MET Significance Study

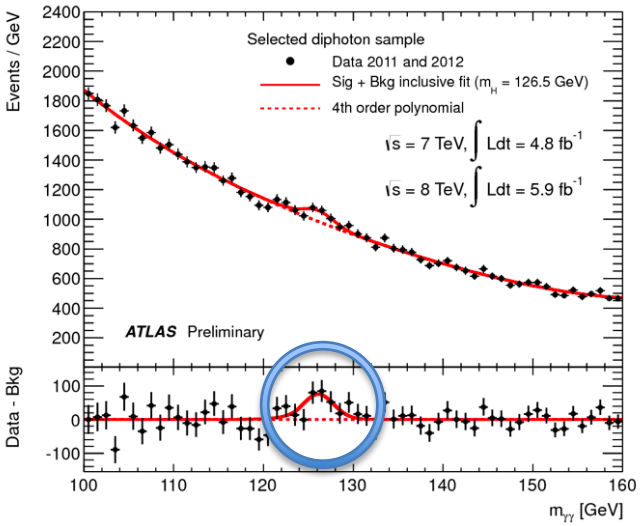
Julia Hinds
CMSE 802 - 001
Michigan State University
April 26th, 2022

Background

- What is particle physics?

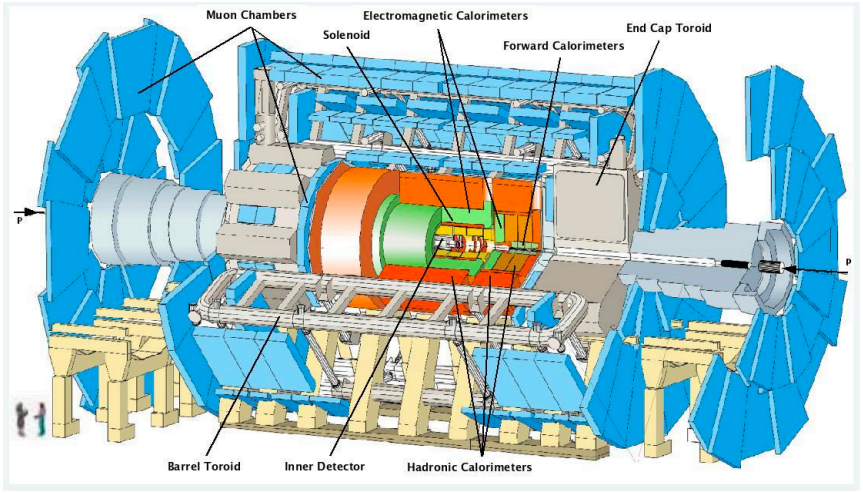
"Like giant toddlers … smashing things together to see what comes out"

-Anonymous



Background

- The European Organization for Nuclear Research (CERN), the Large Hadron Collider (LHC), and ATLAS (A Toroidal LHC ApparatuS)

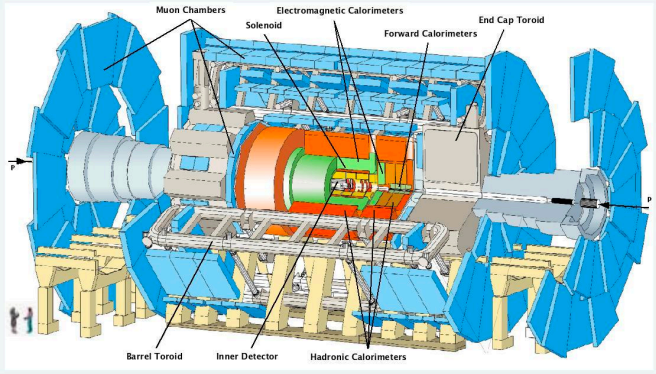


Forbes Magazine. https://www.forbes.com/sites/startswithabang/2021/09/07/the-3-reasons-why-cerns-large-hadron-collider-cant-make-particles-go-faster/?sh=3e360abf2aae [Accessed October 6, 2021].

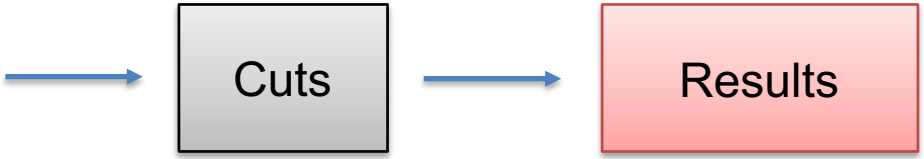
UCI Energy Blogs. https://sites.uci.edu/energyobserver/2012/11/26/introduction-to-the-atlas-detector-at-the-lhc/ [Accessed October 6, 2021].

Motivation

- Improve cuts which improve results!



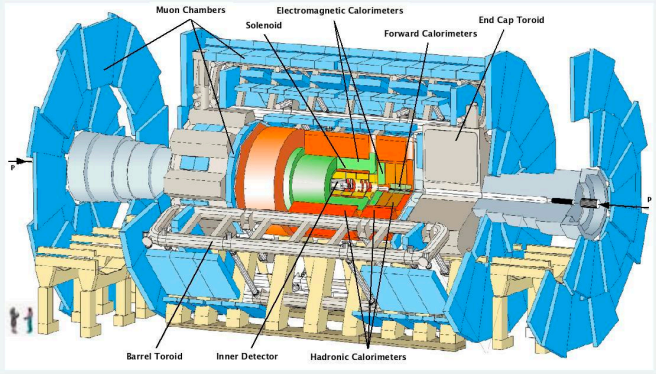
Raw Data



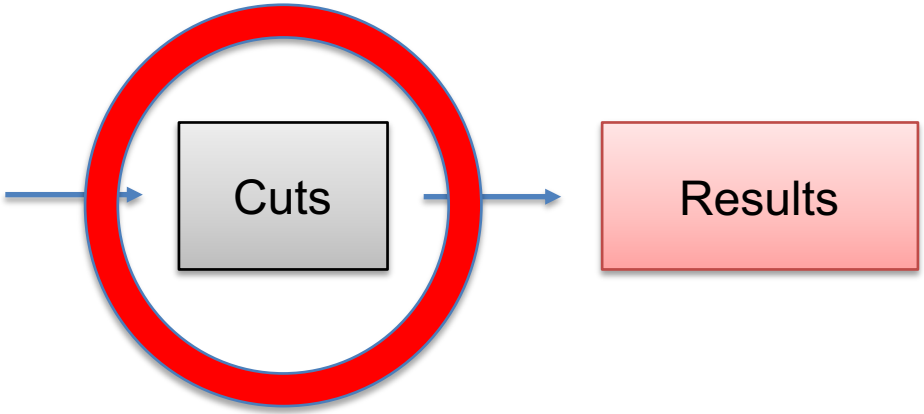
UCI Energy Blogs. https://sites.uci.edu/energyobserver/2012/11/26/introduction-to-the-atlas-detector-at-the-lhc/ [Accessed October 6, 2021].

Motivation

- Improve cuts which improve results!



Raw Data

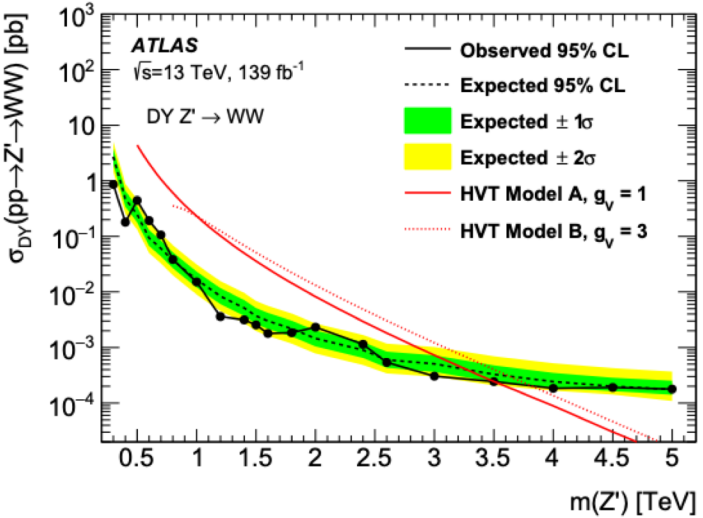


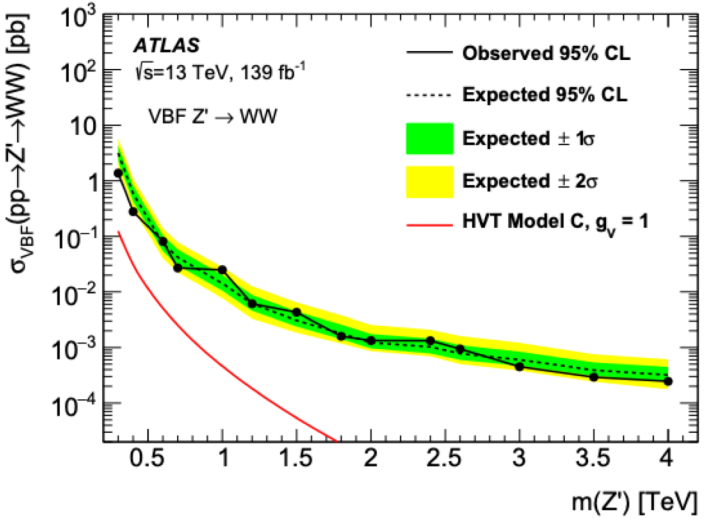
My work applies here!

UCI Energy Blogs.
https://sites.uci.edu/energyobserver/2012/11/26/introduction-to-the-atlas-detector-at-the-lhc/ [Accessed October 6, 2021].

Motivation

- What would improved sensitivity look like?
  - Example: previous analysis results
  - https://arxiv.org/abs/2004.14636





Motivation

- My goal is to apply cuts to the MET Significance (one of the kinematic variables) to try to optimize the signal significance.
  - Introducing a new parameter to clean up end results

Code and Documentation

- Languages and packages
  - C++, ROOT, and Python were used
  - Matplotlib (Jupyter Notebook) and ROOT histograms were used for visualization

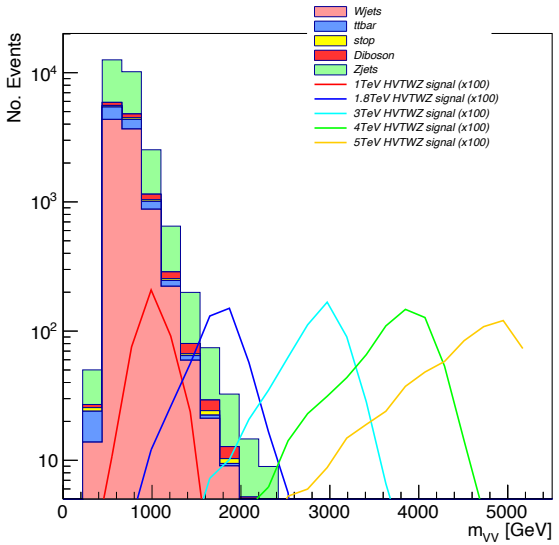
- Analysis Framework
  - METSignificanceStudy
    - https://gitlab.msu.edu/hindsju1/metsignificancestudy
  - Used to study the METSignificance variable and how the cuts impact the significance
  - Code developed in parallel for this project

Code and Documentation

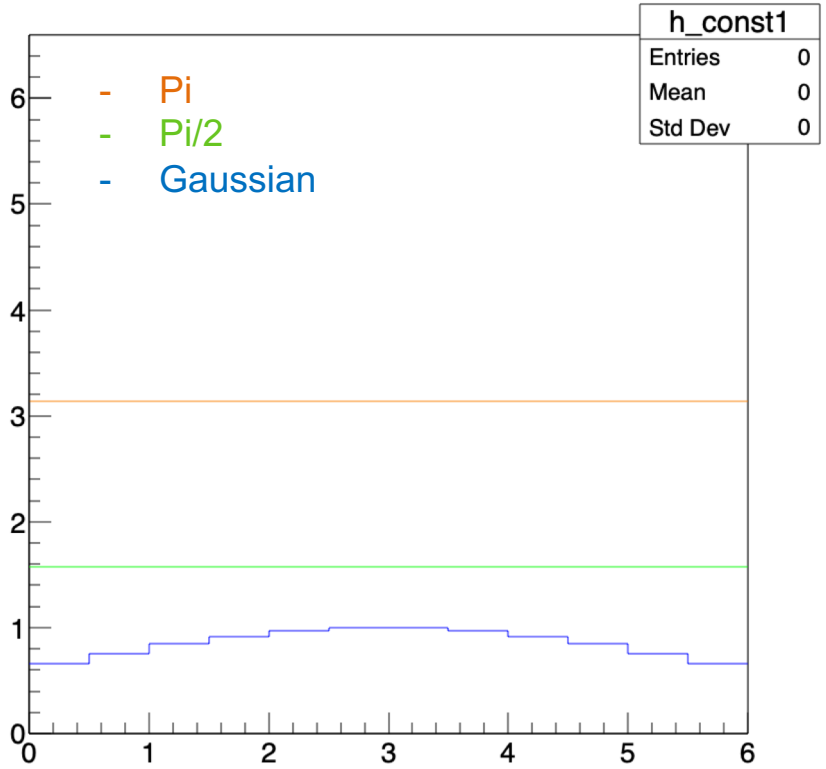
- Analysis Framework
  - CutFlowProduction.C
    - Code that produces the cut flow of the analysis
    - The cut flow is a representation of how many events pass each sequential cut
  - VV_VH_distributions_plotting.ipynb
    - Jupyter Notebook that plots 1D and 2D kinematic distributions
    - Used primarily for visualizations
  - METSigStudy.C
    - Main calculation of significance done
    - Plots the significance vs METSig cut for different signal masses

