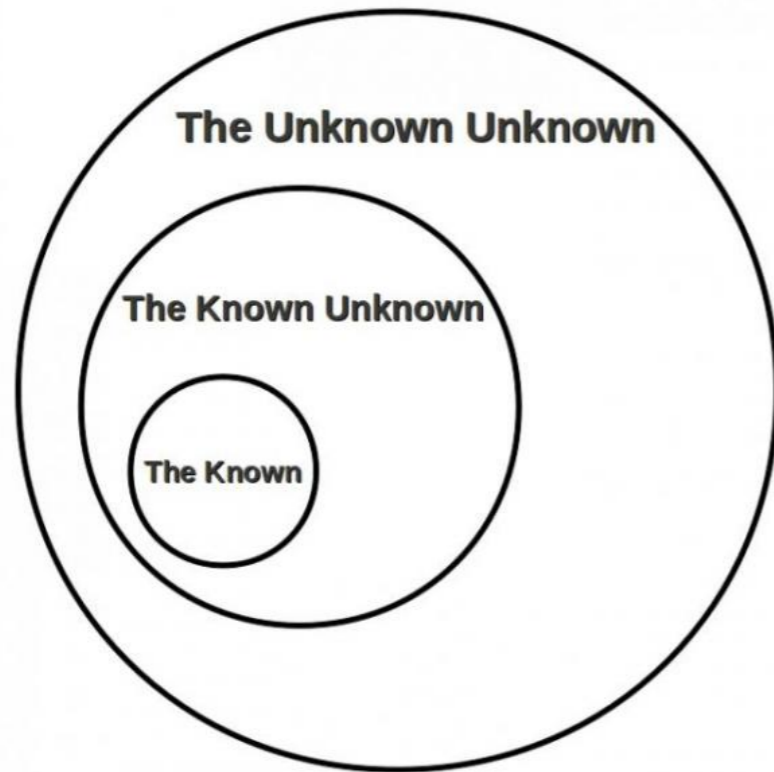


# ML in (LHC) Experiments

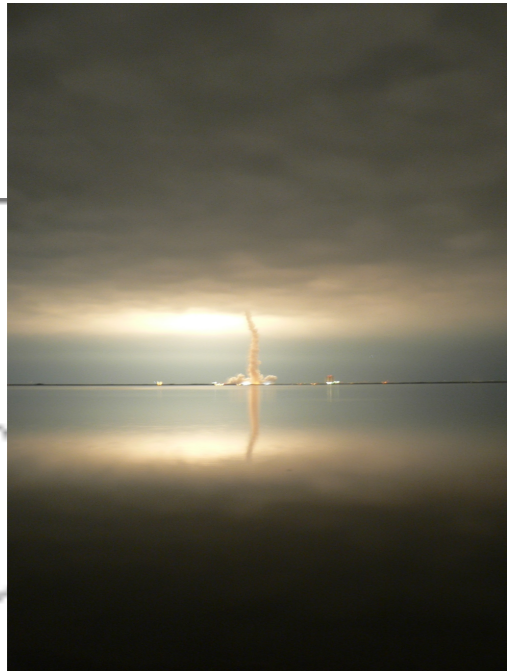


**Deep Thinking**  
**vs**  
**Deep Learning**

**Deepak Kar**  
**[deepak.kar@cern.ch](mailto:deepak.kar@cern.ch)**



# Who am I?



**Gainesville, FL,  
USA**  
2003-2008



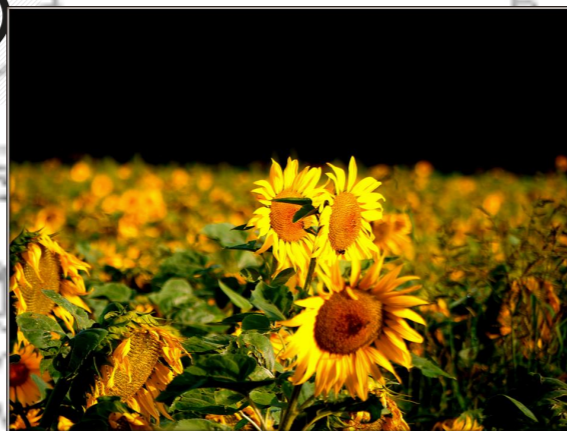
**Glasgow, UK**  
2012-2014



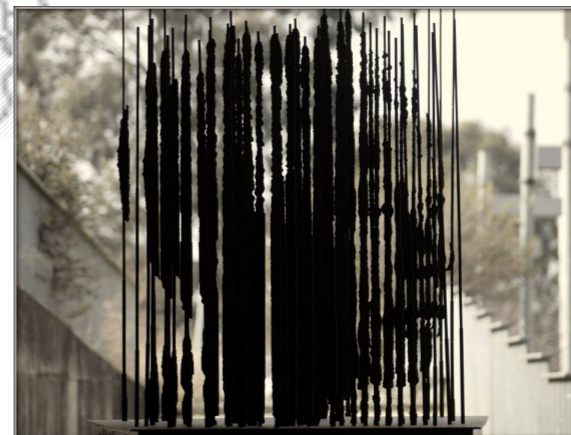
**Dresden, Germany**  
2009-2011



**Calcutta, India**  
till 2003



**Geneva,  
Switzerland**  
2011-2012



**Johannesburg, SA**  
2015 - 2



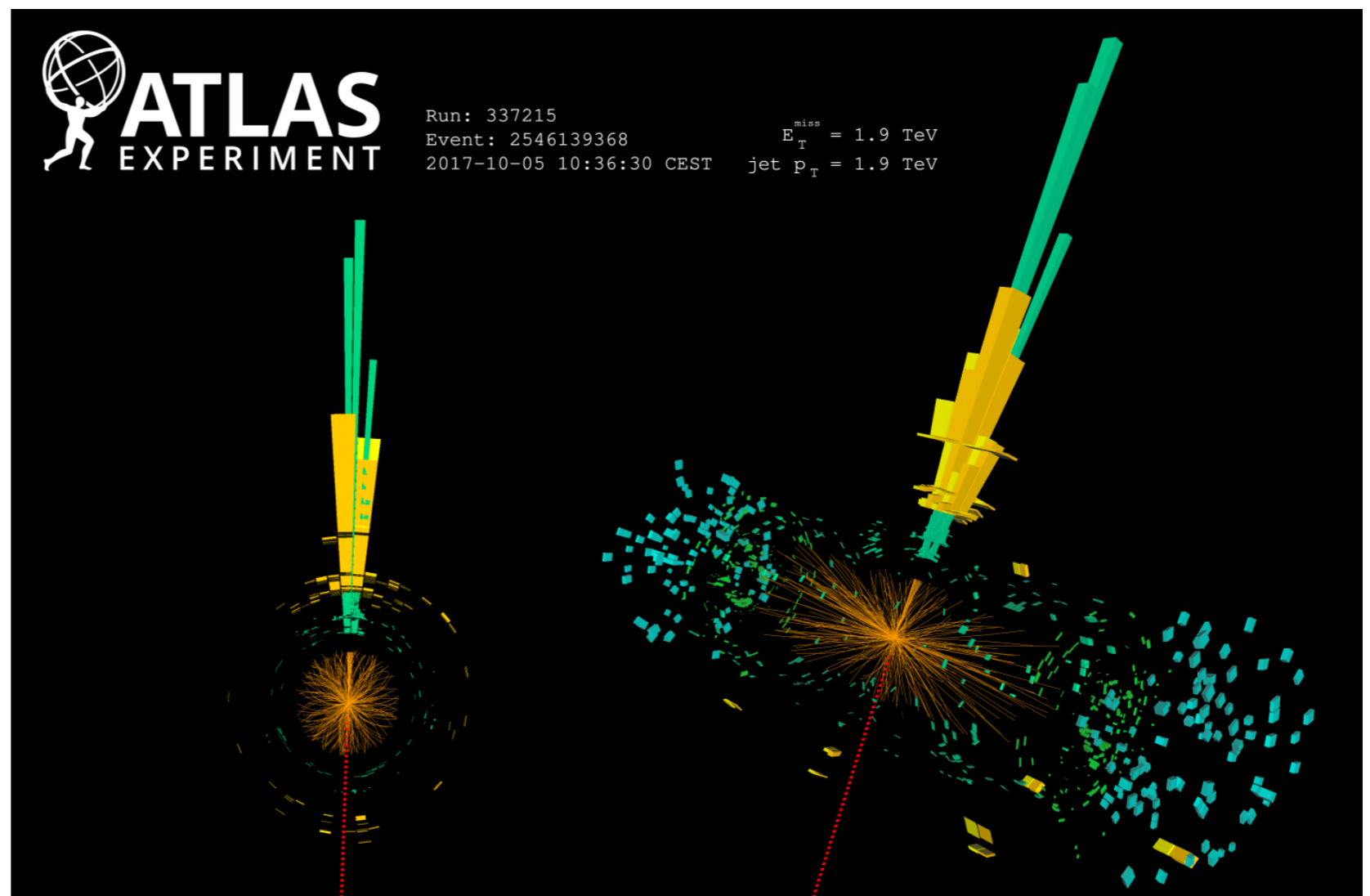
# Semi-visible jets Se

Why this is novel?

So far, almost all dark matter searches in colliders are for WIMPs

So called mono-X signatures, X being any SM particle or object.

Large MET on one side!

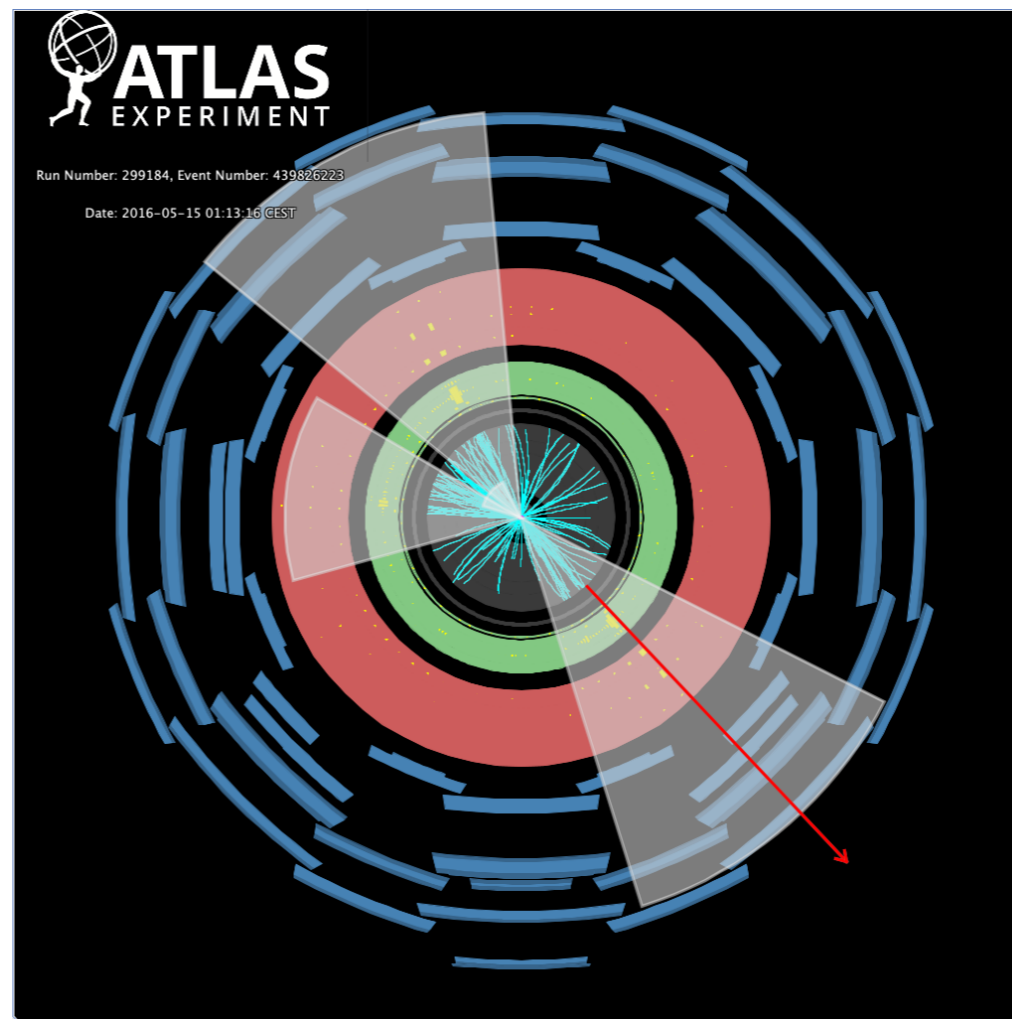


# Semi-visible jets Se

Why this is novel?

So far, almost all dark matter searches in colliders are for WIMPs

We are looking for  
SIMP, where the dark  
sector is considered  
A replica of QCD



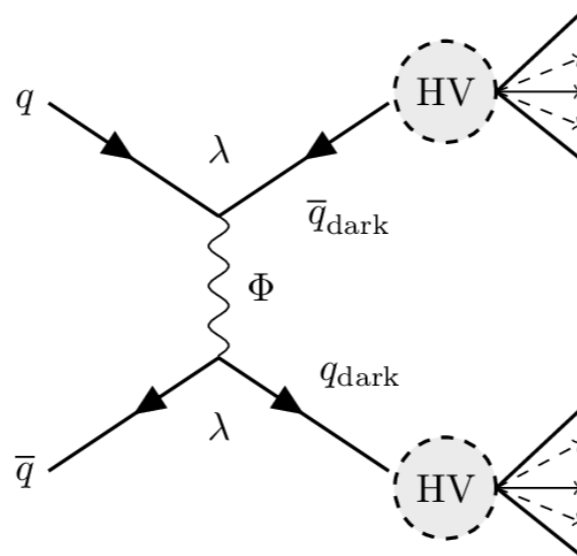


# Semi-visible jets Search

Strongly interacting dark sector:  
bifundamental mediator acts as a portal!

Ratio of the rate of stable dark hadrons over the total number of hadrons in the event is termed  $R_{inv}$

Simulated in Pythia  
Hidden Valley Module



Results in jets interspersed with dark hadrons, with missing transverse momentum direction aligned with one of the SVJs in leading order. Not so for events with extra jets and large boost.

Events with two central jets, MET trigger, leading jet  $p_T > 250$  GeV,  $H_T > 600$  GeV, MET  $600 > \text{GeV}$ , jet closest to MET with  $\Delta\Phi < 2$

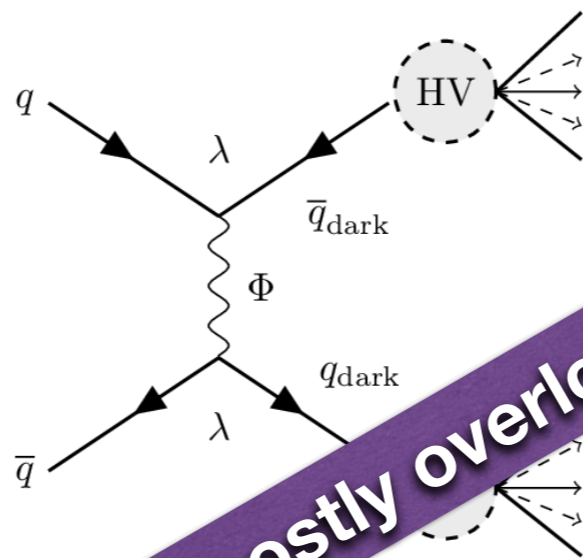
Define: SR (muon veto), and three CRs, 1L, 1L1B, 2L<sub>5</sub> (with muons and b-tagged jets)

# Semi-visible jets Search

Strongly interacting dark sector:  
bifundamental mediator acts as a portal!

Ratio of the rate of stable dark hadrons over the total number of hadrons in the event is termed  $R_{\text{had}}$

Simulated in Pythia  
Hidden Valley Module



Results in jets  
interpersed with dark  
hadrons with missing  
transverse momentum  
direction aligned with  
one of the SVJs in  
leading order. Not so  
for events with extra  
jets and large boost.

**Unique collider topology - mostly overlooked in searches!**

Events with two central jets, MET trigger, leading jet  $p_{T} > 250$  GeV,  $H_{T} > 600$  GeV, MET  $600 > \text{GeV}$ , jet closest to MET with  $\Delta\Phi < 2$

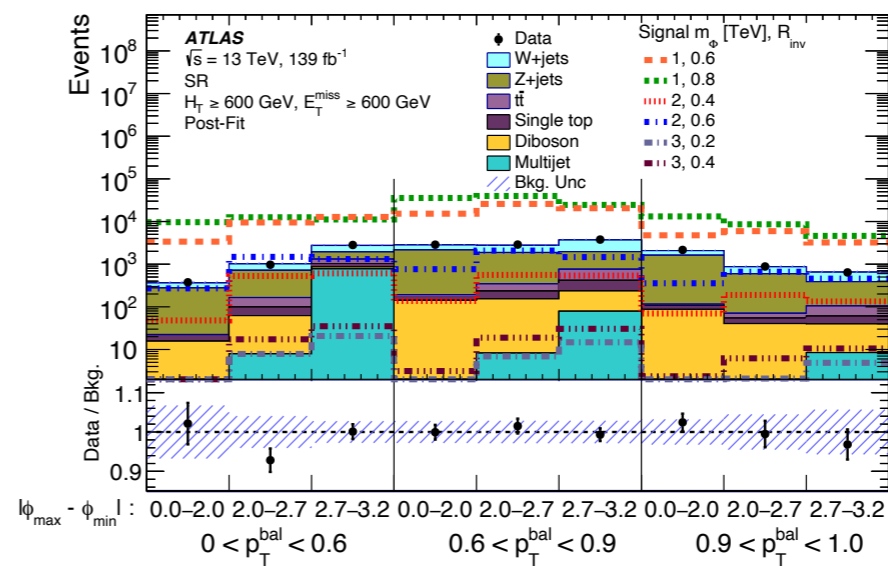
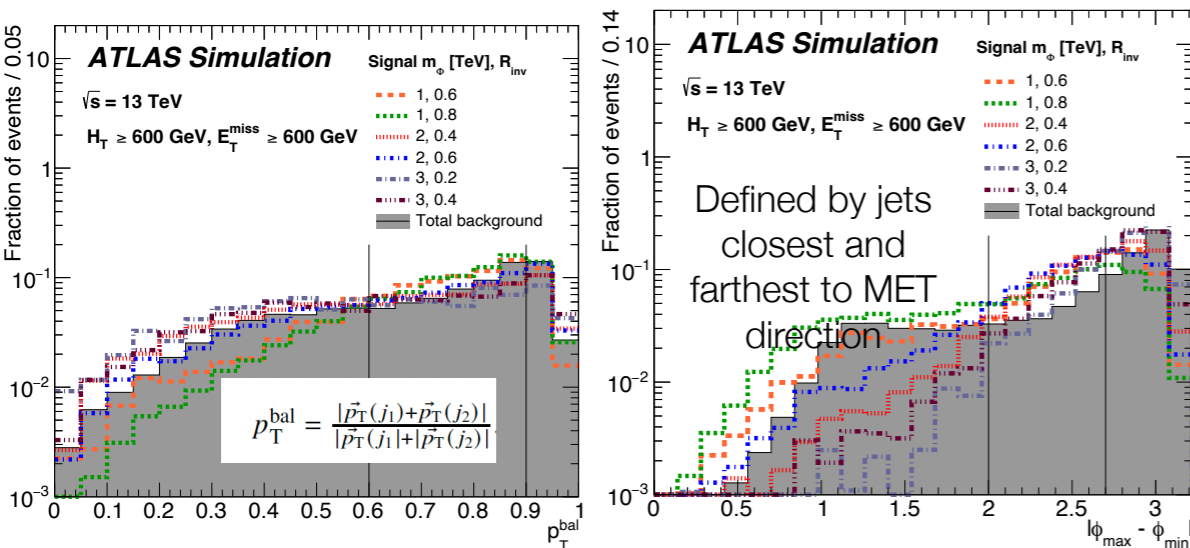
Define: SR (muon veto), and three CRs, 1L, 1L1B, 2L<sub>6</sub> (with muons and b-tagged jets)



# Background Estimate

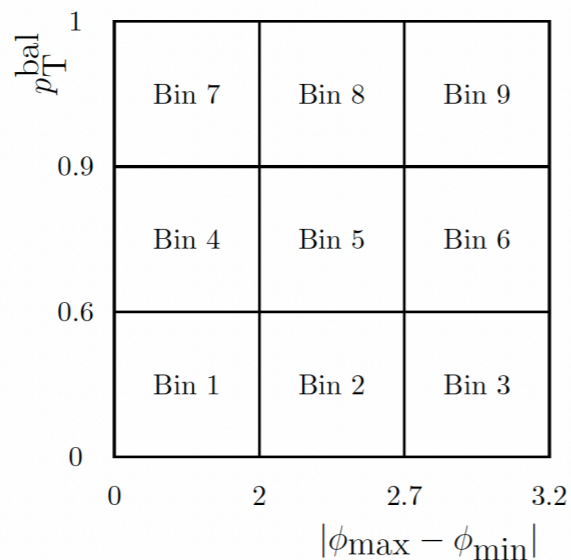
Two sensitive observables:

Partially data-driven method, simultaneously fit SR and three CRs to obtain scale factors for each bg process:



Process	$k^{\text{SF}}$
Z+jets	$1.18 \pm 0.05$
W+jets	$1.09 \pm 0.04$
Top processes	$0.64 \pm 0.04$
Multijet	$1.10 \pm 0.04$

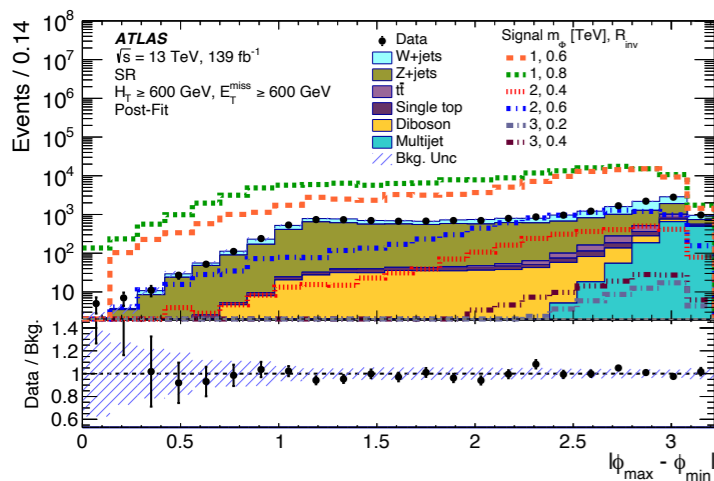
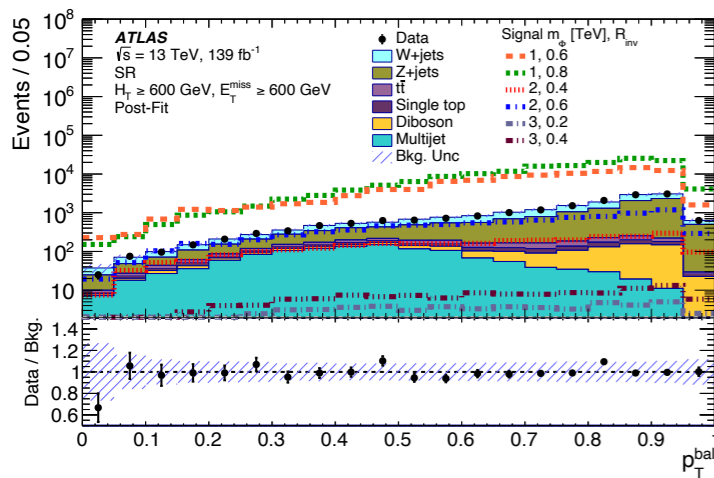
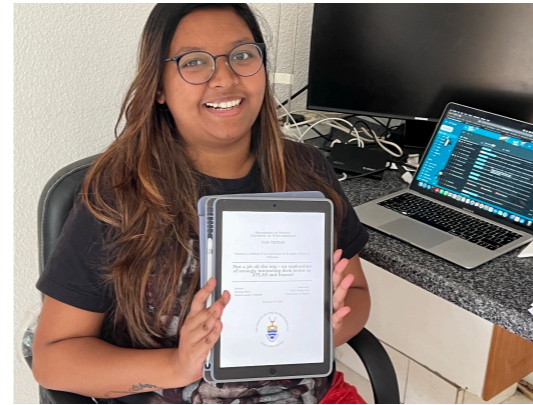
Used to Form a 9-bin grid, with yields in each bin treated as observables:



Absence of signal, good post-fit agreement :(

Multijet reweighed in using a dedicated VR given by MET within 250 to 300 GeV, then fitted

