Dissertation Defense

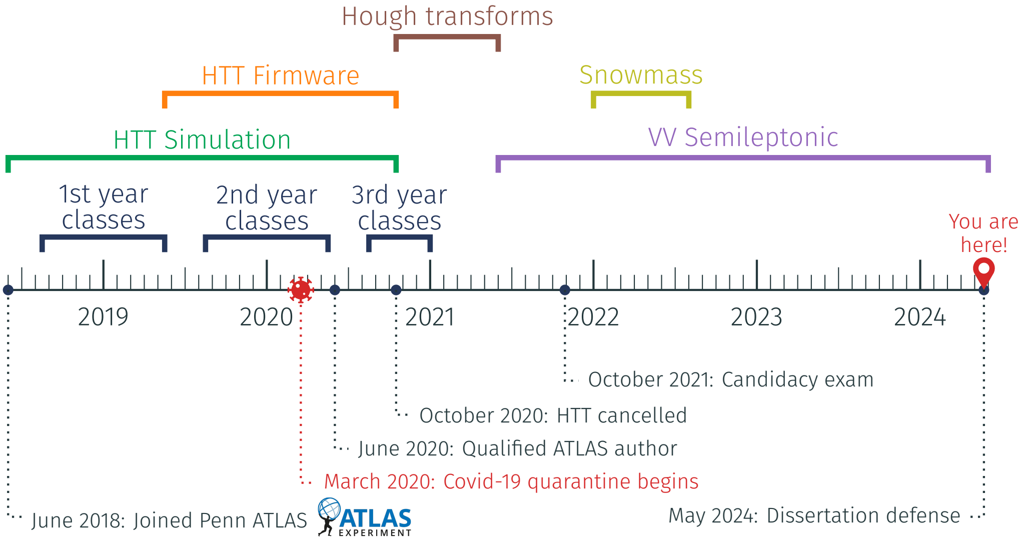
Riley Xu
Advisor: Elliot Lipeles

May 21, 2024

University of Pennsylvania
ATLAS Experiment

Outline

A Timeline


Hough transforms
HTT Firmware
Snowmass
HTT Simulation
VV Semileptonic
1st year classes
2nd year classes
3rd year classes
You are here!
2019
2020
2021
2022
2023
2024
October 2021: Candidacy exam
October 2020: HTT cancelled
June 2020: Qualified ATLAS author
March 2020: Covid-19 quarantine begins
June 2018: Joined Penn ATLAS
May 2024: Dissertation defense

A Summary

Tracking in the Trigger

- HTT Simulation
  - Software framework development
  - Functional simulation of hardware
  - Generating and optimizing patterns
- HTT Firmware
  - Sensor drivers
  - Power-up handshakes and protocols
  - Gigabit transceivers
  - Booting Linux on SoCs
- Hough Transforms
  - Converting HTTSim and maintenance
  - Implementing Hough algorithms
  - Algorithm optimizations

Analysis

- Snowmass 2021
  - Estimating future collider reaches
  - Summary plots for SUSY particles
- VV Semileptonic
  - Unfolding studies
    - Fiducial region definition
    - Fit validation
  - Sensitivity studies
    - ANN W/Z tagger mass window
    - $t\bar{t}$ background rejection
    - EFT interference resurrection
  - V+jets background modeling
    - Functional fits
    - Gaussian process regression

A Summary

Tracking in the Trigger

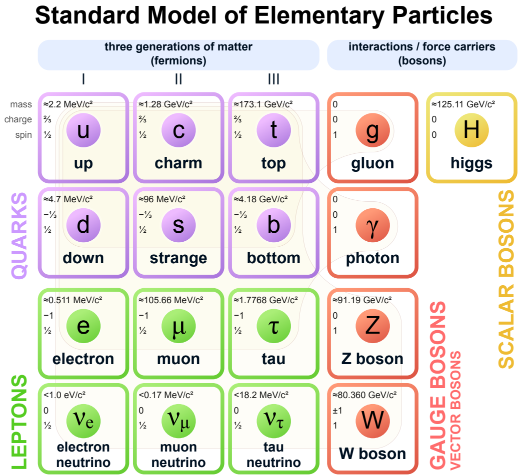
- HTT Simulation
  - Software framework development
  - Functional simulation of hardware
  - Generating and optimizing patterns
- HTT Firmware
  - Sensor drivers
  - Power-up handshakes and protocols
  - Gigabit transceivers
  - Booting Linux on SoCs
- Hough Transforms
  - Converting HTTSim and maintenance
  - Implementing Hough algorithms
  - Algorithm optimizations