Significance Calculation

- Simplified Significance Equation: $\sigma = \sum \sigma_i^2$
  - Where $\sigma_i = \frac{s_i}{\sqrt{s_i + b_i + \delta^2}}$ significance per bin; and, s_i = signal per bin, b_i = total background per bin, δ is the uncertainty in the bin

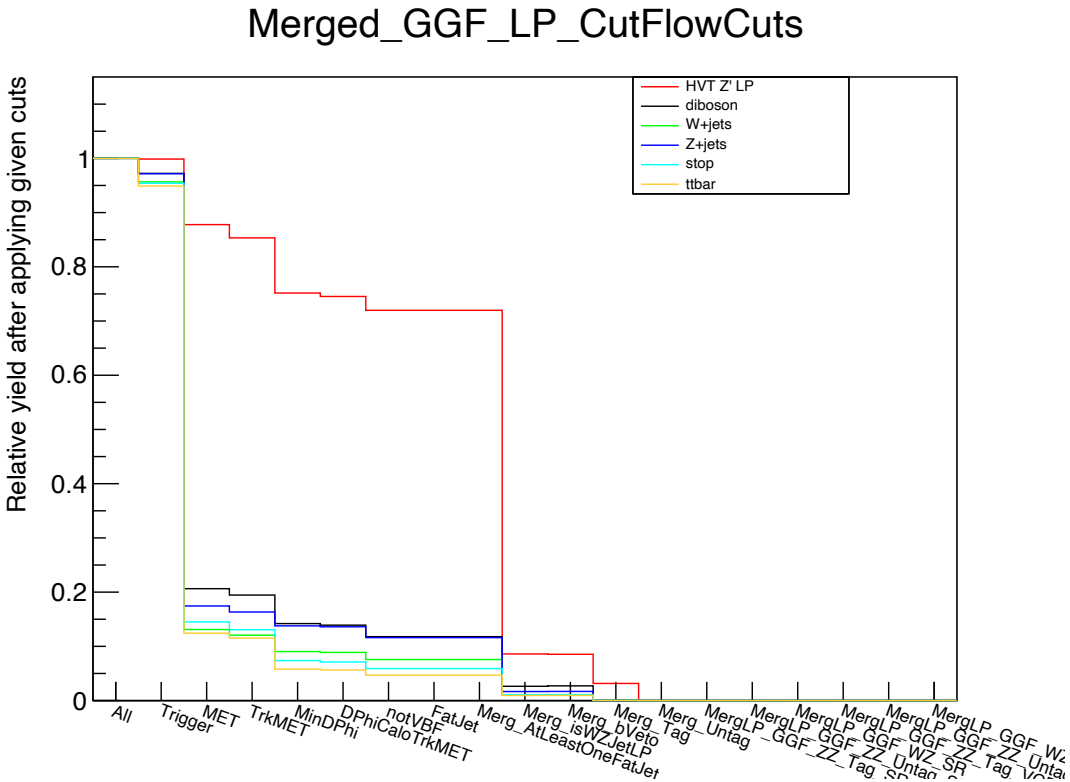


“Unit testing”



- Unit testing for ROOT and C++ is very difficult – created own “pseudo-tests”
  - Tested edge cases (empty histograms being passed, signal histogram not passed, etc.)
  - Tested values for specific cases
    - Used generic ROOT histograms and calculated the significance by hand and compared