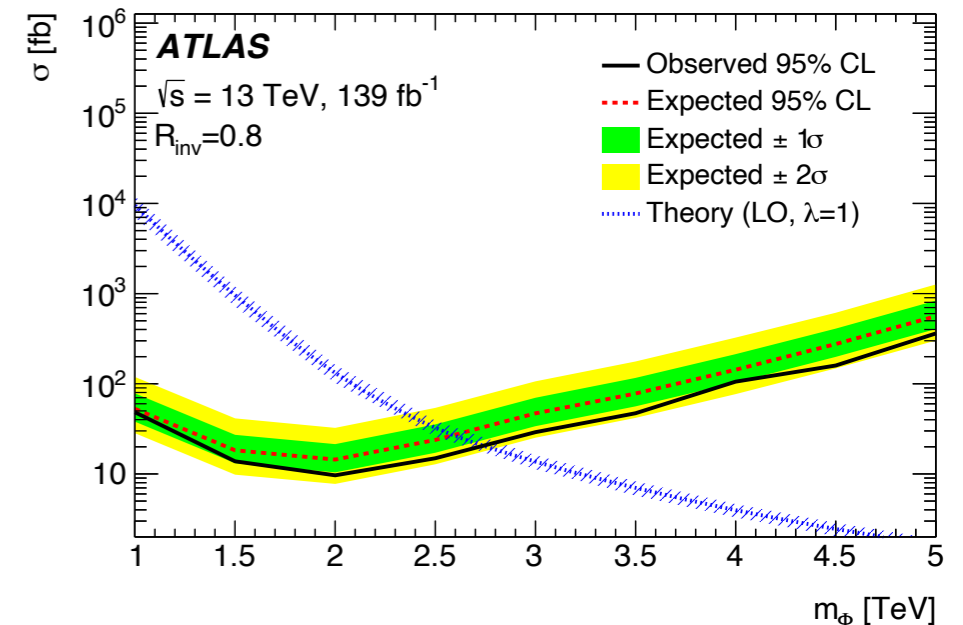
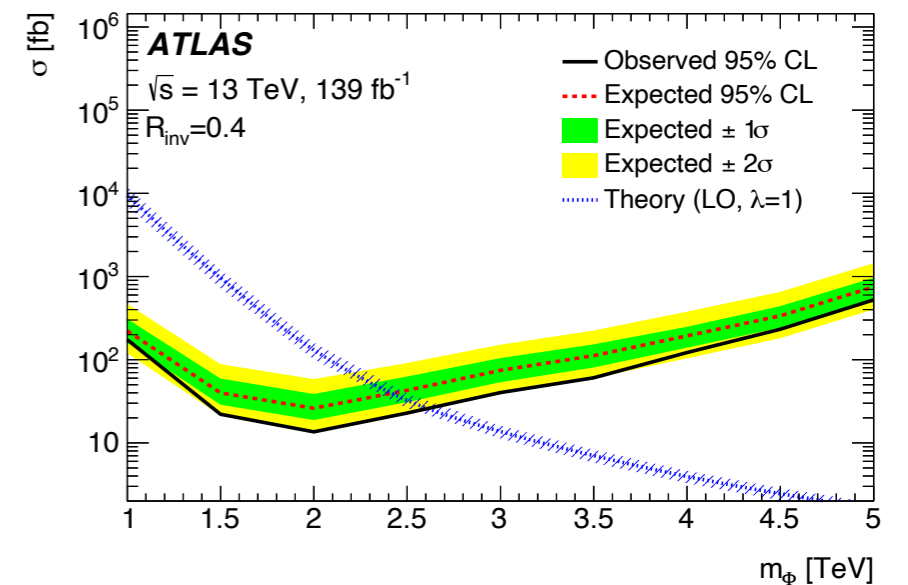
# Results



Excellent agreement between data and background prediction

Limits on mediator mass separately for each  $R_{\text{inv}}$

For mediator mass of 2.5 TeV or higher can also express the limits in terms of the  $q$ - $q_d$ - $\phi$  vertex coupling strength  $\lambda$ , with the XS scaling as  $\lambda^4$

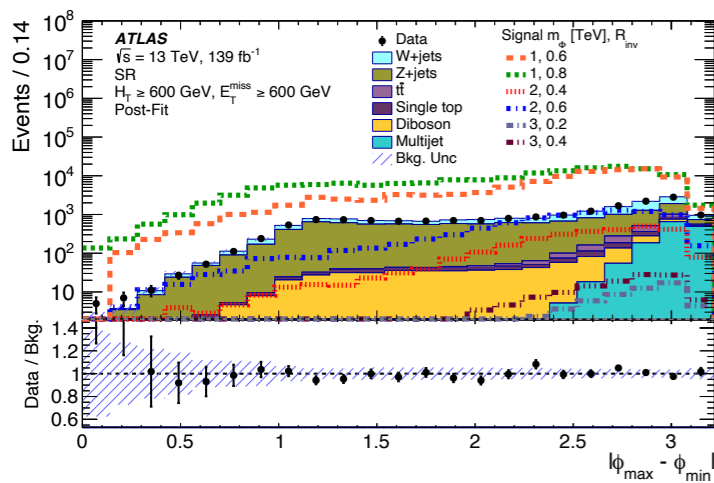
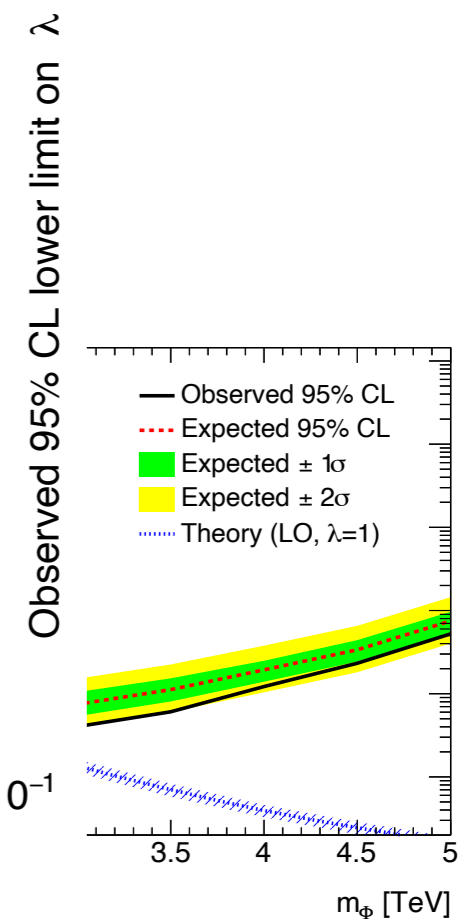
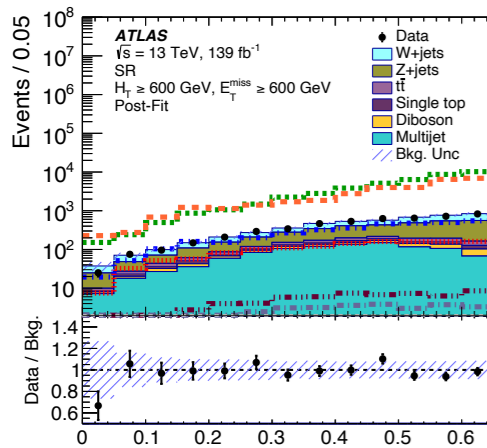
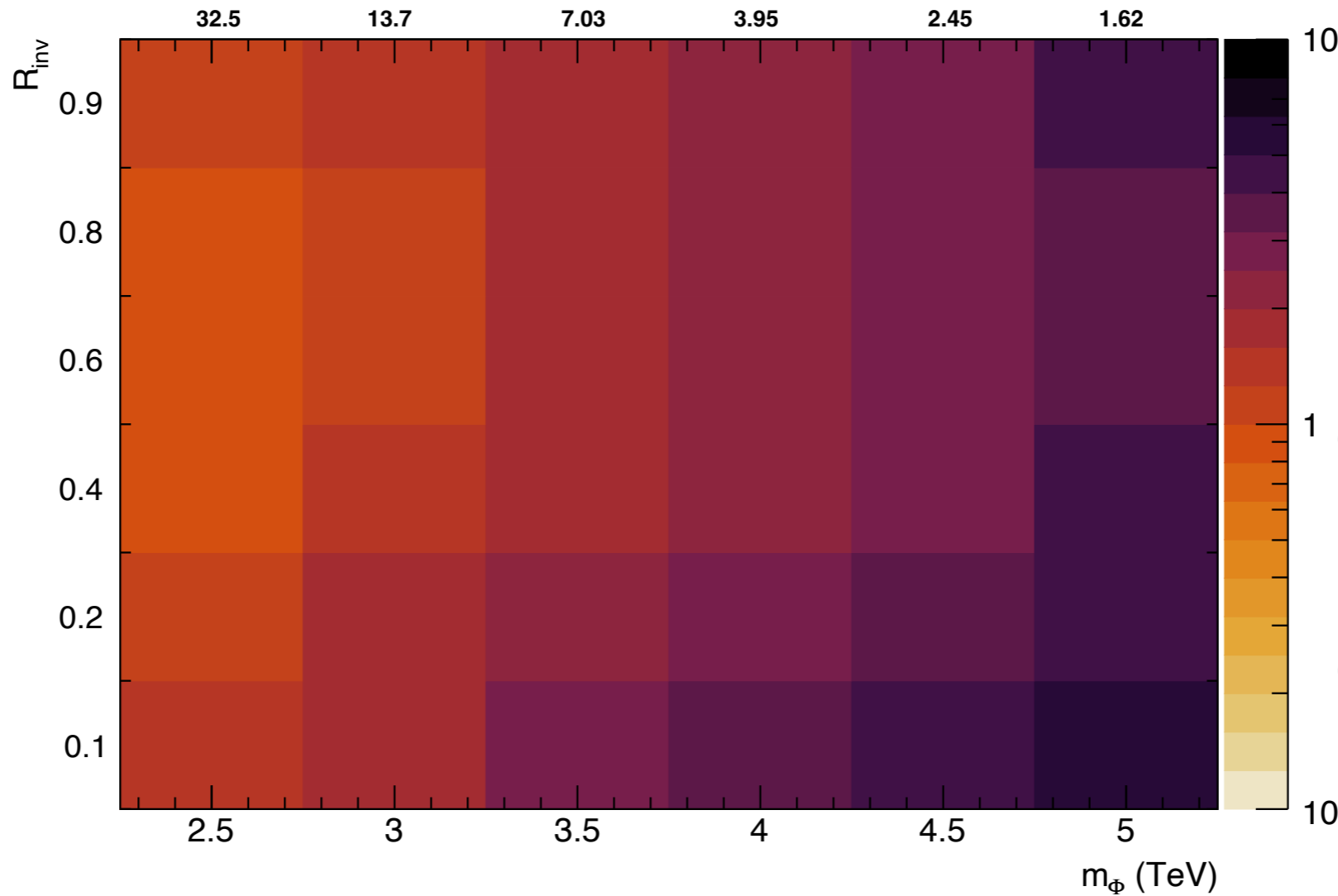




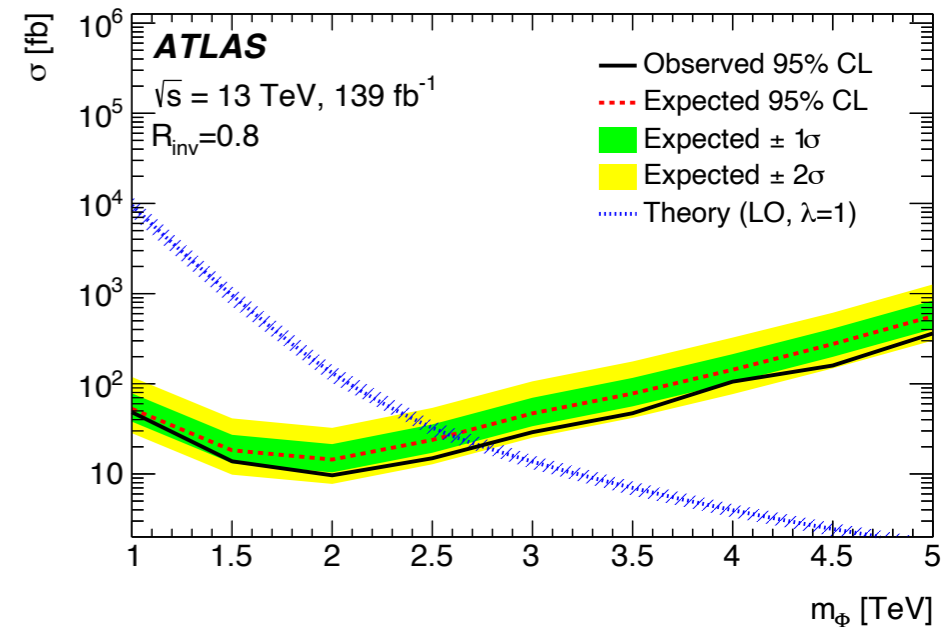
ATLAS  $\sqrt{s} = 13 \text{ TeV}, 139 \text{ fb}^{-1}$

Cross-section for  $\lambda=1$  (fb):

**R**



For mediator mass of 2.5 TeV or higher can also express the limits in terms of the  $q$ - $q_d$ - $\phi$  vertex coupling strength  $\lambda$ , with the XS scaling as  $\lambda^4$



# **Session 1:**

# **Setting the Scene/Objects**



# Poll

**When was “Machine Learning” first used in (experimental) Particle Physics?**

- A. Duh, during the Higgs boson discovery.
- B. Must be at Tevatron, say mid-2000's, its always the Americans.
- C. Before I was born. Like way before that.

# Poll

**When was “Machine Learning” first used in (experimental) Particle Physics?**

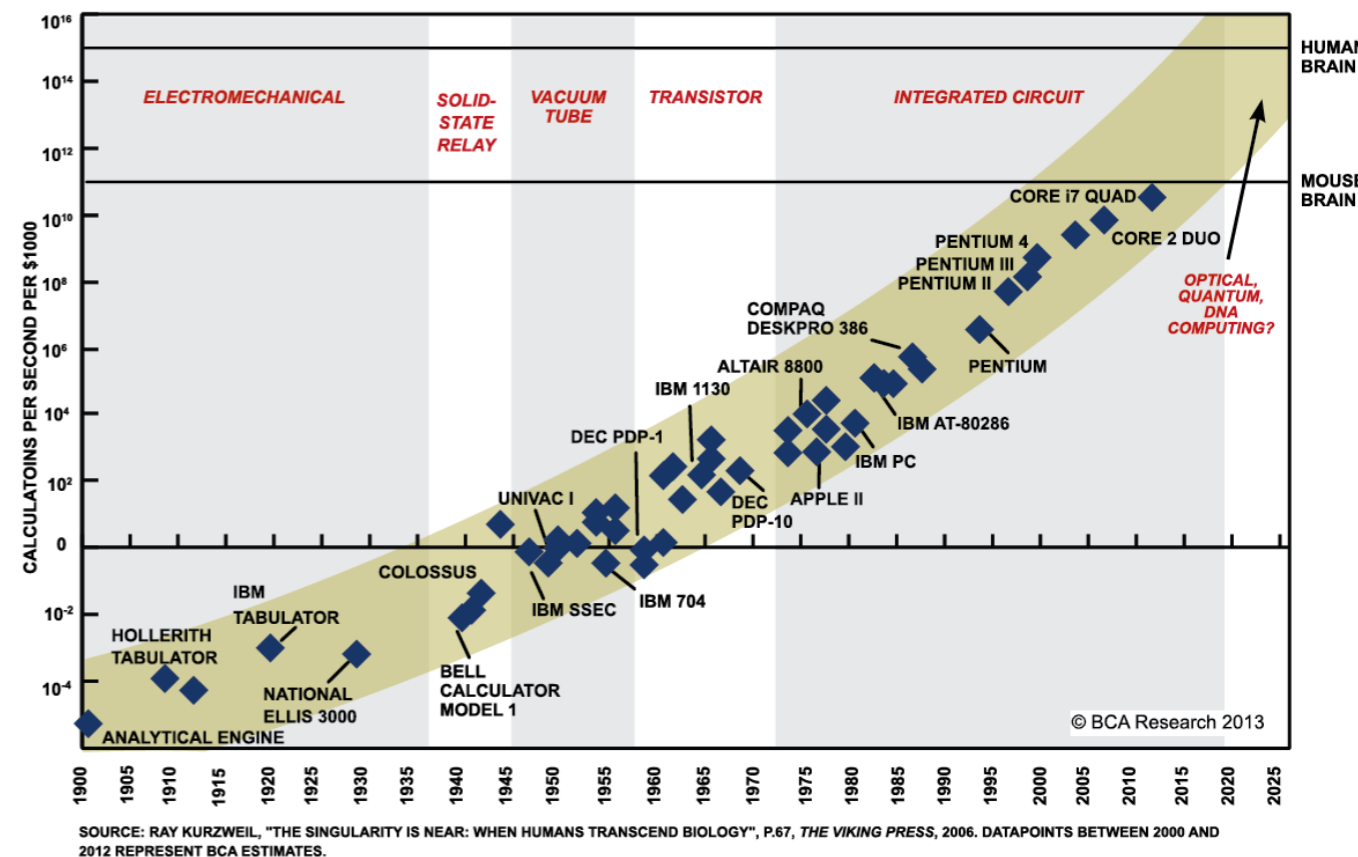
- A. Duh, during the Higgs boson discovery.
- B. Must be at Tevatron, say mid-2000's, its always the Americans.
- **C. Before I was born. Like way before that.**

# Machine Learning: Prologue

Moore's law: the number of transistors that can be packed into a given unit of space will roughly double every two years.

=

computing power tends to approximately double every two years



# But ...

Search for the neutral Higgs bosons of the MSSM  
in  $e^+e^-$  collisions at  $\sqrt{s}$  from 130 to 172 GeV

The ALEPH Collaboration\*)

20 June 1997

## Abstract

The process  $e^+e^- \rightarrow hA$  is used to search for the Higgs bosons of the Minimal Super-symmetric Standard Model. The search is performed in the data sample at centre-of-mass energies between 130 and 172 GeV. A total of 10.9 events are found in either side of the signal region. This results in a 95% C.L. upper limit on the cross-section  $\sigma(e^+e^- \rightarrow hA)$  for  $\tan \beta > 1$ .

The network architecture is multilayer feed-forward, consisting of four layers and is based upon the JETNET 3.4 package [12]. Detailed descriptions of theoretical aspects of neural networks are available elsewhere [13]. The neural network was trained, with the backward propagation method, using both b and non-b jets in radiative returns to the Z from a sample of 400,000 Monte Carlo  $q\bar{q}$  events generated at a centre-of-mass energy of 161 GeV. Radiative returns to the Z were used because the jets in such events are produced in a kinematic configuration similar to that of the signal; this was preferred to training the network using simulated signal events in order to reduce the associated systematic error in the signal efficiency.

### 3.2 b tagging neural network

Six variables which discriminate between b jets and light quark jets are combined using neural networks to tag b quark jets. The first two variables are lifetime-based; the third is based

on the transverse momentum of identified leptons and the last three are based on jet-shape properties. The quantities used are as follows:

1.  $\mathcal{P}_{\text{jet}}$ : probability of the jet being a light quark (uds) jet based upon impact parameters of tracks in the jet, similar to that described in Ref. [9] with modifications for the new VDET;
2.  $\Delta\chi_{\text{svx}}^2$ : the  $\chi^2$  difference between fitting tracks in the jet both to secondary and primary vertices compared to assuming all tracks come from the interaction point. This is based upon a secondary vertex pattern recognition algorithm which searches for displaced



# Digging Deeper ...



LU TP 93-29  
CERN-TH.7135/94  
December 1993

## JETNET 3.0 – A Versatile Artificial Neural Network Package

Carsten Peterson and Thorsteinn Rögnvaldsson

Department of Theoretical Physics, University of Lund,  
Sölvegatan 14 A, S-223 62 Lund, Sweden

Leif Lönnblad

### CERN's Know-How

- Particle physicists were among the first to use machine learning (ML)
- First AI HENP seminar in 1990
- Already in 2010, the CMS and LHCb experiments successfully introduced a neural network based trigger system
- Higgs boson discovery earlier than expected (2012), also with help of ML

1993 CERN SCHOOL OF COMPUTING

Scuola Superiore G. Reiss Romoli, L'Aquila, Italy  
12-25 September 1993

### Neural Networks

*S.R.Amendolia*  
University of Sassari and INFN of Pisa, Italy

#### Abstract

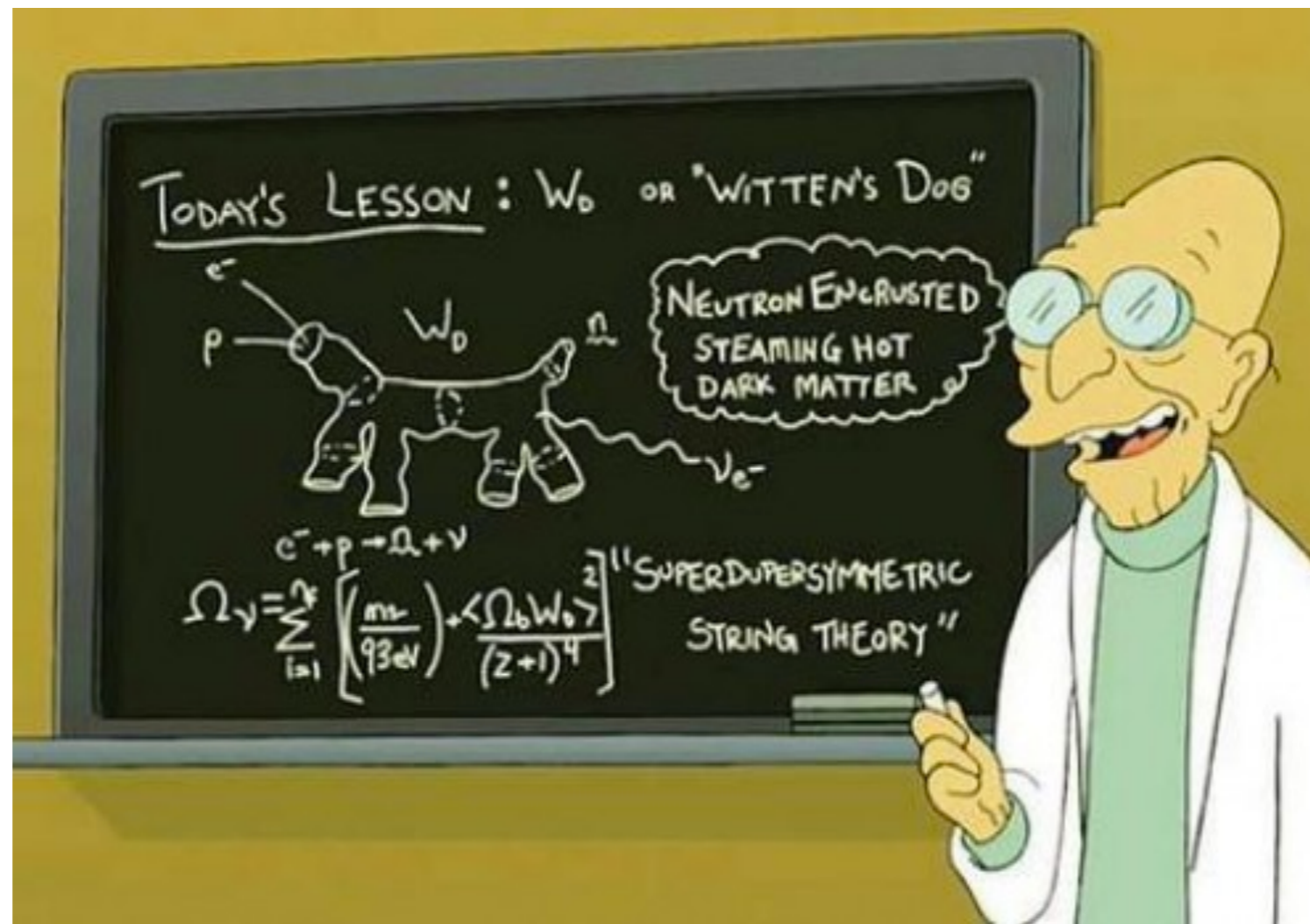
An introductory treatment of the subject of Neural Networks will be given. Topics covered will mostly be relevant for the use of Neural Networks in High Energy Physics, especially for triggering, and examples will be given of this application.

