General Exam: Two Higgs to 4b (and beyond)

Sean Gasiorowski
University of Washington

12th August 2019

Outline

- **General introduction:**
  - ATLAS, coordinates, jets, and b-tagging
- **ATLAS qualification task:**
  - Description and results
- **The HH→4b analysis:**
  - Motivation and benchmarks
  - The analysis: selection, background, and results
  - Looking forward
- **Thesis Timeline**

General Introduction

Introduction: LHC and Timeline

HL-LHC: High Luminosity LHC
LS: Long Shutdown
TeV: Tera electron Volt

Source ATLAS EXPERIMENT

- The Large Hadron Collider (LHC) is a proton-proton collider near Geneva, Switzerland, operating at center of mass energy $\sqrt{s} = 13$ TeV

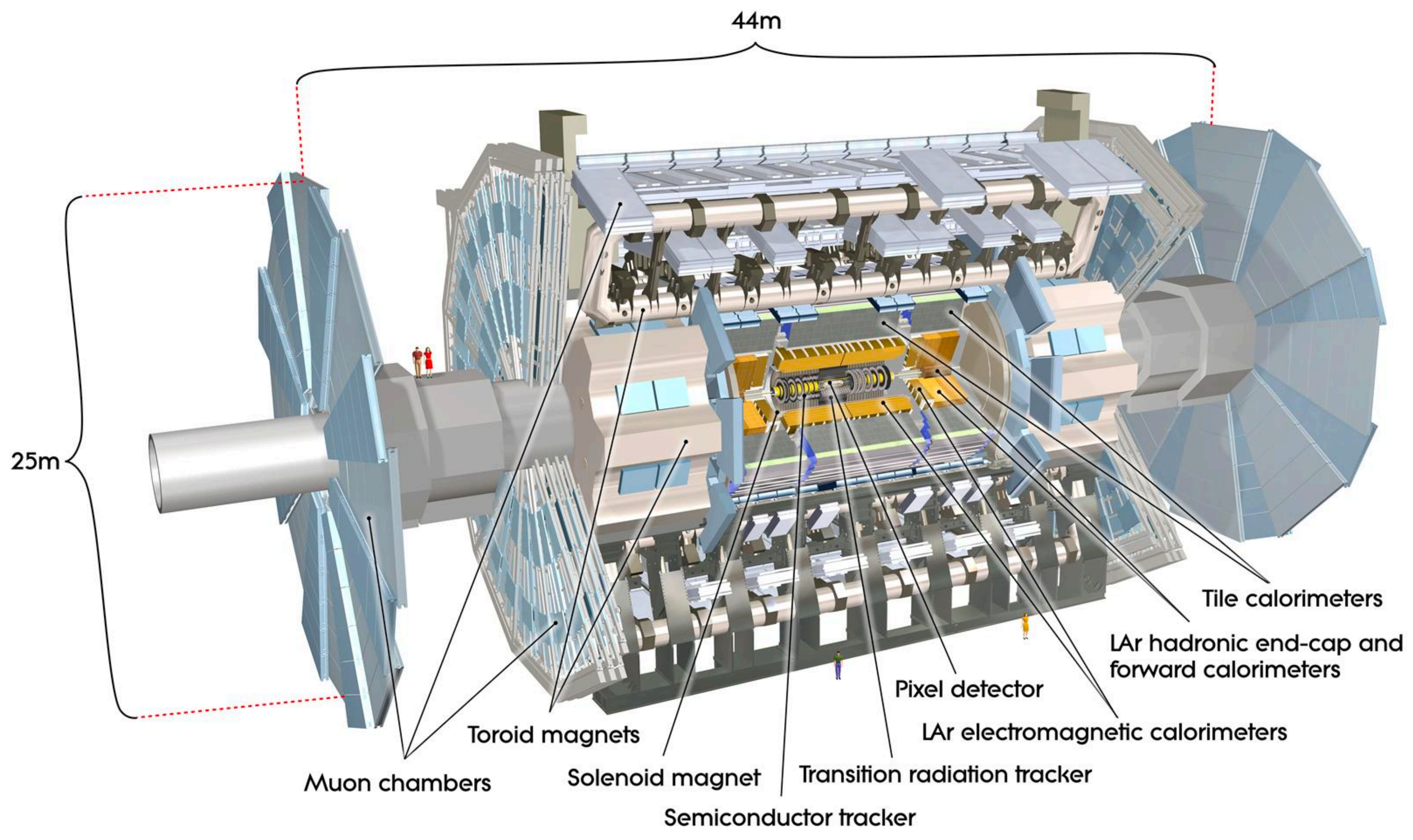
- We are currently in Long Shutdown 2 (LS2), during which we will analyze data from Run 2 (2015-2018) and make plans for Run 3 (2021-2023).

Introduction: ATLAS

- The high energy reach of the LHC gives access to many physics processes not available in other contexts

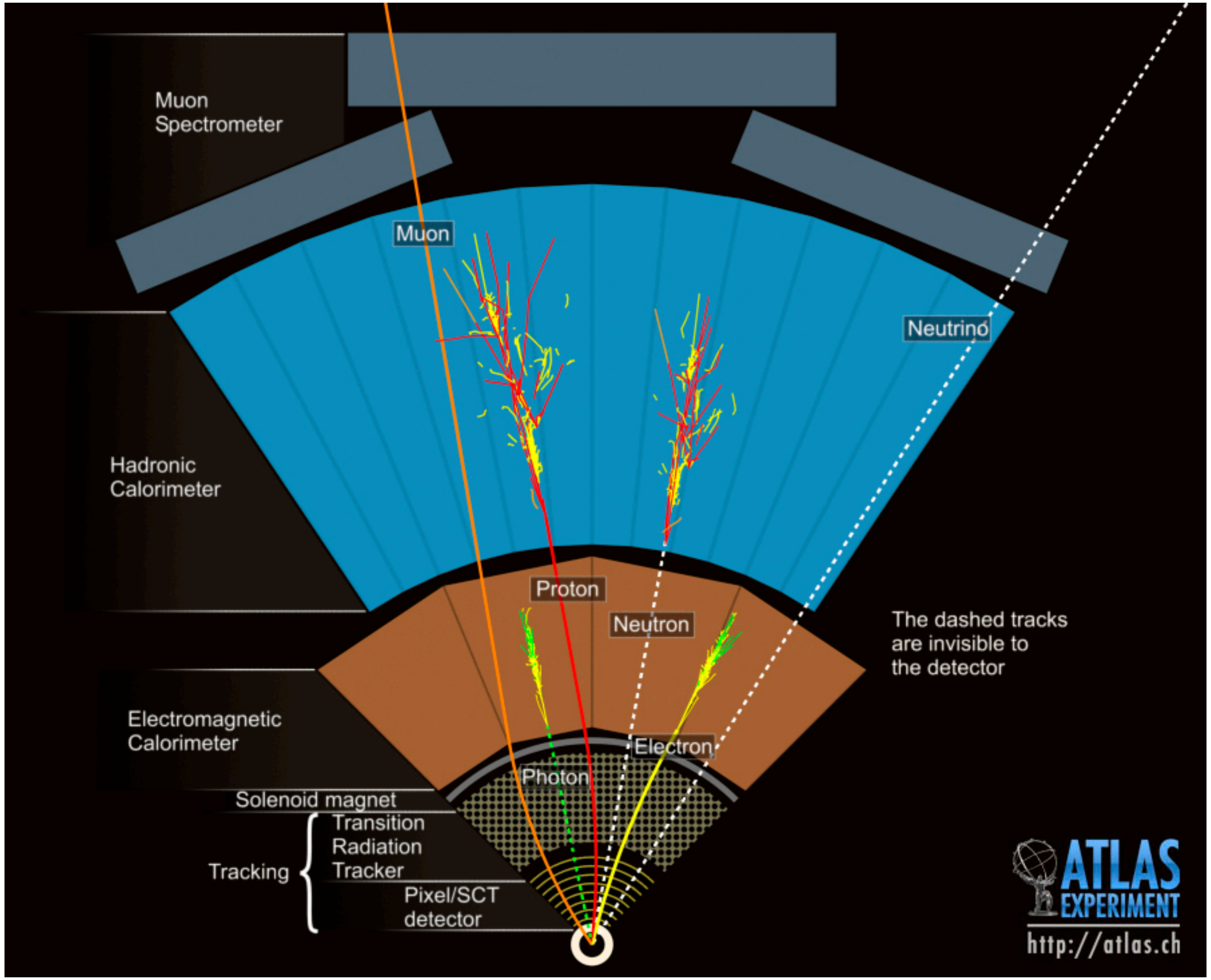
- The ATLAS experiment is one of two “general purpose” experiments at the LHC, designed to observe these processes

    - “General purpose” here means a rich and wide program of searches for new physics, precision measurements, and confirmations/ tests of the Standard Model



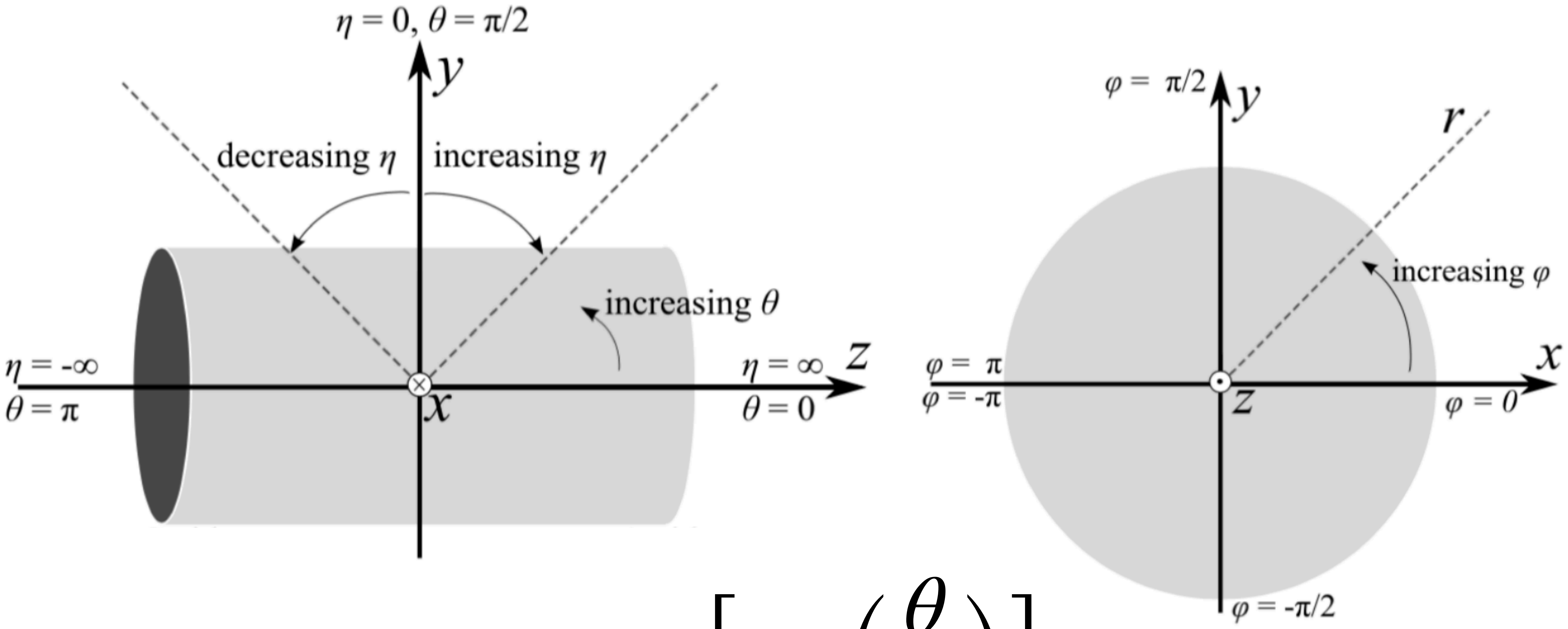
The ATLAS Detector

- The ATLAS detector is composed of several sub-detectors
  - A **tracking** detector, which, through ionization of closely spaced layers, is able to measure the path of a charged particle
    - In a magnetic field => information about the particle momentum and charge
  - A **calorimeter** which measures particle energy
    - Granularity of the calorimeter allows a fine grained look at how particles behave on interaction with the material => physics information!
  - A **muon spectrometer** which measures muon momentum/trajectory, since the muons usually pass through the calorimeter layers



A Note on Coordinates

- ATLAS has a cylindrical geometry
- Coordinates used are η (parameterizes position along beam line) and ϕ (location around the cylinder)
- Transverse momentum (p_T) is the component of momentum perpendicular to the beam line
    - Used, e.g., to remove contribution to momentum from the beam itself

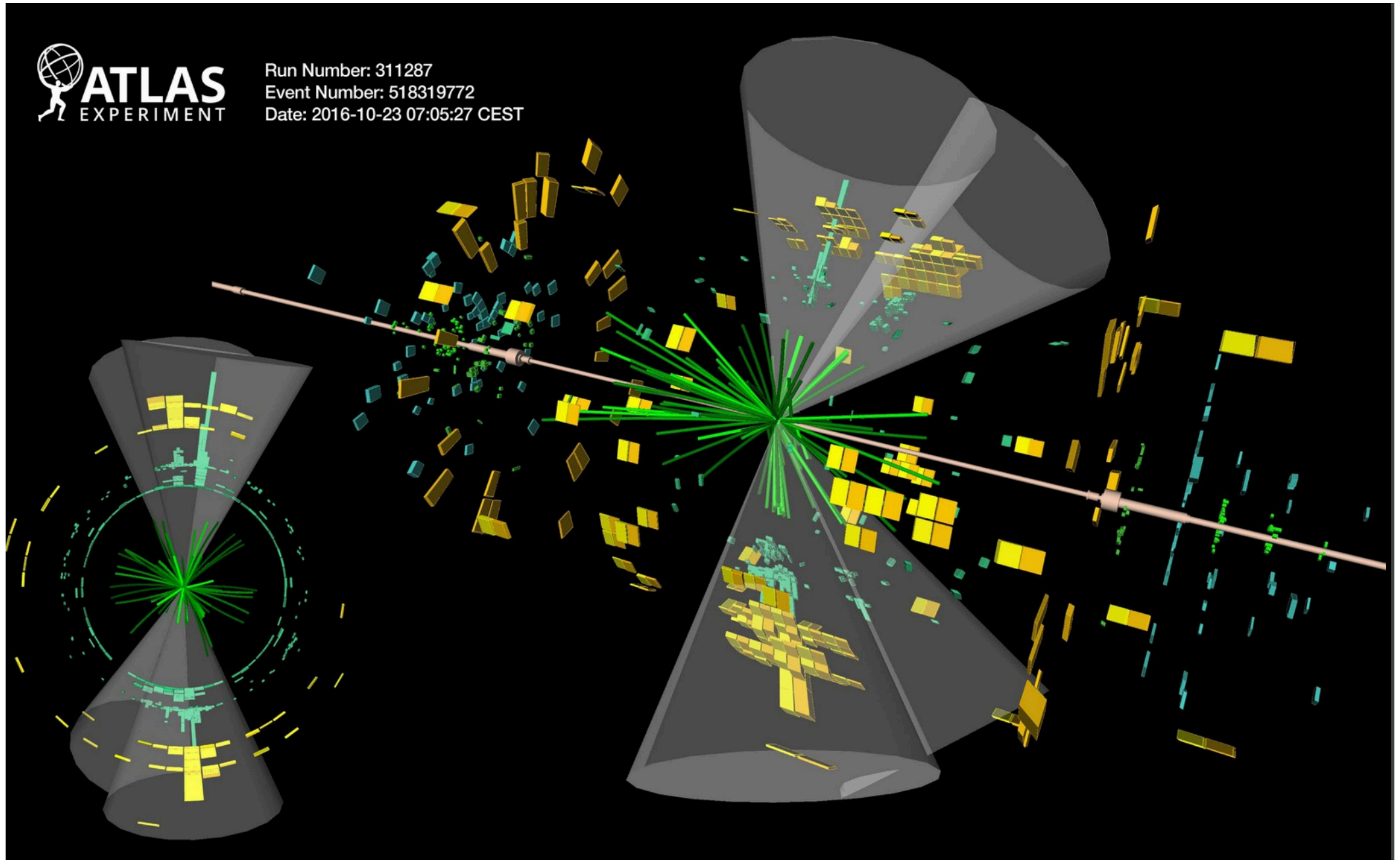


$$\eta \equiv -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]$$

$$\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$$

A Note on Jets

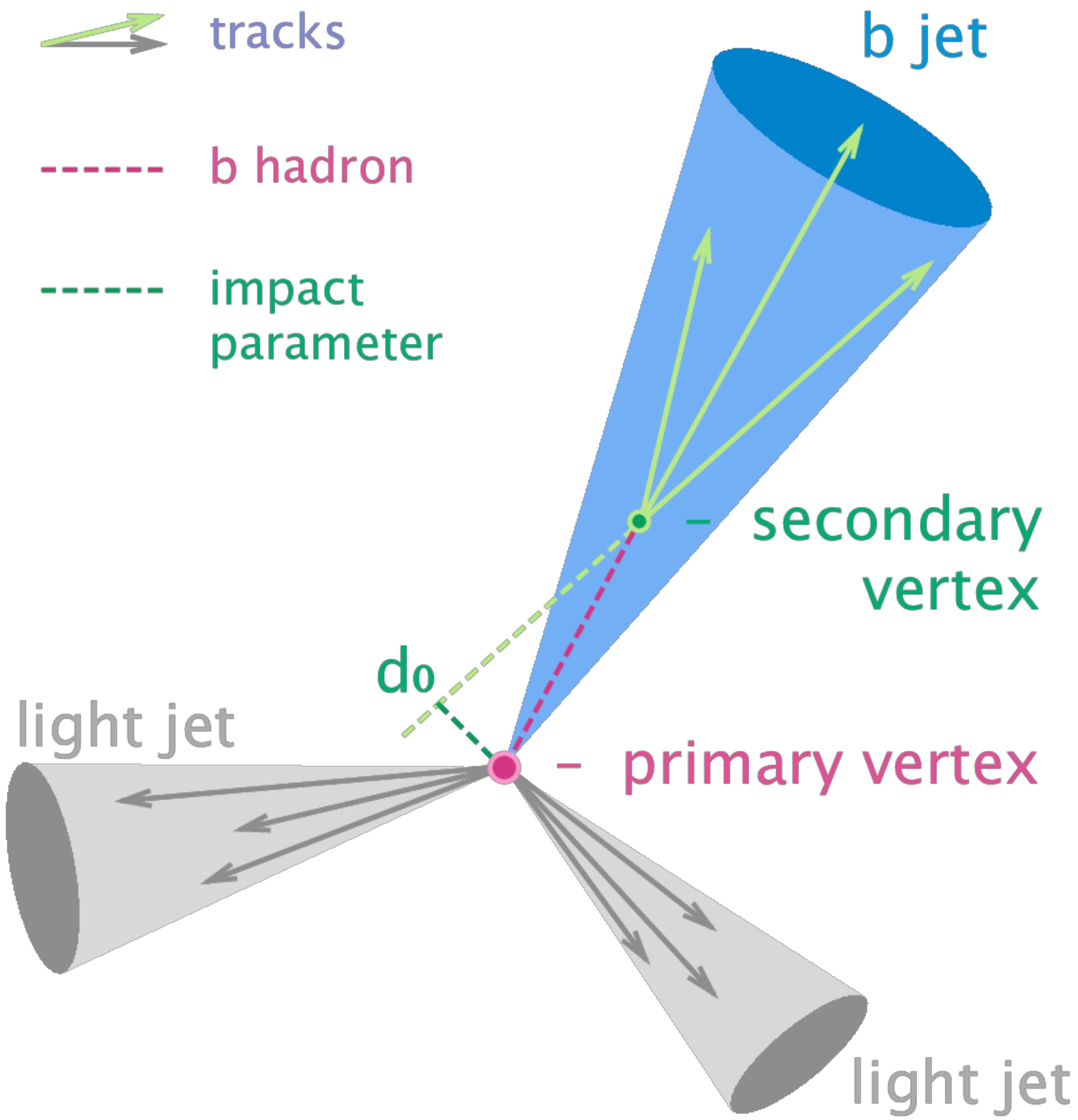
- Proton-proton collisions at the LHC produce a variety of different particles
- Free quarks and gluons cannot exist on their own (QCD confinement), but rather form into hadrons, which then decay
  - This is what we see in our detector
- Group these decay products into objects called **jets**, which serve as a reasonable proxy for the original partons
  - There are a variety of algorithms to group energy deposits into jets
  - Commonly used is the anti-k_t algorithm [ref], which constructs (roughly) a cone of a certain radius in the detector



Candidate HH→4b event from Jana Schaarschmidt, HH Workshop, Fermilab

A brief word on b-tagging

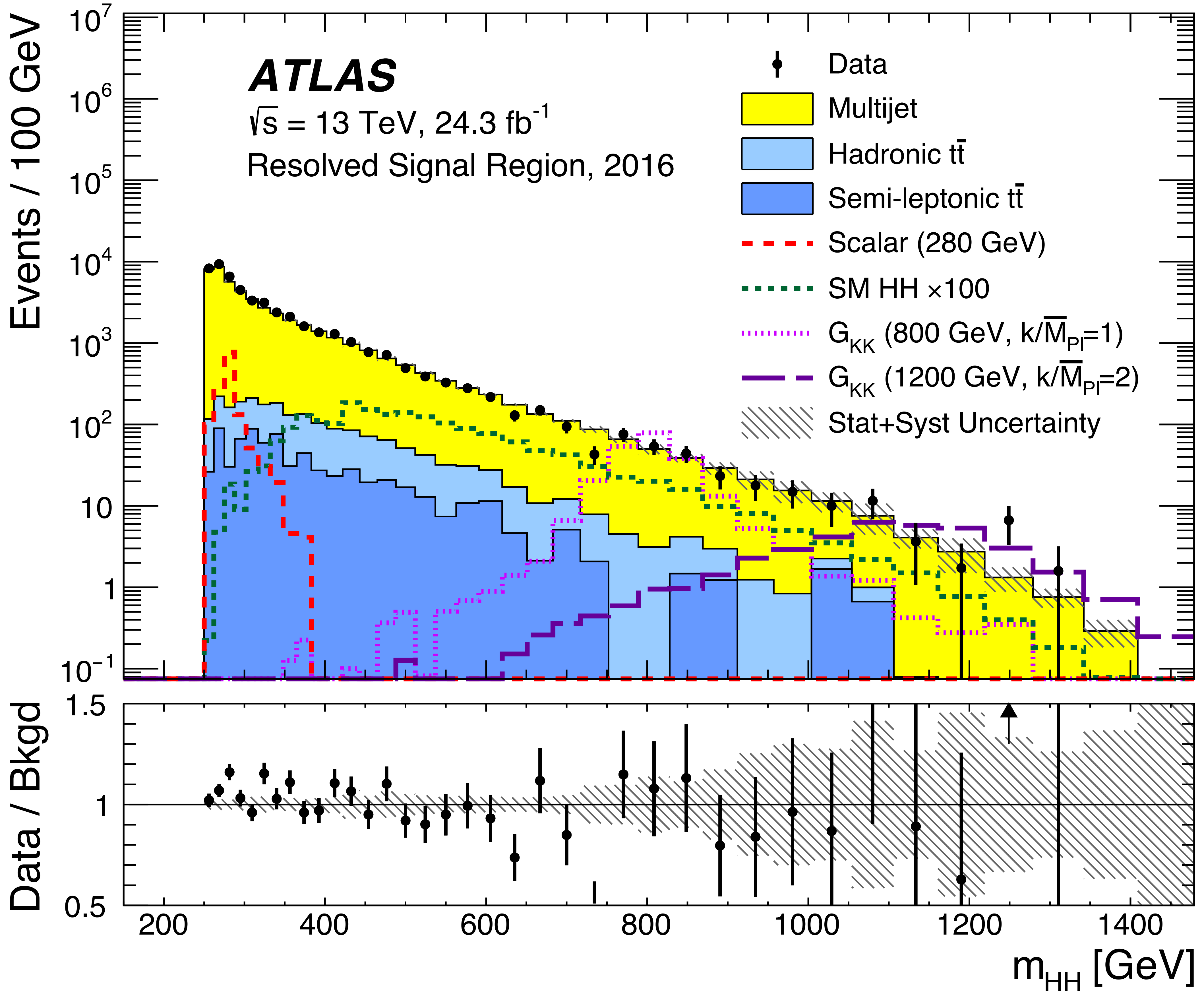
- b quarks are important in the HH→4b analysis. They are identified by looking at jets containing b-hadrons ("b-jets").
- b-hadrons have some important distinguishing features:
  - Long lifetimes => travel a measurable amount away from collision point before decaying ("displaced vertex")
  - Large masses => high decay multiplicities
- These features make it possible to identify b-jets in ATLAS
  - This is called **b-tagging** and is a very active area of research
  - Current standard uses machine learning tools for b jet identification. These output a **b-tagging score**, which is then used for selection



ATLAS Qualification

ATLAS Qualification Task: FastCaloSim

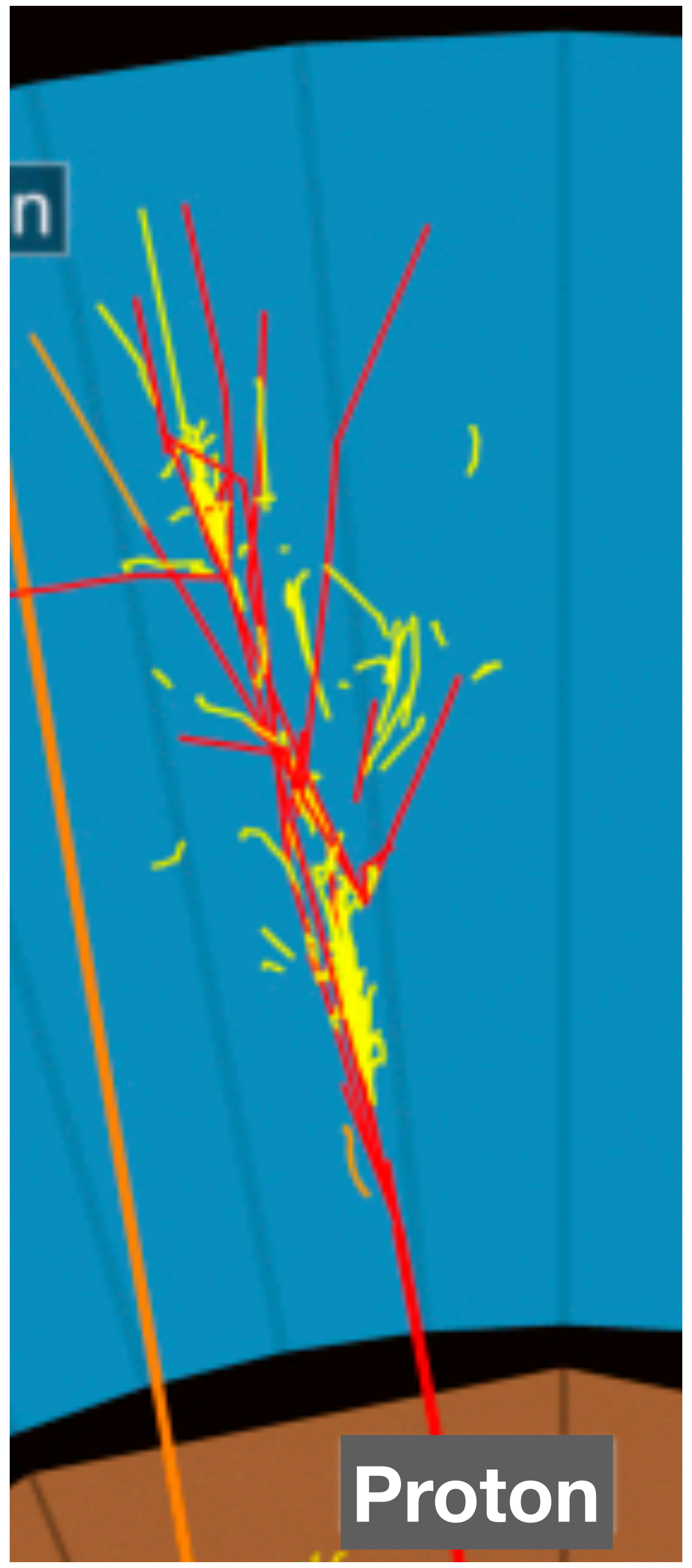
- Physics simulation is an important part of the research program in ATLAS
    - What do our signal models look like? How do we model our known (irreducible) backgrounds?
- However, it is computationally expensive! The largest chunk of CPU time is spent on the calorimeter [ref]
- **FastCaloSim** is an effort to reduce the load by parameterizing the calorimeter response in various ways
    - Goal is to have a lightweight simulation tool that still provides a high quality simulation
- NB: FastCaloSim will likely be the default for simulated samples in run 3

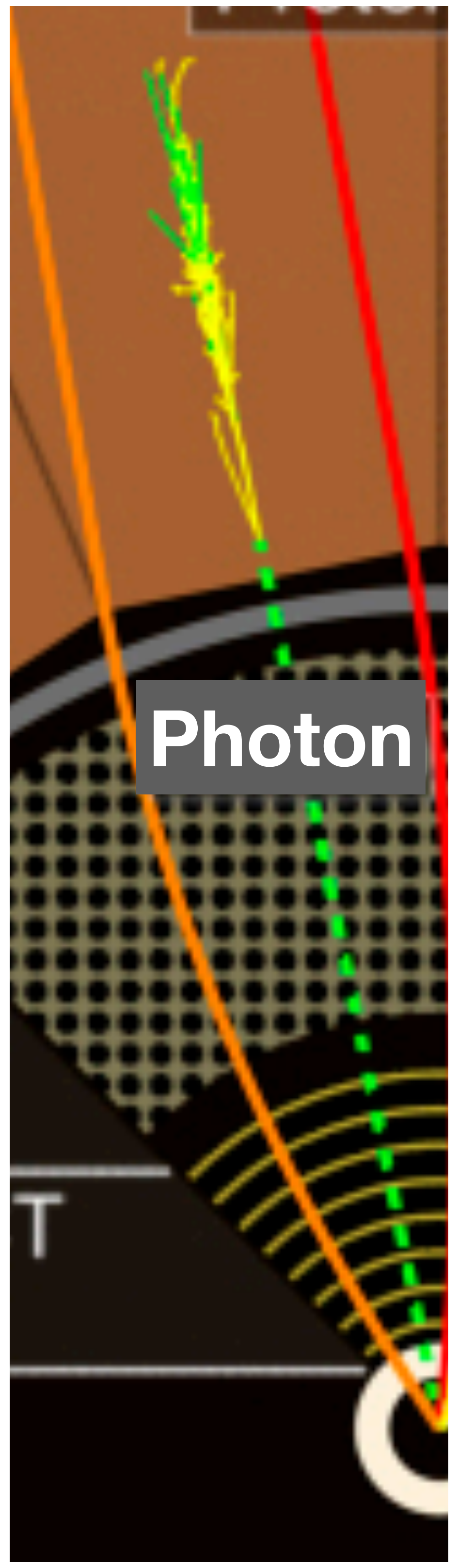


Example plot from the 2016 HH→4b analysis. $t\bar{t}$ (blue) and all signal models (colored lines) taken from Monte Carlo simulation.

FastCaloSim: The Problem

- Showering of particles in the ATLAS detector is very different for different types of particles
- Electromagnetic showers (photons, electrons, e.g.) are in general simpler, easy to model
- Hadronic showers (pions, e.g.) are much more complicated in general => very non-trivial detector response
- Modeling shower shape well is an important part of a good detector simulation
  - Allows for the use of **substructure** – how is energy distributed (e.g., within a jet), and what does this mean for physics?
  - This modeling is one of the final major issues to be resolved before broader adoption of FastCaloSim

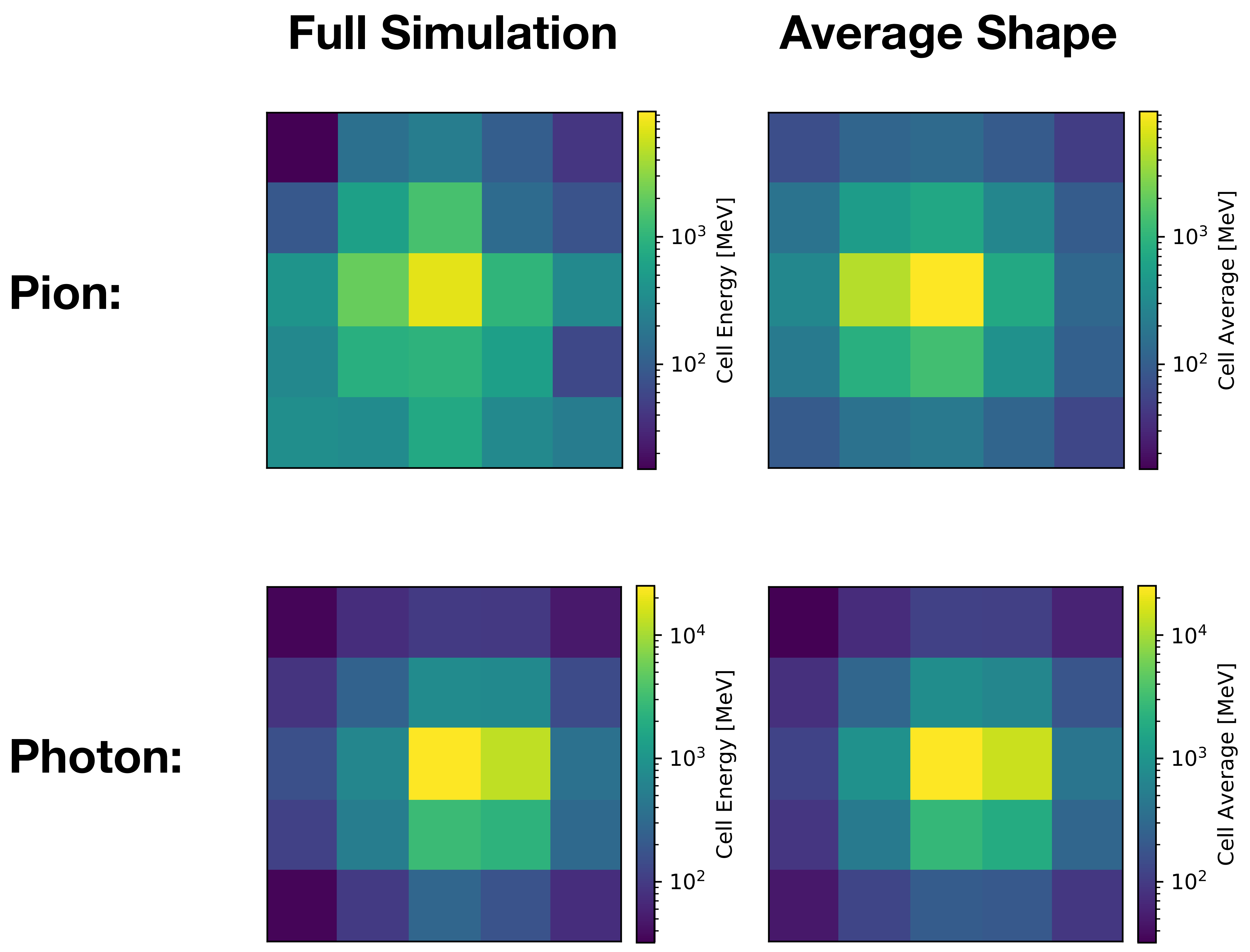




Example hadronic (proton) and EM (photon) showers

FastCaloSim: Current Status

- FastCaloSim is based on parameterization of the full simulation
- Idea is to parameterize the shower shape in some way
- Current approach:
    - Construct an average shape (for, e.g. a pion/photon)
    - Randomly fluctuate about that shape (Poisson noise)
    - Neglects correlations between “fluctuations” away from the average - for hadronic showers, these can be quite non-trivial!



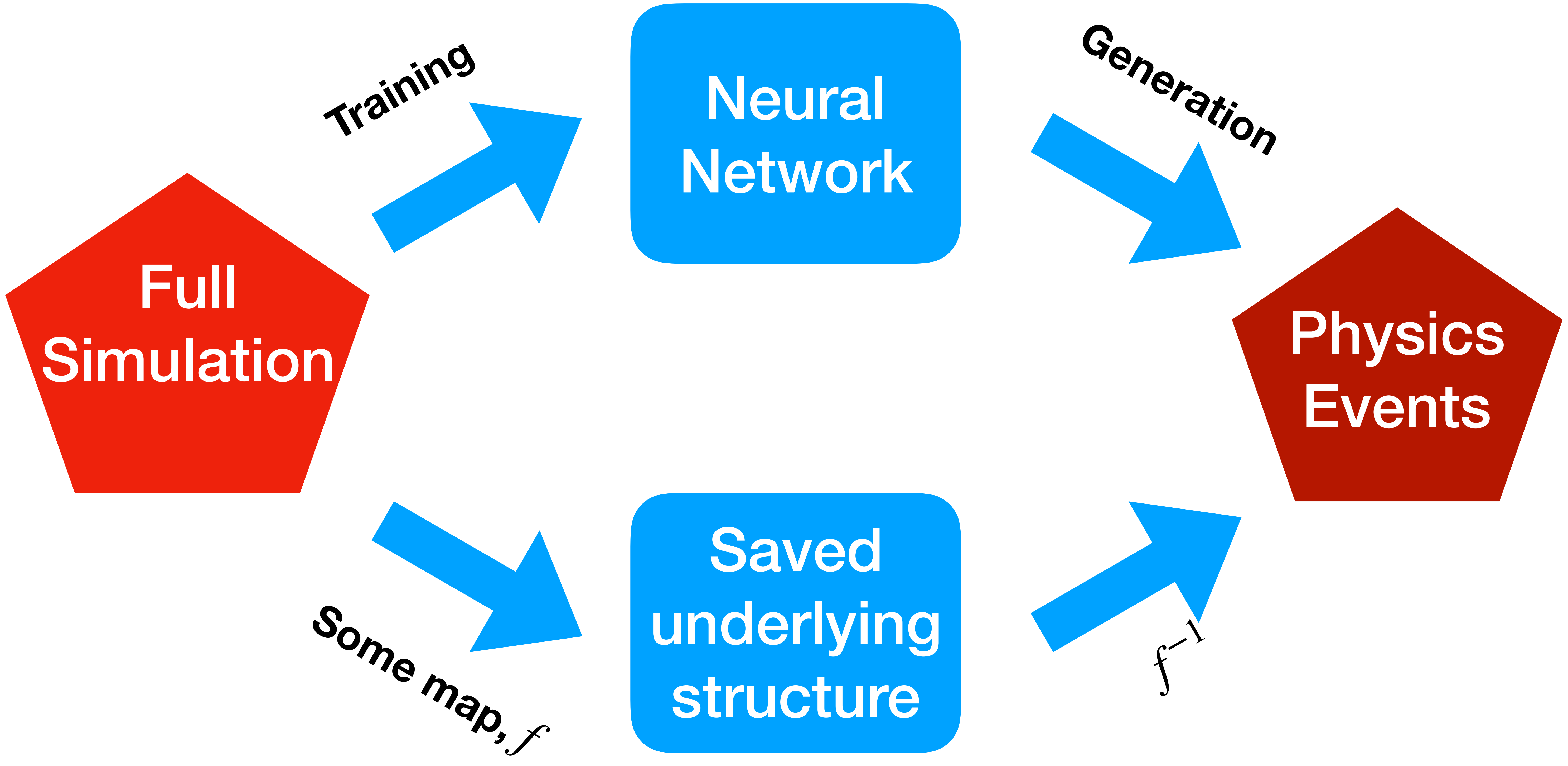
Example of a pion and a photon event in one of the calorimeter layers. Photon is much more similar to the average

FastCaloSim: New Approaches

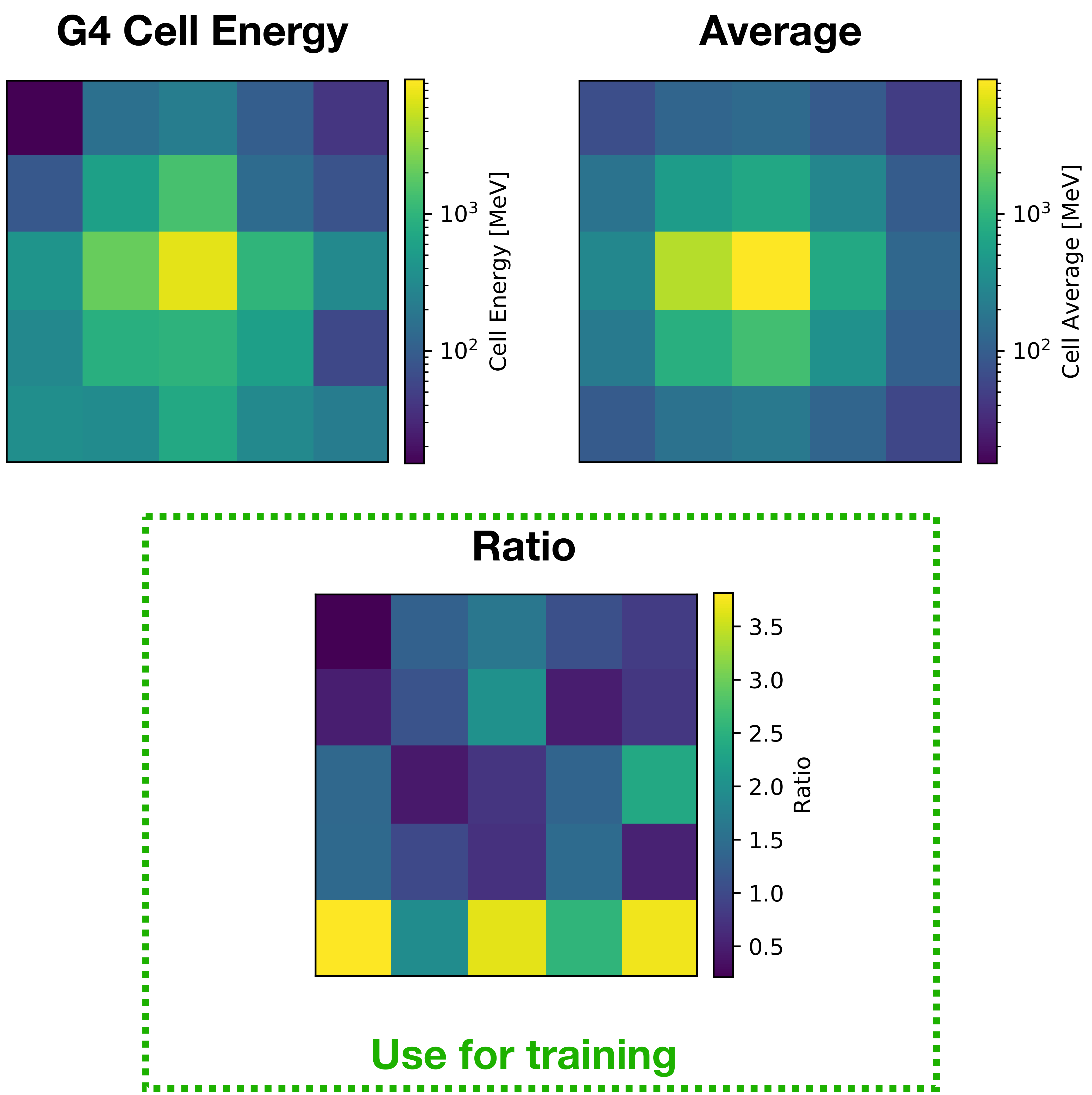
- Two approaches:

  1. Machine learning! Train a neural network to reproduce realistic structure

  2. Non ML-based! Find a convenient representation of the input data for generating new events

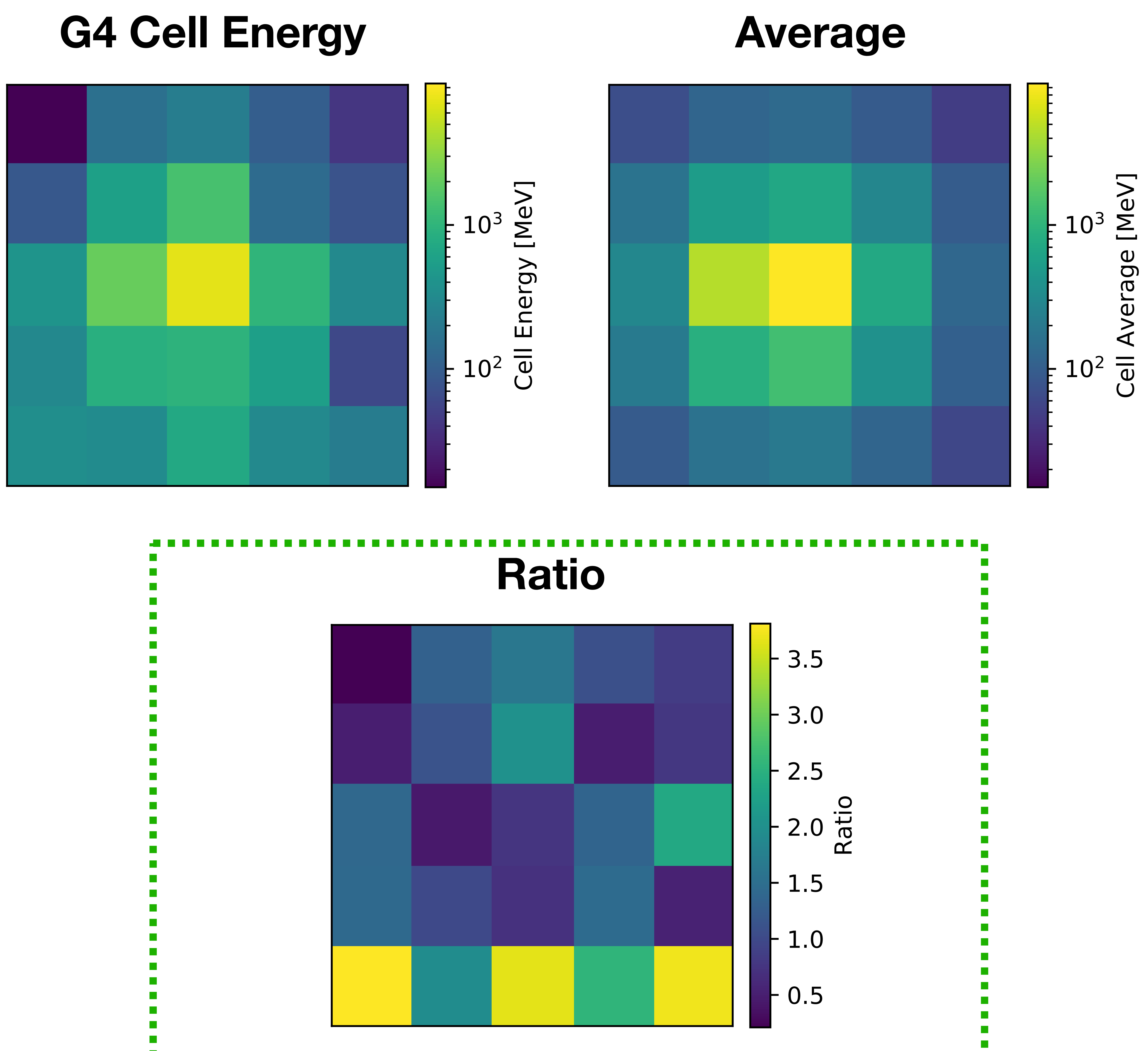


FastCaloSim: Inputs and Goal



- For both methods, we use the full physics simulation as an input
    - This full simulation is from a tool called Geant4
    - Roughly: Follows each particle in steps through the detector, simulating interactions with the material
- We then look at ratios of full simulation events to the corresponding average shape (what we call “fluctuations”)
- These ratios are the inputs to our methods

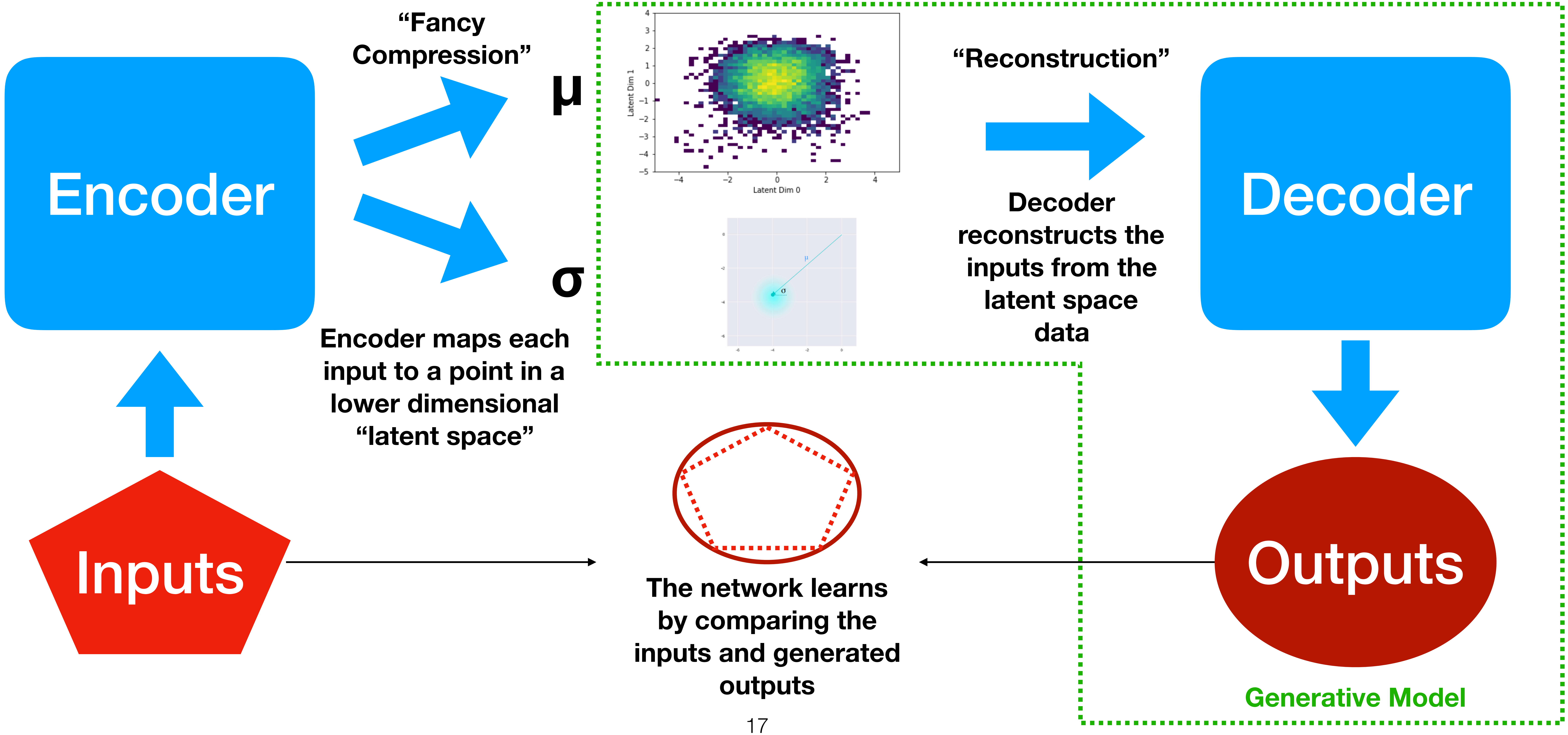
FastCaloSim: Inputs and Goal



Use for training

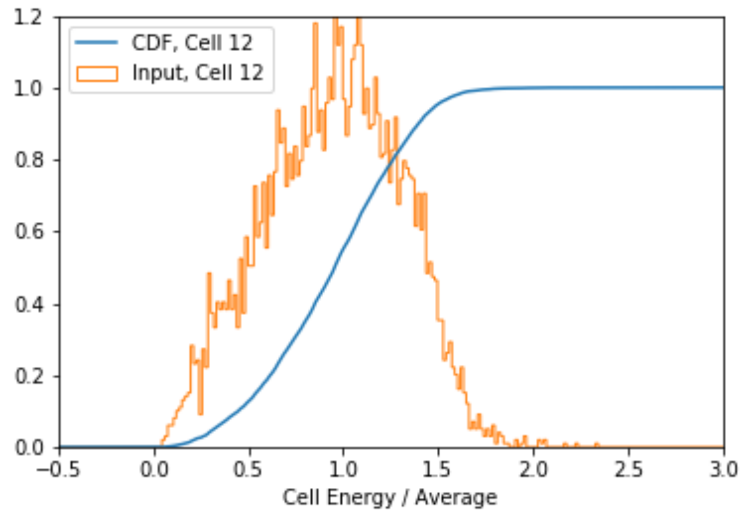
- Each event is a grid of n_x x n_y calorimeter cells
  - A calorimeter cell corresponds to the finite granularity of our physical detector
- We currently examine layers of the calorimeter individually

FastCaloSim: Variational Autoencoder (VAE)

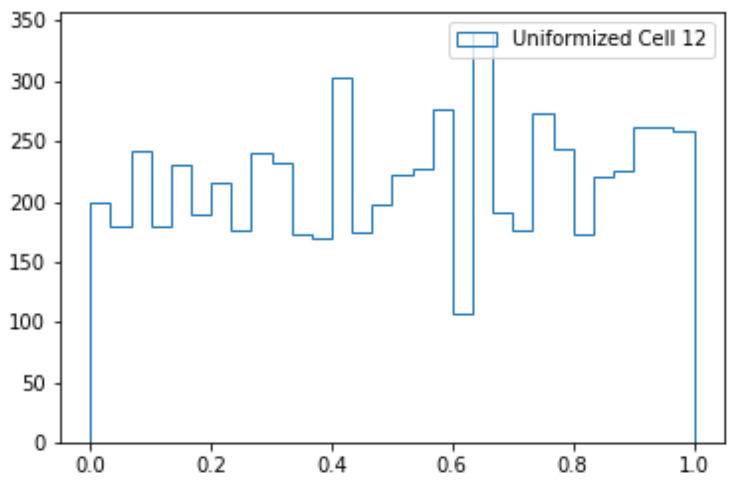


FastCaloSim: Gaussian Method

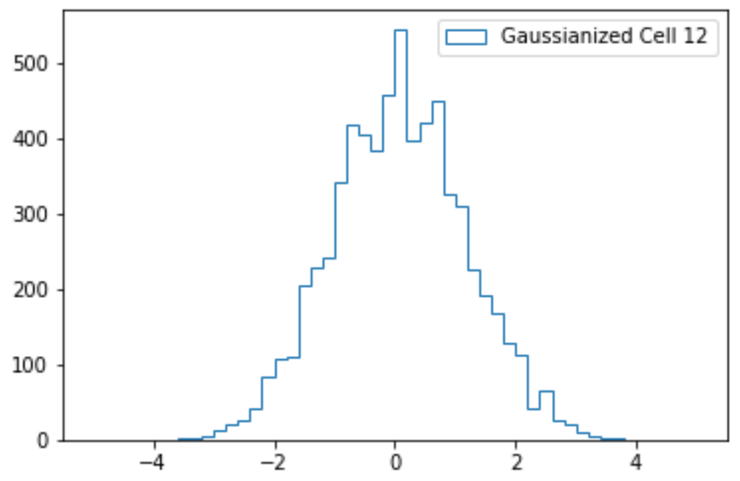
1. Construct CDF from input distribution

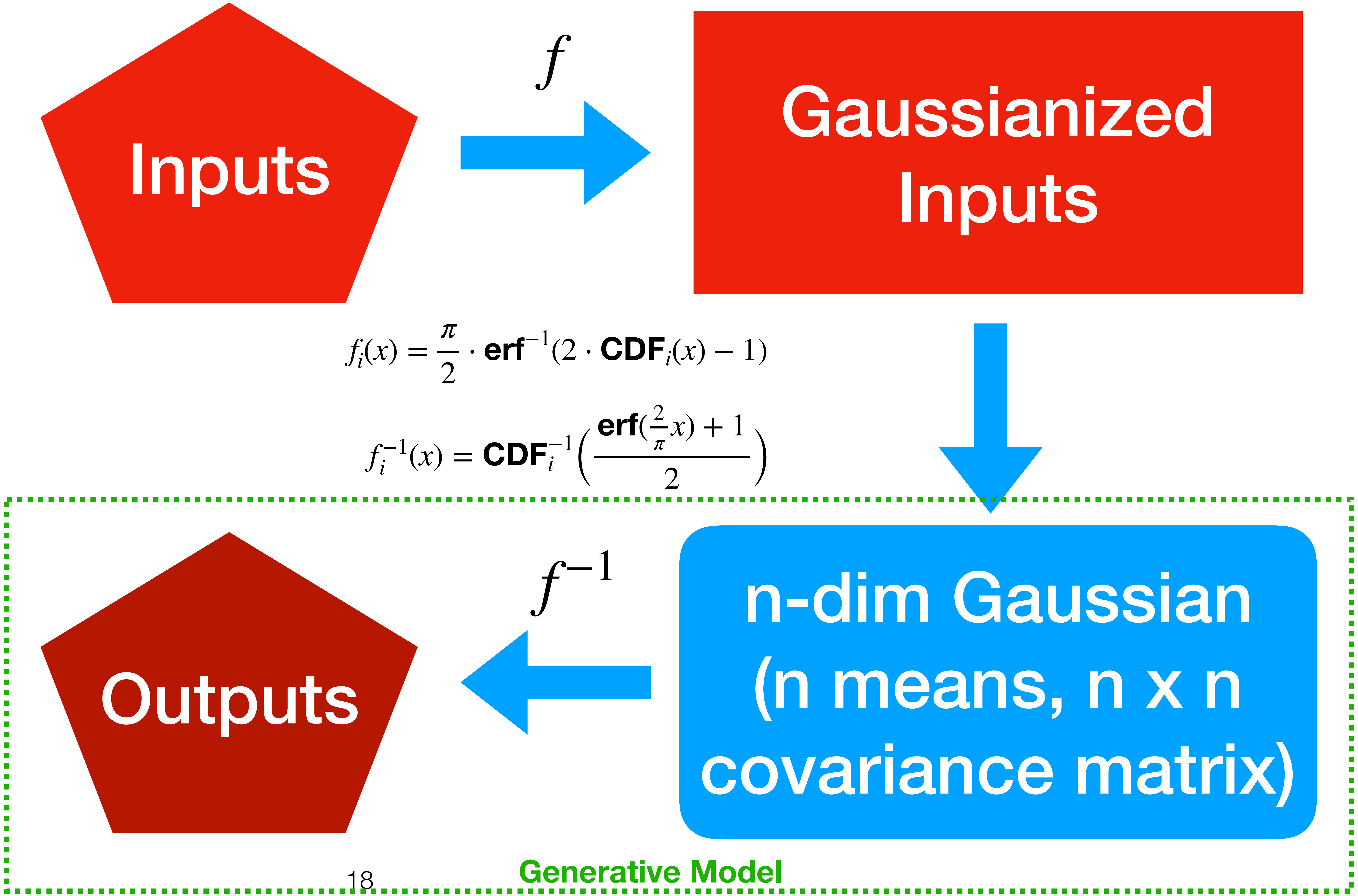


2. Uniformize by sampling from CDF (CDF(x) for each x value in input)



3. erfinv(uniform) = Gaussian



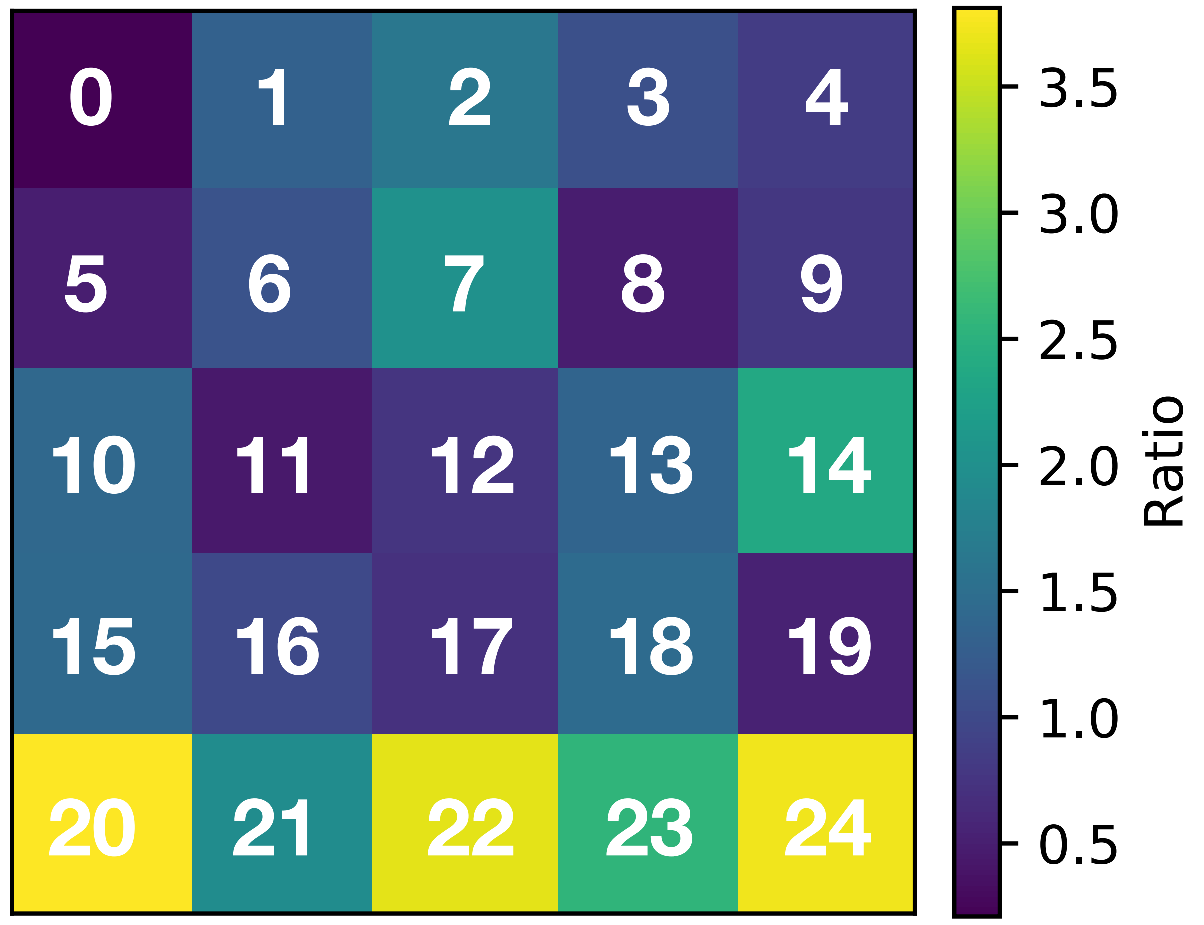


$$f_i(x) = \frac{\pi}{2} \cdot \mathbf{erf}^{-1}(2 \cdot \mathbf{CDF}_i(x) - 1)$$