Analysis

- Snowmass 2021
  - Estimating future collider reaches
  - Summary plots for SUSY particles
- VV Semileptonic
  - Unfolding studies
    - Fiducial region definition
    - Fit validation
  - Sensitivity studies
    - ANN W/Z tagger mass window
    - $t\bar{t}$ background rejection
    - EFT interference resurrection
  - V+jets background modeling
    - Functional fits
    - Gaussian process regression

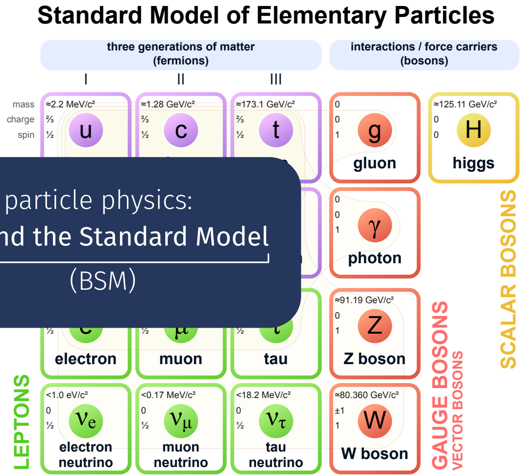
The Standard Model

- The Standard Model (SM):
  - Current model of the fundamental universe
  - Elementary particles and 3/4 forces
  - Well tested and accepted
- But incomplete! Missing:
  - Gravity
  - Neutrino masses / oscillations
  - Dark matter
- And some looming experimental discrepancies:
  - CDF W mass measurement
  - Muon $g-2$



Wikimedia

- The Standard Model (SM):
  - Current model of the fundamental universe
  - Elementary particles and 3/4 forces
  - Well tested
- But incomplete
  - Gravity
  - Neutrino ma
  - Dark matter
- And some looming experimental discrepancies:
  - CDF W mass measurement
  - Muon $g-2$

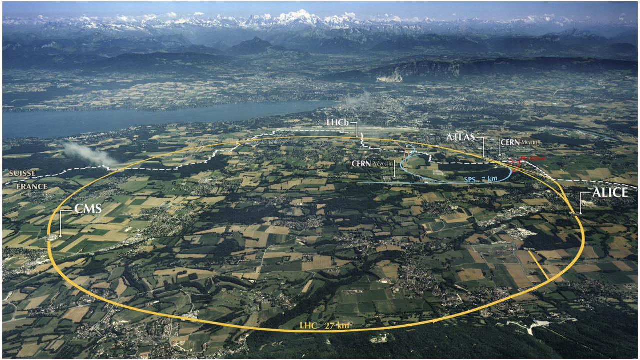


Wikimedia

Goal of modern particle physics:
Discover physics **beyond the Standard Model**
(BSM)

Searching for BSM Physics: Large Hadron Collider

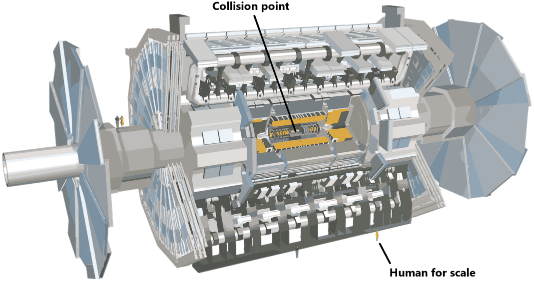
- What does BSM physics look like?
  - We don't know…but many theories
  - Usually postulate new, massive particles
  - Need to probe high energies
- Large Hadron Collider (LHC):
  - 27km underground ring by Geneva
  - Accelerates protons to 99.9999991% speed of light
  - Collided head-on to create massive particles
  - Detectors monitor collision products



Maximilien Brice

The ATLAS Detector

- General purpose detector on LHC ring
  - Monitors proton-proton collision products
  - Several layered systems to identify variety of particles
- A 7000 tonne machine
  - Approved in 1995, finished in 2008
  - 46 meters long, 25 meters diameter
- Collaboration of >6000 people from 47 countries