#### 1. Basics of Neural Networks

##### 1.1 Introduction

There exist many papers and books on the subject of Neural Networks in the literature, and this cannot be an attempt at making a better or more comprehensive treatment. We send to the relevant references [1,2,3,4,5]. The difficulty in this work, compared to

# What you do/want to do?





# What we do/want to do



# **Steps from a model/calculation to experimental (non) observation**

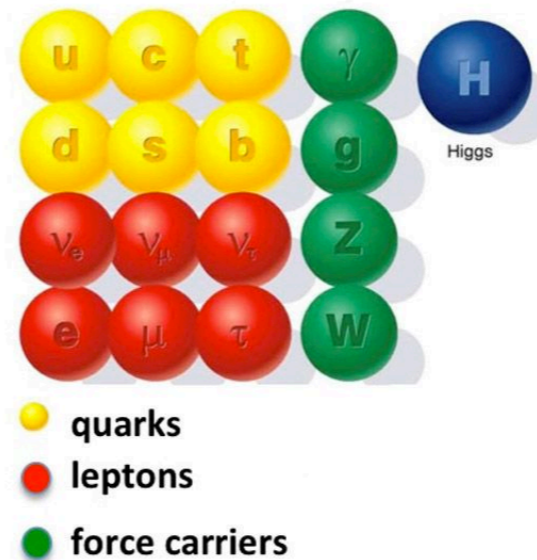


**Searches: two broad  
(overlapping) categories ...**

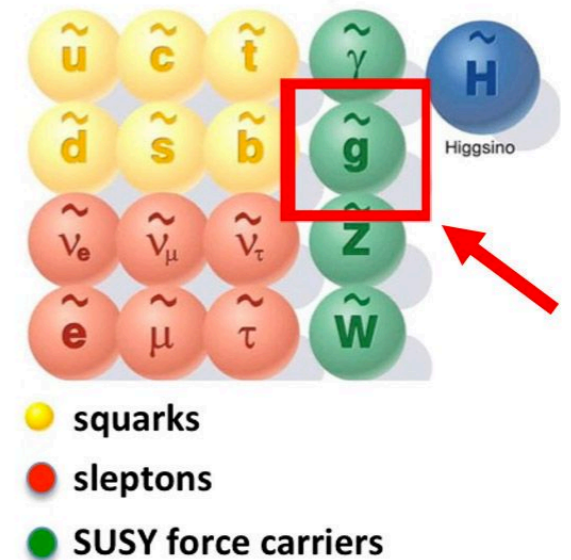
# Theoretical model driven

- SUSY
- UED incl QBH
- Compositeness
- LRSM

The known world of Standard Model particles



The hypothetical world of SUSY particles

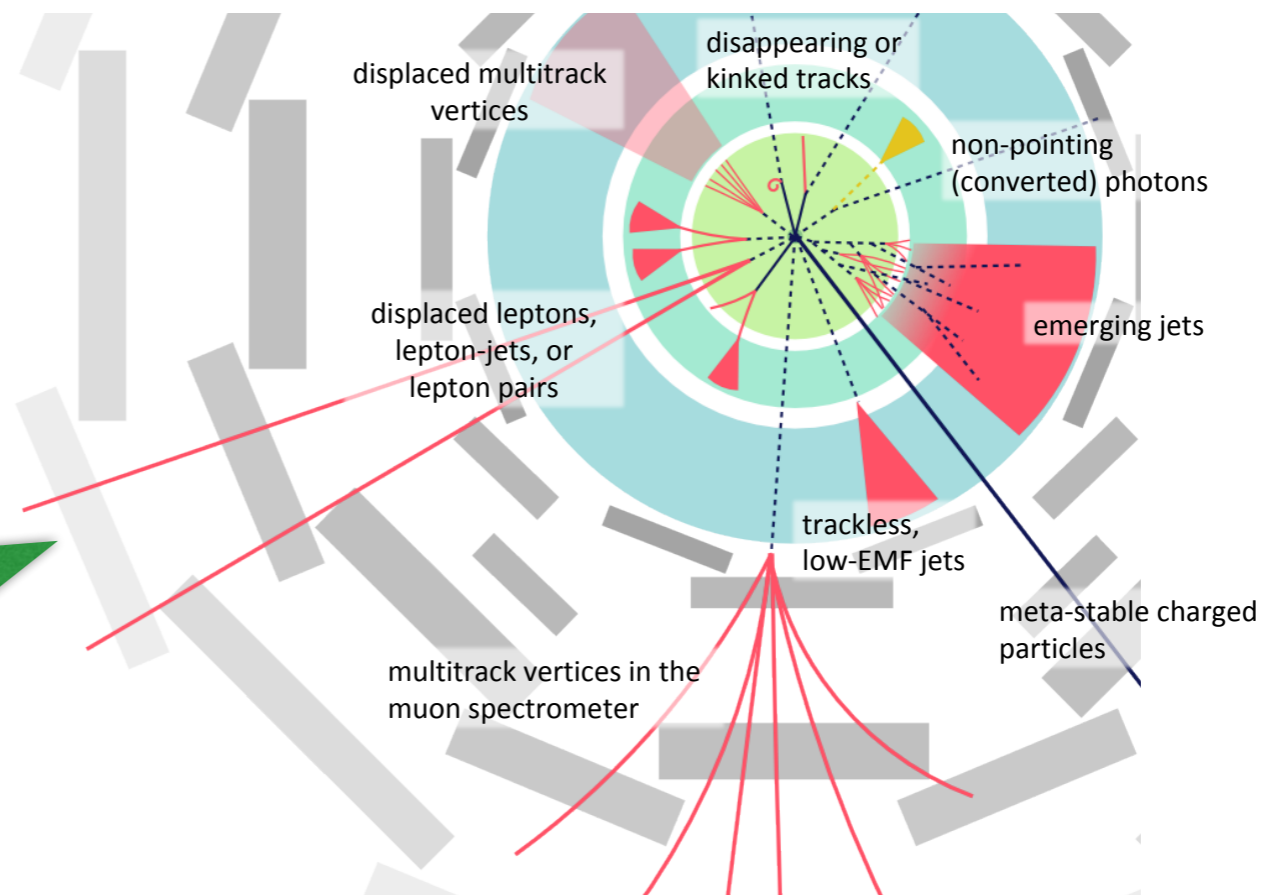


We can only exclude based on what our detectors see!

# Phenomenological model/signature driven

- DM/WIMP: *mono-everything*
- Dark Photon
- Extended Higgs sector: 2HDM
- 4th generator quarks/top partners
- Leptoquarks
- Heavy  $W'$  or  $Z'$
- Diboson resonances
- LFV
- LLP

+ SIMP!



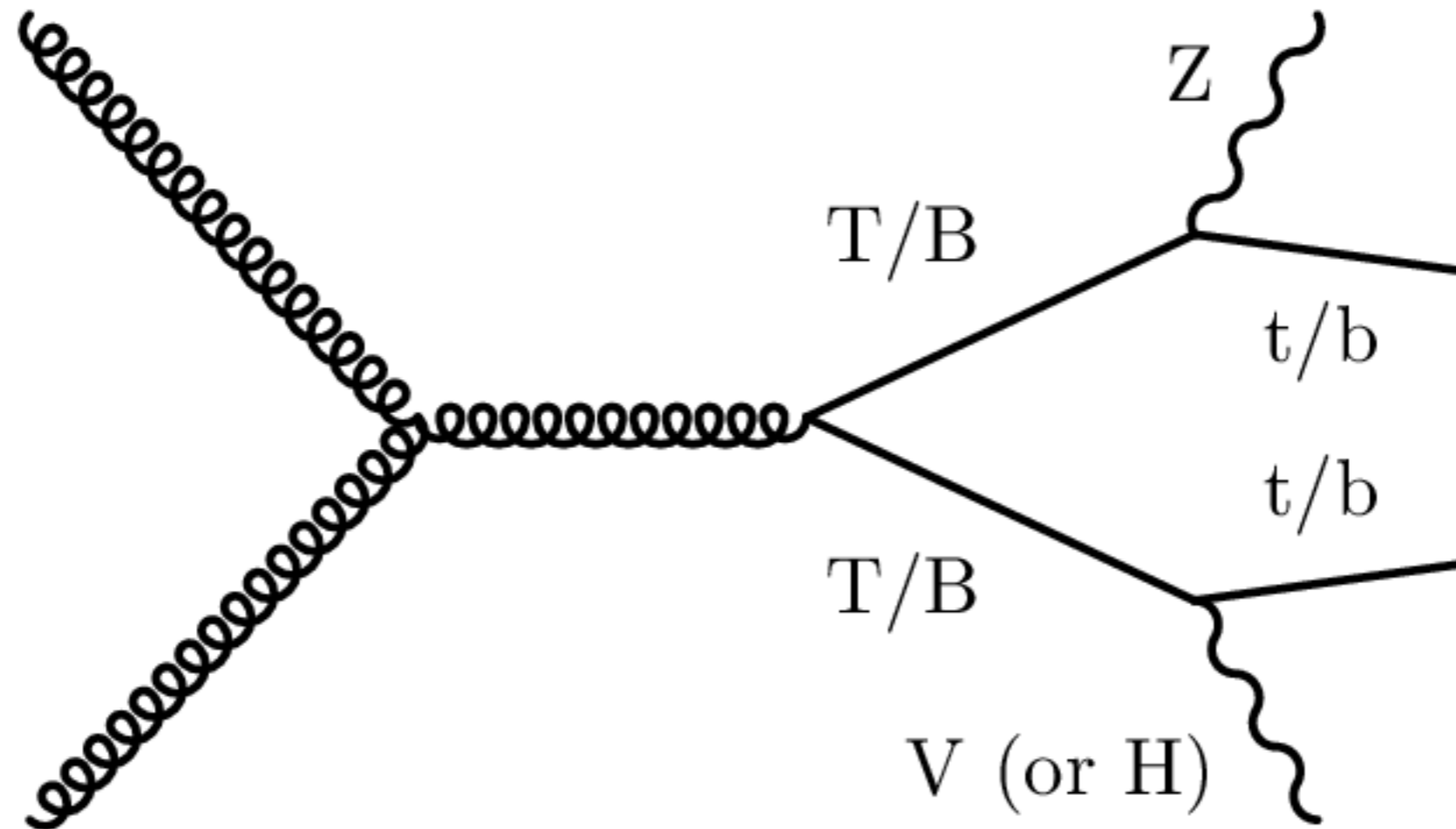
**Let's start with an  
example:**

<https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/EXOT-2018-58/>



# Search for:

Pair  
production of  
VLQs

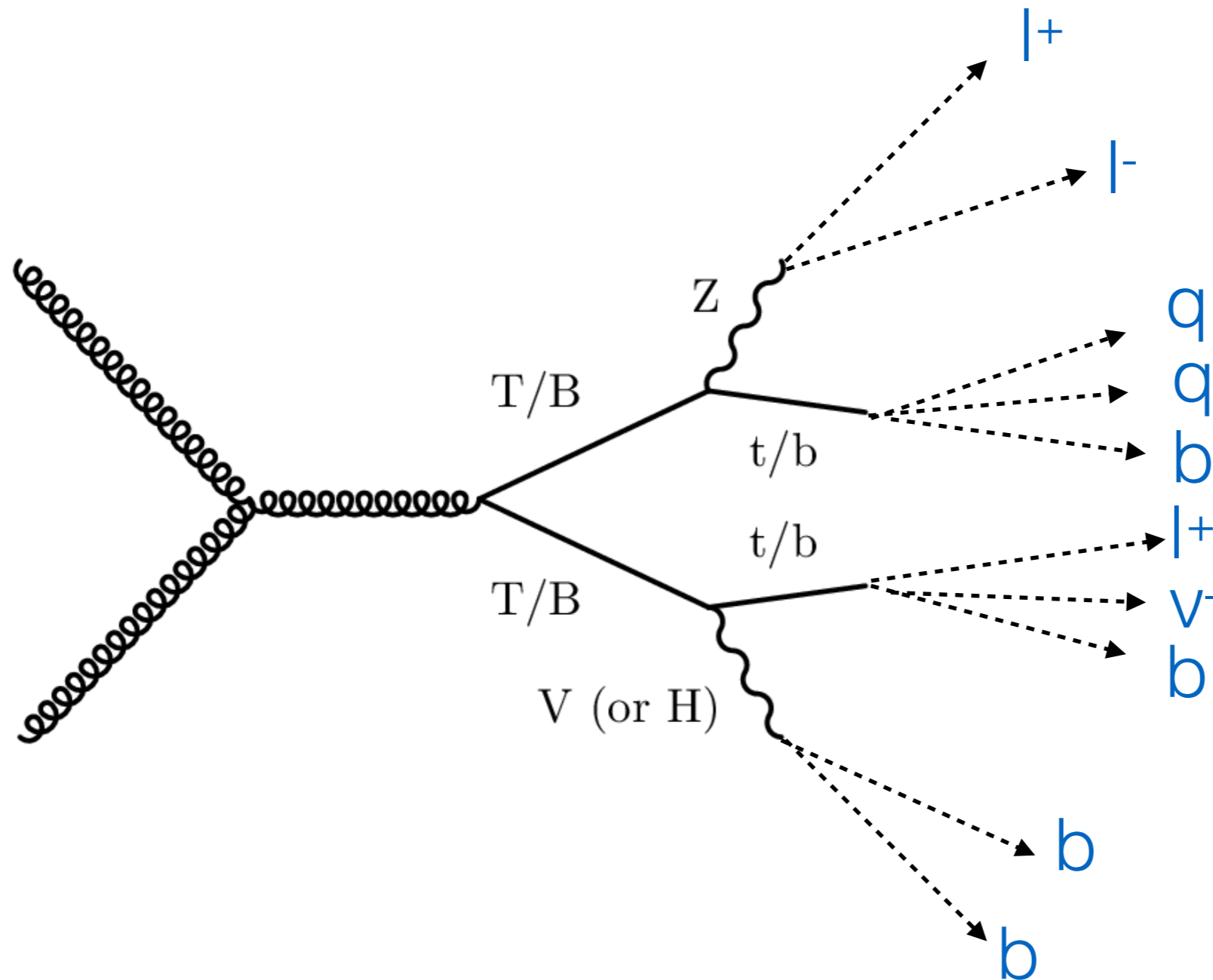


**Question: what would we see in the detector?**

# What you see is not what you have!

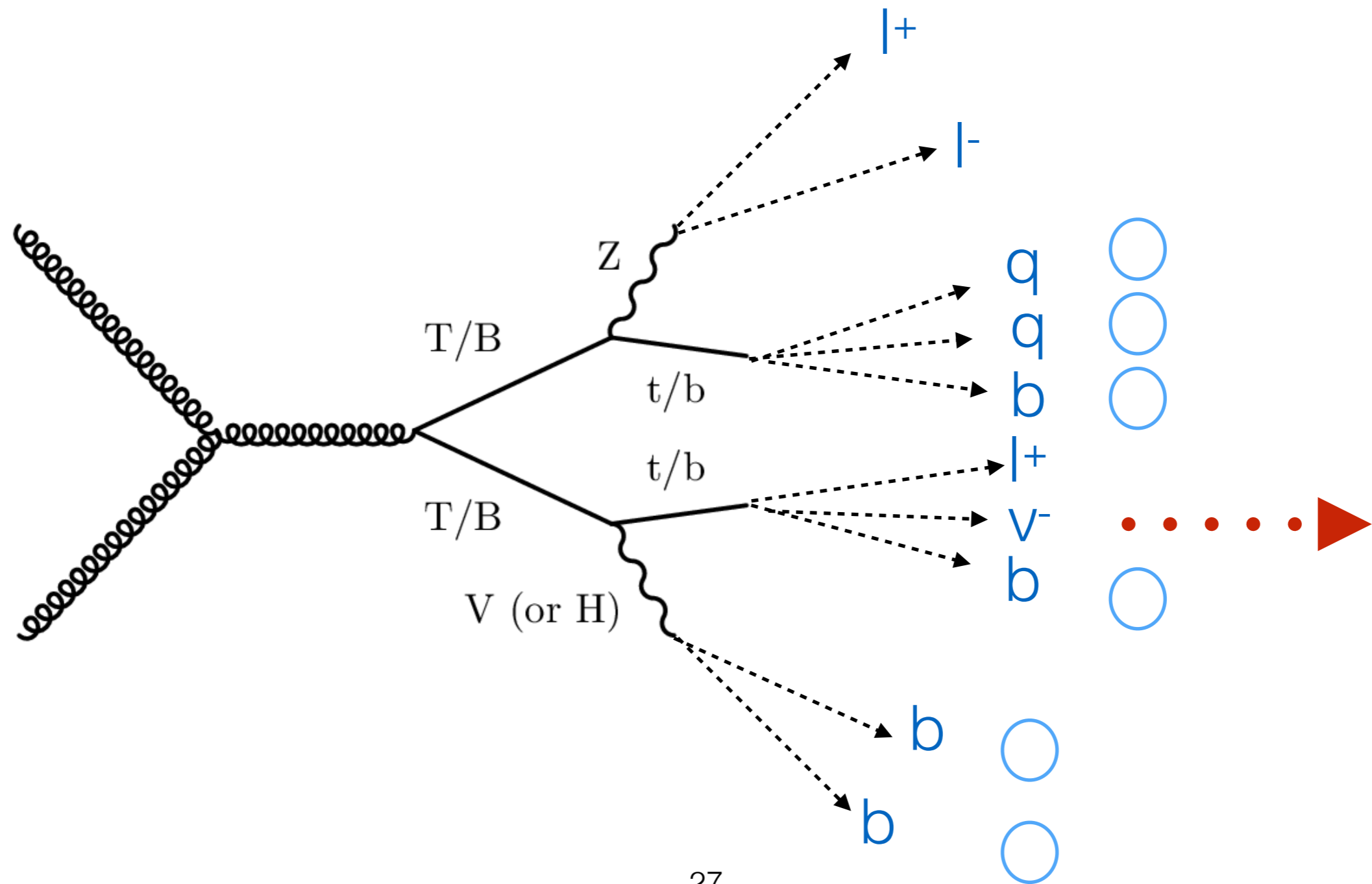
- We don't get what is coming out of the collisions.
- Finite lifetime of particles, decays before reaching the detector.
- Detectors have finite resolution, less than perfect response and efficiency.