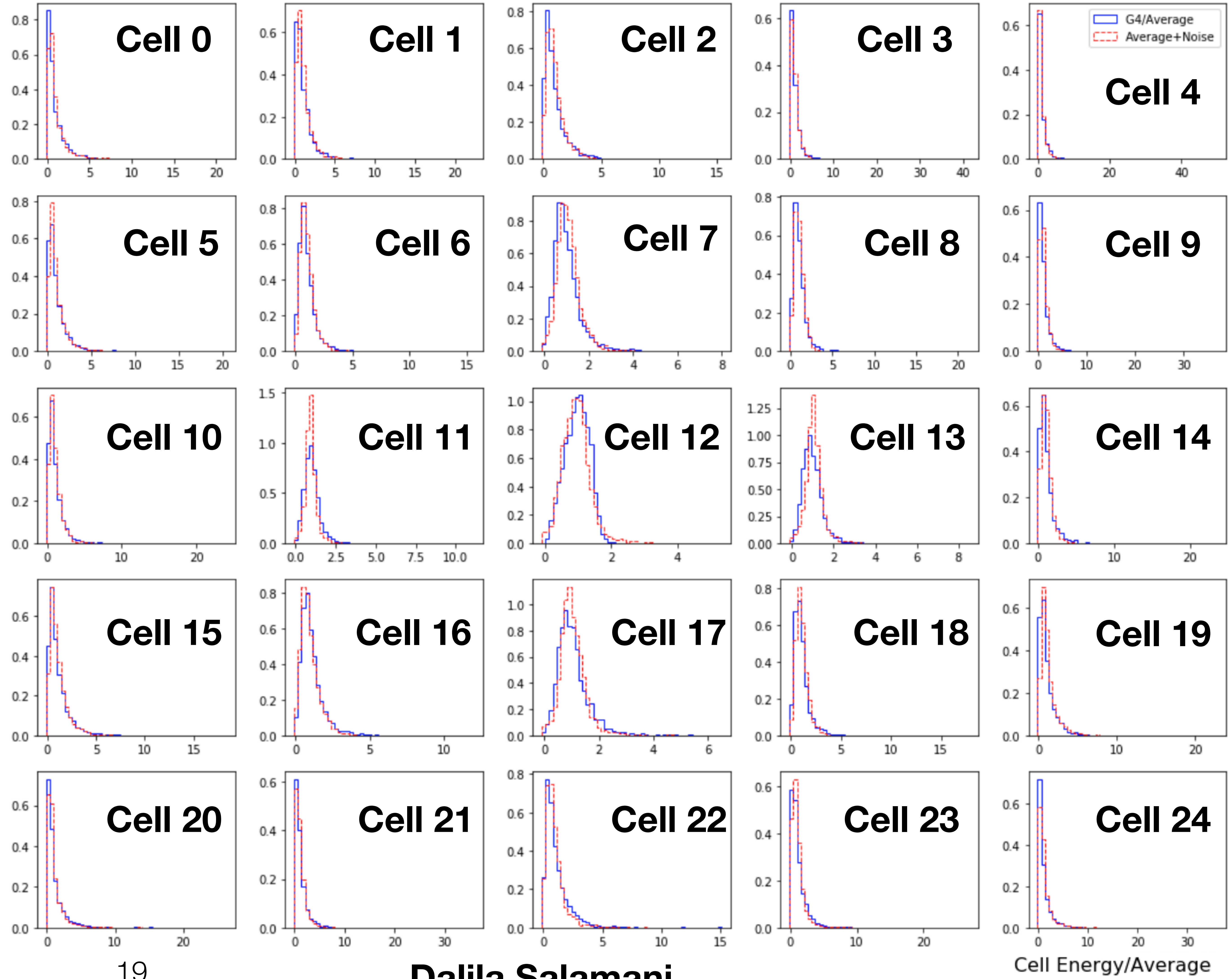
$$f_i^{-1}(x) = \mathbf{CDF}_i^{-1}\left(\frac{\mathbf{erf}(\frac{2}{\pi}x) + 1}{2}\right)$$

VAE Results

Ratio (Full Sim Cell Energy / Average Sim Cell Energy, 1 event)

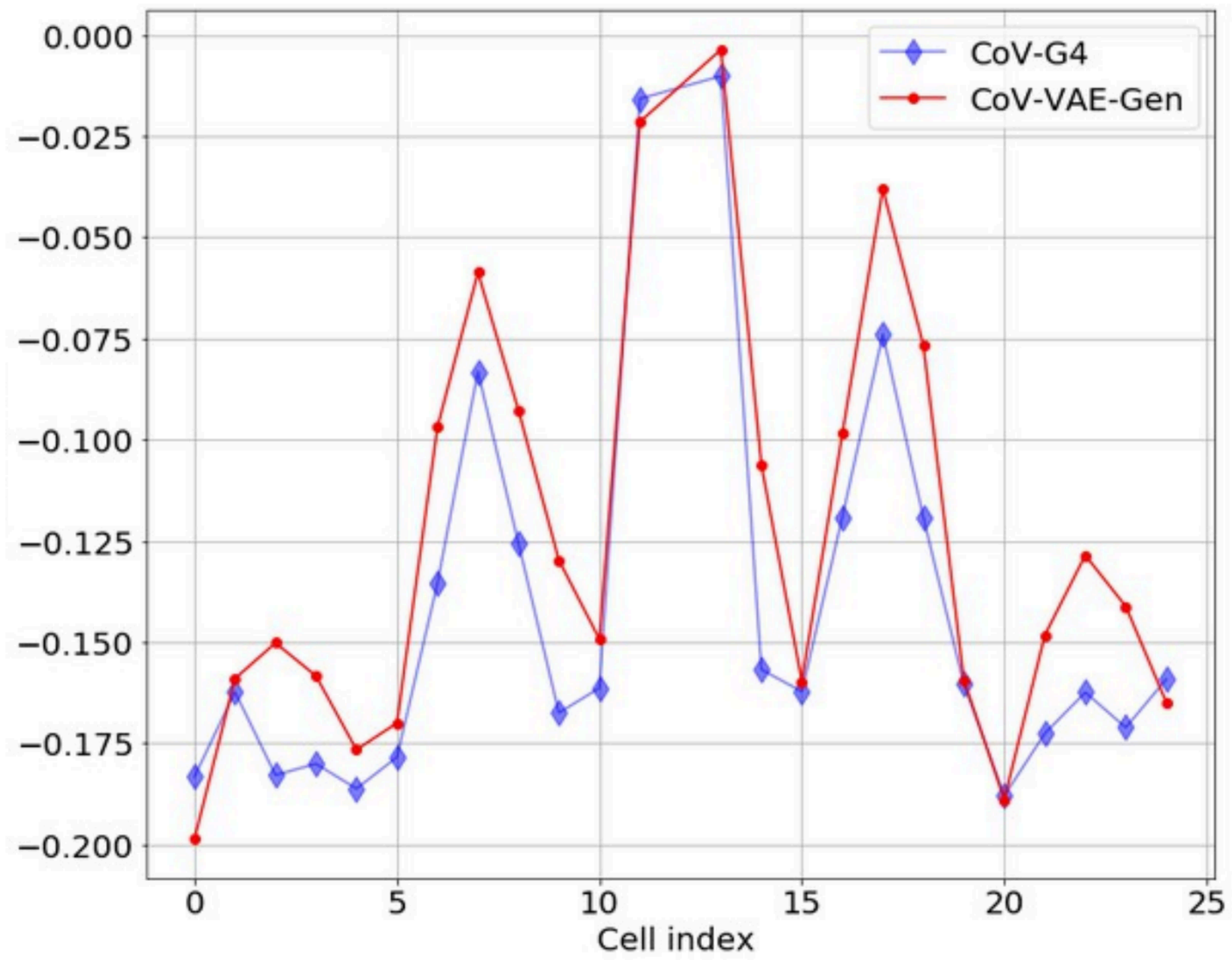


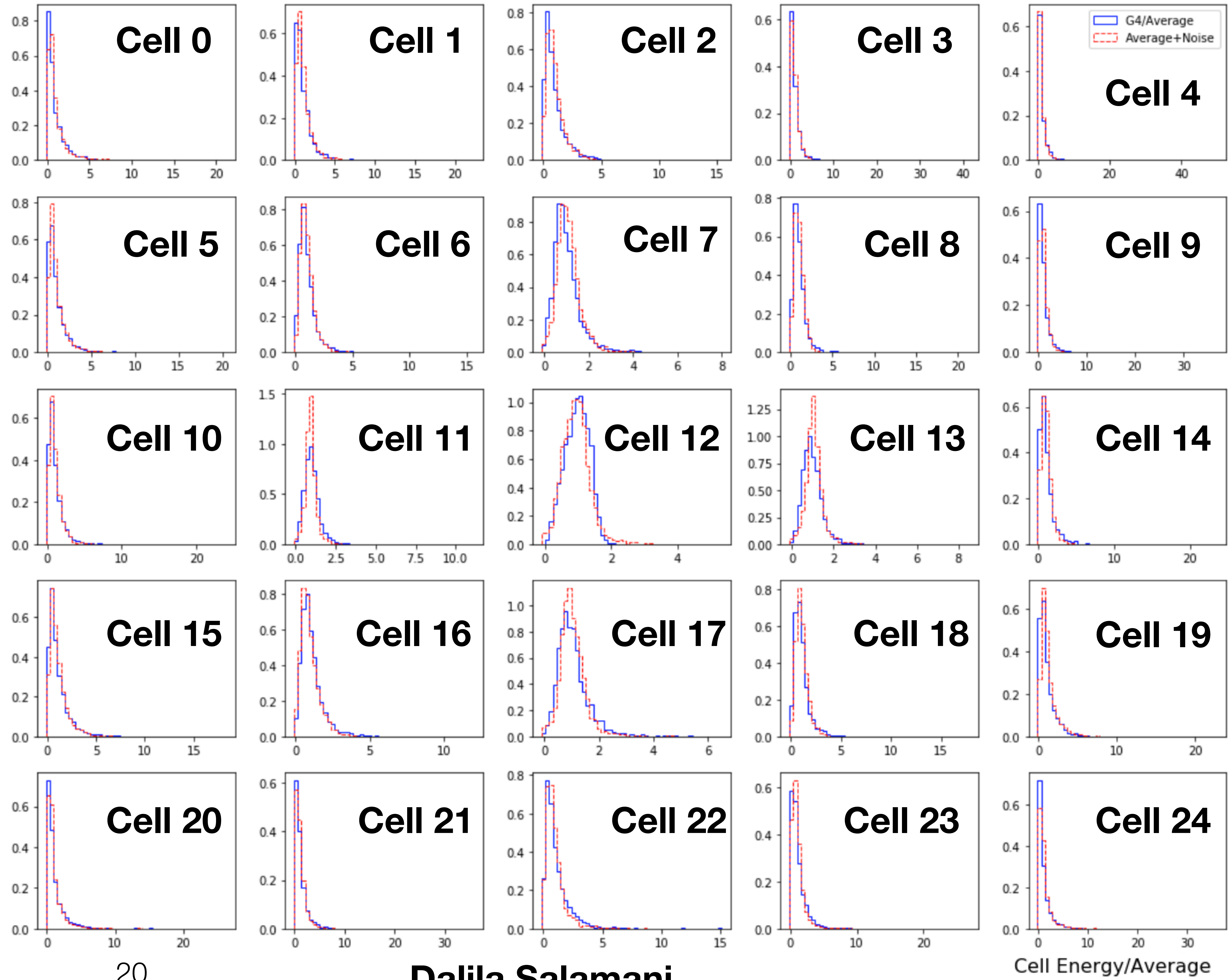
Histogram over set of events



- 65 GeV pions, EMB2, 5x5 cell grid
- Distributions shown are the ratios (in each cell) to the average shape

VAE Results

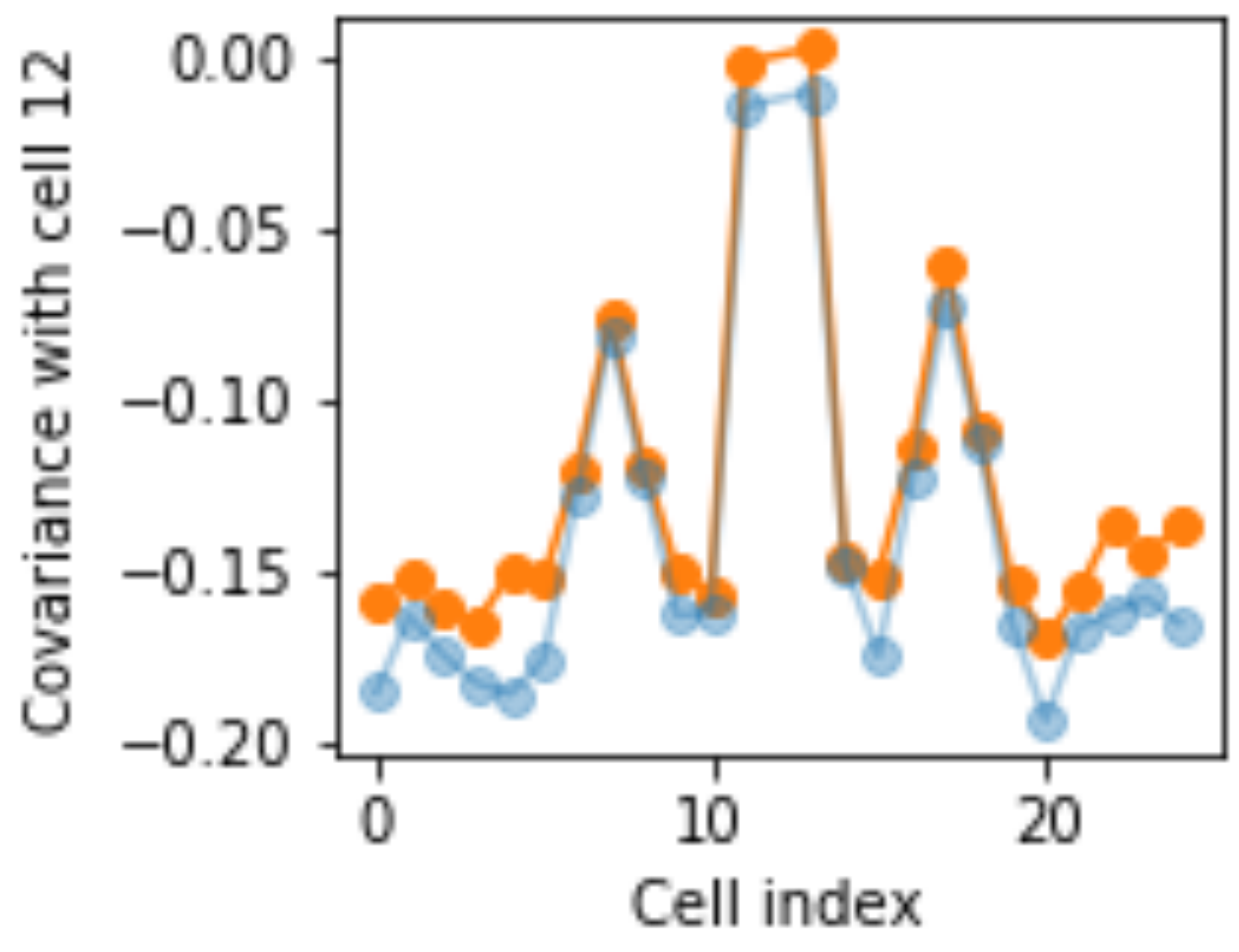


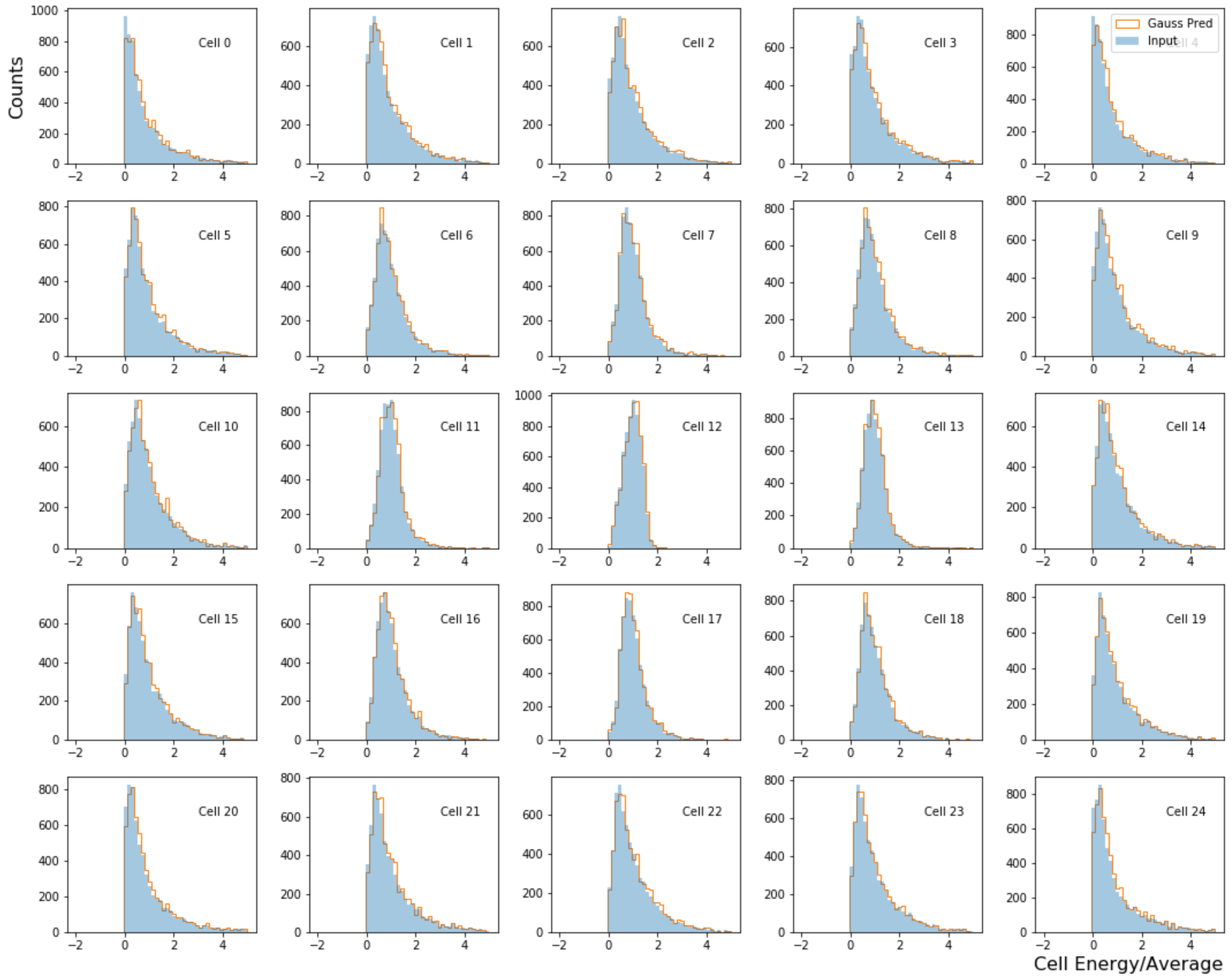


- Performs well! Covariance and distributions are well modeled

Dalila Salamani

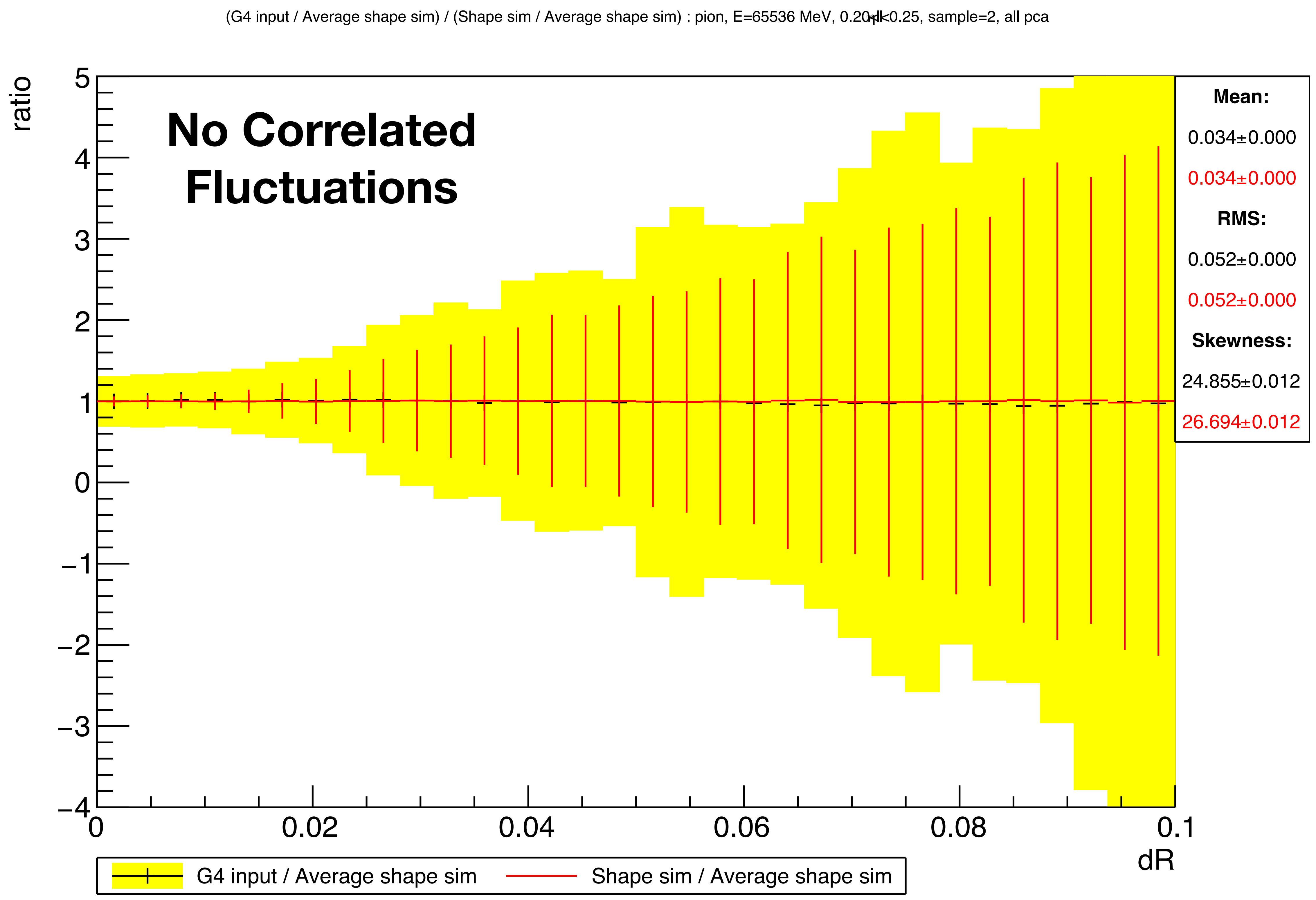
Gaussian Method Results

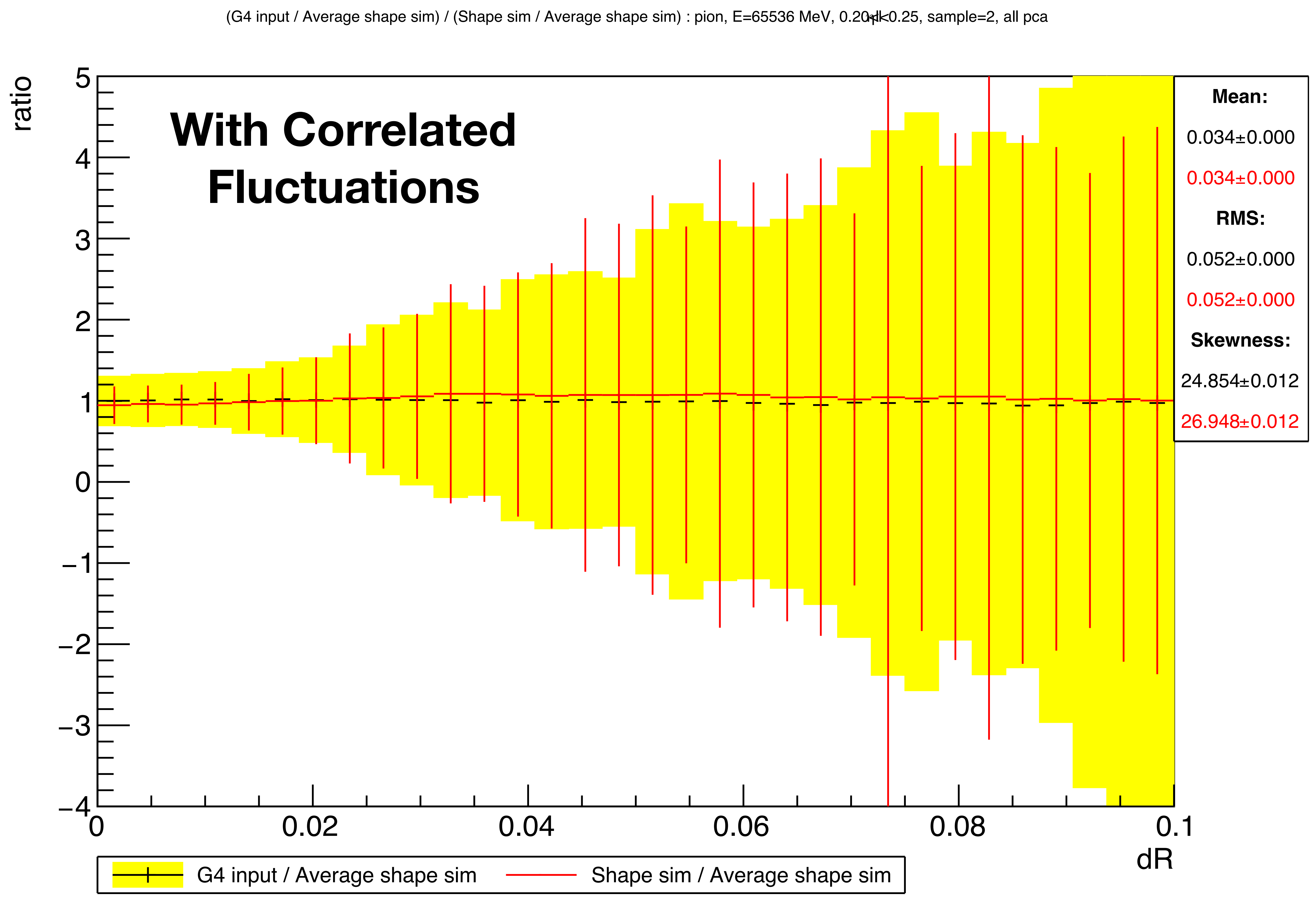




- Again 65 GeV pions, EMB2, 5x5 cell grid
- Performance is also very good!

Layer 2 (EMB2): Ratios





- Plots from the central FastCaloSim shape validation
- Nice improvement in the core

ATLAS Qualification: Status

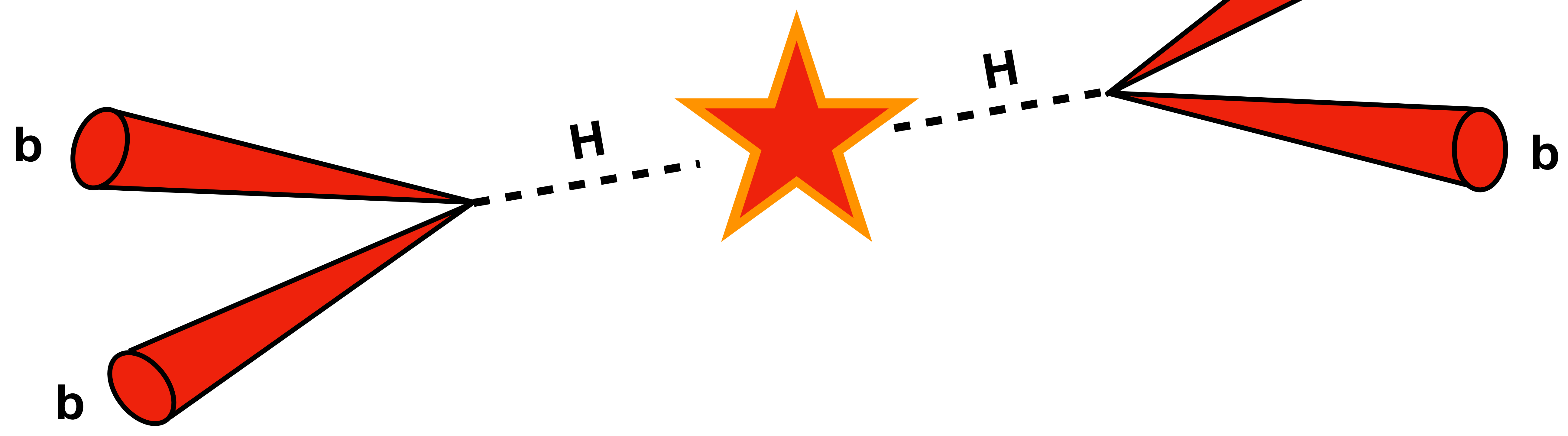
- My Contributions:
  - **VAE method**: Initial studies on VAE vs other methods. Developed much of final network structure.
  - **Gaussian Method**: Current focus, developed and studied almost entirely by me.
- I'm qualified!
  - Gaussian method is implemented in standalone FastCaloSim code and has been studied for pions across various calorimeter layers
    - VAE studies ongoing
- Future work:
  - Need to examine effect at different energies/eta points
  - Need to examine the physics impact - what does a simulated sample look like with/without this modeling?

Physics Analysis: $HH \rightarrow b\bar{b}b\bar{b}$

HH→4b: Motivation and Benchmarks

HH→4b: Introduction

- General concept:
  - Two Higgs bosons are produced from a proton-proton collision
  - Each of these decay into two b quarks
- Important questions:
  - How is the HH produced?
    - Production modes, benchmarks
  - Why 4b?



Standard Model HH Production

- Gluon-gluon fusion (ggF) accounts for more than 90% of Standard Model HH production at the LHC
- We consequently focus on ggF production for this analysis
- For reference, the single Higgs ggF production cross section is ~46.86 pb [ref] = 46860 fb
  - ~1500 x the HH cross section!

HH Production Mode	$\sqrt{s}$ = 13 TeV Cross Sections [fb]
ggF HH	31.05
VBF HH	1.73
ZHH	0.363
W^+ HH	0.329
W^- HH	0.173
ttHH	0.775
tjHH	0.0289

LHCHXSWG

HH Physics Interest

- Show here the relevant Feynman diagrams for ggF production

- Triangle diagram => signal models:

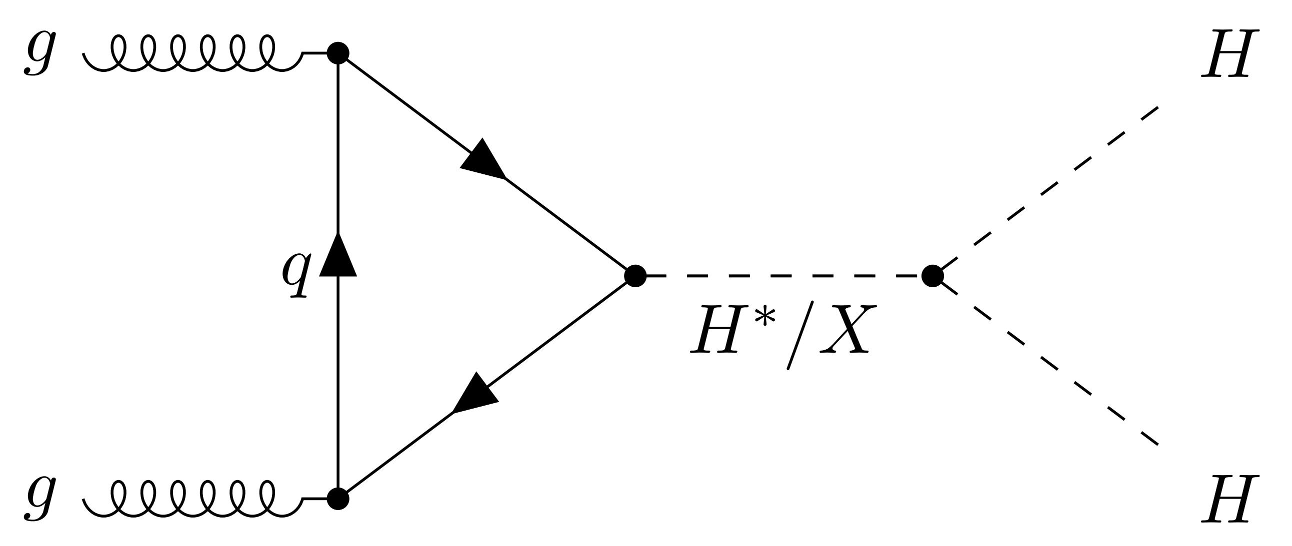
  - HH production via decay of a heavy resonance (X)

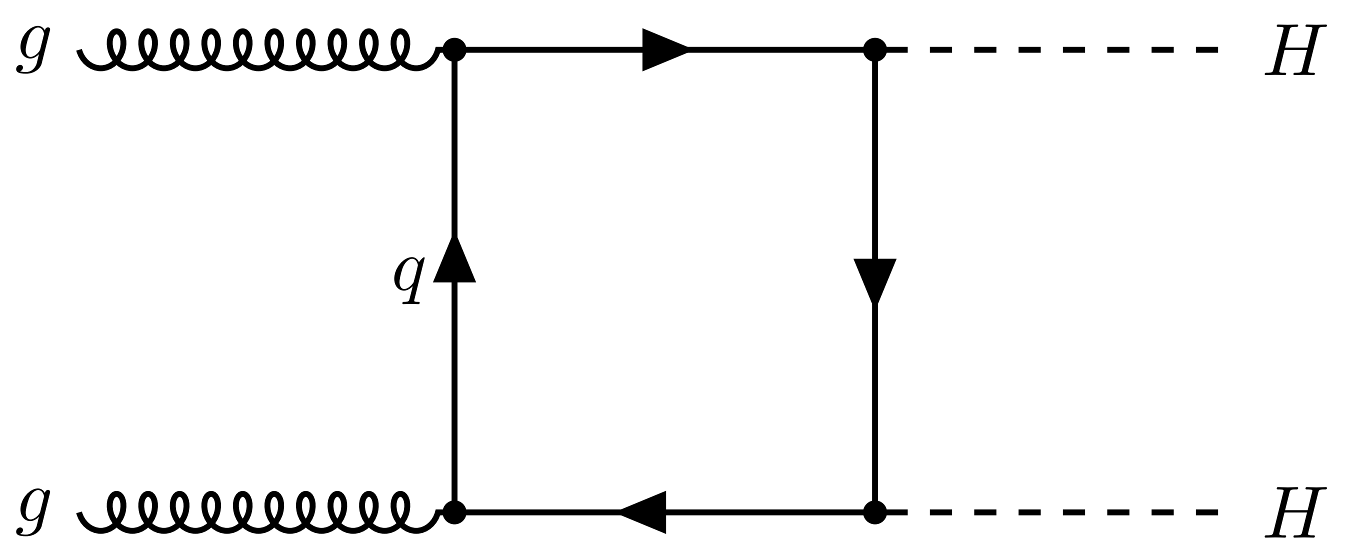
  - Non-resonant (off-shell) HH production

- Box diagram => interference

  - Extra fermion line => relative minus sign between the two diagrams

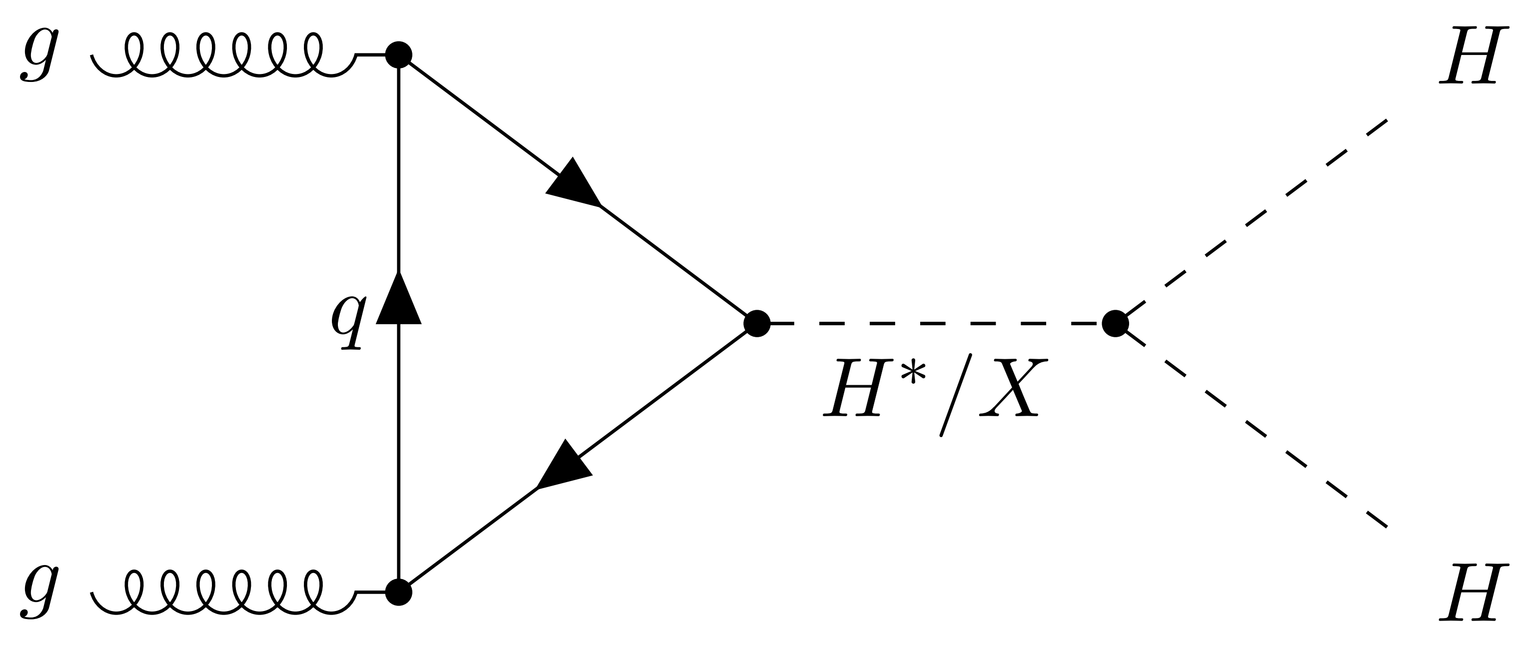
  - Summing diagrams to get the cross section then leads to a cancellation due to this negative contribution





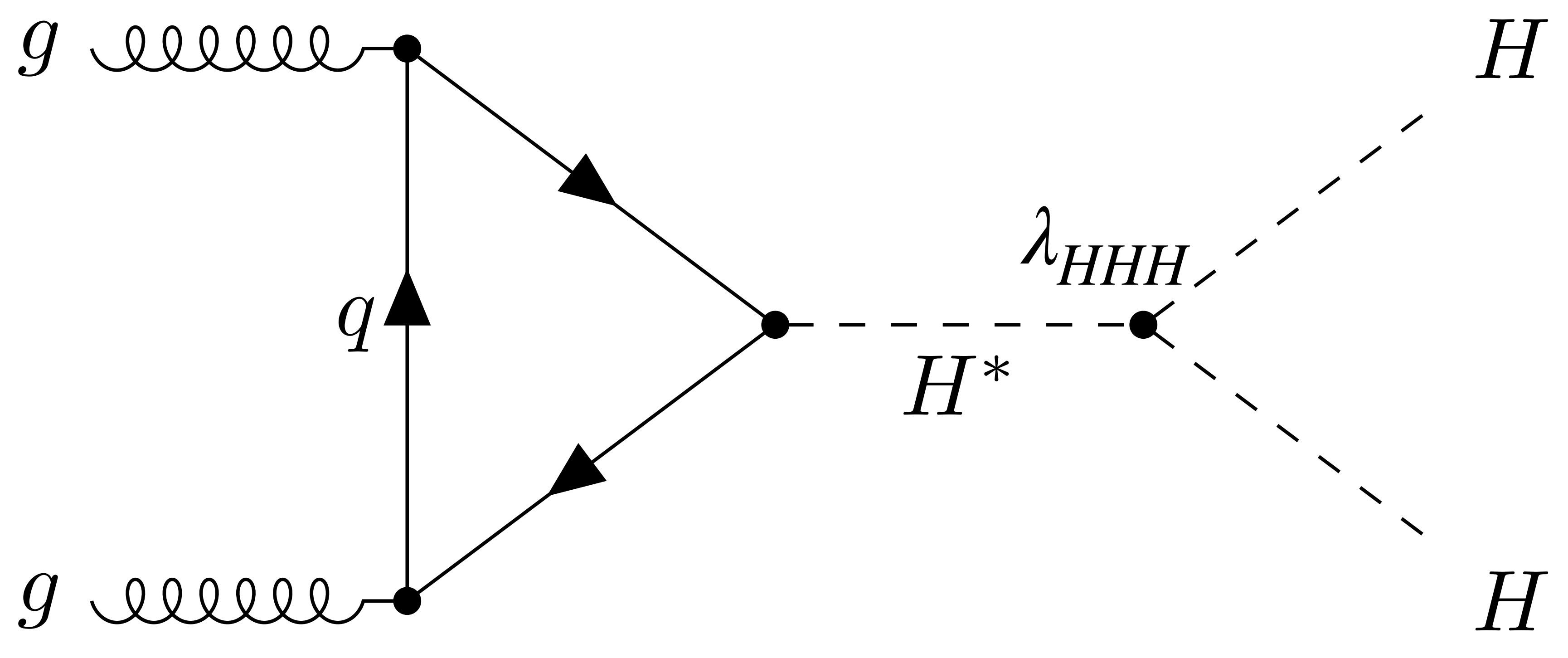
HH Signal Models and Benchmarks

- Searches are split into two modes:
  - **Non-resonant searches:**
    - Standard model HH production
  - **Resonant searches:**
    - X = Heavy scalar (S)
    - X = Spin 2 graviton (G_{KK}^*)



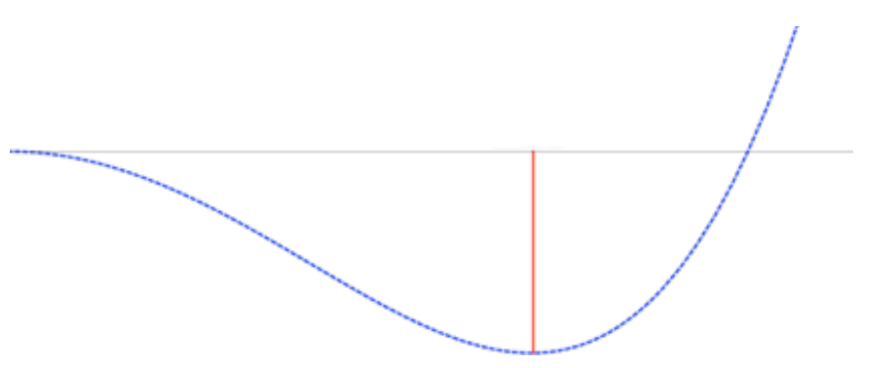
Non-resonant HH

- Contribution from exchange of a virtual (off-shell) Standard Model Higgs boson
- Why do we care?
  - Standard Model process! Allows a probe of the three Higgs coupling, λ_{HHH}
    - The relationship (on right) between λ_{HHH} and m_H, v, comes directly from the shape of the Higgs potential
    - Measurement of λ_{HHH} is the only experimental way to reconstruct the Higgs potential



Higgs Potential: $V(\phi) = -\mu^2\phi^2 + \lambda\phi^4$

Expand about minimum: $V(\phi) \rightarrow V(v+h)$



$$V = V_0 + \frac{1}{2}m_H h^2 + \frac{m_H^2}{2v}h^3 + \frac{m_H^2}{8v^2}h^4$$

$$\boxed{\lambda_{HHH}^{SM} = \frac{m_H^2}{2v}}$$

$m_H^2 = 2\lambda v^2 \approx 125$ GeV

$v = \frac{\mu}{\sqrt{\lambda}} \approx 246$ GeV

Non-resonant HH: Beyond the Standard Model

- Observing Standard Model HH production and measuring λ_{HHH} are crucial to verifying that electroweak symmetry breaking is due to a Standard Model-like Higgs sector

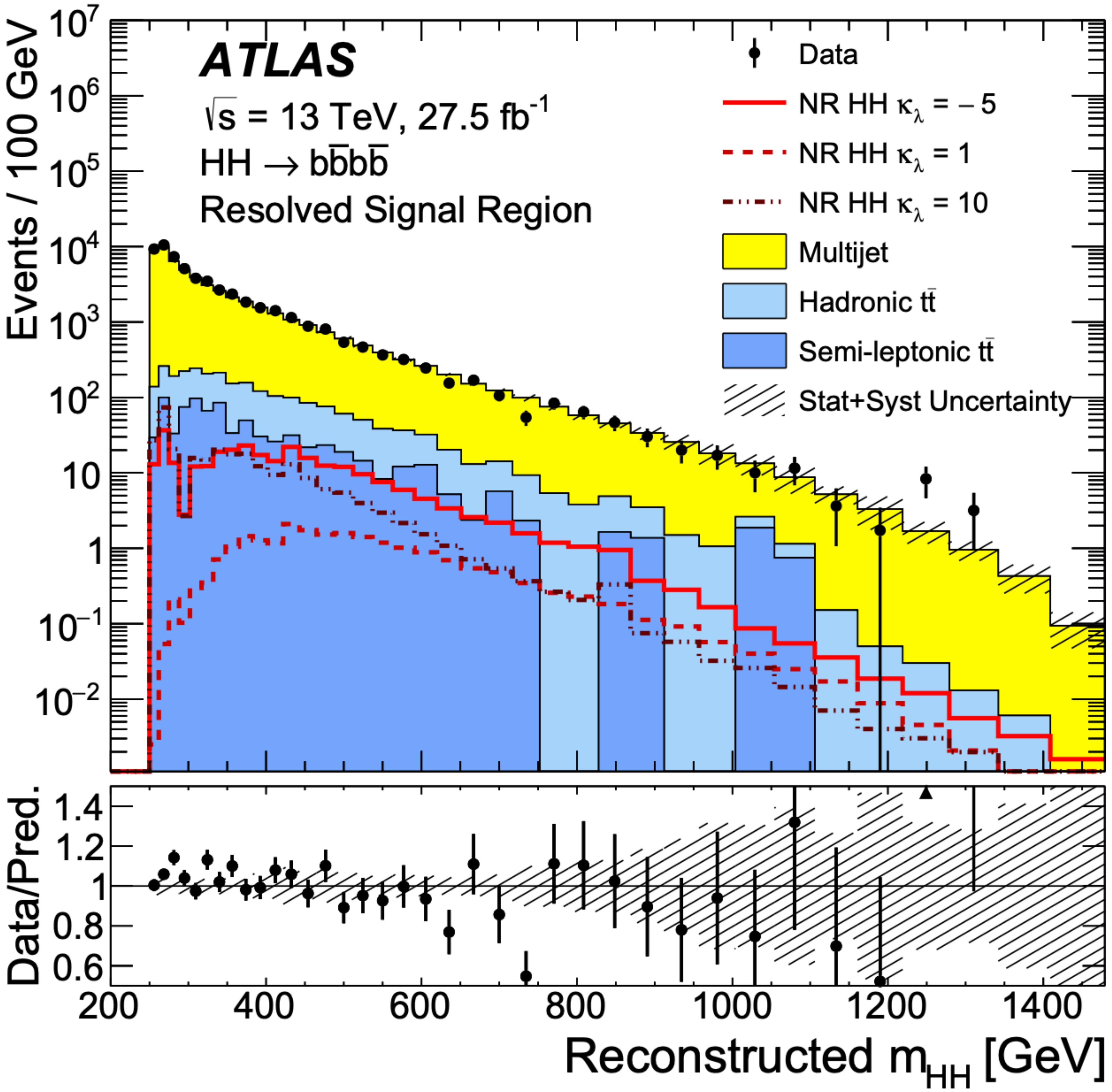
- However, there are a variety of models that predict modifications to λ_{HHH}

  - New degrees of freedom => mixing with other Higgs doublets, loop modifications, etc

- To probe this, we perform searches as a function of λ_{HHH}, usually parameterized as a function of

$$\kappa_\lambda = \frac{\lambda_{HHH}}{\lambda_{HHH}^{SM}}$$

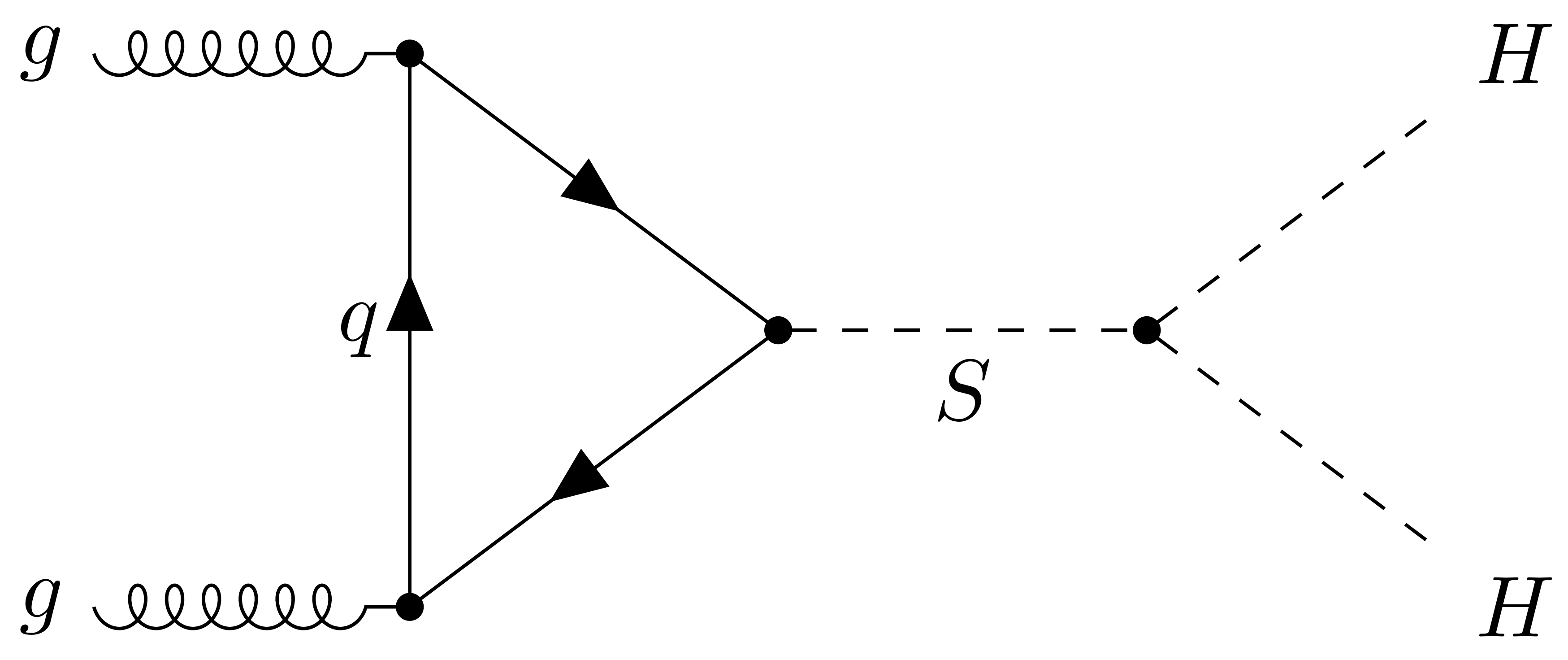
- Varying κ_λ impacts not only the cross section, but also the kinematic distributions! Both effects contribute to our final constraints on κ_λ

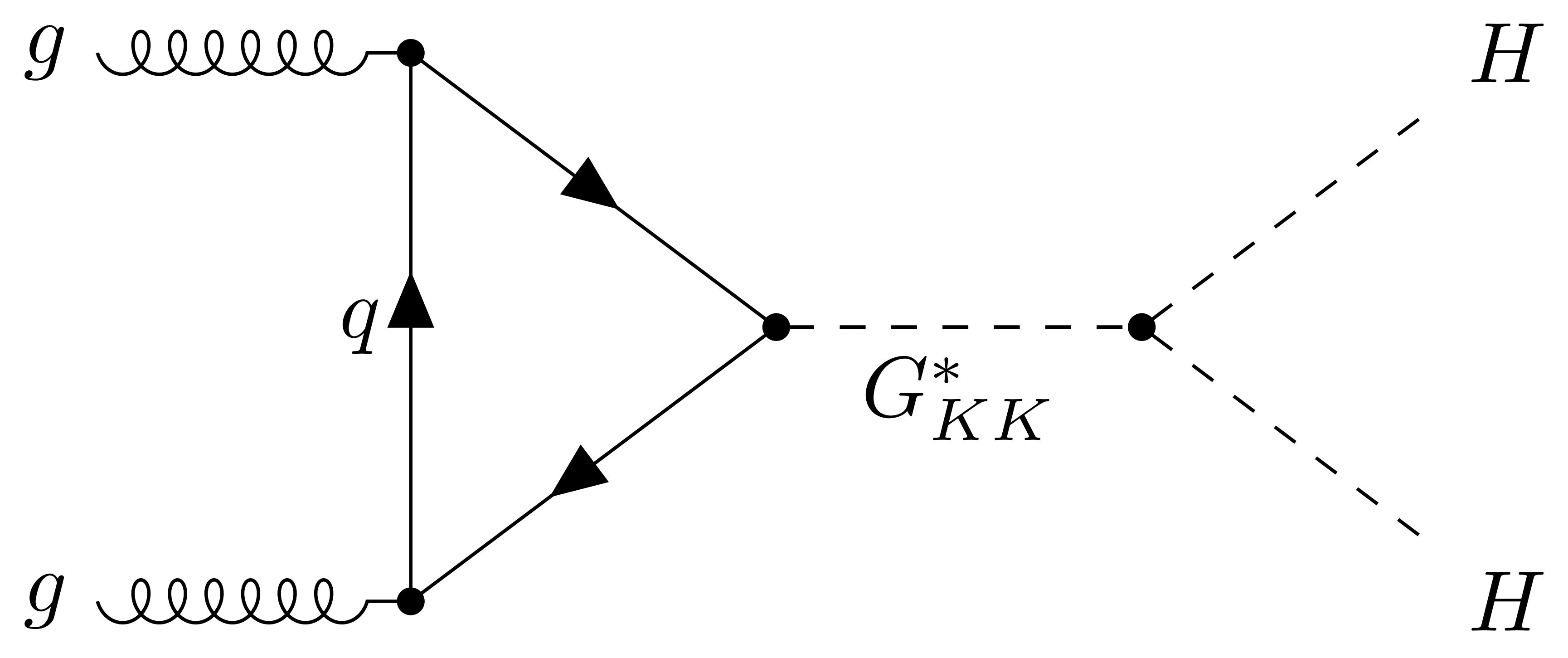


HH→4b Signal Region m_{HH}. The colored lines show the shapes of the non-resonant HH signal for κ_λ = -5, 1 (SM value), and 10 [ref]

Resonant Searches

- **Spin 0**: Generic search for a heavy scalar resonance (no specific model assumed)
  - Applicable to, e.g., 2 Higgs doublet models, where the scalar is a heavy Higgs
- **Spin 2**: Kaluza-Klein Graviton in the Randall Sundrum model
  - Graviton arising from a warped extra dimension
  - Model parameter $c = k/\overline{M}_{Pl}$, where k is the curvature of the extra dimension, $\overline{M}_{Pl}$ is the Planck mass





4b or not 4b?

- The Higgs boson decays to a variety of different particles
- The most common decay mode is to $b\bar{b}$ (58%)
- Thus, it is natural to look for HH decays involving $b\bar{b}$

H Decay Mode	Fraction
bb	58.24%
WW	21.37%
gg	8.187%
ττ	6.272%
cc	2.891%
ZZ	2.619%
γγ	0.227%
Total	99.806%

Branching Ratios for m_H = 125 GeV

4b or not 4b?