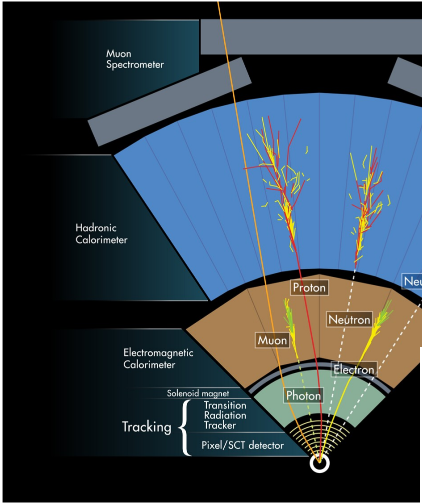


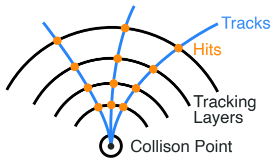
atlas.cern

Particle Reconstruction

- Particles fly through detector and leave **hits**
- Reconstruction: hits → particles
  - "Connect-the-dots" → **tracks** (particle momentum)
  - Calorimeter showers → **jets** (particle energy)
  - Where hits occur → particle type

CERN-EX-1301009





Tracking in the Trigger: Hough Transforms

Trigger System

- BSM physics is rare...
  - LHC collides **lots** of protons, once every 25ns
  - ATLAS generates 120 TB of data every second
  - Impossible to save everything
- Trigger:
  - Filter out "boring" (low-energy) data on the fly
  - Must filter down by a factor of $\approx 40,000$
  - Must be efficient: don't throw out interesting events on accident
- Need tracking in the trigger!
  - Currently using software tracking in CPUs
  - Very power intensive, unscalable after current run (ending circa 2025)

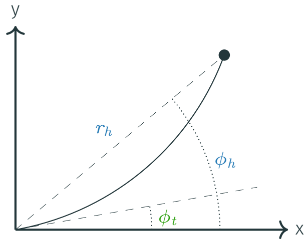
Hough Transform

- Efficient tracking method run on FPGAs
  - Cheaper and less power hungry than CPUs
- Track parameters are related to hit coordinates:

$$\phi_t = \frac{B r_h}{2} \frac{q}{p_T} + \phi_h$$

  - Measure: hit coordinates ϕ_h, r_h
  - Want: track parameters ϕ_t, q/p_T

Beam at origin, into page



Hough Transform

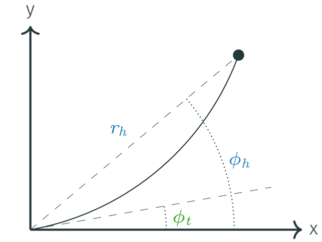
- Efficient tracking method run on FPGAs
  - Cheaper and less power hungry than CPUs
- Track parameters are related to hit coordinates:

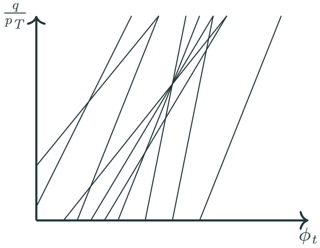
$$\phi_t = \frac{B r_h}{2} \frac{q}{p_T} + \phi_h$$

$$y = m \quad x + b$$

  - Measure: hit coordinates ϕ_h, r_h
  - Want: track parameters ϕ_t, q/p_T
  - For each hit: line of valid track parameters

Beam at origin, into page





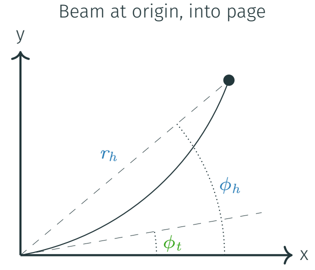
Hough Transform

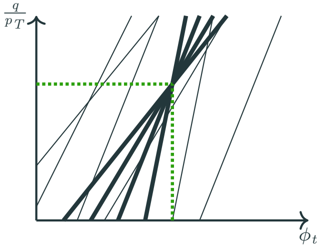
- Efficient tracking method run on FPGAs
  - Cheaper and less power hungry than CPUs
- Track parameters are related to hit coordinates:

$$\phi_t = \frac{B r_h}{2} \frac{q}{p_T} + \phi_h$$

$$y = m \; x + b$$

  - Measure: hit coordinates ϕ_h, r_h
  - Want: track parameters ϕ_t, q/p_T
  - For each hit: line of valid track parameters
  - Hits from same track intersect at true parameters
- Algorithm:
  1. Plot the lines for each hit in the detector
  2. Points where multiple lines intersect == tracks!