# Decays:

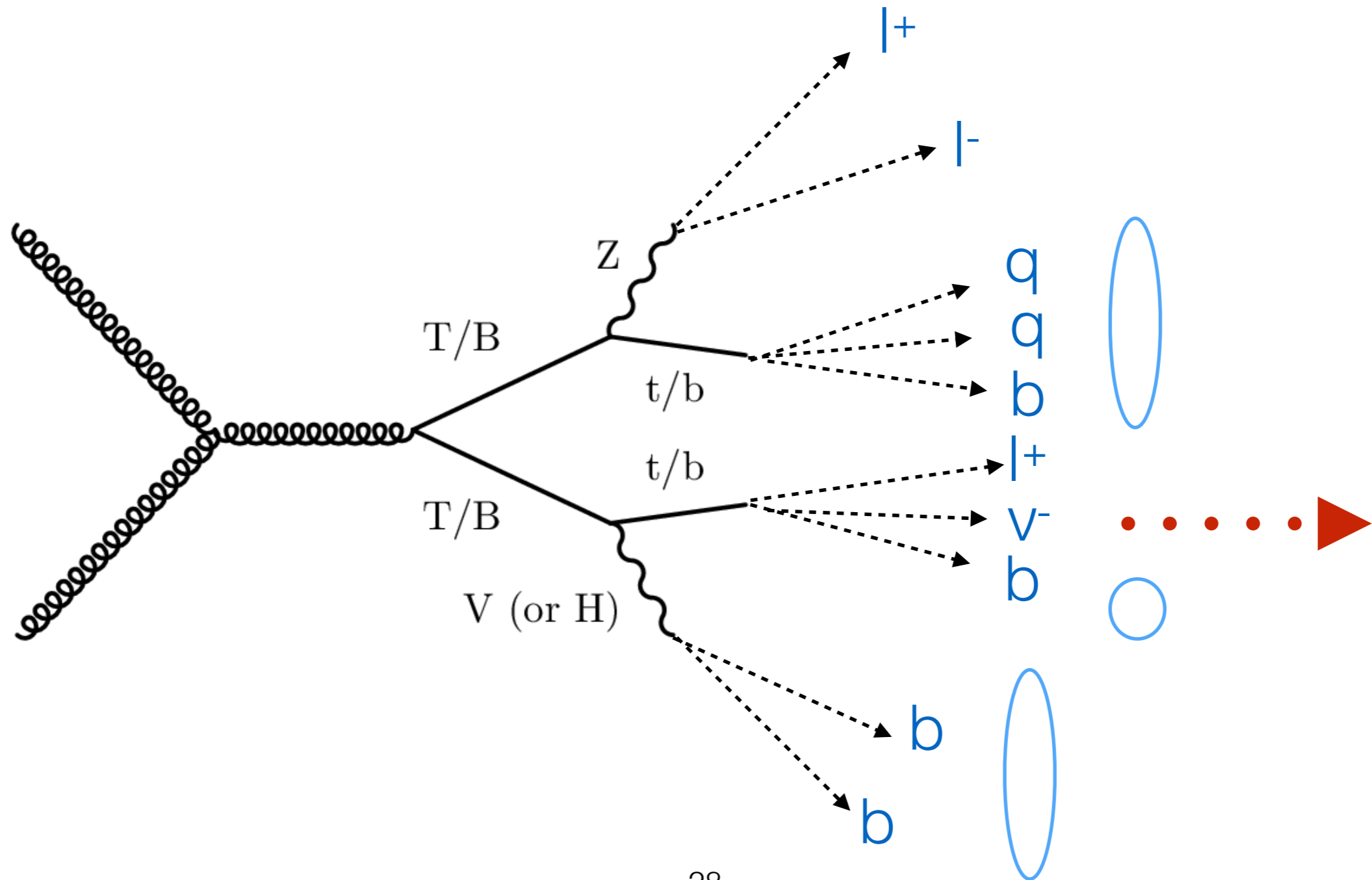




# Decays: Detector Objects



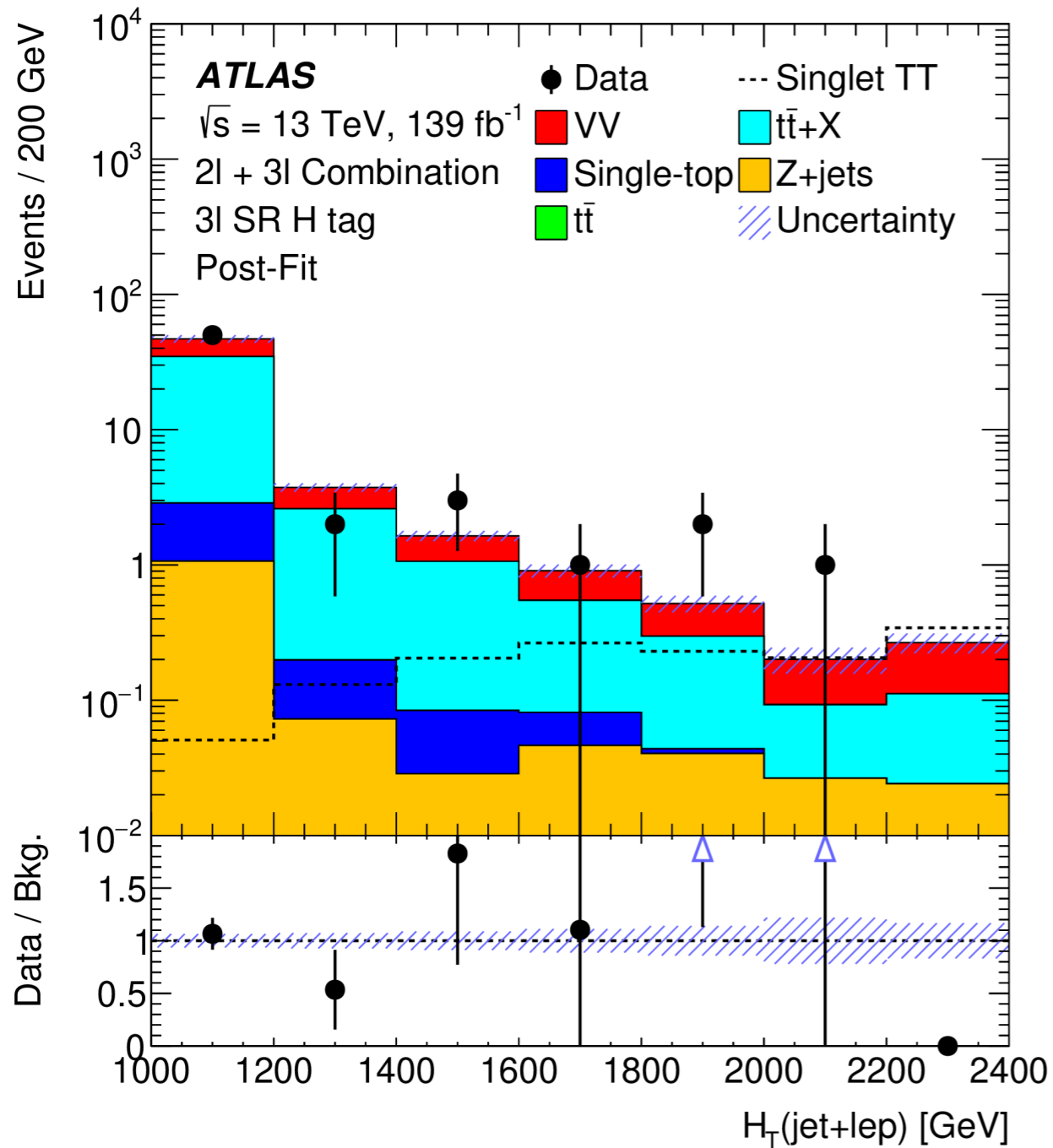
# Decays: Detector Objects



# Details

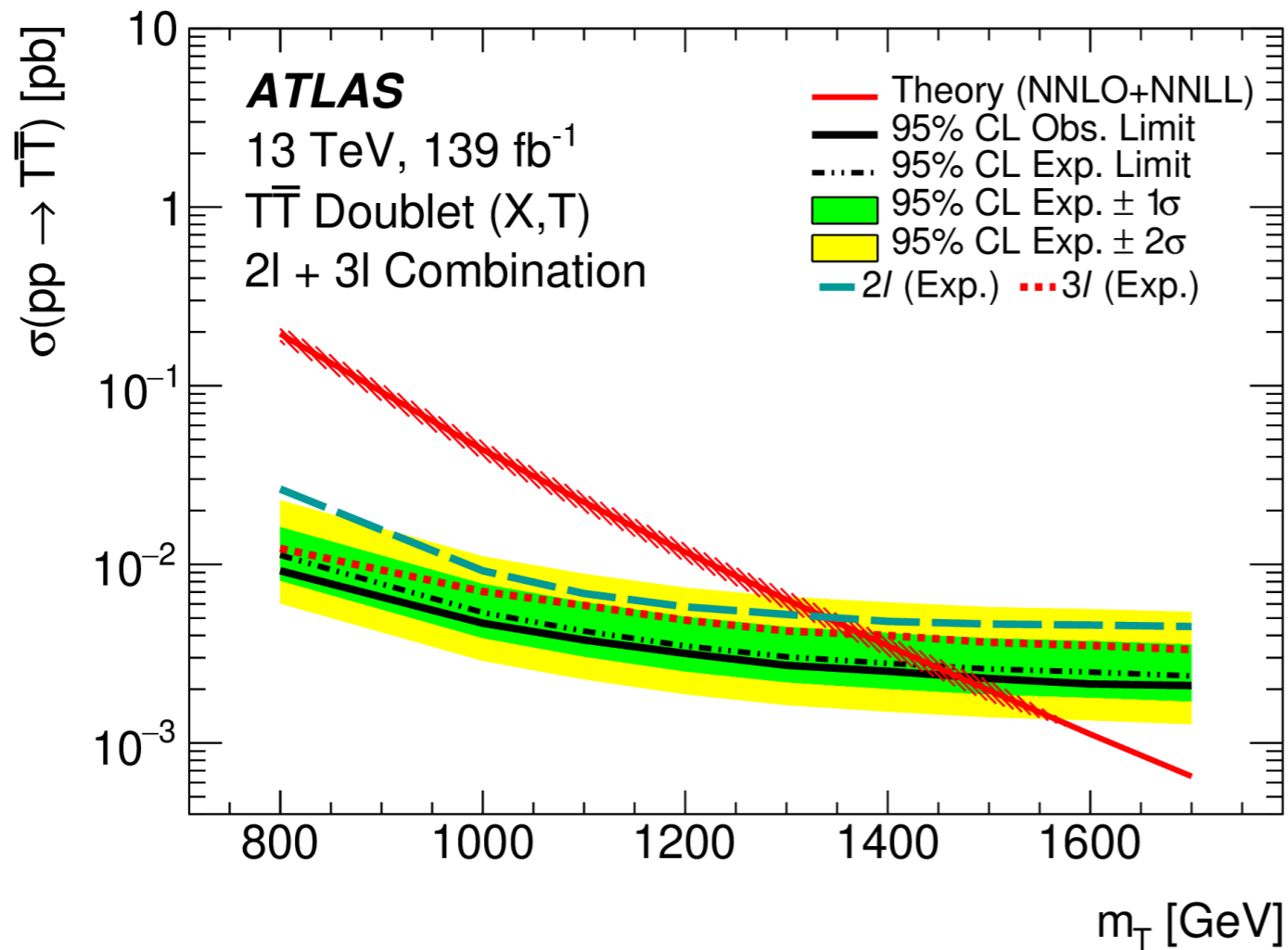
Preselection	$\geq 2$ central jets at least two SF leptons with $p_T > 28$ GeV at least one pair of OS-SF leptons $ m(\ell\ell) - m_Z  < 10$ GeV					
Channel definitions	$2\ell$ $= 2\ell$ $p_T(\ell\ell) > 300$ GeV $H_T(\text{jet}) + E_T^{\text{miss}} > 920$ GeV				$3\ell$ $\geq 3\ell$ $p_T(\ell\ell) > 200$ GeV $H_T(\text{jet} + \text{lep}) > 300$ GeV	
Region definitions	$1b$ SR $H_T(\text{jet}) + E_T^{\text{miss}} > 1380$ GeV $= 1$ $b$ -jet	$2b$ SR $H_T(\text{jet}) + E_T^{\text{miss}} > 1380$ GeV $\geq 2$ $b$ -jet	$1b$ CR $H_T(\text{jet}) + E_T^{\text{miss}} < 1380$ GeV $= 1$ $b$ -jet	$2b$ CR $H_T(\text{jet}) + E_T^{\text{miss}} < 1380$ GeV $\geq 2$ $b$ -jet	SR – $\geq 1$ $b$ -jet	VV CR – $= 0$ $b$ -jet
MCBOT categories	7	7	–	–	5	–
Fitted variable	$m(Zb_1)$	$m(Zb_2)$	$H_T(\text{jet}) + E_T^{\text{miss}}$		$H_T(\text{jet} + \text{lep})$	

# From this:



Construct observables,  
 Signal and Control regions,  
 estimate background...

# From this:



Discover or set limits: When signature of a new model is not found, the model is excluded up to a certain parameter value

Presented in terms of 95% confidence level, which the associated probability of that observation being correct 95% of the time. In other words, if the measurement is made repeatedly on independent datasets, the measured value will be obtained at least 95% of the time.



# Components of a limit plot

- Expected line: from MC simulation (SM), usually using same luminosity as data. The 1 and 2  $\sigma$  bands are from MC uncertainty. (Brazil plot!)
- Observed line: from data, actual number of events seen, with statistical uncertainties. It is expected to stay within the expected bands if the simulation is accurate.
- Theory line: calculated from new model, often with associated systematic uncertainties.

# Interpreting a limit plot

- As long as the expected and observed lines are below the theory prediction, the conclusion is no evidence of the new particle is seen. By this argument, the expected and observed limits are respectively 5.1 and 5.2 TeV, from where the theory prediction line intersects the expected and observed lines. If at any point, the observed line goes beyond the expected brazil-bands, that may indicate data contains more events than SM predicts. However, the threshold for an observation is  $3\sigma$  and a discovery is  $5\sigma$ . Only in the region where observed is higher than the theory line, and beyond the statistically allowed deviations from expected, this particular new model can be confirmed.

# An Aside: Why Limits?

We all want to find new physics.

# An Aside: Why Limits?

We all want to find new physics.

But out of 100 new physics models, at least 99 are wrong,  
possibly all 100 are!

So null results also tell us a lot.

And techniques/methods developed can help in a future  
discovery!

# Recap:

- Reconstruct objects from detector information
- Decide on sensitive observables for the specific final state we are interested in
- Estimate (SM) background
- Look for new physics/*measure at particle level*



One collision == EVENT

Outcome of the collisions is probabilistic, no exhaustive list of possibilities!

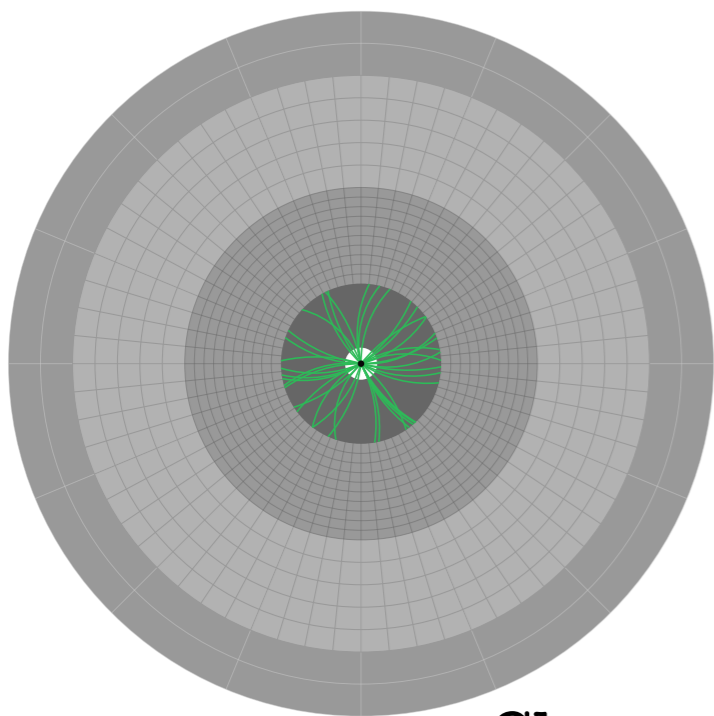
Cross section: how often a particular process occurs, measured as an effective area the target particle presents to projectile particles.

Actual number for a process:

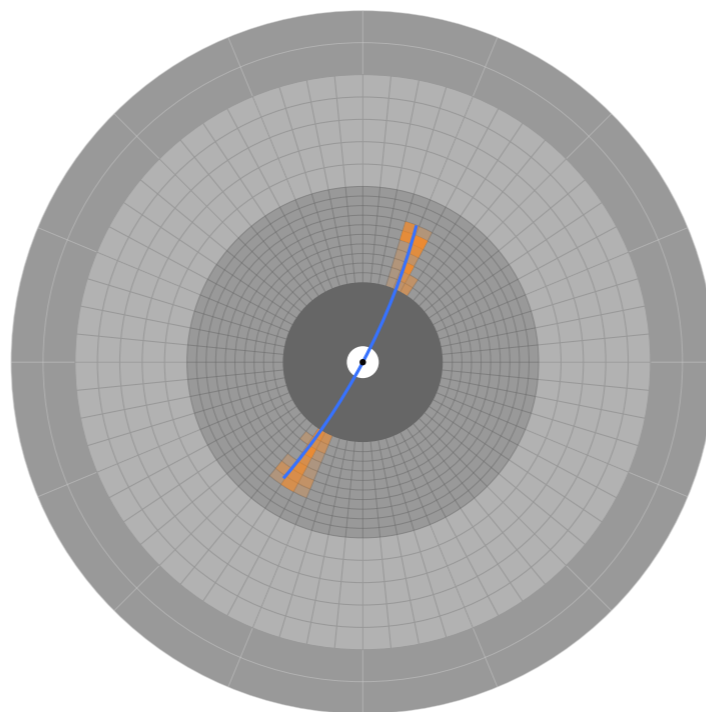
$$N_{process} = \sigma_{process} \int L dt$$

Luminosity measured in units of 1/area

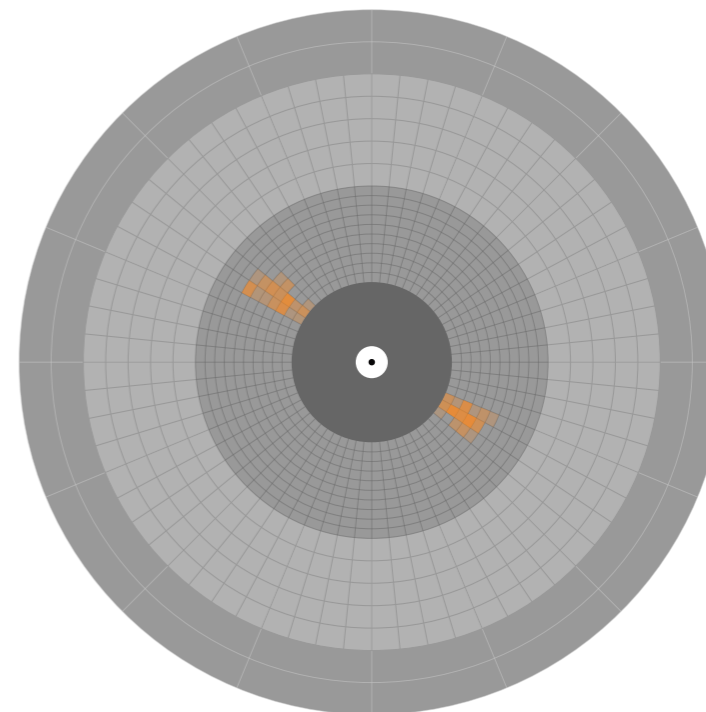
# Objects from Collisions



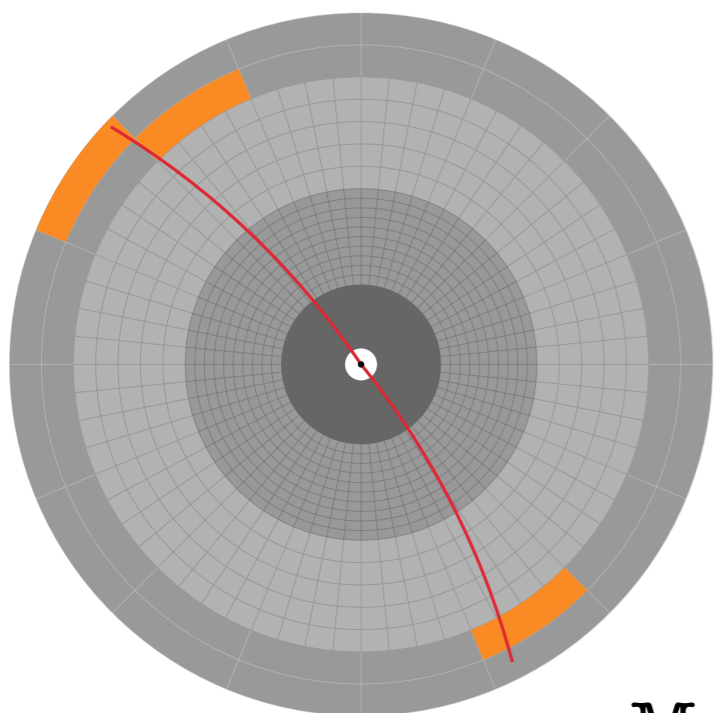
Charged particles



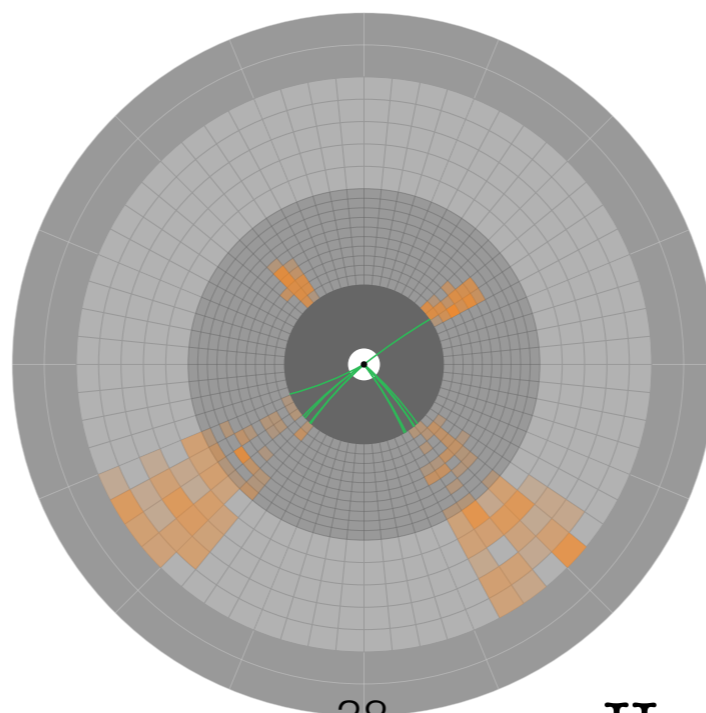
Electrons



Photons

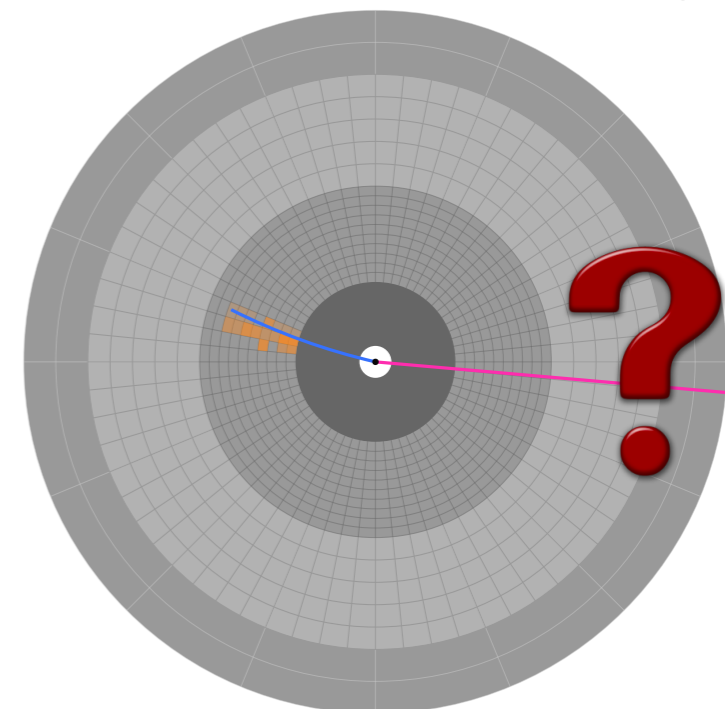


Muons



38

Hadrons



Neutrinos

# Missing (Transverse) Energy

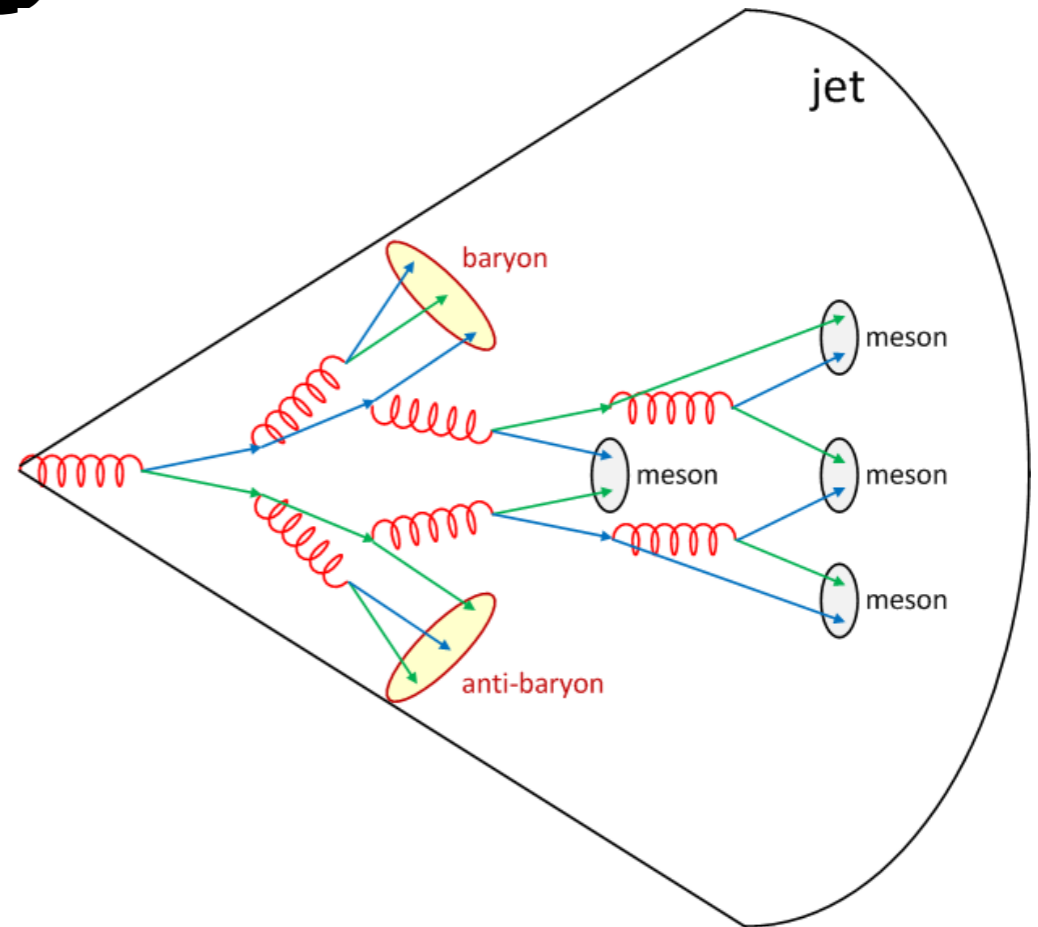
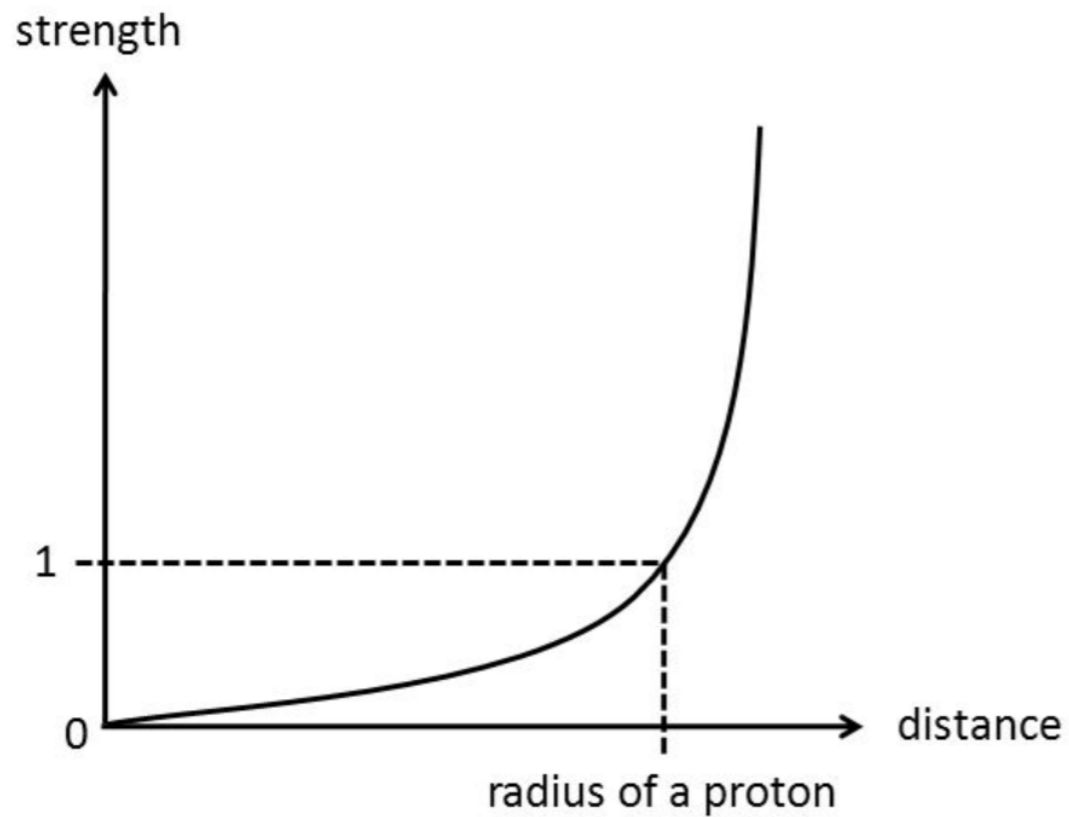


- Do not interact with the detector
- Imbalance of  $p_T$
- Can be signs of new physics as well!

Only invisible SM particles are the neutrinos.  
DM, SUSY particles have not been seen yet!



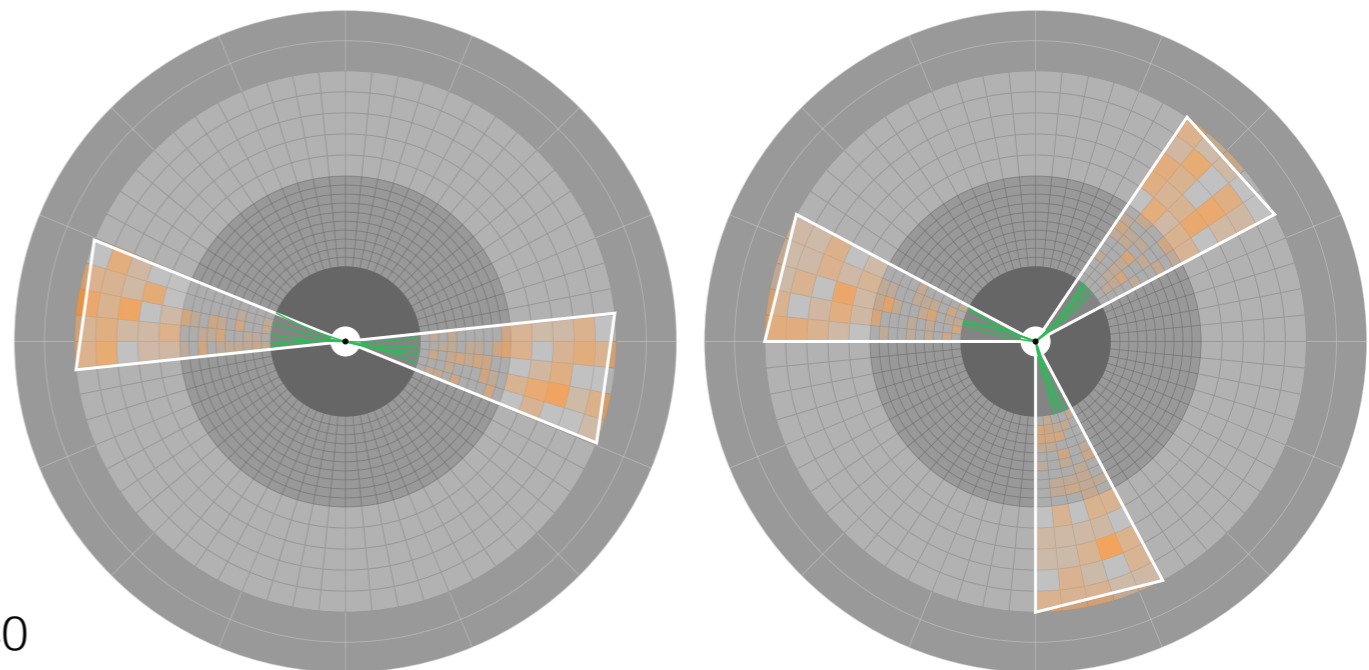
# Jets



Strong interaction works like a rubber band!