- Show on the right HH branching fractions (derived from the previous slide)
- As expected 4b has the largest branching fraction (~34% of HH decays) – great place to look!
- There are also a variety of other channels with significant contributions

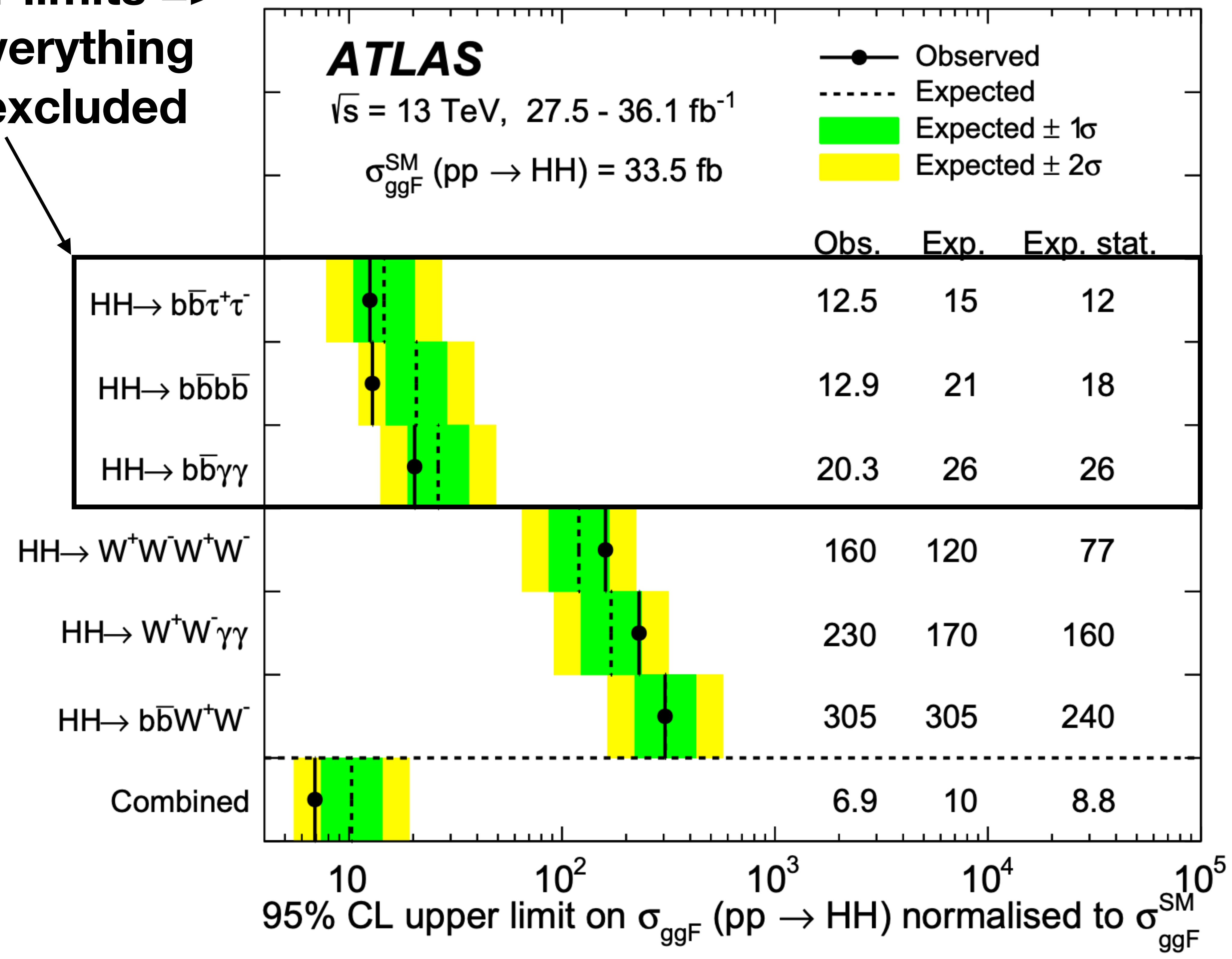
HH Branching Fractions	bb	WW	ττ	ZZ	γγ
bb	34%				
WW	25%	4.6%			
ττ	7.3%	2.7%	0.39%		
ZZ	3.1%	1.2%	0.33%	0.069%	
γγ	0.26%	0.10%	0.028%	0.012%	0.0005%

Branching Ratios

Beaten by Background

- Can often gain in other channels from lower amounts of (irreducible) background
  - LHC is a hadronic collider => lots of events from generic QCD processes
  - Hadronic HH processes are difficult to distinguish from these generic QCD processes
  - Other objects (τ, γ) help to distinguish signal from these generic events => smaller overall background

3 most sensitive channels. Upper limits => left is better, everything to the right is excluded

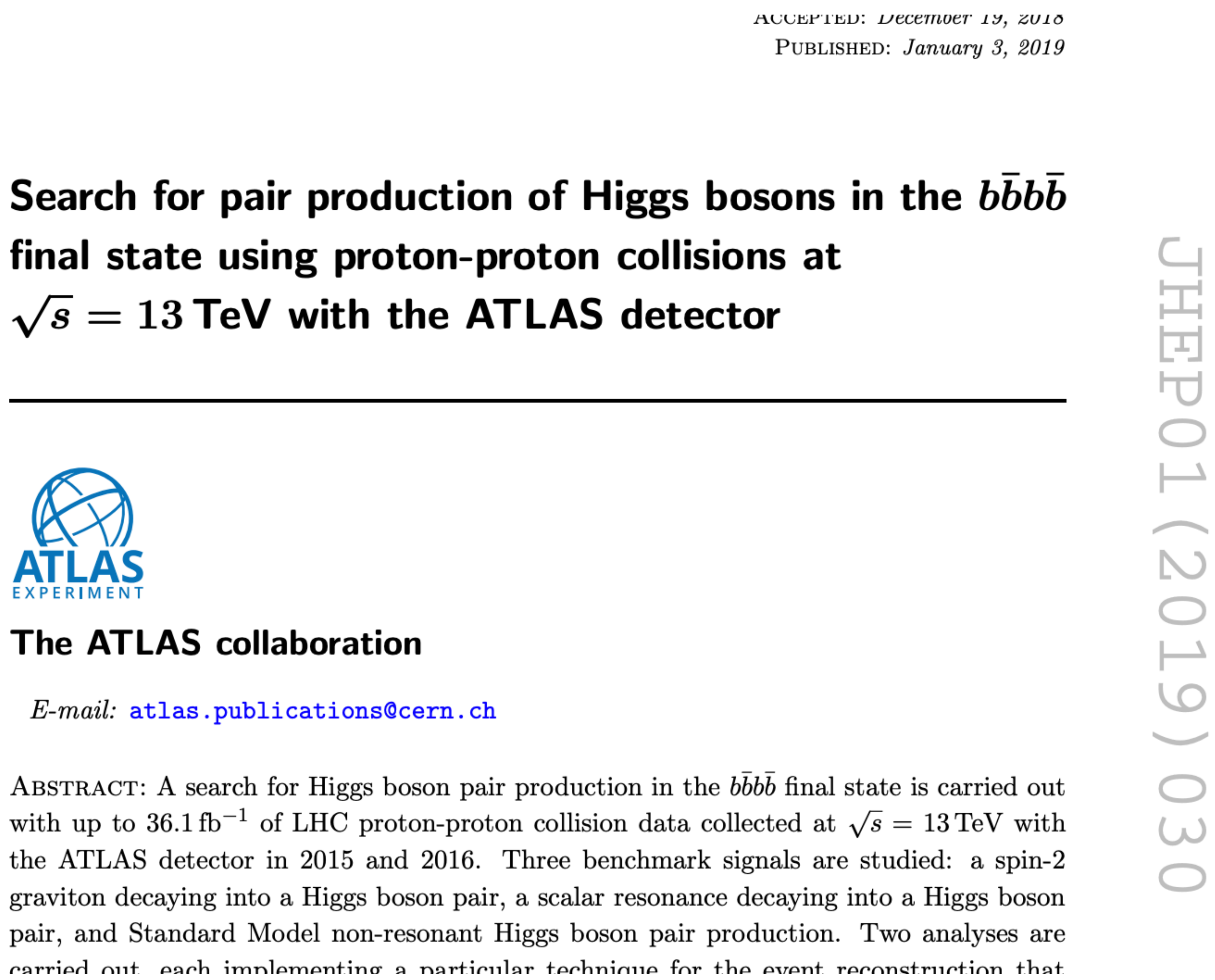


Limits on Standard Model HH production from some of the most sensitive channels [ref]

HH→4b: Analysis and Selection

Some Context

- A paper was published with the ATLAS data from 2015/2016 (36.1 fb^{-1})
- Our target is to publish on inclusive Run II data (15/16/17/18 = 139 fb^{-1})
- Important baseline steps for this:
  - Understanding/reproducing the previous analysis
  - Understanding the effect of changes to physics objects (e.g. in b-tagging)



ACCEPTED: December 19, 2018
PUBLISHED: January 3, 2019

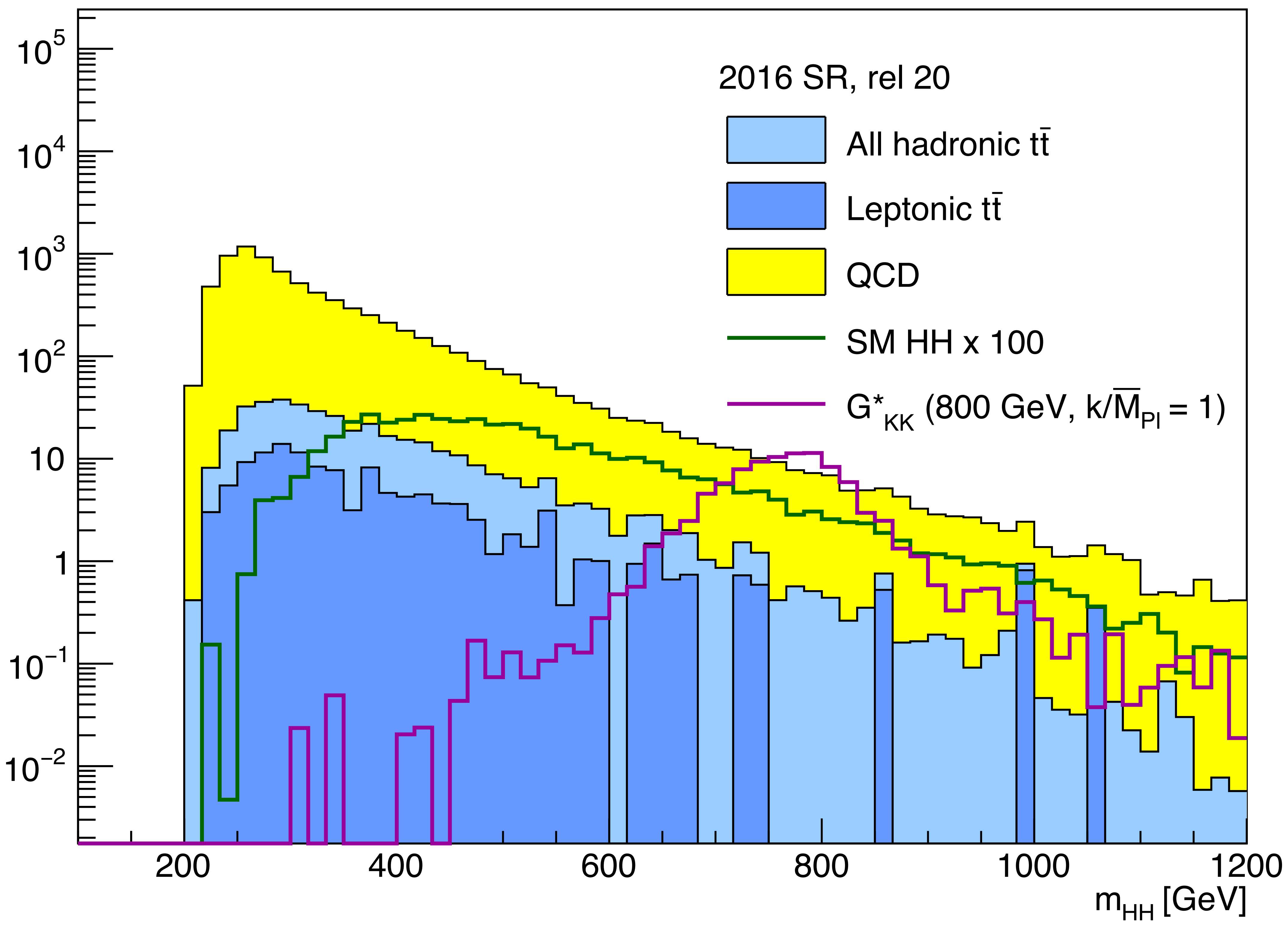
Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

ATLAS EXPERIMENT
The ATLAS collaboration
E-mail: atlas.publications@cern.ch

ABSTRACT: A search for Higgs boson pair production in the $b\bar{b}b\bar{b}$ final state is carried out with up to 36.1 fb^{-1} of LHC proton-proton collision data collected at $\sqrt{s} = 13$ TeV with the ATLAS detector in 2015 and 2016. Three benchmark signals are studied: a spin-2 graviton decaying into a Higgs boson pair, a scalar resonance decaying into a Higgs boson pair, and Standard Model non-resonant Higgs boson pair production. Two analyses are carried out, each implementing a particular technique for the event reconstruction that

JHEP01(2019)030

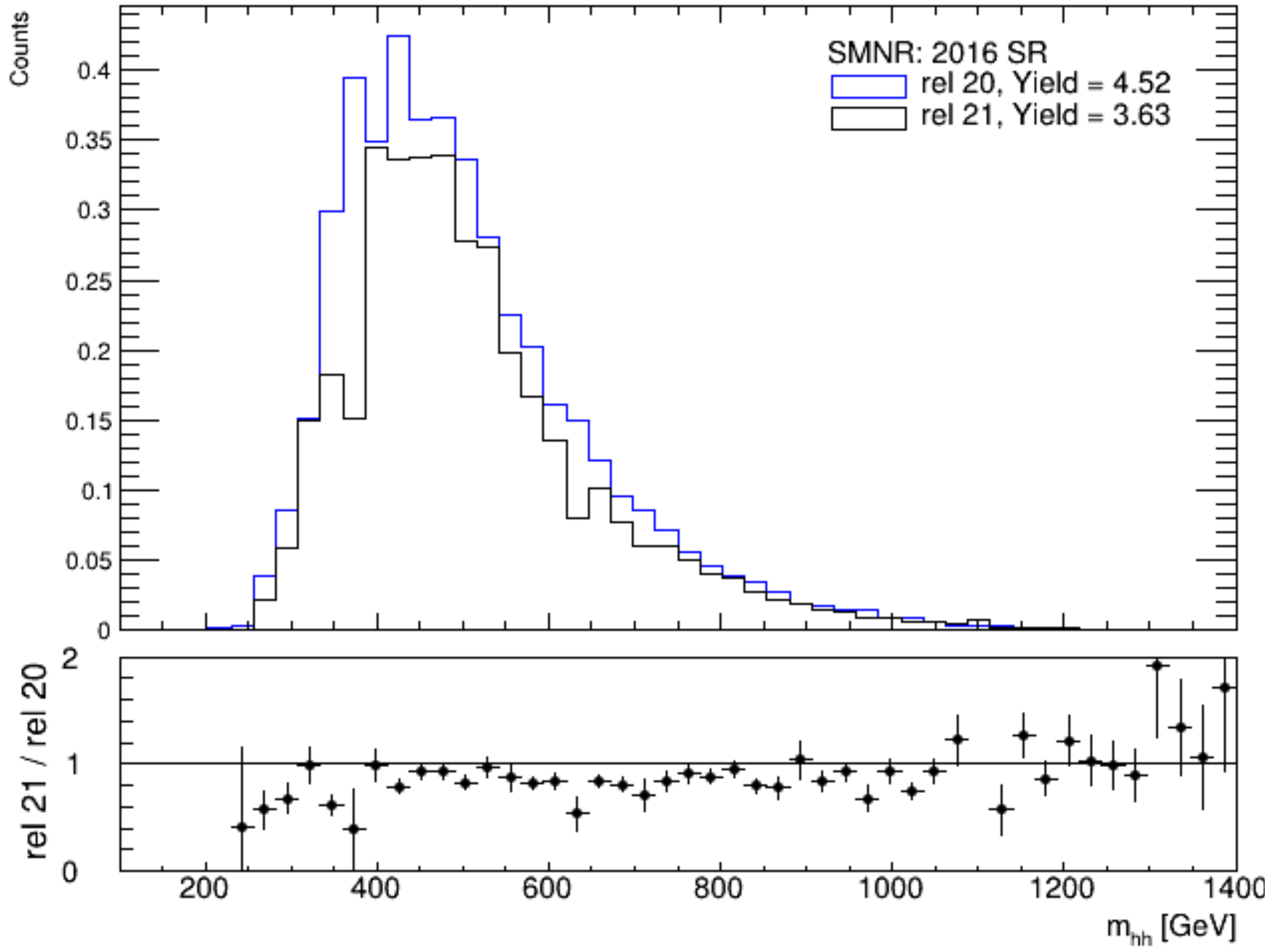
Remembering the Past



- I have reproduced (almost) entirely the baseline 15/16 analysis in my own code framework
  - Includes object selection, background estimation, statistical limit setting
  - In particular focused on the background estimation, writing a much more flexible and understandable code base
- In progress: reproducing and re-evaluating systematic uncertainties
  - Will complete this soon

Preparing for the Future

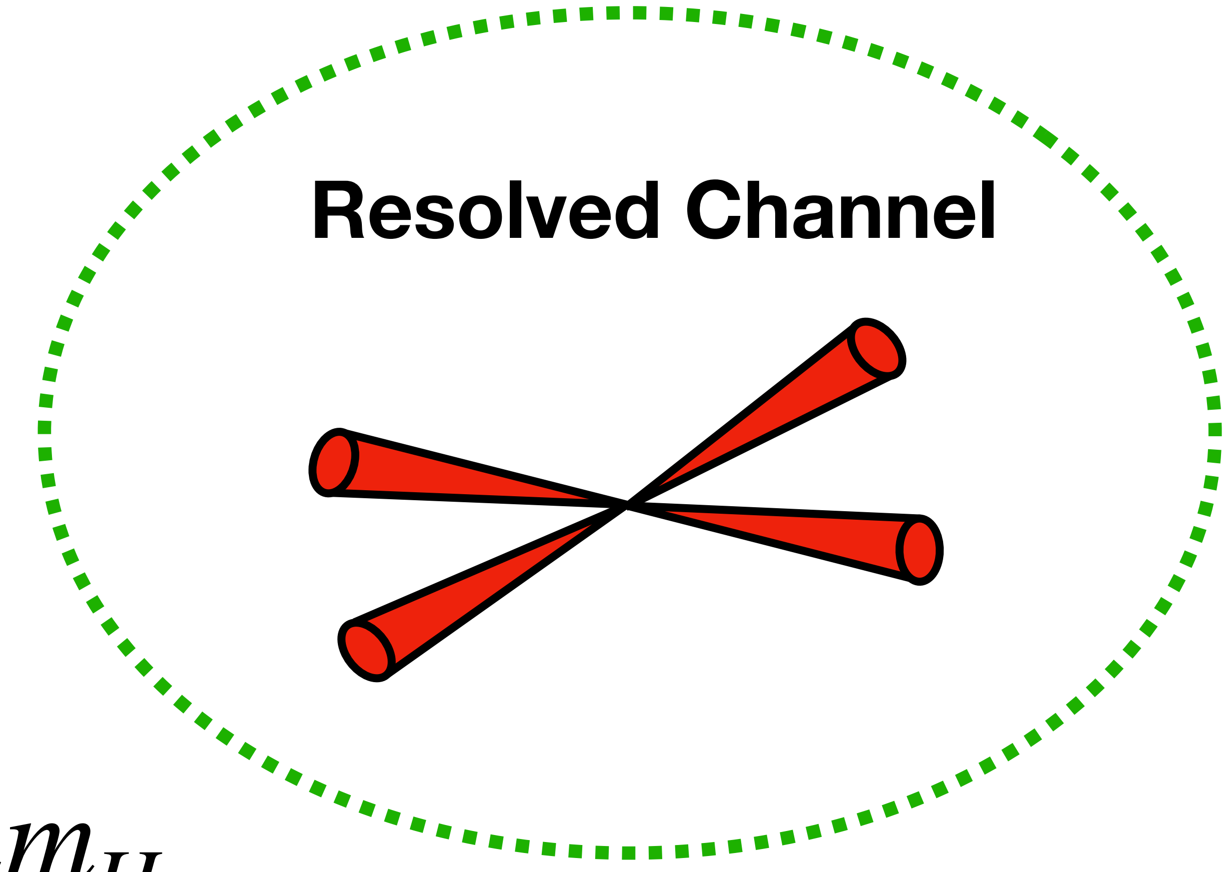
- I have also performed an extensive set of cross-checks
  - Between the new framework and the published analysis
  - Between old (rel 20) and new (rel 21) ATLAS software releases
    - Includes, e.g., changes due to b-tagging modeling improvements
      - Results in up to 25% difference in some MC yields!



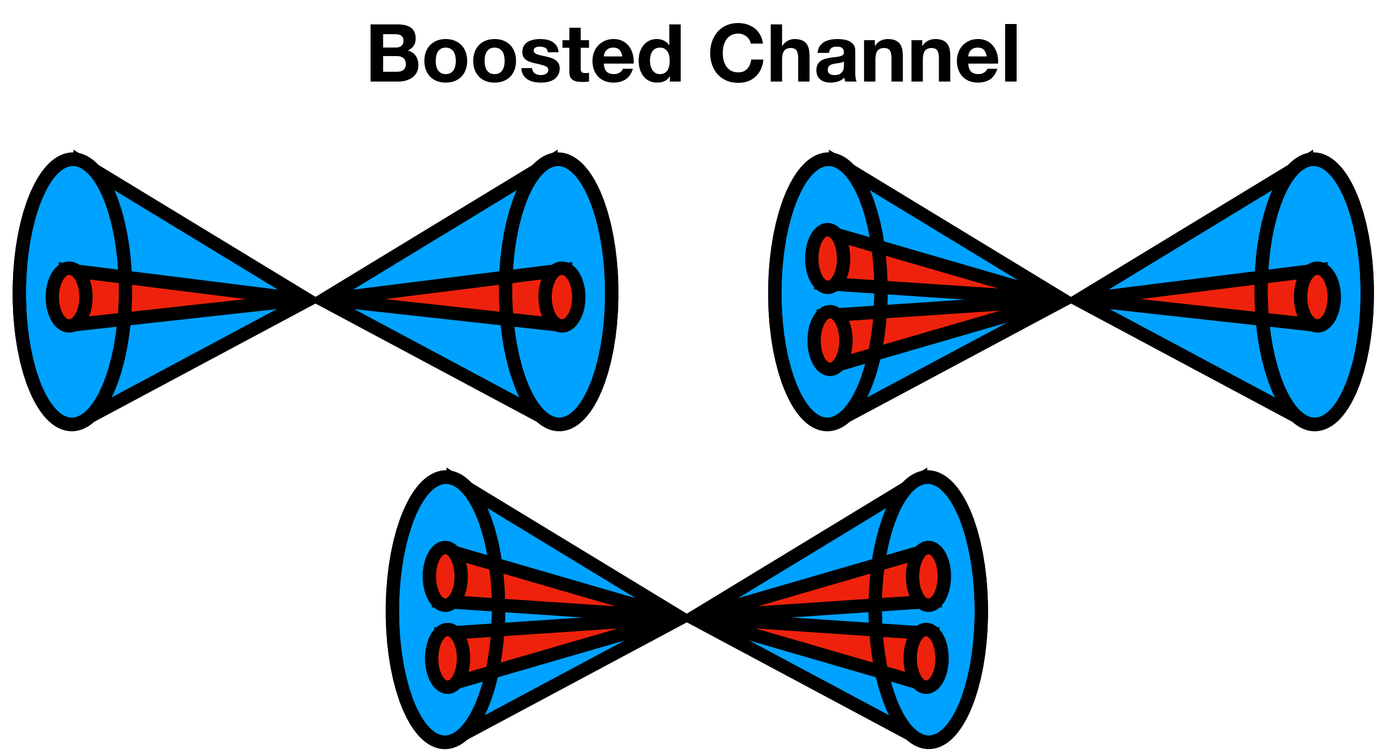
	Expected Signal Events
rel 20	4.52
rel 21	3.63
% Difference	21.84

The Analysis: Overview

- HH→4b is split into two regimes:
  - **Resolved:** All four b-quarks can be distinguished in the detector (lower mass regime)
  - **Boosted:** b-quarks are too close to distinguish - group into larger radius jets and rely on b-tagged track jets
- I am focusing on the resolved analysis
  - Greater sensitivity to Standard Model HH production



$$\Delta R_{bb,H} \approx \frac{2m_H}{p_{T,H}}$$

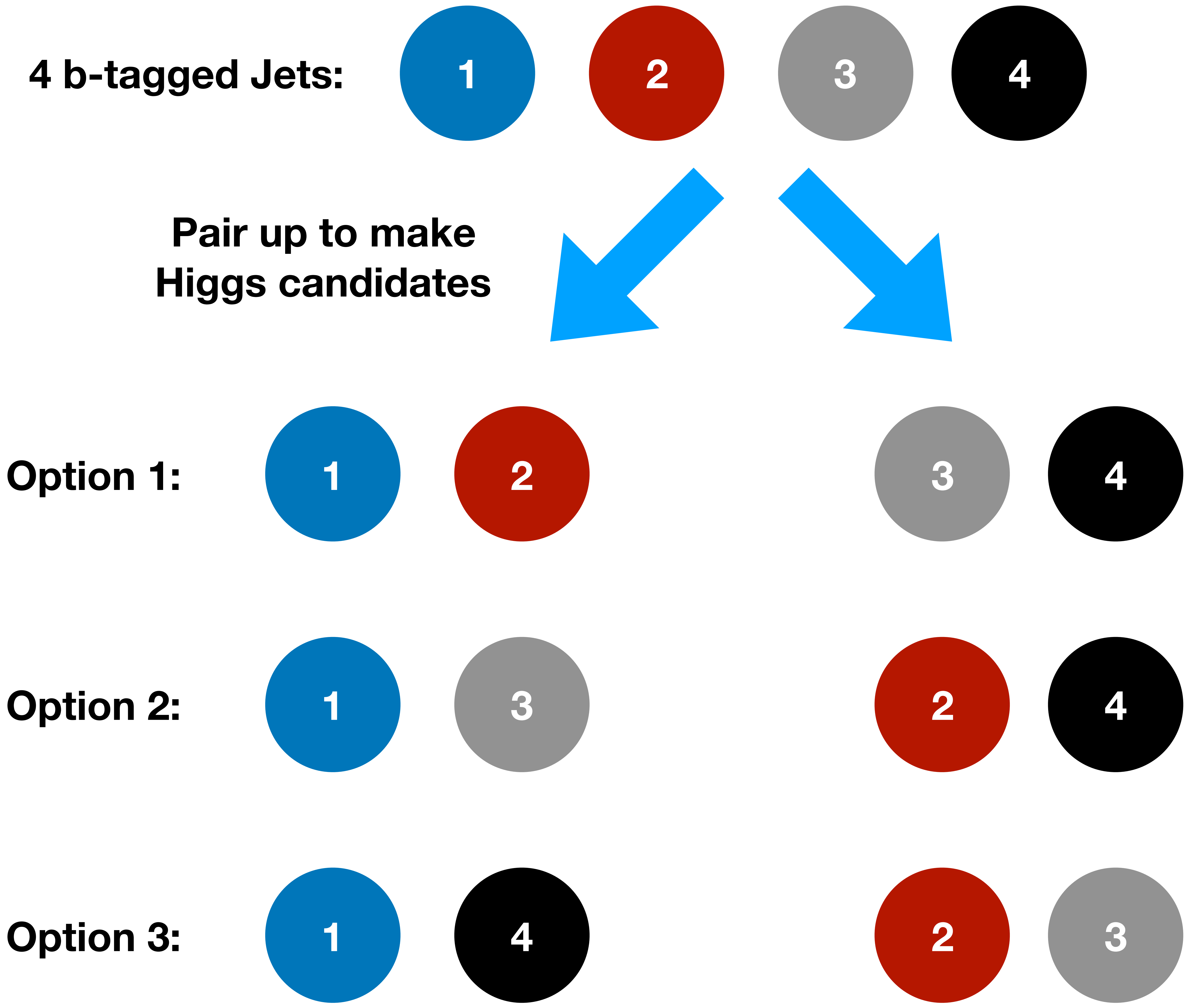


Triggers for Resolved HH→4b

- **Triggers** in ATLAS are used to decide which events to keep and which to throw away (before analysis)
    - This is a combination of online (during data taking) processing and offline (after data taking) selection
- For resolved HH→4b, we used a combination of b-jet triggers, requiring
    - 1 b-tagged jet with $p_T > 225$ GeV
    - OR 2 b-tagged jets, both with $p_T > 35$ GeV or $p_T > 55$ GeV
    - Some additional jets may be required

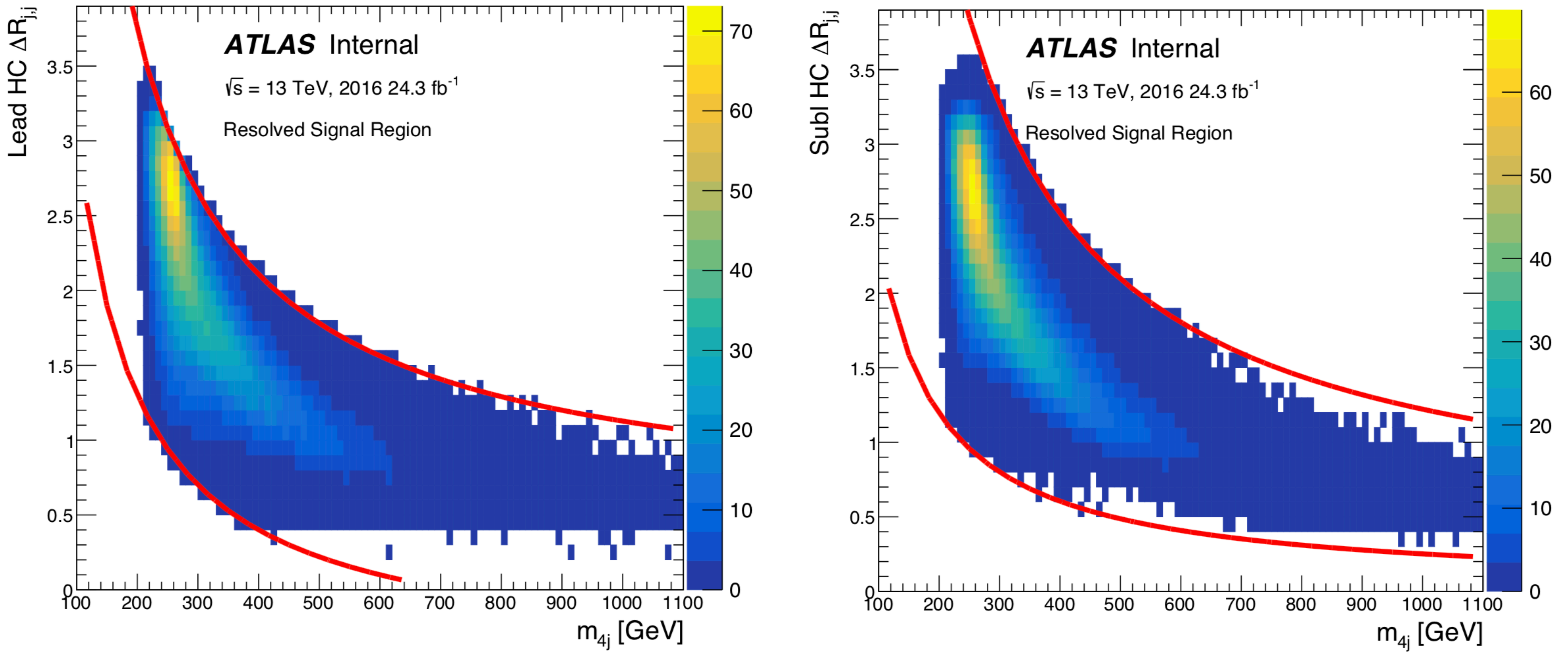
Resolved Analysis: Selection and Pairing

- Require 4 b-tagged jets (R=0.4 Anti-k_t) with $p_T > 40$ GeV, $|\eta| < 2.5$
    - b-tagging uses MV2c10 (a BDT based tagger) at the 70% working point
        - This means that there is a 70% chance that a b-jet will be tagged
- Need to then pair them up into Higgs candidates. 3 possibilities for this.
- However, not all pairs are consistent with expected kinematics for HH→4b



Resolved Analysis: Pairing Kinematics

- We expect the b-jets from each Higgs to get closer together at high Lorentz boosts
- Idea: make selections on the ΔR between our b-jets as a function of the 4 jet mass, requiring consistency with $HH \to 4b$
- Corresponding selection shown on right - pairings which pass this selection are considered "valid"



2016 Multijet after full selection (15/16 analysis). Red lines show bounds on ΔR_{jj}

m_{4j} < 1250 GeV

$$\frac{360\ \text{GeV}}{m_{4j}} - 0.5 < \Delta R_{jj,\text{lead}} < \frac{653\ \text{GeV}}{m_{4j}} + 0.475$$

$$\frac{235\ \text{GeV}}{m_{4j}} < \Delta R_{jj,\text{subl}} < \frac{875\ \text{GeV}}{m_{4j}} + 0.35$$

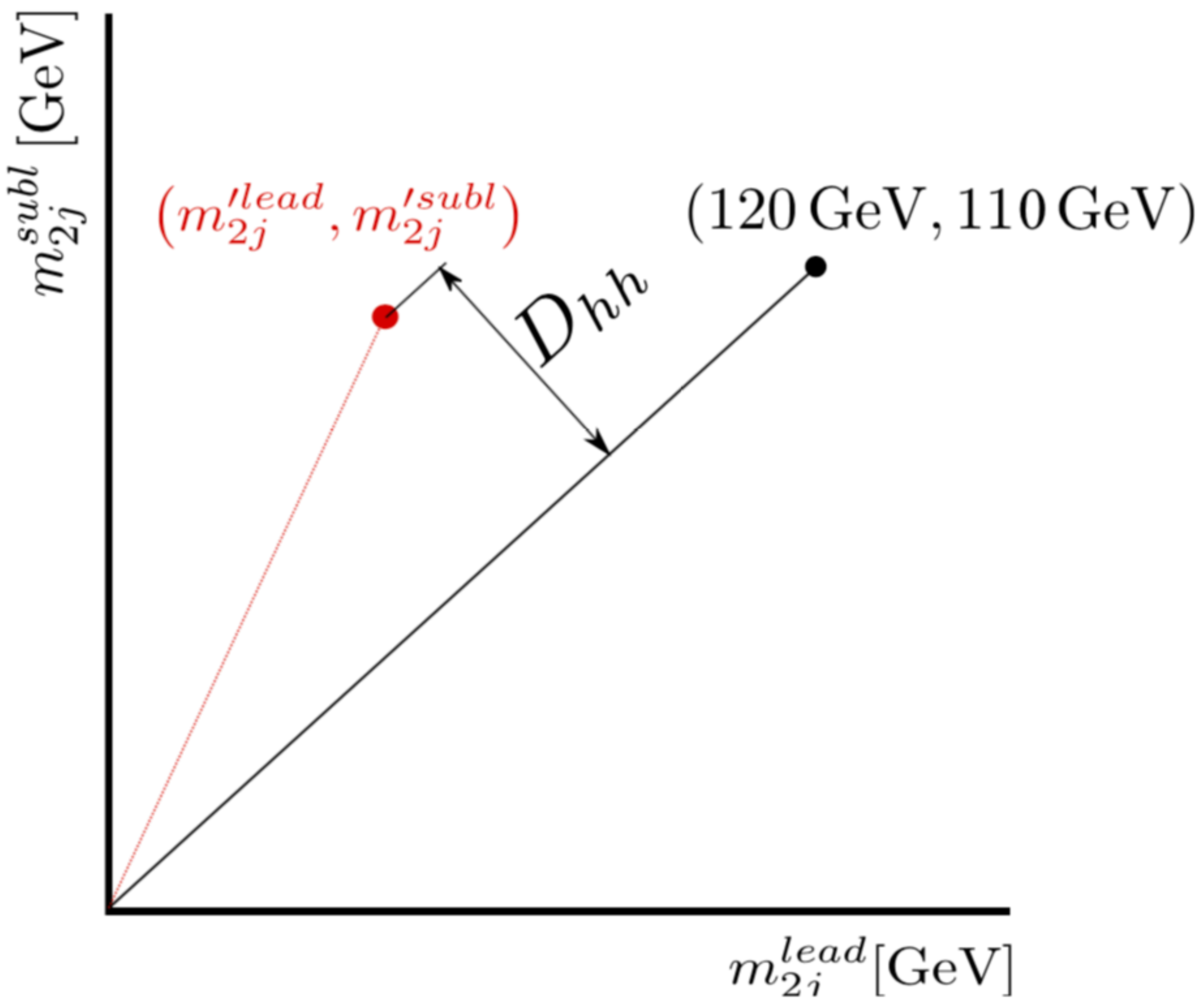
m_{4j} > 1250 GeV

$$0 < \Delta R_{jj,\text{lead}} < 1$$

$$0 < \Delta R_{jj,\text{subl}} < 1$$

Resolved Analysis: Pairing Ambiguity

$$D_{HH} = \frac{\left| m_{2j}^{\text{lead}} - \frac{120}{110} m_{2j}^{\text{subl}} \right|}{\sqrt{1 + \left(\frac{120}{110}\right)^2}}$$



- If we only have one valid pairing, we're good!
- If more than one, still need to choose!
- Current method: D_{HH} minimization
    - Pick pairing with smallest distance from the line shown on left
- Works well! Often > 90% correct

Resolved Analysis: Kinematic Selections

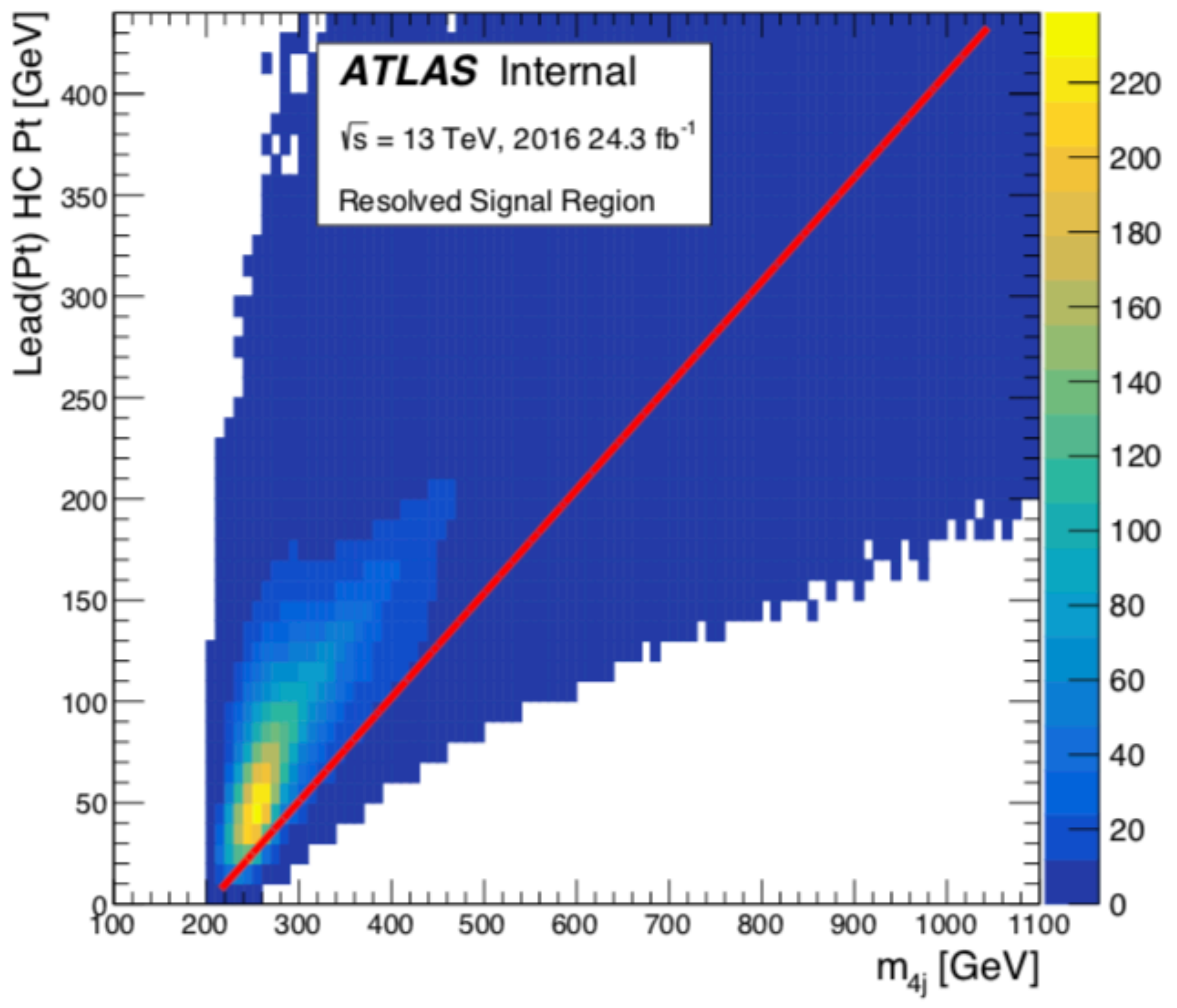
- Mass dependent p_T cuts on the Higgs candidates:

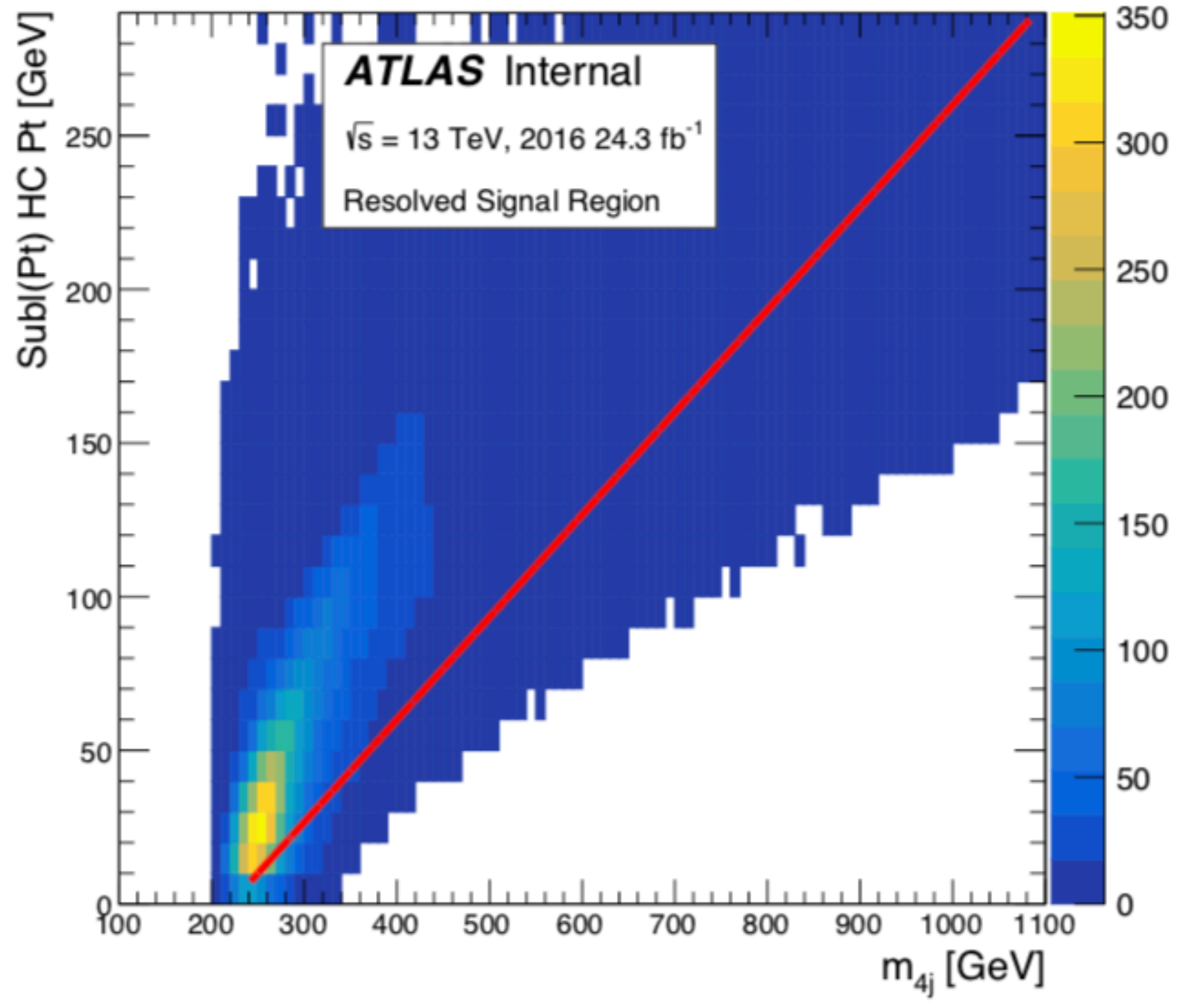
$$p_T^{lead} > 0.5 m_{4j} - 103 \text{ GeV}$$
$$p_T^{subl} > 0.33 m_{4j} - 73 \text{ GeV}$$

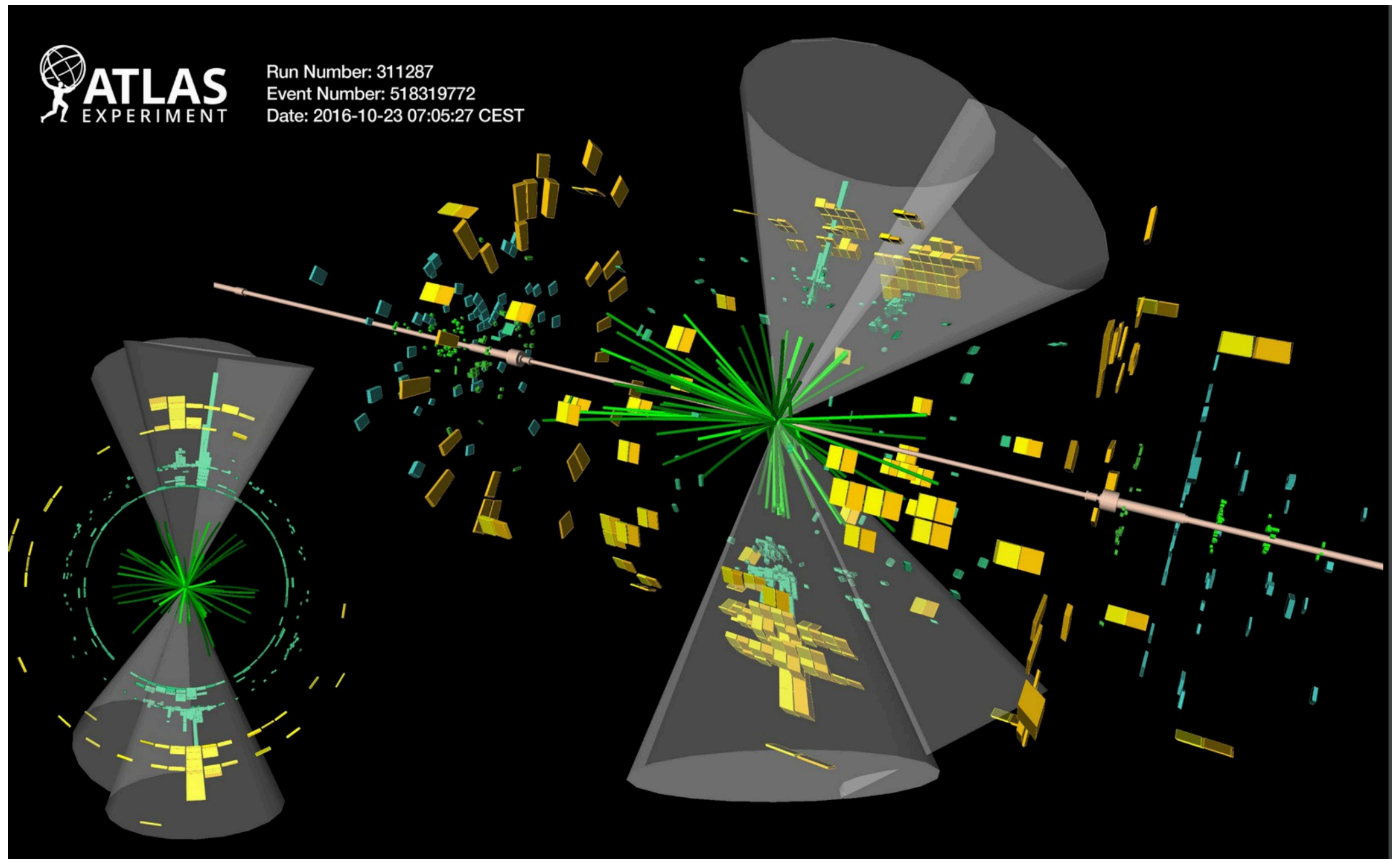
  - Consistency with X→HH

- $|\Delta\eta_{HH}| < 1.5$

  - Somewhat back to back. Rejects QCD multijet background







Resolved Analysis: Top Veto

- Reject events consistent with (hadronic) top quark decay
- Build top candidates using three jets (with at least one jet from a Higgs candidate)
- Order by b-tagging score: highest is the b quark, other two make the W candidate
- Make all possible candidates. Choose the one that minimizes X_{Wt} (consistency with W, top mass)
- If minimal $X_{Wt} < 1.5$, reject the event
- Reduces hadronic $t\bar{t}$ by ~60%, semi-leptonic by ~45%

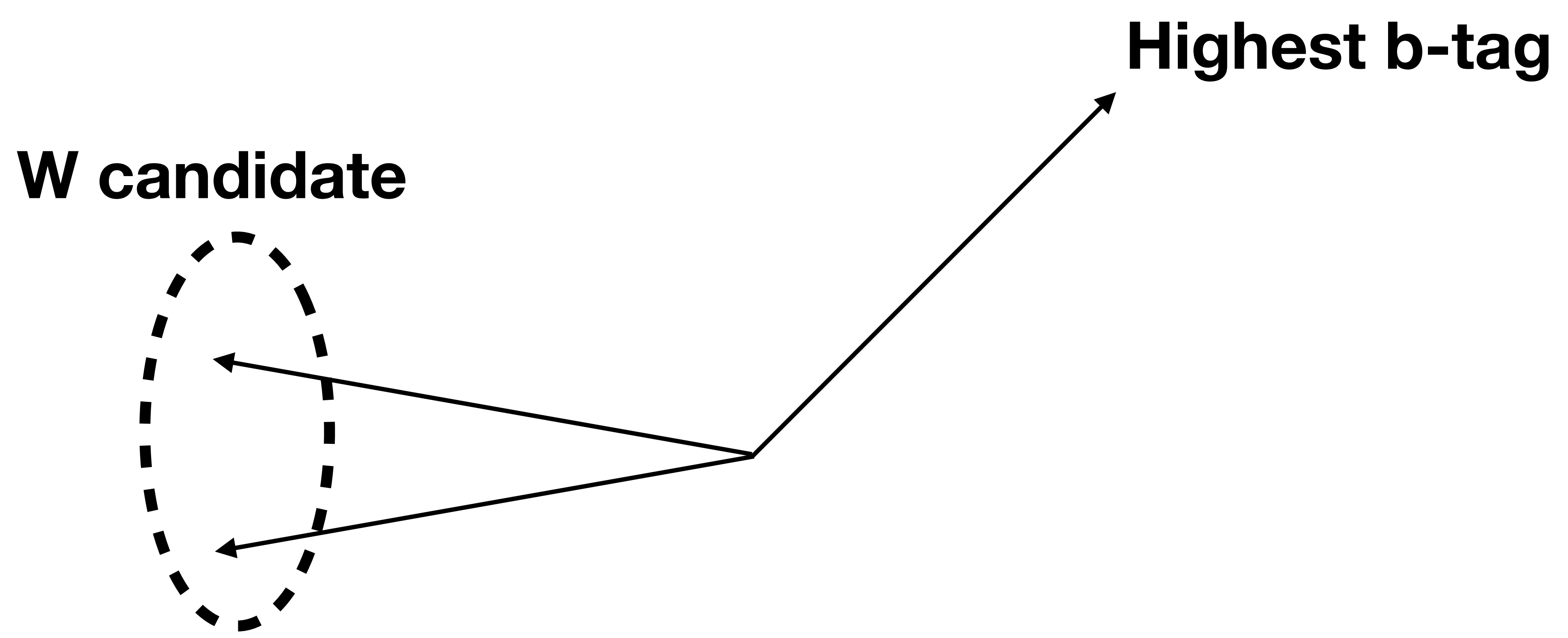


Illustration of top candidate building

$$X_{Wt} = \sqrt{\left(\frac{m_W - 80\ \text{GeV}}{0.1 m_W}\right)^2 + \left(\frac{m_t - 173\ \text{GeV}}{0.1 m_t}\right)^2}$$

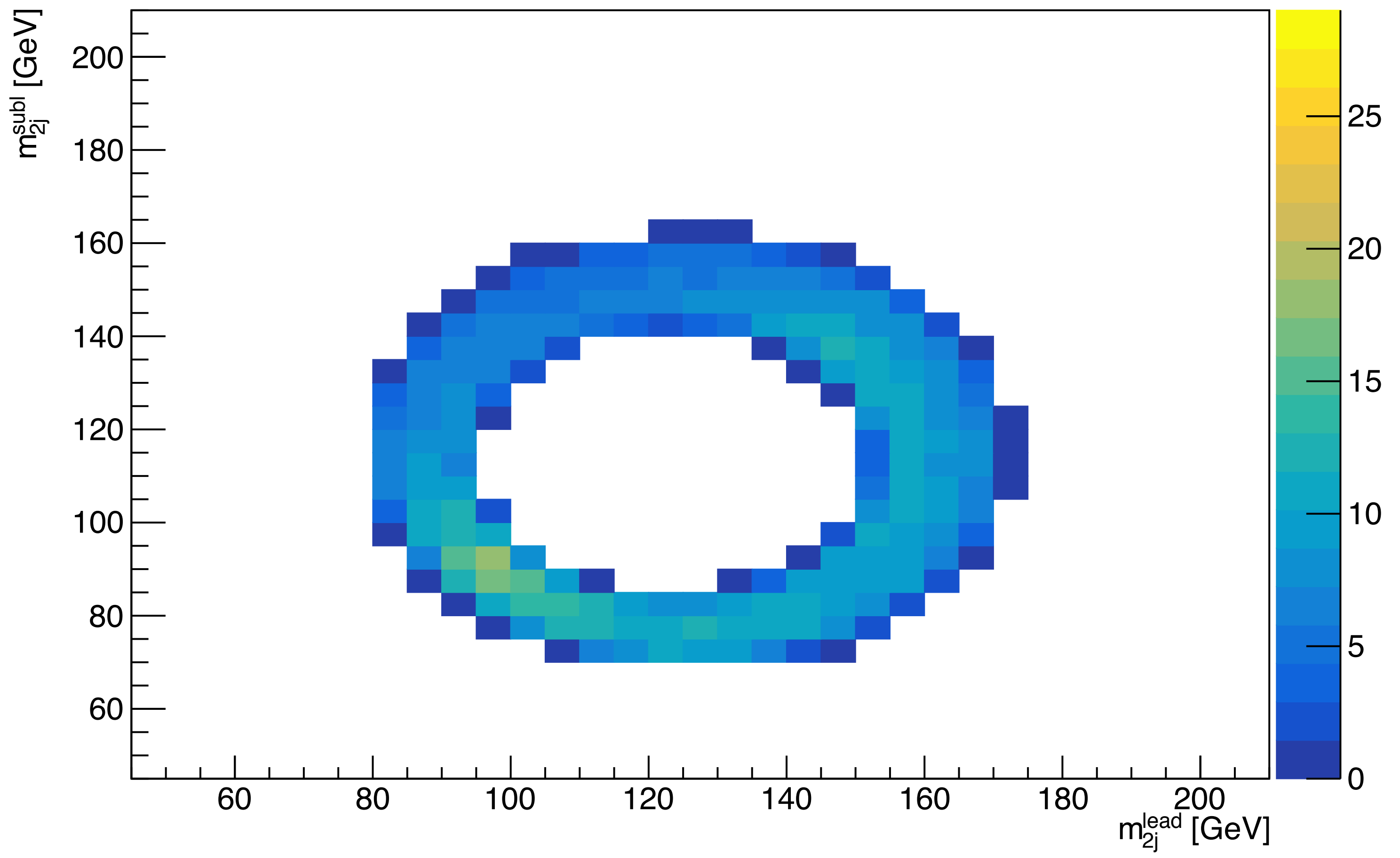
Resolved Analysis: Mass Regions

$$\sqrt{(m_{2j}^{\text{lead}} - (120 \times 1.05)\text{GeV})^2 + (m_{2j}^{\text{subl}} - (110 \times 1.05)\text{GeV})^2} < 45 \text{ GeV}$$

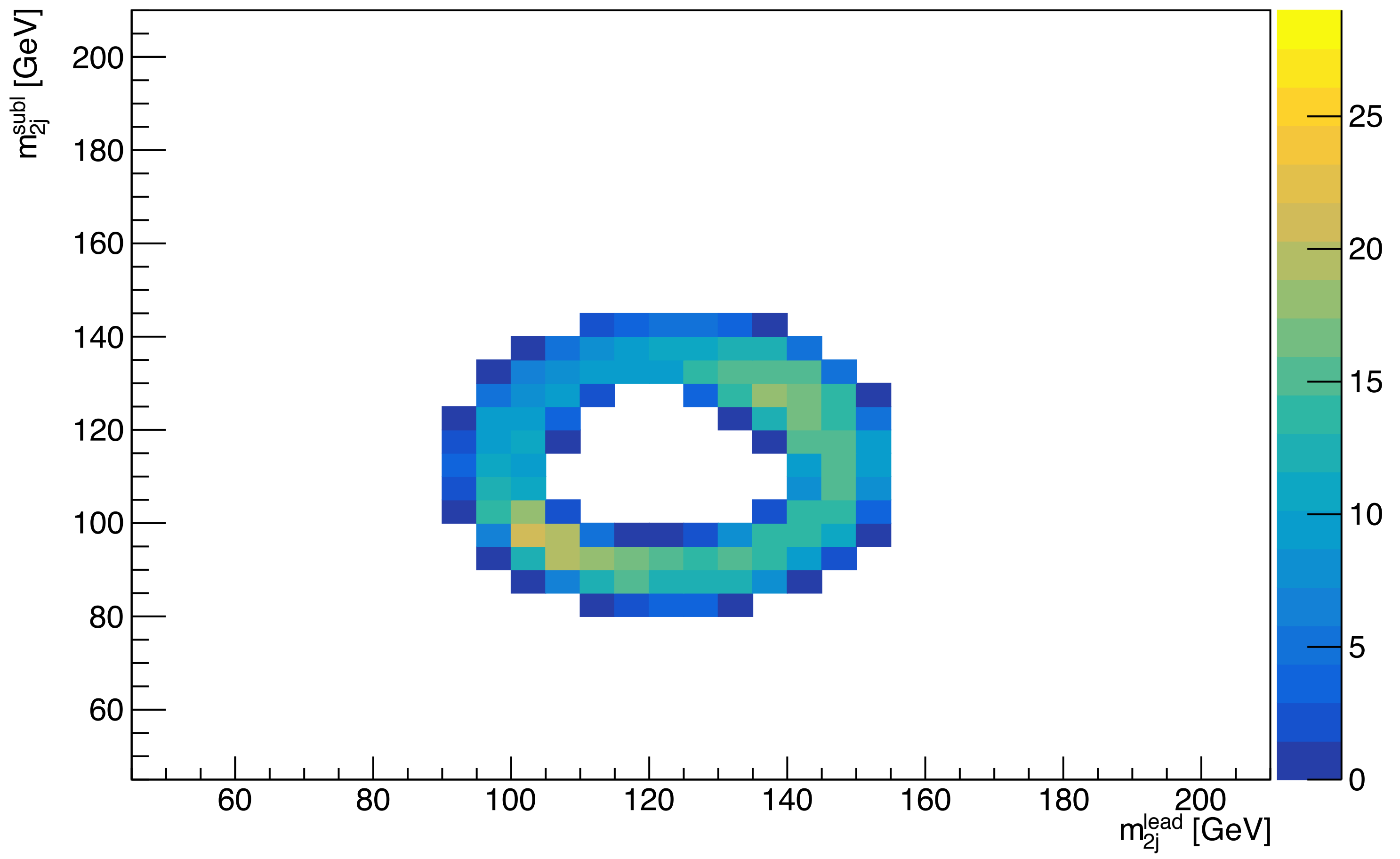
$$\sqrt{(m_{2j}^{\text{lead}} - (120 \times 1.03)\text{GeV})^2 + (m_{2j}^{\text{subl}} - (110 \times 1.03)\text{GeV})^2} < 30 \text{ GeV}$$

$$X_{HH} = \sqrt{\left(\frac{m_{2j}^{\text{lead}} - 120 \text{ GeV}}{0.1 m_{2j}^{\text{lead}}}\right)^2 + \left(\frac{m_{2j}^{\text{subl}} - 110 \text{ GeV}}{0.1 m_{2j}^{\text{subl}}}\right)^2} < 1.6$$

