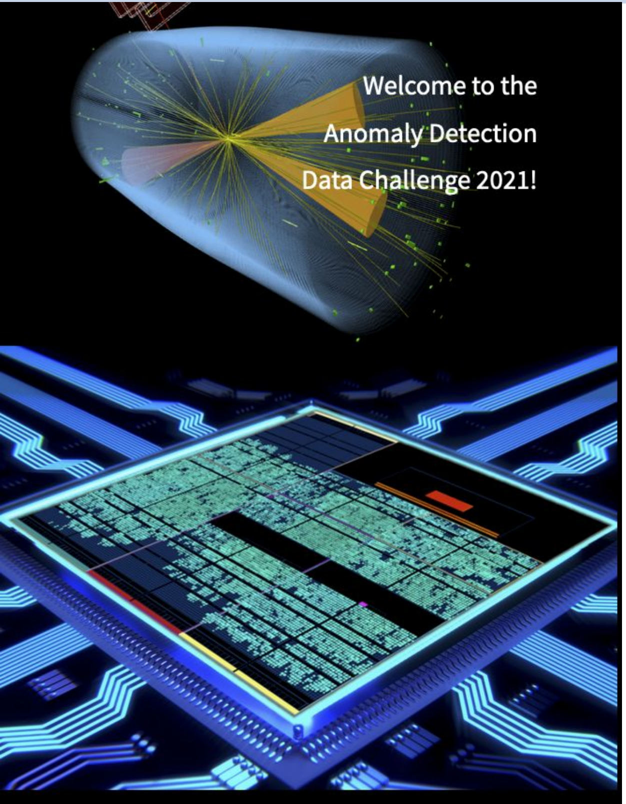


Kasieczka, G., Mastandrea, R., **Mikuni, V.**, Nachman, B., Pettee, M., & Shih, D. (2023). *Physical Review D*, *107*(1), 015009.

- Choice of **representation of inputs** also affects the performance!
- Differences in performance for autoencoders when using **m_1, m_2** as inputs or **$\log(m_1)$, $\log(m_2)$**



Anomaly detection at trigger level



- There has also been a number of exciting efforts to do anomaly detection at **trigger level**
- **Microsecond inference time** needed to achieve this goal at the LHC
- Number of works showing the feasibility of the idea with even a data challenge to compare methods
- See More at **Javier's talk!**

Govorkova, Ekaterina, et al. *Scientific Data* 9.1 (2022): 118.
Govorkova, Ekaterina, et al. *Nature Machine Intelligence* 4.2 (2022): 154-161.
Mikuni, Vinicius, Benjamin Nachman, and David Shih. *Physical Review D* 105.5 (2022): 055006.
Dillon, Barry M., et al. *2206.14225* (2022).

Data challenges

- A number of **data challenges** were proposed to raise awareness of this new and complementary way of **searching for new physics**

The LHC Olympics 2020

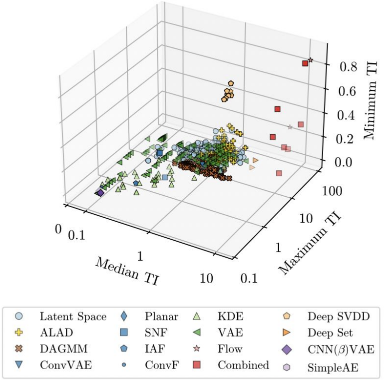
A Community Challenge for Anomaly Detection in High Energy Physics

Kasieczka, Gregor, et al. *Reports on progress in physics* 84.12 (2021): 124201.

The Dark Machines Anomaly Score Challenge II

Based on arXiv: 2105.14027

https://github.com/bostdiek/DarkMachines-UnsupervisedChallenge

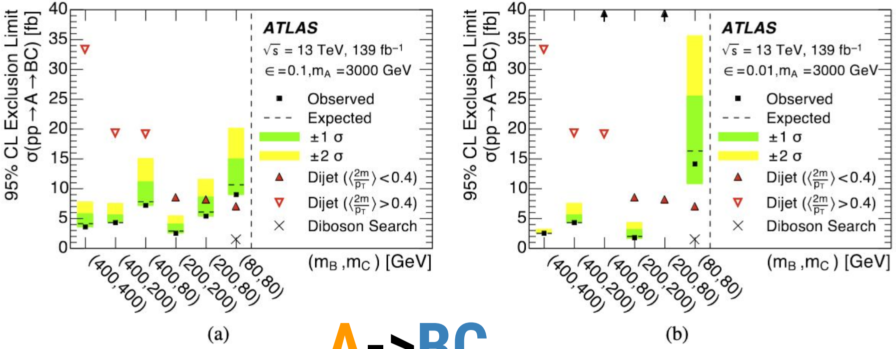


Aarrestad, Thea, et al. *SciPost Physics* 12.1 (2022): 043.