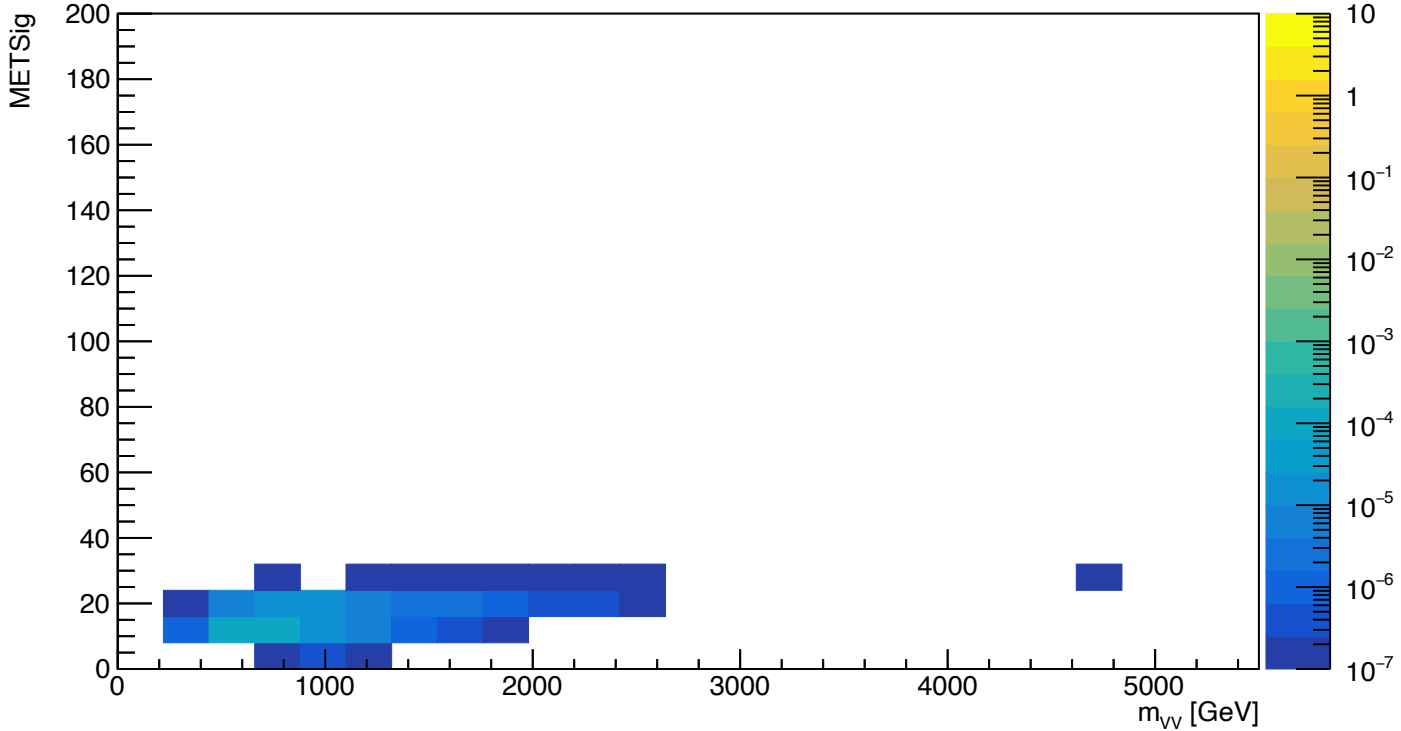
Results – VV PFlow+UFO HVT LP no-tag



Results – VV PFlow+UFO HVT LP no-tag

MICHIGAN STATE UNIVERSITY

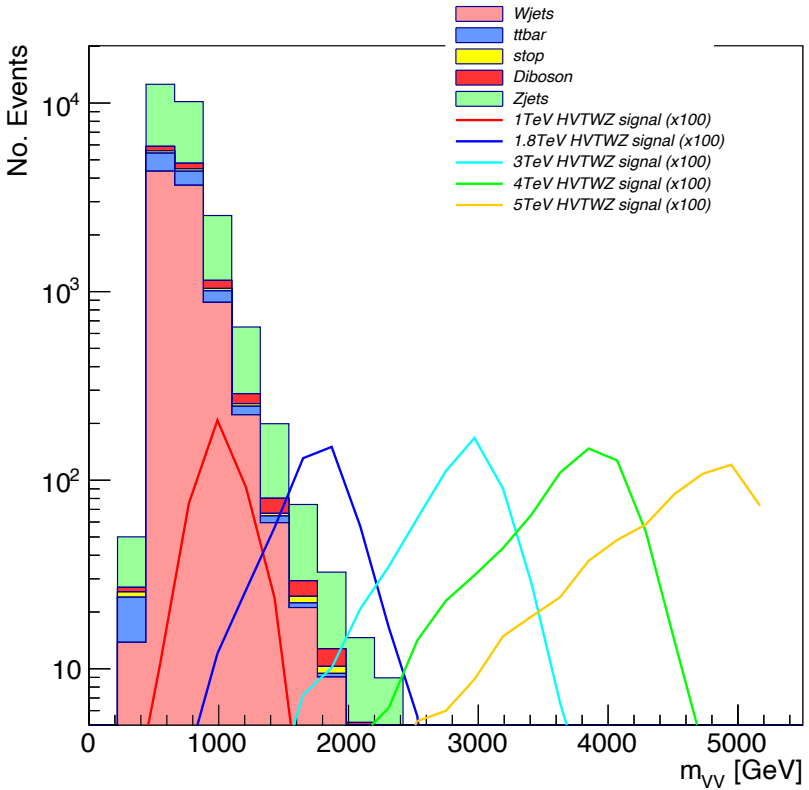
Z+jets only



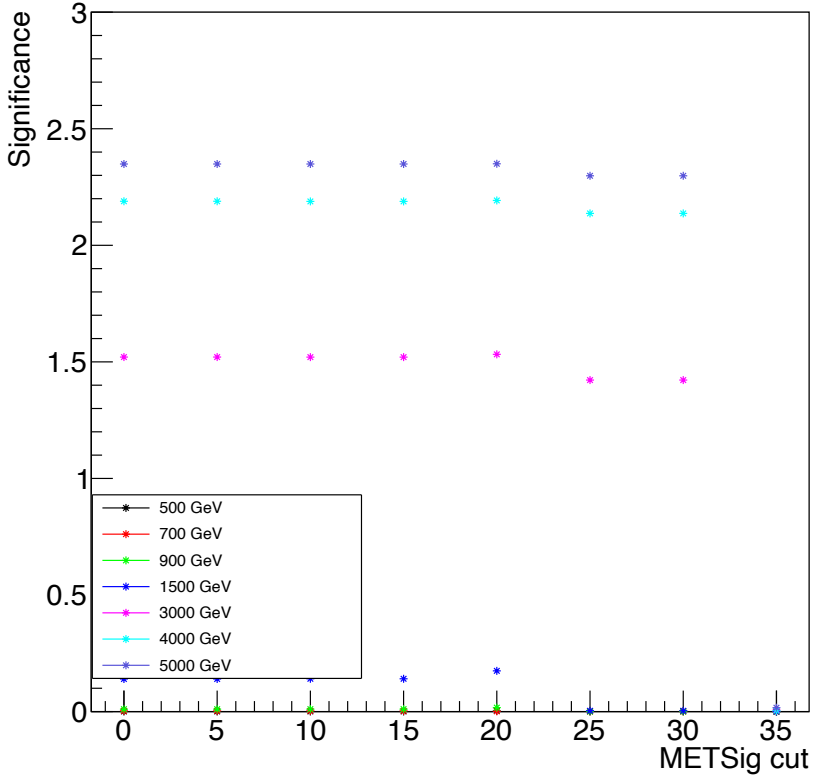
Results – VV PFlow+UFO HVT LP no-tag

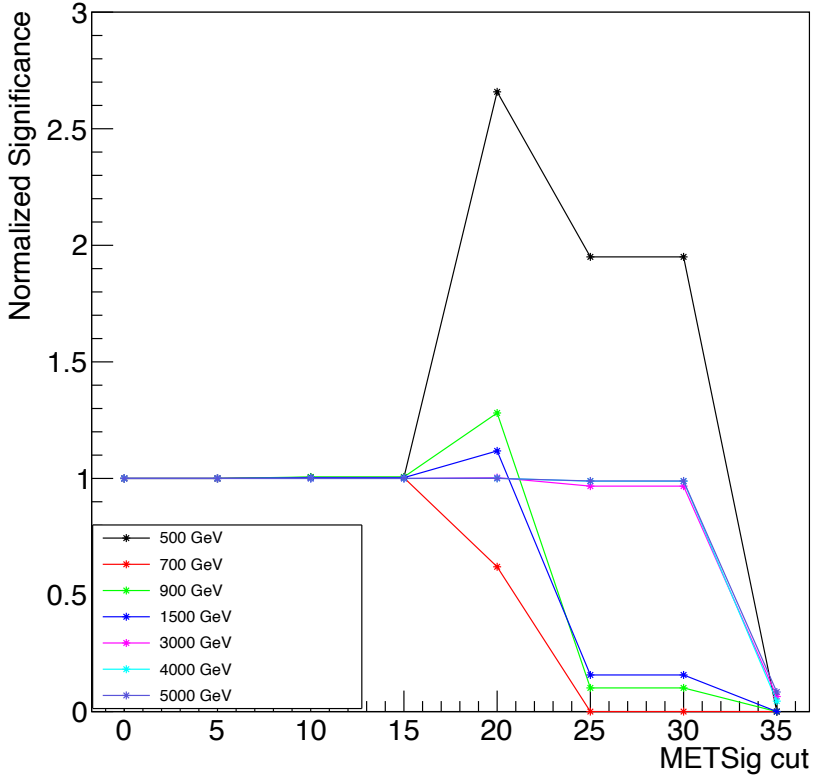
MICHIGAN STATE UNIVERSITY

METSig cut = 0



Results – VV PFlow+UFO HVT LP no-tag



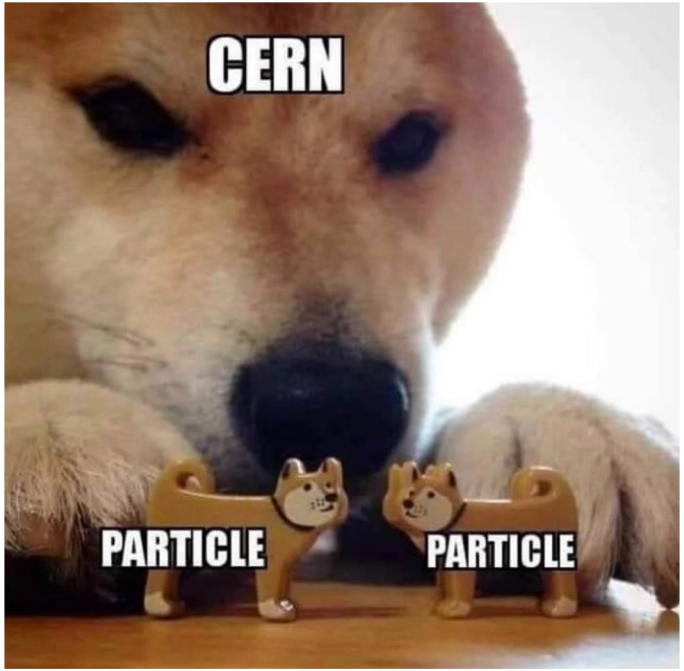


Conclusion and Next Steps

- Developed code to analyze MET Significance cuts affect on the signal significance
- Possible improved sensitivity between a METSig cut of 15 – 25

- Next Steps,
  - Produce finer binning to find an approximate cut
  - Apply ML to see if it can find an optimal cut with all signal masses
  - Replace previous cuts with this improved cut to see if sensitivity is gained

Questions?

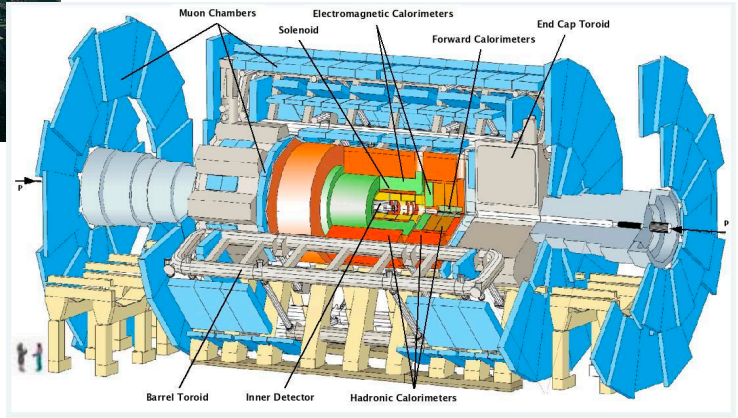


Back up

Background

- The European Organization for Nuclear Research (CERN), the Large Hadron Collider (LHC), and ATLAS (A Toroidal LHC ApparatuS)

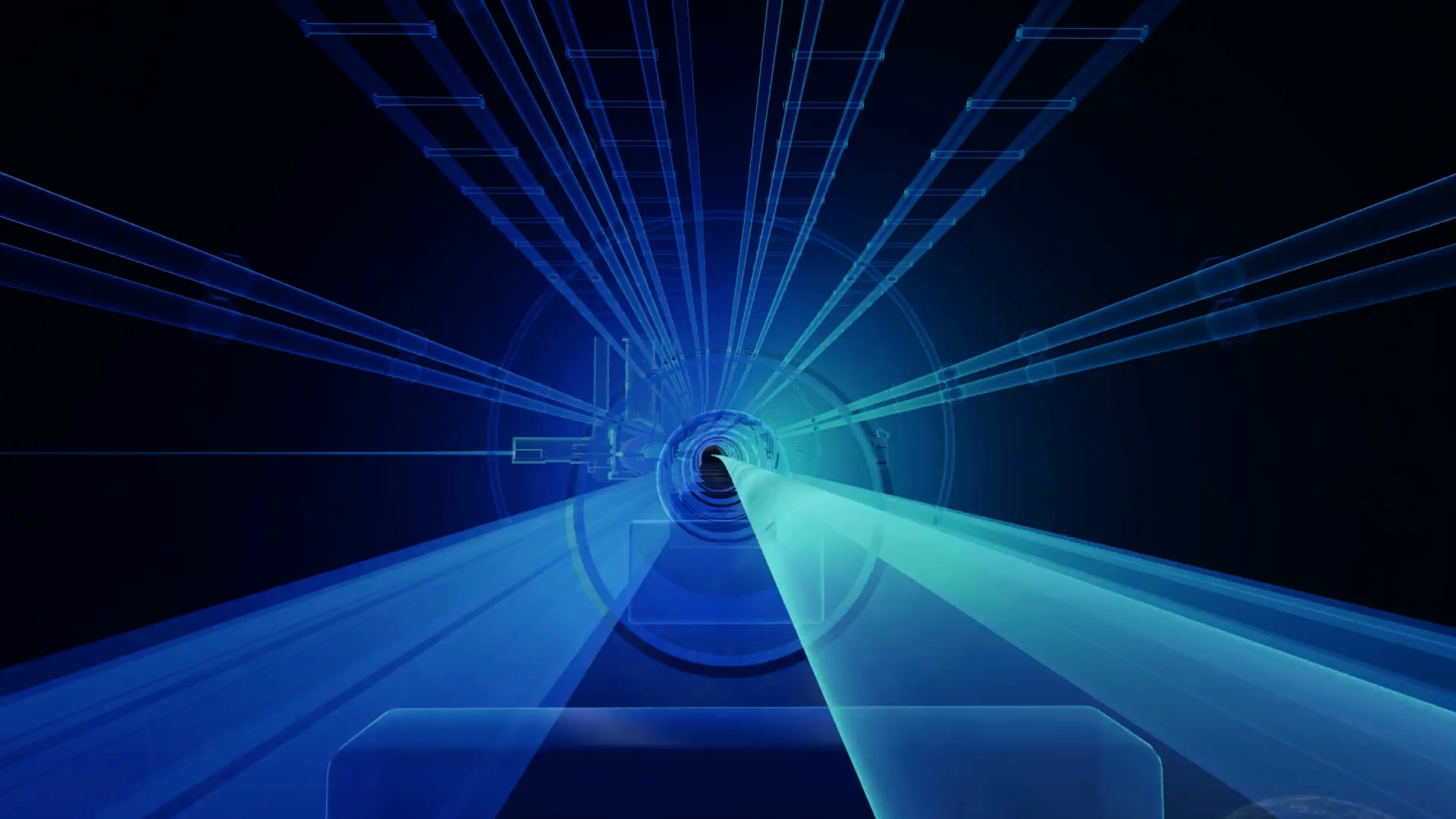
- CERN & LHC
  - Largest particle physics laboratory in the world
  - 27 kilometer circumference
  - 4 major detectors (**ATLAS**, CMS, ALICE, LHCb)
  - No. collisions per second: 1 billion



- ATLAS
  - General purpose detector at the LHC
  - 45 meters in length
  - 25 meters in diameter
  - Weighs over 7000 tons
  - 100 meters below ground

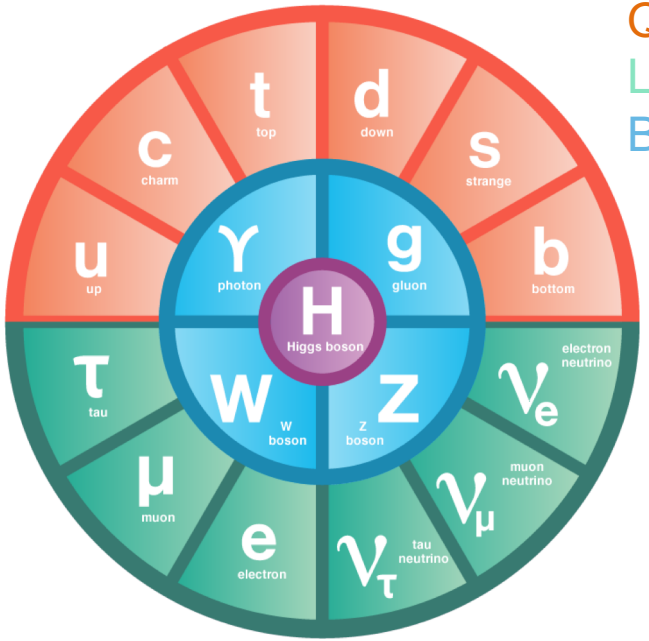
Forbes Magazine. https://www.forbes.com/sites/startswithabang/2021/09/07/the-3-reasons-why-cerns-large-hadron-collider-cant-make-particles-go-faster/?sh=3e360abf2aae [Accessed October 6, 2021].

UCI Energy Blogs. https://sites.uci.edu/energyobserver/2012/11/26/introduction-to-the-atlas-detector-at-the-lhc/ [Accessed October 6, 2021].



Background

- Standard model (SM) theory
  - Theory that describes all of the fundamental particles that we have observed so far
  - Can be categorize and describe interactions
- Need for beyond standard model (BSM)
  - Are we missing any particles?
  - Is there a gravity force carrier?
  - How does dark matter fit in?
  - Is there anything more fundamental?
  - Is there a point where we have all the forces unified together?

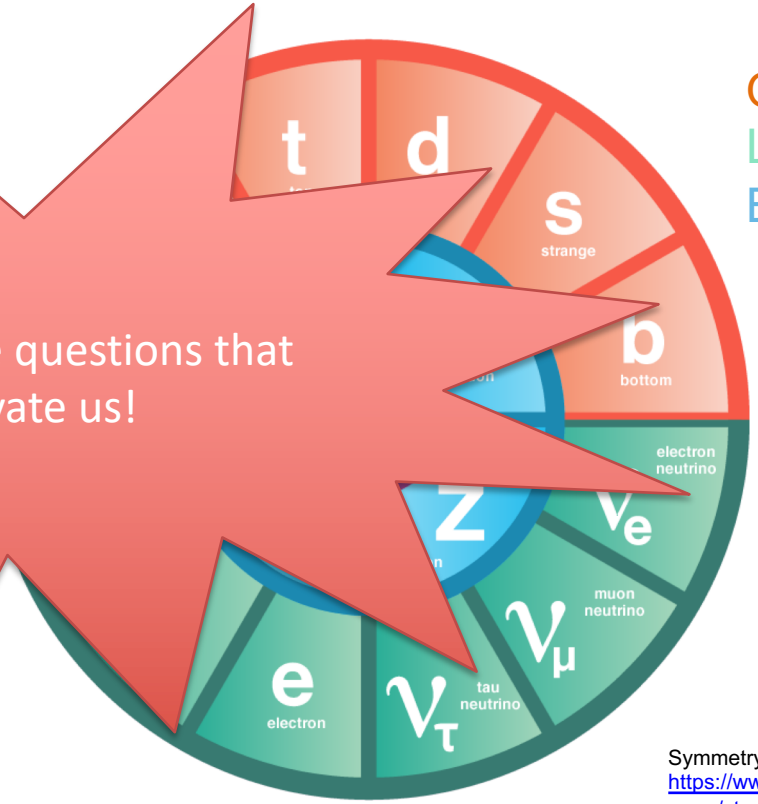


Symmetry Magazine. https://www.symmetrymagazine.org/standard-model/ [Accessed October 6, 2021].