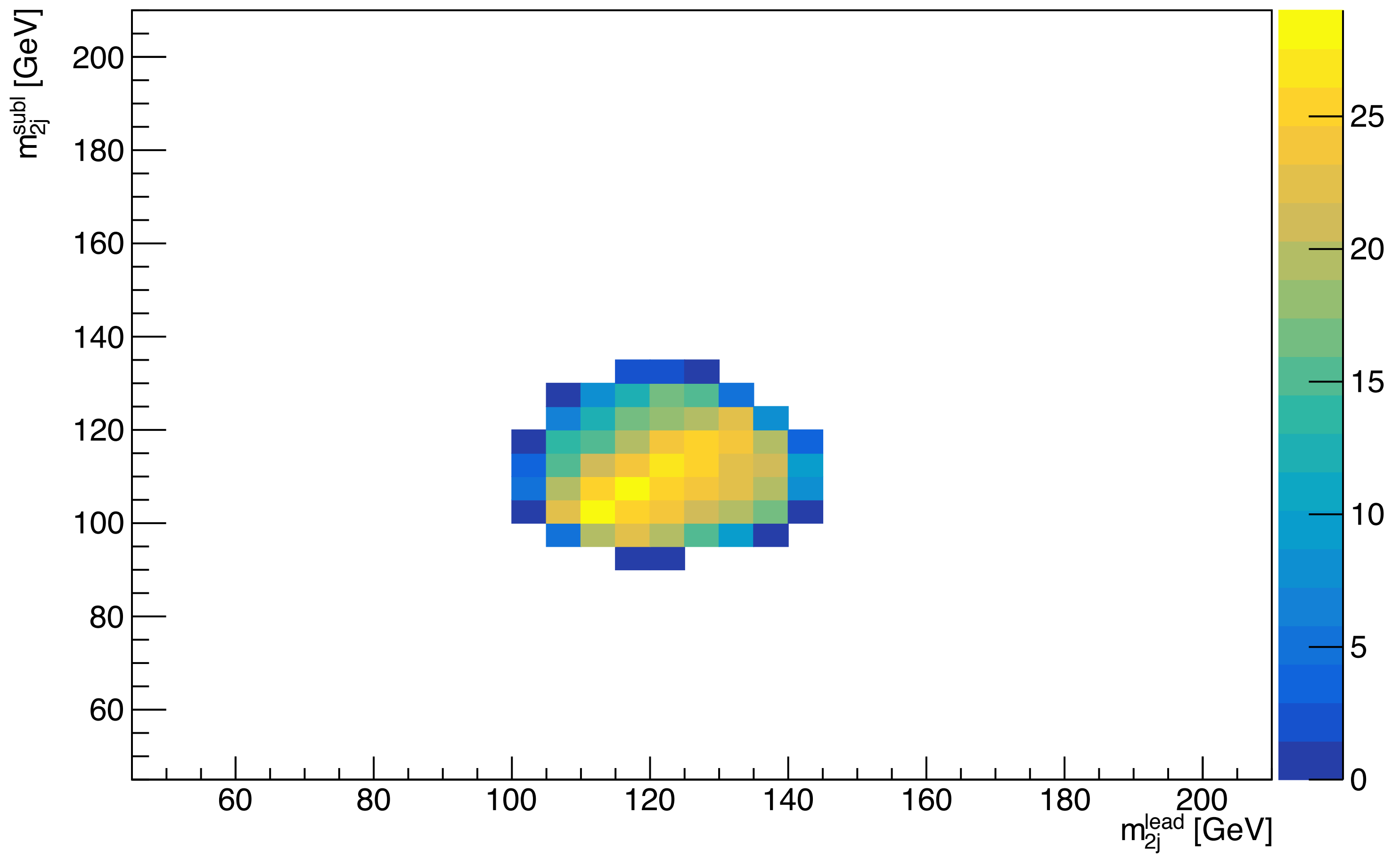
Sideband

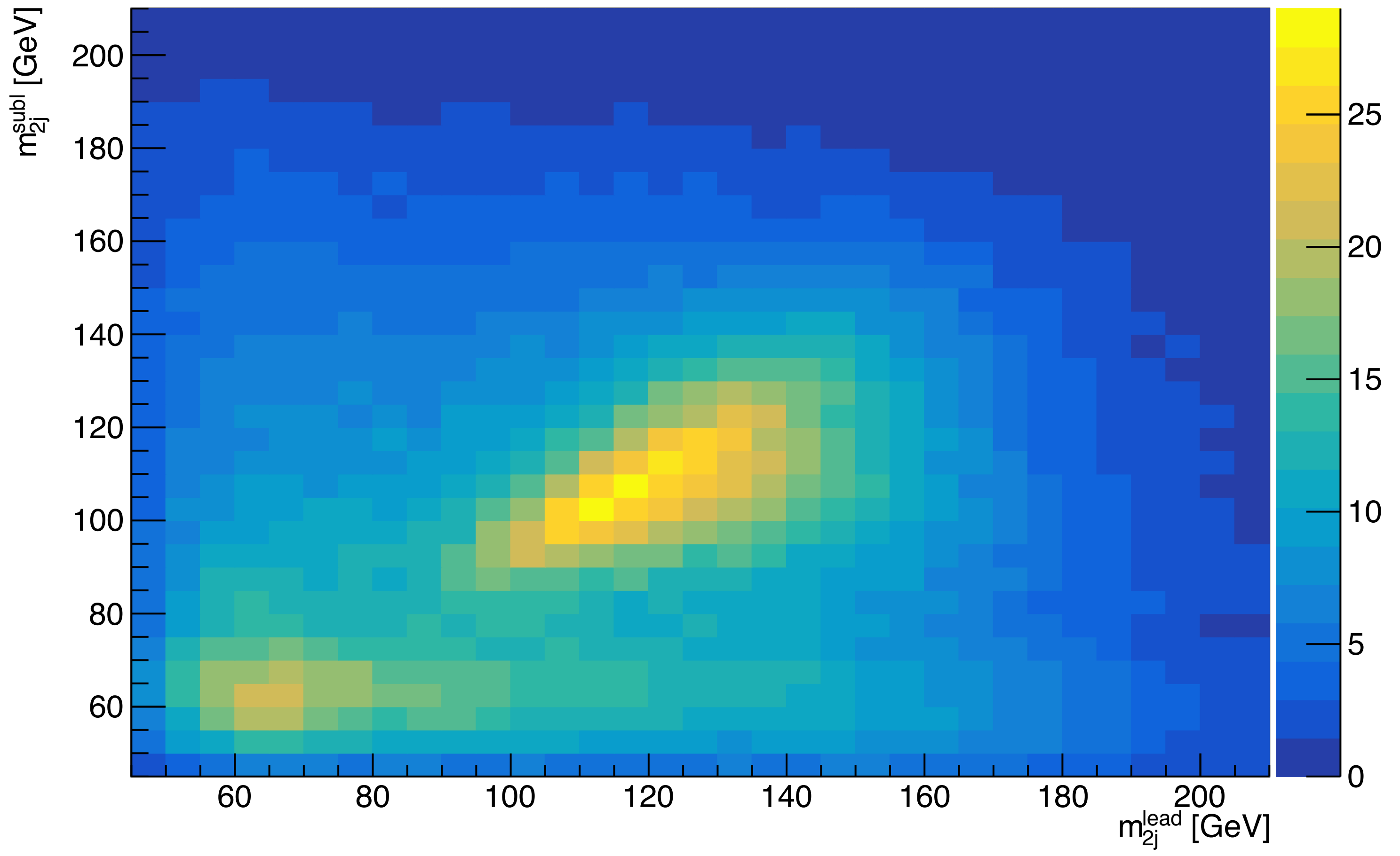


Control



Signal





- Our signal region is then defined by a requirement that the Higgs candidate masses are close to the Higgs mass (X_{HH} above)
- We further define a sideband and control region in this Higgs candidate mass plane - assumed to be similar to, but still orthogonal from, our signal region

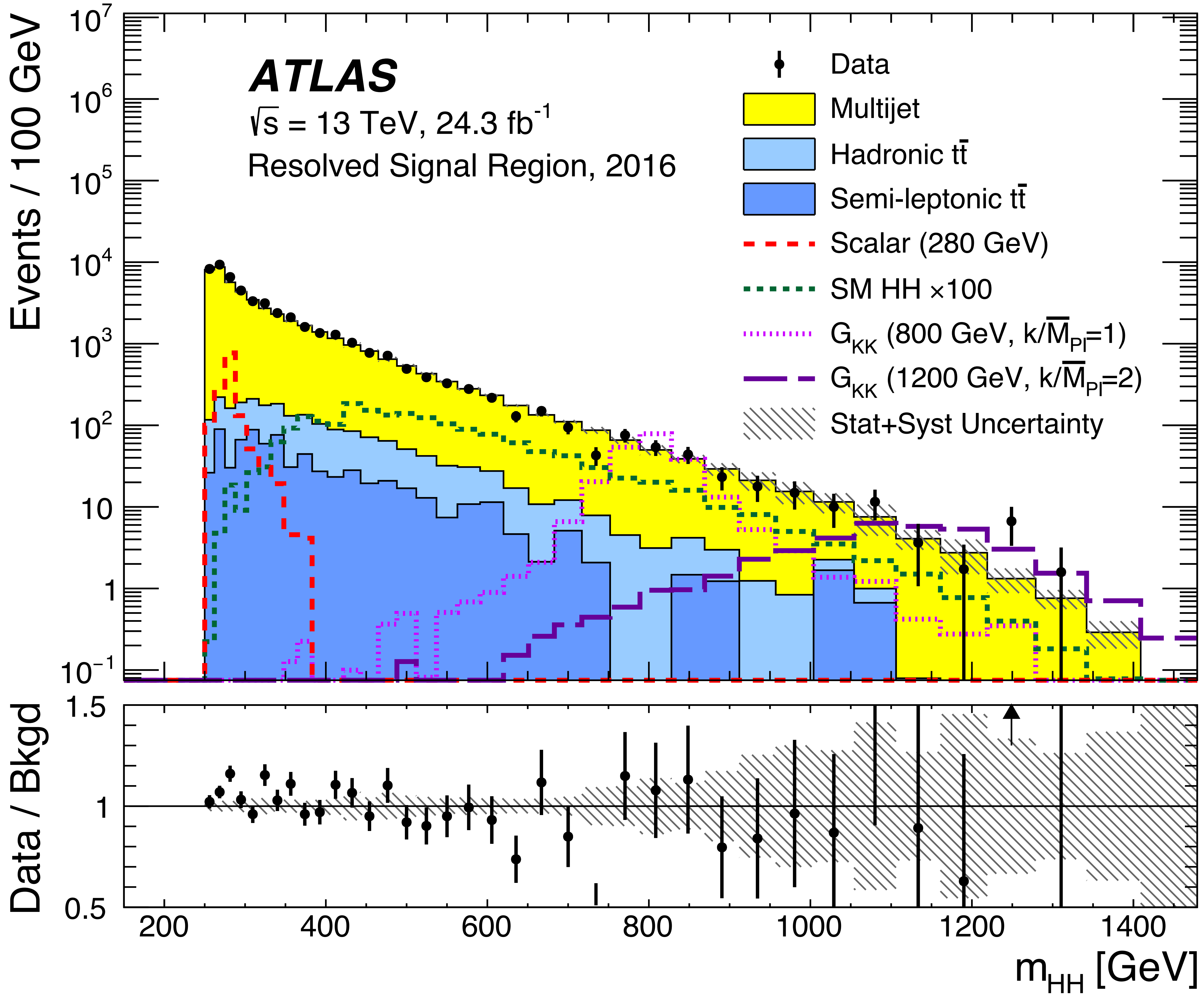
HH→4b: Background Estimation

The Analysis: Background Processes

- We have a set of criteria chosen to select only those events consistent with our signal model

- However, there are some background processes which result in final states that are indistinguishable from our signal

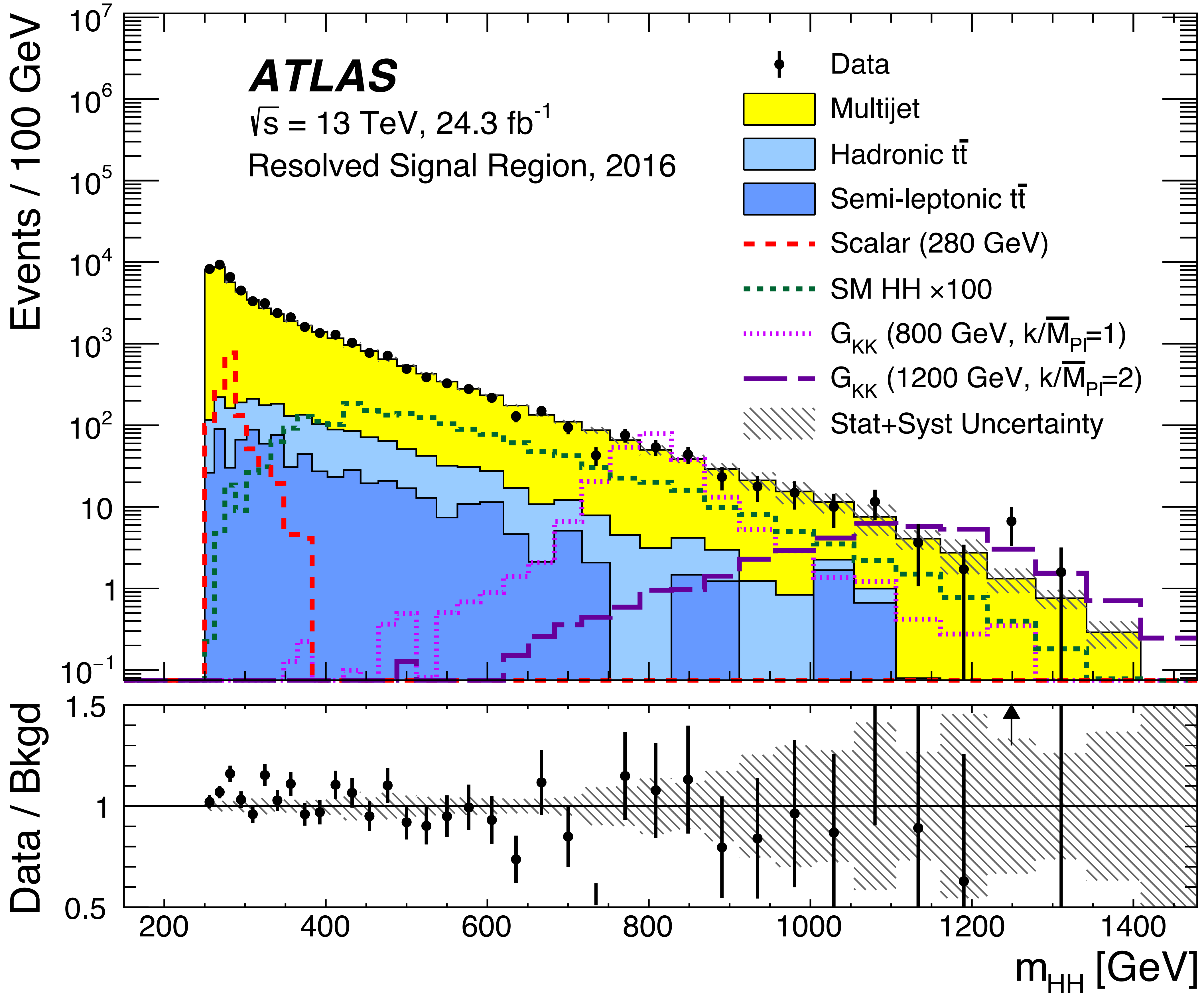
- An important part of our analysis is understanding these backgrounds well!



m_{HH} spectrum from 2016 HH→4b analysis. $t\bar{t}$ (blue) and QCD multijet (yellow) are the two dominant backgrounds (QCD ~95%)

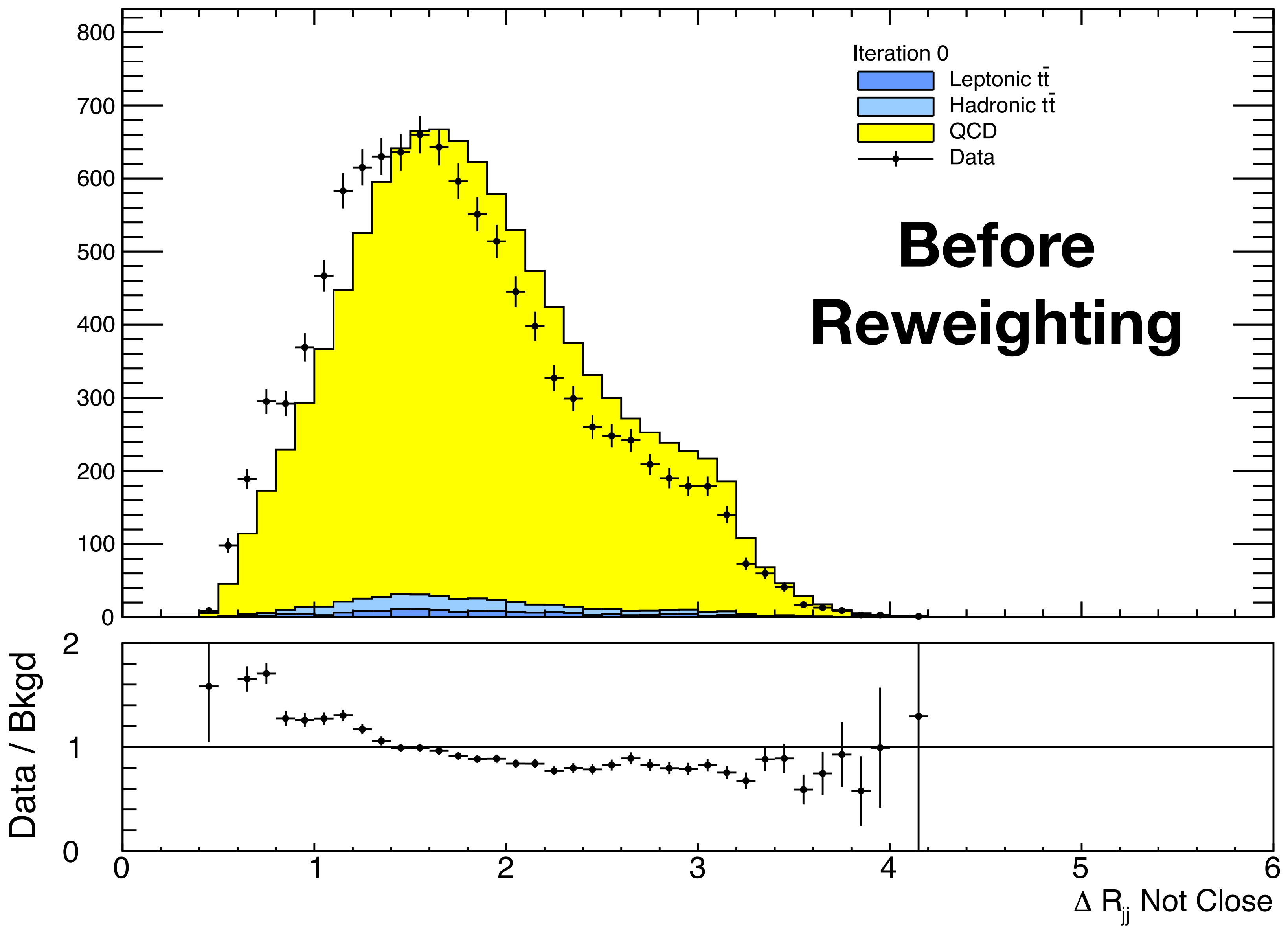
The Analysis: Background Processes

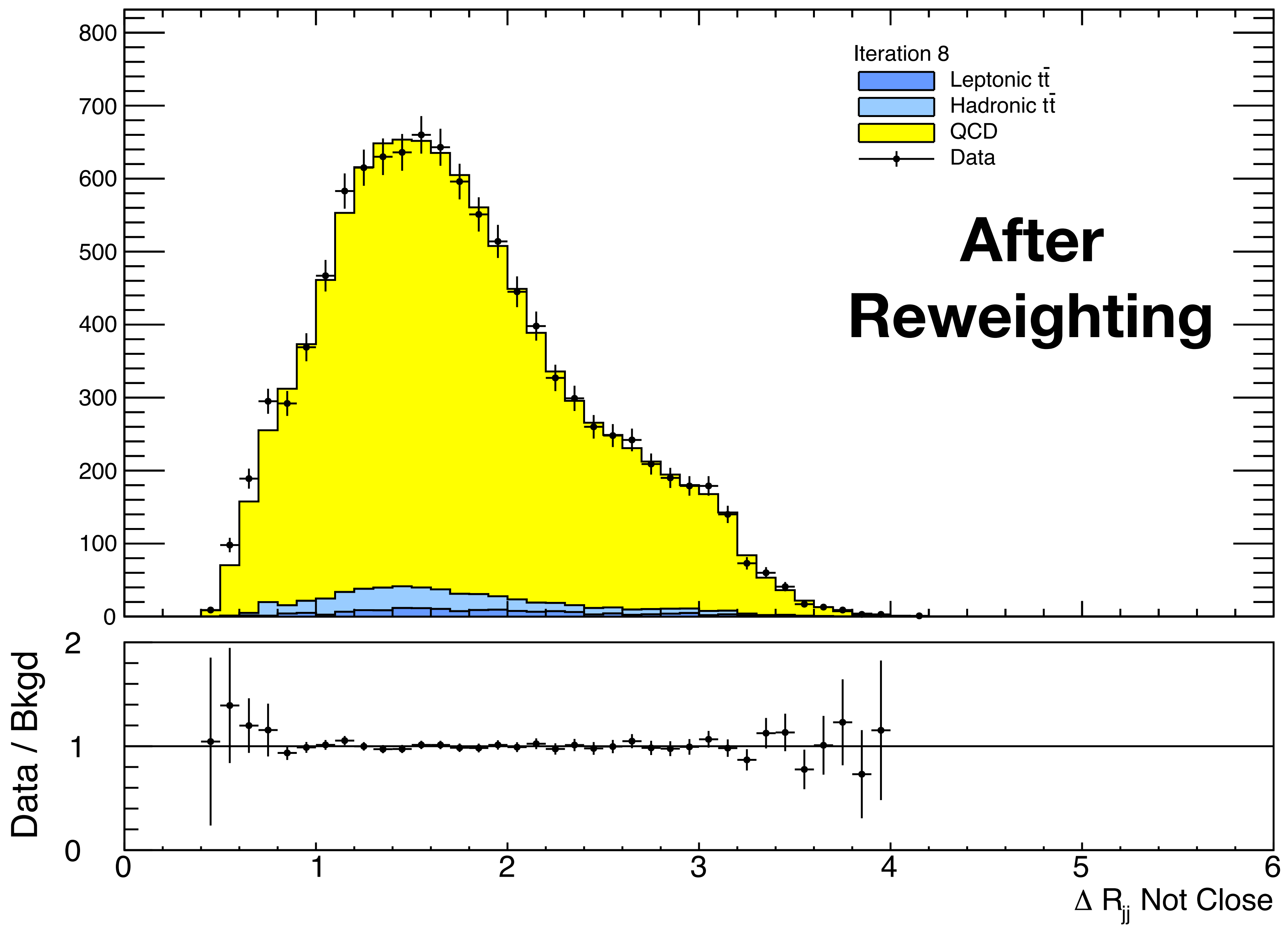
- **QCD Multijet** (~94.6%): Estimated with a data-driven method
  - Difficult to get high statistics, reliable simulation for QCD with our 4 b-jet signal region
- **Hadronic** (~3.6%) and **semi-leptonic** $t\bar{t}$ (~1.7%): Shape from MC simulation, normalization from data driven fit



m_{HH} spectrum from 2016 HH→4b analysis. $t\bar{t}$ (blue) and QCD multijet (yellow) are the two dominant backgrounds (QCD ~95%)

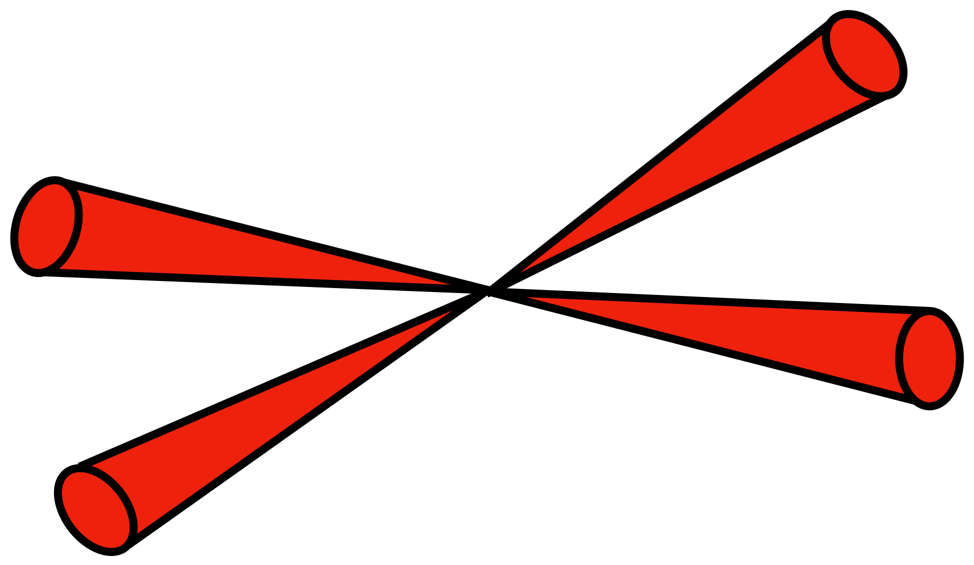
Data Driven QCD



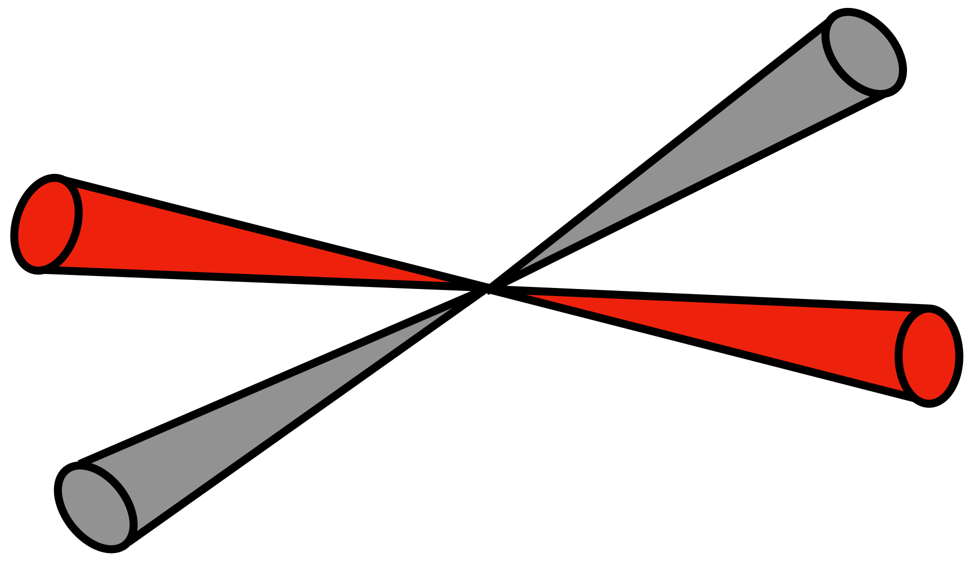


- Hypothesis: events with two b-tags are somewhat similar to events with four b-tags
    - We can correct for differences with a reweighting
- Idea: use reweighted 2 tag data as our QCD background estimate

Data-driven QCD: Procedure



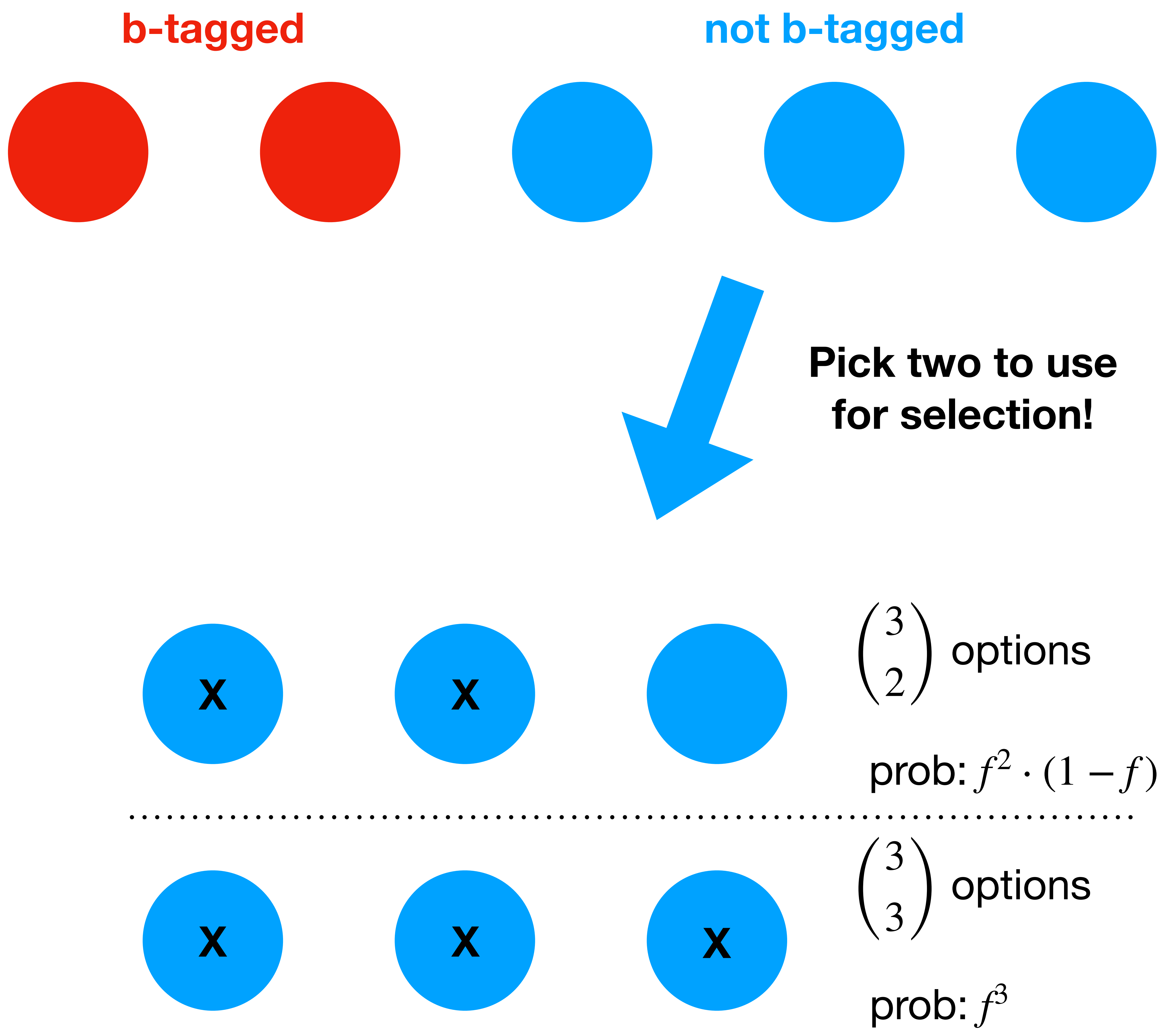
Signal selection
(≥ 4 b-tagged jets)



Background selection (2 b-tagged jets, ≥ 4 total jets)

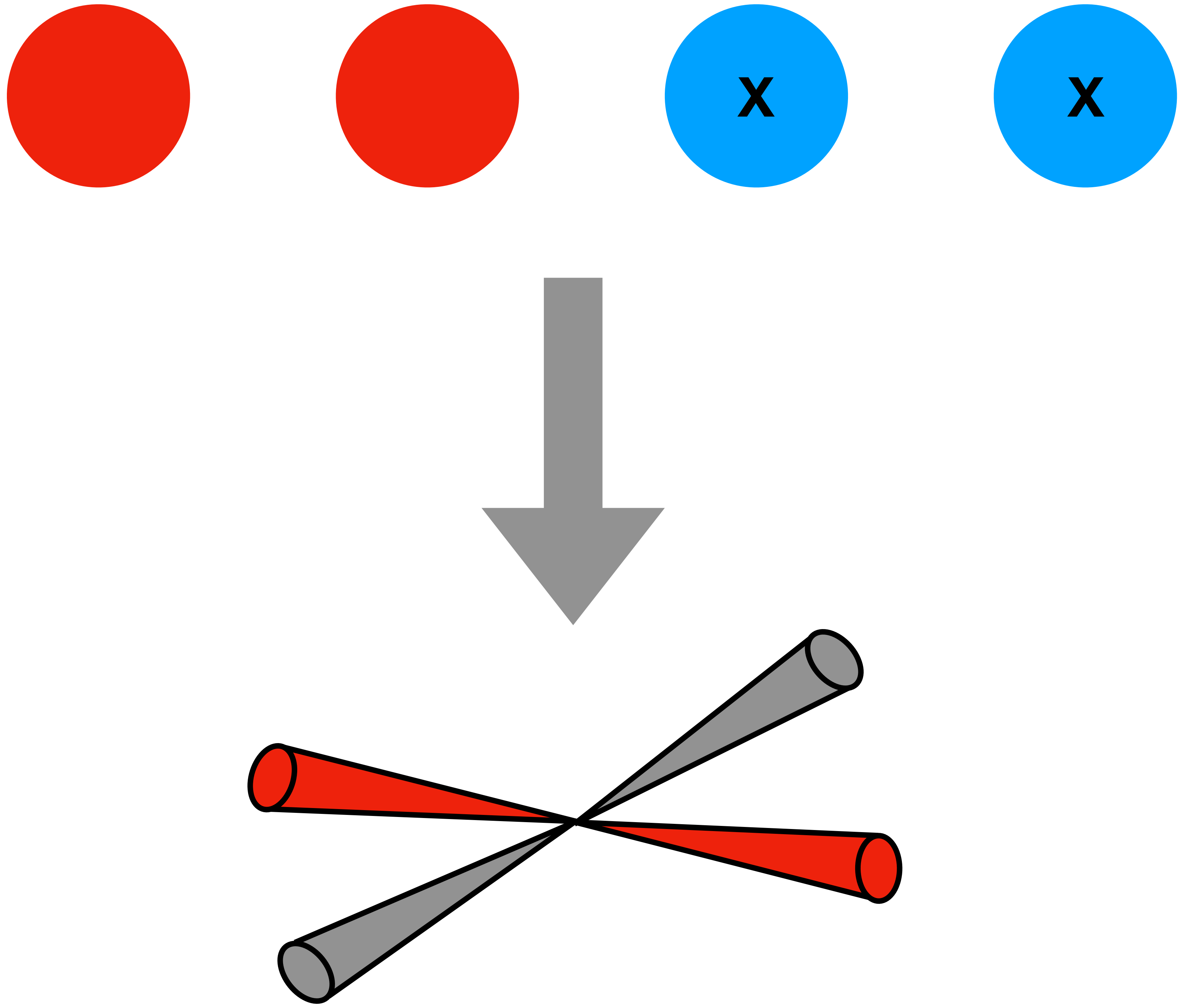
- First, select events with 4 jets (passing basic kinematic requirements), exactly 2 of which are b-tagged

Data-driven QCD: Procedure



- We want to use the same signal region selection => need two more jets!
- Assign "pseudo-tags" among the remaining jets so that tagged+pseudo-tagged >= 4
  - Flip a coin for each jet with probability f

Data-driven QCD: Procedure



$$\text{nJetWeight} = \binom{3}{2}\cdot f^2\cdot(1-f)+\binom{3}{3}\cdot f^3 = 3f^2\cdot(1-f)+f^3$$

Weight for the case of 3 non-tagged jets

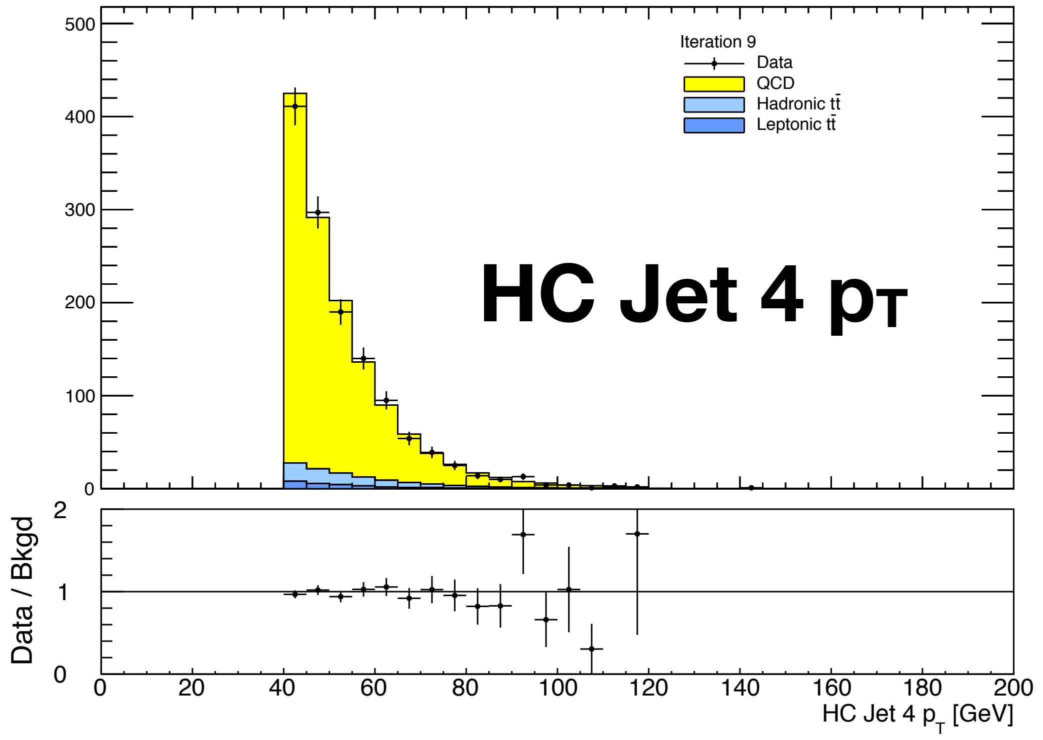
- Use the 4 jets (tagged+pseudo-tagged) with the highest b-tagging score to form Higgs candidates and do the rest of the cutflow
- Assign a weight to the event:

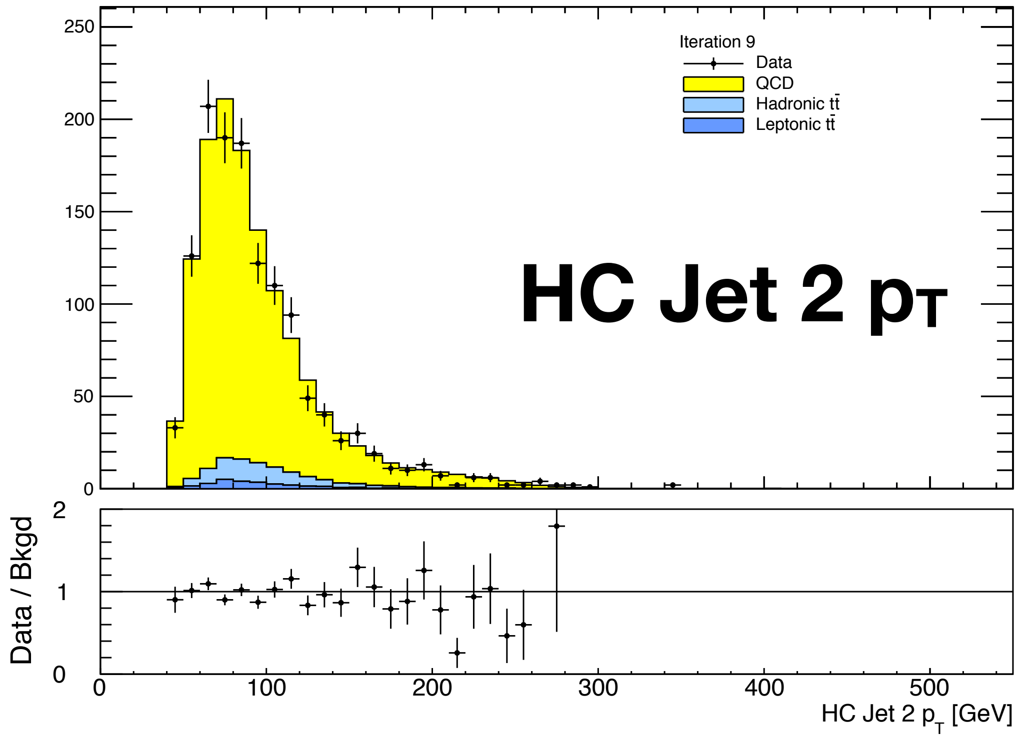
$$\text{nJetWeight} = \sum_{i=2}^{n}\binom{n}{i}f^i\cdot(1-f)^{n-i}$$

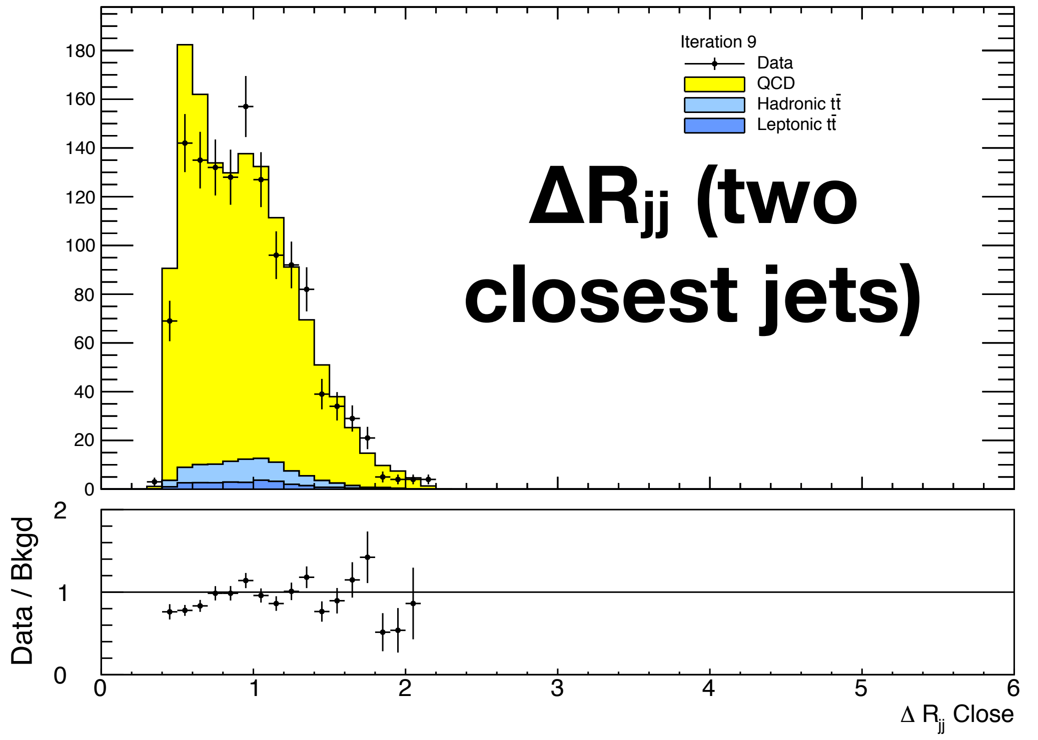
= sum of probabilities of each choice

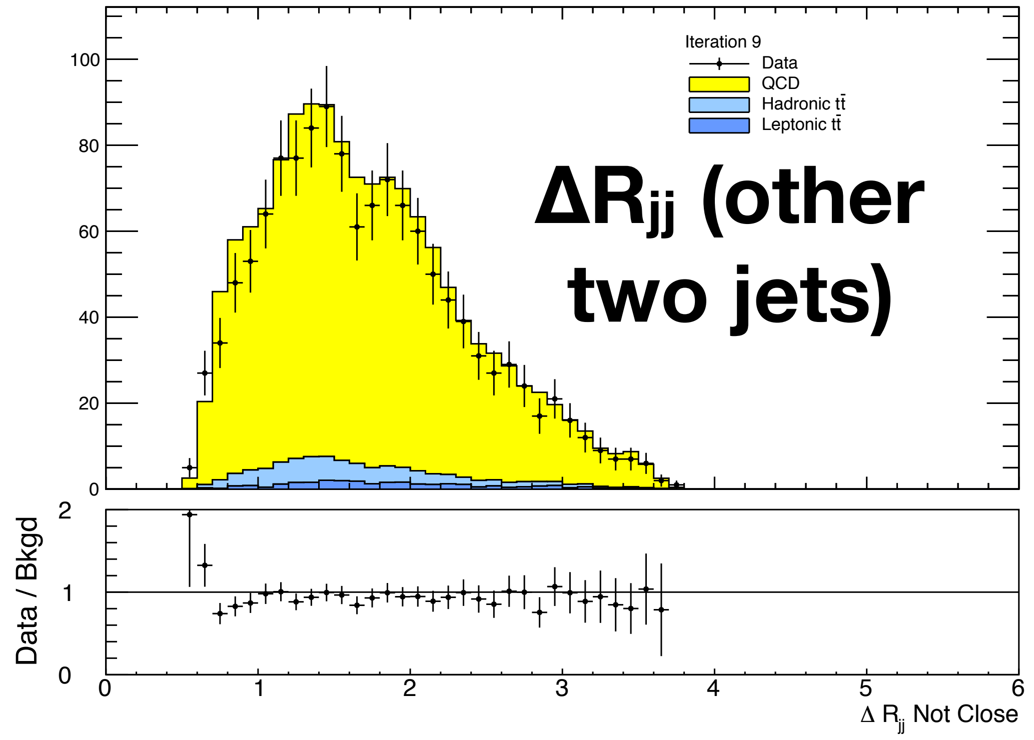
Kinematic Reweighting

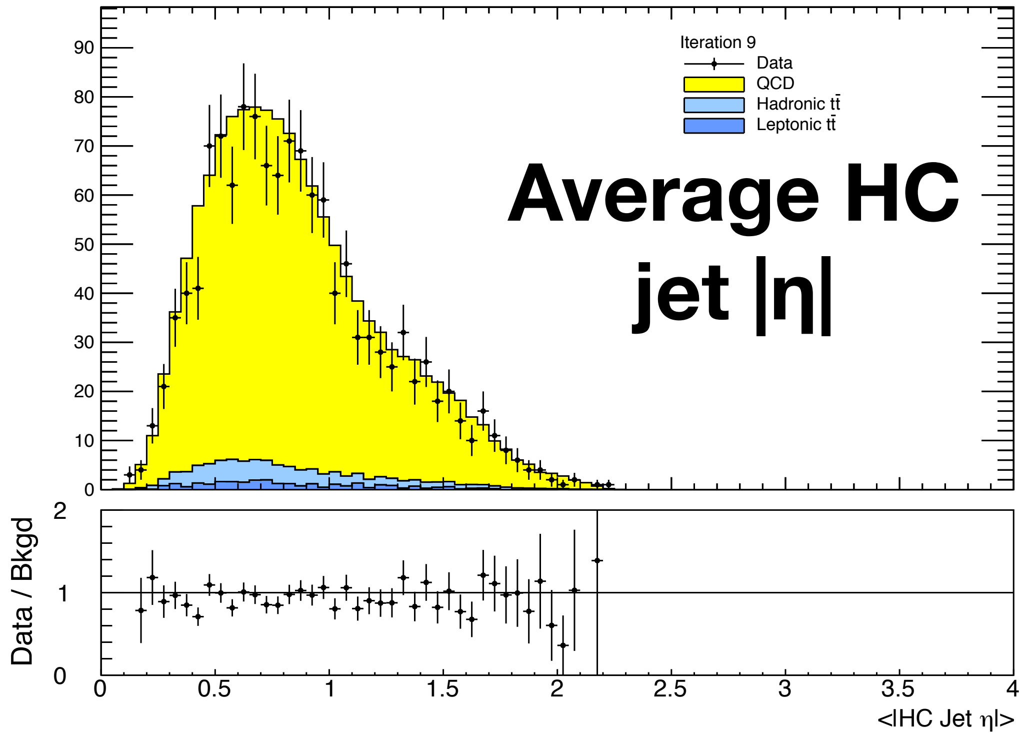
- Now we have 2 tag events in our signal region
- We then correct kinematic differences by deriving weights in our **sideband region**
- Reweighting procedure:
    - Pick distributions sensitive to 2 vs 4 tag differences
    - Subtract simulated ttbar component for 2 and 4 tag
    - Take ratios of remaining distributions and form splines











The Analysis: Background Estimation

- Conceptually, weight for an event is then

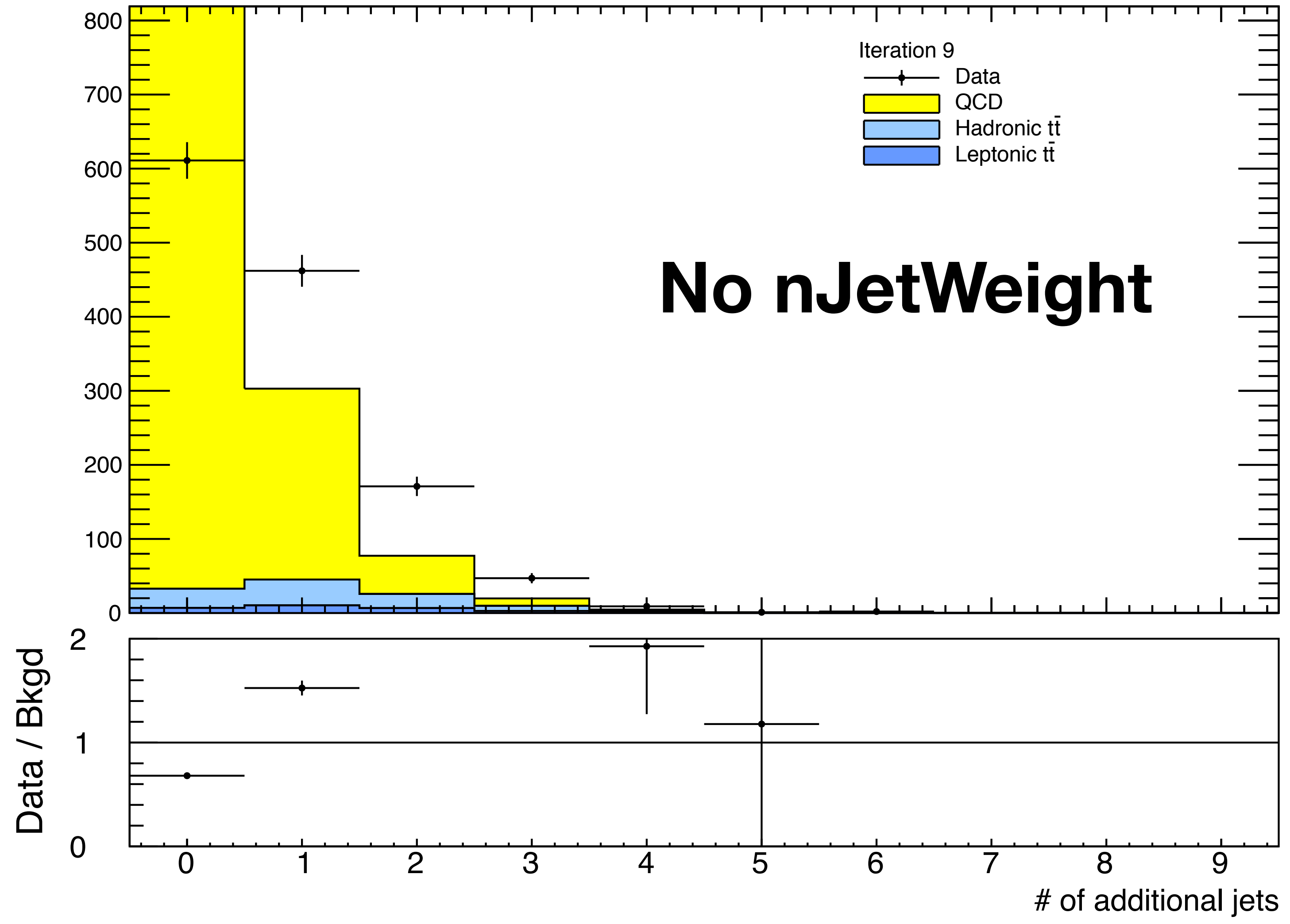
$$w = \text{nJetWeight} \times \prod_{a \in A} f^a(x_a)$$

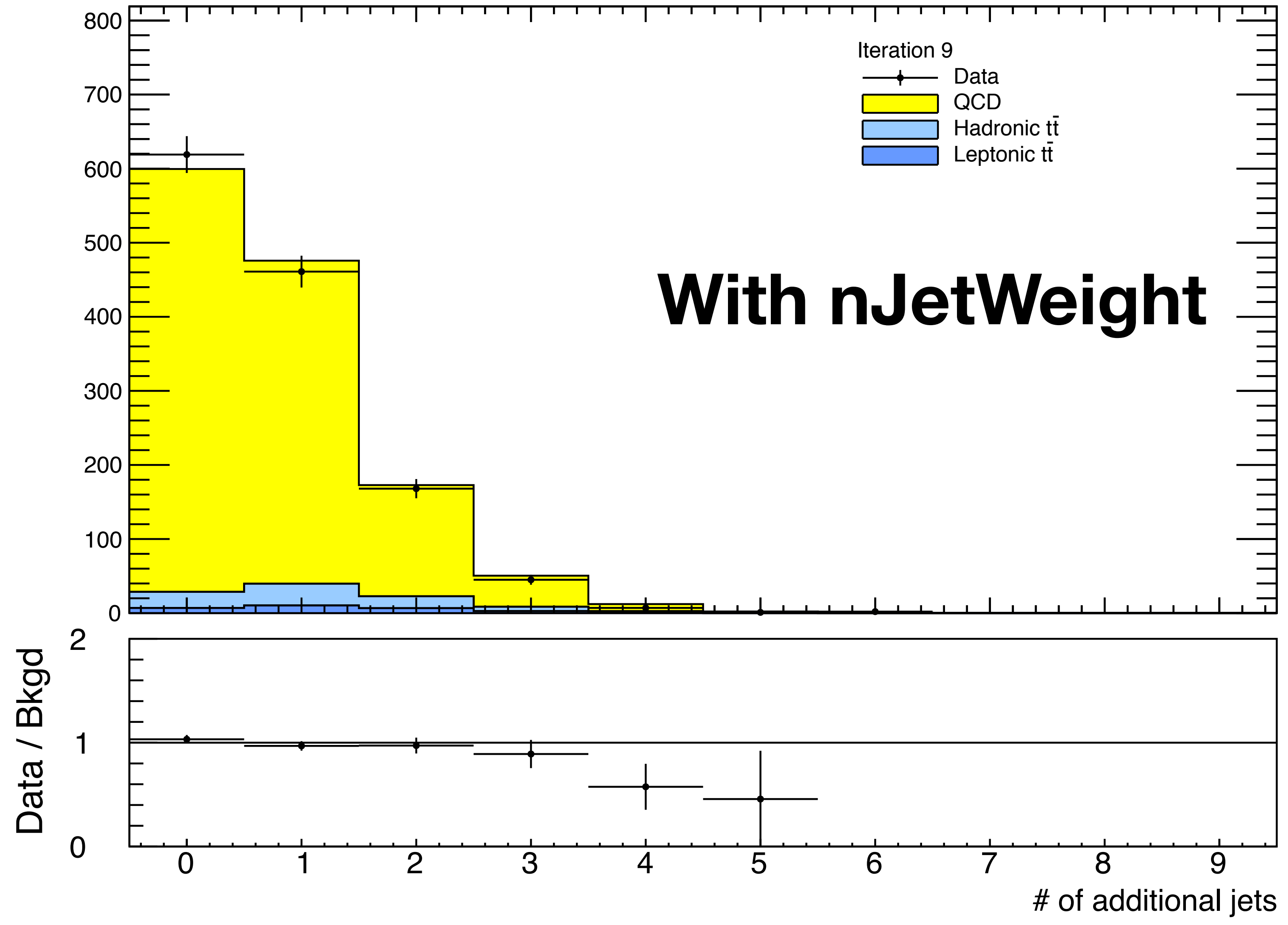
  - $a \in A$ runs over the reweighting distributions
  - $f^a(x_a)$ is taken from the splines
  - nJetWeight depends on a fit of the pseudo-tag probability, f, to match the distribution of the number of jets

- In practice, this is a bit ad hoc, so we need to iterate. Adding a factor that approaches 1 for later iterations, our event weight is then

$$w = \text{nJetWeight} \times \prod_{i=0}^{i<I} \prod_{a \in A} [(f_i^a(x_a) - 1) \times (1 - 2^{-i-1}) + 1]$$

where i runs over iterations

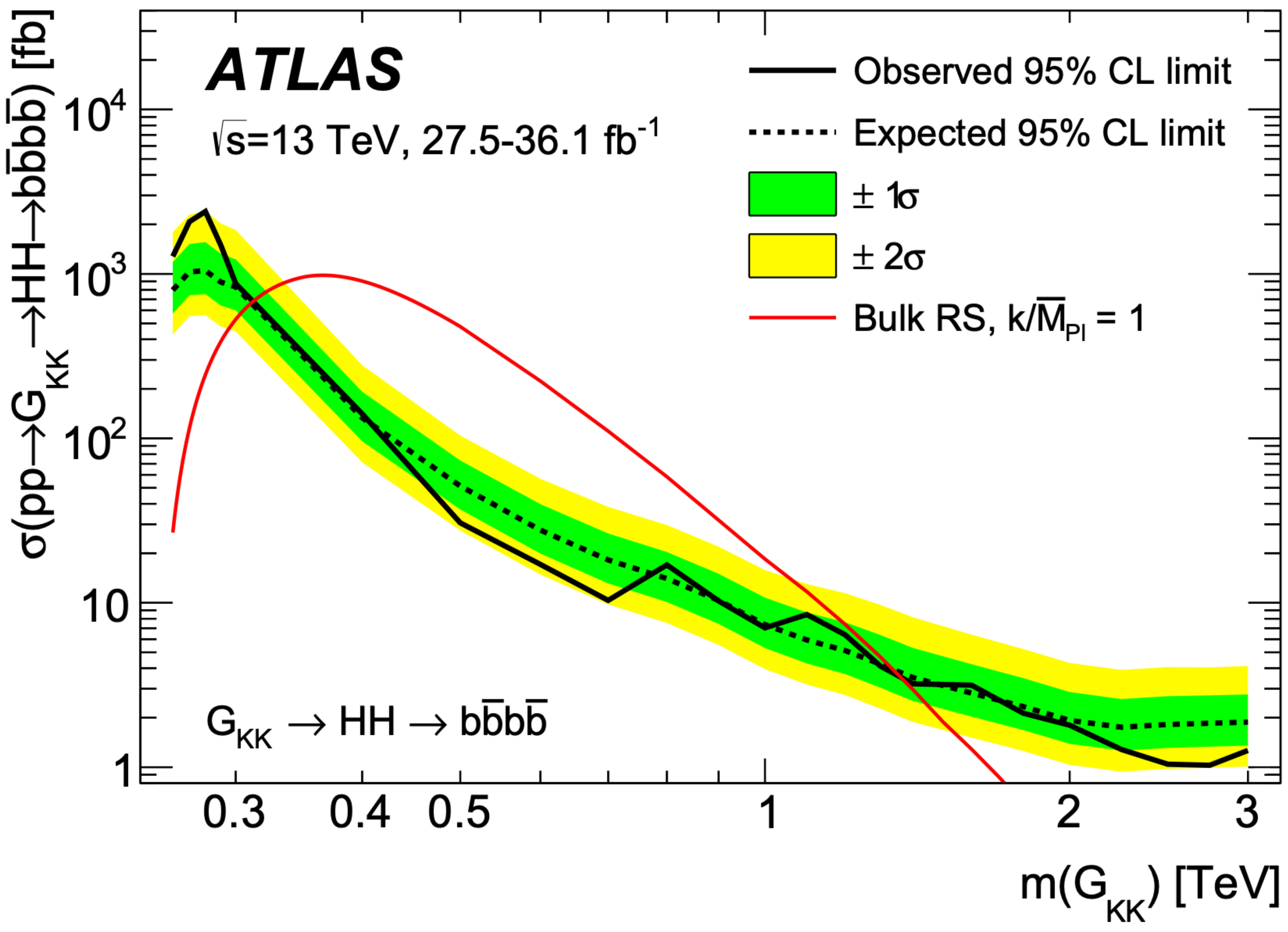




HH→4b: Results

The Analysis: Results

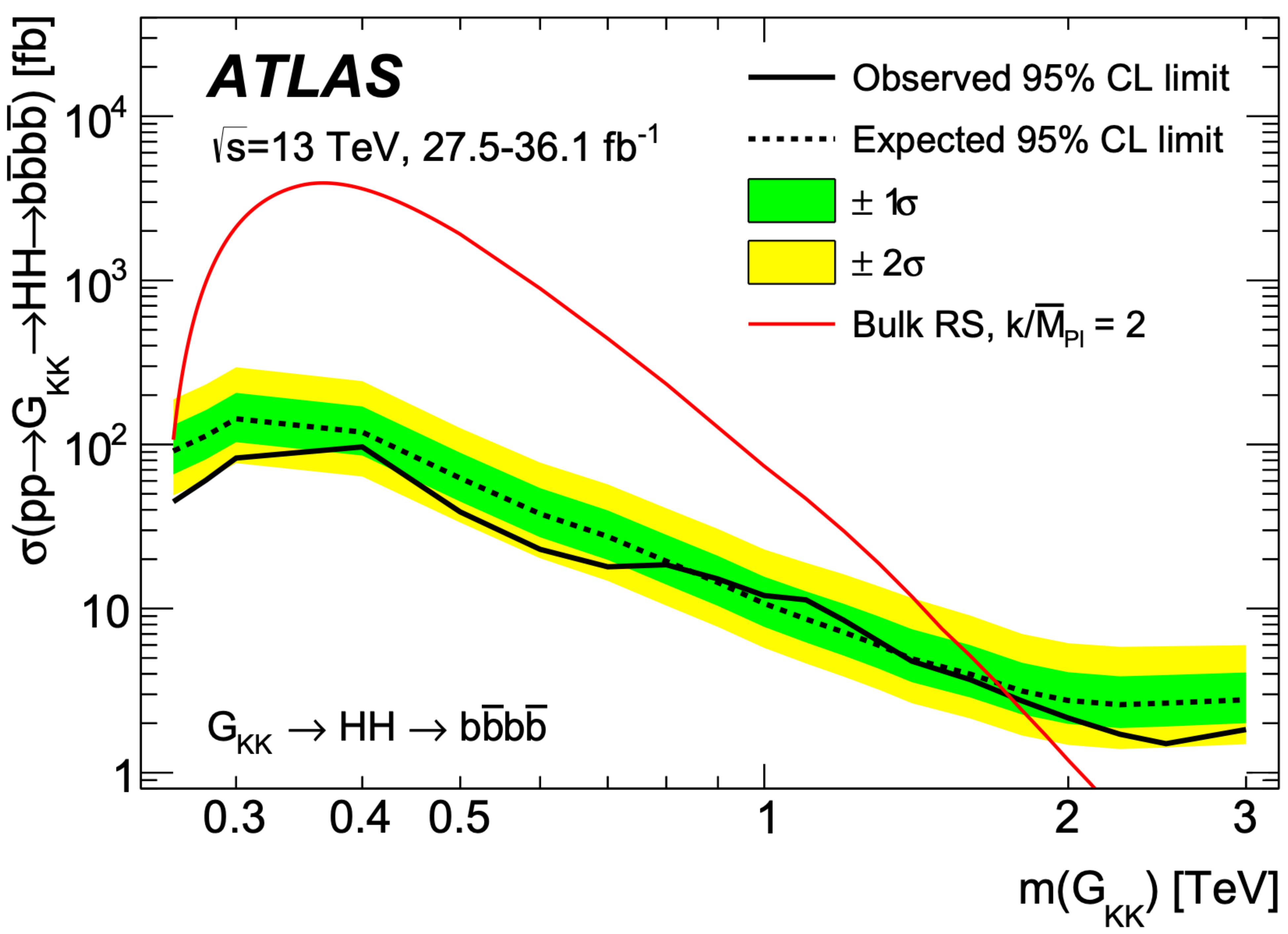
- We have not yet observed anything in the HH→4b channel
- We thus set upper limits on the cross sections for models we consider
- We report here the published results for the combined (boosted and resolved) analyses
    - Limits for c=1.0 graviton and Standard Model HH shown here
    - All cross sections above the line are excluded