No free quarks/gluons!

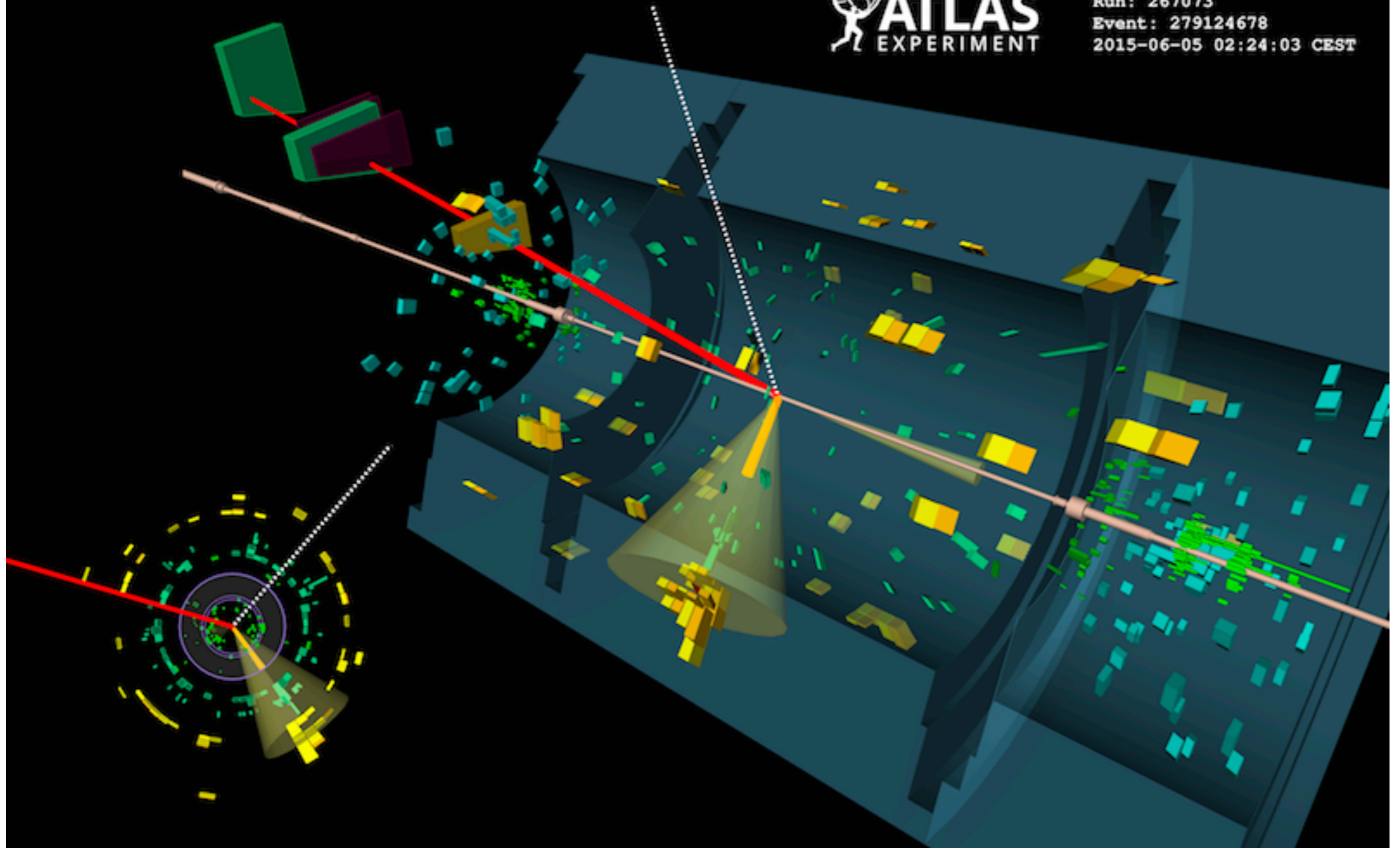
Collected in jets!







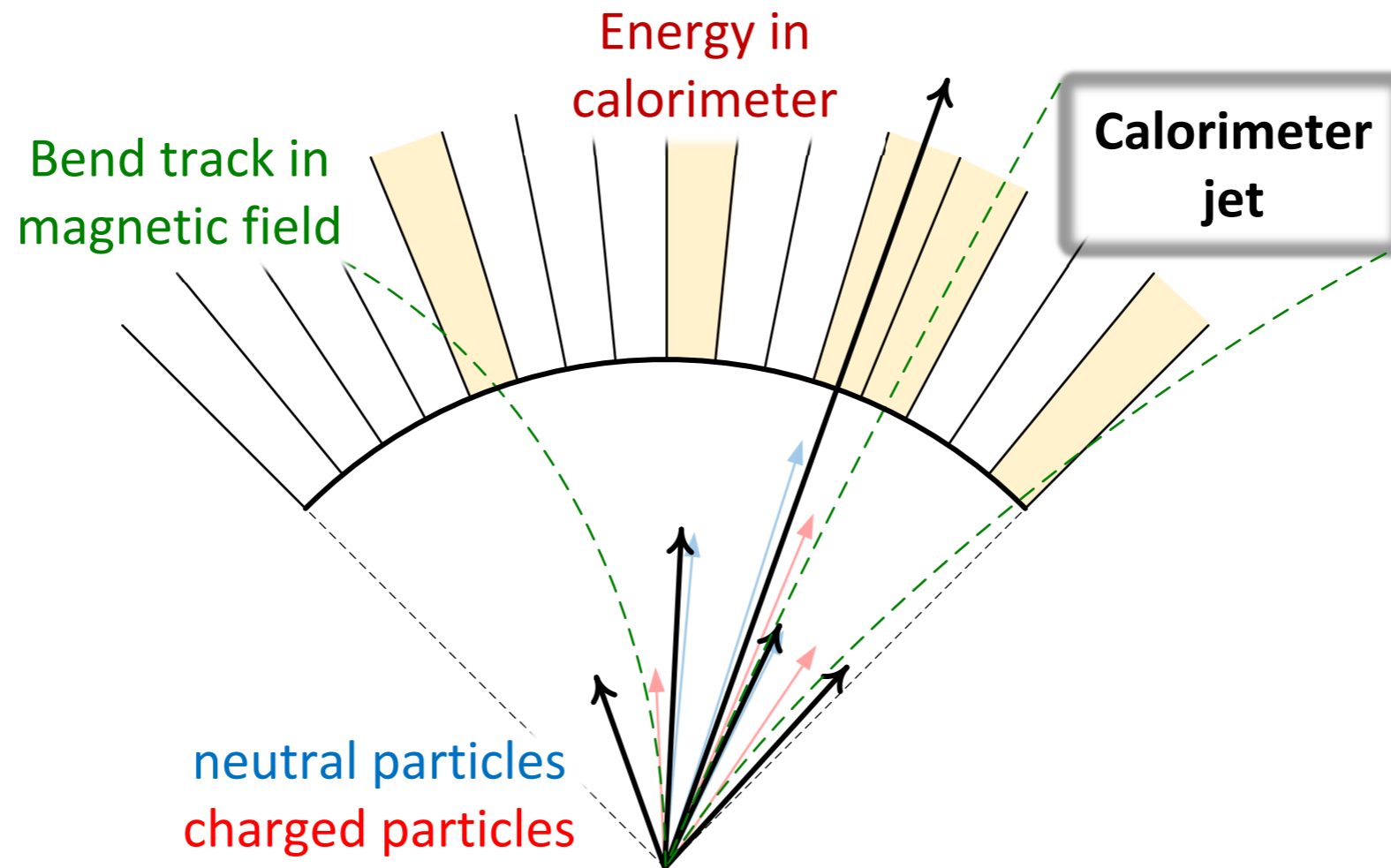




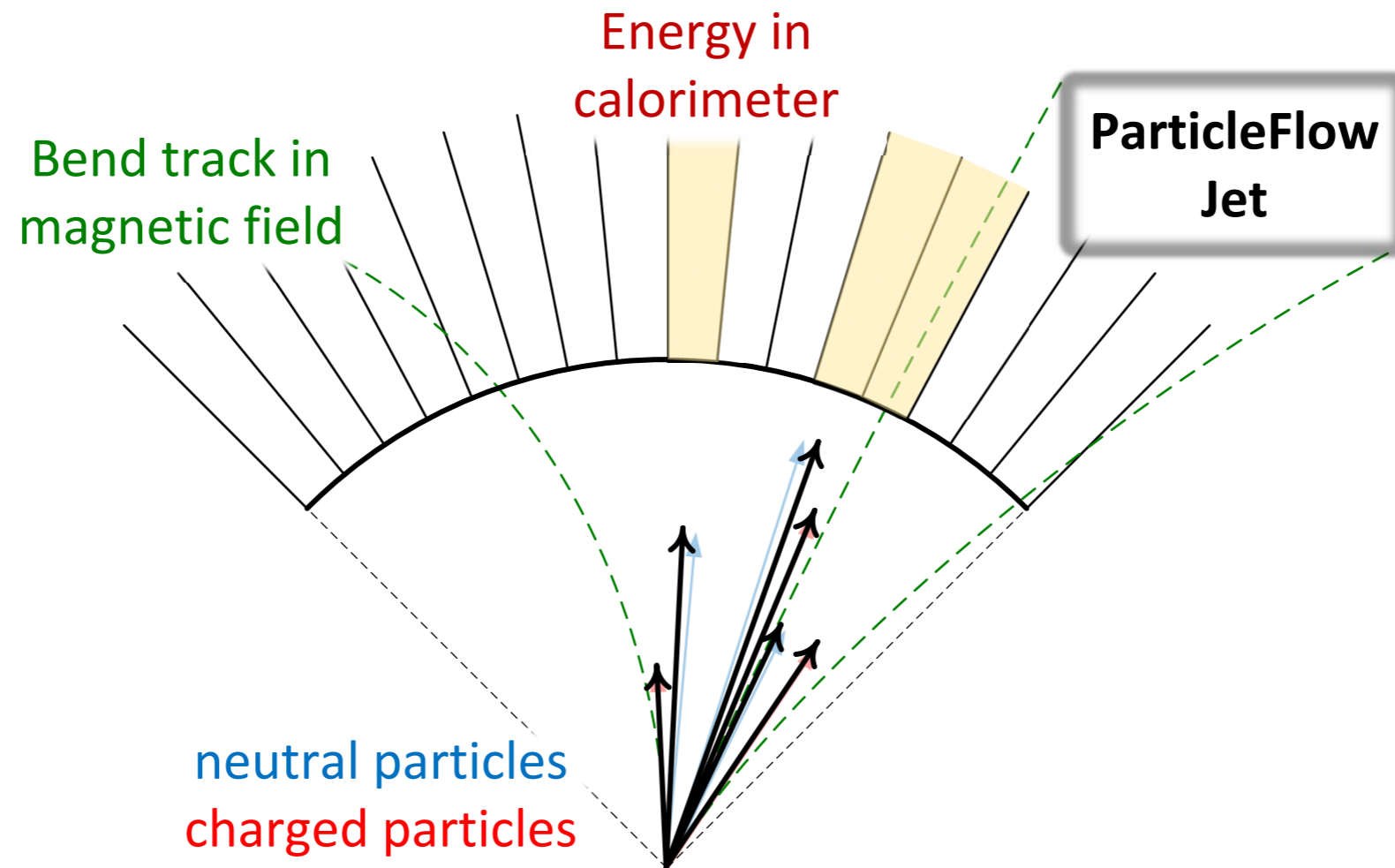
# Jet Making

- Defined by input objects, combination algorithm, and the radius.
- Usual algorithm in LHC experiments: anti- $k_t$  algorithm, which combines inputs in momentum-space, starting with *hardest* inputs.
- Algorithms need to be theoretically robust!

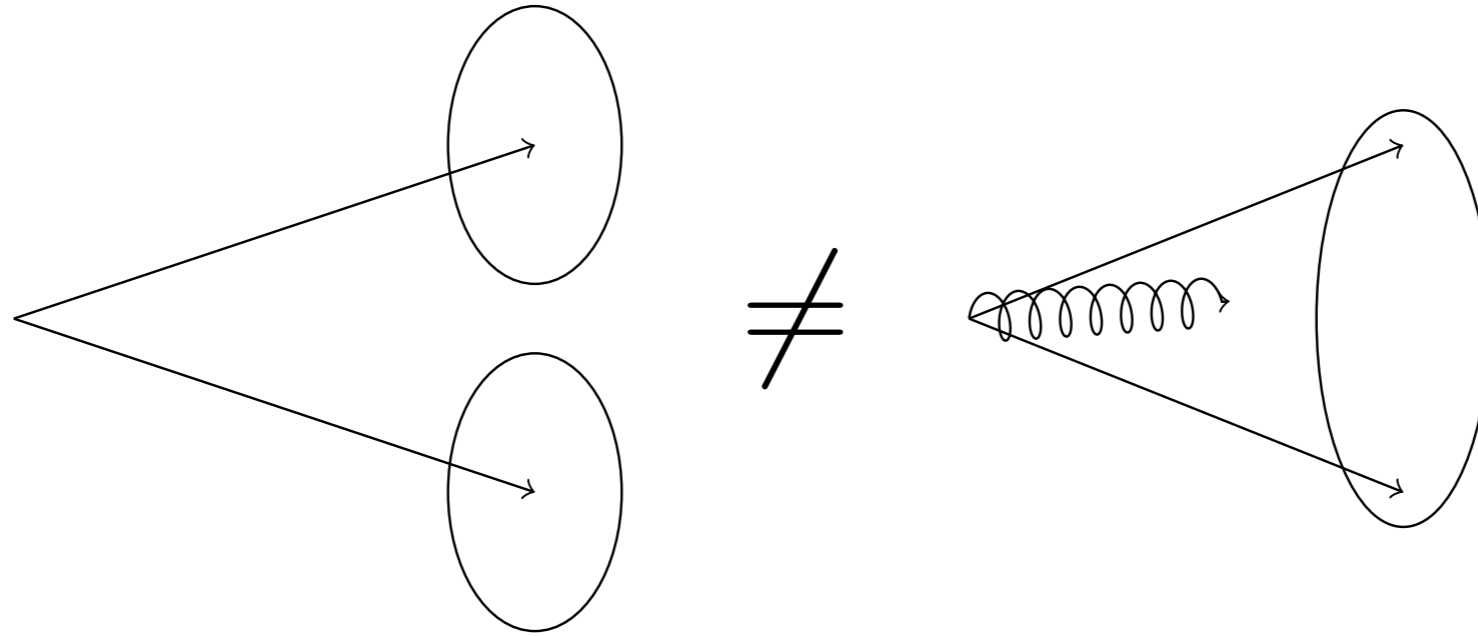
# Calorimeter Objects



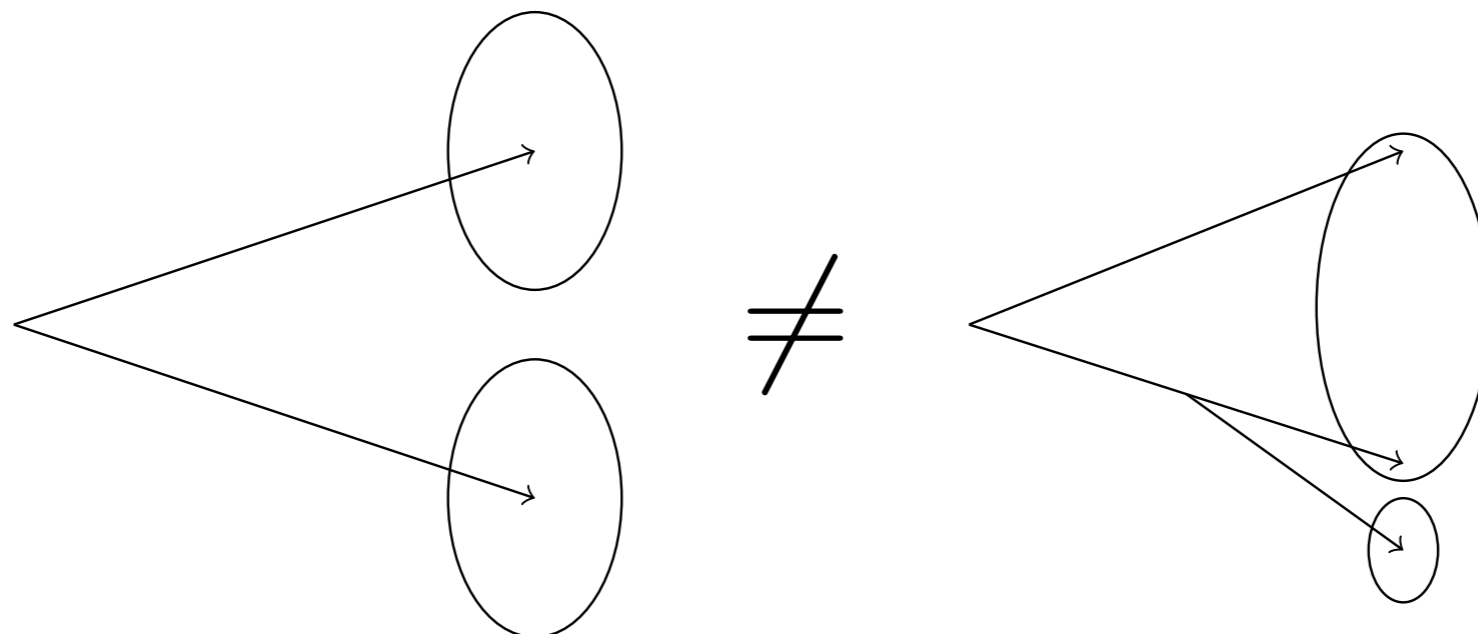
# ParticleFlow Objects



# IRC Safety



Infrared safety



46 Collinear safety



# Jet making

Bottom Up

Sequential recombination algorithms (momentum space): iteratively pairwise combination of the inputs till a minimum inter-jet distance is reached.

Top Down

Cone algorithms (coordinate space): Collect all inputs within a cone such that the cone axis is the vector sum of momenta in it.



# Jet making

Distance between two input objects

Distance between each input object and beam

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta y^2 + \Delta \phi^2}{R^2}; \quad d_{iB} = k_{ti}^{2p}; \quad p = \begin{cases} 1 & k_t \\ 0 & \text{Cambridge/Aachen} \\ -1 & \text{anti-}k_t \end{cases}$$

Intrinsic transverse momentum

Fixed “radius” parameter

- Find the smallest of all  $\{d_{ij}, d_{iB}\}$
- If this is one of the  $d_{ij}$  values, inputs  $i$  and  $j$  are merged.
- If it is one of the  $d_{iB}$  values,  $i^{\text{th}}$  input is considered a jet.
- Continue till all inputs are merged into jets.

# Jet making

Distance between two input objects

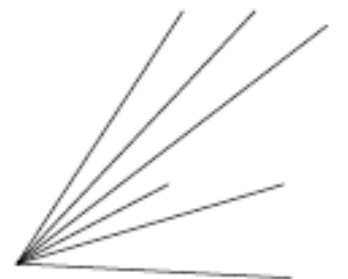
Distance between each input object and beam

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta y^2 + \Delta \phi^2}{R^2}; \quad d_{iB} = k_{ti}^{2p}; \quad p = \begin{cases} 1 & k_t \\ 0 & \text{Cambridge/Aachen} \\ -1 & \text{anti-}k_t \end{cases}$$

Intrinsic transverse momentum

Fixed “radius” parameter

- Find the smallest of all  $\{d_{ij}, d_{iB}\}$



# Jet making

Distance between two input objects

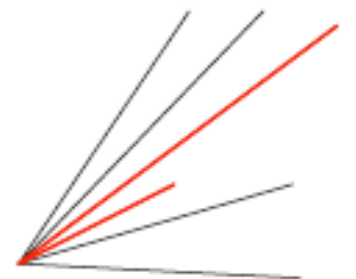
Distance between each input object and beam

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta y^2 + \Delta \phi^2}{R^2}; \quad d_{iB} = k_{ti}^{2p}; \quad p = \begin{cases} 1 & k_t \\ 0 & \text{Cambridge/Aachen} \\ -1 & \text{anti-}k_t \end{cases}$$

Intrinsic transverse momentum

Fixed “radius” parameter

- Find the smallest of all  $\{d_{ij}, d_{iB}\}$



# Jet making

Distance between two input objects

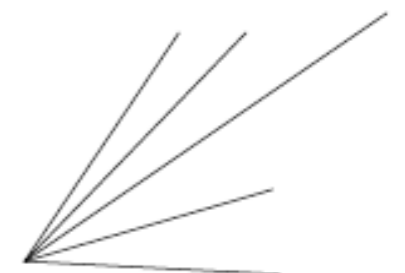
Distance between each input object and beam

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta y^2 + \Delta \phi^2}{R^2}; \quad d_{iB} = k_{ti}^{2p}; \quad p = \begin{cases} 1 & k_t \\ 0 & \text{Cambridge/Aachen} \\ -1 & \text{anti-}k_t \end{cases}$$

Intrinsic transverse momentum

Fixed “radius” parameter

- Find the smallest of all  $\{d_{ij}, d_{iB}\}$
- If this is one of the  $d_{ij}$  values, inputs  $i$  and  $j$  are merged.



# Jet making

Distance between two input objects

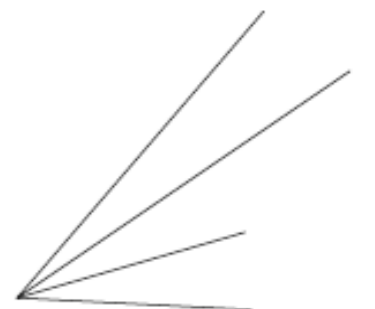
Distance between each input object and beam

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta y^2 + \Delta \phi^2}{R^2}; \quad d_{iB} = k_{ti}^{2p}; \quad p = \begin{cases} 1 & k_t \\ 0 & \text{Cambridge/Aachen} \\ -1 & \text{anti-}k_t \end{cases}$$

Intrinsic transverse momentum

Fixed “radius” parameter

- Find the smallest of all  $\{d_{ij}, d_{iB}\}$
- If this is one of the  $d_{ij}$  values, inputs  $i$  and  $j$  are merged.
- If it is one of the  $d_{iB}$  values,  $i^{\text{th}}$  input is considered a jet.
- Continue till all inputs are merged into jets.





# Jet making

Distance between two input objects

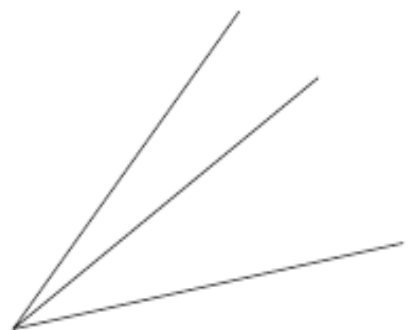
Distance between each input object and beam

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta y^2 + \Delta \phi^2}{R^2}; \quad d_{iB} = k_{ti}^{2p}; \quad p = \begin{cases} 1 & k_t \\ 0 & \text{Cambridge/Aachen} \\ -1 & \text{anti-}k_t \end{cases}$$

Intrinsic transverse momentum

Fixed “radius” parameter

- Find the smallest of all  $\{d_{ij}, d_{iB}\}$
- If this is one of the  $d_{ij}$  values, inputs  $i$  and  $j$  are merged.
- If it is one of the  $d_{iB}$  values,  $i^{\text{th}}$  input is considered a jet.
- Continue till all inputs are merged into jets.



# Jet making

Distance between two input objects

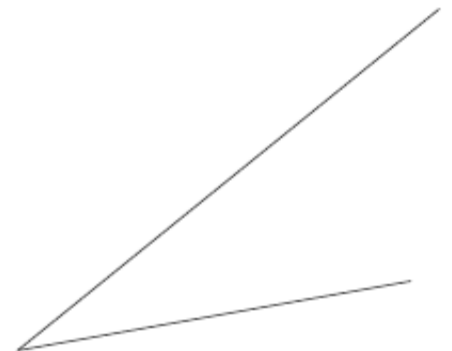
Distance between each input object and beam

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta y^2 + \Delta \phi^2}{R^2}; \quad d_{iB} = k_{ti}^{2p}; \quad p = \begin{cases} 1 & k_t \\ 0 & \text{Cambridge/Aachen} \\ -1 & \text{anti-}k_t \end{cases}$$

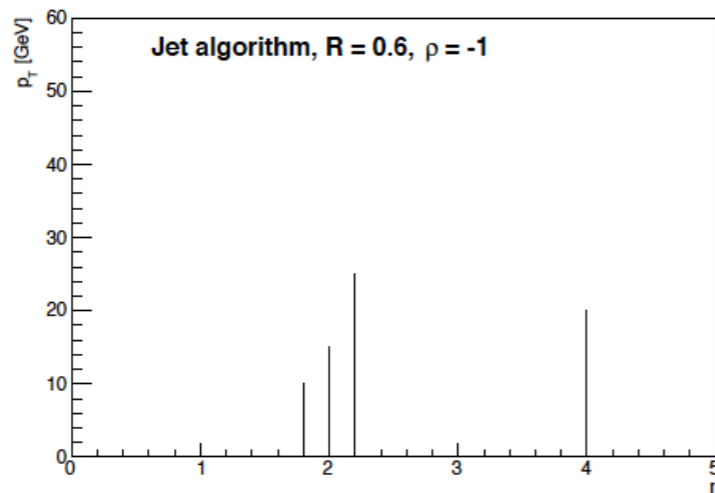
Intrinsic transverse momentum

Fixed “radius” parameter

- Find the smallest of all  $\{d_{ij}, d_{iB}\}$
- If this is one of the  $d_{ij}$  values, inputs  $i$  and  $j$  are merged.
- If it is one of the  $d_{iB}$  values,  $i^{\text{th}}$  input is considered a jet.
- Continue till all inputs are merged into jets.