Background

- Standard model
  - Like the per[illegible]table for fundamental pa[illegible]
- Questions you could[illegible]
  - Are we mis[illegible]
  - Is there a gravity [illegible]
  - How does d[illegible]
  - Is there anything more fundamental?
  - Is there a point where we have a[illegible] the forces unified together?

These are the questions that motivate us!



Quarks
Leptons
Bosons

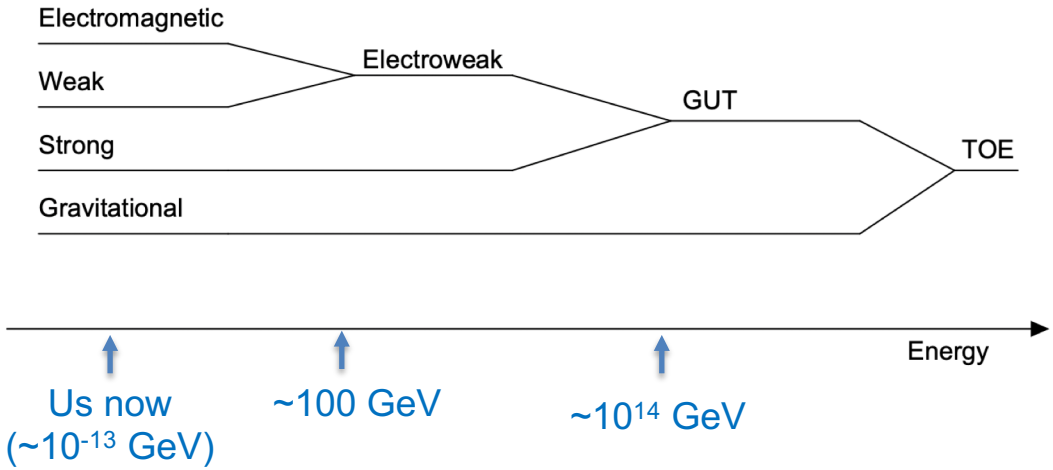
Symmetry Magazine. https://www.symmetrymagazine.org/standard-model/ [Accessed October 6, 2021].

Motivation

- My questions: Is there a point where we have all the SM forces unified together? Is there a gravity force carrier?
- To answer this, we'll need new physics models (experimental inputs)
  - Grand Unified Theory
  - Quantum gravity/Graviton
- We need experimental input to understand these models better
  - Maybe exclude some of the theories (need measurements to determine which theories are wild, and which ones are wild and plausible!)
  - Important to note, even if we don't find something, that can still help exclude certain parts of the theory!

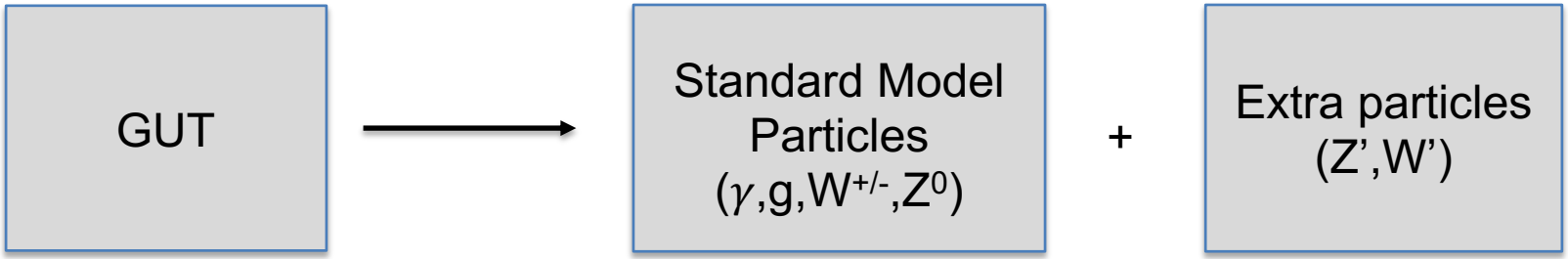
Motivation

- Question: Is there a point where we have all the SM forces unified together?
- Grand Unified Theory (GUT)
  - A model where at high energies, the electromagnetic, weak, and strong forces merge into a single force



Motivation

- Question: Is there a point where we have all the SM forces unified together?
- Grand Unified Theory (GUT)
  - A model where at high energies, the electromagnetic, weak, and strong forces merge into a single force



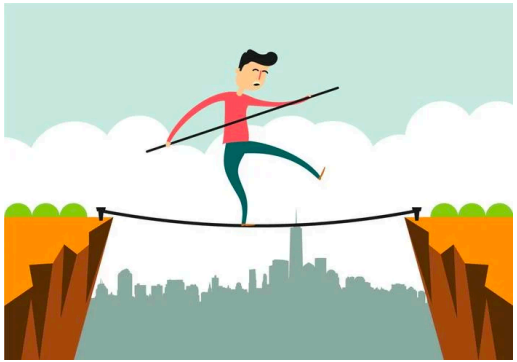
- Observing these new particles could provide evidence of unification at a higher energy

Motivation

- Question: Is there a gravity force carrier?
- Quantum Gravity/Graviton
  - Hypothetical spin-2 particle that mediates the gravitational force
  - But first, let's talk about dimensionality…

Motivation

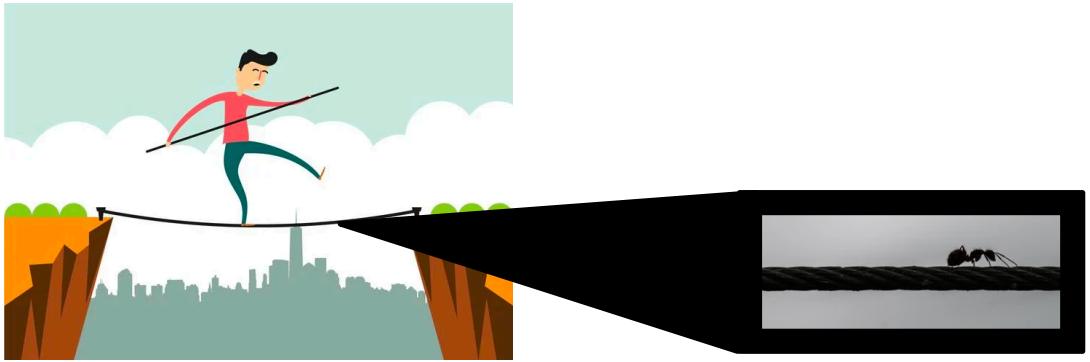
- Question: Is there a gravity force carrier?
- Quantum Gravity/Graviton
  - Hypothetical spin-2 particle that mediates the gravitational force
  - But first, let's talk about dimensionality…



Vecteezy. https://www.vecteezy.com/vector-art/152799-tightrope-walker-over-city-buildings [Accessed October 13 2021].

Motivation

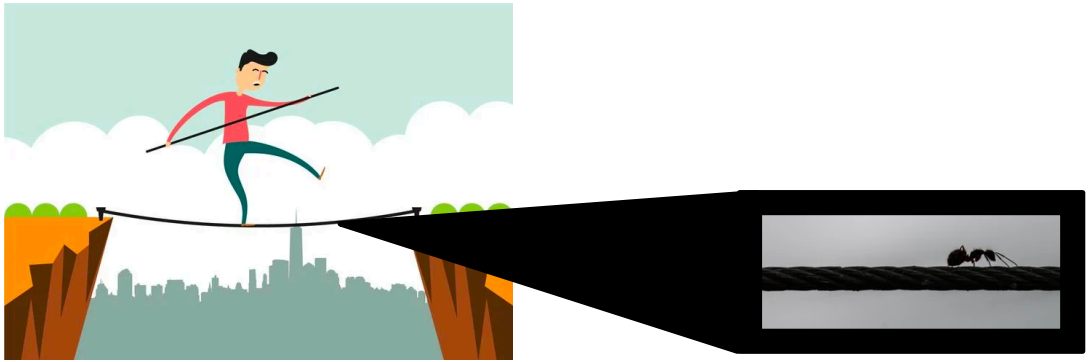
- Question: Is there a gravity force carrier?
- Quantum Gravity/Graviton
  - Hypothetical spin-2 particle that mediates the gravitational force
  - But first, let's talk about dimensionality…

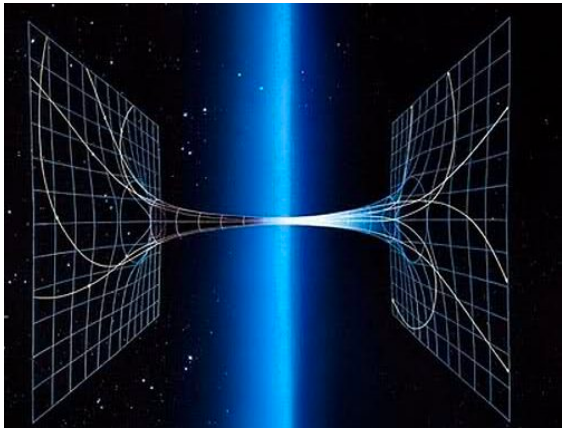


Vecteezy. https://www.vecteezy.com/vector-art/152799-tightrope-walker-over-city-buildings [Accessed October 13 2021].
Shutterstock. https://www.shutterstock.com/image-photo/ant-on-tightrope-1445223776 [Accessed October 13 2021].

Motivation

- Question: Is there a gravity force carrier?
- Quantum Gravity/Graviton
  - Hypothetical spin-2 particle that mediates the gravitational force
  - But first, let's talk about dimensionality…

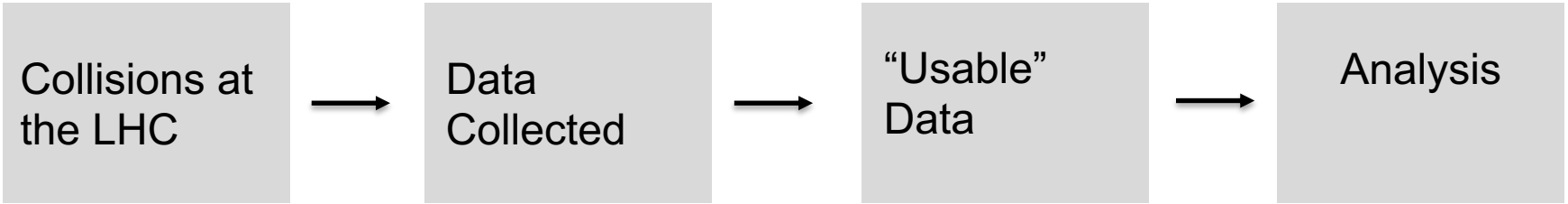




Vecteezy. https://www.vecteezy.com/vector-art/152799-tightrope-walker-over-city-buildings [Accessed October 13 2021].
Shutterstock. https://www.shutterstock.com/image-photo/ant-on-tightrope-1445223776 [Accessed October 13 2021].
Scribal Multiverse. http://www.scribalmultiverse.com/brane-worlds-and-higher-dimensions/ [Accessed October 13 2021].

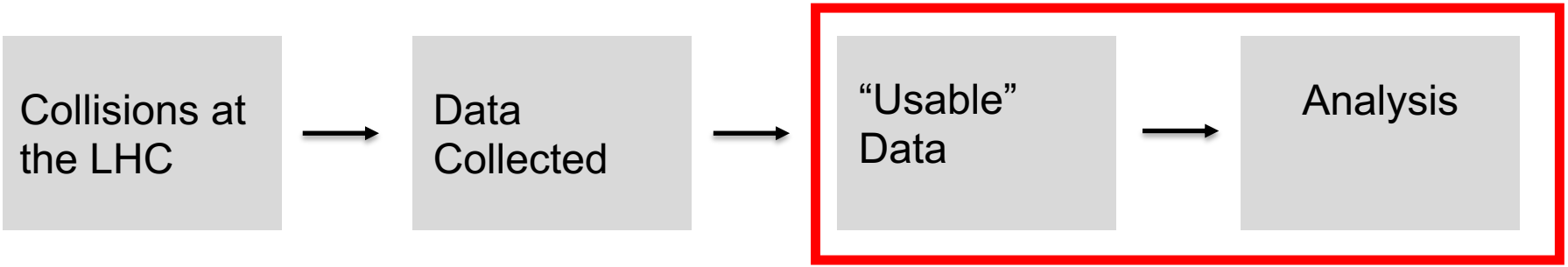
Motivation

- How do you “detect” this new physics?



Motivation

- How do you “detect” this new physics?