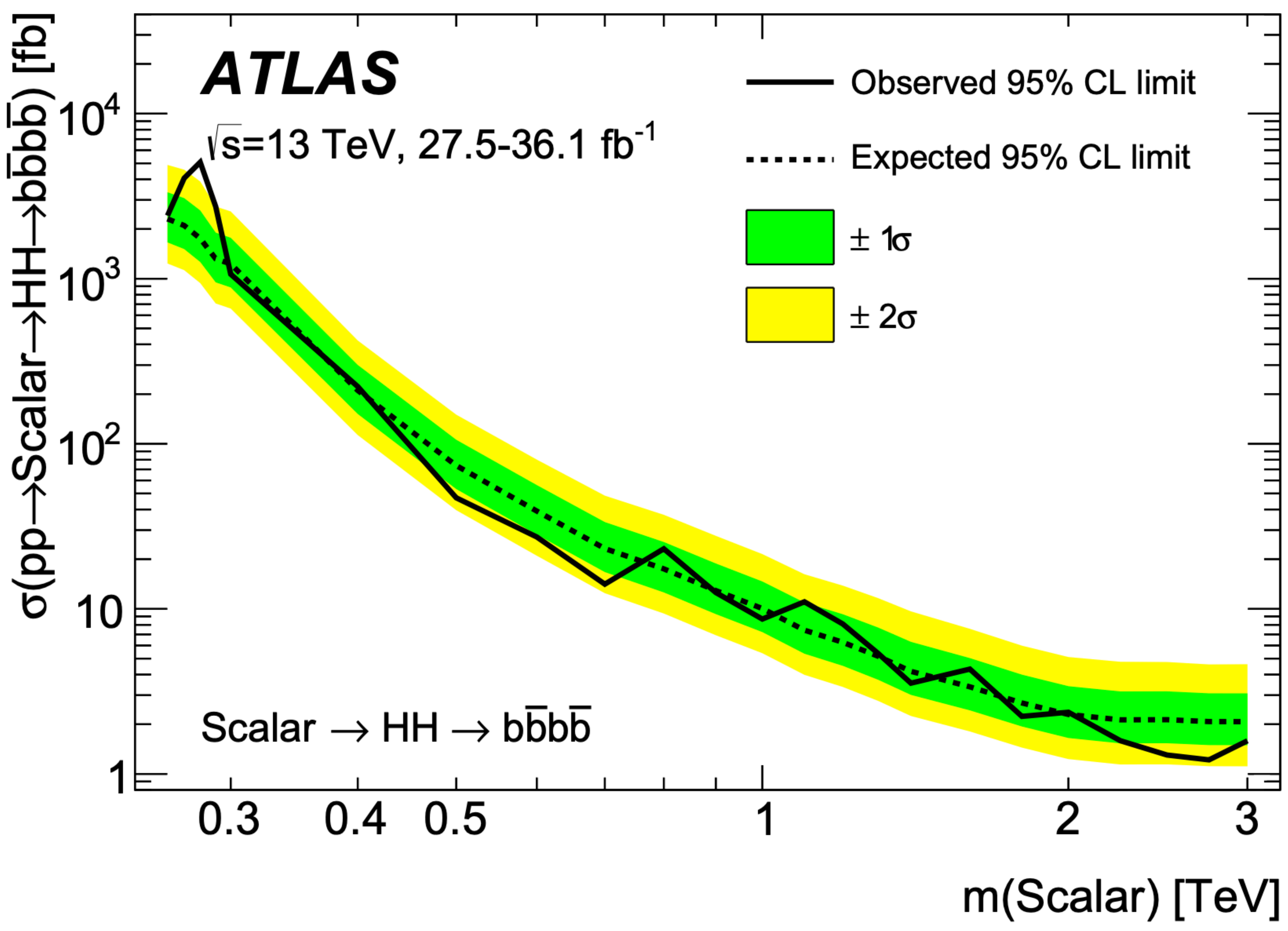


Observed	-2σ	-1σ	Expected	$+1\sigma$	$+2\sigma$
12.9	11.1	14.9	20.7	30.0	43.6

SM HH limits (as multiples of the SM cross section)

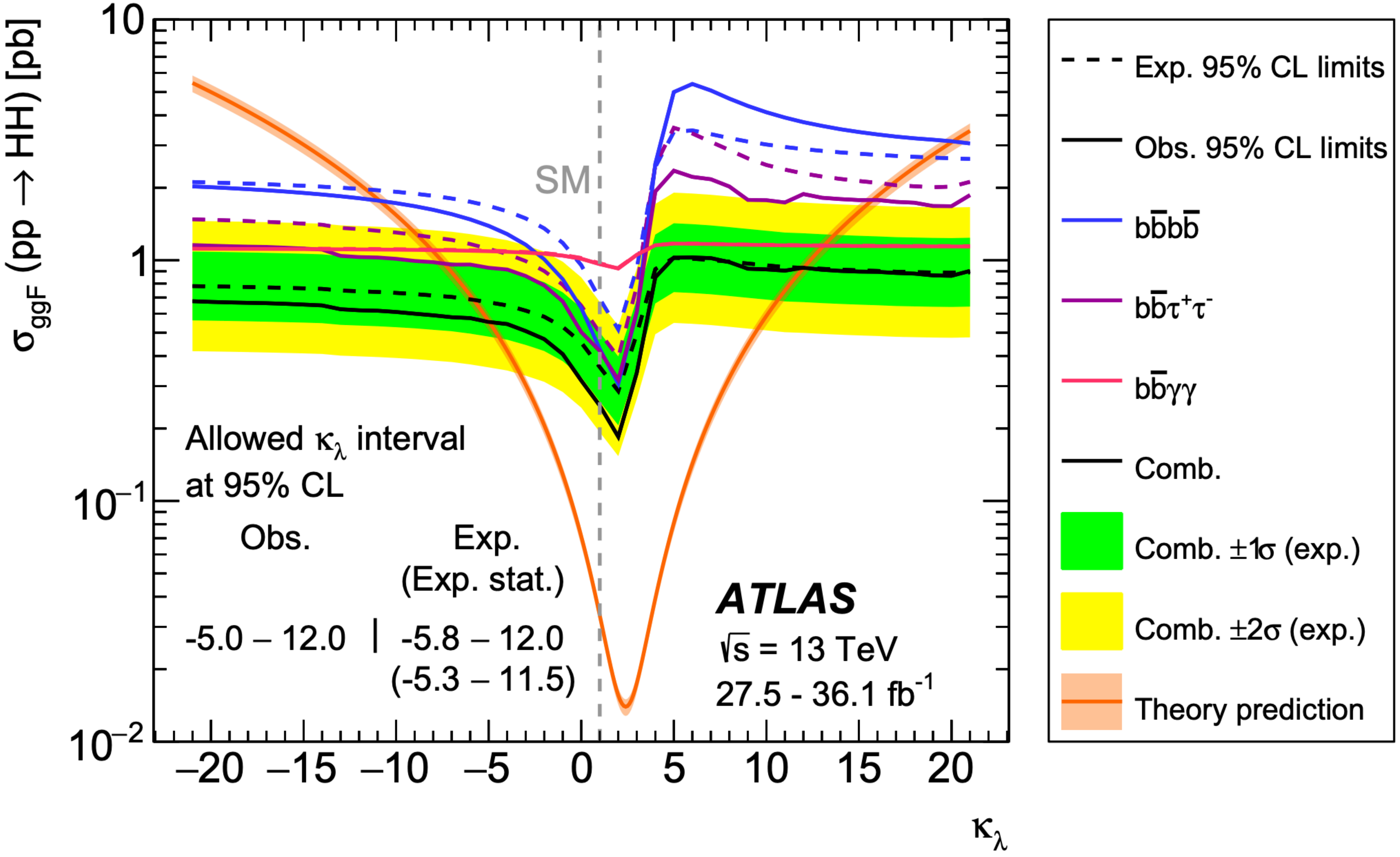
The Analysis: Results





- Limits for c=2.0 graviton and generic narrow width scalar

The Analysis: κ_λ Scan

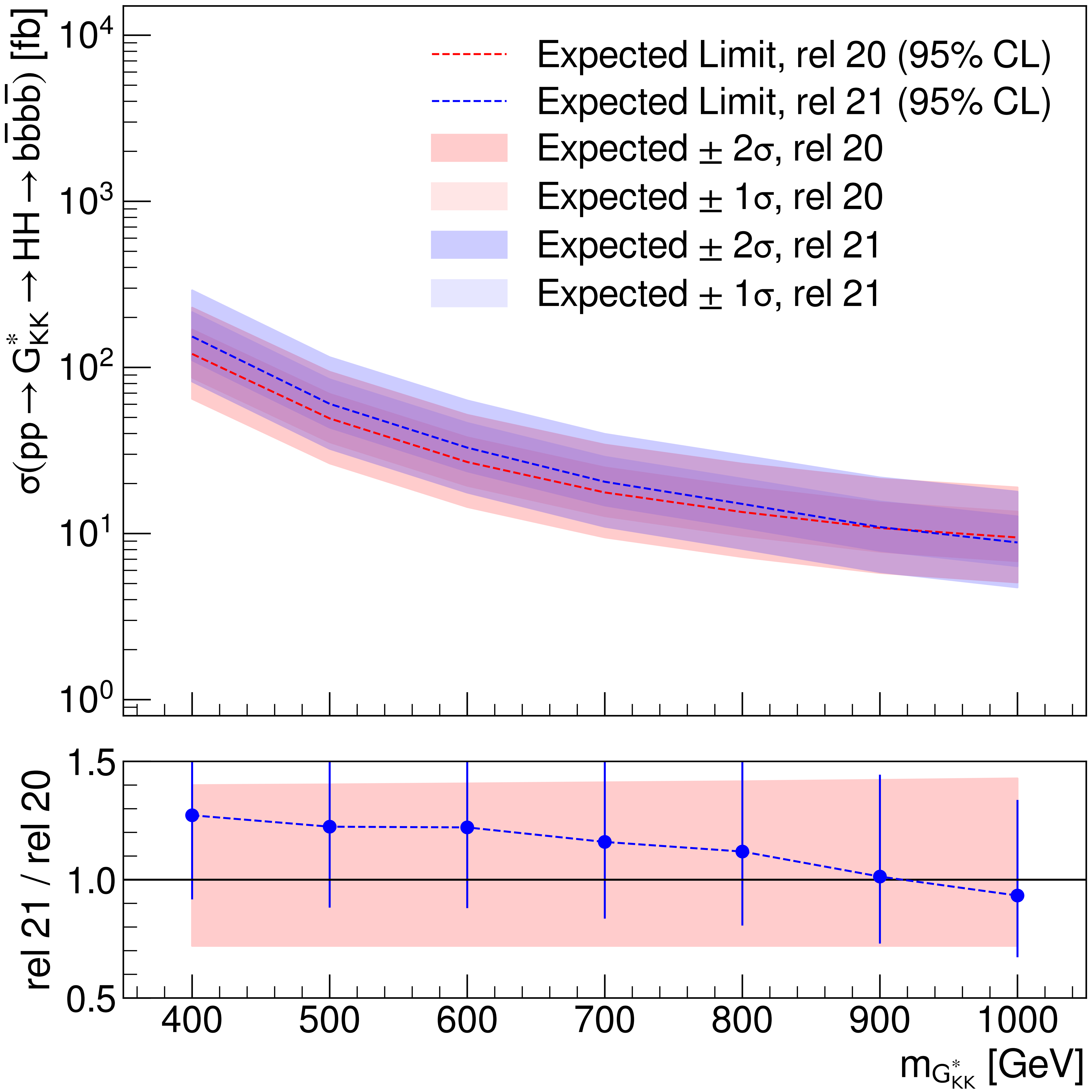


- Results from a scan over values of κ_λ for combined non-resonant HH production channels [ref]
  - Recall $\kappa_\lambda = 1$ is the Standard Model value
- Allowed values are restricted to be between -5.0 and 12.0 (observed limits)
- 4b alone restricts the allowed range to be between -10.9 and 20.1

HH→4b: Looking Forward

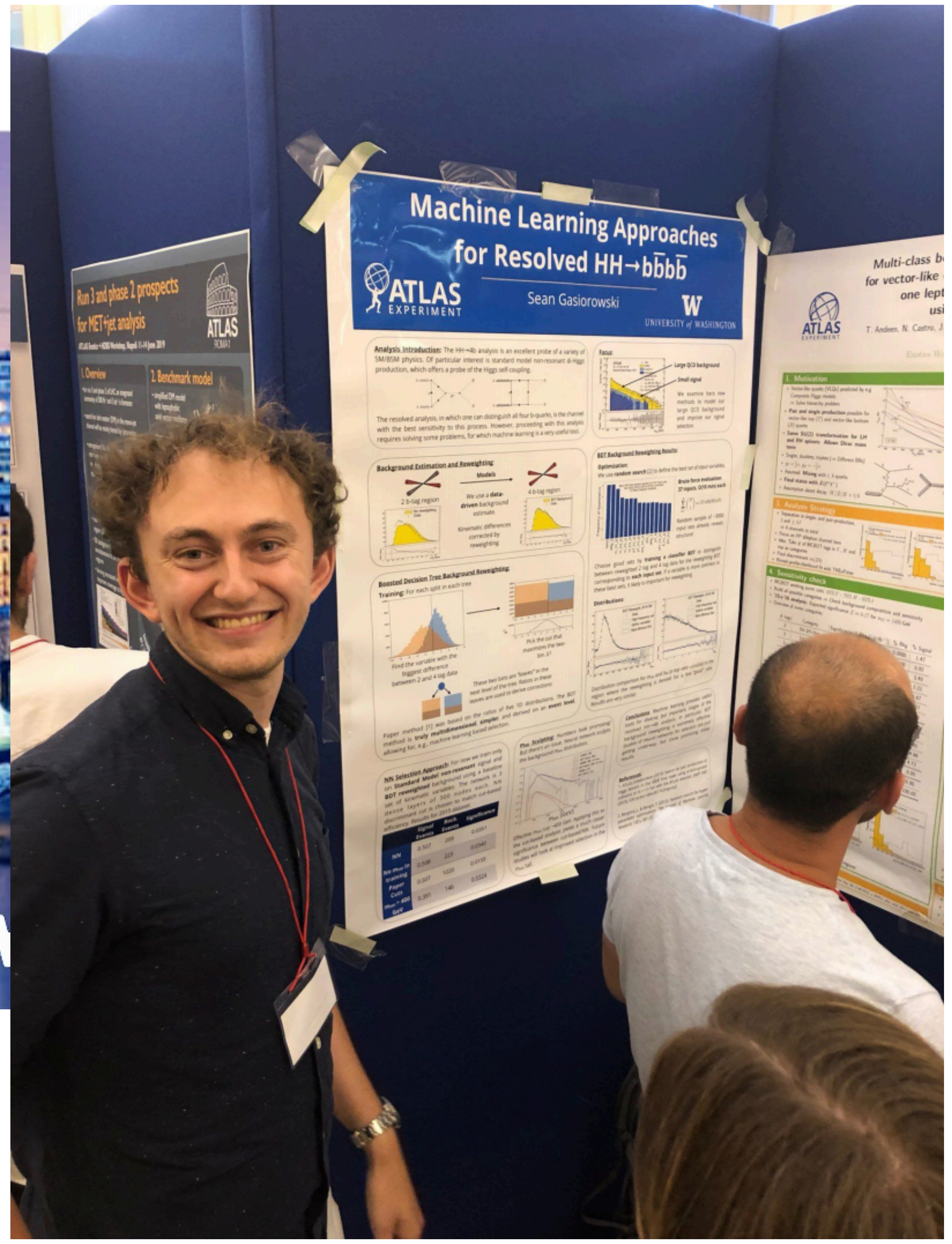
Old Analysis, New Software

- As mentioned, I've reproduced the published resolved analysis and performed an extensive set of cross checks
- Show on the right limits for the previous (red) and the new (blue) ATLAS software for the c=1.0 graviton
- Limits are ~25% worse!
  - Consistent with observed changes in MC yields
- Changes are understood to be due to b-tagging modeling improvements



Towards the Future

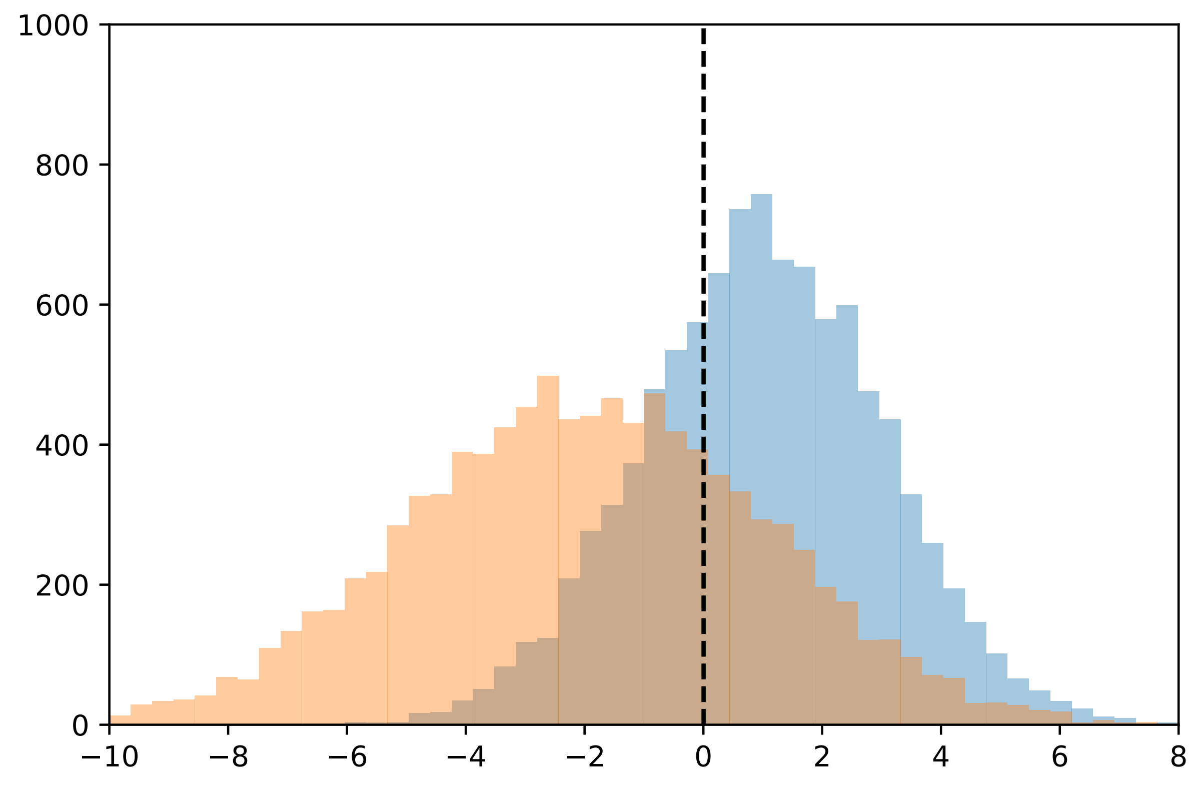
- Lightning talk and poster at ATLAS Exotics/HDBS workshop (Naples, Italy, June 2019)
    - Winner of best poster (and a cool orange backpack)!
- Poster at ATLAS week (CERN, June 2019)



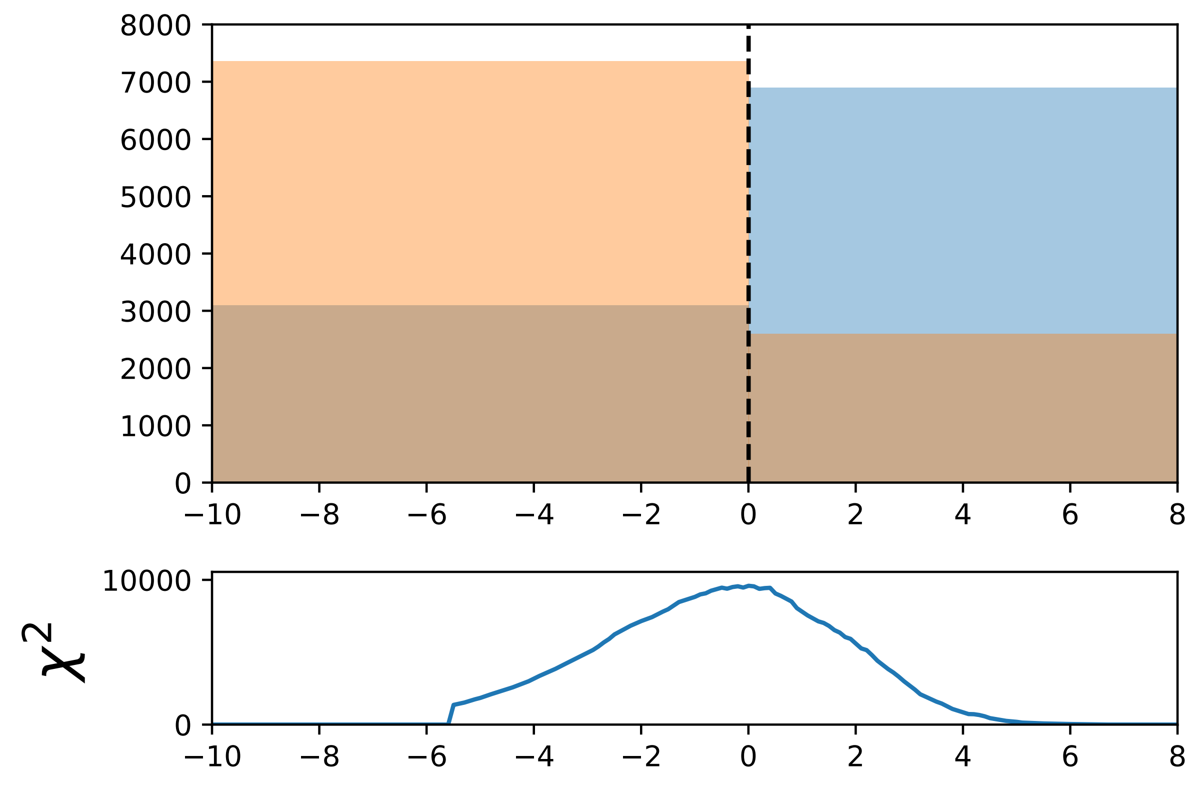
Towards the future: Areas of Focus

- **Background estimation**:
  - Previous method is complicated, ad hoc
    - Multiplying splines together doesn't properly account for correlations between variables
  - New idea: Boosted Decision Tree (BDT) for background estimation!
- **Selection**:
  - Machine learning tools have a lot to offer - what can they do for our analysis?
- Much activity in the group on several other areas of optimization. I will, in addition, be working on several aspects of the baseline analysis (systematics, fits, limits, etc)

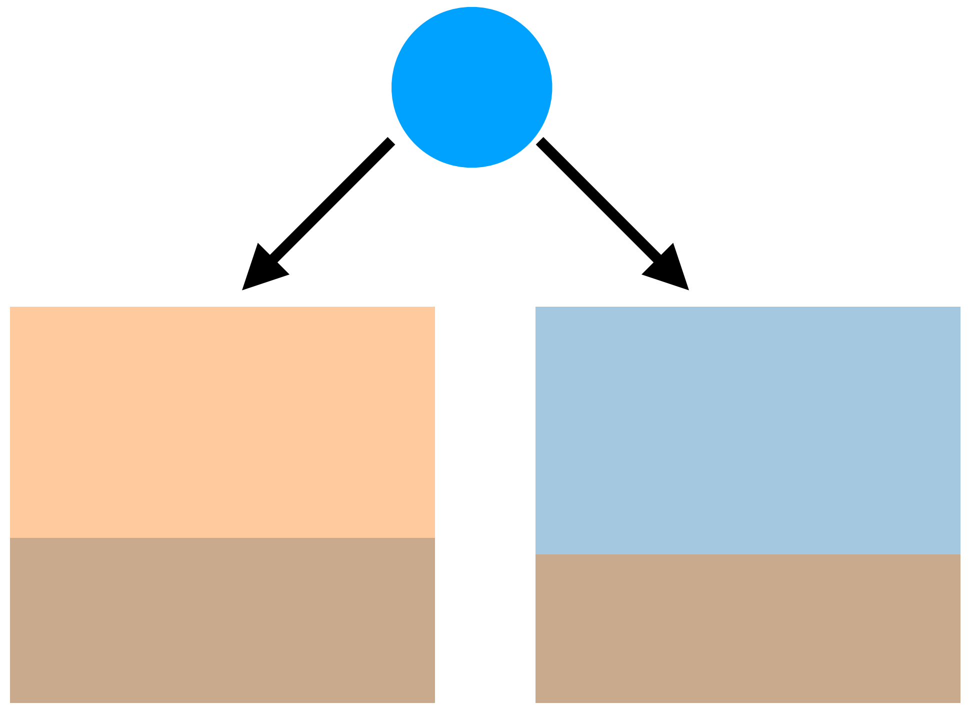
Reweighting BDT: The Approach



Find the variable with the biggest difference between 2 and 4 tag data

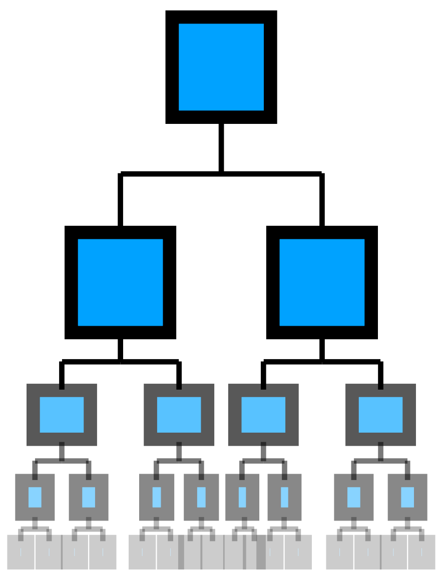


Pick the cut that maximizes the two-bin χ^2

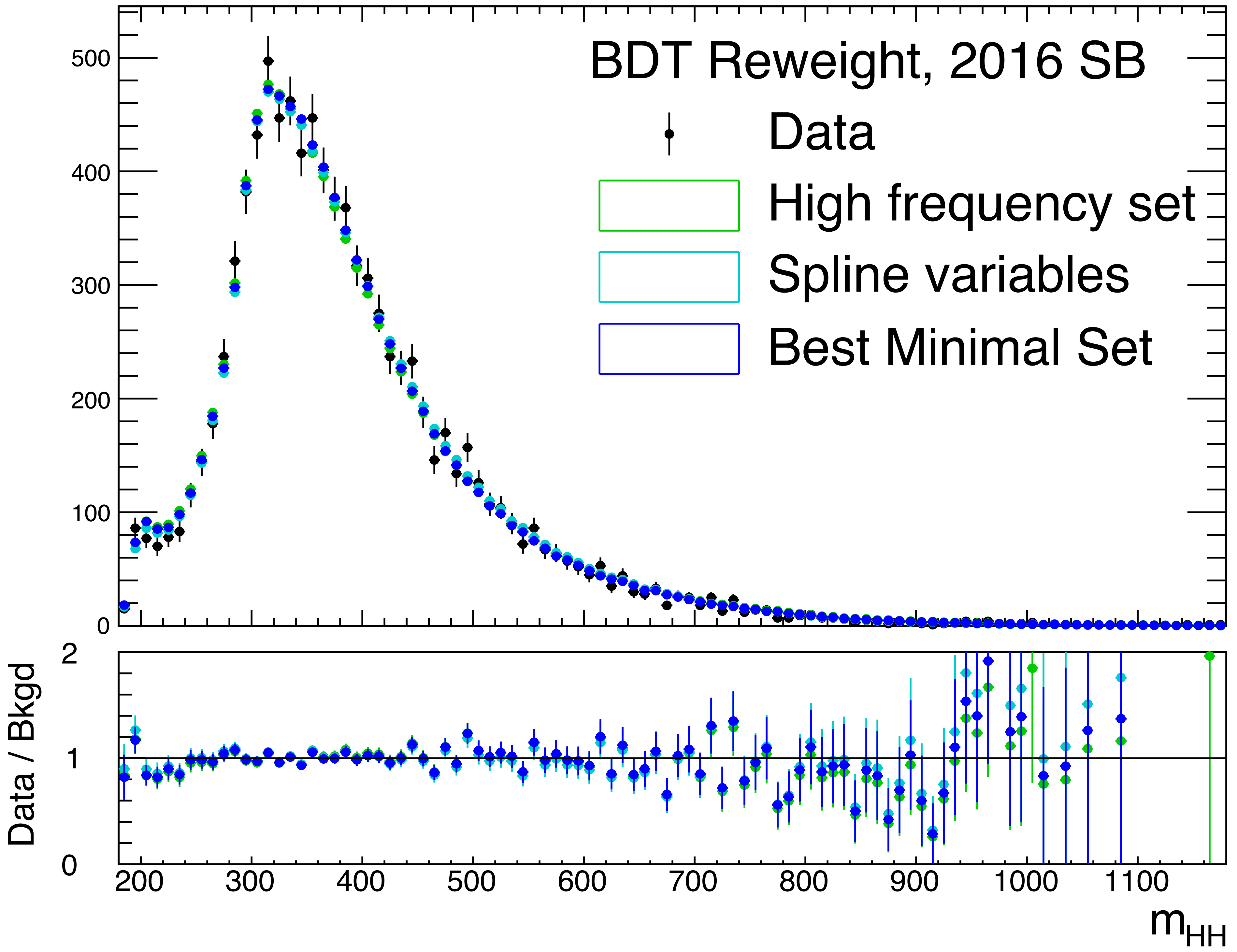


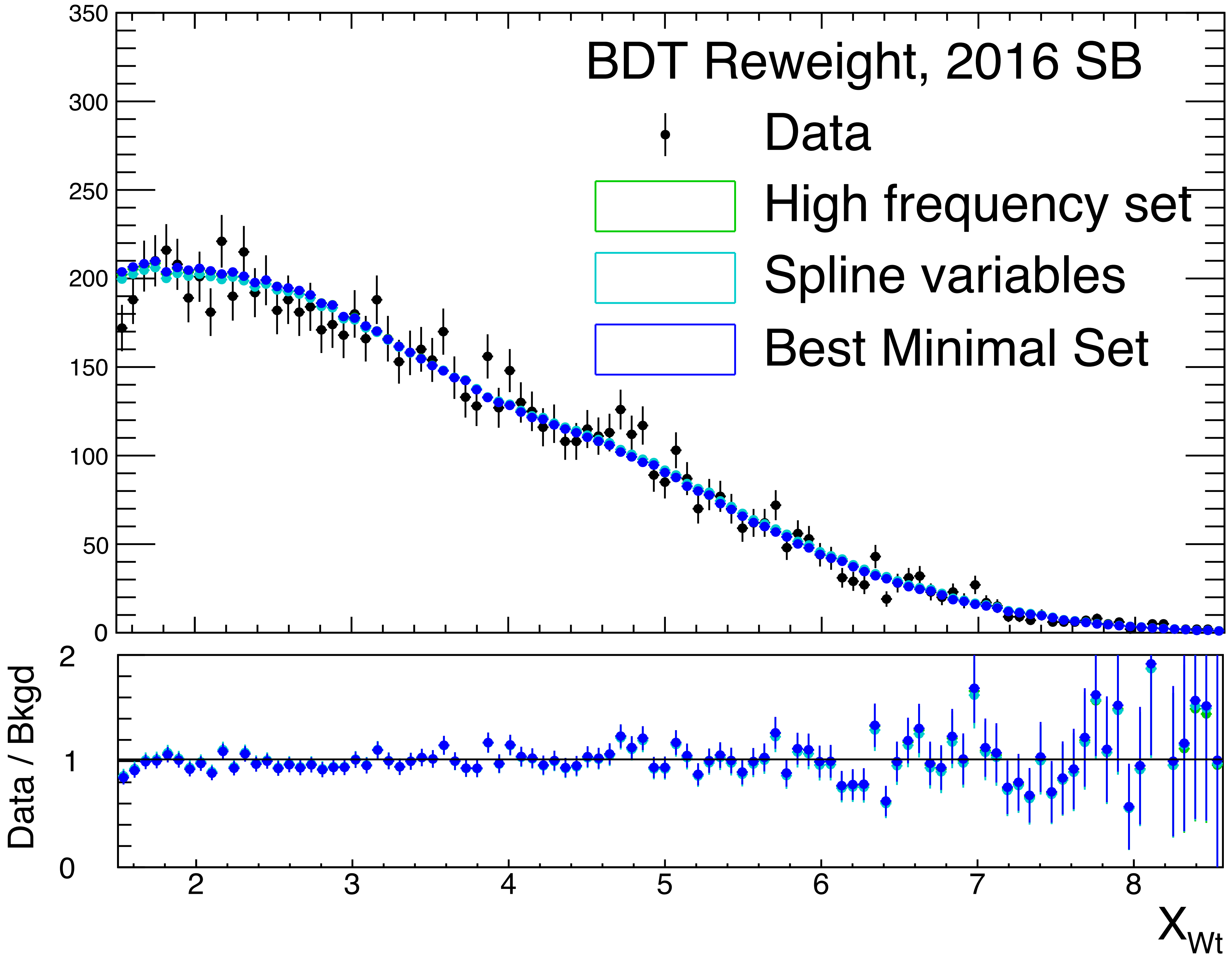
These two bins are "leaves" in the next level of the tree. Ratios in these leaves are used to derive corrections

- The general procedure is shown above for some given set of input variables
- We consider here a **fully data driven** background (recall, QCD ~95% of the background)
- The BDT method is **simpler** and **truly multidimensional**



Reweighting BDT: Results

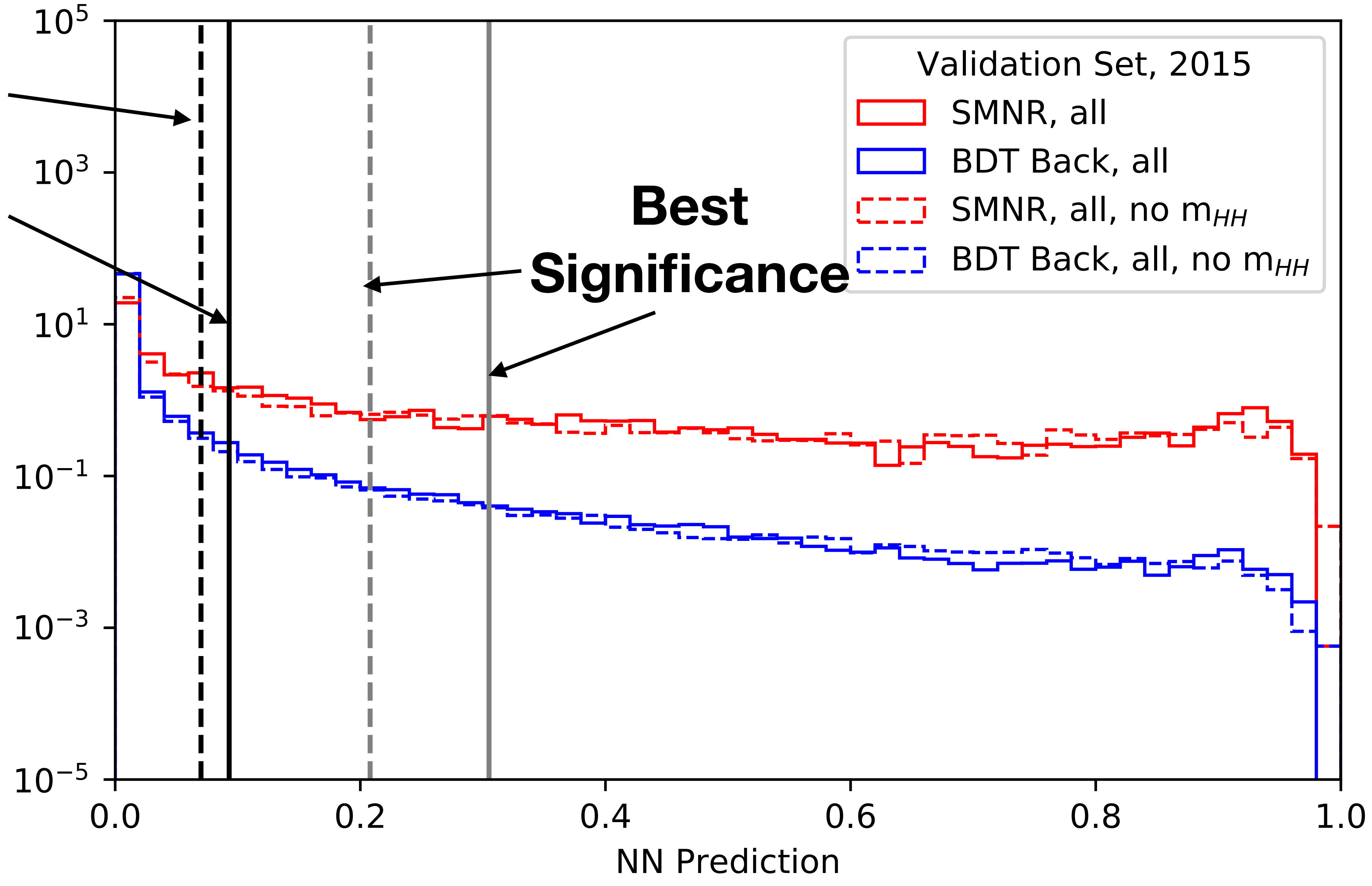




- Work is ongoing on optimization of the BDT - in particular, choosing which variables are sensitive to 2 vs. 4 tag differences
- Results for a few different input set choices are shown above. Agreement with data is very good

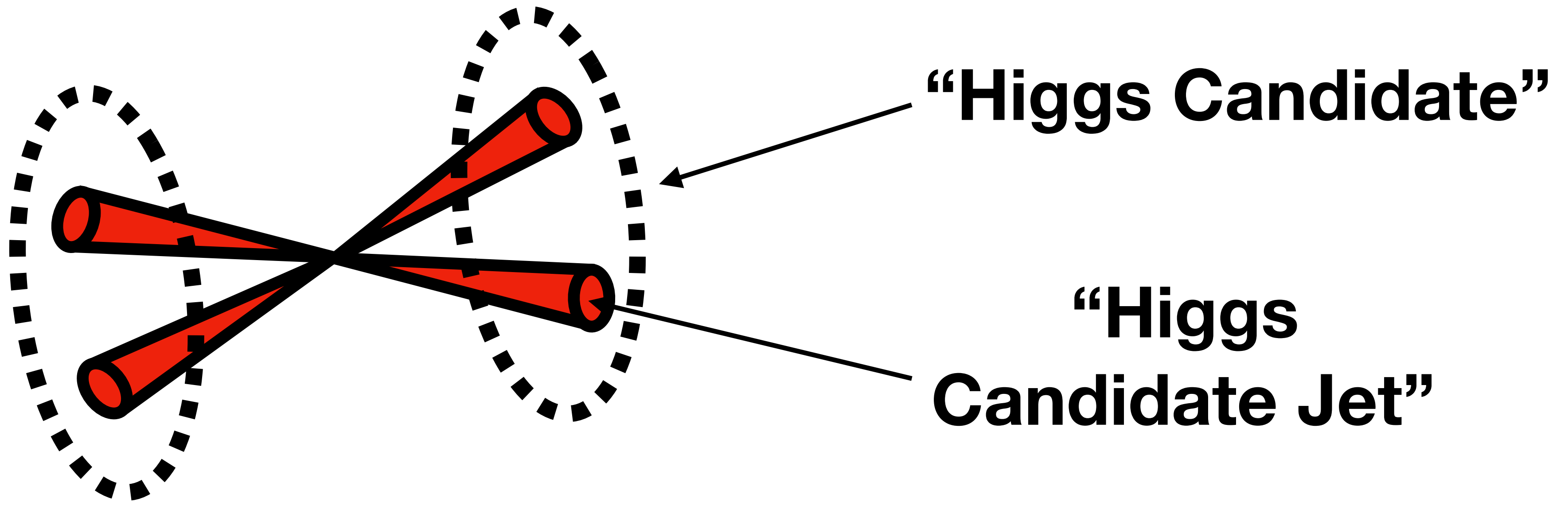
Neural Network Selection

Match Paper Analysis Signal Efficiency (results on next slide)



NN discriminant. Solid lines trained with variables on right, dashed lines same, but without m_{HH}. Histograms are normalized to 1.

$$\text{Significance} = \sqrt{2 \cdot ((s+b) \cdot \ln\left(1 + \frac{s}{b}\right) - s)}$$



- Try to improve on the cut-based selection by using a neural network. Looking at Standard Model non-resonant signal here
- Input variables:
  - m_{HH} (unless specified)
  - p_T, η, ϕ, m, and E for each Higgs candidate and HC jet
  - X_{Wt} (top veto variable) and X_{hh} (distance from (120 GeV, 110 GeV) in Higgs candidate mass plane)

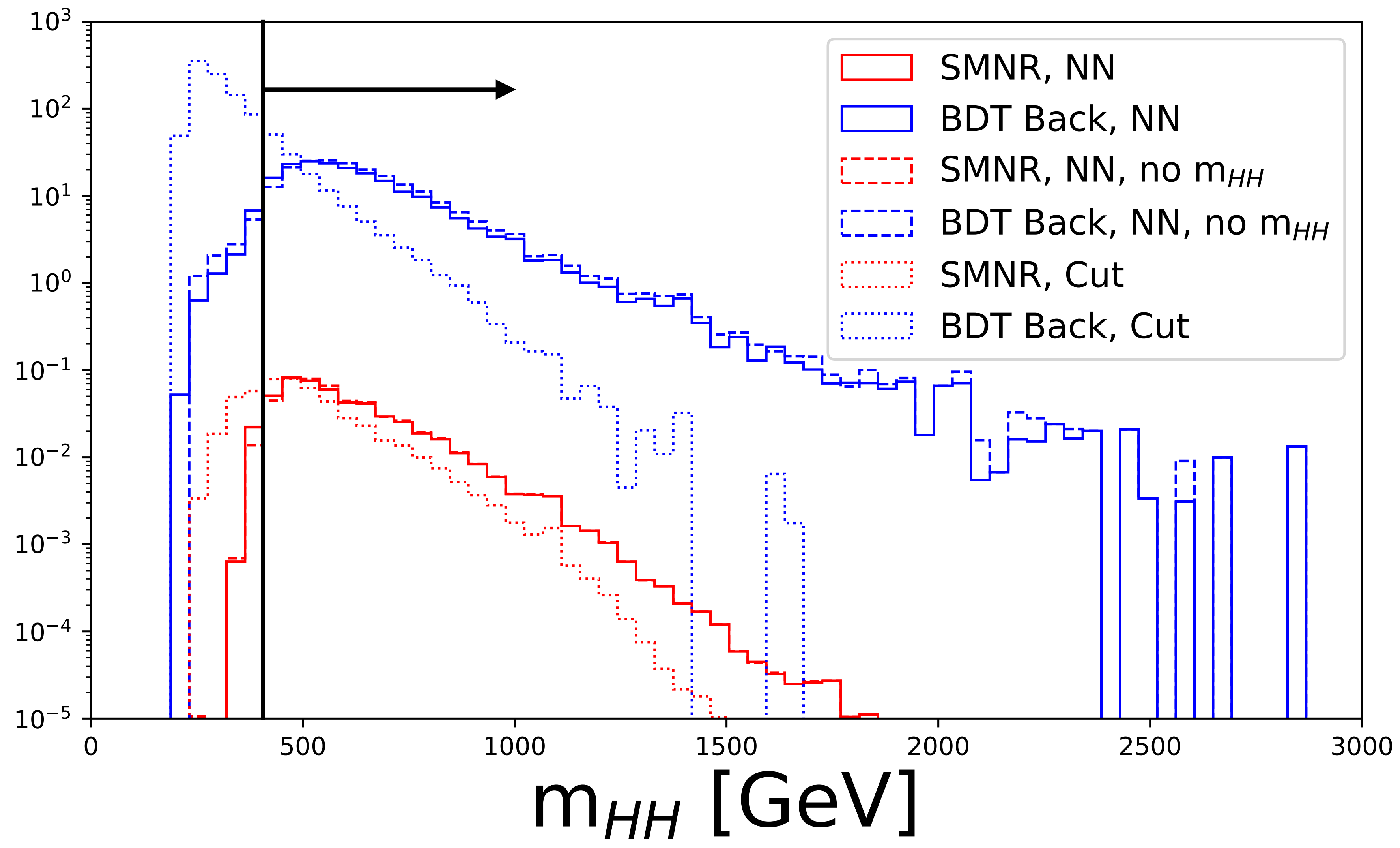
S vs. B: Neural Networks and m_{HH}

- NN improves significance, but sculpts the m_{HH} distribution (used for limit setting)

- This results in an "effective m_{HH} cut" at around 400 GeV

    - Applying this cut on top of the paper cuts makes the NN/cut-based results much closer - selection power in m_{HH}!

- Future interest: Explore parameterizing selection as a function of m_{HH}



	Signal Events	Back. Events	Significance
NN	0.507	209	0.0351
Paper Cuts	0.507	1020	0.0159
m_{HH} > 400 GeV	0.391	146	0.0324

Timeline and Conclusions

Summary of Progress

- **Completed ATLAS qualification**
  - Developed and implemented method to improve shower shape modeling
- **Reproduced the baseline published HH→4b analysis**
  - Involved writing a more flexible, understandable code base
- **Performed extensive cross-checks between ATLAS software releases**
  - Up to 25% change in limits due to changes in b-tagging
- **Began detailed work on improvements to the HH→4b analysis**
  - In particular, focusing on selection and background estimation

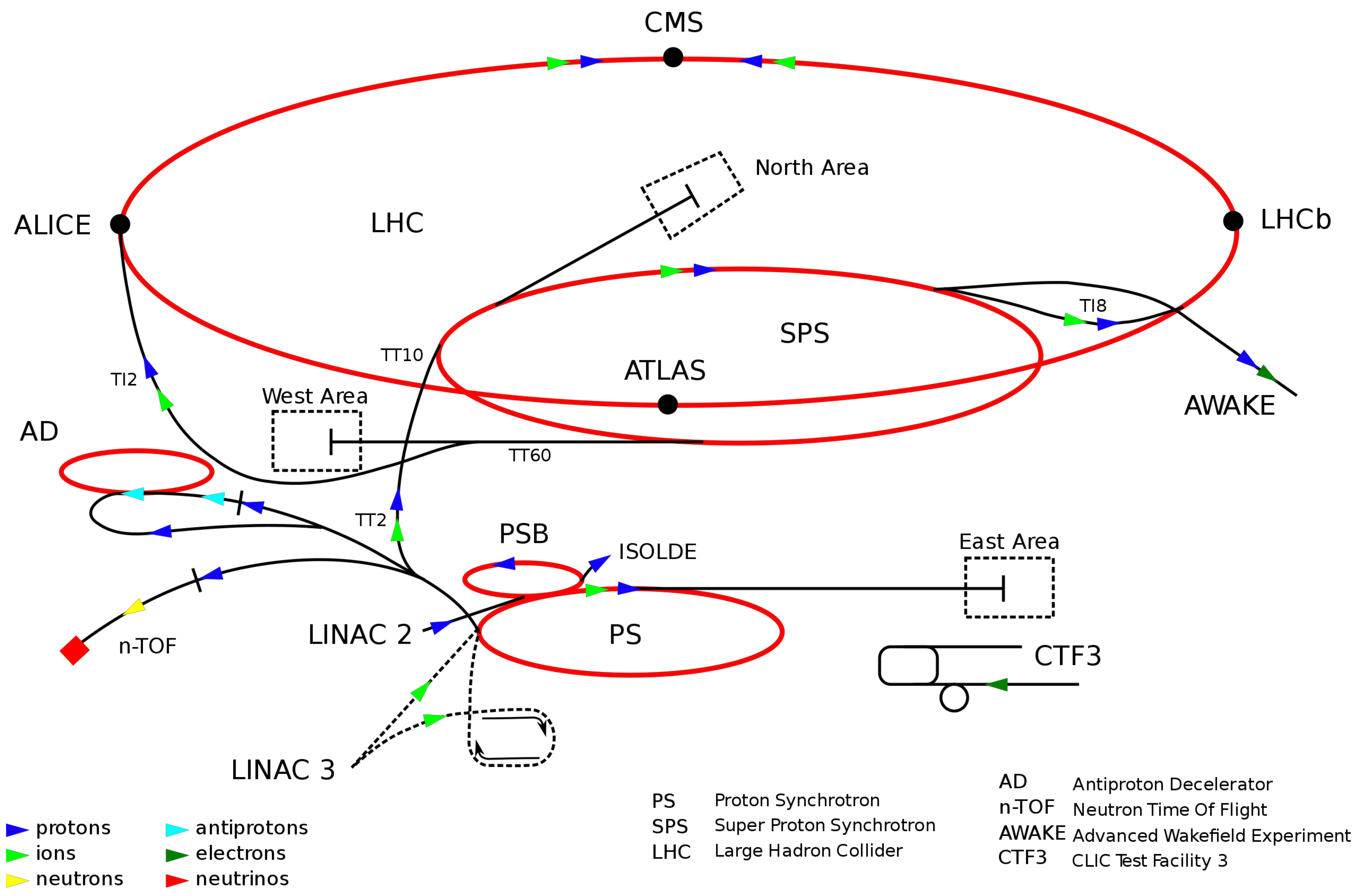
Timeline

Now - Early 2020	Finalize background estimation, selection methods, systematics for 4b. Start writing paper. Test new FastCaloSim for physics, finish implementation.
Early 2020	Resonant 4b result published
Summer 2020	Non-resonant 4b result published
Fall 2020 – Spring 2021	Work on 4b reinterpretation paper ($X \rightarrow SH$, e.g.)
Spring 2021 – Summer 2021	Write thesis, finish up reinterpretation if need be
Summer 2021 – Fall 2021	Graduate

Thanks!

