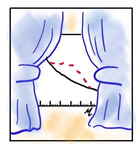
Draw the Curtains, Funnel, DREAM

- AI Physics Day -

CURTAINs

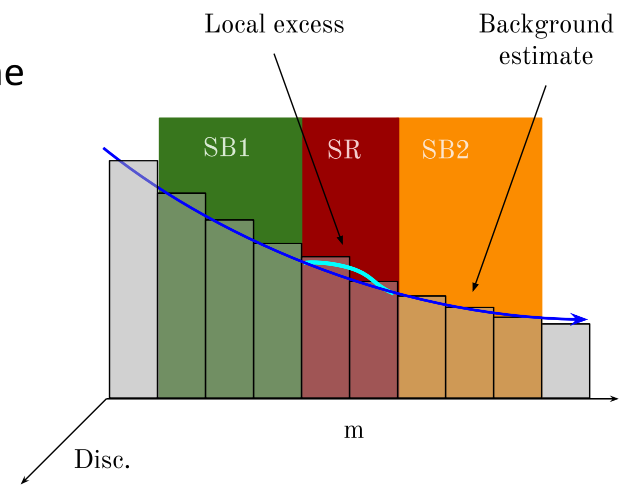


Anomaly Detection

- In high energy physics we are looking for a tiny signal in an overwhelming background.
- We want to develop **model independent** methods for searching for new physics.
- We can assume that new resonances are localised in some feature.
- **How can we do a model independent search using this simple assumption?**

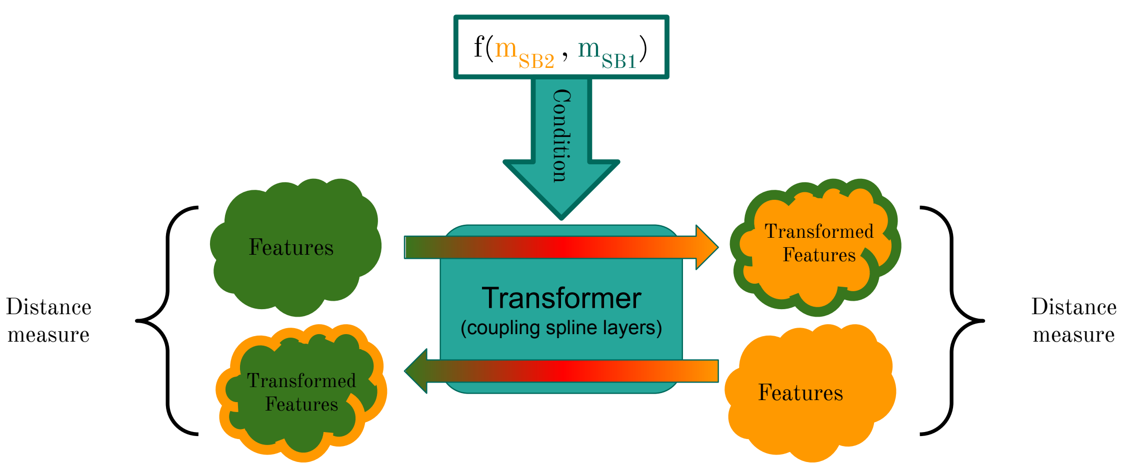
Bump Hunts

- Need background estimate in the Signal Region (SR).
- Extrapolate from the sidebands (SB).
- Look for a bump!
- Need additional observables to increase sensitivity.
- **How to extend interpolation?**



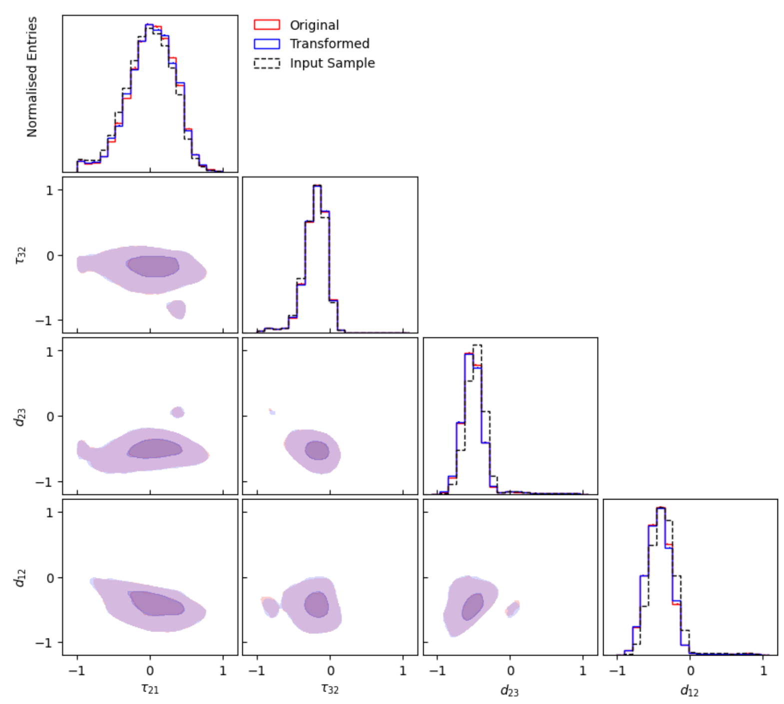
Method

- **Transform** data from the sidebands into the signal region.
- Transformed side bands provide a background template.
- Train a classifier to separate signal region samples from background template.
- Training on mixtures of samples results in the optimal classifier.