Motivation

VV-Semileptonic Analysis

- Search for new particles which decay into a semileptonic final state
- VV – one hadronic and one leptonic decay
  - Where VV can be WW, WZ, or ZZ (V = W,Z)
  - Hadronic decay allows for higher cross section and leptonic decay allows for a clearer signal
- Signal models
  - Spin-0: Scalar “Radion”
  - Spin-1: Heavy vector triplet (HVT) W'/Z'
  - Spin-2: Randall-Sundrum Graviton
- Production
  - Gluon-gluon fusion (ggF)
  - Quark-quark annihilation (Drell-Yan)
  - Vector-boson fusion (VBF)

Motivation

VV-Semileptonic Analysis

- Search for new particles which decay into a semileptonic final state
- VV – one hadronic and one leptonic decay
  - Where VV can be WW, WZ, or ZZ (V = W,Z)
  - Hadronic decay allows for higher cross section and leptonic decay allows for a clearer signal
- Signal models
  - Spin-0: Scalar “Radion”
  - Spin-1: Heavy vector triplet (HVT) W'/Z'
  - Spin-2: Randall-Sundrum Graviton

Main focus for now

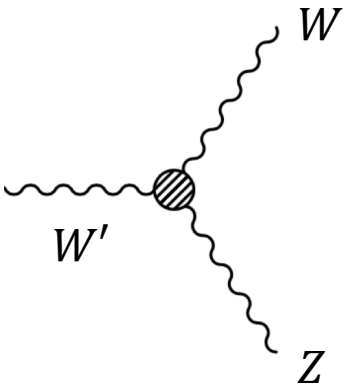
- Production
  - Gluon-gluon fusion (ggF)
  - Quark-quark annihilation (Drell-Yan)
  - Vector-boson fusion (VBF)

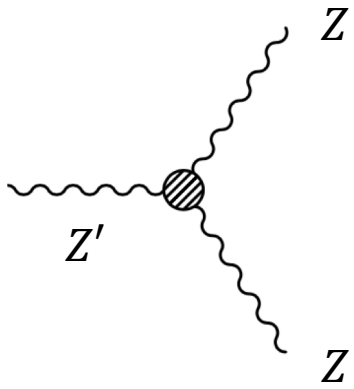
Motivation

- Signal models
  - Heavy vector triplet (HVT) model
    - Heavy gauge bosons (W' , Z')

Three distinct channels:

0-lepton: $ZV \rightarrow \nu\nu qq$
1-lepton: W$V \rightarrow l\nu qq$
2-lepton: Z$V \rightarrow llqq$





$Z \rightarrow \nu\nu$
$Z \rightarrow ll$
$W \rightarrow l\nu$

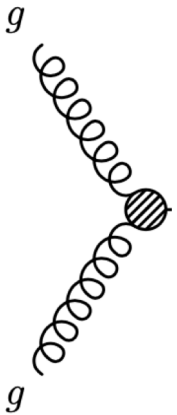
Motivation

VV-Semileptonic Analysis

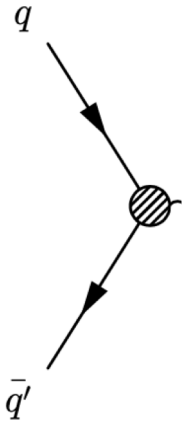
- Search for new particles which decay into a semileptonic final state
- VV – one hadronic and one leptonic decay
  - Where VV can be WW, WZ, or ZZ (V = W,Z)
  - Hadronic decay allows for higher cross section and leptonic decay allows for a clearer signal
- Signal models
  - Spin-0: Scalar “Radion”
  - Spin-1: Heavy vector triplet (HVT) W'/Z'
  - Spin-2: Randall-Sundrum Graviton

- Production
  - Gluon-gluon fusion (ggF)
  - Quark-quark annihilation (Drell-Yan)
  - Vector-boson fusion (VBF)

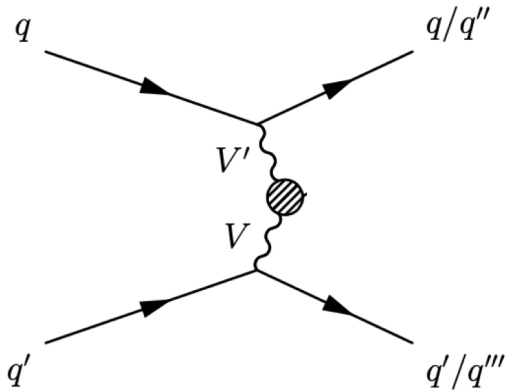
- VV-Semileptonic Analysis



gluon-gluon fusion (ggF)



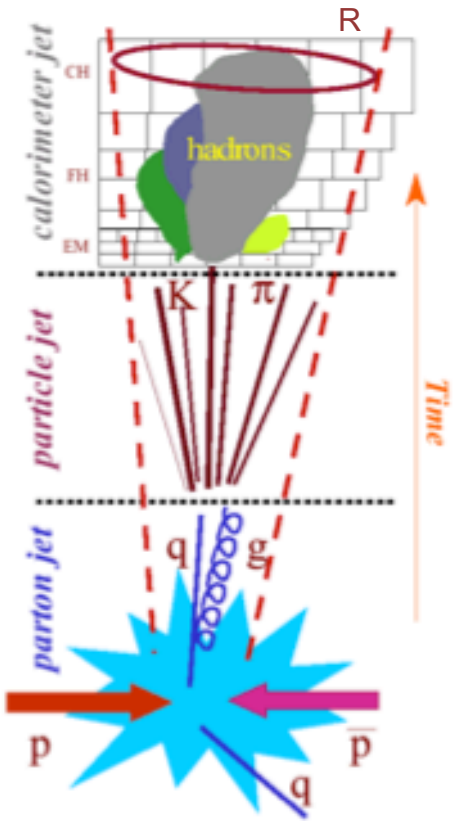
Drell-Yan (DY)



Vector-boson fusion (VBF)

Definitions

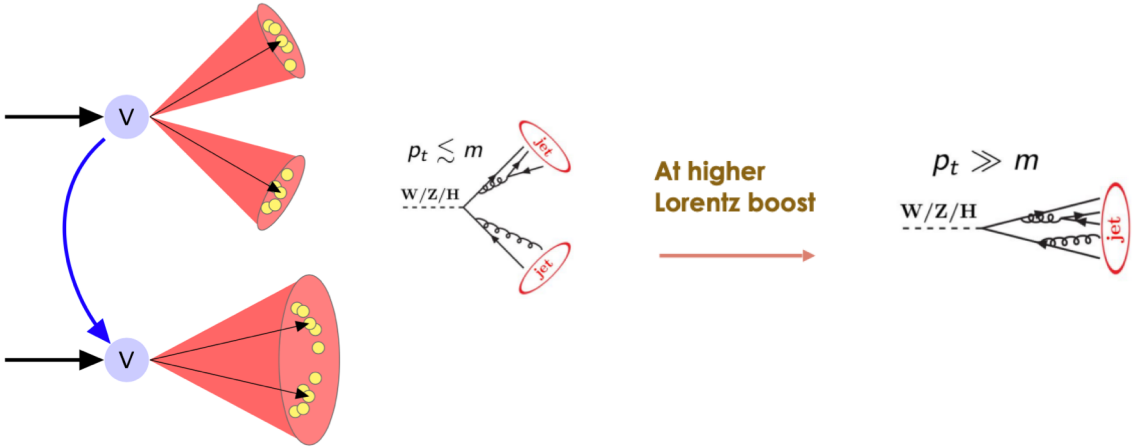
- Collecting final state particles
  - **Jets** are particles collected from the same source
  - Jet size is dependent on cone size, R
    - "rule of thumb" $R = 2m/p_T$
    - Large-R jet R=1.0
    - Small-R jet R=0.4
  - Parton jet: physics interaction
  - Particle jet: added tracks from underlying event
  - Calorimeter: measuring energy of neutral or charged particle



Loch, Peter. US-ATLAS mini-course on Jets.
https://indico.cern.ch/event/1054786/ [Accessed October 1, 2021].

Definitions

- Resolved
  - Two small R-jets
- Merged
  - One large R-jet
  - Only merged used here
- Missing E_T (MET or E_T^{miss})
  - Reconstructed by negative sum of the all physics objects and ID tracks not associated with them in the transverse direction (should sum to zero)
  - Neutrinos don't show up in detector, so when look at energy conservation there will be energy missing, i.e. neutrino energy

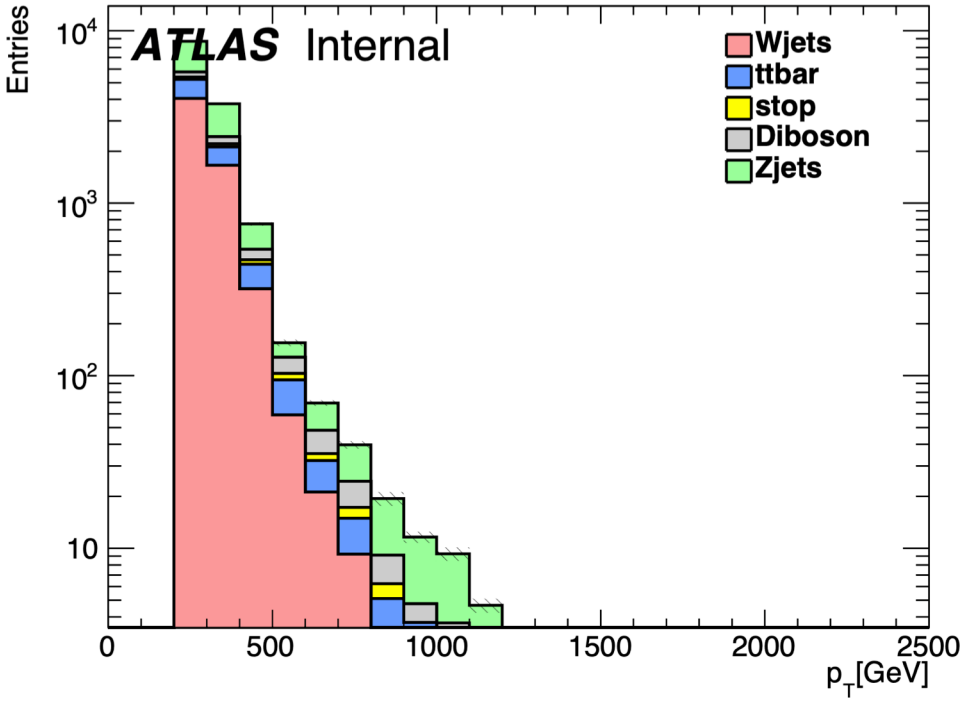


Definitions

- Purity
  - High (HP) or Low (LP) depends on passing and/or failing the boson tagging requirement
- Signal Regions (SR)
  - Try to maximize signal, minimize background
  - Strict cuts, with one of those cuts requiring boson tagging.
- Control Regions (CR or VCR)
  - Regions which look very similar to the signal region but have some slightly different selection to enrich a particular background (i.e. V boson), which is orthogonal to the signal region

Definitions

- Backgrounds



- SM processes can “mimic” signals, i.e. have same final states (pass event selection)

Definitions

- Jet Collections
  - Small-R jet
    - EMTopo (Electromagnetic Topological)
    - PFlow (Particle Flow)
  - Large-R jet
    - TCC (Track Calo Clusters)
    - LCTopo (Local Call Topological)
    - UFO (Unified Flow Object)
  - Compare the 3 large-R jet collections with respect to each other only

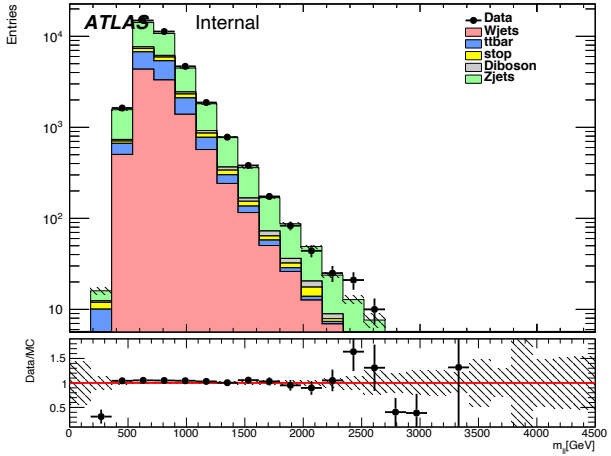
H.C. Berg. Private communication. 2021.