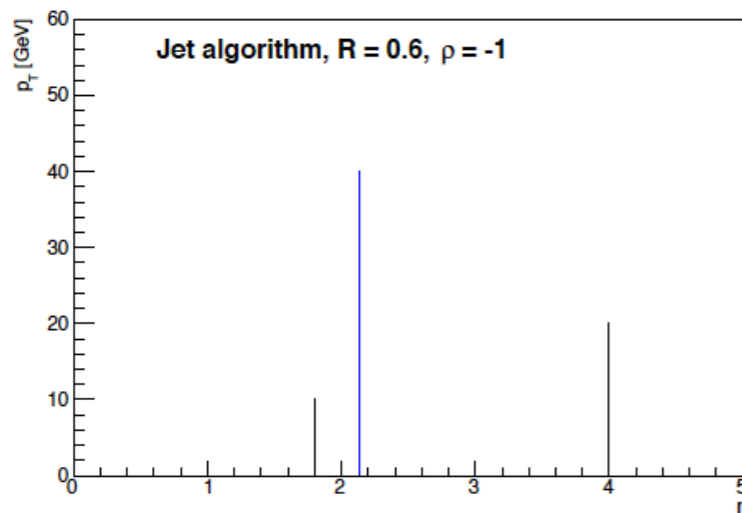
# Step 1



$d_{ij}$	1	2	3	4
1	-	0.00049	0.00071	0.03361
2	0.00049	-	<b>0.00018</b>	0.02778
3	0.00071	<b>0.00018</b>	-	0.01440
4	0.03361	0.02778	0.01440	-
$d_{iB}$	0.01000	0.00444	0.00160	0.0025

(a) We have 4 input objects as shown. The smaller value is indicated, which dictates 2 and 3 should be merged. The merged  $p_T$  will be 40 GeV, and the position is determined by the  $p_T$ -weighted average:  $(2 * 15 + 2.2 * 25) / 40 = 2.13$ .

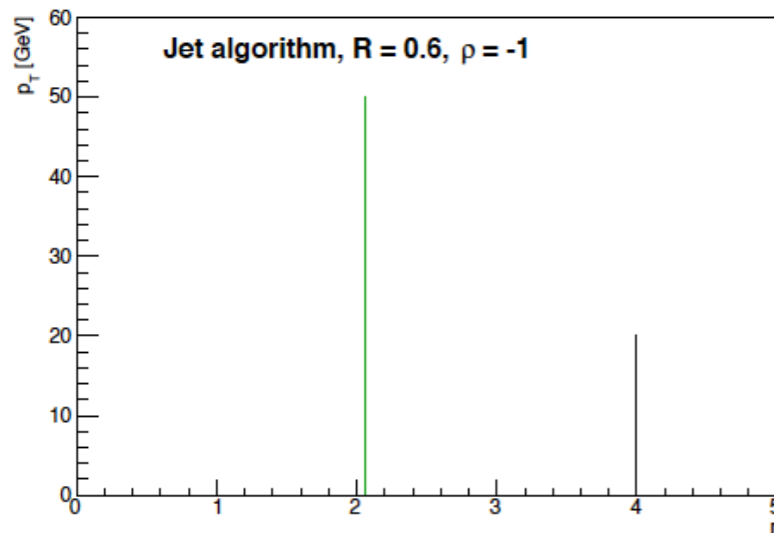
# Step 2



$d_{ij}$	1	23	4
1	-	0.00019	0.03361
23	0.00019	-	0.00607
4	0.03361	0.00607	-
$d_{iB}$	0.00444	0.00063	0.00250

(b) At this step, we indicate the merged input from previous step by 23. The distances indicate that inputs 1 and 23 should be merged. The merged  $p_T$  will be 50 GeV, and the position will be determined by the  $p_T$ -weighted average:  $(1.8 * 10 + 2.13 * 40) / 50 = 2.06$ .

# Step 3



---

$d_{ij}$	123	4
123	-	0.00418
4	0.00418	-

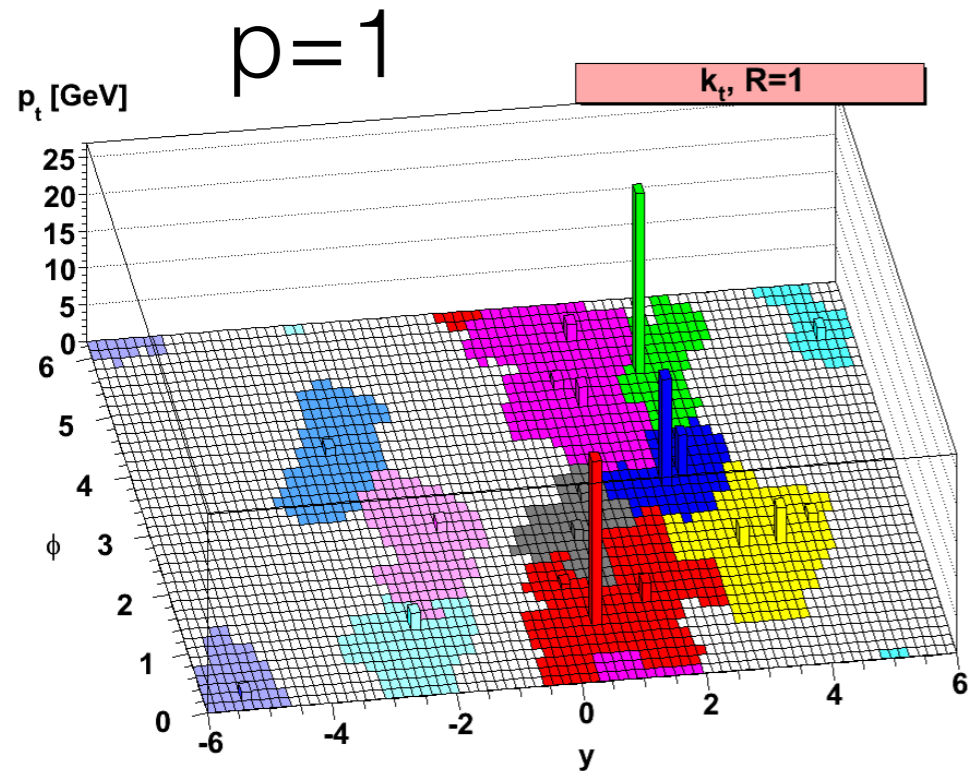
---

$d_{iB}$	0.00040	0.00250
----------	---------	---------

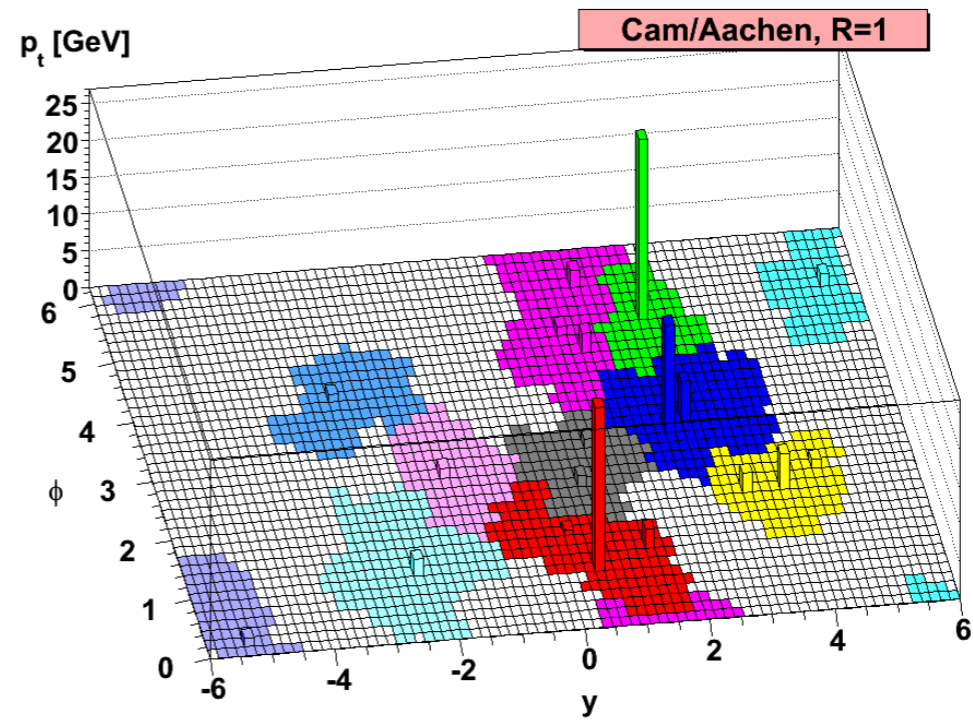
---

(c) At this step, we indicate the merged input from previous step by 123. The distances indicate the input 123 should be classified as a jet itself. Since that leaves input 4, that will be classified as a jet as well.

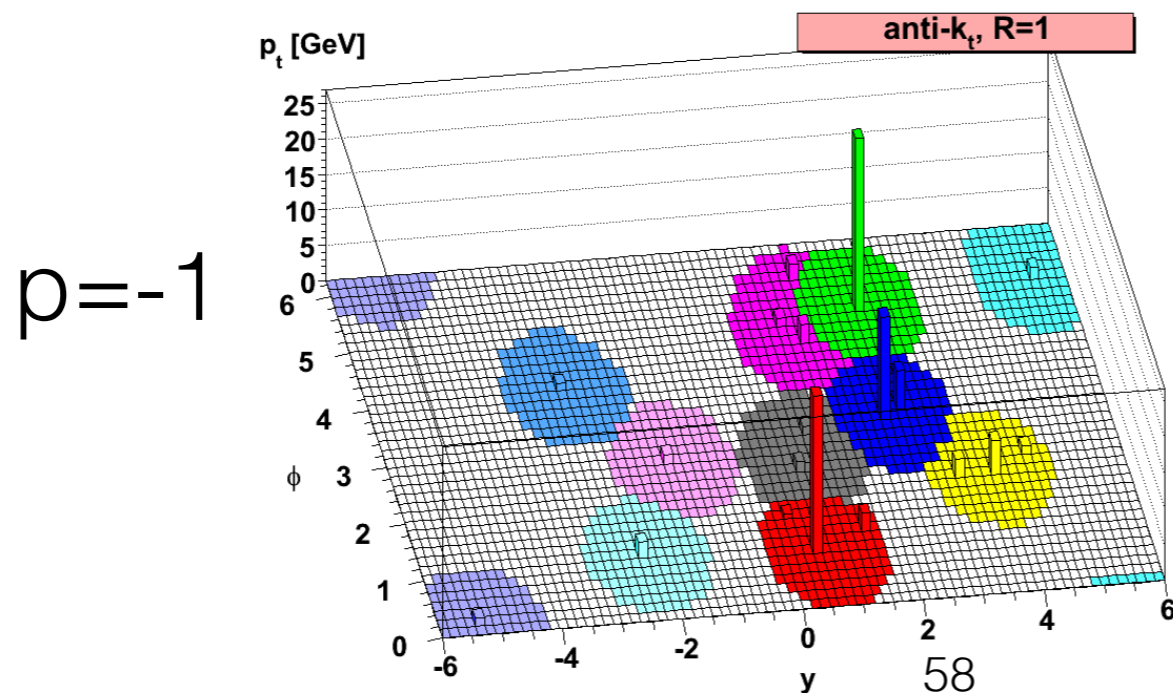
# Finished product!



Irregularly shaped jets



Shape follow angular distribution of components



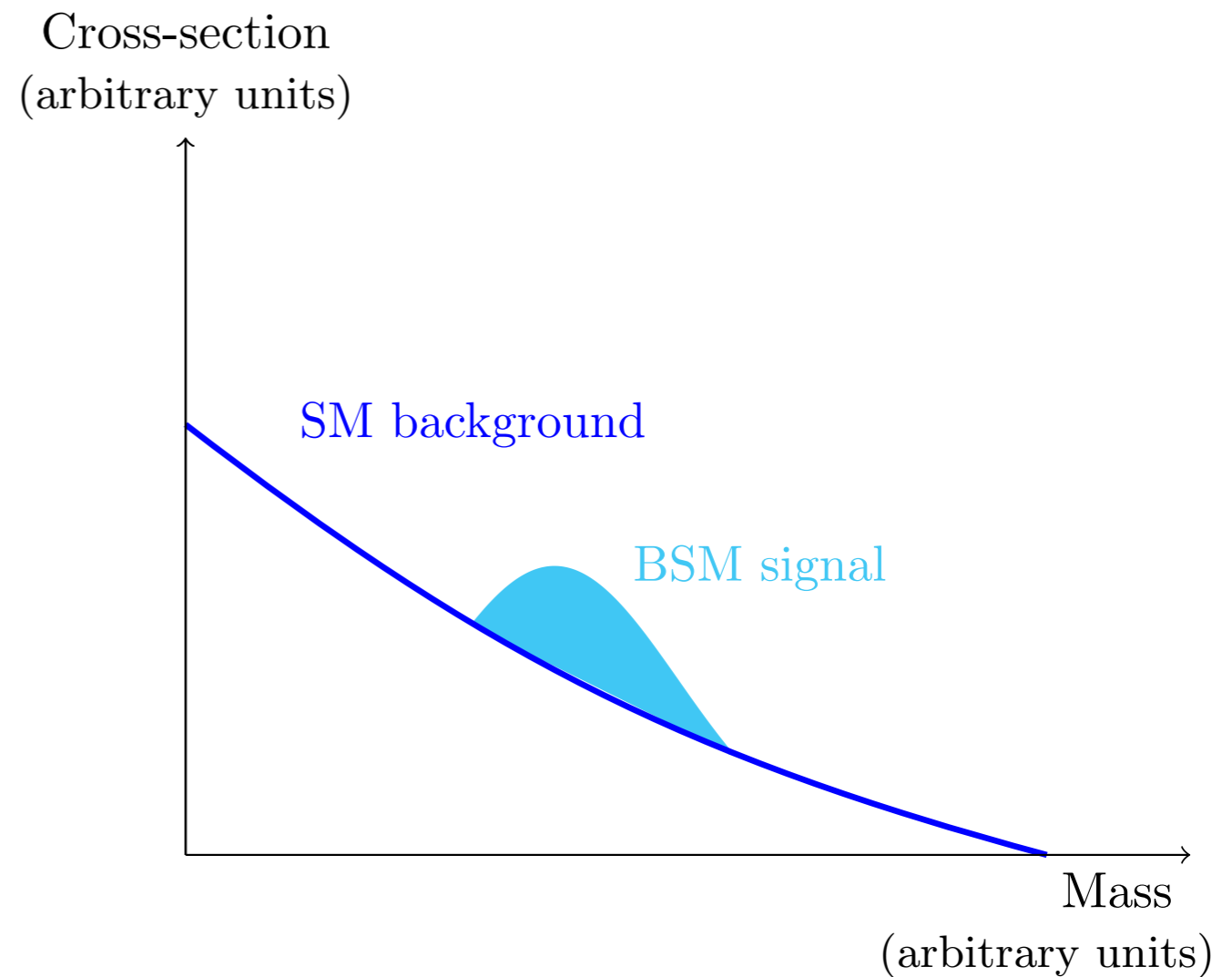
Almost circularly shaped jets



# Searches



- Resonance searches: bump hunting
- Cut and count
- Excess of MET: DM
- Signal strength



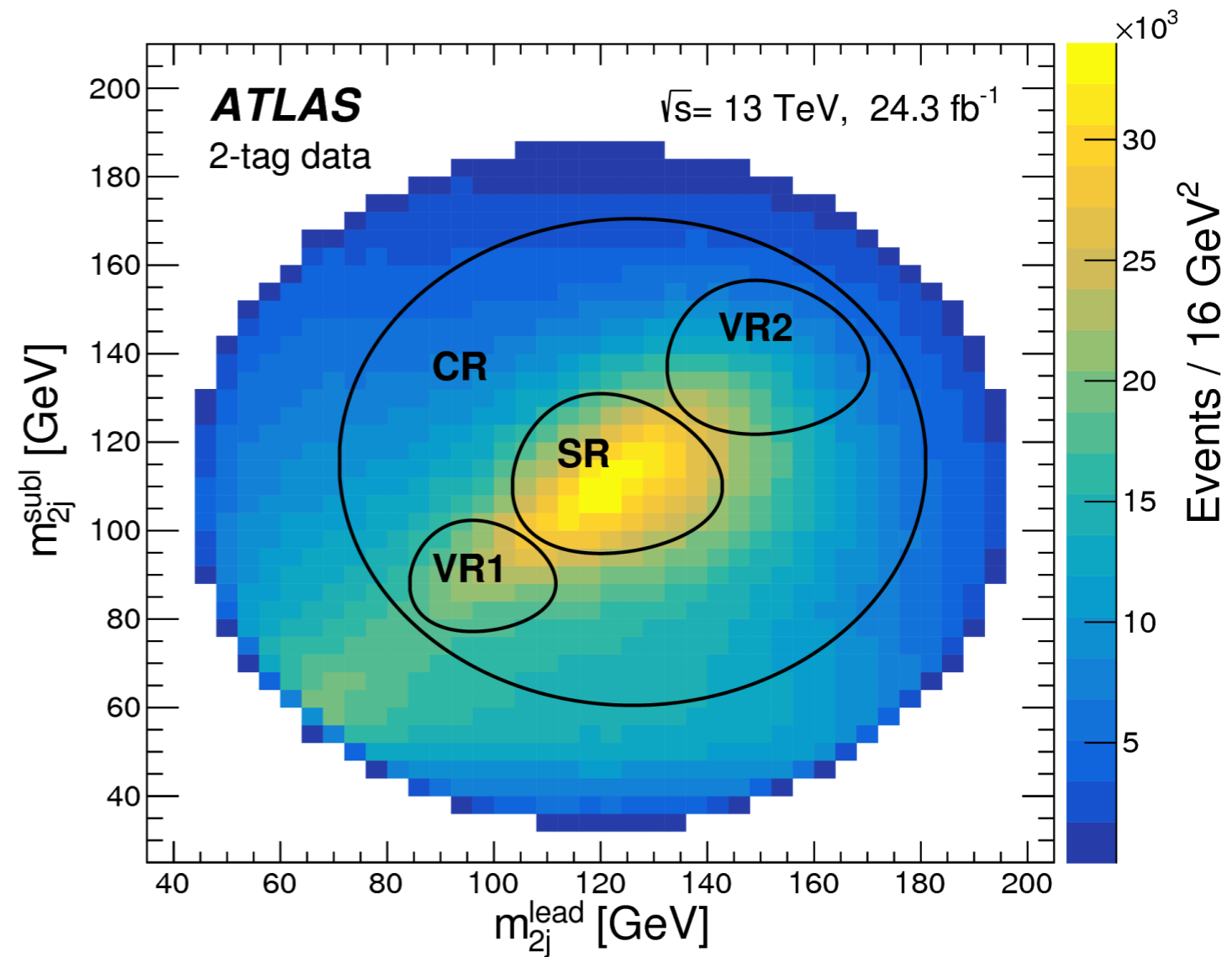
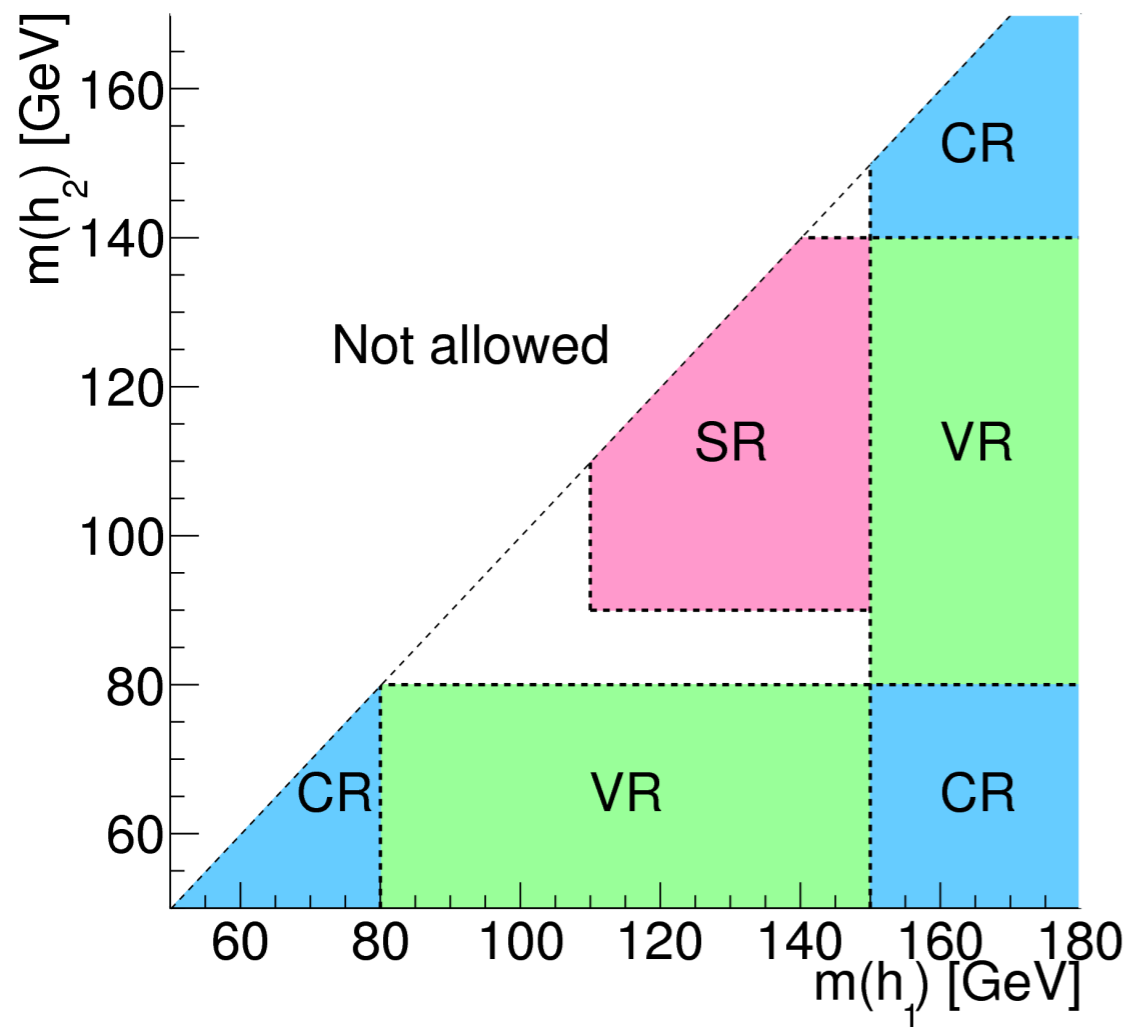
# Types of backgrounds

- Irreducible: same final state. SM ZZ for H to ZZ.
- Reducible: not the same final state, resulting from misreconstructed processes or misidentified objects.  $W(\lnu)+jets$  for  $Z(l\bar{l})+jets$ .
- Combinatorial: random combination of objects looking like the signal. All hadronic  $t\bar{t}$ .