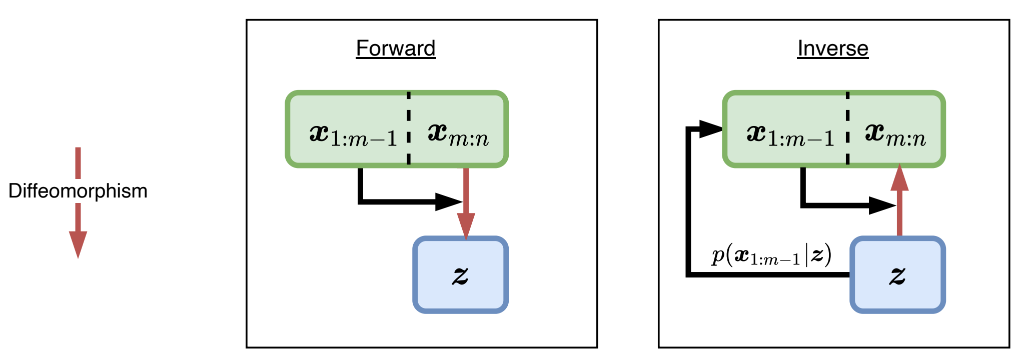
Results

Transformed samples from both side bands are **indistinguishable** from signal region samples



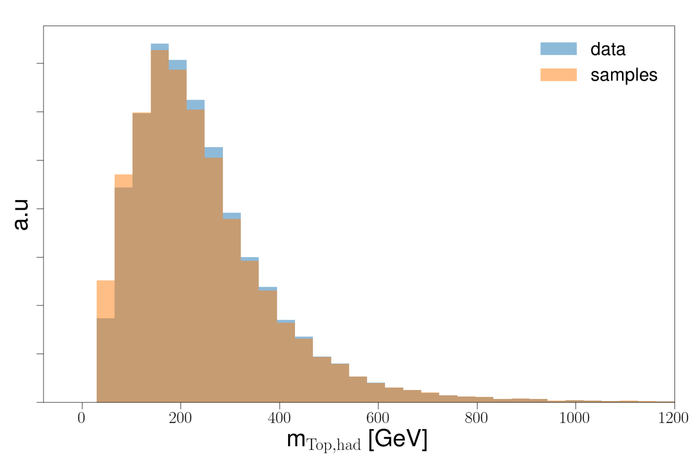
Funnels

$$\log p(\boldsymbol{x}) \simeq \log p(\boldsymbol{z}) + \mathcal{V}(\boldsymbol{x}) + \mathcal{E}(\boldsymbol{x})$$



$$\mathcal{V}(\boldsymbol{x}) = \log|\det(J_\theta(\boldsymbol{x}_{m:n}))| + \log p(\boldsymbol{x}_{1:m-1}|\boldsymbol{z}) \qquad \mathcal{E} = 0$$

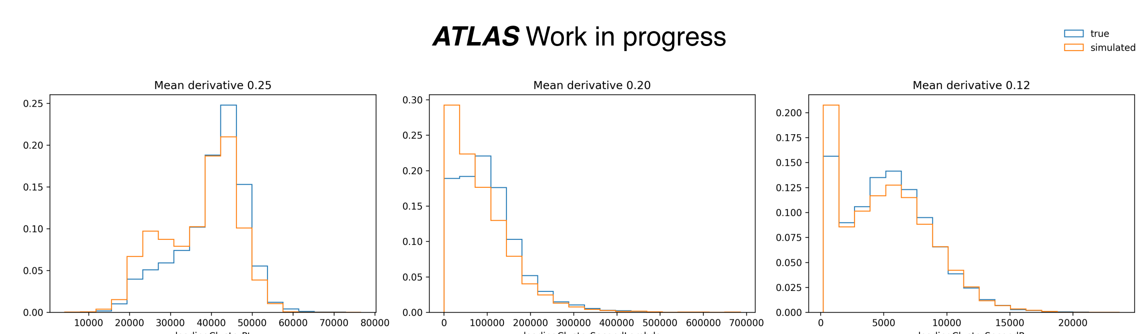
- Unlike auto encoders, **exact likelihoods** can be calculated.
- Unlike normalising flows, the data dimension can be **reduced**.
- Strongly correlated distributions such as those in this $t\bar{t}$ sample can be learned.



This method allows likelihood methods to scale to **large datasets** efficiently.

DREAM

- How to automatically **validate** high dimensional simulations?
- Many variables of interest, different approaches, skewed distributions, low statistics… **Difficult** to validate!
- **D**ensity **R**atio **E**stimation for **A**utomatic **M**onitoring uses a classifier to separate simulated samples from true samples.



The **derivatives** of the classifier can be used to order the features.

Sample	AUC	Model
Pions 262 GeV 1.5η	**0.527**	**FCSGAN**
	0.538	FastCaloGANv1
	0.540	FastCaloGAN
	0.602	FCS

The classifiers ability to separate (AUC) **scores** the model. More separation = worse modelling.

UNIVERSITÉ DE GENÈVE

Samuel Klein samuel.klein@unige.ch Tobias Golling, Slava Voloshynovskiy, Johnny Raine, Sebastian Pina-Otey, Debajyoti Sengupta 03.11.2021