- Previous analysis used EMTopo with LCTopo and TCC, new jet collection algorithm is PFlow with UFO
  - My immediate goal is to determine if there are any improvements to the analysis by switching to Pflow/UFO

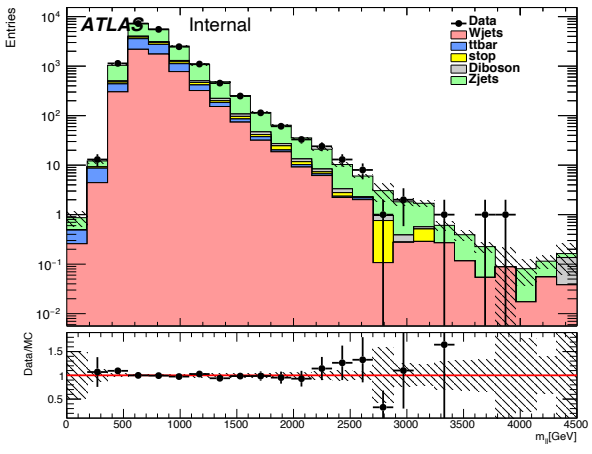
Kinematic comparison of jet collections

- Investigating some of the ways that we can reconstruct jets with the ATLAS detector to see if they bring improvement to the analysis

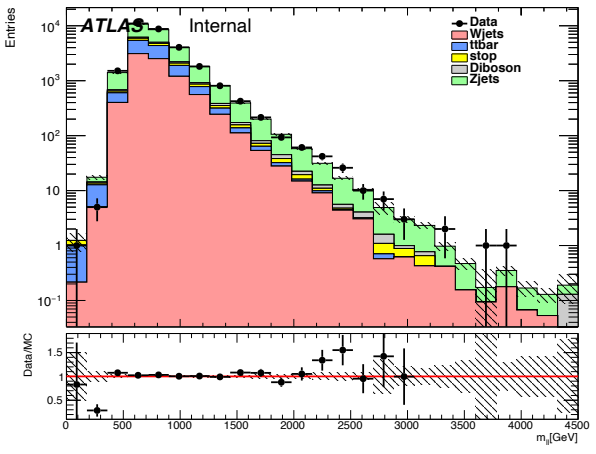
VCR – ggF – HP (Invariant mass)



PFlow UFO



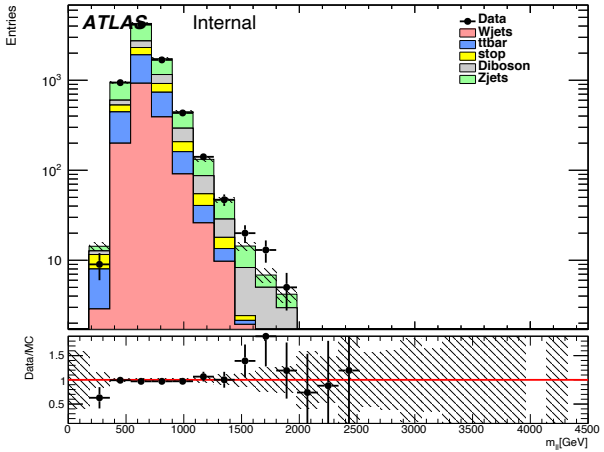
EMTopo TCC



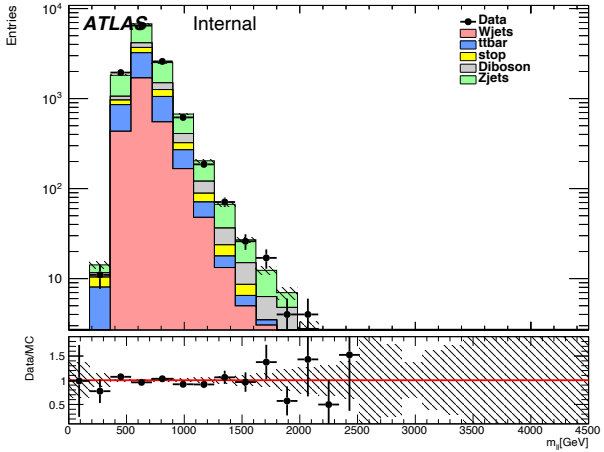
EMTopo LCTopo

Kinematic comparison of jet collections

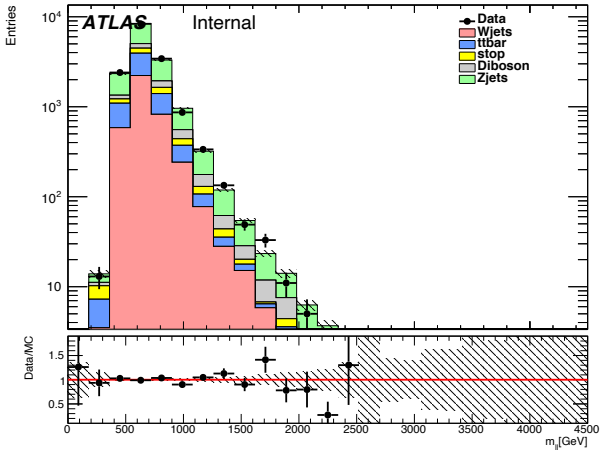
SR – ggF – HP (Invariant mass)



PFlow UFO



EMTopo TCC



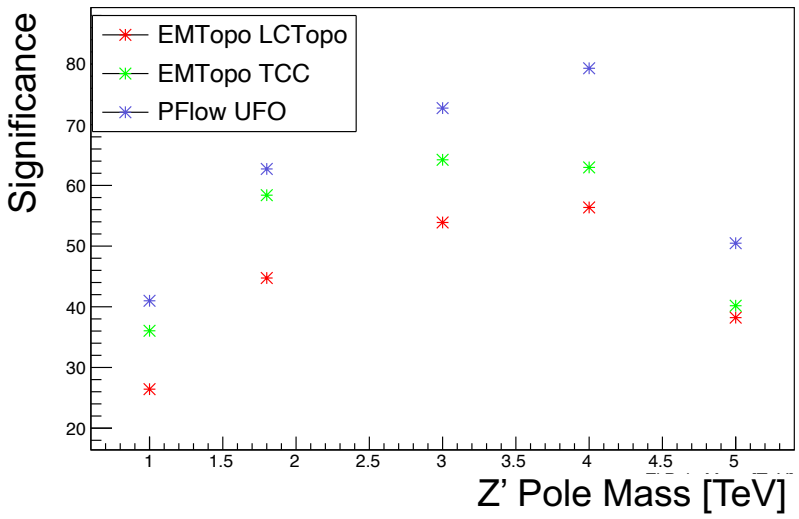
EMTopo LCTopo

- Also compared p_T, m_{jj}, D_2, MET for ggF and VBF (HP and LP)

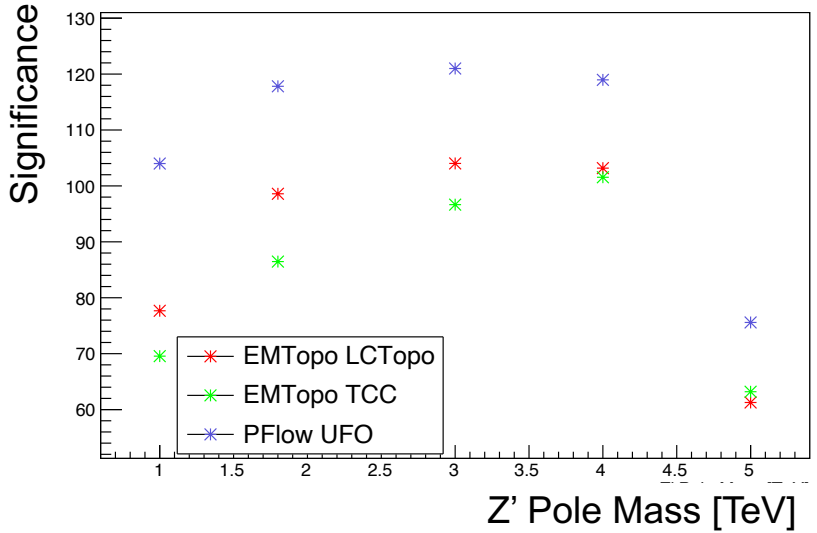
Significance study

- Simplified Significance Equation: $\sigma = \sum \sigma_i^2$
  - Where $\sigma_i = \frac{s_i}{\sqrt{s_i + b_i}}$ significance per bin; and, s_i = signal per bin, b_i = total background per bin

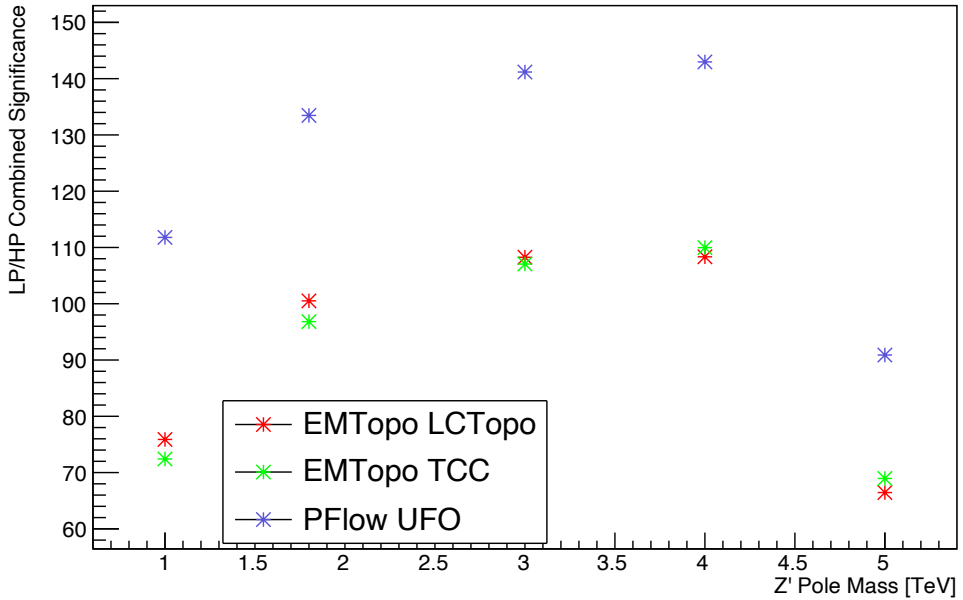
ggF HVTWZ (LP)



ggF HVTWZ (HP)



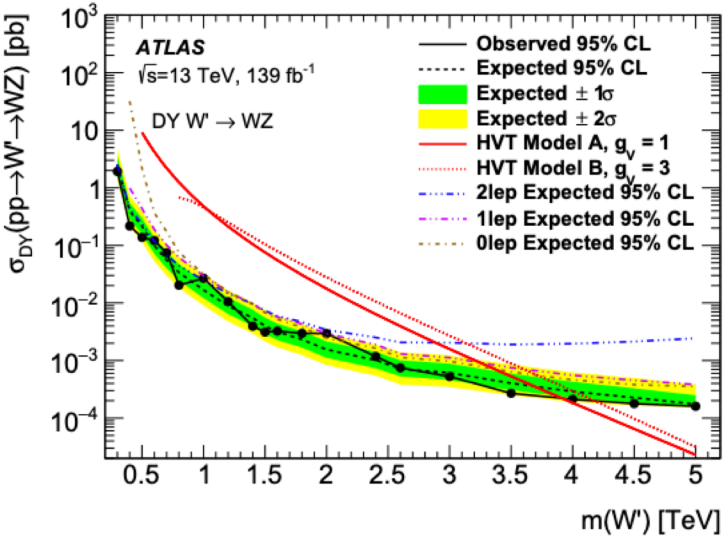
ggF HVTWZ (combined LP/HP)



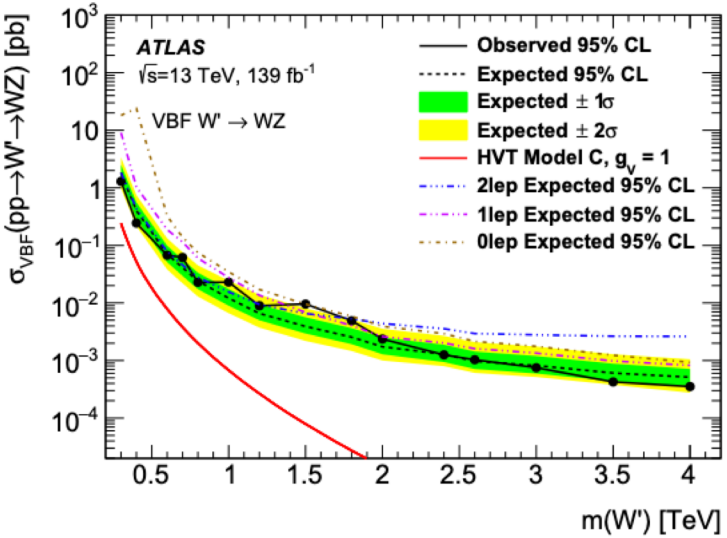
Previous analysis results

Paper: https://arxiv.org/abs/2004.14636

Drell-Yan production



VBF production



- Observed (black solid curve) and expected (black dashed curve) 95% CL upper limits on the DY (left) and VBF (right) production cross-section of an HVT W' boson at $\sqrt{s}$ = 13 TeV in the WZ decay mode as functions of its mass, combining searches in the three leptonic channels

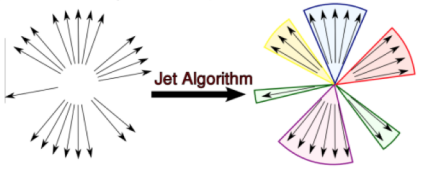
Next Steps

- Immediate
  - Obtain results for each signal models
    - Spin-0: Scalar “Radion”
    - Spin-1: Heavy vector triplet (HVT) W'/Z' (shown in this presentation)
    - Spin-2: Randall-Sundrum Graviton
  - Scale down the signal (picobarn -> femtobarn cross-section)
  - Investigate other 0-lepton improvements
    - MET significance study
      - CMSE 801 final project
- Future
  - Look into other studies as needed – possible improvements from ML approaches
  - Main analyzer for the 0-lep channel
  - Hopefully publish in 2022

JET FINDING ALGORITHM

Use a jet algorithm to cluster objects into a jet: It maps final state particle momenta to jet momenta.