Hough Transform

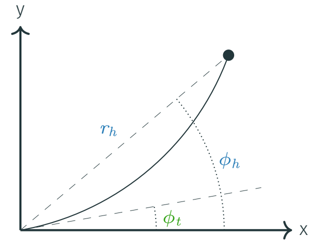
- Efficient tracking method run on FPGAs
  - Cheaper and less power hungry than CPUs
- Track parameters are related to hit coordinates:

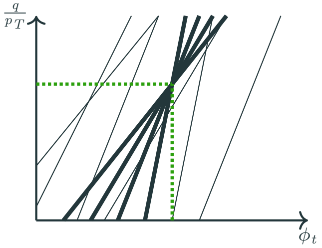
$$\phi_t = \frac{B r_h}{2} \frac{q}{p_T} + \phi_h$$

$$y = m x + b$$

  - Measure: hit coordinates ϕ_h, r_h
  - Want: track parameters ϕ_t, q/p_T
  - For each hit: line of valid track parameters
  - Hits from same track intersect at true parameters
- Algorithm:
  1. Plot the lines for each hit in the detector
  2. Points where multiple lines intersect == tracks!
- Relies on many simplifying assumptions
  - Major assumption: constant B field in tracker

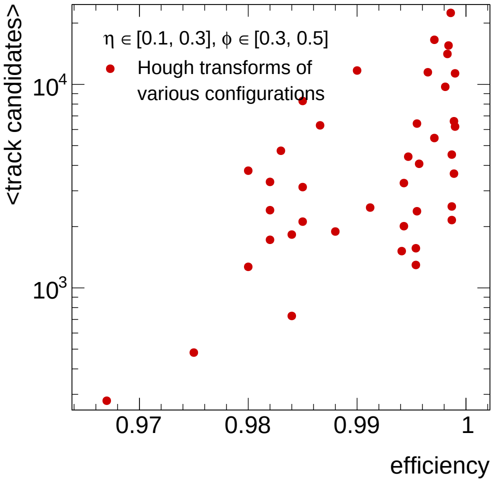
Beam at origin, into page





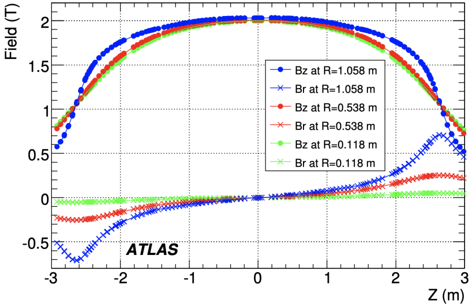
Hough Transform Tests

- Let's test!
  - Lots of tunable parameters
    - Corrections for detector inefficiencies
    - Corrections for Hough assumptions
    - And many more
  - Also try Elliot's alternative version
- Metrics:
  - Efficiency: % of tracks found
  - Fake rate: # of fake tracks per event
- Summary:
  - Great efficiency
  - Fake rate a bit high, but filtered downstream
  - Many suitable working points to choose from



Magnetic Field

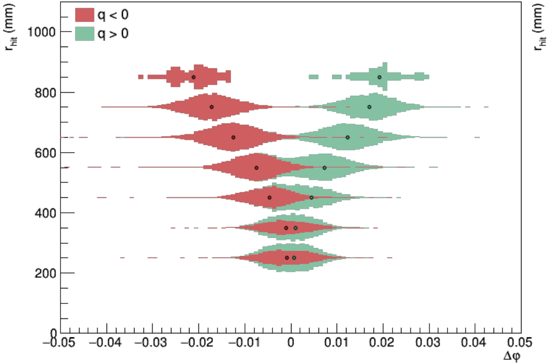
- Problems arise:
  - Previous results were for small slice in center of detector
  - Endcap has much lower efficiency: 99% → 87%
- Hypothesis:
  - Hough equation assumes constant magentic field
  - Actual field in ATLAS decays near endcaps



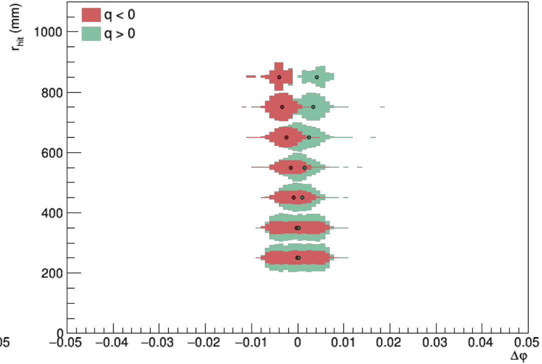
J. Phys.: Conf. Ser. 110 (2008) 092018

Magnetic Field Errors

- Verify: in simulation we know truth track parameters
- Calculate expected ϕ_h from Hough equation, compare vs true ϕ_h
- Big deviation at large r and low p_T, with sign dependent on charge