Slides at ML4Jets 2021

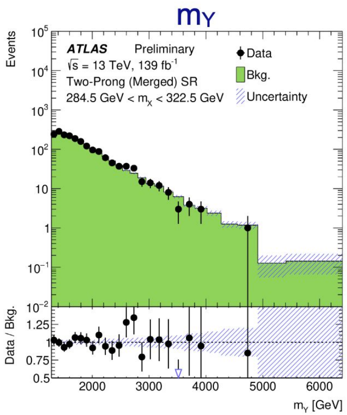
Conclusion

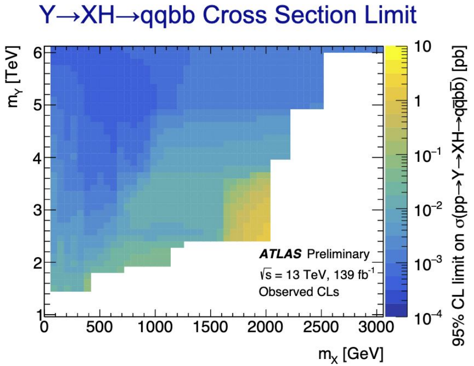
- **Anomaly detection** is and **alternative** and **complementary** way to search for new physics processes
- Different anomaly detection methods still rely on a few assumptions and is important to be **aware of their limitations**
- Nevertheless, **collider results** using some of these methods are already out!



(a) (b)

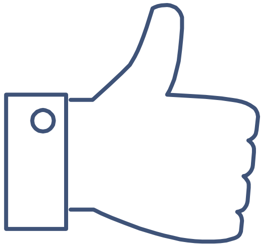
A->BC





ATLAS Collaboration, Phys. Rev. Lett. 125, 131801 (2020)

ATLAS-CONF-2022-045

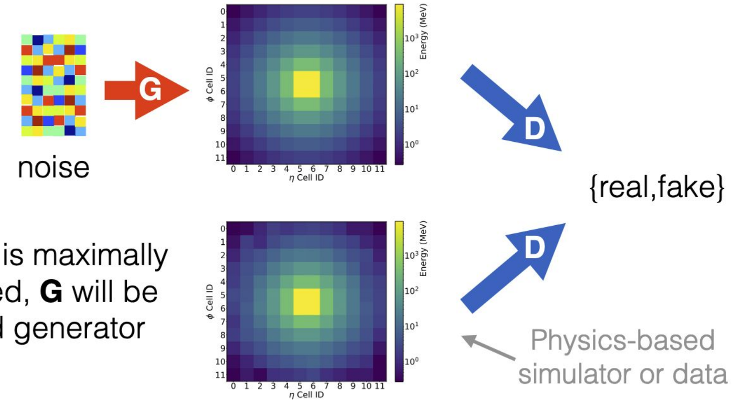


THANKS!

Any questions?

Popular strategies

GANS



When **D** is maximally confused, **G** will be a good generator

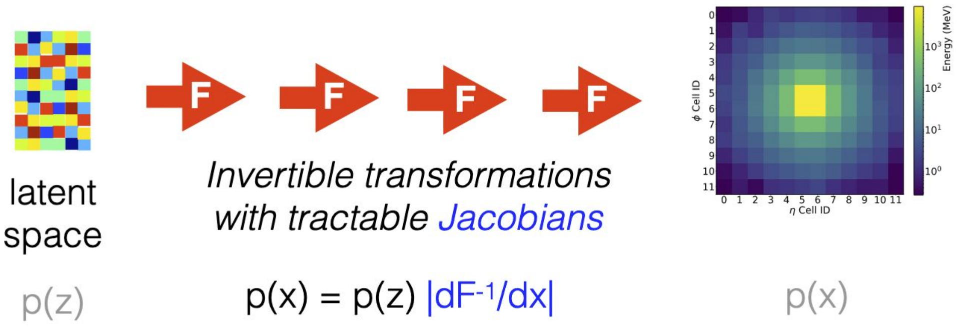
Train the GAN to learn the background distribution

- Compare samples from GAN with the data in a region of interest!

Eur.Phys.J.Plus 136 (2021)

Flows

Optimize via maximum likelihood



Train the NF to learn $p_b(x)$:

- Find regions with low $p_b(x)$
- Sample from the flow and compare with data
- Interpolate multiple background regions

Hallin et al., 2109.00546
PRD 101 (2020) 9, 095004
PRD 101 (2020) 075042
Raine et al., 2203.09470