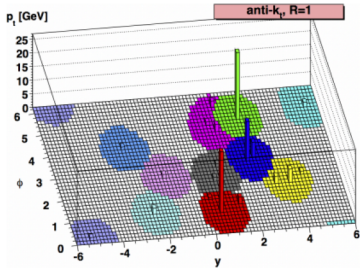
1 set of 4-vectors (constituents)



several subsets (jets)

- The Anti-Kt is the most widely used jet algorithm

Requires an external parameter, the jet radius R, which specifies up to which angle separate partons are recombined into a single jet.



Depending on physics intent, different size of jets radii are useful

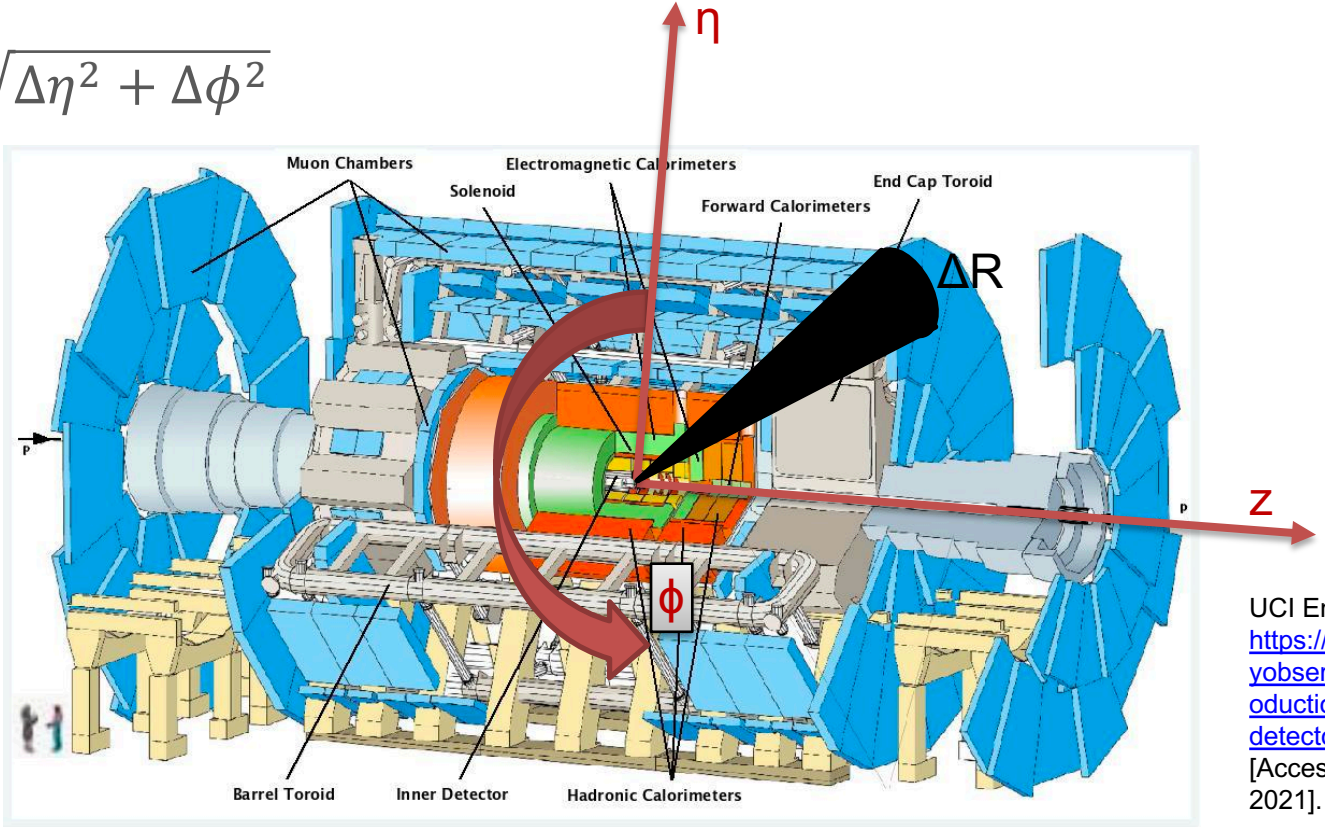
- CMS: 0.4, 0.8
- ATLAS: 0.4, 1.0

4

Dilia Maria Portillo Quintero, DBL Workshop 2021. https://indico.cern.ch/event/1070672/ [Accessed October 16, 2021].

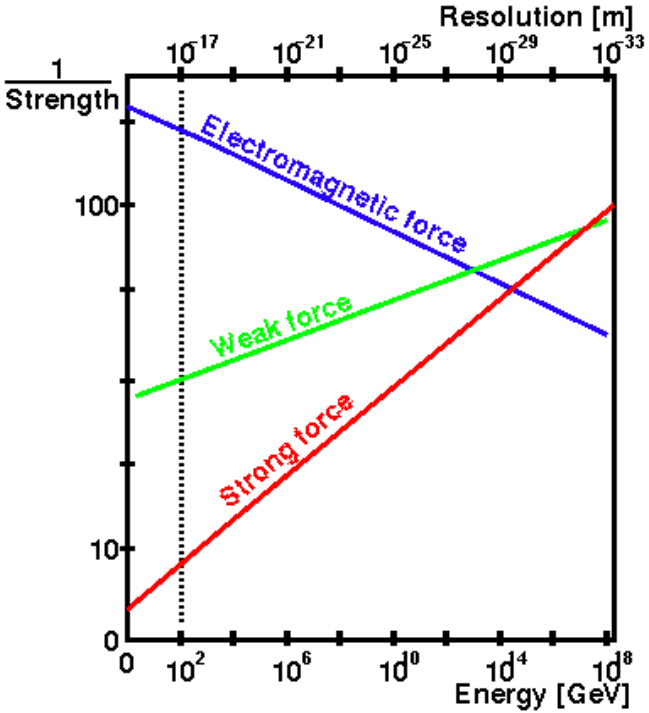
Background

- $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$



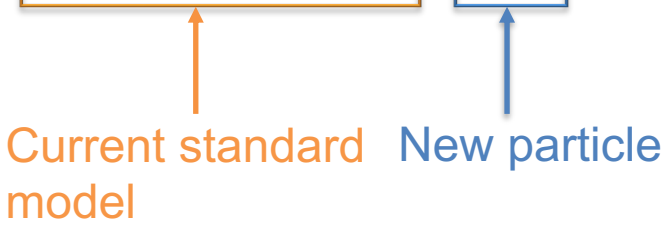
UCI Energy Blogs. https://sites.uci.edu/energyobserver/2012/11/26/introduction-to-the-atlas-detector-at-the-lhc/ [Accessed October 6, 2021].

- Running coupling constants



Motivation

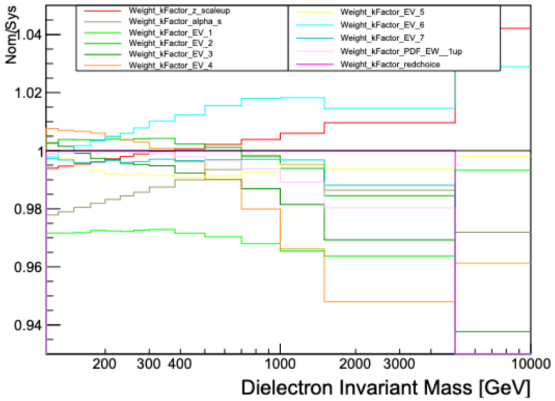
- Example of GUT that has failed
  - SU(5) → SU(3) x SU(2) x U(1) x U'(1)

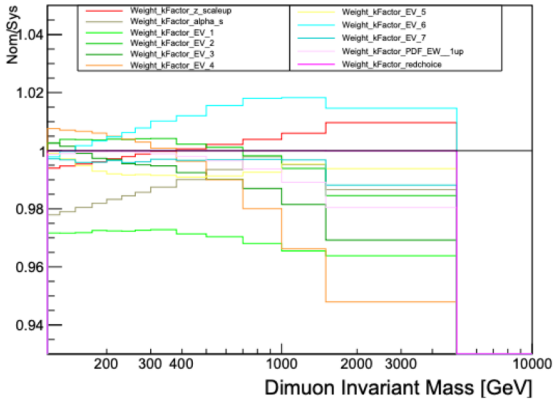


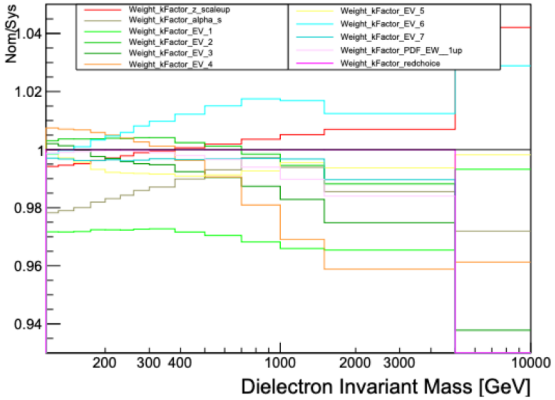
  - False because requires the proton is unstable
    - We have evidence that the proton is stable

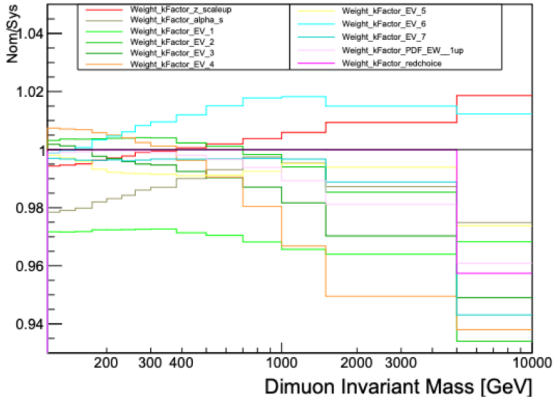
- Support note: https://cds.cern.ch/record/2703975/files/ATL-COM-PHYS-2019-1455.pdf
- Theoretical uncertainties in the Z/γ^* prediction arise from the PDF uncertainty, the renormalisation and factorisation scale choices, as well as the value of the strong coupling constant α_s, and higher-order corrections

HMDY Theory Uncertainties









Code and Documentation

- GitHub repo for CxAODReader:
  https://gitlab.cern.ch/CxAODFramework/CxAODReader_VVSemileptonic
- Derivation: HIGG5D1 (0 lep-channel)
- EasyTreePlotter:
  https://gitlab.cern.ch/rles/easytreeplotter/-/blob/master/README.md
- b-tagger
  - DL1r used for UFO
  - MV2c10 used for LCTTopo and TCC

Regions

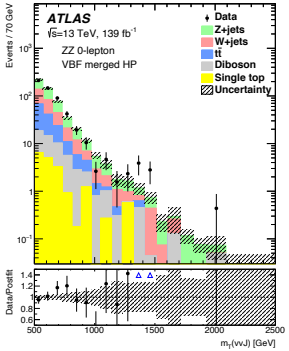
- V - Control Region (VCR)
  - Fails invariant mass cut
  - Low Purity (LP)
    - Fails $D_2^{(\beta=1)}$ requirement of boson tagging
  - High Purity (HP)
    - Passes $D_2^{(\beta=1)}$ requirement of boson tagging
- Signal Region (SR)
  - Passes invariant mass cut
  - Low Purity (LP)
    - Passes $D_2^{(\beta=1)}$ requirement of boson tagging
  - High Purity (HP)
    - Fails $D_2^{(\beta=1)}$ requirement of boson tagging