$p_T = 2$ GeV



$p_T = 10$ GeV

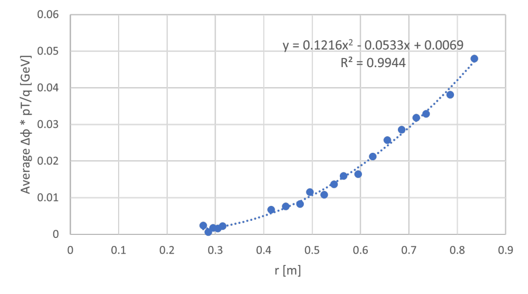
Magnetic Field Correction

- How to fix? Empirical correction!
- Approximate deviation with form

$$\Delta\phi = f(r_h)\frac{q}{p_T}$$

- Empirically fit a quadratic $f(r_h)$
- Add the correction to the Hough equation

$$\phi_t = \left(\frac{Br_h}{2} + f(r_h)\right)\frac{q}{p_T} + \phi_h$$



Plotting $\Delta\phi/\frac{q}{p_T} = f(r_h)$

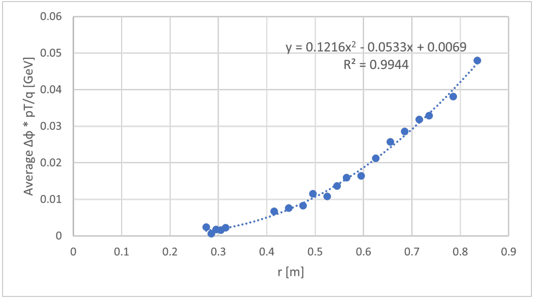
Magnetic Field Correction

- How to fix? Empirical correction!
- Approximate deviation with form

$$\Delta\phi = f(r_h)\frac{q}{p_T}$$

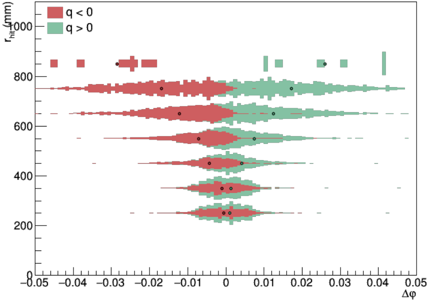
- Empirically fit a quadratic $f(r_h)$
- Add the correction to the Hough equation

$$\begin{aligned} \phi_t &= \left(\frac{Br_h}{2} + f(r_h)\right)\frac{q}{p_T} + \phi_h \\ y &= \qquad m \qquad\quad x + b \end{aligned}$$

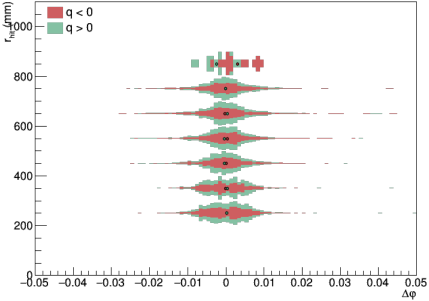


Plotting $\Delta\phi / \frac{q}{p_T} = f(r_h)$

Correction Test



Before correction



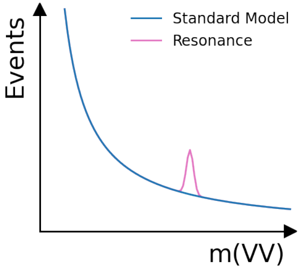
After correction

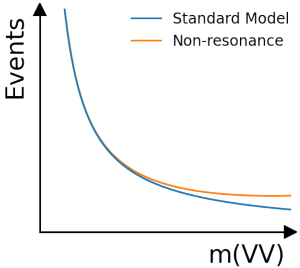
Config	Eff % in p_T bins (GeV) [1-2]	[2-4]	[4+]	Fake Tracks ($\mu = 200$)
No correction	87.0	98.0	99.5	13k
With correction	97.3	98.9	99.5	10k

VV Semileptonic: Gaussian Process Regression

Two Approaches

- Look for BSM physics in two major ways:
  - Direct search: assume a model and look for predicted signatures (often resonance)
  - Indirect measurement: look for shape deviations from SM (non-resonance)
- Indirect requires precise measurement of distributions
  - Need accurate background modeling!