# Signal and control regions

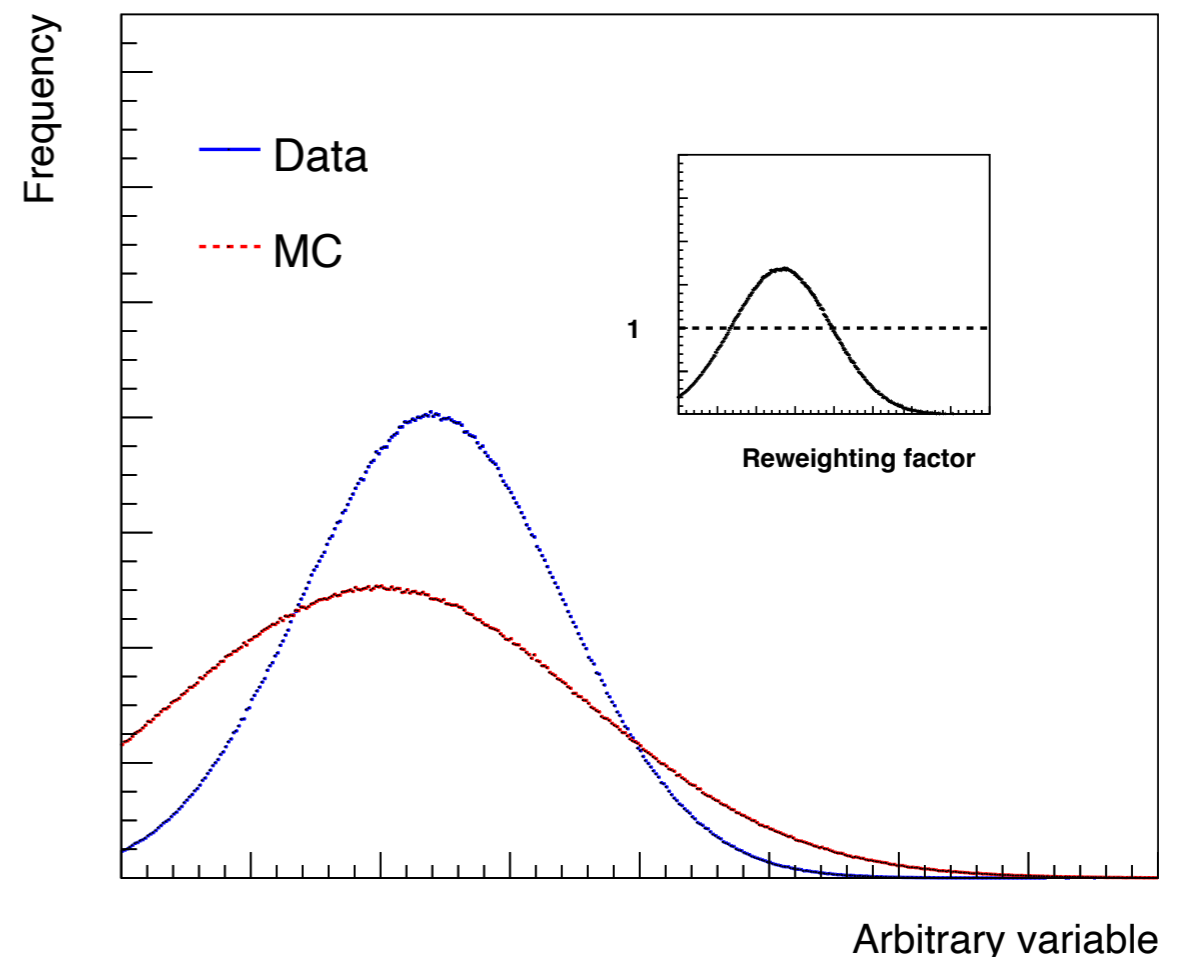
- We apply selection *cuts* on the objects and event topology to maximise signal and minimise background.
- However, when searching for a new physics signal, we do not want to bias ourselves.
- So divide the events into signal region, where data is blinded when we fix analysis strategies, and control region by inverting one (or more signal cuts), where we can check data-MC agreement and estimate background contribution.
- Unblinded after cuts are optimised and **fixed**.

# SR and CR



# Data-MC agreement in CR

- What if simulation does not describe the data in CR?
- Modelling?
- Calibration/efficiency estimates wrong?
- Reweight :-)



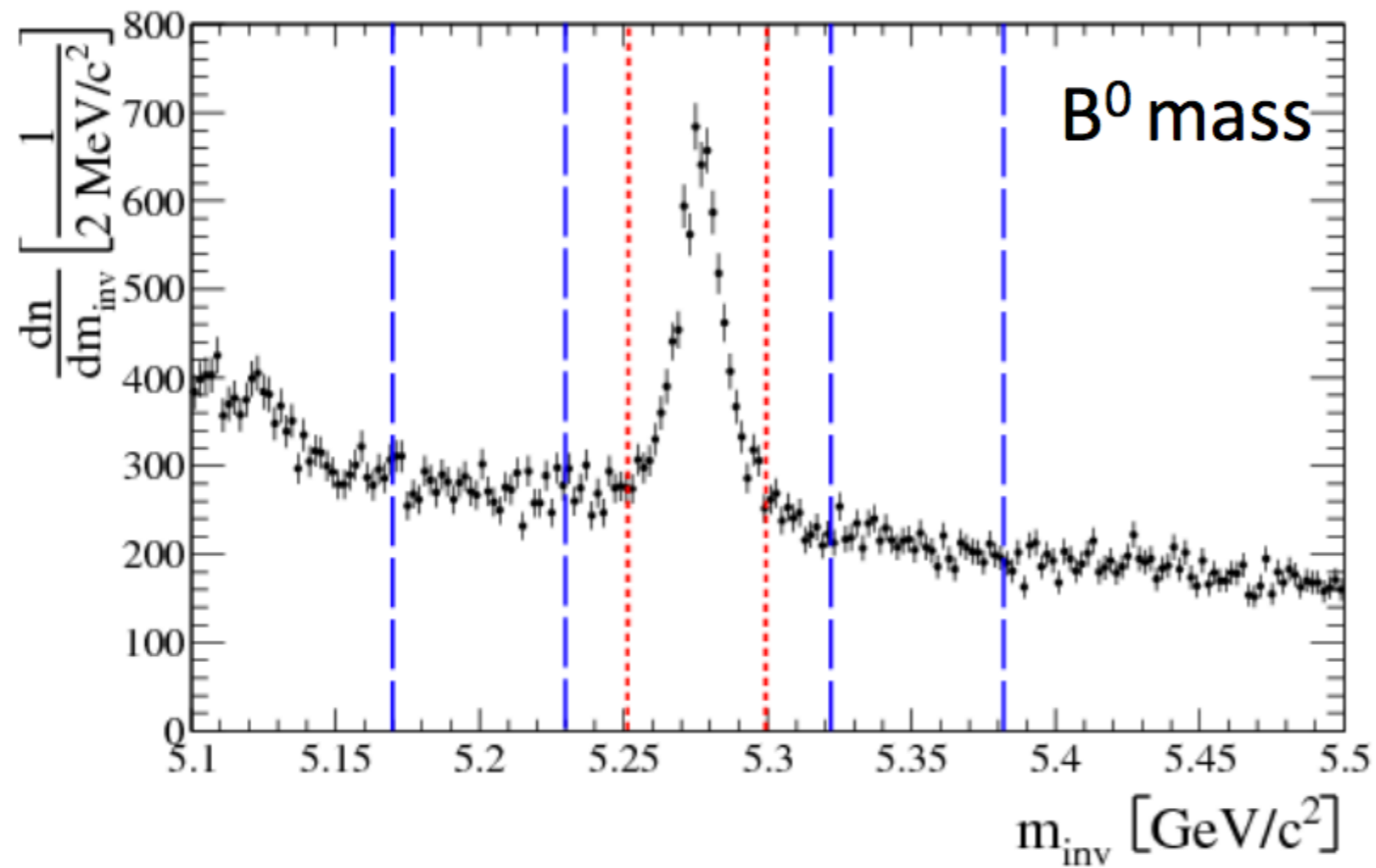
# Estimate the backgrounds

Data and/or  
simulation driven

- Anti-selection/inversion of cuts
- Side-bands/shape extraction by fit
- ABCD method
- ...



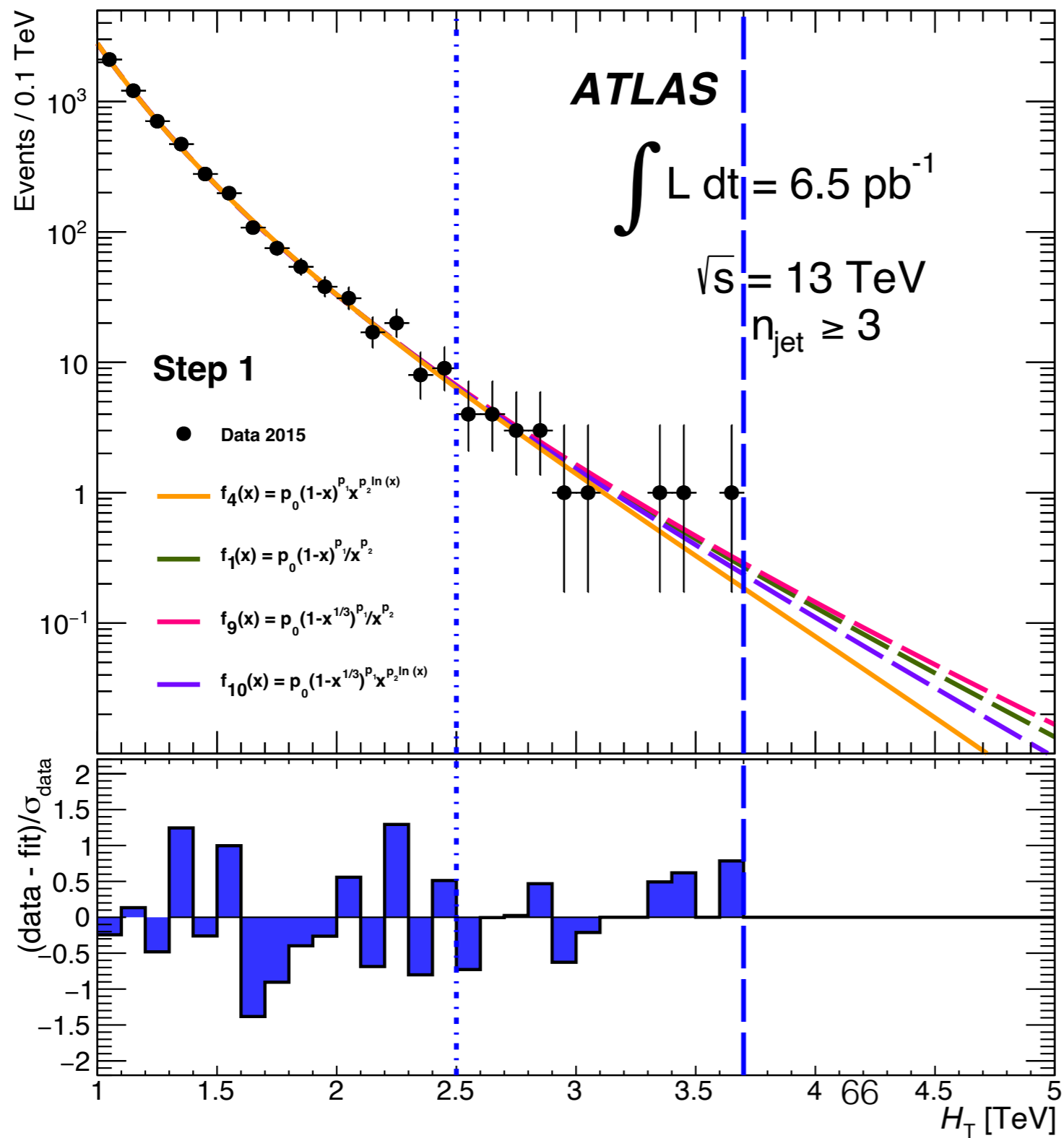
# Sidebands



Estimate background  
under a resonance peak

With or without fitting

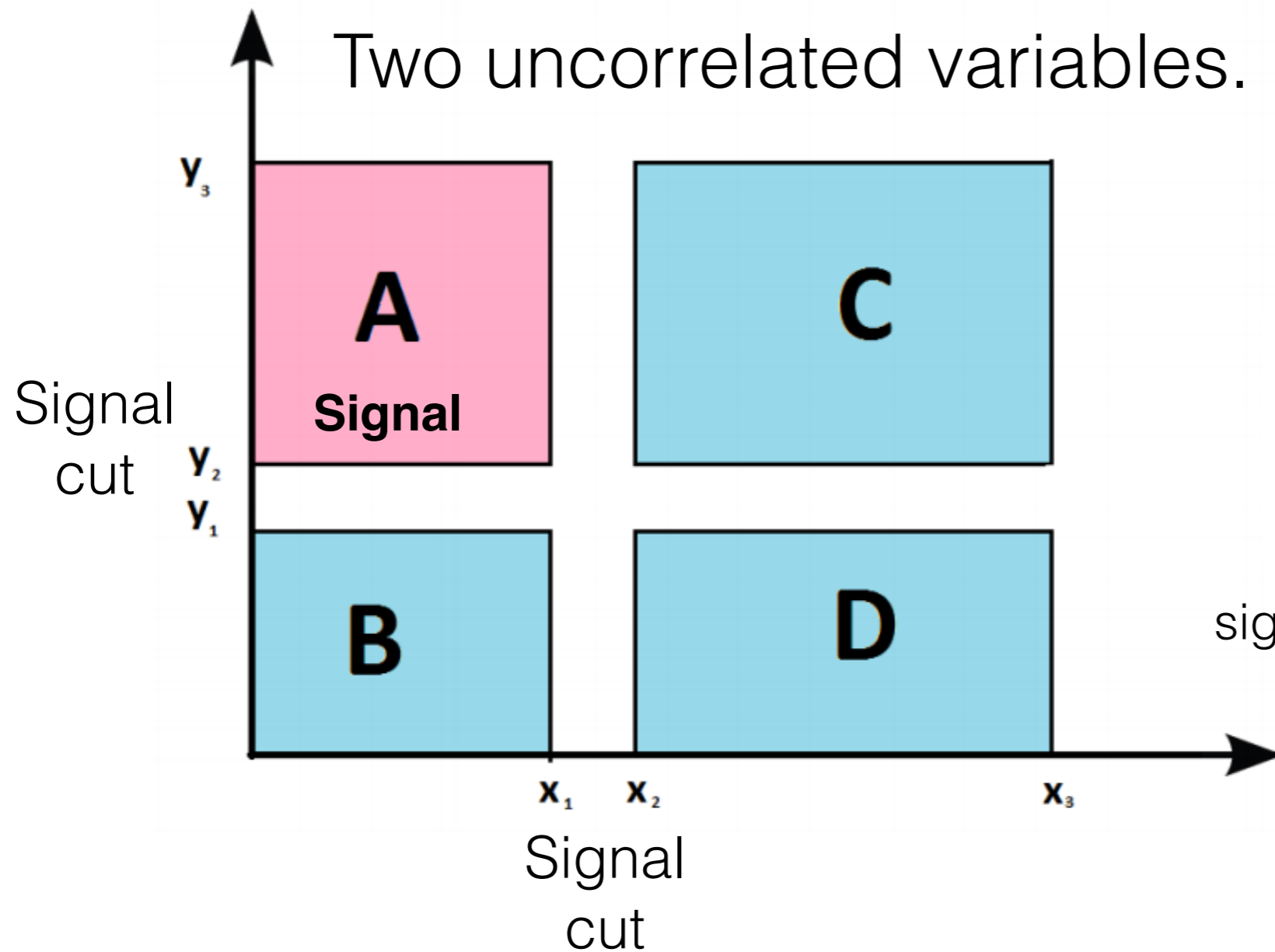
# Sidebands



Estimate background  
under a resonance peak

With or without fitting

# ABCD method



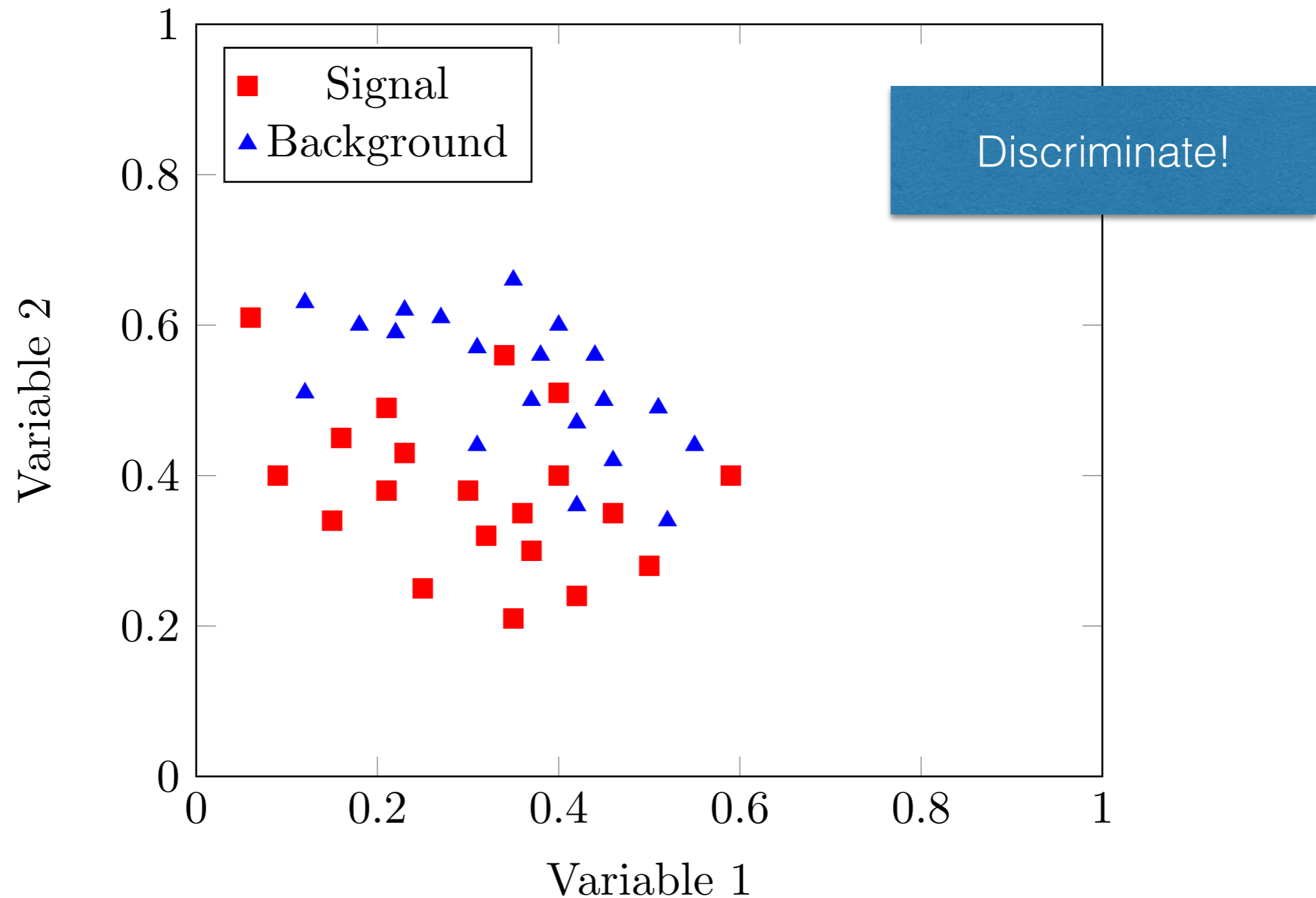
Estimate by:

$$N_A = \frac{N_B \times N_C}{N_D}$$

b.g in  
signal region

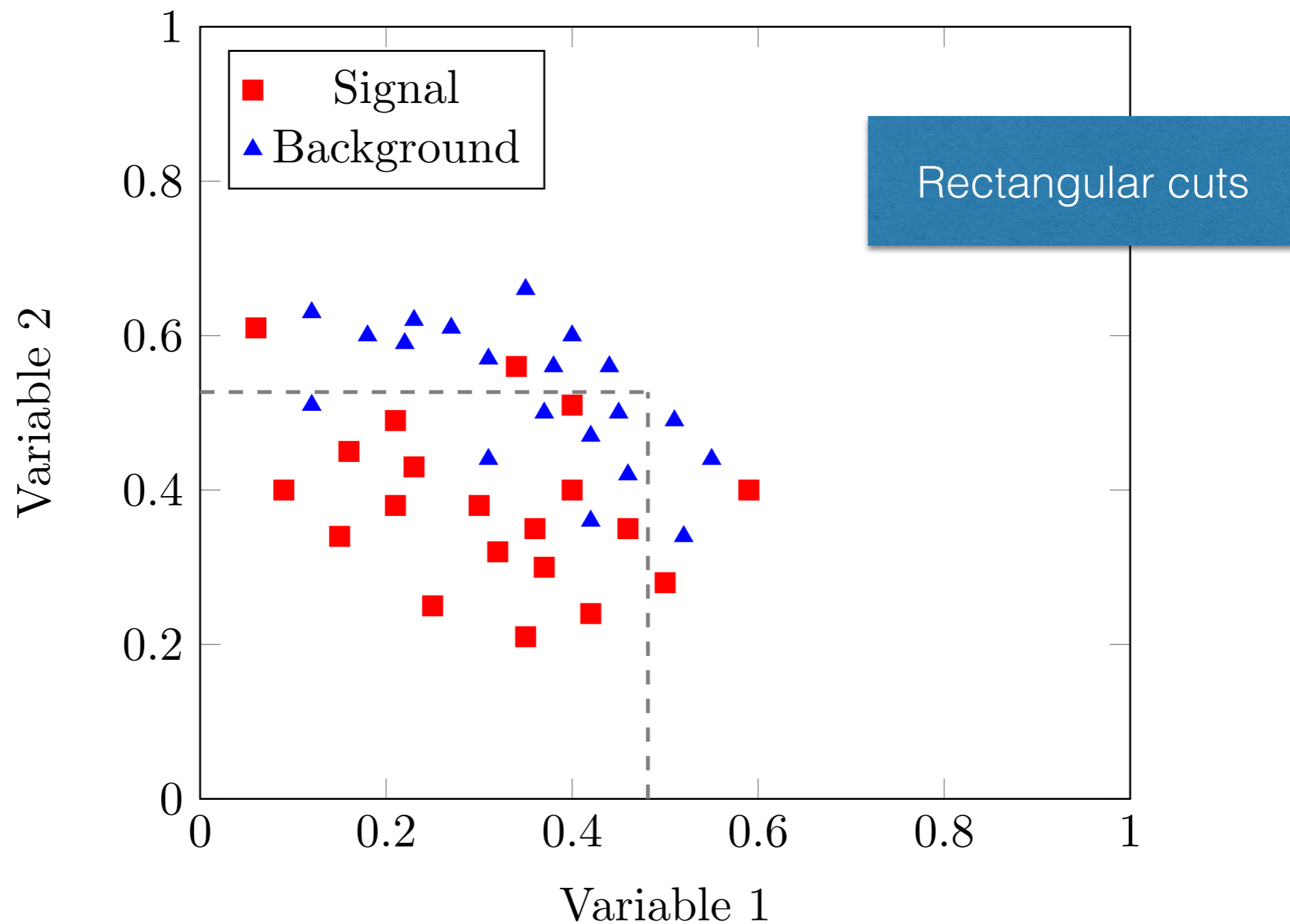
# **Use of ML in object reconstruction**