Backup

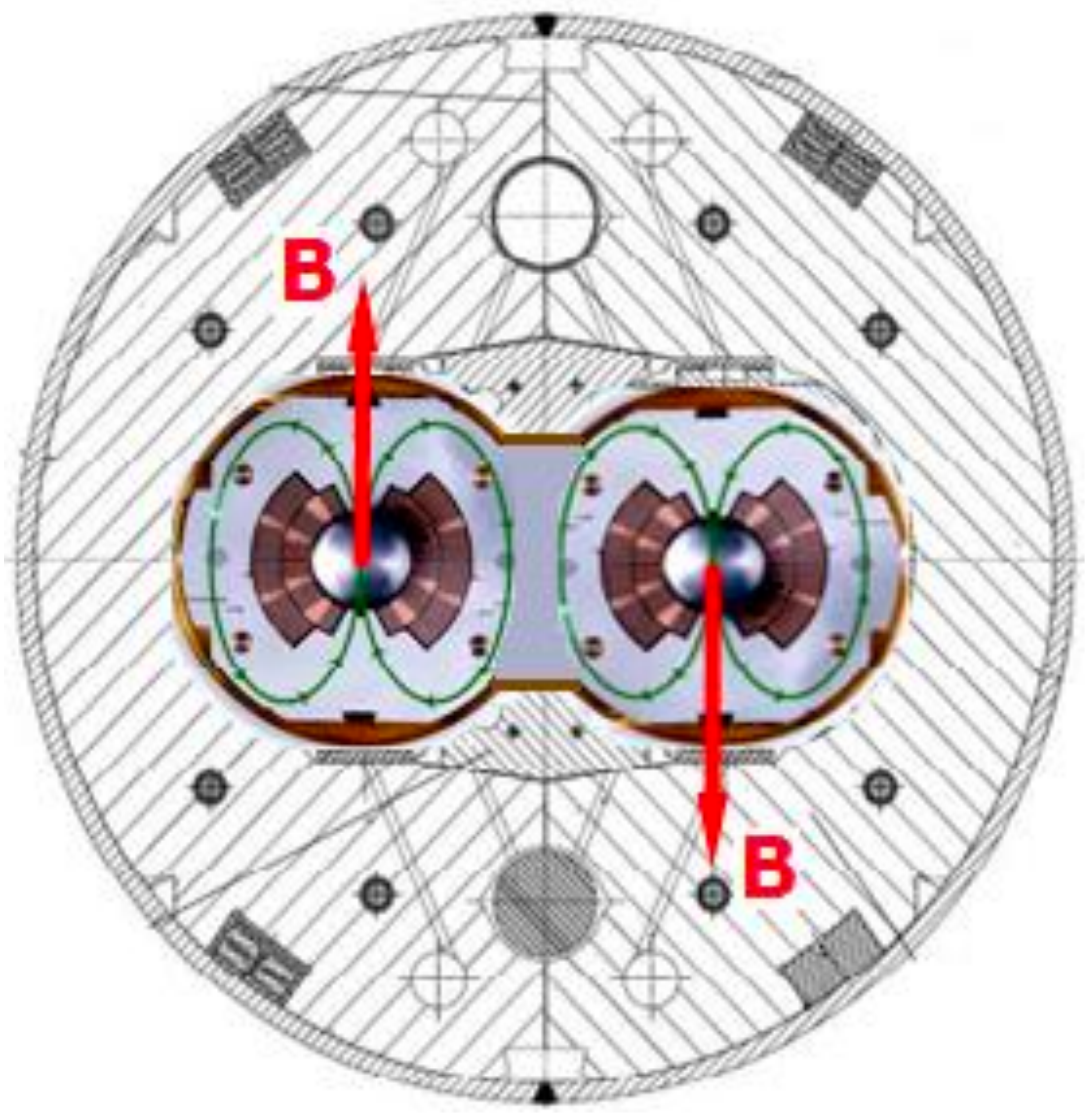
LHC: Ring and Acceleration



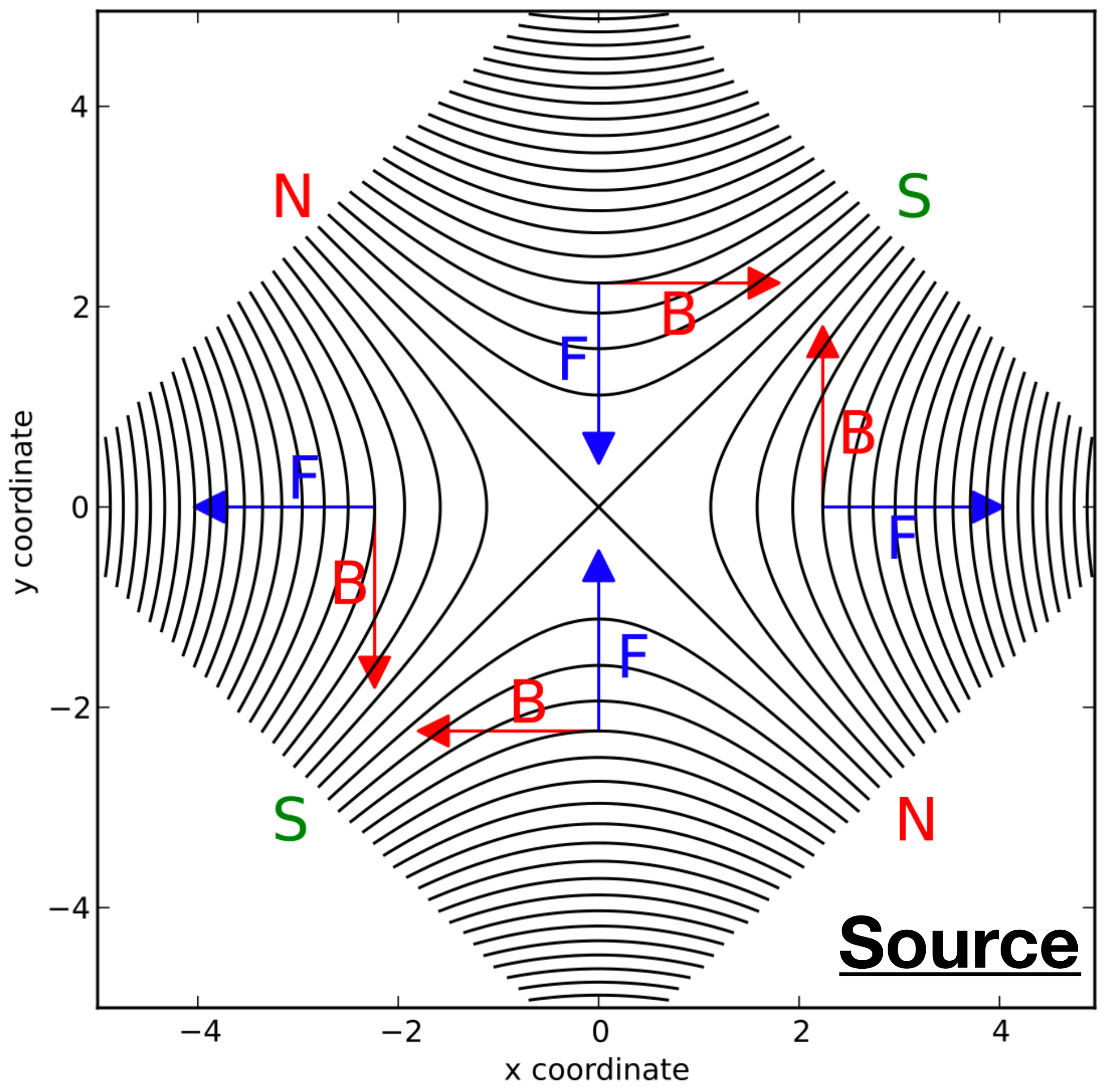
Source, acceleration info

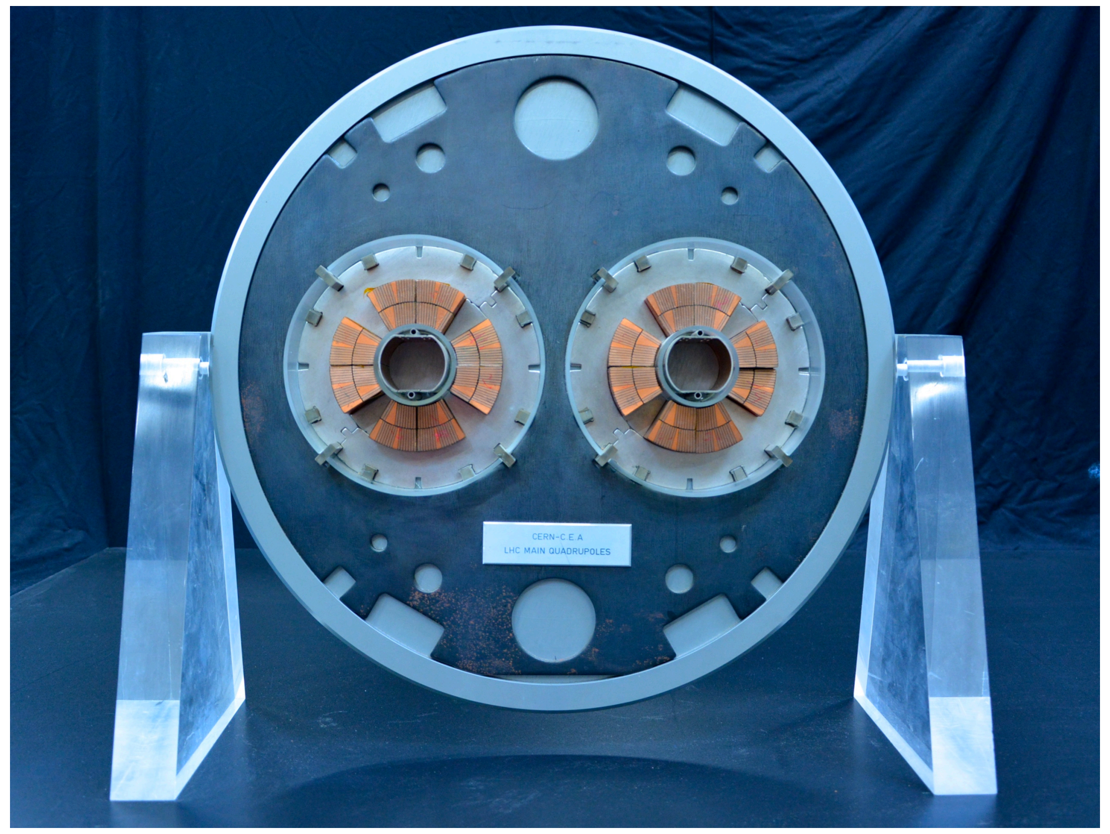
- The LHC is a 27 km circumference ring. There are 4 collision points along the ring for the 4 particle detectors (ATLAS, CMS, LHCb, ALICE)
- Protons move in two counter-circulating beams
- Acceleration chain:
  - Electric field strips hydrogen of its electrons
  - LINAC 2 accelerates protons to 50 MeV
  - Beam is injected into the Proton Synchrotron Booster (PSB), which pushes them to 1.4 GeV, and then the Proton Synchrotron (PS), bringing the beam to 25 GeV
  - From there, protons are transferred to the Super Proton Synchrotron (SPS) => 450 GeV, and then finally to the LHC => 6.5 TeV

LHC: Magnets



Dipole magnets provide a centripetal force to keep protons in the ring [ref]

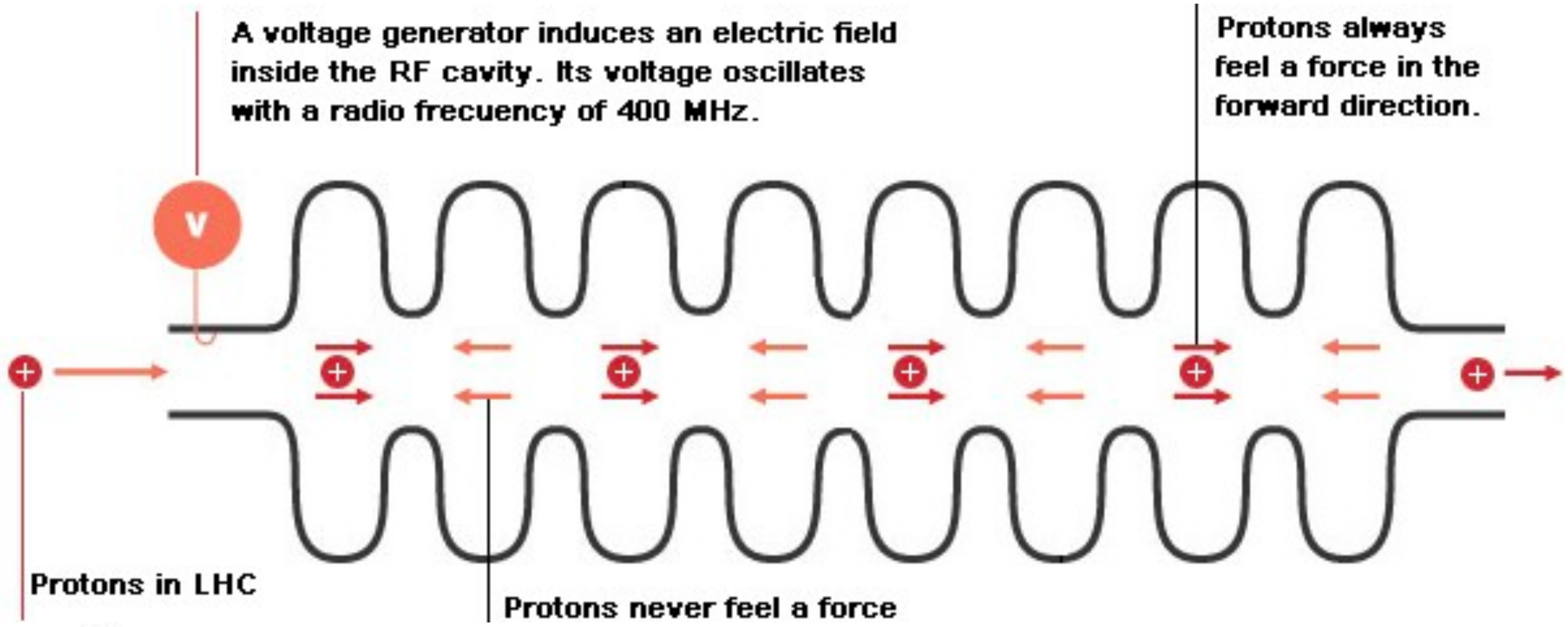


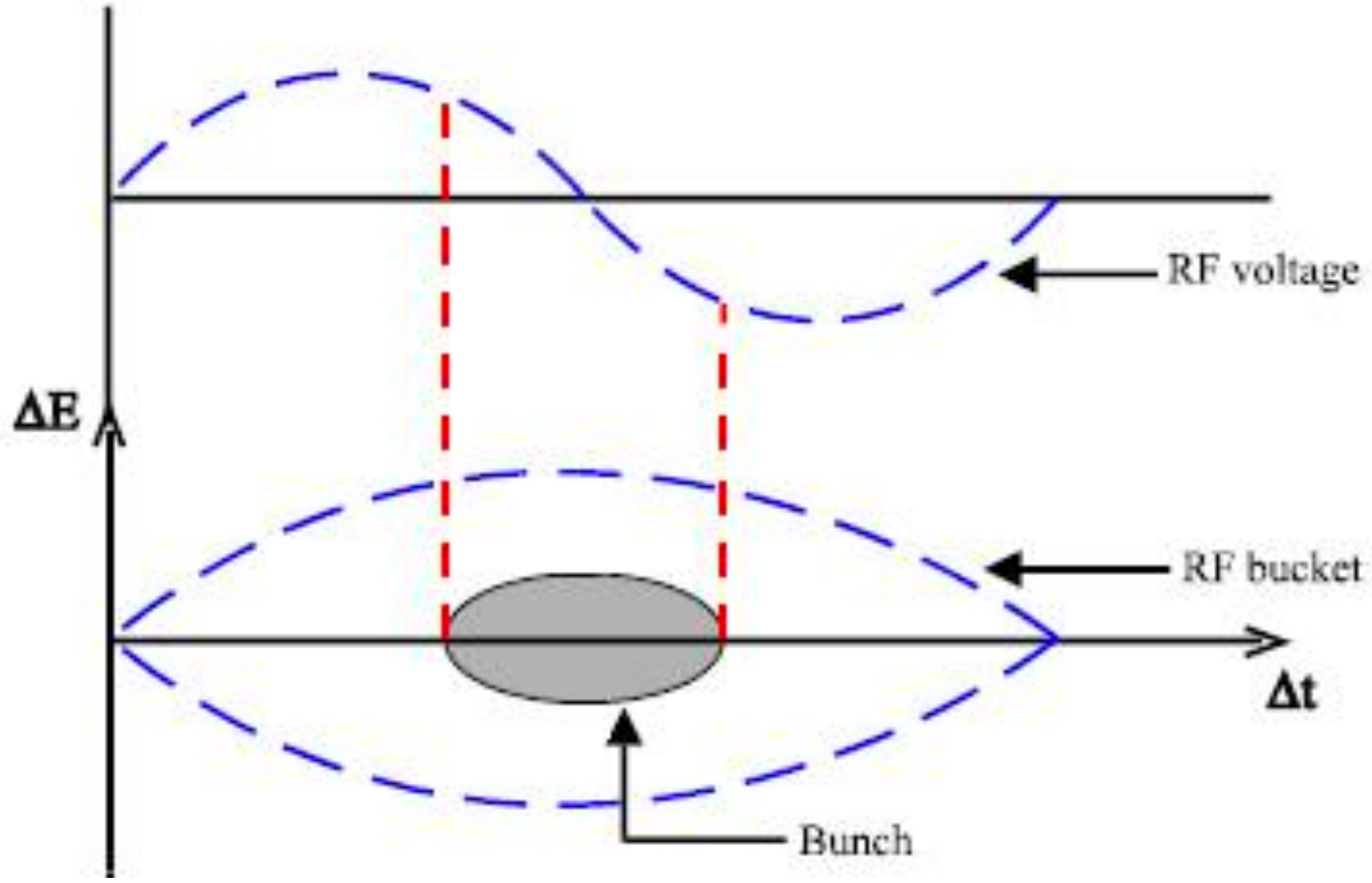


Quadrupole magnets focus (alternate in which direction)

- Superconducting magnets (cooled by liquid helium) are used to keep the protons in the LHC
  - Dipole magnets are used to bend the paths of particles around the ring
  - Quadrupoles are used to keep the particles in a tight beam
  - Higher order magnets are used to correct for small imperfections

LHC: Acceleration



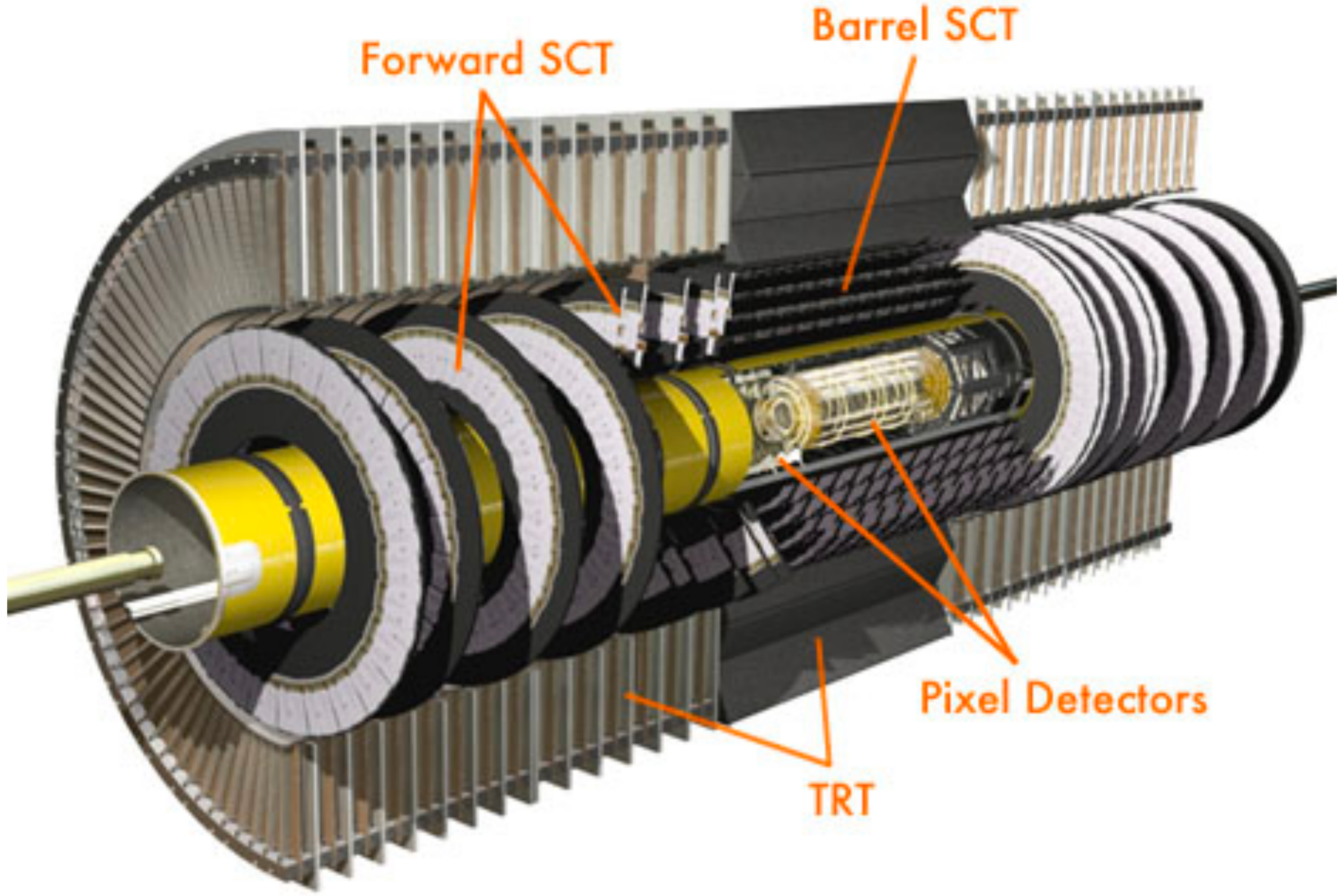


Source, buckets

Buckets are RF groupings. Bunches are groups of protons, filling a bucket. 25 ns bunch spacing is used at the LHC

- Magnetic fields do no work! We need to use electric fields to accelerate the protons
- Most of this job is done by Radiofrequency (RF) cavities
- In these, an electromagnetic field is made to oscillate (switch direction) at a precise rate
- Timing is important so that protons always see an accelerating voltage
- However there's some self correction! There is some finite spread in grouping of particles. If a particle arrives too early, e.g., it will see some decelerating voltage, too late, it will see a higher accelerating voltage

ATLAS Detector: Inner Detector

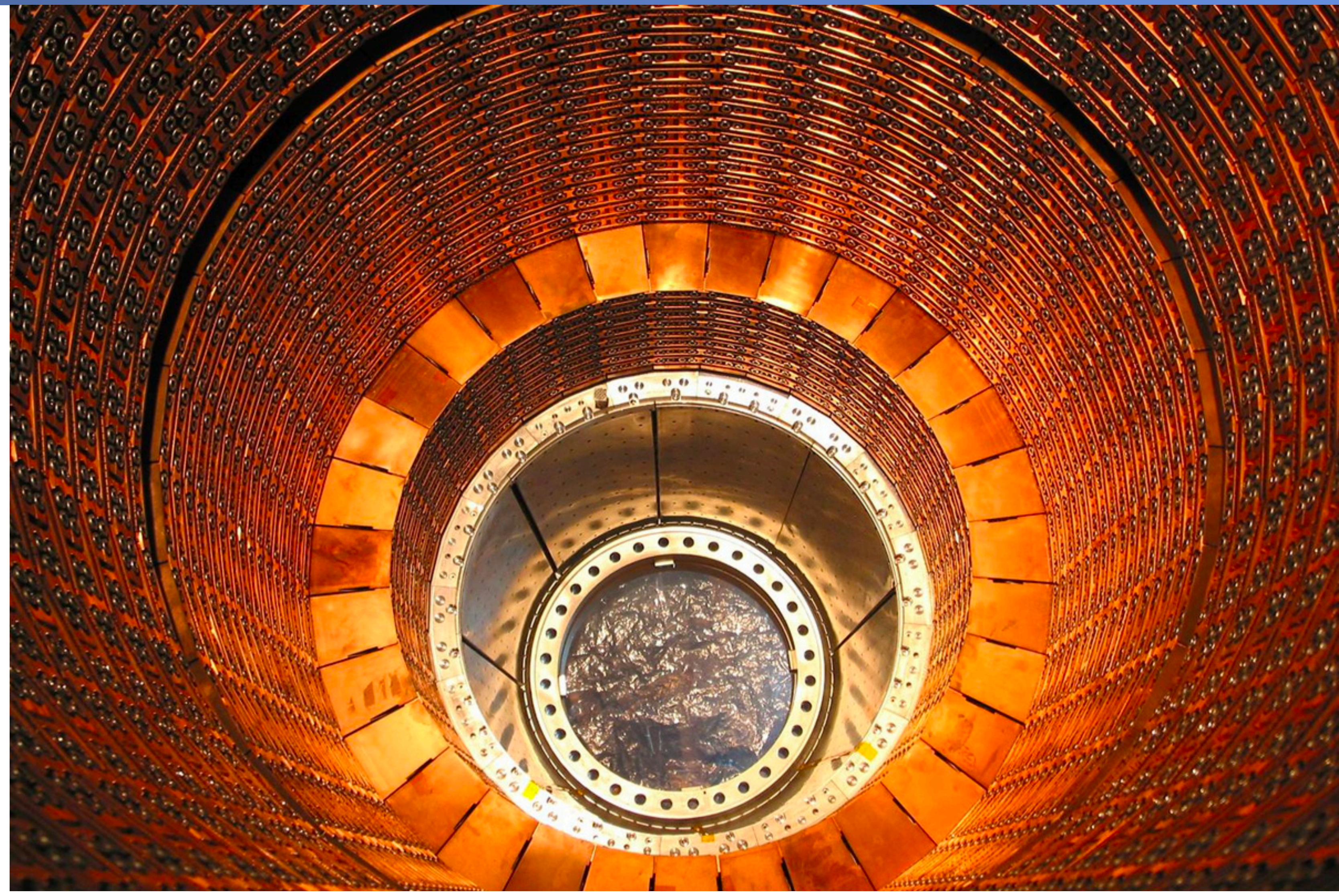


- The Inner Detector contains the tracking elements for ATLAS
  - Everything is inside of a 2 T solenoid magnet, providing information about momentum from curvature (in x, y direction, z is along the beamline)
- Silicon tracking: Silicon is ionized by a charged particle (knocks out electrons), charge is collected
- **Pixel detector:** High granularity, very close to interaction point, high degree of positional information - important for b-tagging!
- **Semiconductor Tracker (SCT):** Lower granularity, but similar in concept to Pixel
- **Transition Radiation Tracker (TRT):** Drift tubes and materials with widely varying indices of refraction
  - Drift tube: straw filled with (Xenon) gas with a wire in the center, which collects electrons displaced by an ionizing particle
  - Varying index of refraction => transition radiation. The amount given off depends on particle speed, lower mass => higher speed for a given energy => helps with particle identification

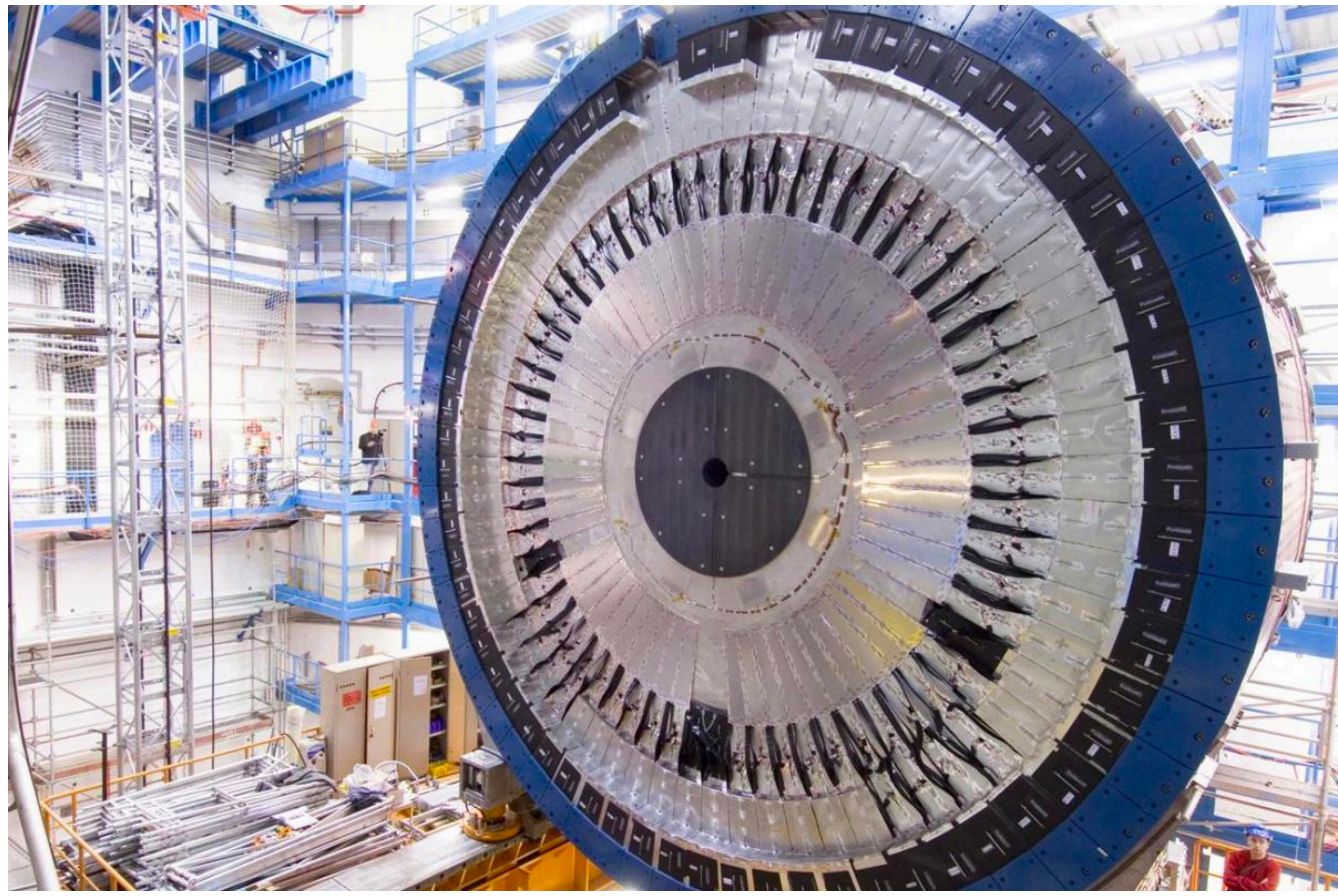
ATLAS Detector: Calorimeters

- Concept: absorb the energy of a particle in order to measure it. Have sensitive material to look at the shape of a particle shower.
- **EM Calorimeter**: Lead absorbers with liquid argon in between
  - EM showers are usually shorter, more compact
    - Electron mean free path, e.g., is short in lead=> usually can be contained in the EM calorimeter
  - Closer to the interaction point, high granularity
  - Charged particles from EM decay ionize liquid argon to give shape
- **Hadronic Calorimeter**: Steel absorbers with plastic scintillating tiles in between
  - Hadrons rely on nuclear (strong) interactions - mean free path is longer => make it through the EM calorimeter
  - Less precise than EM. Scintillator produces light with ionizing particle, that is then collected

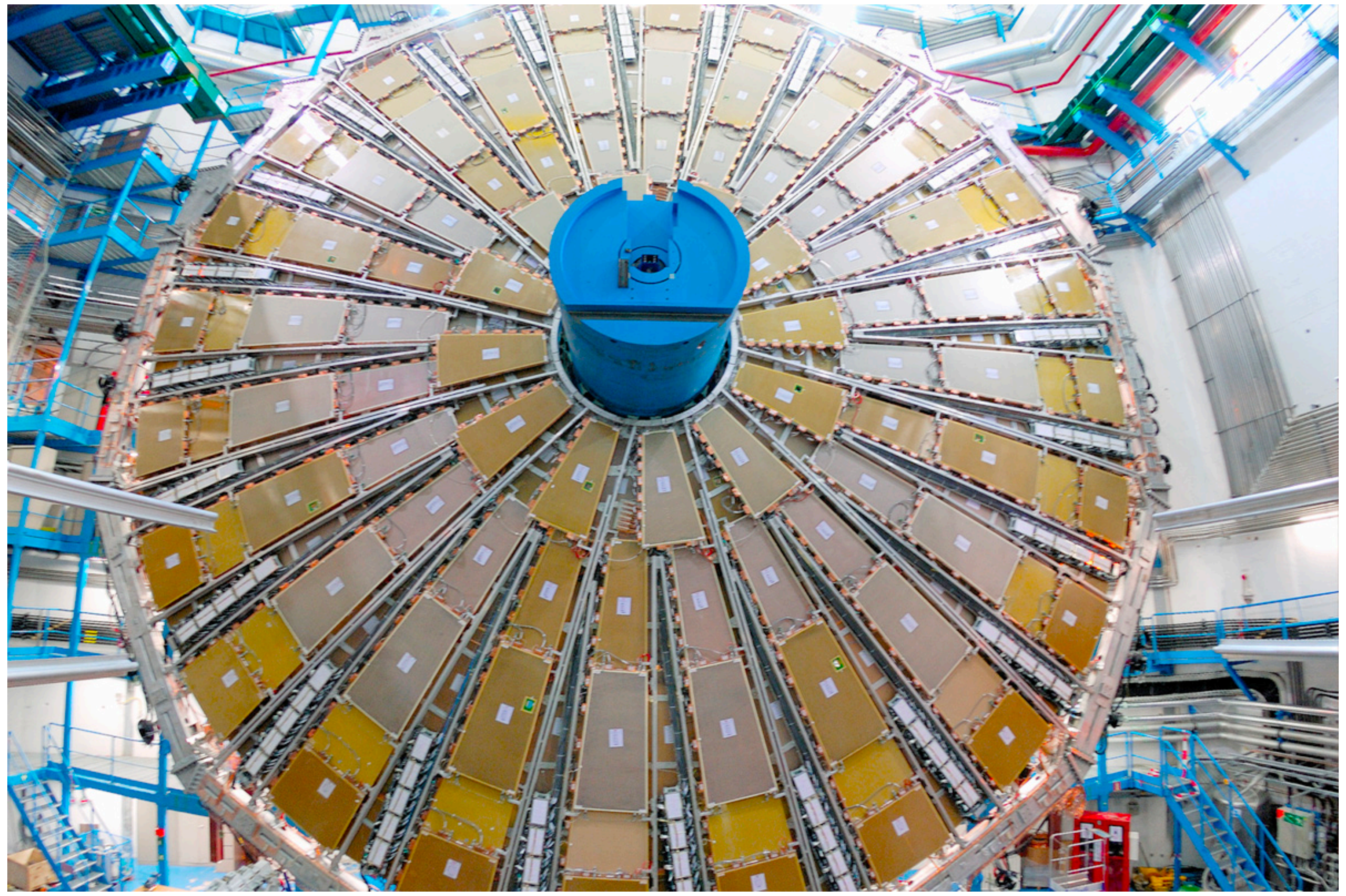
Electromagnetic Calorimeter (Liquid Argon):



Hadronic Calorimeter (Tile):



ATLAS Detector: Muon Spectrometer

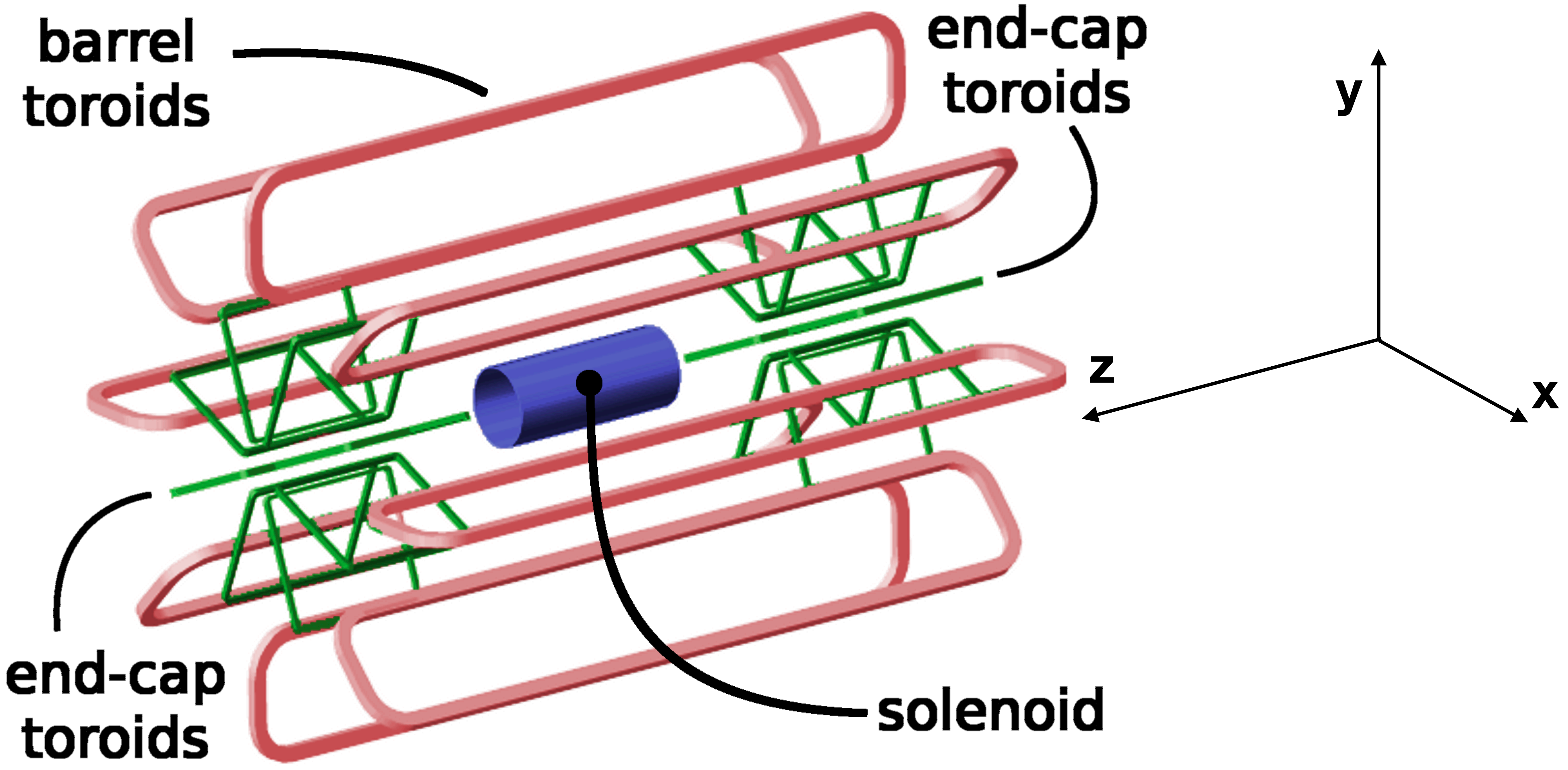


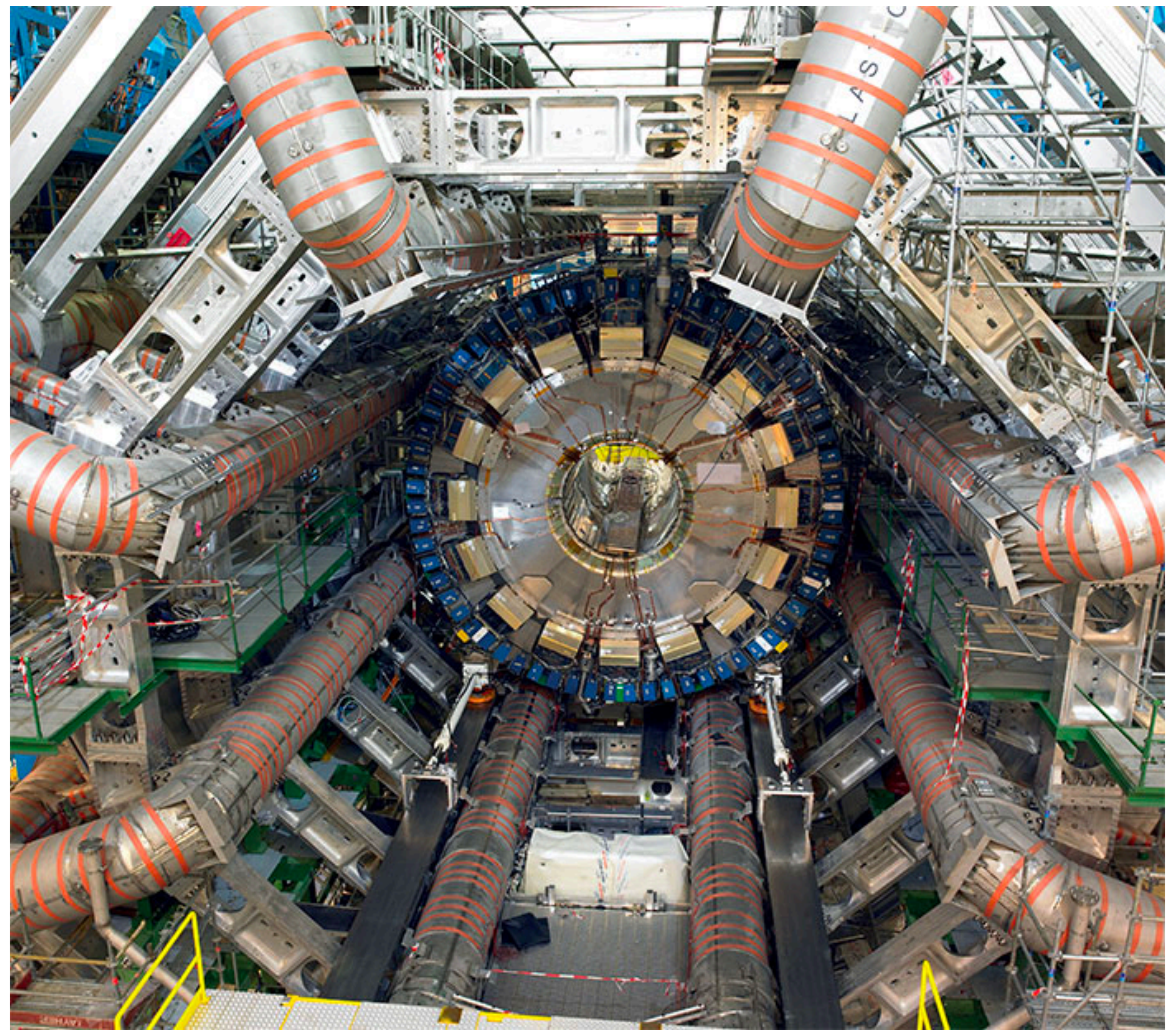
m_μ = **106 MeV**

m_e = **0.510 MeV**

- Outermost layer of ATLAS
- Muons are heavy and don't interact strongly
  - 200 x heavier than electrons => are not stopped by EM interactions with, e.g., electrons in the absorbers of the calorimeter
  - Don't interact strongly, so hard scattering with nuclei is rare
- Three parts: **Triggering chambers** (detect if muon, non-bending direction coordinate measurement), **drift tubes** (tracking system - measure path/curve of muons), and **toroid magnets** (provide the magnetic field for curve/momentum measurement. Similar concept to solenoid, different configuration)

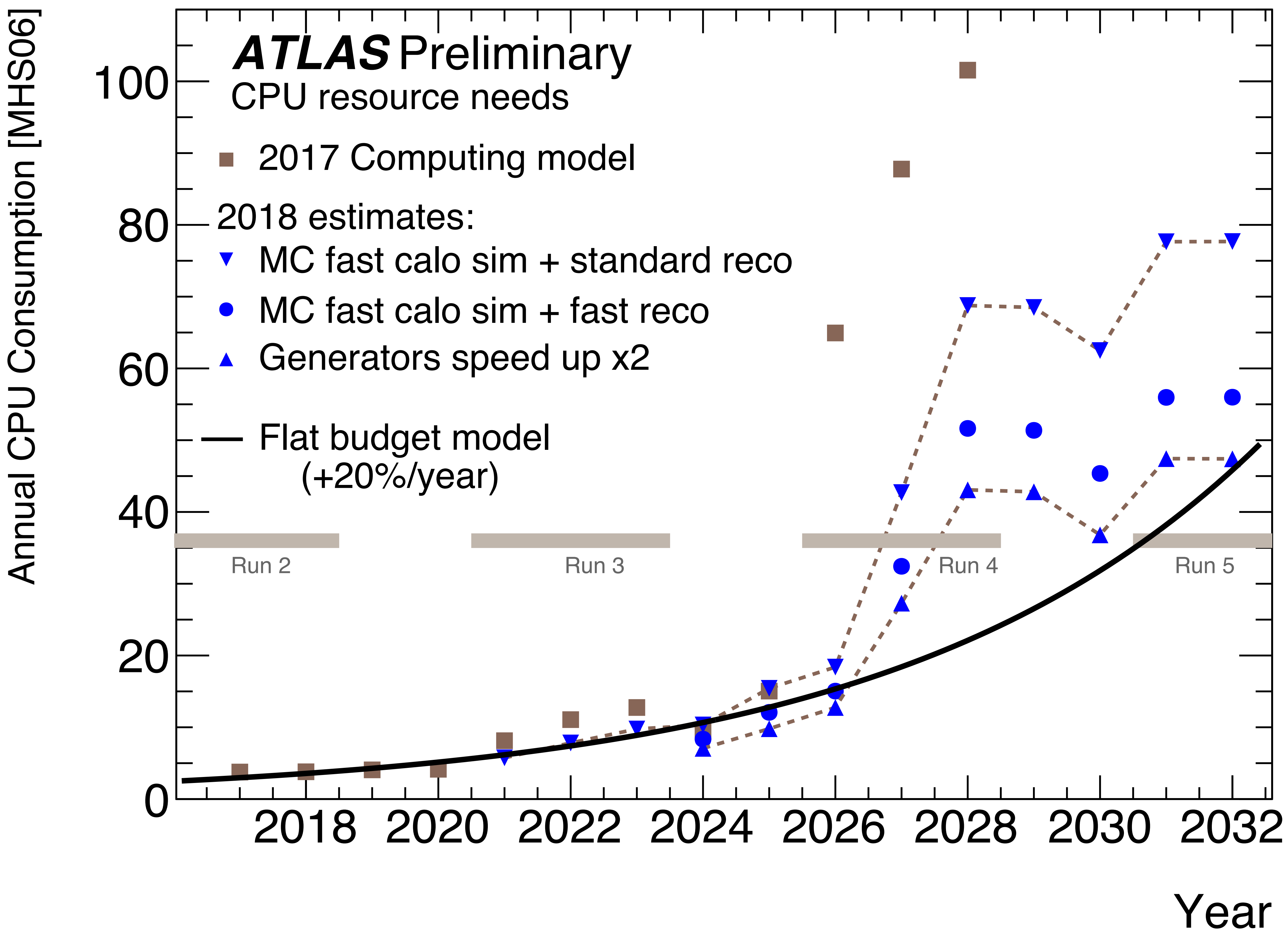
ATLAS Detector: More on Magnets





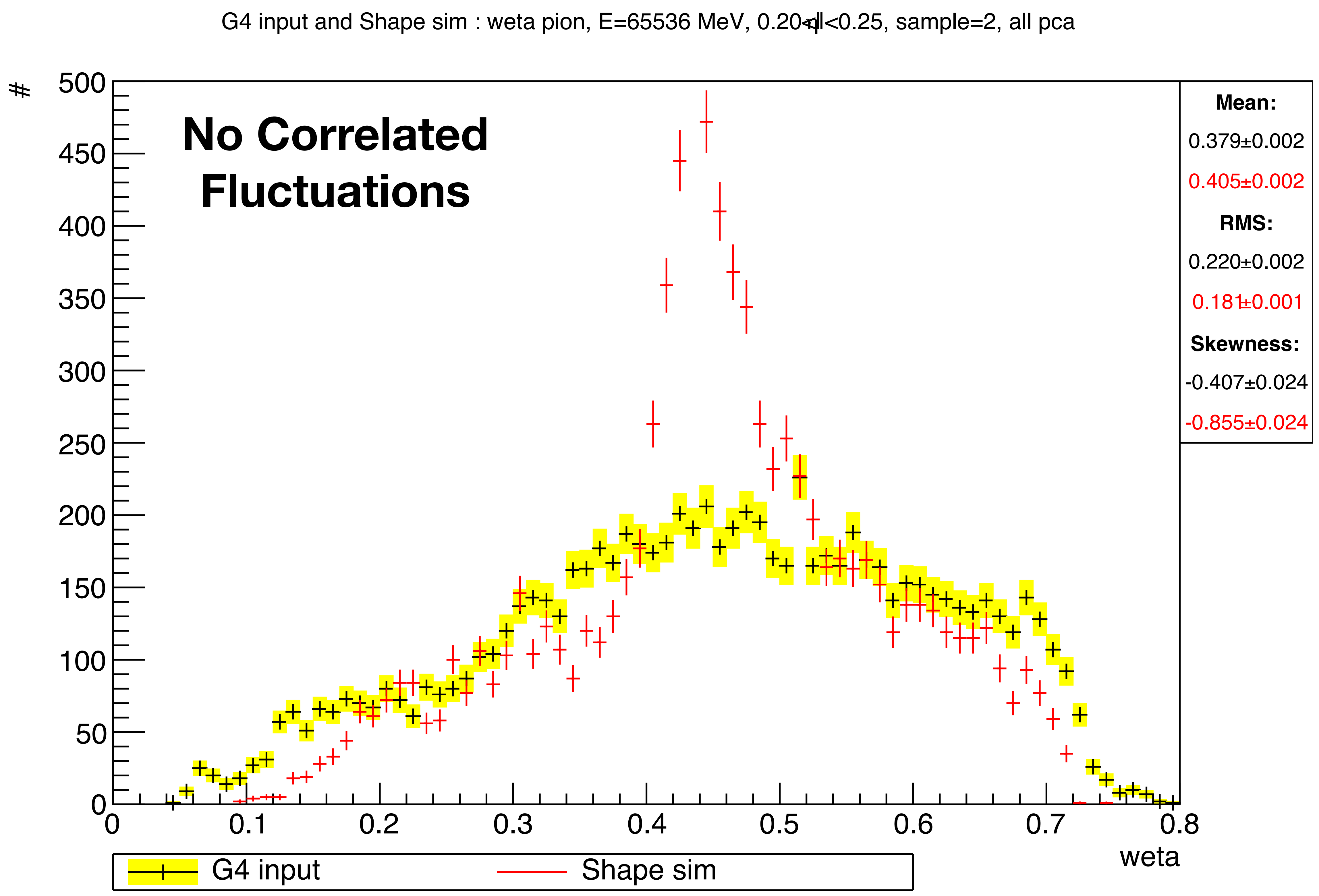
- Particles come from the interaction point ~ radially
- Solenoid: Magnetic field along z-axis
  - => particles bent in x-y plane (e.g. s6)
- Toroid: magnetic field ~circle in x-y plane around beam line
  - => particles bent along the z-axis

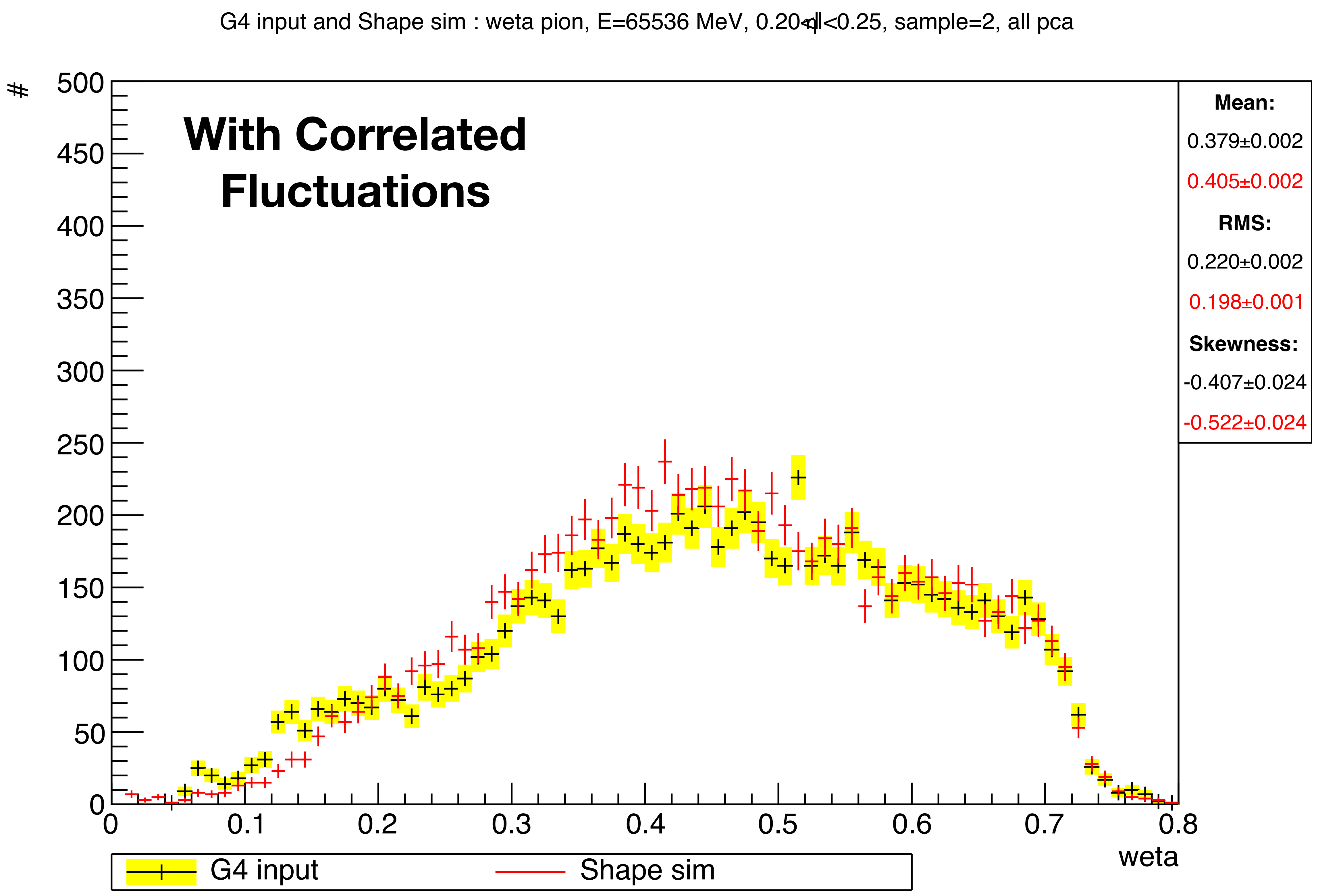
Need for Fast Simulation



- CPU consumption for ATLAS. Current model is unsustainable, fast simulation improves the situation substantially

Layer 2 (EMB2): Weta

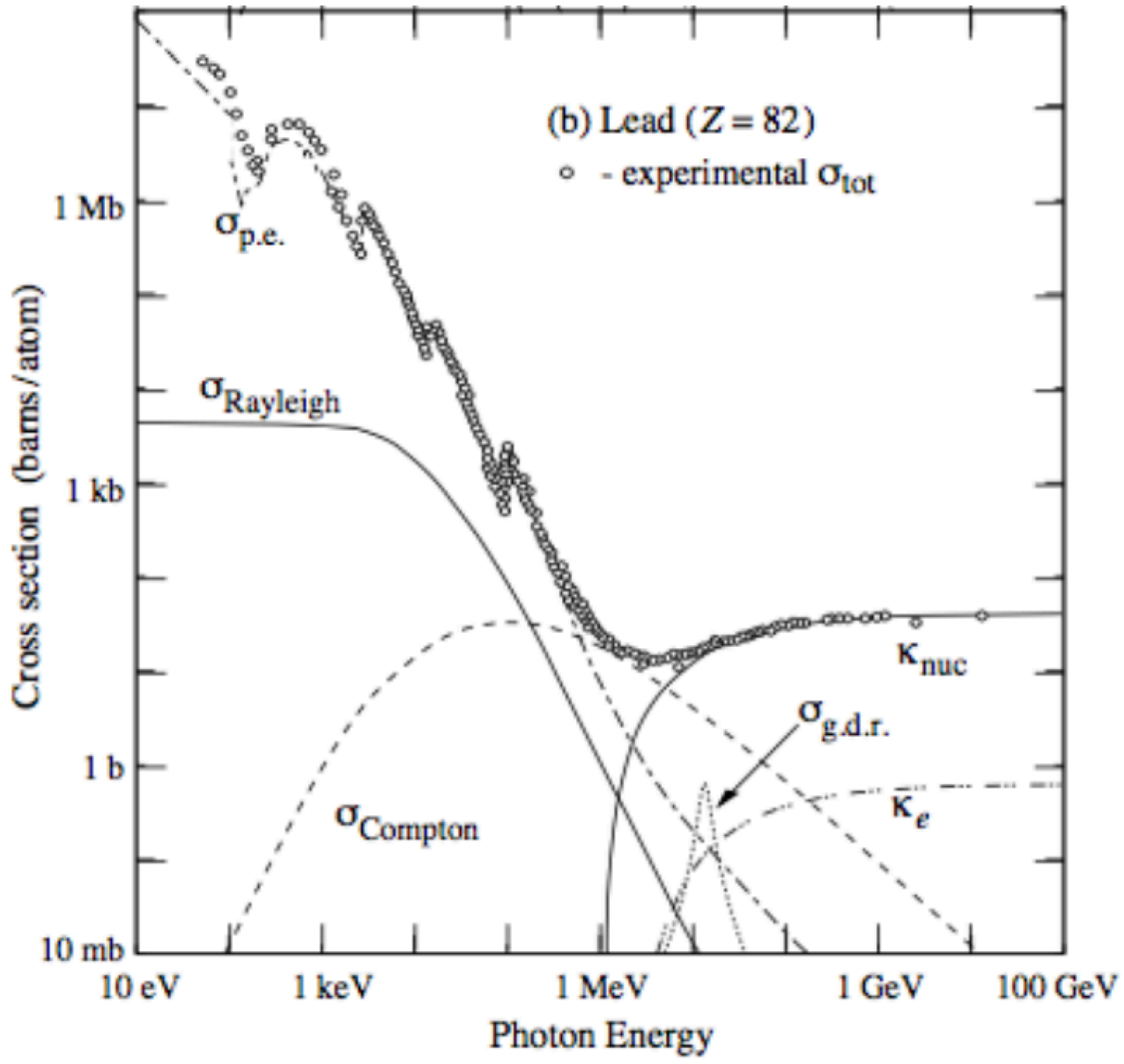




- Large peak present in shape sim without correlated fluctuations. This disappears with them - much better agreement!

EM Showering

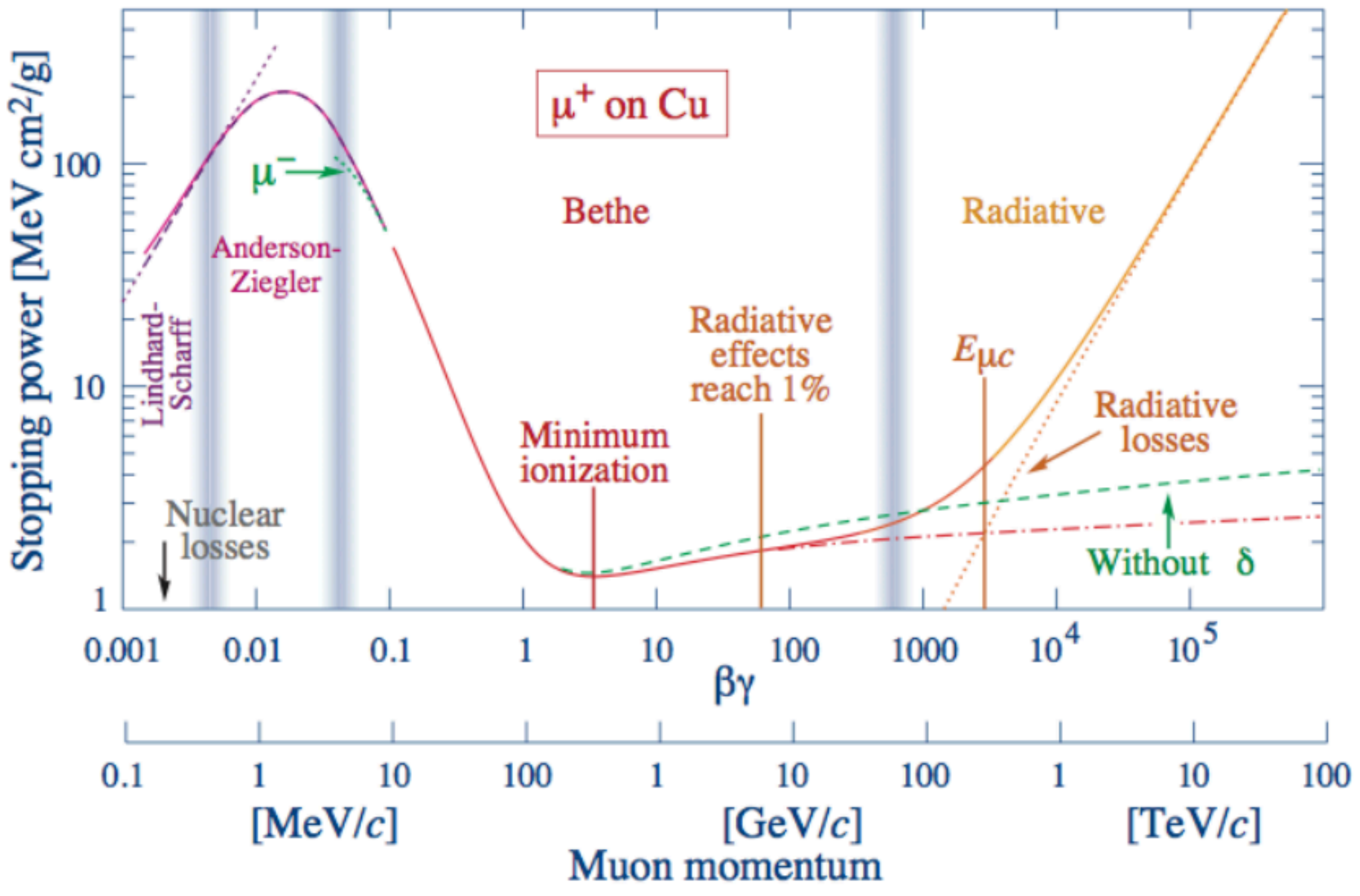
- EM showers: mostly due to
  - Bremsstrahlung (electron): high energy electron emits a photon with some (potentially significant) amount of energy
    - Due to scatter off of field of heavy nucleus
  - and pair production (photon)
    - Energy split in production of electron-positron pair
  - Cross section for photon to scatter is quite large for lead1



Source

Muons and EM Showers

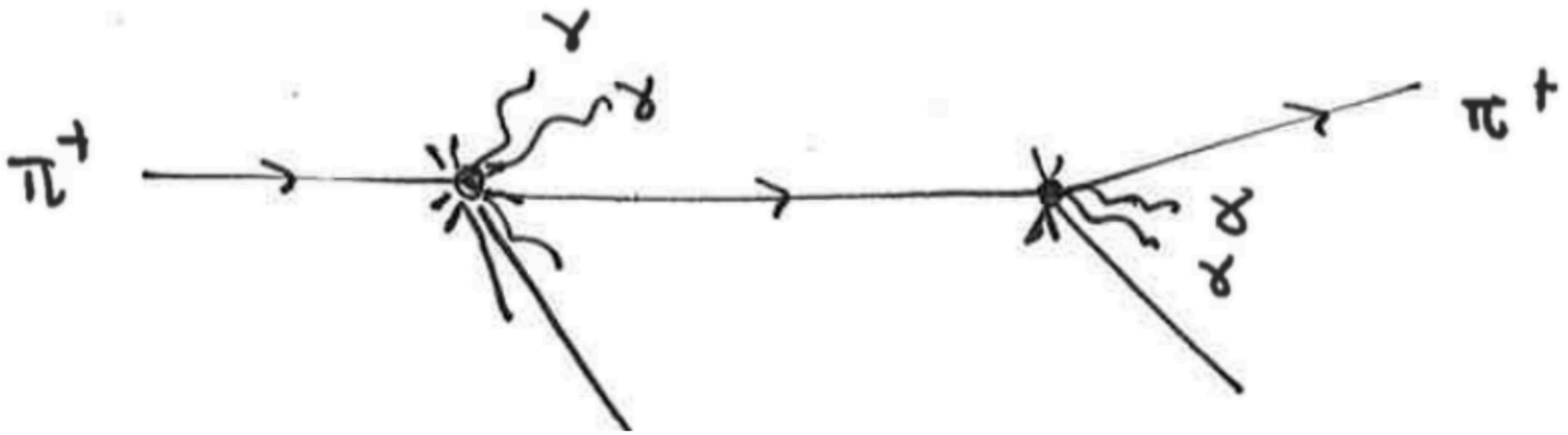
- For heavier particles (muons) radiative losses don't contribute as much
  - See in copper on right - radiative losses don't contribute until high energy
  - Coulomb scattering from a nucleus can change particle direction



Source

Hadronic Showers

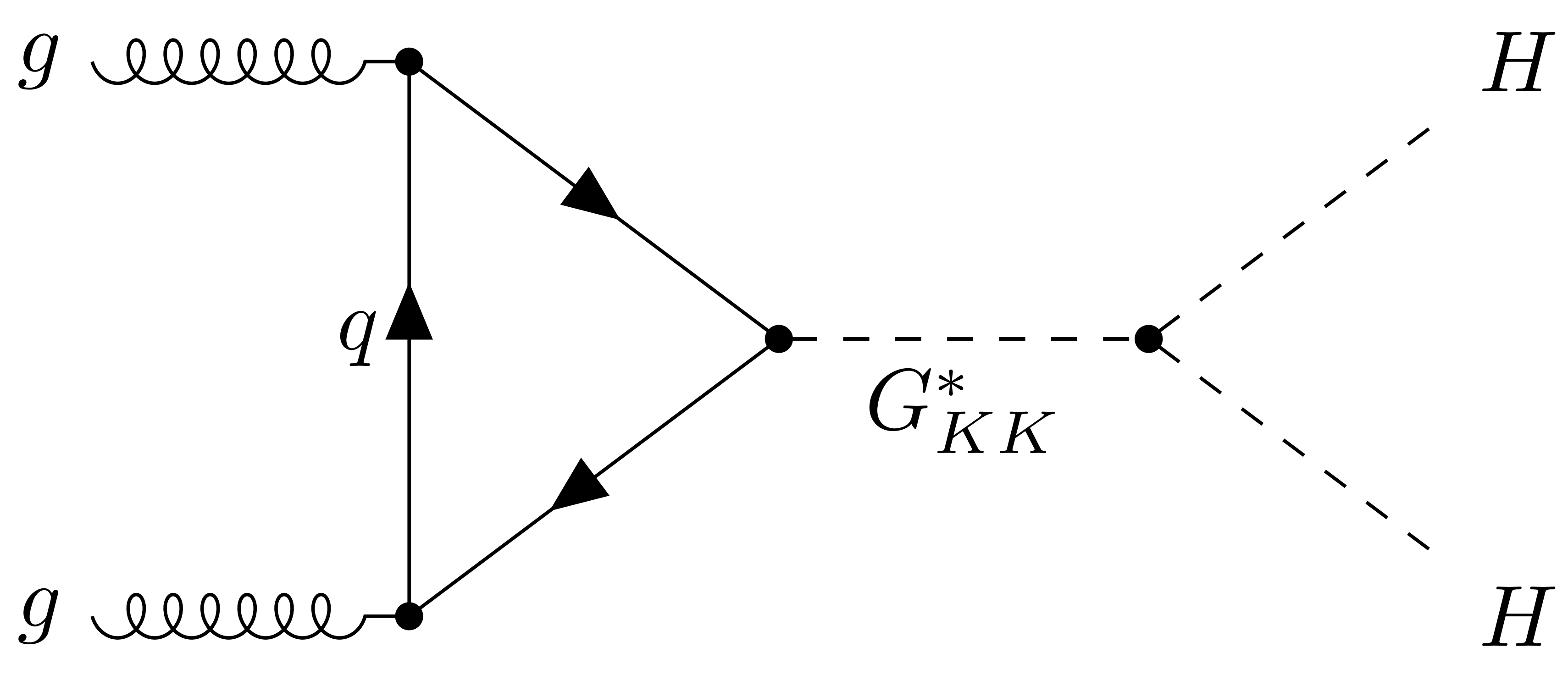
- Hadrons can interact strongly with atomic nuclei
- More complex because they involve a wider variety of processes with different length scales
- Consider, e.g., π^+:
  - Scatter creates, e.g., π^0, π^+, π^-
  - π^0 decays very quickly to two photons, which then shower
  - π^+, π^- are longer lived, and may continue for some distance (interaction length)



Source

Resonant Searches: Spin 2

- Spin 2 Randall-Sundrum Kaluza Klein Graviton
  - Graviton arising from a warped extra dimension
  - Decay width to HH is a function of $c = k/\overline{M}_{Pl}$
    - k is the curvature of the warped extra dimension
    - $\overline{M}_{Pl} = 2.4 \times 10^{18}$ GeV is the reduced Planck mass
  - Consider here c=1.0, 2.0
  - Model beginning to be disfavored by the community - much of the parameter space is already excluded



SM Higgs Mechanism

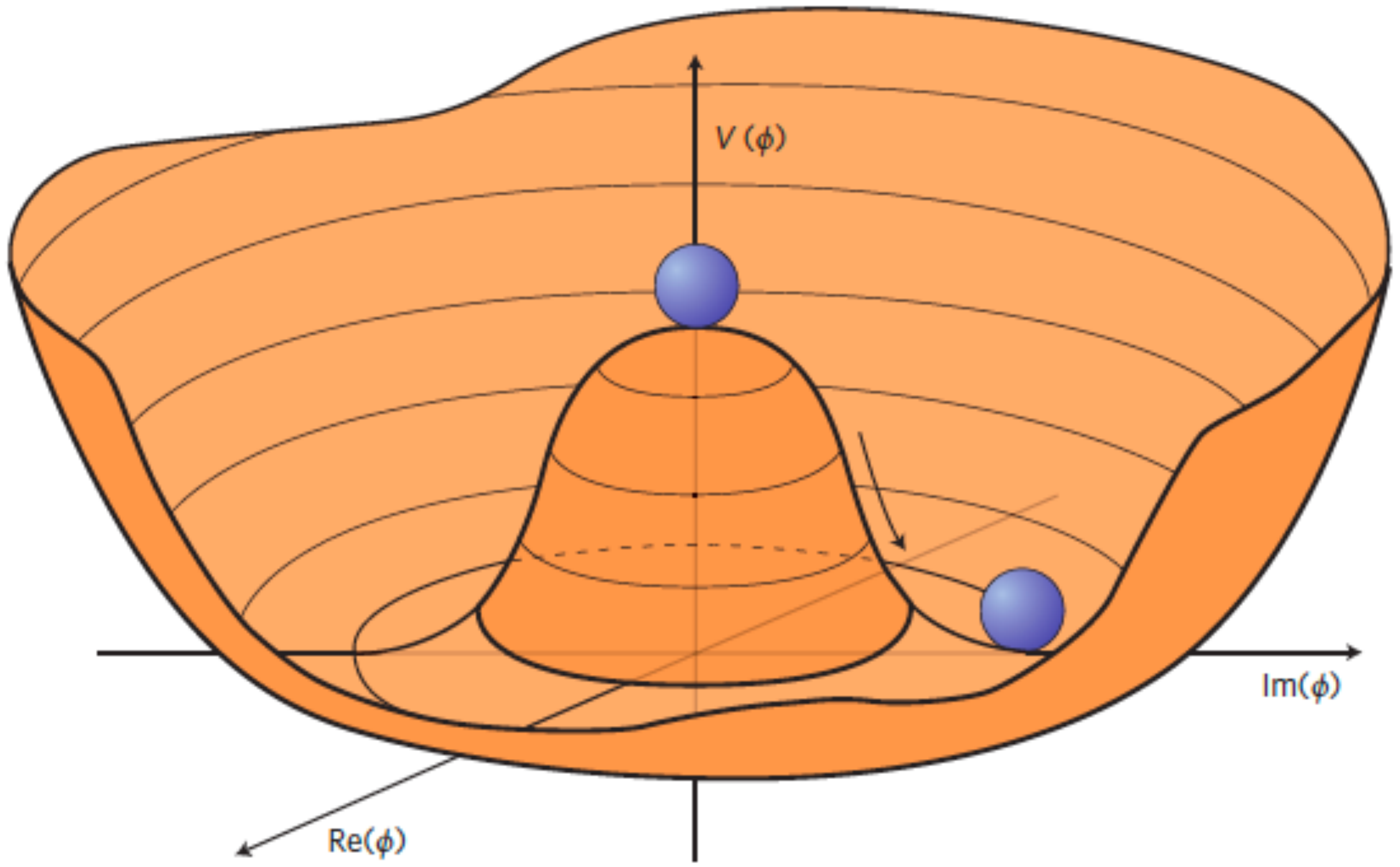


Illustration of the Higgs potential in a U(1) model. Minima are a circle at the bottom of the hat. Choosing one point spontaneously breaks the symmetry [ref]

Higgs Potential: $V(\phi) = -\mu^2\phi^2 + \lambda\phi^4$

Minimum satisfies $0 = -2\mu^2\phi + 4\lambda\phi|\phi|^2$

$$\implies |\phi|^2 = \frac{\mu^2}{2\lambda}$$

Define $v = \sqrt{2}\langle|\phi|\rangle = \frac{\mu}{2\lambda}$

spontaneously breaks SU(2) x U(1) symmetry.