Selection	SR		VR	
	HP	LP	HP	LP
Number of Loose leptons	0			
$E_{\mathrm{T}}^{\mathrm{miss}}$	> 250 GeV			
p_{T}^{miss}	> 50 GeV			
min($\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}}$,small-R jets))	> 0.4			
$\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{miss})$	< 1			
Number of large-R jets	$\geqslant$ 1 large-R jets			
D_2 cut	pass	fail	pass	fail
W/Z mass window cut	pass	pass	m_J > 50 GeV, fail mass window cut	
Numb. of associated VR track jets b-tagged	For $Z \rightarrow J$: $\leq$ 1 (= 2) for untagged (tagged) category			

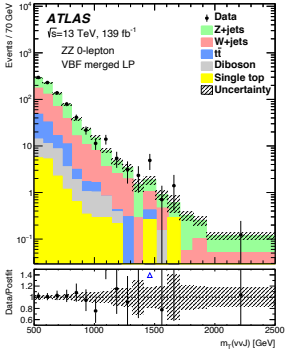
Cuts

Event selection	0-lepton ($ZV \to \nu\nu V_h$)	1-lepton ($WV \to \ell\nu V_h$)	2-lepton ($ZV \to \ell\ell V_h$)
V_ℓ selection	No *Loose* lepton $E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV $p_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV	1 *Tight* electron or 1 *Medium* muon with $p_{\mathrm{T}}^{\ell} > 30$ GeV $E_{\mathrm{T}}^{\mathrm{miss}} > 60$ GeV $p_{\mathrm{T}}^{V_\ell} > 75$ GeV	2 *Loose* leptons with $p_{\mathrm{T}}^{\ell} > 30$ GeV from the $Z \to \ell\ell$ candidate
Event veto	No additional *Loose* leptons Veto events with b-jets not associated with the $Z \to qq$ candidate		
Event categorisation	≥ 1 large-R jets or ≥ 2 small-R jets VBF and ggF/DY classification according to RNN score		
V_h selection (Merged)		$E_{\mathrm{T}}^{\mathrm{miss}} > 100$ GeV $p_{\mathrm{T}}^{V_\ell} > 200$ GeV	
	≥ 1 large-R jets The leading jet passing p_{T}-dependent m_J requirement		
		$\mathcal{R}_{p_{\mathrm{T}}/m} > 0.35$ (ggF/DY) $\mathcal{R}_{p_{\mathrm{T}}/m} > 0.25$ (VBF)	$\mathcal{R}_{p_{\mathrm{T}}/m} > 0.35$ (ggF/DY) $\mathcal{R}_{p_{\mathrm{T}}/m} > 0.25$ (VBF)
V_h selection (Resolved)	Not Performed	Failed merged selection ≥ 2 small-R jets with $\lvert\eta\rvert < 2.5$ $62 < m_{jj} < 97$ GeV for $W \to jj$ $70 < m_{jj} < 105$ GeV for $Z \to jj$	
		$\mathcal{R}_{p_{\mathrm{T}}/m} > 0.35$ (ggF/DY) $\mathcal{R}_{p_{\mathrm{T}}/m} > 0.25$ (VBF)	$\mathcal{R}_{p_{\mathrm{T}}/m} > 0.35$ (ggF/DY) $\mathcal{R}_{p_{\mathrm{T}}/m} > 0.35$ (VBF)

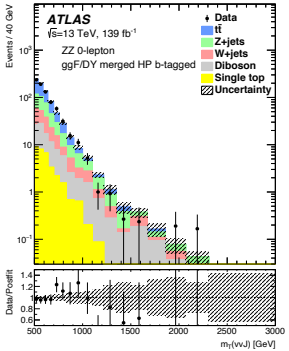
Previous analysis results



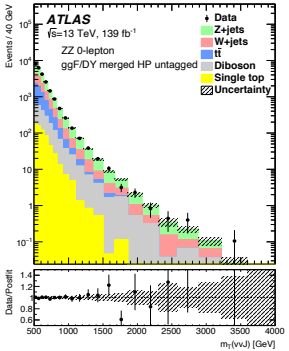
(a) VBF, merged HP



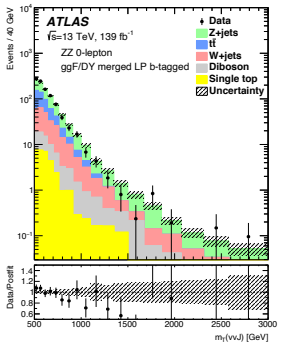
(b) VBF, merged LP



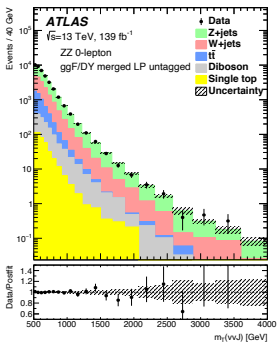
(c) ggF/DY, merged HP tagged



(d) ggF/DY, merged HP untagged



(e) ggF/DY, merged LP tagged

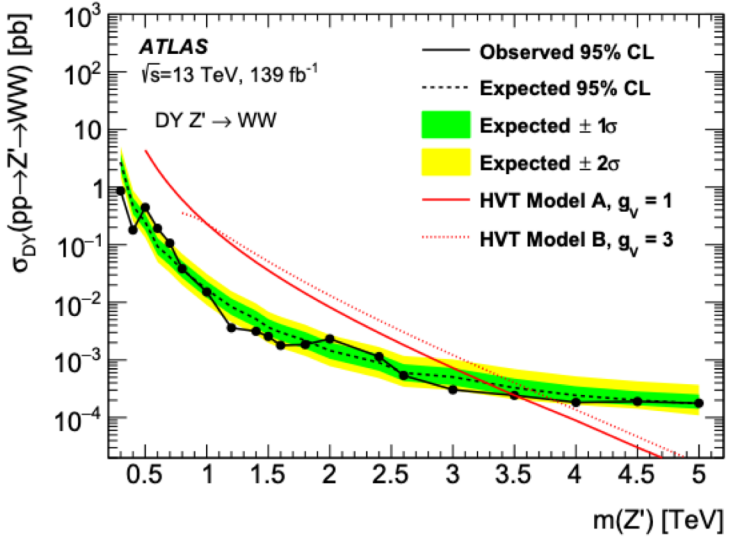


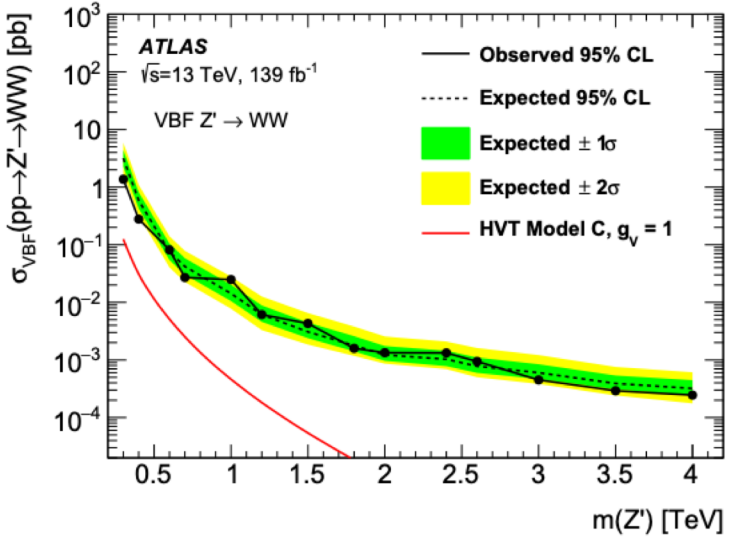
(f) ggF/DY, merged LP untagged

0-lep invariant mass
Paper: https://arxiv.org/abs/2004.14636

Previous analysis results

Paper: https://arxiv.org/abs/2004.14636



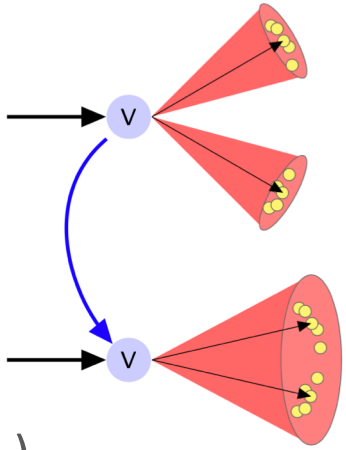


- Observed (black solid curve) and expected (black dashed curve) 95% CL upper limits on the (a) DY and (b) VBF production cross-section of an HVT Z' boson at √ s = 13 TeV in the WW decay mode as functions of its mass, combining searches in the three leptonic channels

Significance

- Simplified Significance Equation: $\sigma = \sum \sigma_i^2$
  - Where $\sigma_i = \frac{s_i}{\sqrt{s_i + b_i}}$ significance per bin; and, s_i = signal per bin, b_i = total background per bin
- More standard definition: $$Z = \begin{cases} +\sqrt{2\left(n\ln\left[\frac{n}{b}\right] - (n-b)\right)} & \text{if } n \geqslant b \\ -\sqrt{2\left(n\ln\left[\frac{n}{b}\right] - (n-b)\right)} & \text{if } n < b. \end{cases}$$
  - http://cdsweb.cern.ch/record/2736148/files/ATL-PHYS-PUB-2020-025.pdf (Equation 2)
  - n = number of events, b = predicted number of events

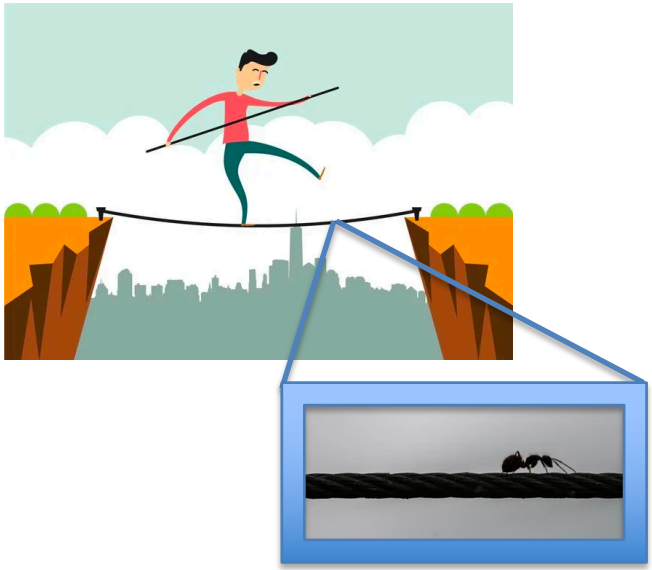
Definitions

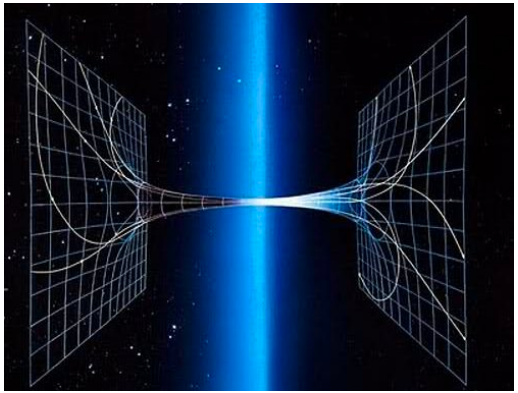


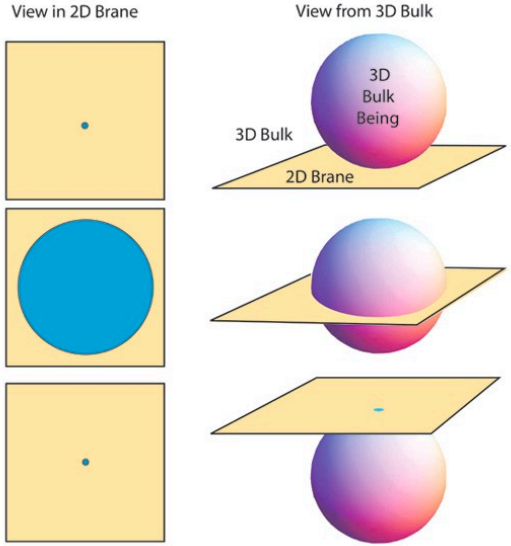
- Resolved
  - Two small R-jets
- Merged
  - One large R-jet
  - Only merged used here
- Jet substructure variable (D_2)
  - Large R-jet selection
  - Reconstructed by the energy correlation functions
- Missing E_T (MET or E_T^{miss})
  - Reconstructed by negative sum of the all physics objects and ID tracks not associated with them

Motivation

- Graviton
  - Hypothetical spin-2 particle that mediates the gravitational force







Definitions

- Pre-selection
  - Very limited cuts to reduce the data size of the CxAOD's
  - Zero reconstructed electrons/muons, some MET threshold, and one large-R jet
- Signal Regions
  - Several stricter cuts than pre-selection, with one of those cuts requiring boson tagging.
  - Try to maximize signal, minimize background
- Control Regions
  - Regions which look very similar to the signal region but have some slightly different selection to enrich a particular background (i.e. V boson), which is orthogonal to the signal region