# Precursor: Multivariate analysis

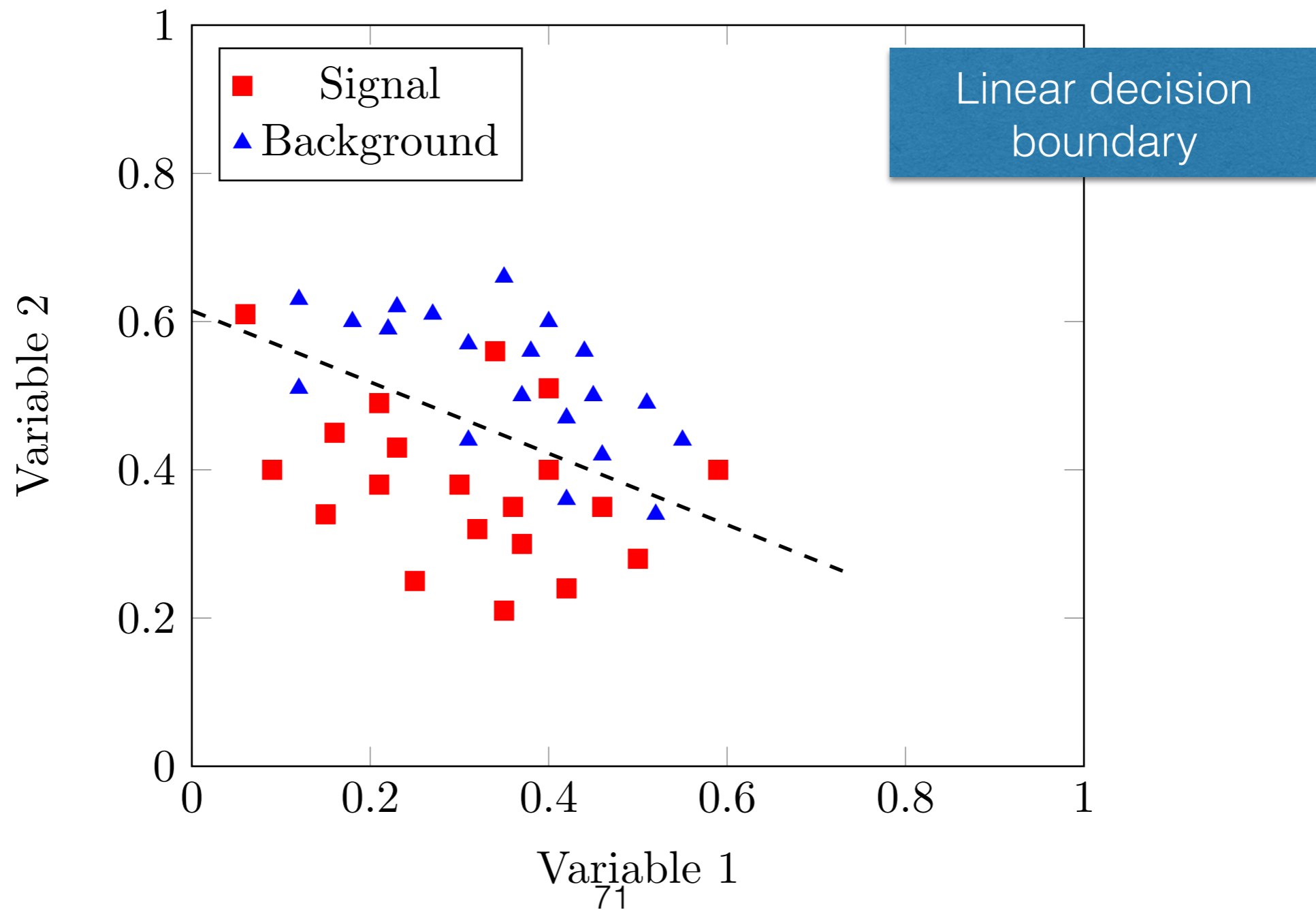




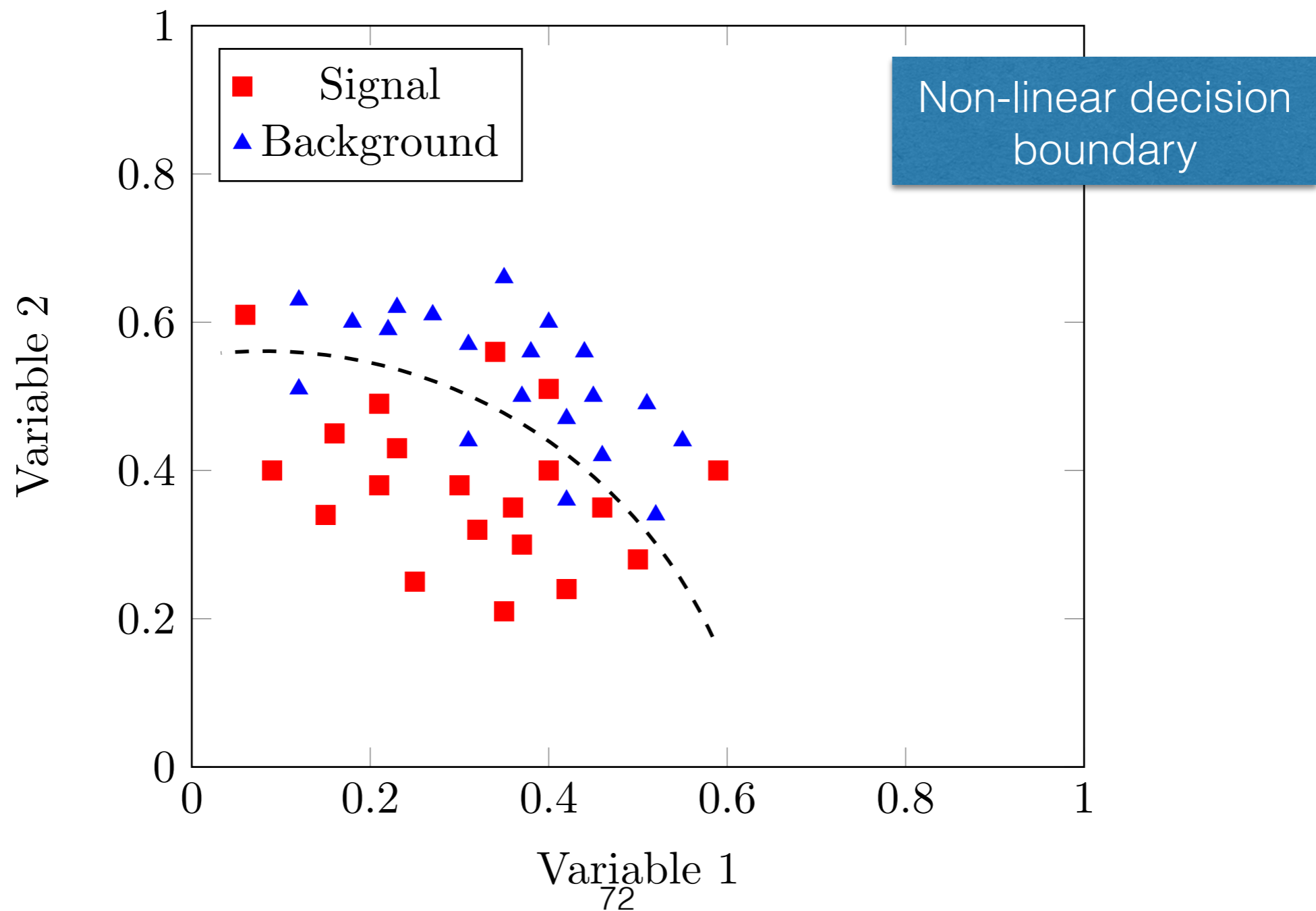
# Precursor: Multivariate analysis



# Precursor: Multivariate analysis



# Precursor: Multivariate analysis



# Challenge:

Find the optimal decision boundary  
in N-dimension!

# MVA

- N variables used in classification: feature variables
- Correlation reduces dimensionality
- N dimensional constant surface  $\longrightarrow$  mapping to a single discriminating variable
- Actual cut on the variable, as before!

# Preprocessing

- Combine or transform the variables to bring out physical features —> called feature extraction
- Example:  $W$ -boson transverse mass, scaling by  $\sqrt{s}$ ,




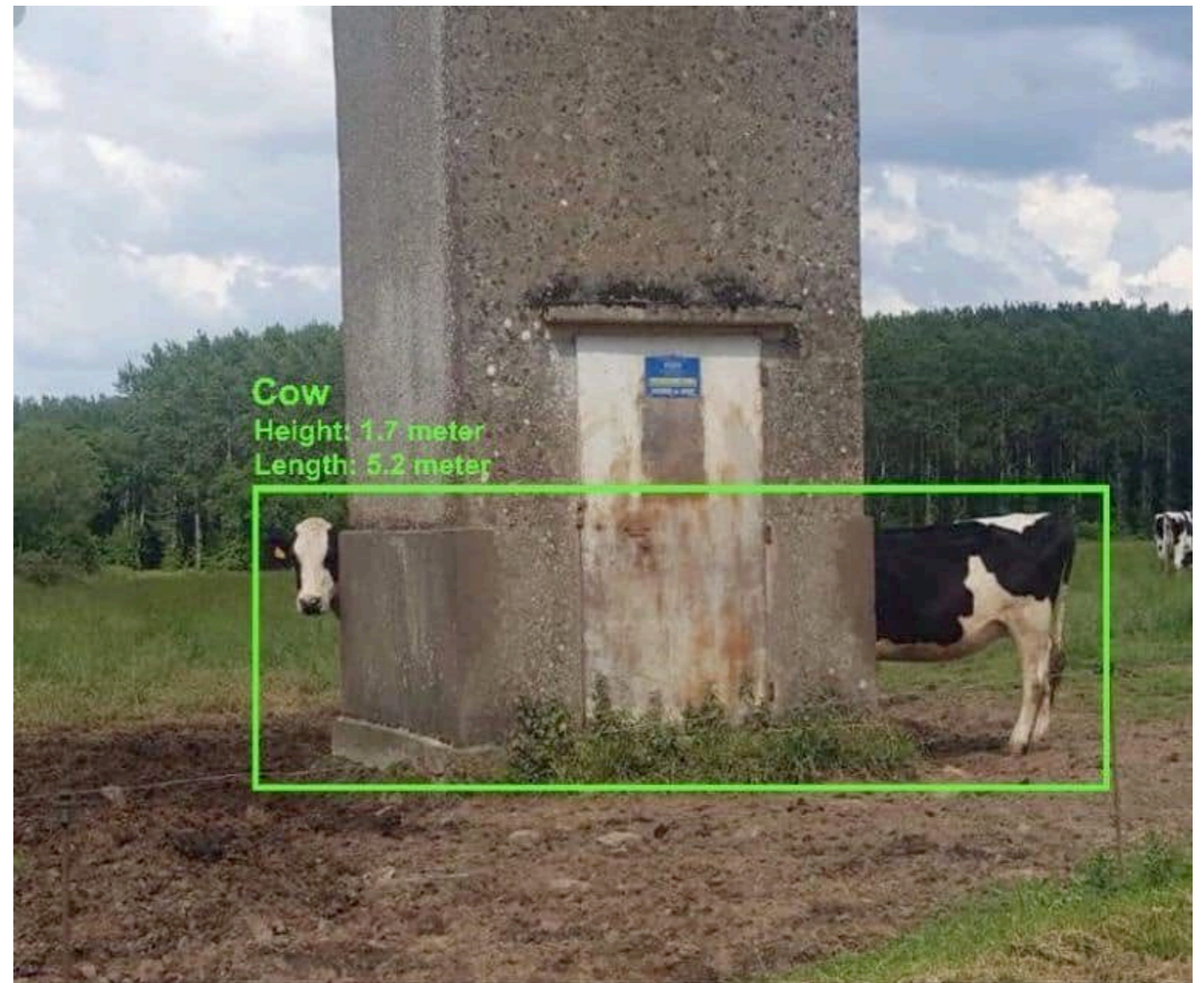
# Machine Learning

- Algorithmically find this decision boundary based on *data* (without explicitly programmed)
- Learning: represent the data by an approximate functional form (whether or not such a form exists is immaterial) between input variables  $x$  and output variables  $y$
- Use that predict behaviour of future similar datasets from  $x'$  to  $y'$ .

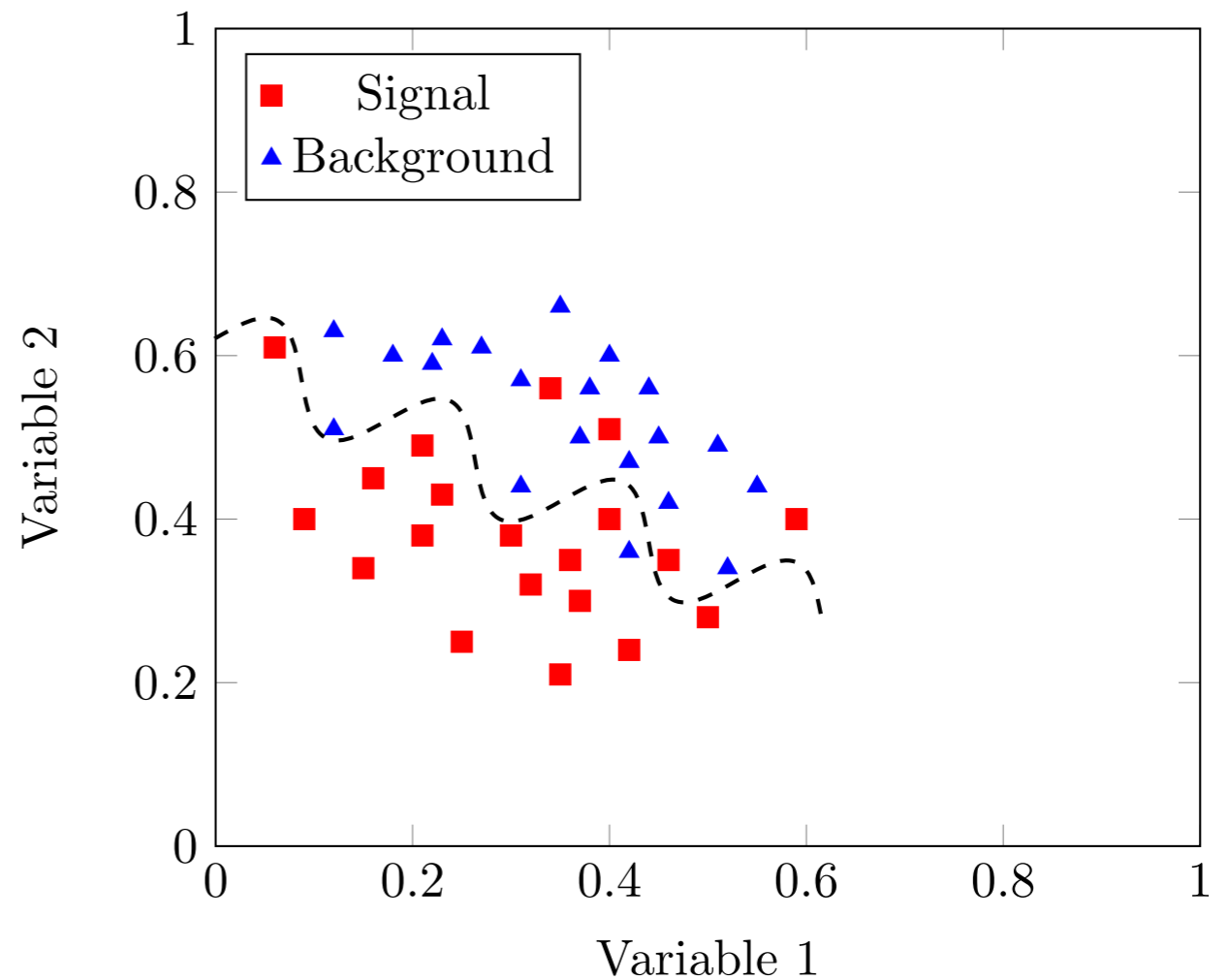
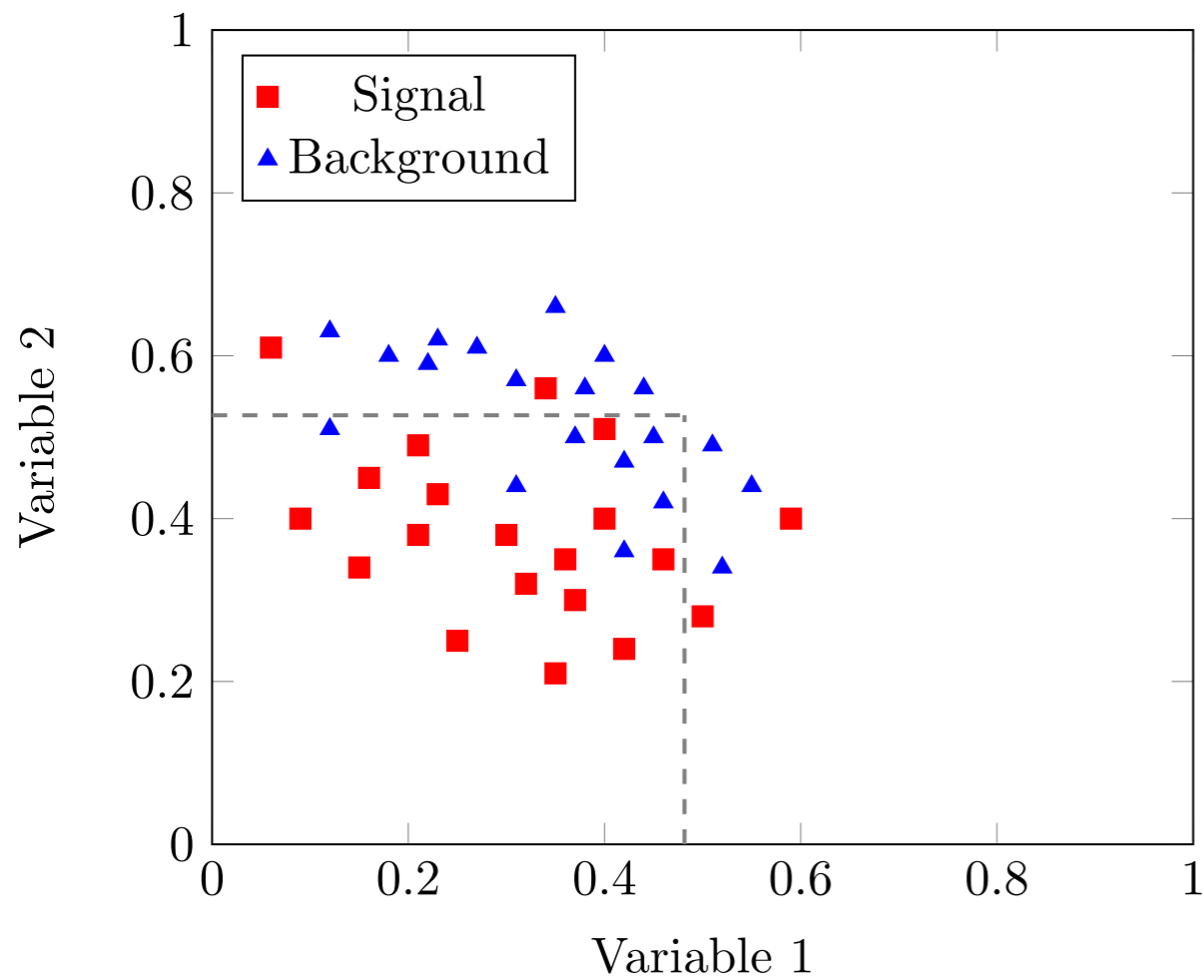
# But!

Math break!  
1, 3, 5, 7, ...  
What number will be next?

 217341, because if  
 $f(x) = 18111/2 \cdot x^4 - 90555 \cdot x^3 + 633885/2 \cdot x^2 - 452773 \cdot x + 217331$ ,  
then:  
 $f(1) = 1$   
 $f(2) = 3$   
 $f(3) = 5$   
 $f(4) = 7$   
 $f(5) = 217341$



# Underfitting and Overfitting



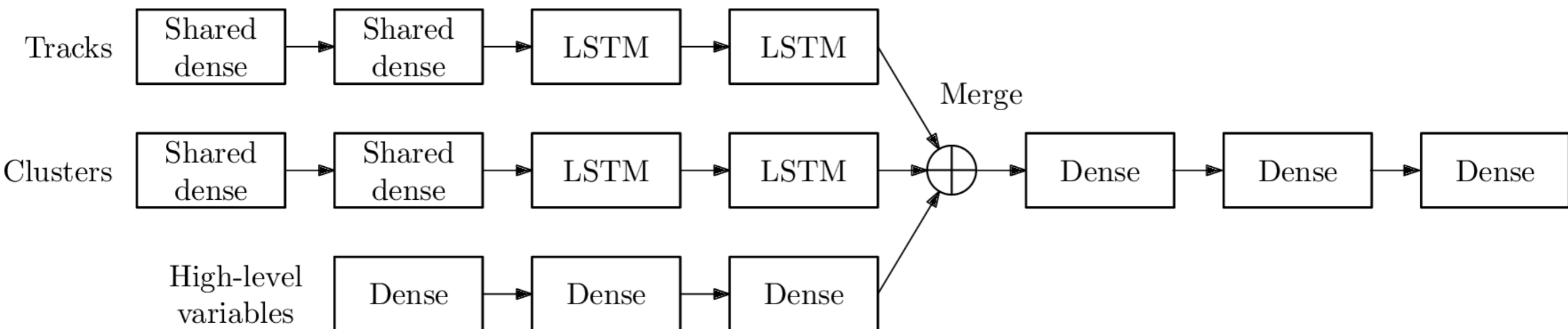
# Back to the Objects

# Quiz

Why the (hadronic) decay of tau contain only odd number of charged particles tracks?

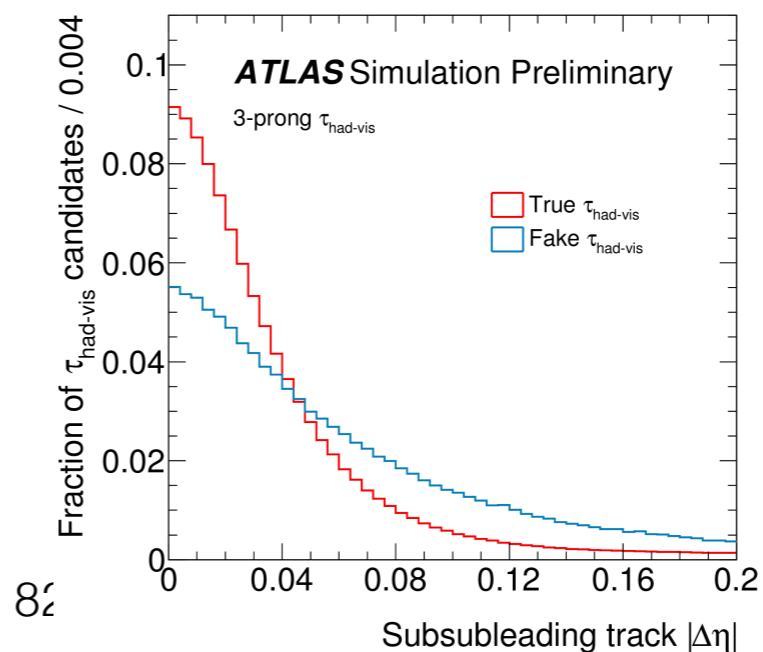
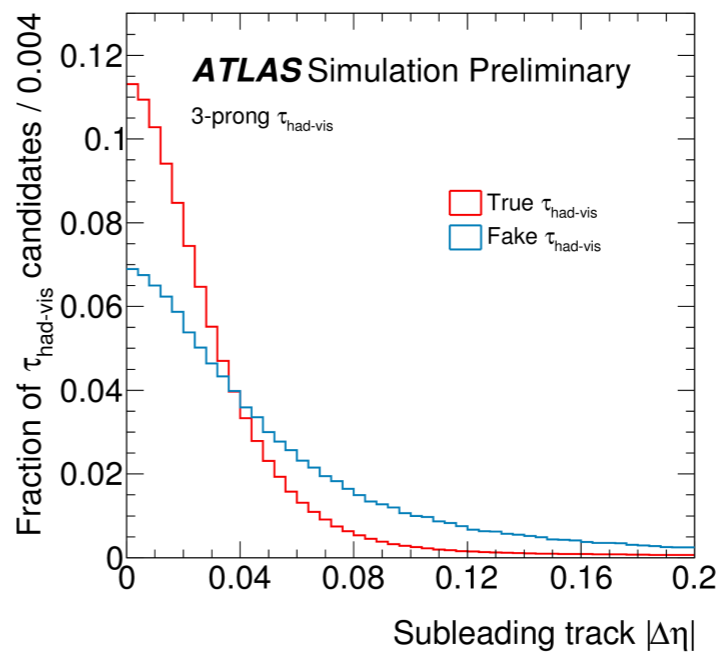
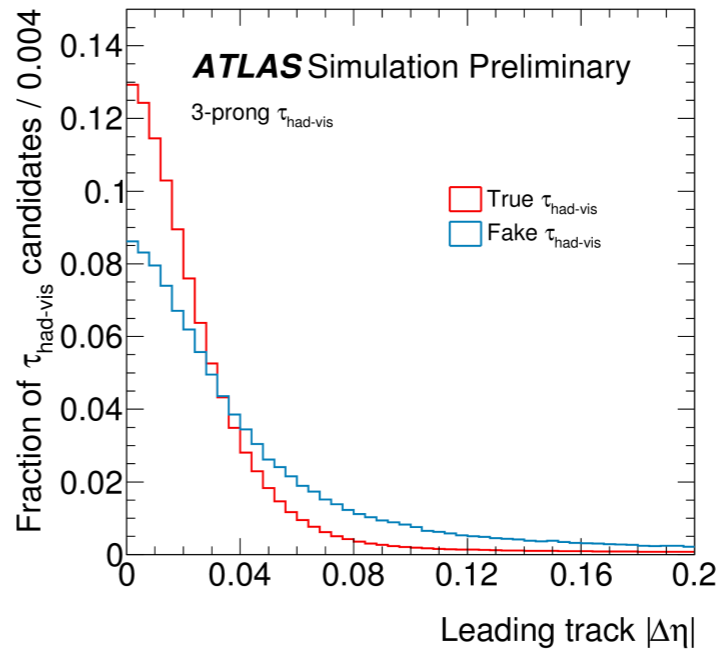
# Example: ATLAS hadronic tau reconstruction

- The challenge: discriminating non-tau jets
- Use a RNN
- Trained sep for 1/3 prongs using simulated samples





	Observable	1-prong	3-prong
Track inputs	$p_T^{\text{seed jet}}$	•	•
	$p_T^{\text{track}}$	•	•
	$\Delta\eta^{\text{track}}$	•	•
	$\Delta\phi^{\text{track}}$	•	•
	$ d_0^{\text{track}} $	•	•
	$ z_0^{\text{track}} \sin\theta $	•	•
	$N_{\text{IBL hits}}$	•	•
	$N_{\text{Pixel hits}}$	•	•
Cluster inputs	$p_T^{\text{jet seed}}$	•	•
	$E_T^{\text{cluster}}$	•	•
	$\Delta\eta^{\text{cluster}}$	•	•
	$\Delta\phi^{\text{cluster}}$	•	•
	$\lambda_{\text{cluster}}$	•	•
	$\langle\lambda_{\text{cluster}}^2\rangle$	•	•
	$\langle r_{\text{cluster}}^2\rangle$	•	•
High-level inputs	$p_T^{\text{uncalibrated}}$	•	•
	$f_{\text{cent}}$	•	•
	$f_{\text{leadtrack}}^{-1}$	•	•
	$\Delta R_{\text{max}}$	•	•
	$ S_{\text{leadtrack}} $	•	•
	$S_T^{\text{flight}}$	•	•
	$f_{\text{track}}^{\text{iso}}$	•	•
	$f_{\text{track}}^{\text{EM}}$	•	•
	$p_T^{\text{EM+track}}/p_T$	•	•
	$m^{\text{EM+track}}$	•	•
	$m^{\text{track}}$	•	•



# Example: ATLAS hadronic tau reconstruction