Minima form a sphere in 4d space (Higgs doublet => 4 degrees of freedom). All minima equivalent by SU(2) rotation.

Look at vacuum state with $\langle\phi\rangle = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix}$

SU(2) x U(1) gives 4 vector bosons, A^1, A^2, A^3, B

Expanding around this gives $\phi(x) = \begin{pmatrix} \pi^+(x) \\ (v + h(x) + i\pi^3(x))/\sqrt{2} \end{pmatrix}$

Source

SM Higgs Mechanism

For $\phi(x) = \begin{pmatrix} \pi^+(x) \\ (v + h(x) + i\pi^3(x))/\sqrt{2} \end{pmatrix}$**,** $\pi^+ = (\pi^1 + i\pi^2)/\sqrt{2}$ **and** π^3 **are Goldstone bosons. We can set these to 0 by an SU(2) gauge transformation, leaving a real valued scalar field** $h(x)$**, the field of the Higgs boson**

Covariant derivative on Fermions: $D_\mu \Psi = (\partial_\mu - igA_\mu^a I^a - ig' B_\mu Y)\Psi$

Apply to the Higgs field. Get terms: $(\frac{gv}{2})^2 \, W_\mu^+ W^{\mu -} = m_W^2 W_\mu^+ W^{\mu -}$ $\quad W^\pm = \frac{1}{\sqrt{2}}(A_\mu^1 \mp i A_\mu^2)$

$$c_w = \cos\theta_w = \frac{g}{\sqrt{g^2 + g'^2}} \, , \qquad s_w = \sin\theta_w = \frac{g'}{\sqrt{g^2 + g'^2}}$$

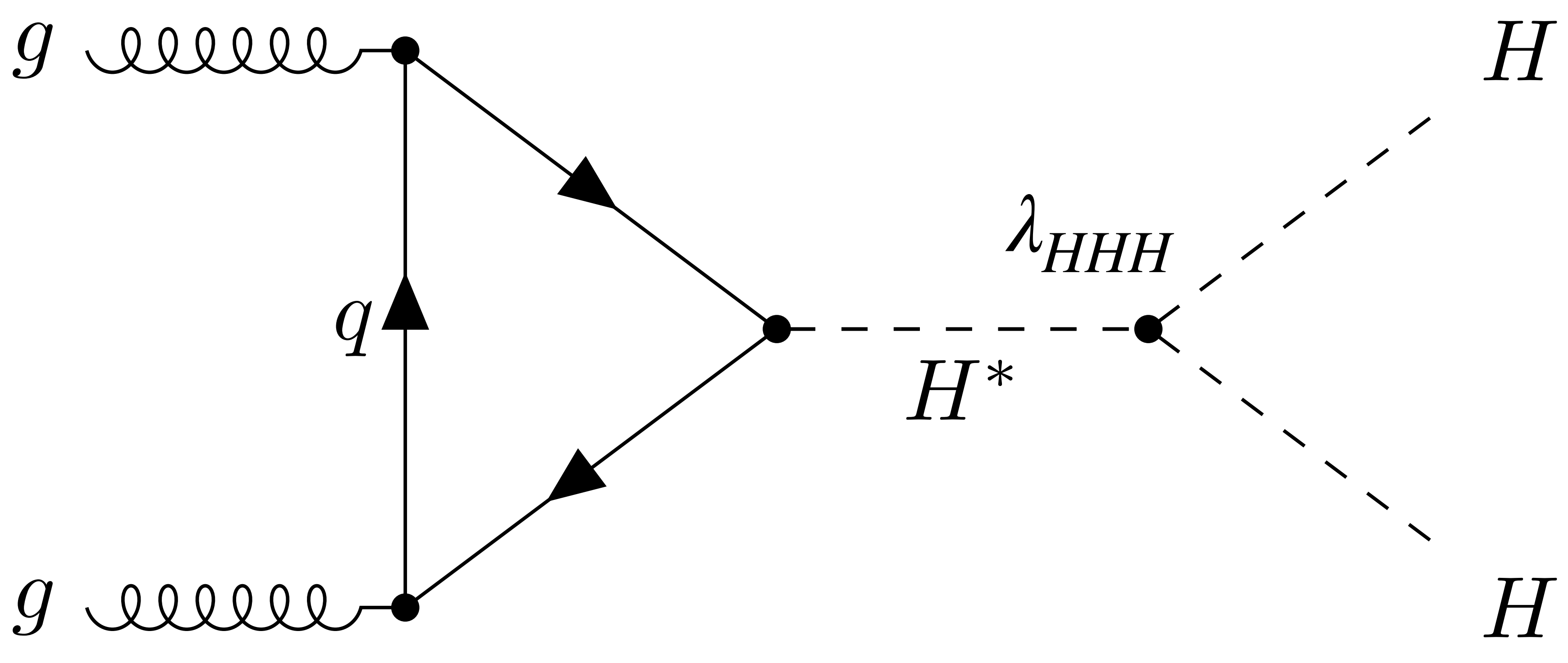
$$Z_\mu = c_w A_\mu^3 - s_w B_\mu$$
$$A_\mu = s_w A_\mu^3 + c_w B_\mu$$

$$m_Z^2 = \frac{(g^2 + g'^2)v^2}{4}$$

A_μ **remains massless - identify with the photon**

Source

Non-resonant HH



$$m_H^2 = 2\lambda v^2 \approx 125 \text{ GeV}$$

$$v = \frac{\mu}{\sqrt{\lambda}} \approx 246 \text{ GeV}$$

$$V(\phi) = -\mu^2\phi^2 + \lambda\phi^4$$

This is the most general, renormalizable potential for a single Higgs doublet consistent with electroweak symmetry (SU(2) x U(1))

[V] = 4
[ϕ] = 1
[μ] = 1
[λ] = 0
A ϕ^5 term would have a coupling with negative mass dimension

Expand about minimum: $V(\phi) \to V(v+h)$

$$V = V_0 + \frac{1}{2}(-\mu^2 + 3\lambda v^2)h^2 + \lambda v h^3 + \frac{\lambda}{4}h^4$$

$$= V_0 + \lambda v^2 h^2 + \lambda v h^3 + \frac{\lambda}{4}h^4$$

$$= V_0 + \frac{1}{2}m_H h^2 + \frac{m_H^2}{2v}h^3 + \frac{m_H^2}{8v^2}h^4$$

Fermi effective theory/ SM:

$$m_W = \frac{gv}{2}$$

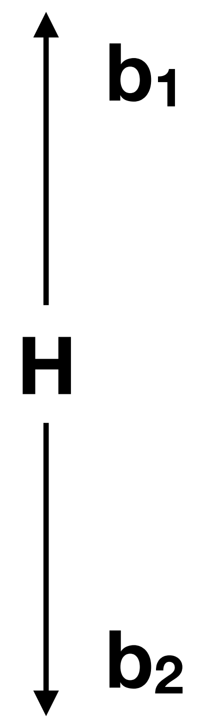
$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8m_W^2} \implies G_F = \frac{1}{\sqrt{2}v^2} \implies v = \frac{1}{\sqrt{\sqrt{2}G_F}}$$

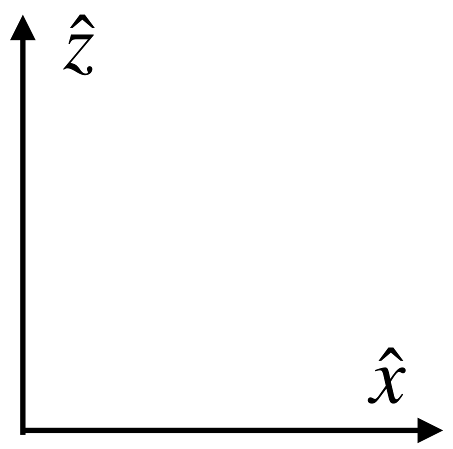
$$G_F \approx 1.166 \times 10^{-5} \text{ GeV}^{-2}$$

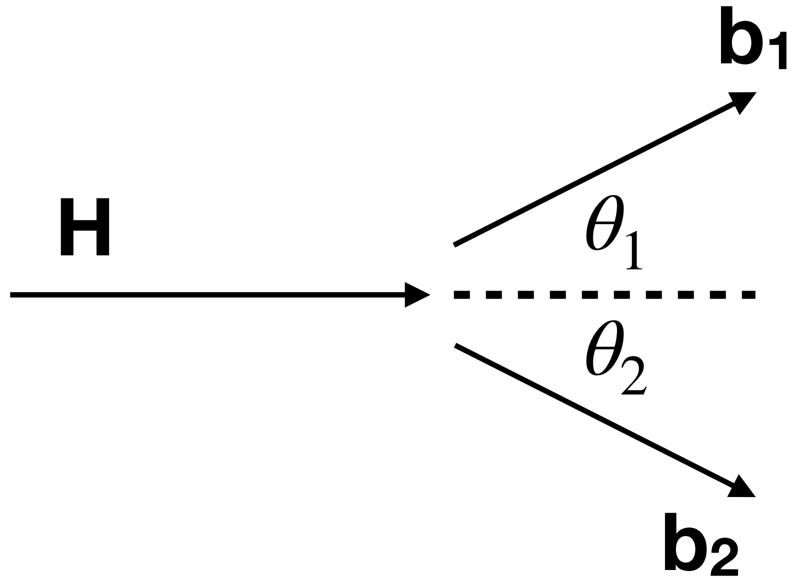
Best determination from muon lifetime

$$\Gamma_\mu = \frac{1}{\tau_\mu} = \frac{G_F^2 m_\mu^5}{192\pi^3}$$

1→2 ΔR Relation







$$\tan\theta_1 = \frac{p_{1_x}}{p_{1_z}} = \frac{\sqrt{m_H^2 - 4m_b^2}}{P}$$

$$\tan\theta_2 = \frac{p_{2_x}}{p_{2_z}} = \frac{\sqrt{m_H^2 - 4m_b^2}}{P}$$

Higgs Rest Frame

$$E_H^0 = m_H$$

$$\vec{p}_1^{\,0} = p_1^0\hat{z} = p_2^0\hat{z} = -\vec{p}_2^{\,0}$$

$$\vec{p}_1^{\,0} = -\vec{p}_2^{\,0} \implies (\vec{p}_1^{\,0})^2 = (\vec{p}_2^{\,0})^2 \equiv (\vec{p}_b^{\,0})^2$$

$$m_1 = m_2 = m_b$$

$$E_b^0 \equiv E_1^0 = E_2^0 = \sqrt{(\vec{p}_b^{\,0})^2 + m_b^2} = \frac{m_H}{2}$$

$$p_b^0 = \sqrt{E_b^0 - m_b^2} = \sqrt{\frac{m_H^2}{4} - m_b^2} = \frac{1}{2}\sqrt{m_H^2 - 4m_b^2}$$

Boost

$$\vec{P}_H = P\hat{x}$$

$$P = \gamma m_H \implies v = \frac{P}{E_H}$$

$$E_H = \gamma m_H \implies \gamma = \frac{E_H}{m_H}$$

$$p_{1_z} = p_{1_z}^0 = p_{2_z}^0 = p_{2_z} = \frac{1}{2}\sqrt{m_H^2 - 4m_b^2}$$

$$p_{1_x} = \gamma(p_{1_x}^0 + vE_1^0) = \gamma v\frac{m_H}{2} = \frac{P}{2}$$

$$p_{2_x} = \gamma(p_{2_x}^0 + vE_2^0) = \gamma v\frac{m_H}{2} = \frac{P}{2}$$

$$\theta_1 + \theta_2 = 2\arctan\left(\frac{\sqrt{m_H^2 - 4m_b^2}}{P}\right) \approx 2\frac{\sqrt{m_H^2 - 4m_b^2}}{P} \approx \frac{2m_H}{P}$$

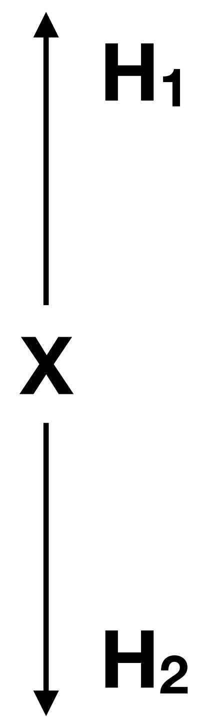
$$m_b \ll m_H$$

$$\theta_1 + \theta_2 \approx \Delta R$$

$$P \approx p_{T,H}$$

$$\implies \boxed{\Delta R_{bb} \approx \frac{2m_H}{p_{T,H}}}$$

1→2 ΔR Relation



$$E_X = m_X$$

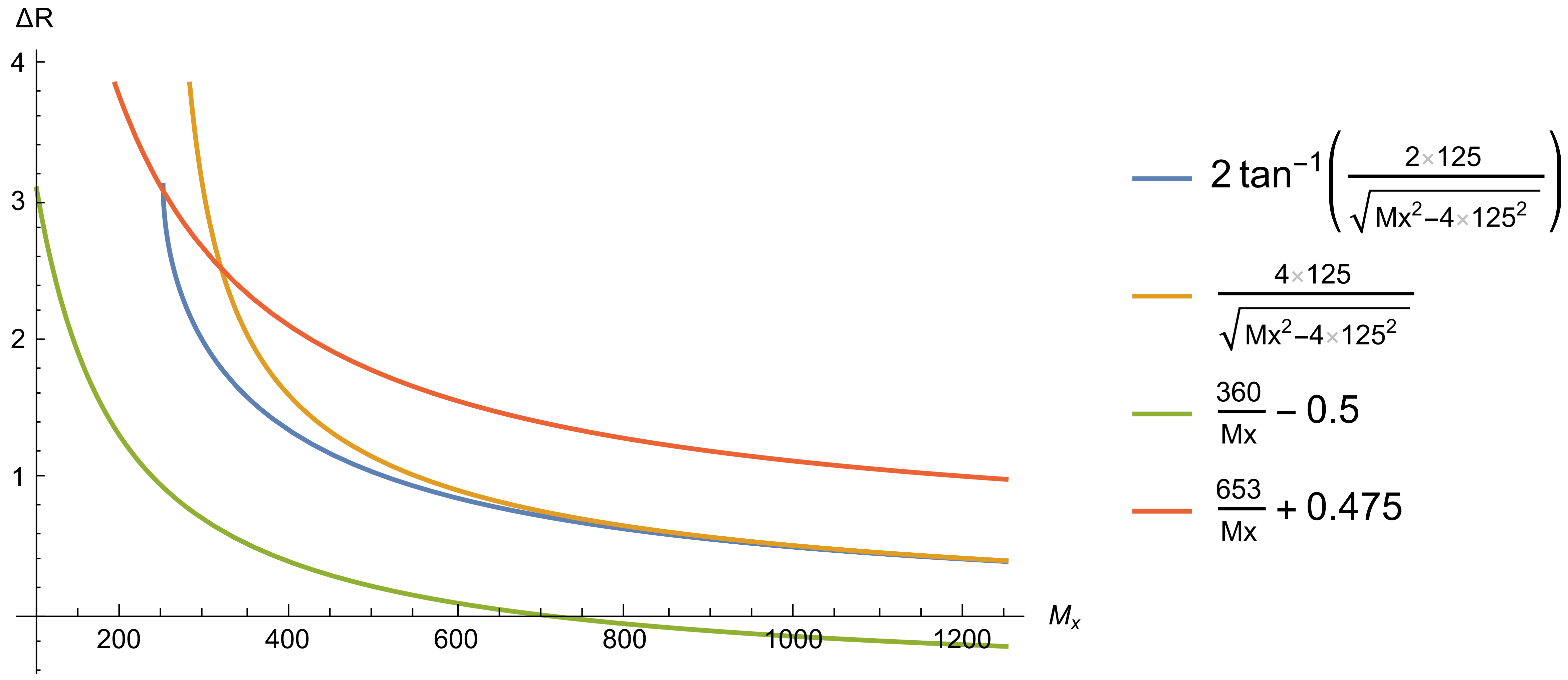
$$E_H \equiv E_{H_1} = E_{H_2} = \frac{m_X}{2}$$

$$p_H \equiv p_{H_1} = p_{H_2} = \sqrt{E_X^2 - m_H^2} = \frac{1}{2}\sqrt{m_X^2 - 4m_H^2}$$

In terms of resonance mass, X produced roughly at rest, Higgses are roughly back to back

```
In[245]:= Plot[{2 * ArcTan[2 * 125 / Sqrt[Mx ^ 2 - 4 * 125 ^ 2]],
   4 * 125 / Sqrt[Mx ^ 2 - 4 * 125 ^ 2], 360 / Mx - 0.5, 653 / Mx + 0.475}, {Mx, 100, 1250},
  AxesLabel → {Mx, ΔR}, PlotLegends → LineLegend["Expressions"]]
```

Out[245]=



$$\Longrightarrow \Delta R_{bb} \approx \frac{2m_H}{p_{T,H}} \approx \frac{4m_H}{\sqrt{m_X^2 - 4m_H^2}}$$

or (not Taylor series-ing the arctan)

$$\Delta R_{bb} \approx 2\arctan\left(\frac{2m_H}{\sqrt{m_X^2 - 4m_H^2}}\right)$$

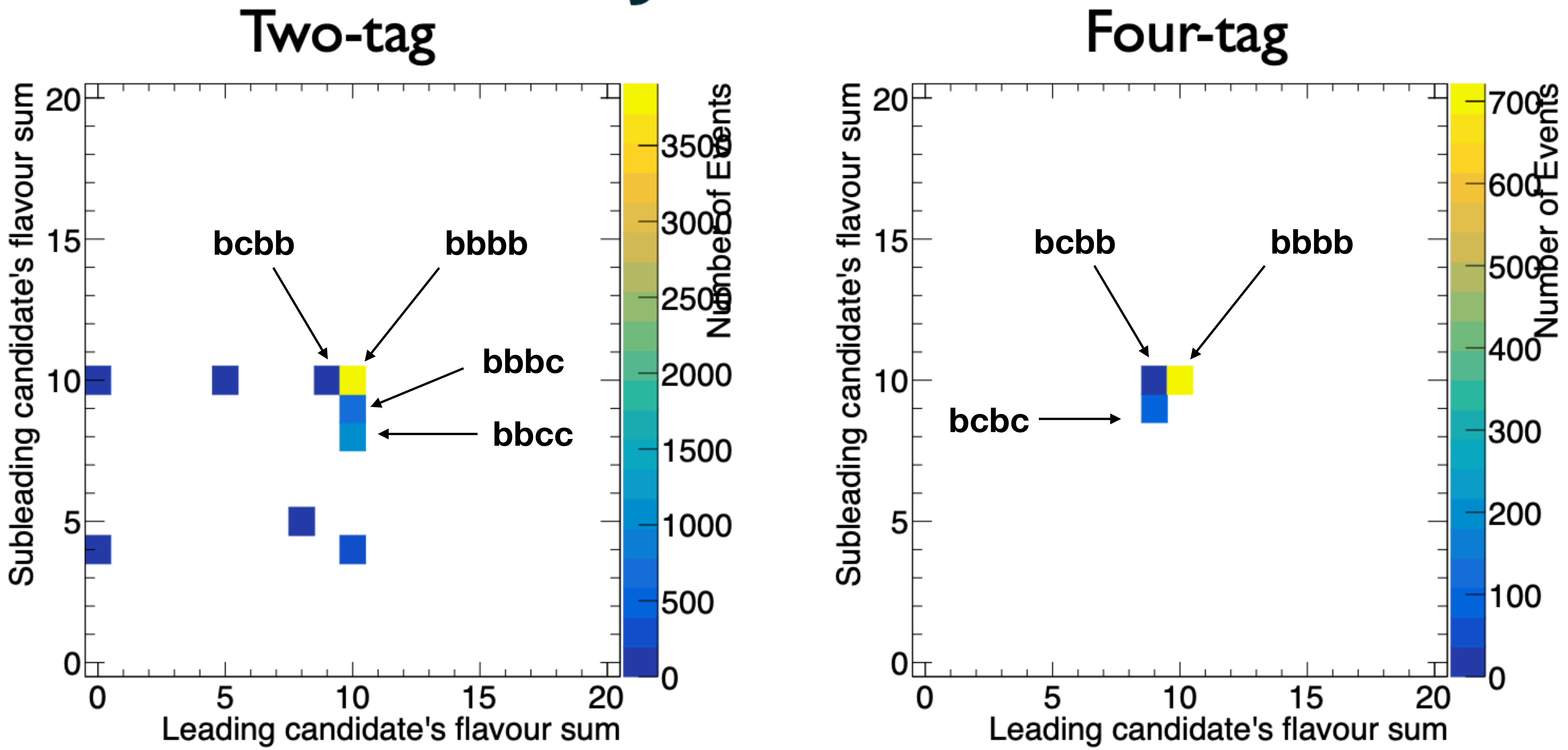
The Analysis: Background Processes

- What can these be? Need processes that result in a final state with 4 b-quarks
  - QCD processes: gg→4b, 2c2b e.g.
  - $t\bar{t} \to bW^+\bar{b}W^-$
    - Leptonic $t\bar{t}$: e.g. $W^+ \to c\bar{b}$, $W^- \to l^-\bar{\nu}_l$ => $b\bar{b}c\bar{b}$ in the final state, c fakes a b
    - Hadronic $t\bar{t}$:
      - e.g. $W^+ \to c\bar{s}$, $W^- \to \bar{c}s$ => $b\bar{b}c\bar{c}$ in the final state, where the c's fake b's
      - or (CKM suppressed) $W^+ \to c\bar{b}$, $W^- \to \bar{c}b$ => $b\bar{b}b\bar{b}$ in the final state
  - Other processes (ZZ→4b, etc.) are expected to have much smaller contributions, as we select the Higgs candidate masses to be close to m_H = 125 GeV

The Analysis: Background Composition

Flavor: 0 = light, 4=c, 5=b, 15=τ
(e.g. 10,10 is bbbb)

JZ4W



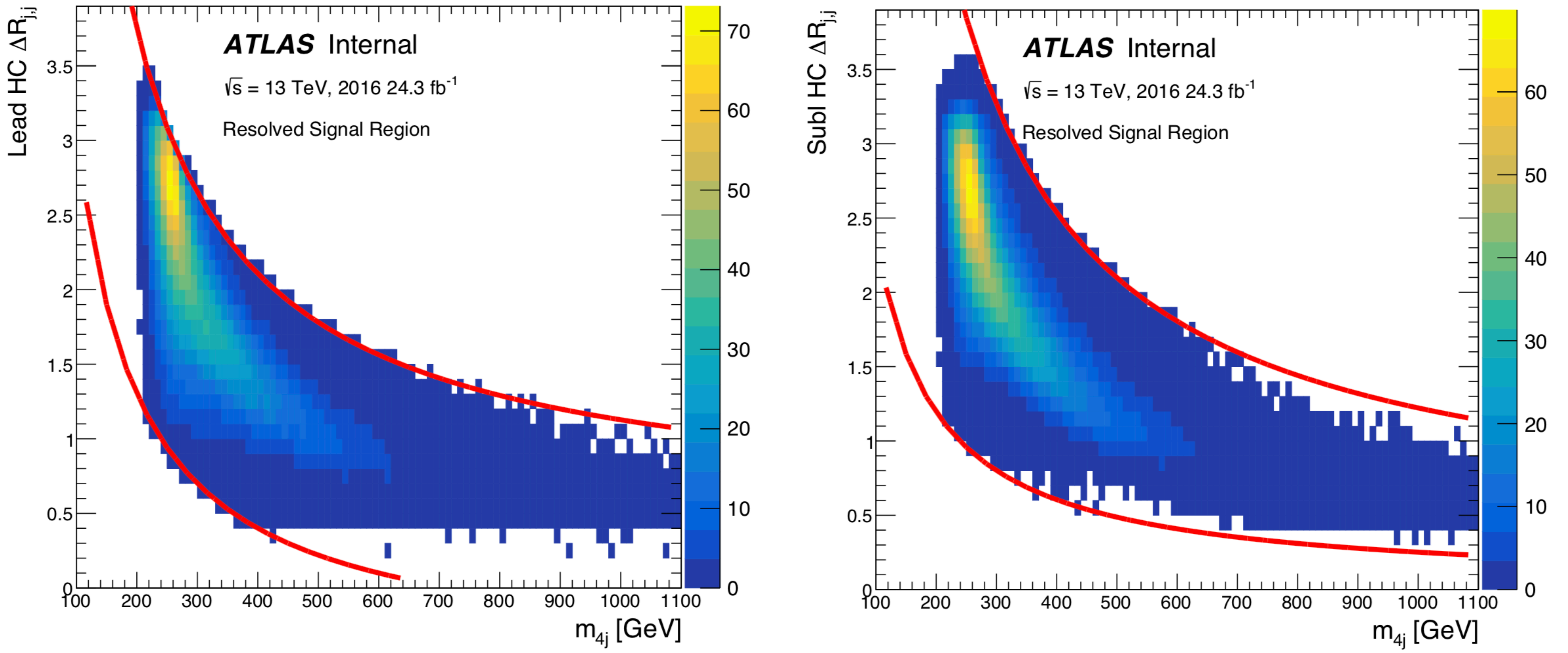
This slice makes biggest contribution (by far) to total background

Resulting from combination of acceptance and cross-section

Some plots and a paper, thanks to David Wardrope

Resolved Analysis: Selection and Pairing

- Require 4 b-tagged jets (R=0.4 Anti-k_t) with $p_T > 40$ GeV, $|\eta| < 2.5$
- Need to then pair them up! 3 possibilities
- Not all pairings are consistent with HH→4b
  - Angle between the decay products of the Higgs depends on m_{4j} (m_{HH})
  - Requirements on the right efficiently reject pairings in which one b-tagged jet is not consistent with coming from a Higgs boson decay
  - Pairings which pass this selection are considered "valid"



2016 Multijet after full selection (15/16 analysis). Red lines show bounds on ΔR_{jj}

m_{4j} < 1250 GeV

$$\frac{360\ \text{GeV}}{m_{4j}} - 0.5 < \Delta R_{jj,\text{lead}} < \frac{653\ \text{GeV}}{m_{4j}} + 0.475$$

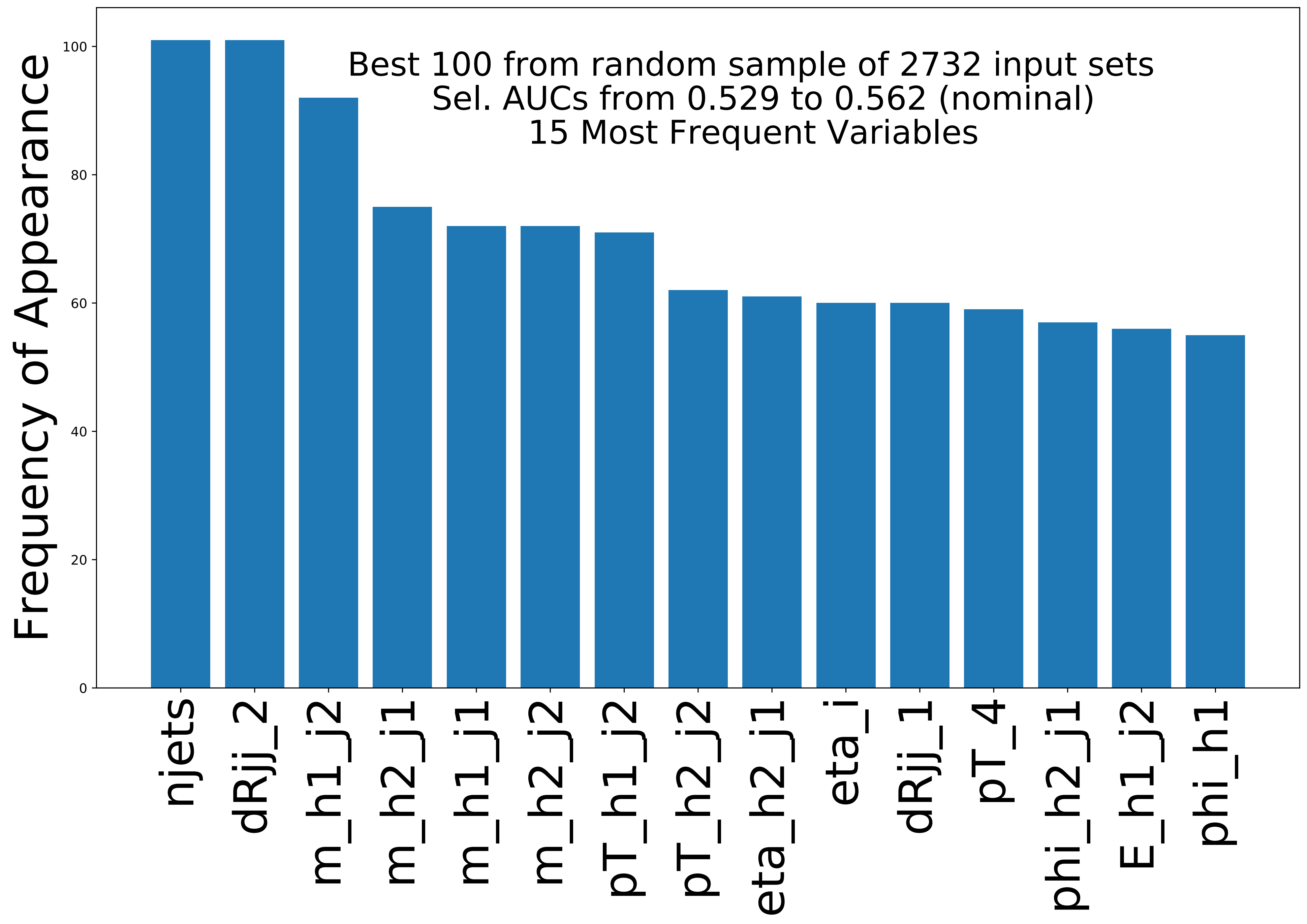
$$\frac{235\ \text{GeV}}{m_{4j}} < \Delta R_{jj,\text{subl}} < \frac{875\ \text{GeV}}{m_{4j}} + 0.35$$

m_{4j} > 1250 GeV

$$0 < \Delta R_{jj,\text{lead}} < 1$$

$$0 < \Delta R_{jj,\text{subl}} < 1$$

Reweighting BDT: Optimization



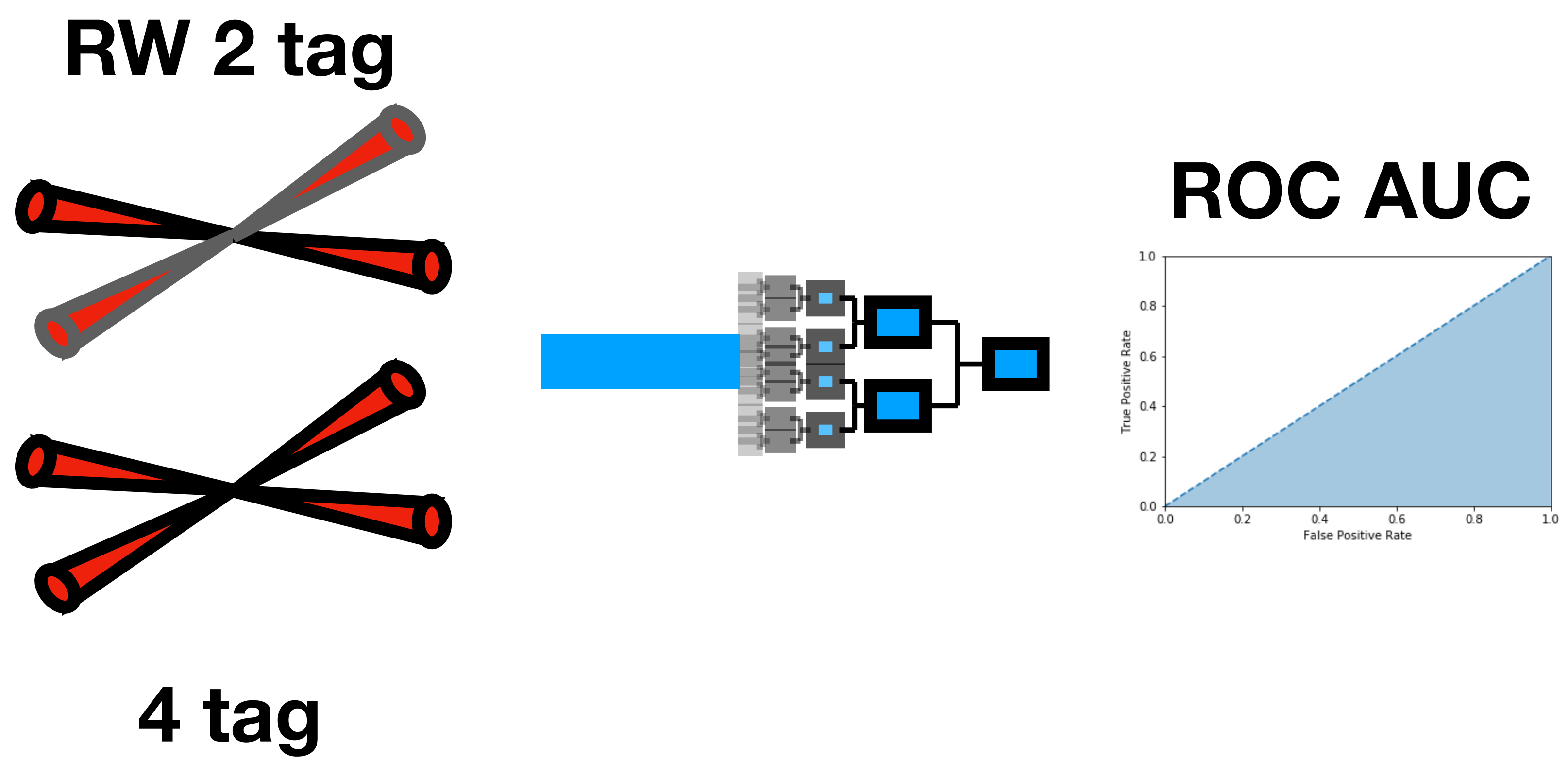
Frequency of variable appearance in best 100 input sets (of ~3000 sampled). More frequent => likely more important

- Optimization:
  - **Focus**: input variable set. A lot of options!
  - **Method**: Random search through possible input sets

$$\sum_{i=1}^{37} \binom{37}{i} = 137{,}438{,}953{,}471$$

Number of possible input sets given ~37 relevant variables

<u>Metric (Classifier)</u>:



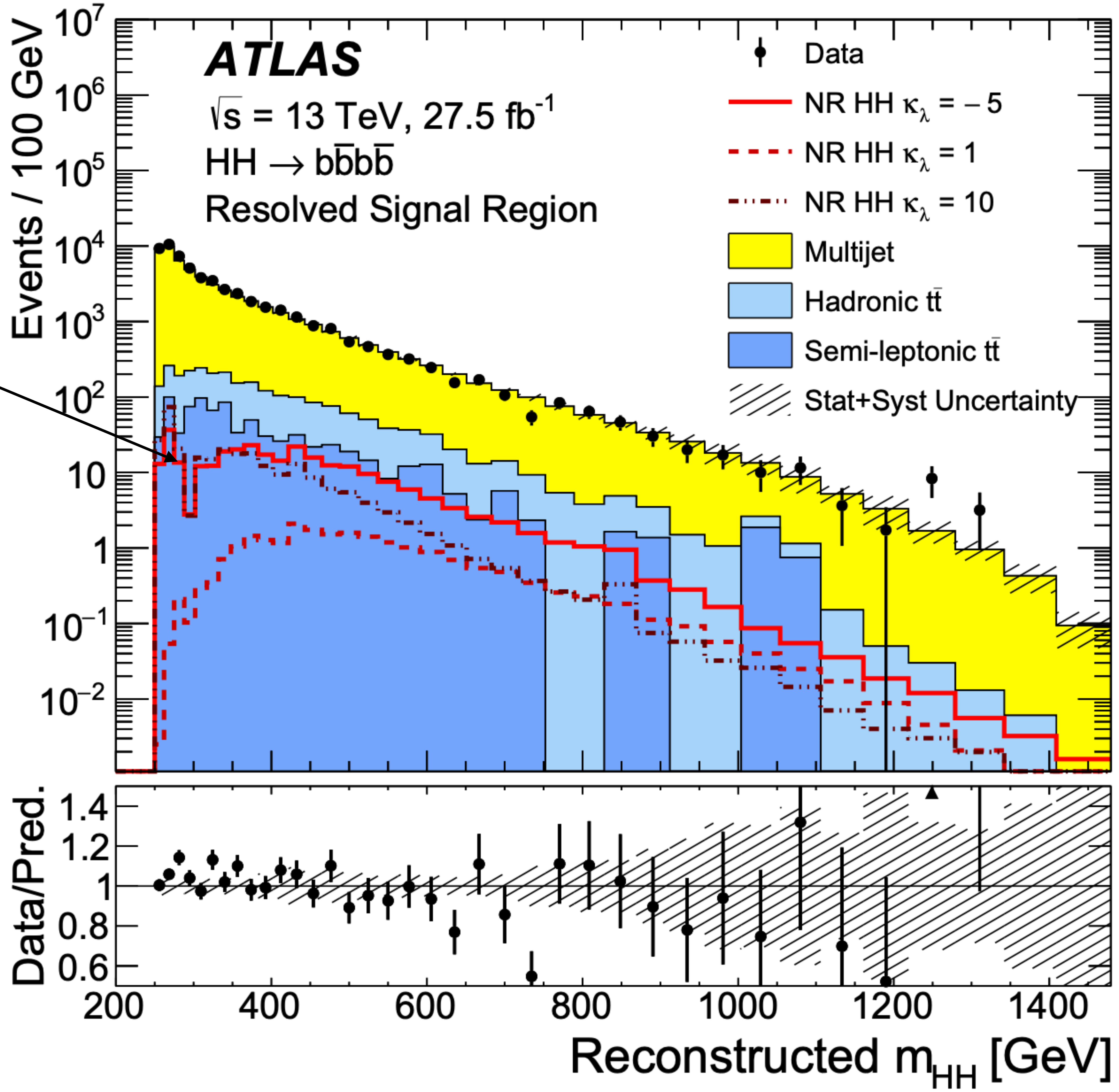
Non-resonant HH: Beyond the Standard Model

- κ_λ samples generated by computing the m_{HH} spectrum for each value of κ_λ at the **generator level** (no detector, just physics)

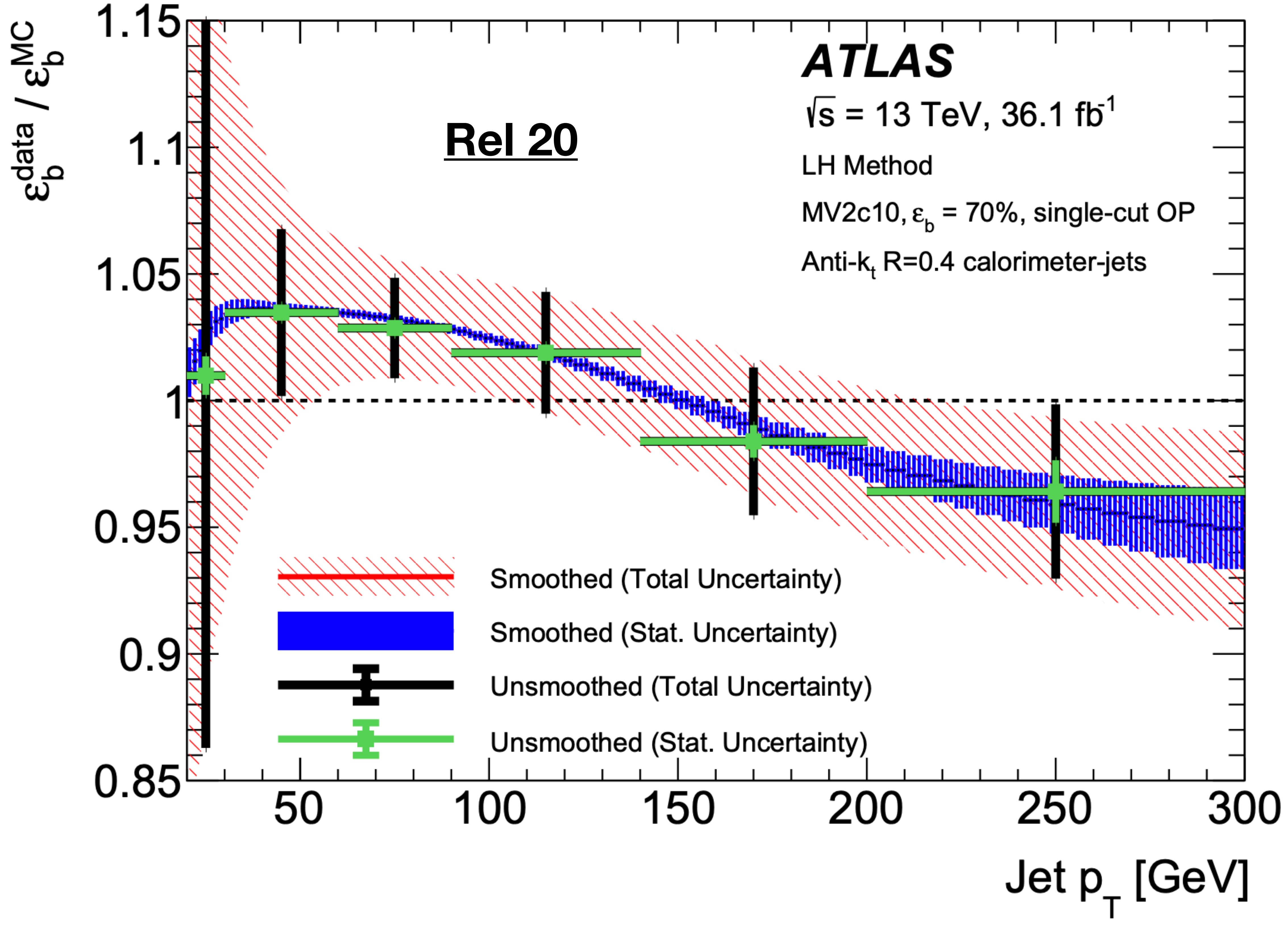
- Binned ratios to the Standard Model value m_{HH} spectrum are then computed and used to reweight events from a full detector simulation of NLO SM HH samples

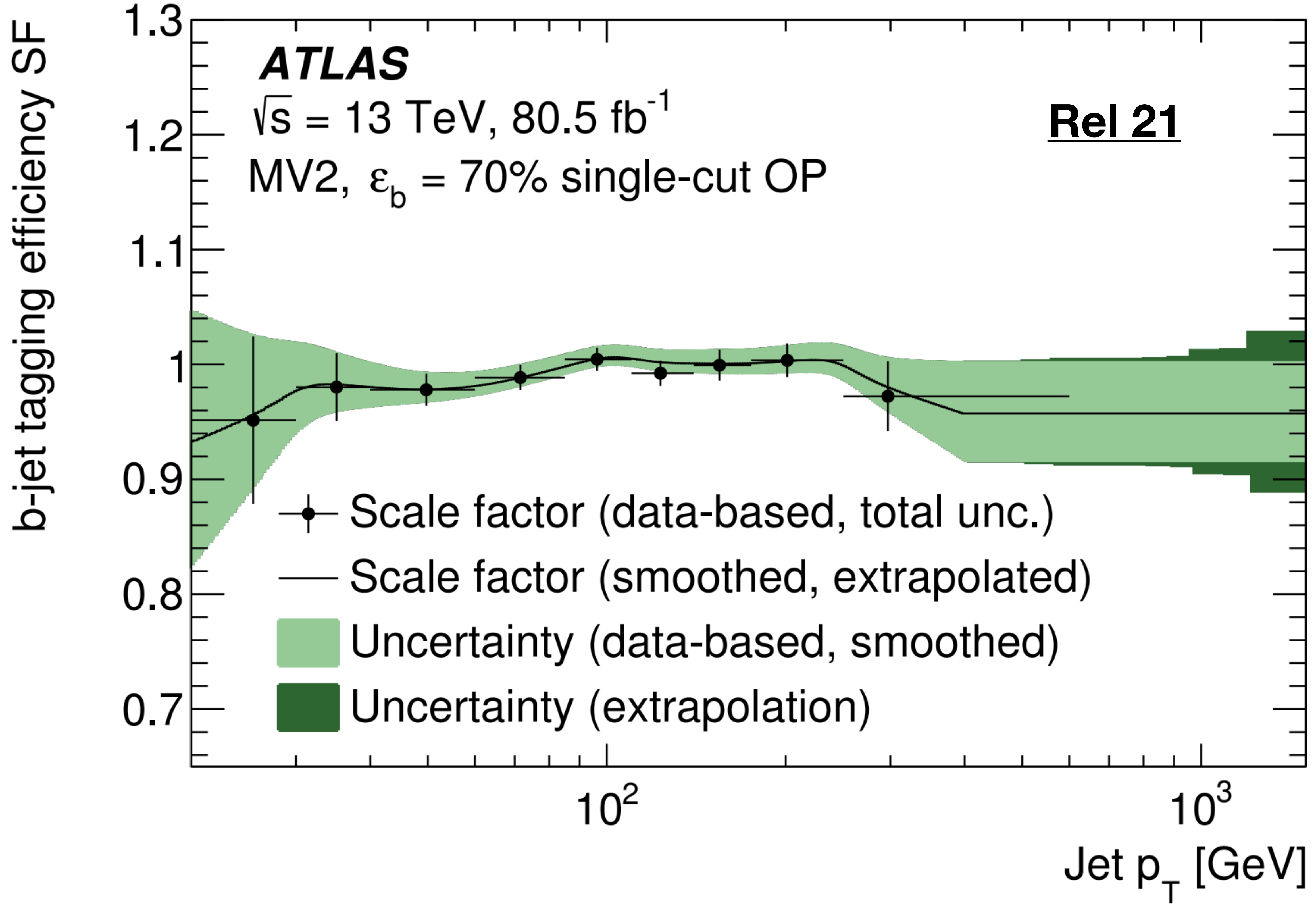
- Thus, statistical fluctuations translate across the samples, resulting in the dip seen on right

Fluctuation in κ_λ = -5 and 10 purely statistical



b-tagging SF’s



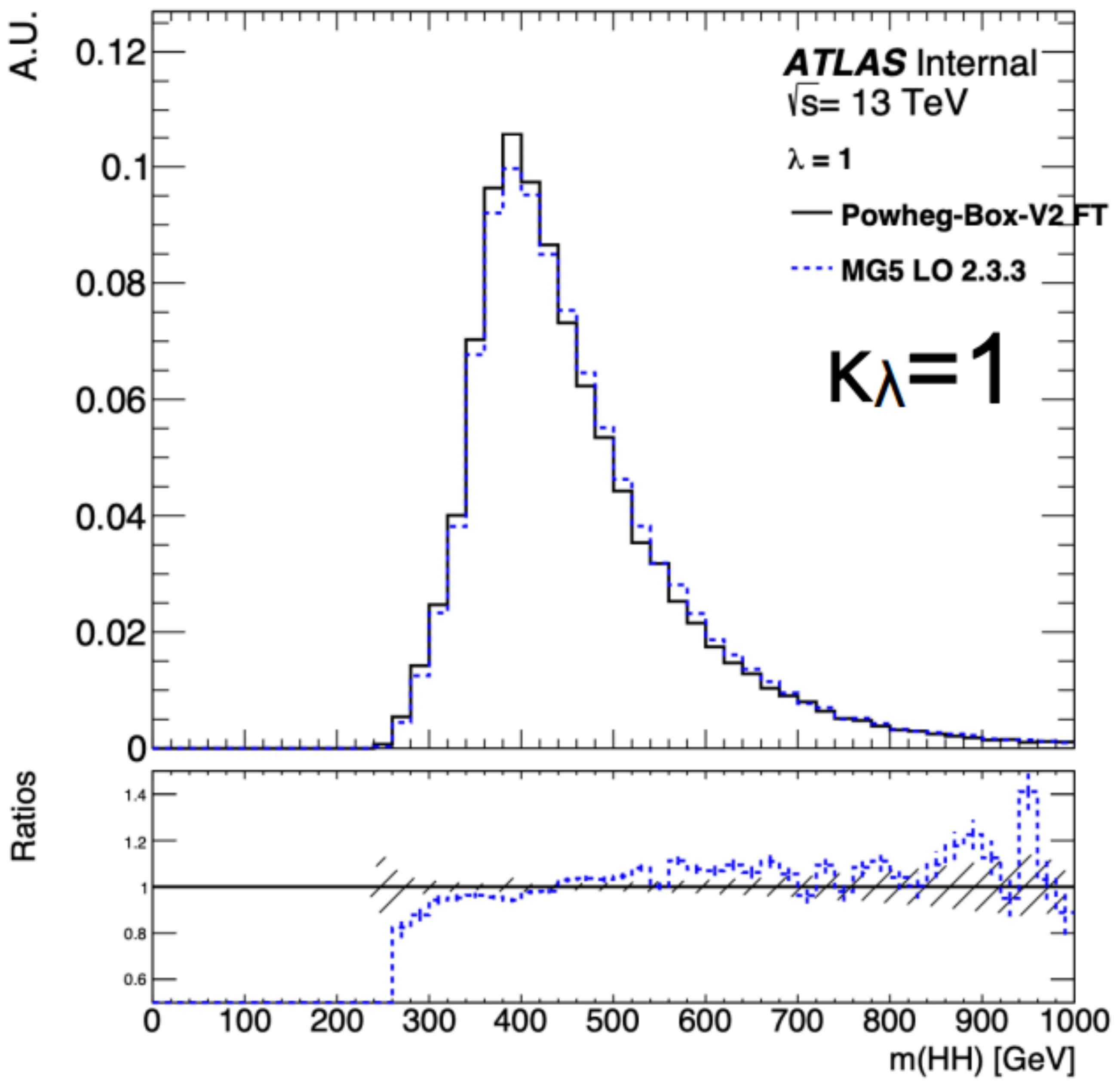


1.03^4 = 1.12

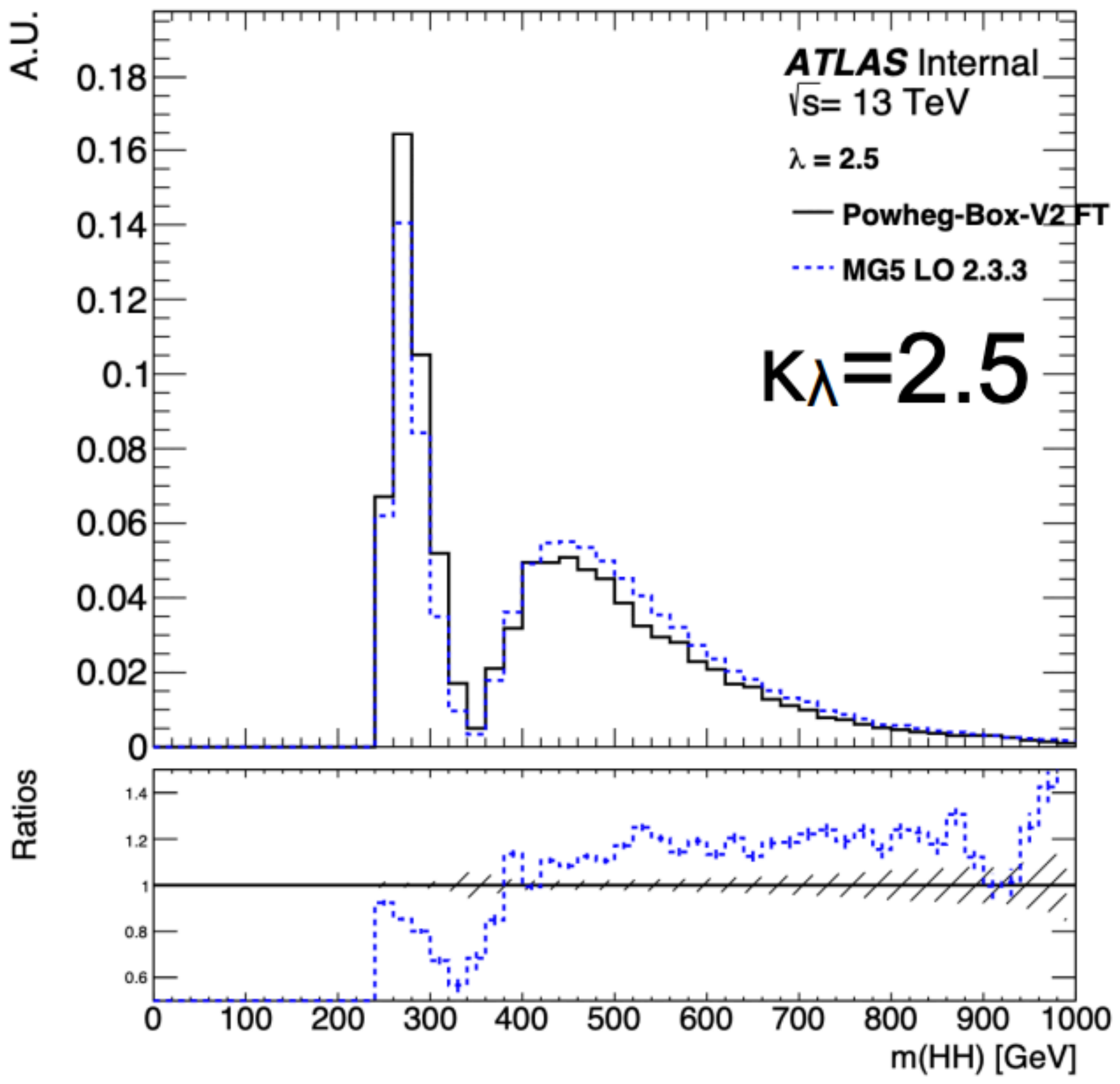
0.97^4 = 0.89

Consistent with ~20-25% difference in yields

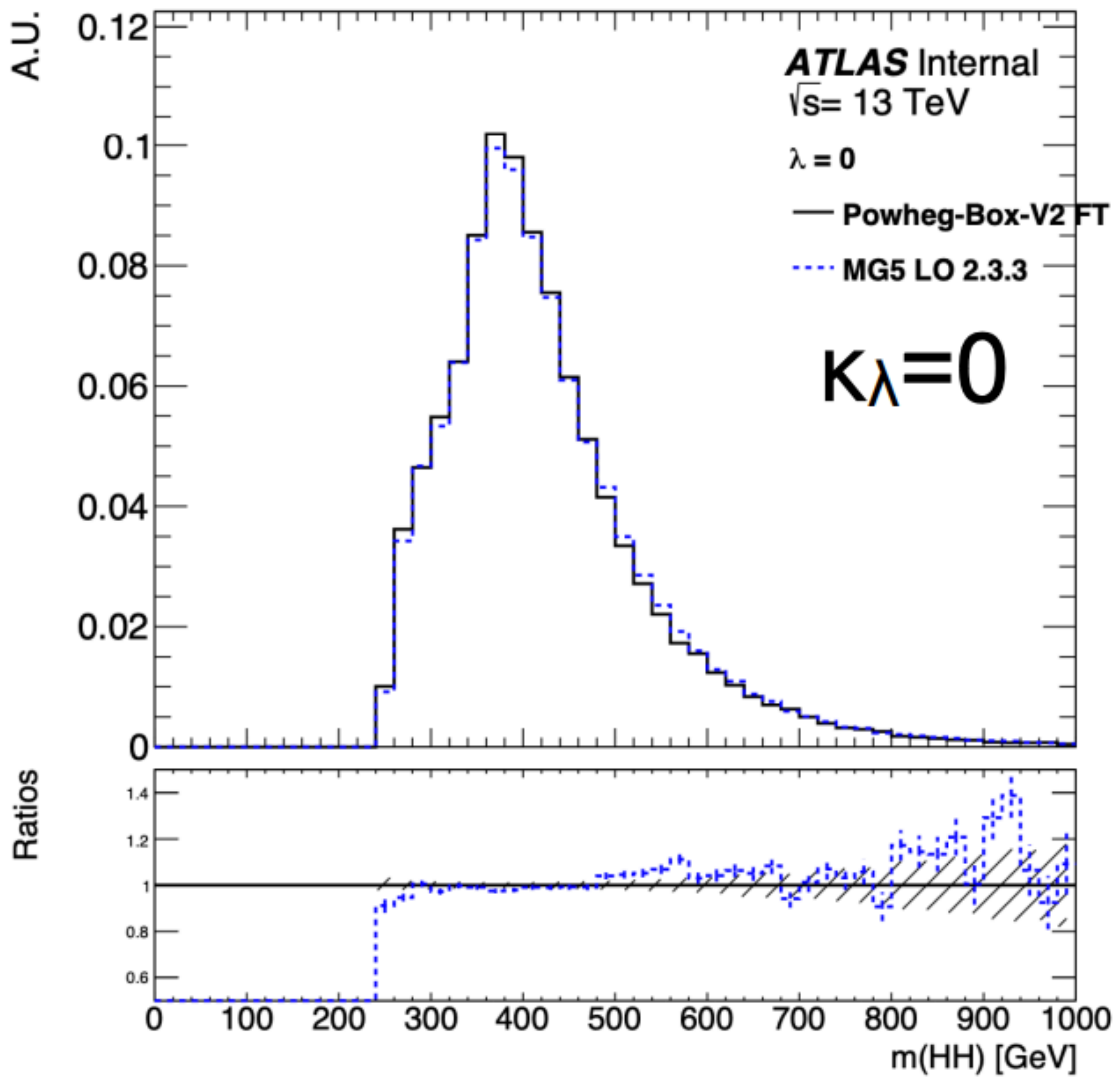
More κ_λ variations



SM



Close to max interference of the triangle and the box diagrams



Box diagram only

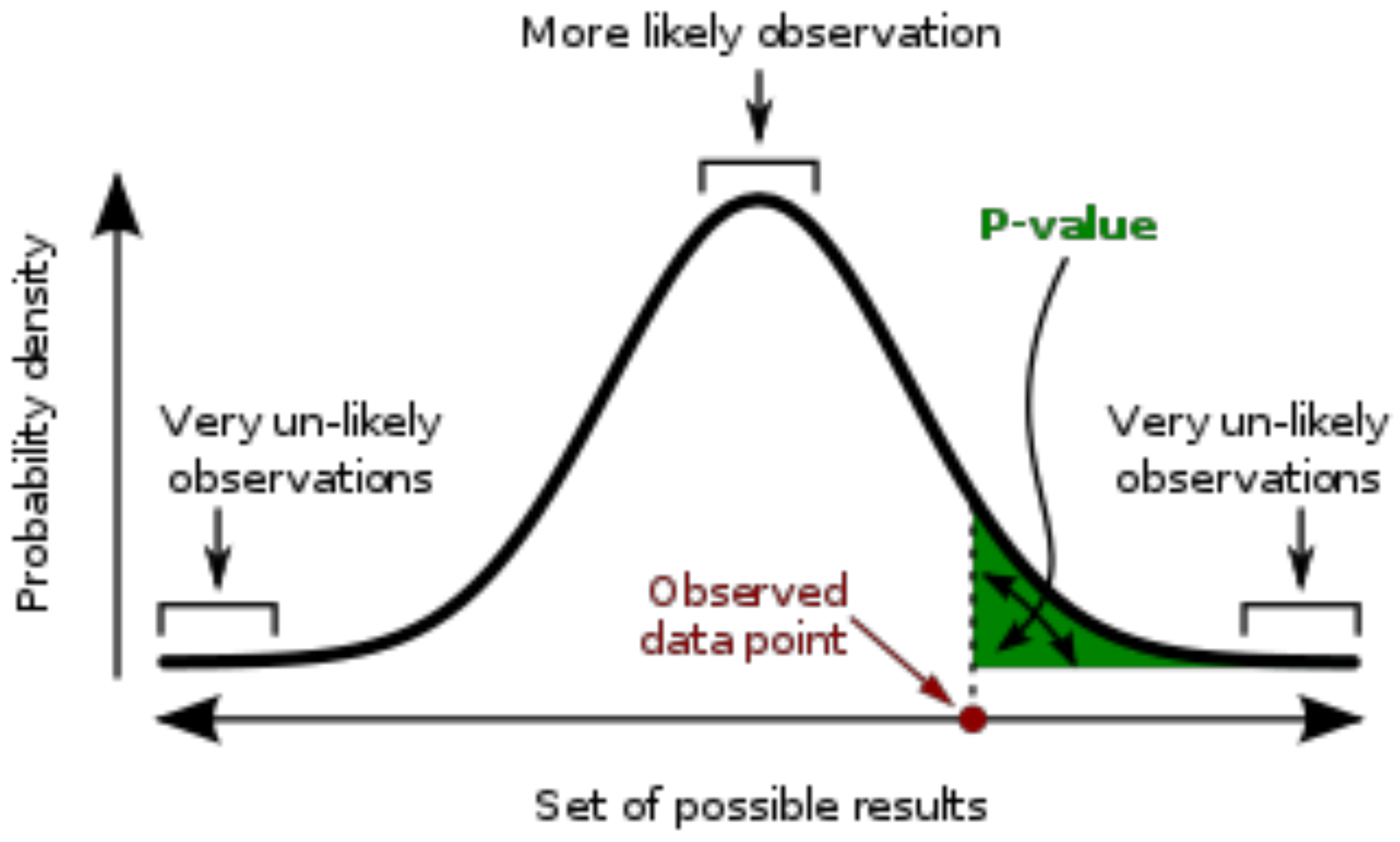
Source (Xiaohu Sun)

Limit Setting Overview: Significance

- Null hypothesis H_0 (background only), alternative hypothesis, H_1 (signal + background)
- Quantify agreement with H by computing a probability of finding data with equal or greater incompatibility with H (p-value)
- Particle physics defines significance

$$Z = \Phi^{-1}(1 - p)$$

where Φ^{-1} is the inverse of the cumulative distribution for the standard Gaussian



Source

Standard for Discovery: Z=5, p = 2.87 × 10^{-7}
Standard for exclusion: p=0.05 (95% CL), Z=1.64

See, e.g., Cowan, et al

Limit Setting Overview: Binned Analyses

- We use a binned analysis (histograms)
  - We measure some variable x (e.g. m_{HH}) and create a histogram $\mathbf{n} = (n_1, \ldots, n_N)$
  - Then the expectation value in each bin can be written

    $$E[n_i] = \mu s_i + b_i$$

    with s_i corresponding mean number of entries in each bin from signal, b_i from background, and μ as the signal strength

$$s_i = s_{\text{tot}} \int_{\text{bin } i} f_s(x; \boldsymbol{\theta}_s)\, dx$$

$$b_i = b_{\text{tot}} \int_{\text{bin } i} f_b(x; \boldsymbol{\theta}_b)\, dx$$

A way of representing s_i and b_i, where $f_{s(b)}$ corresponds to the probability density functions of the variable x with nuisance parameters (systematics, e.g.) θ, which can impact the shape

See, e.g., Cowan, et al

Limit Setting Overview: Likelihood

- Most searches are based on **likelihood ratios**

- Likelihood is defined on right. This is the product of the Poisson probabilities for all bins

- The second contribution (u_k) is from measurements to constrain nuisance parameters (see **m** on right)

- Likelihood ratios are then defined as on right for a given value of the signal strength, μ. The hats denote maximum likelihood (ML) parameters (values that maximize the likelihood function)

$$L(\mu, \boldsymbol{\theta}) = \prod_{j=1}^{N} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)} \prod_{k=1}^{M} \frac{u_k^{m_k}}{m_k!} e^{-u_k}$$

$$\mathbf{m} = (m_1, \ldots, m_n)$$

$$E[m_i] = u_i(\boldsymbol{\theta})$$

Histogram for constraints on nuisance parameters. The u_i are calculable from θ.

ML parameter for given μ

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})}$$

Maximized likelihood

See, e.g., Cowan, et al

Limit Setting Overview: Test Statistics

- p-value calculation is based on some **test statistic** in which the deviation from hypothesis H is measured
  - We use $CL_s = \frac{CL_{s+b}}{CL_b} = \frac{P_{s+b}(q \leq q_{obs})}{P_b(q \leq q_{obs})}$ for test statistic q (ratio of probabilities to produce a value of q less than observed). Cross sections excluded if $CL_s \leq 0.05$
- For 4b we use the statistics shown on right
  - Limit setting takes into account that, for resonances on top of background, $\mu < 0$ is unphysical
  - For upper limits, data with $\hat{\mu} > \mu$ would not be less compatible with μ than the data obtained (μ is “below” the limit)

$$q_0 = \begin{cases} -2ln\frac{L(0,\hat{\hat{\theta}}(0))}{L(\hat{\mu},\hat{\theta})} & \hat{\mu} > 0 \\ 0 & \hat{\mu} < 0 \end{cases}$$

Test statistic used for searches (compatibility with background only)

$$\widetilde{q_\mu} = \begin{cases} -2ln\frac{L(\mu,\hat{\hat{\theta}}(\mu))}{L(0,\hat{\hat{\theta}}(0))} & \hat{\mu} < 0 \\ -2ln\frac{L(\mu,\hat{\hat{\theta}}(\mu))}{L(\hat{\mu},\hat{\theta})} & 0 \leq \hat{\mu} < \mu \\ 0 & \hat{\mu} > \mu \end{cases}$$

Test statistic used for upper limit setting

See, e.g., Cowan, et al

Limit Setting Overview: Asymptotic Approx

- Limits are calculated in the **asymptotic approximation**
  - Likelihood ratios can be written in a simple form as on right, where N is the data sample size, σ is taken from the covariance matrix of estimators of the parameters
  - Assuming $\hat{\mu}$ is Gaussian distributed and ignoring the $1/\sqrt{N}$ term, the ratio as on right follows a non-central chi-squared distribution
    - We can then calculate significances numerically
    - Note: ignoring $1/\sqrt{N}$ term is a good approximation even for N ~ 10

$$-2\ln\lambda(\mu) = \frac{(\mu-\hat{\mu})^2}{\sigma^2} + \mathcal{O}(1/\sqrt{N})$$

Asymptotic form of likelihood ratios

$$\tilde{q}_\mu = \begin{cases} \frac{\mu^2}{\sigma^2} - \frac{2\mu\hat{\mu}}{\sigma^2} & \hat{\mu} < 0 , \\ \frac{(\mu-\hat{\mu})^2}{\sigma^2} & 0 \leq \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu . \end{cases}$$

$$Z_\mu = \begin{cases} \sqrt{\tilde{q}_\mu} & 0 < \tilde{q}_\mu \leq \mu^2/\sigma^2 \\ \frac{\tilde{q}_\mu + \mu^2/\sigma^2}{2\mu/\sigma} & \tilde{q}_\mu > \mu^2/\sigma^2 . \end{cases}$$

Asymptotic form of upper limit test statistic and significance

See, e.g., Cowan, et al

Counting Experiment Significance

$$L(\mu) = \frac{(\mu s + b)^n}{n!} e^{-(\mu s + b)}$$

Regard b as known. Data is then just n.

$$q_0 = \begin{cases} -2 \ln \frac{L(0)}{L(\hat{\mu})} & \hat{\mu} \geq 0, \\ 0 & \hat{\mu} < 0 \end{cases}$$

Test statistic for discovery

$$Z_0 = \sqrt{q_0} = \begin{cases} \sqrt{2\left(n \ln \frac{n}{b} + b - n\right)} & \hat{\mu} \geq 0, \\ 0 & \hat{\mu} < 0. \end{cases}$$

Asymptotic approximation for significance

$$\mathrm{med}[Z_0|1] = \sqrt{q_{0,\mathrm{A}}} = \sqrt{2\left((s+b)\ln(1 + s/b) - s\right)}$$

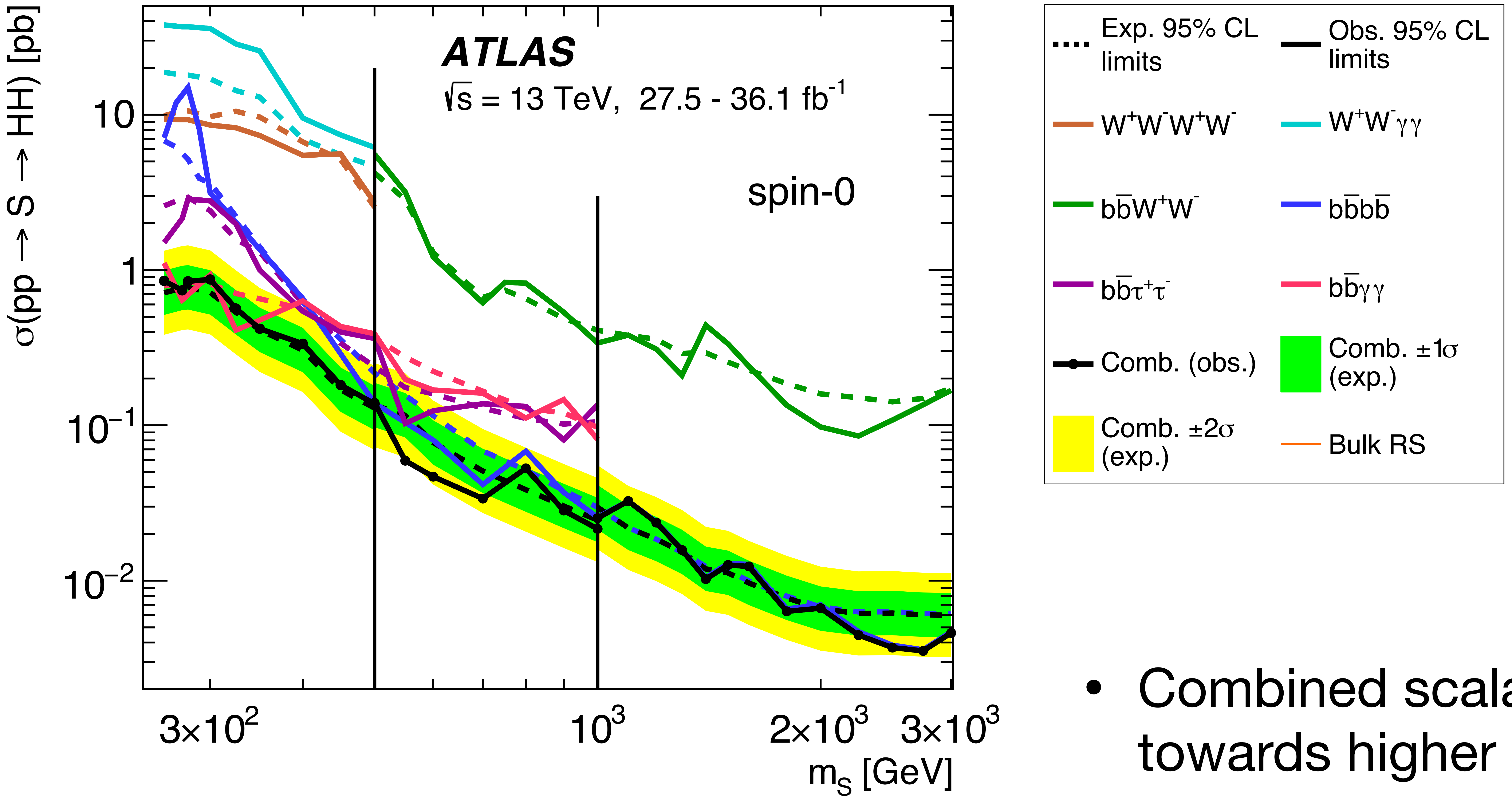
Assume nominal signal hypothesis ($\mu = 1$). Replace n by its Asimov value (expectation)

$$\mathrm{med}[Z_0|1] = \frac{s}{\sqrt{b}}\left(1 + \mathcal{O}(s/b)\right)$$

Expand log in s/b

See, e.g., Cowan, et al

Combination (Spin 0)

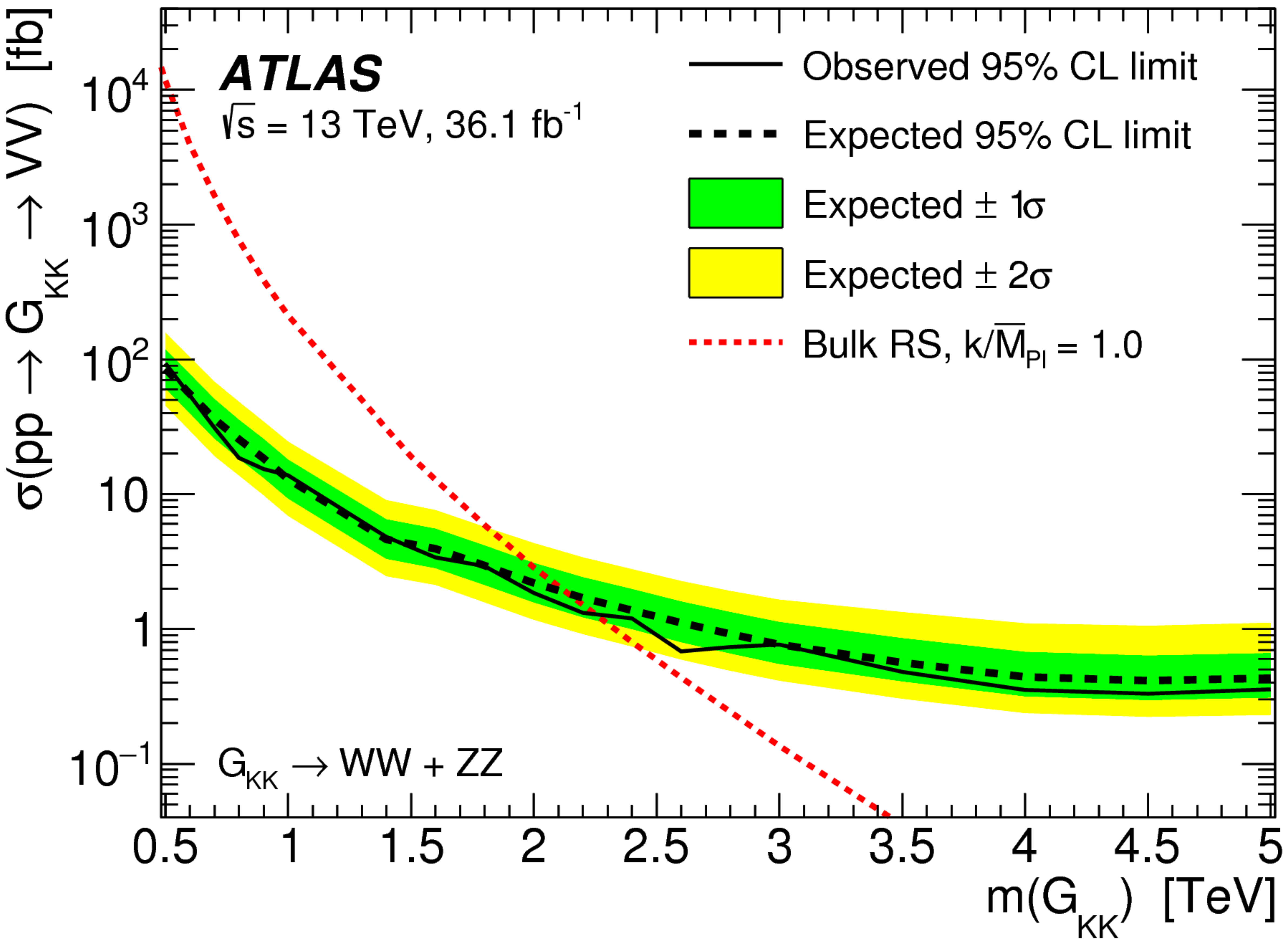


- Combined scalar limits. 4b dominates towards higher mass (> 400 GeV)

Combination (Spin 2)

- Combined spin 2 limits. Again, 4b dominates

c=1.0 Graviton Exclusion



Source

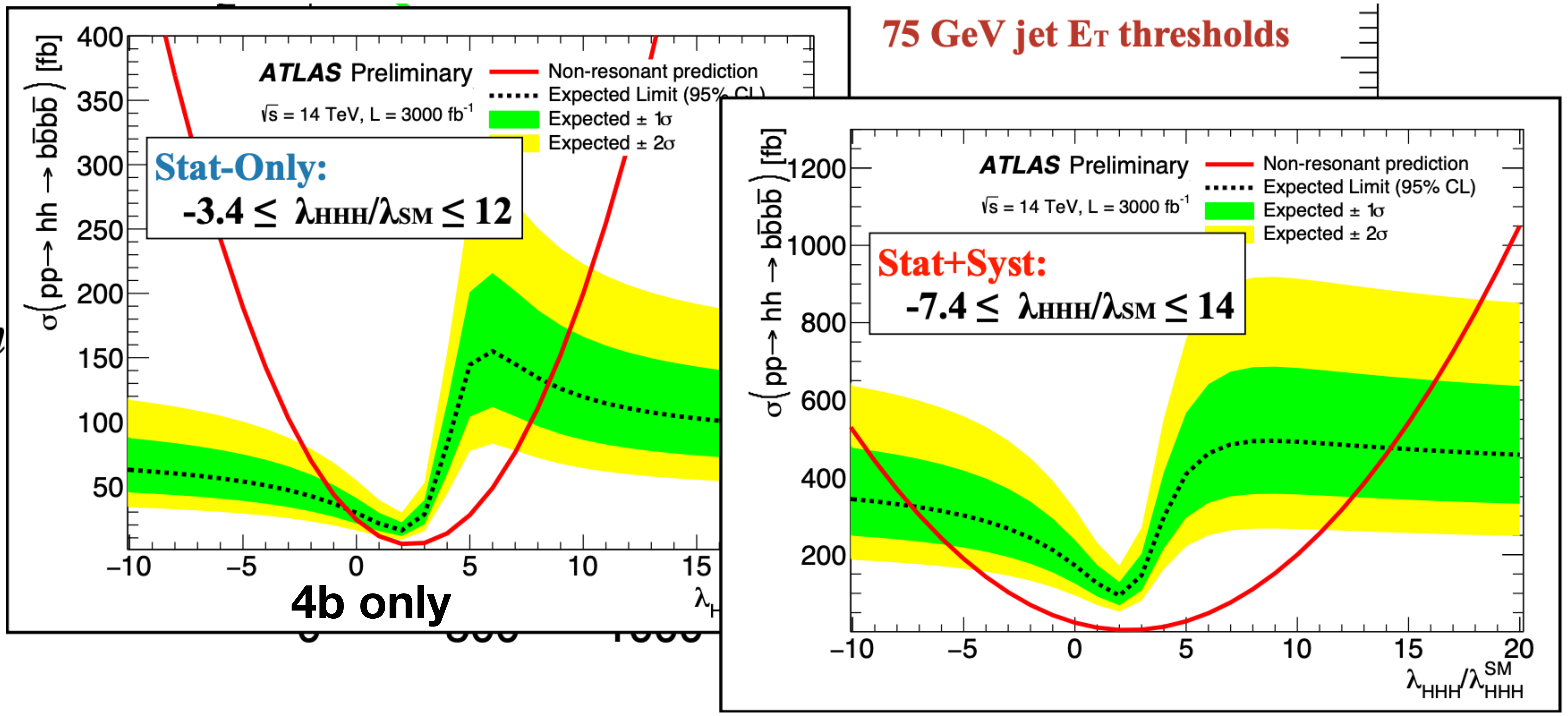
- Strong limits set on c=1.0 RS KK graviton by bosonic+leptonic final state searches
- Excludes the model over most of the 4b mass range
  - Exclusion here from ~500 to ~2300 GeV
  - 4b mass range from ~300 to ~3000 GeV

HL-LHC Projections

95% CL limits on μ_{HH}

14 TeV, 3000/fb

channel	CMS stat.-only (stat+sys)	ATLAS stat.-only (stat+sys)
bbbb	2.9 (7.0)	2.0 (11.5)
bbγγ	1.3 (1.3)	- (2.6)
bbττ	3.9 (5.2)	- (4.3)
bbWW	4.6 (5.8)	-

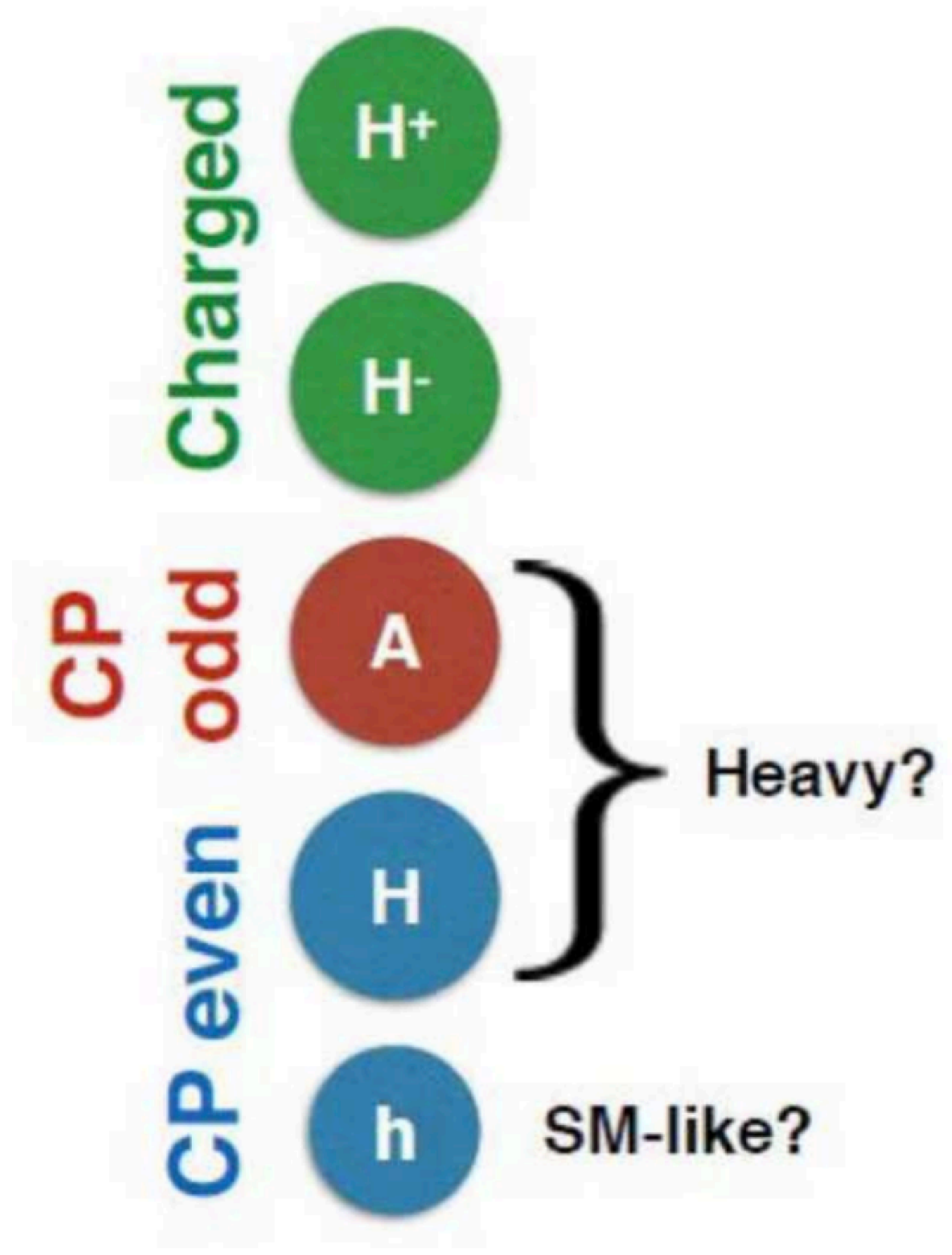


Source

- Projections shown with current analysis/systematics model
- At the end of the HL-LHC, we will start to be sensitive to larger deviations from the standard model
    - Estimated ~30% precision on λ_{HHH} after combining channels/experiments
- As always, however, the goal is to beat the projection - with new methods, we will hopefully be able to push further

2HDM

- Describes a class of theories with two Higgs doublets instead of one
- This results in a total of 5 physical Higgses (mass eigenstates):
  - Recall: for 1 Higgs doublet, we had 4 degrees of freedom, 3 of which are “eaten” by giving the W and Z masses
  - For 2 Higgs doublets, we then have 8 degrees of freedom, 3 of which, again, are “eaten”
- Under some assumptions (real VEVs, e.g.) these are: two CP even scalars (h, H) one CP odd scalar (A), and two charged scalars ($H^{\pm}$)



Source

Overview

2HDM continued

- Parameters then include the 4 masses of the physical Higgs bosons as well as angles α and β, where
  - α describes the mixing between the two CP even scalars (h and H)
  - $\tan\beta = \frac{v_2}{v_1}$, the ratio of the VEVs of the two doublets
- Additionally, we have a parameter $v = \sqrt{v_1^2 + v_2^2}$, which is set to be 246 GeV, if one of the scalars is the SM Higgs

$$\Phi_i = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2} w_i^+ \\ (h_i + v_i) + i z_i \end{pmatrix}$$

Mass Eigenstates

Neutral Goldstone and CP Odd Higgs: $\begin{pmatrix} \zeta \\ A \end{pmatrix} = \begin{pmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$

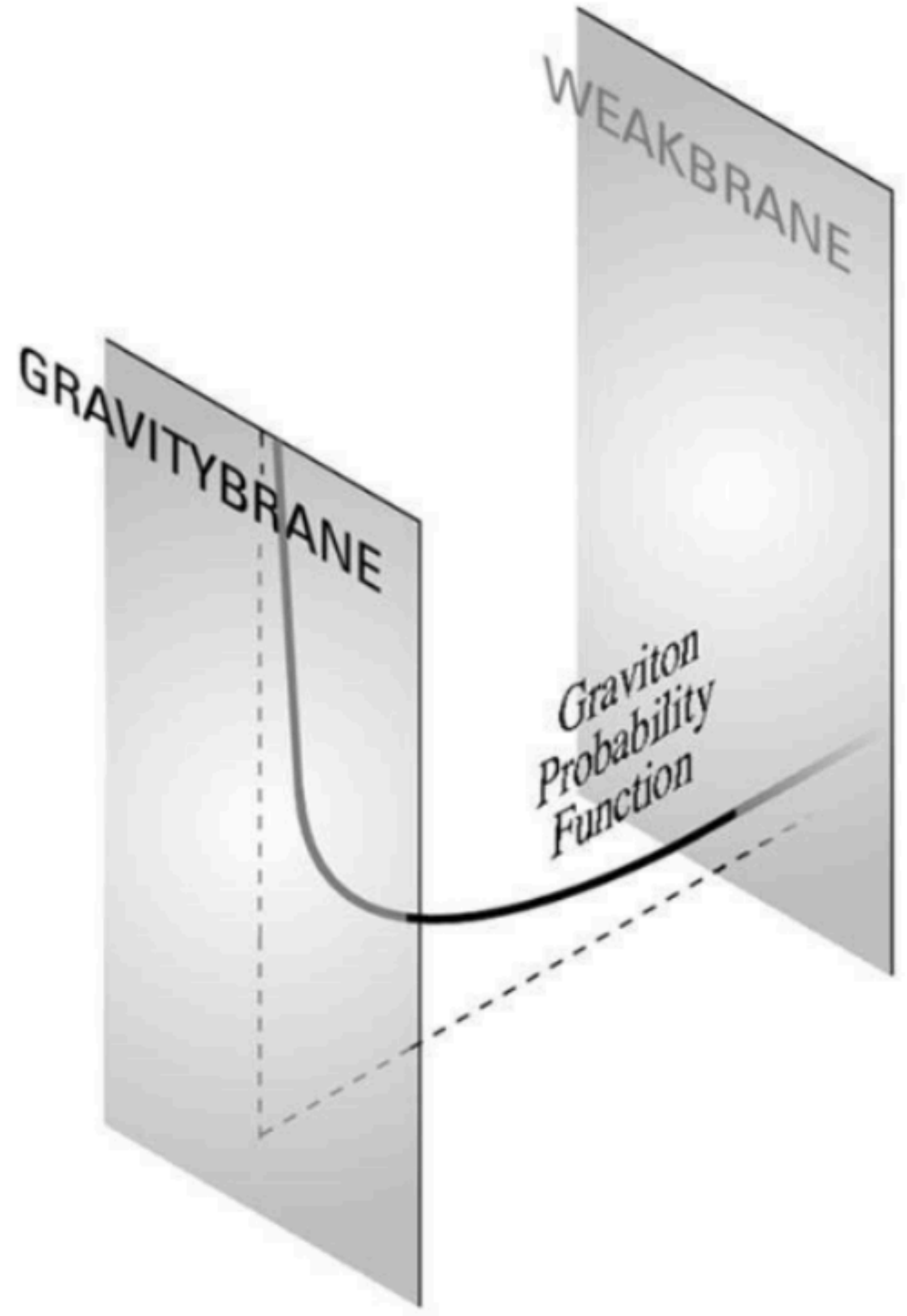
CP Even Higgses: $\begin{pmatrix} H \\ h \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$

Charged Goldstone and Charged Higgs: $\begin{pmatrix} \omega^\pm \\ H^\pm \end{pmatrix} = \begin{pmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{pmatrix} \begin{pmatrix} w_1^\pm \\ w_2^\pm \end{pmatrix}$

Source

Randall Sundrum Model

- Idea: we live in 3+1 dimensions
- The Randall Sundrum model postulates a warped fifth dimension which contains
  - A 3+1 dimensional "gravitybrane" (where gravity is strong - also called the Planckbrane, scale ~M_{Pl})
  - A 3+1 dimensional "weakbrane" (where the Standard Model lives - also called the TeVbrane, TeV scale)
  - This warping solves the **hierarchy problem** (why the weak force is so much stronger than gravity) as it generates a large ratio of energy scales
    - In particular TeV/M_{Pl} ~ $e^{-k\pi R}$ gives the ratio of scales, where k is the curvature scale and R is the proper size of the extra dimension. For the right scale, $kR \approx 11$



Schematic of the Randall Sundrum model [ref]

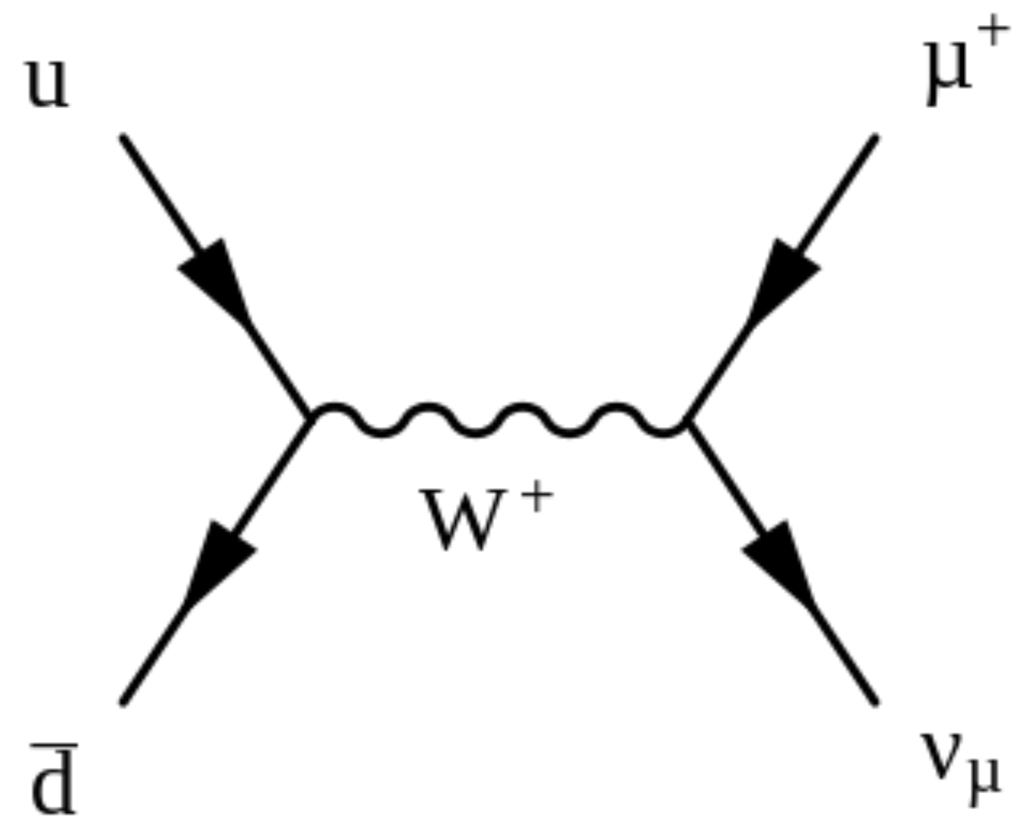
RS and Gravitons

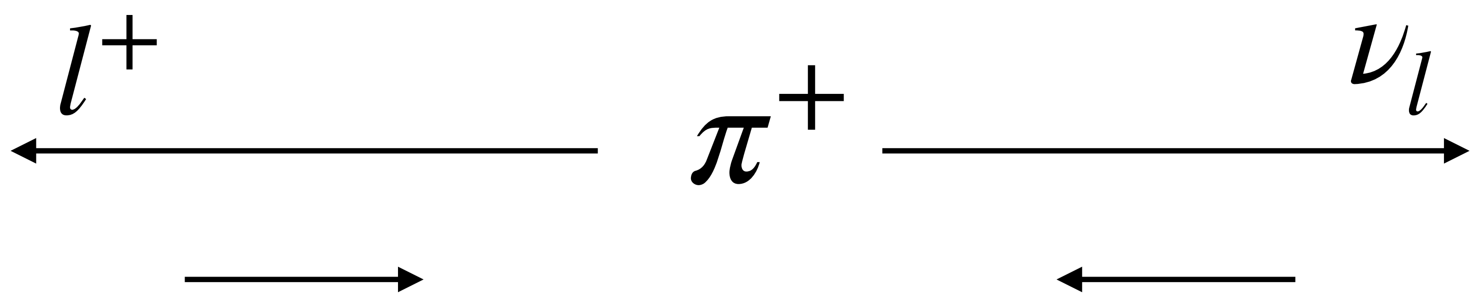
- Kaluza-Klein theory (very roughly) provides a prescription for 5D theories by decomposing them in terms of their 4D parts, plus some extras
- Why do we care about branes and warped extra dimensions?
  - Masses and couplings of Kaluza-Klein modes in this theory are of TeV scale - perfect for the LHC!
  - Idea is that the large Planck scale (weak gravity) arises from the small overlap of the graviton wave function in the fifth dimension with our brane
  - However, all other scales are TeV scale, which we can regard as fundamental to the theory
  - Theory predicts SM couplings - can search with HH!

Source

Misc: Pion Decay

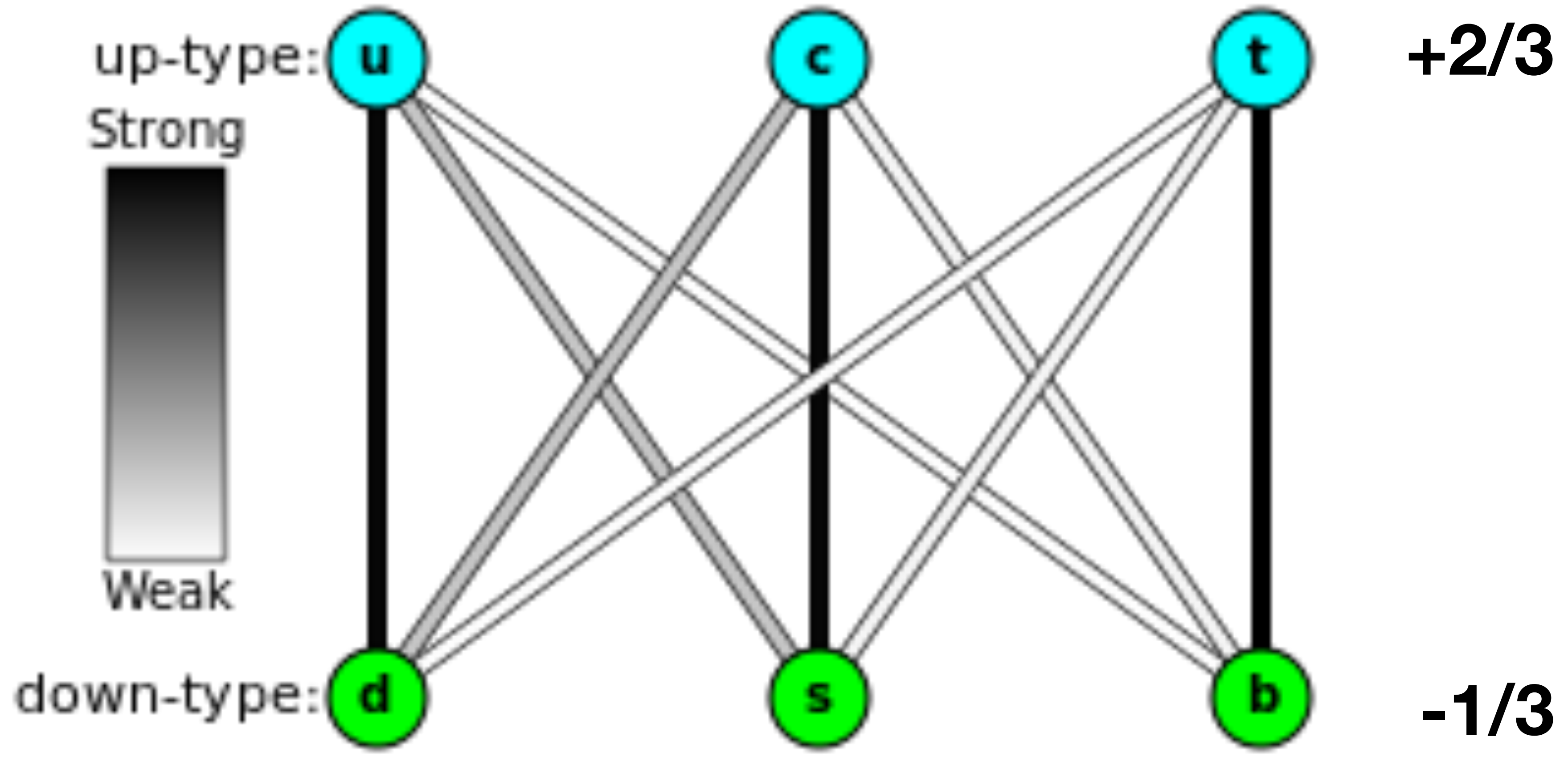
- Pion is spin 0
- l^+ and ν_l are spin 1/2
- Weak decay! W couples to left handed particles, right handed anti-particles
  - ν_l is left-handed
- Angular momentum conservation => l^+ is also left handed
- Violates weak interaction preference that l^+ is right handed! Also, helicity conservation (helicity before = 0, helicity after = -1)
  - => Helicity suppression! If massless, this is prohibited, so decay rate must be proportional to mass
    - => $\pi^{\pm}$ decays more often to muons than electrons





Misc: CKM Matrix

- Describes mixing between quark mass eigenstates
  - For e.g., W decay, widths are proportional to the square of the CKM element
  - For us:
    - $W^+ \to c\bar{b} \propto |V_{cb}|^2 \approx 0.0018$
    - $W^+ \to c\bar{s} \propto |V_{cs}|^2 \approx 0.994$
    - So much more likely to get c, s from top decay in our 4b final state
      - bbcc, rate of c's faking b's is ~10%

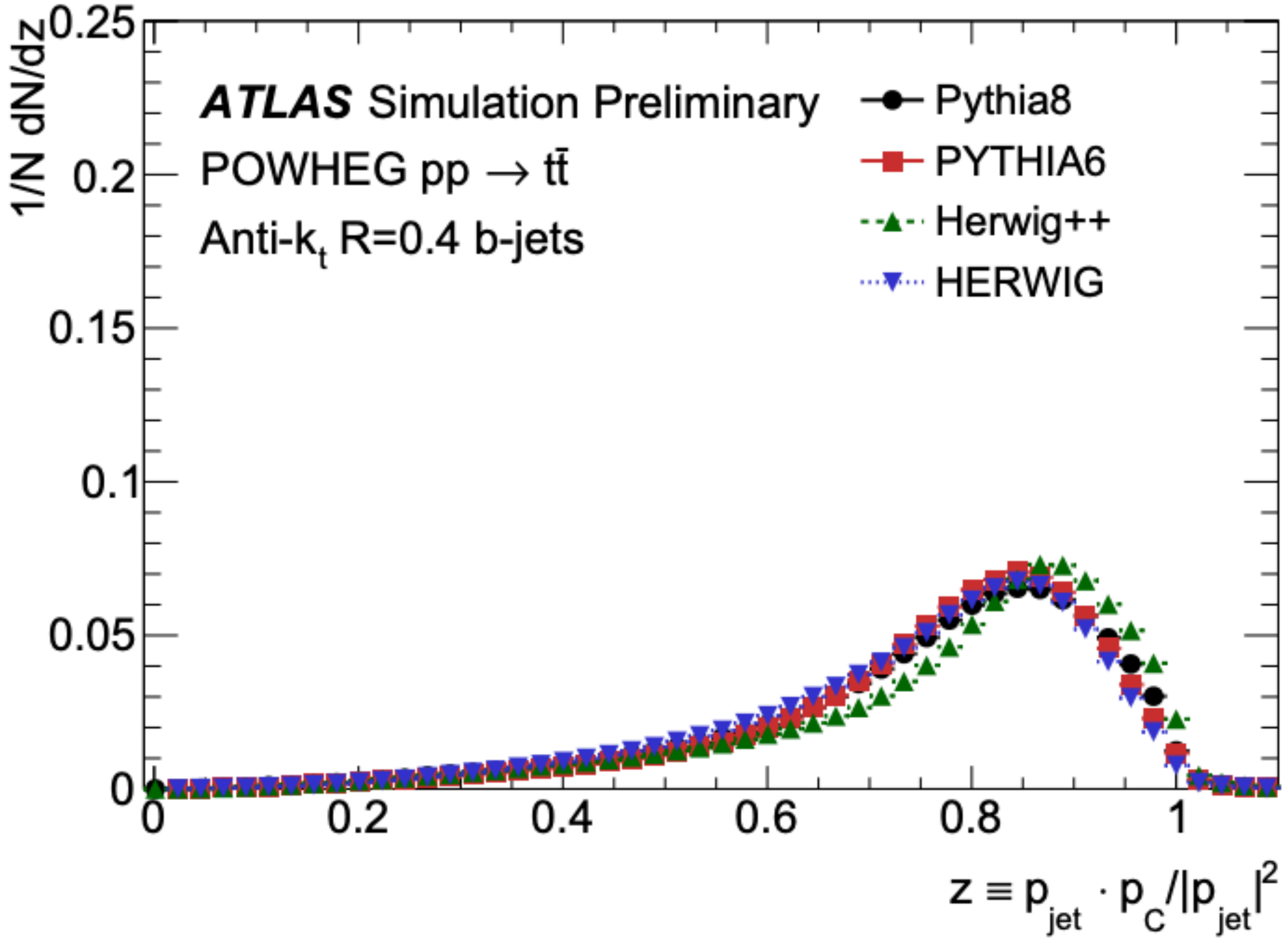


$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} \approx \begin{pmatrix} 0.97420 & 0.2243 & 0.00394 \\ 0.218 & 0.997 & 0.0422 \\ 0.0081 & 0.0394 & 1.019 \end{pmatrix}$$

CKM Values (PDG)

Misc: More on b-tagging

- Long Lifetimes:
  - b-hadron lifetimes ~1.5 ps
  - $c\tau = 1.5 \text{ ps} \cdot 3 \times 10^8 \text{m/s} = 0.00045 \text{ m} = 0.45 \text{ mm}$
  - π^0 e.g. has a lifetime of 8.4e-17 s = 8.4e-5 ps, around 20,000 times shorter!
  - Why? b->c CKM suppressed, e.g.
- Large Masses:
  - B meson mass ~ 5.2 GeV ($B^+ = u\bar{b}$, $B^0 = d\bar{b}$, e.g.)
  - π^0 mass ~ 134 MeV = 0.134 GeV
  - => High multiplicity. b hadrons have an average of 5 charged particles per decay
- Hard fragmentation function:
  - b-hadrons contribute to an average around 75% of the b-jet energy

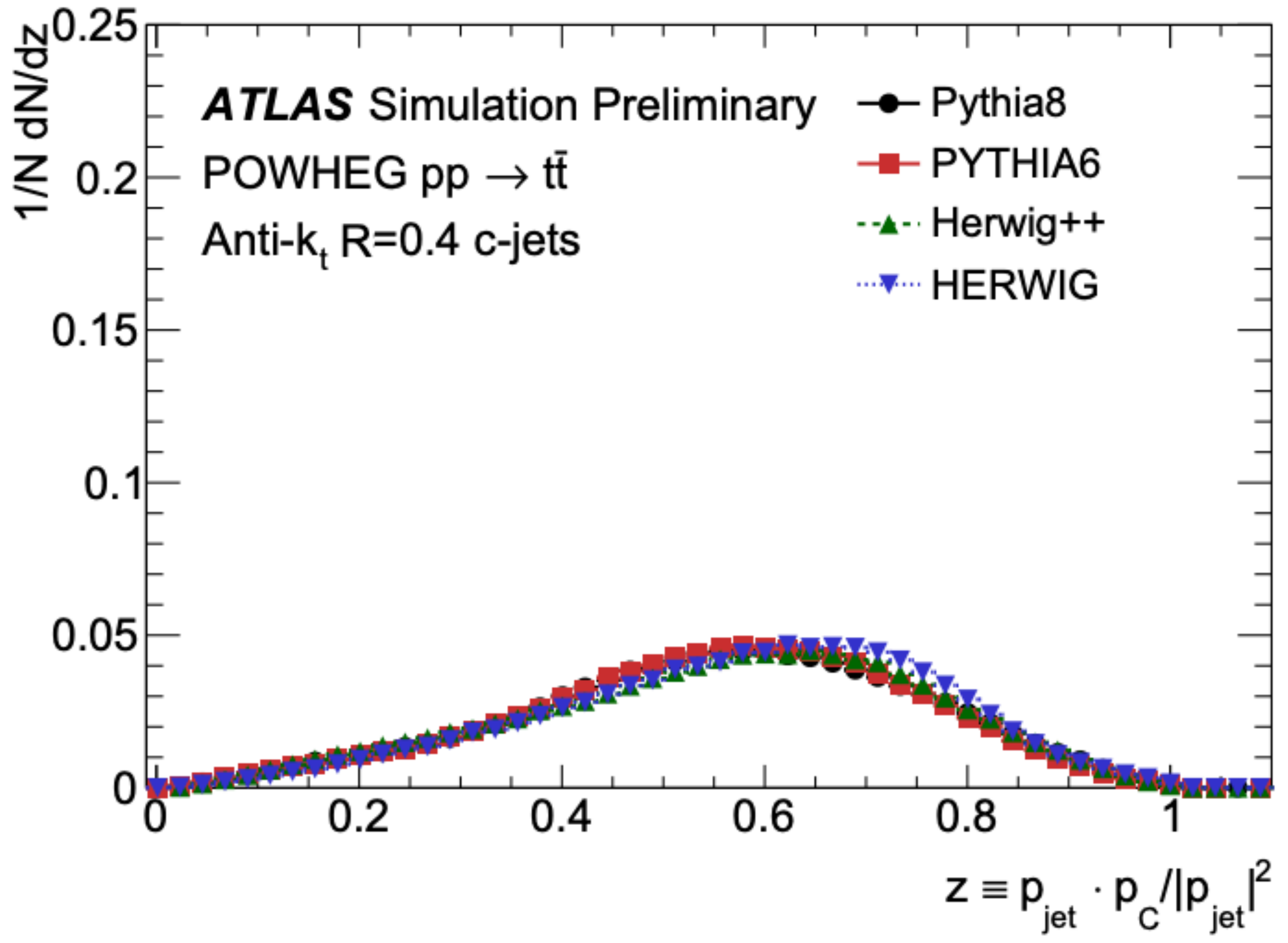


Flavor tagging paper

Flavor tagging introduction

Misc: b vs c

- c-hadrons ($D^+ = c\bar{d}$, $D^0 = c\bar{u}$, e.g.)
  - Lifetime ~ 0.5-1 ps (factor of 2 smaller than b-hadron)
  - Mass ~1.9 GeV (~2-3x smaller than b-hadron)
  - Fragmentation: c-hadrons carry an average of ~55% of jet energy



Flavor tagging paper

Flavor tagging introduction

Misc: b vs c and MV2c10

- MV2c10 is a BDT based tagger
  - Takes as inputs lower level tagging information
    - Impact parameter based
    - Secondary vertex finding
    - Jet properties
- Outputs a b-tagging score, used for selection
  - We use the 70% working point
    - 70% of the real b-jets will be tagged as such
    - ~11% of c-jets will be pass the b-jet selection
    - ~0.33% of light jets will pass

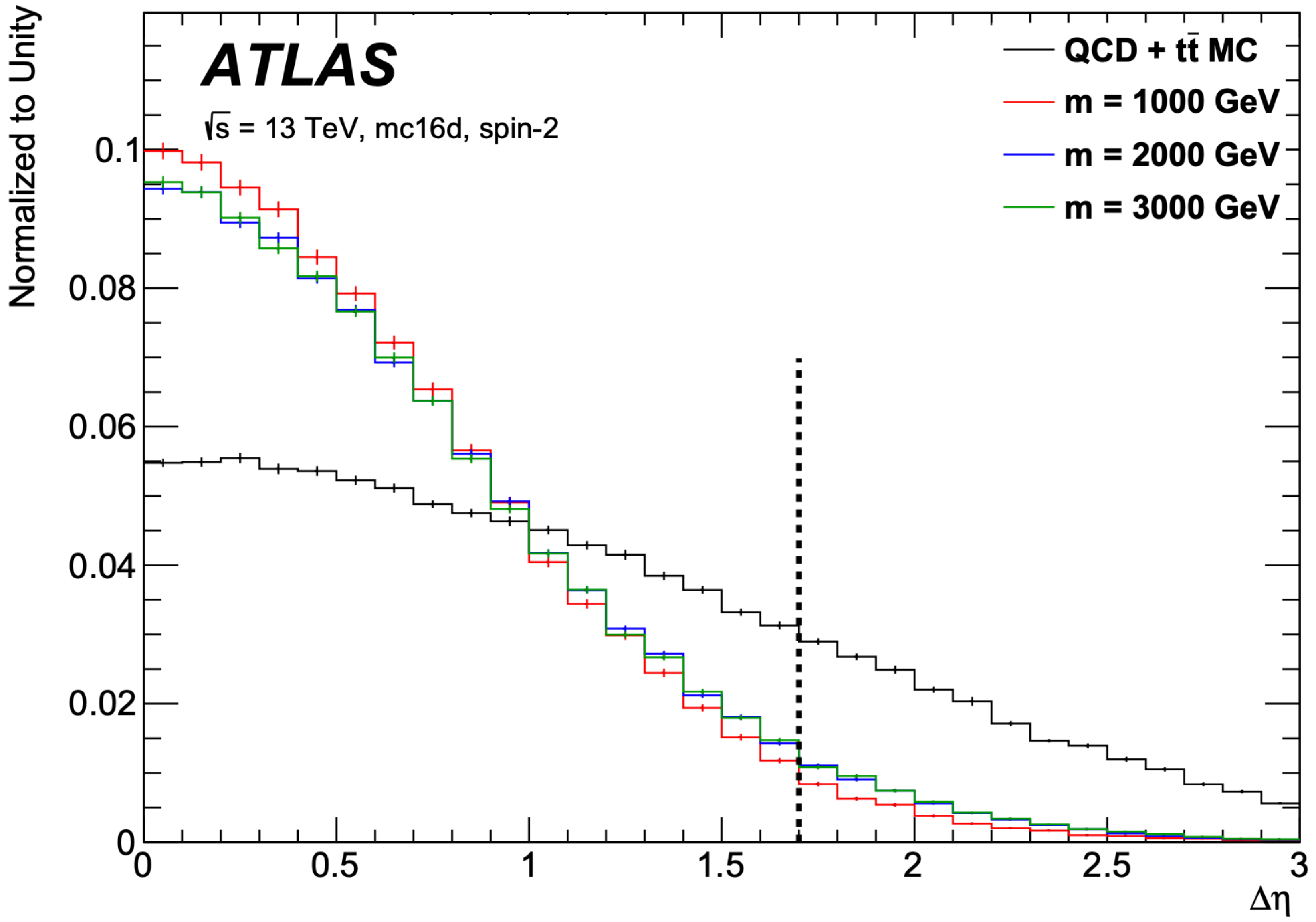
ϵ_b	MV2			
	Selection	Rejection		
		c-jet	τ-jet	Light-flavour jet
60%	> 0.94	23	140	1200
70%	> 0.83	8.9	36	300
77%	> 0.64	4.9	15	110
85%	> 0.11	2.7	6.1	25

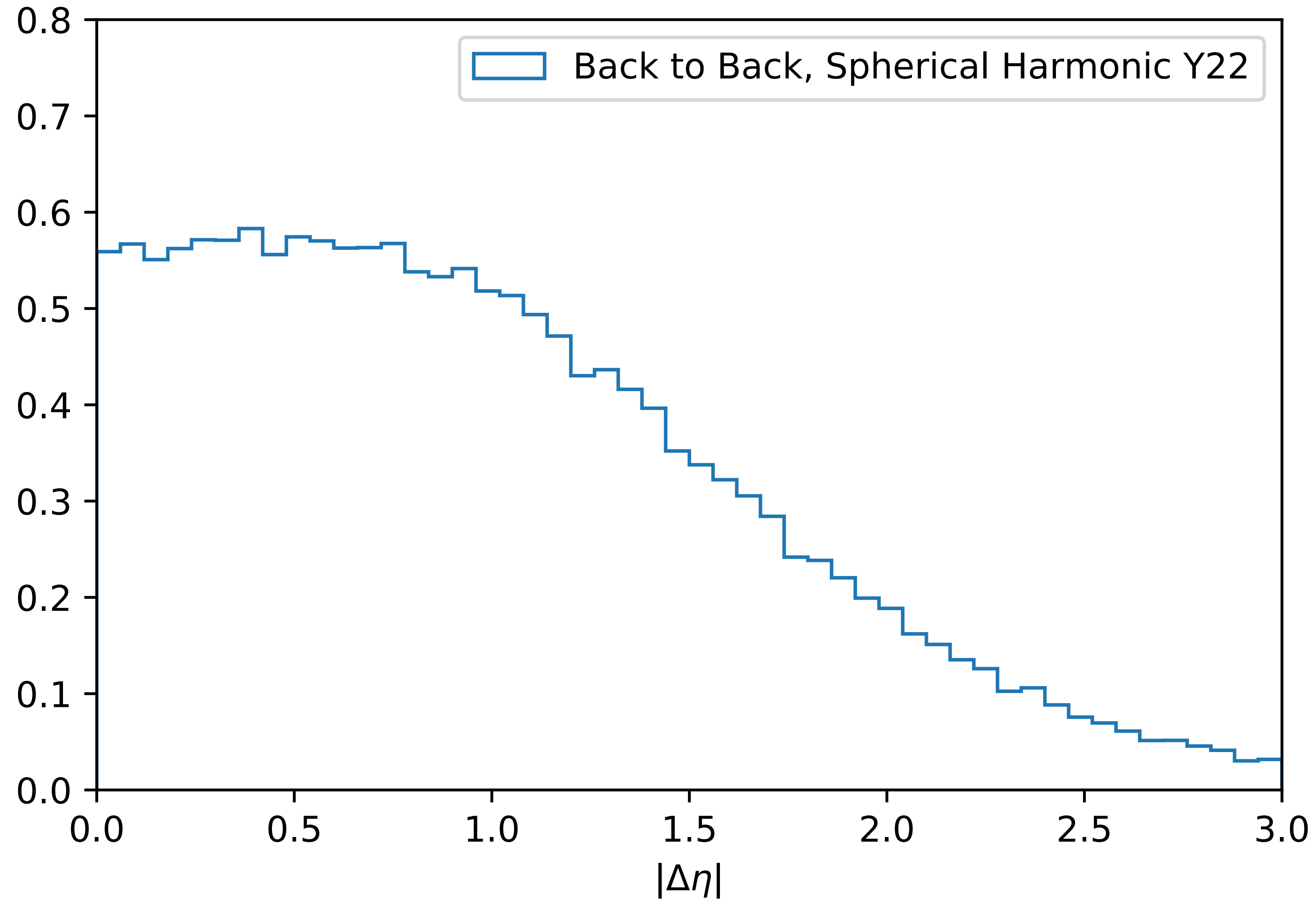
b-tagging efficiency and rejection (1/eff for each category)

Flavor tagging paper

Flavor tagging introduction

Misc: Angular Distributions





- Very roughly, we would expect spin 0 to be isotropic (proportional to $|Y_0^0|^2$), spin 2 to have some angular dependence (which might go like $|Y_2^2|^2$)
- Checking this, we sample from corresponding distributions and look at $|\Delta\eta|$
- Uniform (on right) matches scalar shape well (see, e.g. green "back to back") though the shape is quite different for m=1000 GeV
  - Assumption of at rest production may break down
- Spherical harmonic shape matches spin 2 a bit worse, but we would expect some mixture of other harmonics here. We do see an inflated low $|\Delta\eta|$, consistent with observed, relative to the scalar