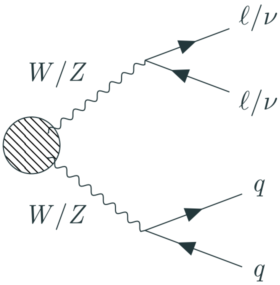
Analysis Overview

- VV Semileptonic: Pair of vector bosons ($V = W, Z$)
  - One decays to leptons: three channels 0ℓ, 1ℓ, 2ℓ
  - Other decays to quarks: single jet
- Why?
  - Dibosons are specially sensitive to certain BSM effects (EFT c_W operator)
  - ATLAS has only measured fully leptonic dibosons
  - Semileptonic

    $\Longrightarrow$ Higher branching ratios

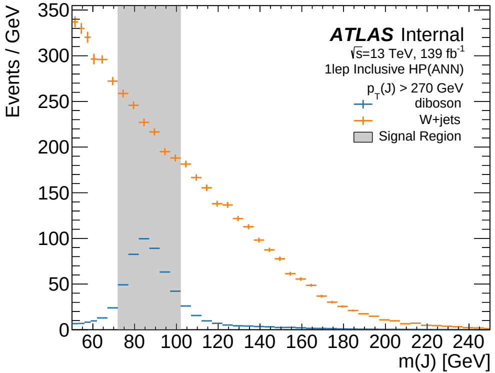
    $\Longrightarrow$ More data at higher energies

    $\Longrightarrow$ Better sensitivity to BSM effects
- But: much messier backgrounds



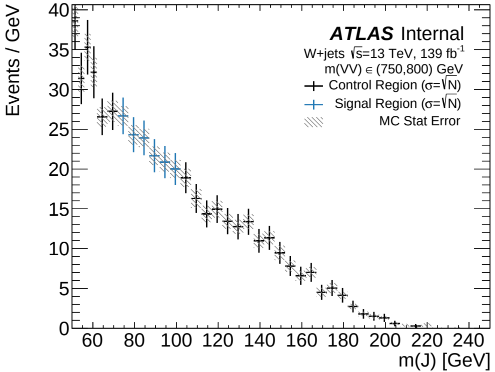
V+jets Background

- Main background is V+jets
  - Real leptonic boson
  - Extraneous jet mimics the hadronic boson
- Want data-driven approach
  - MC has large systematic errors
- Method: Sideband interpolation
  - Real bosons have a mass peak: signal region $m(J) \in [72, 102]$ GeV
  - V+jets is smooth distribution
  - Interpolate V+jets contribution from sidebands



First Attempts

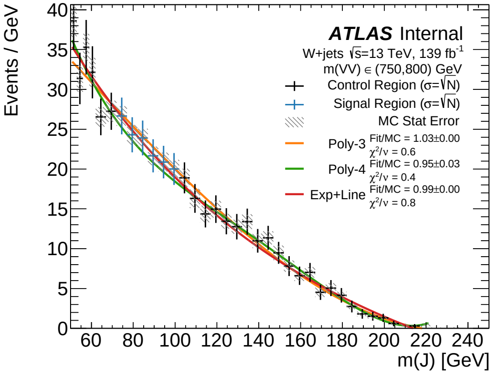
- Common method: functional fits



$m(VV) \in [750, 800]$ GeV

First Attempts

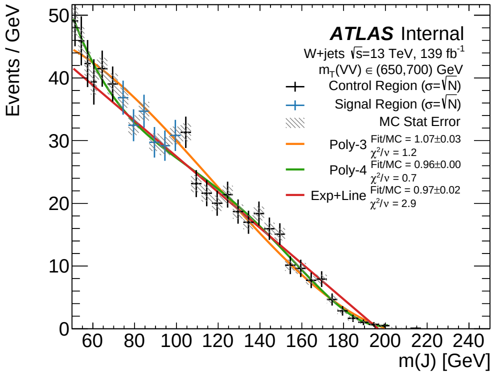
- Common method: functional fits
  1. 3rd order polynomial
  2. 4th order polynomial
  3. Exponential + linear
- Many problems:
  - Arbitrary choice of function leads to bias!
  - In this example, 8% difference between poly-3 and poly-4



$m(VV) \in [750, 800]$ GeV

First Attempts

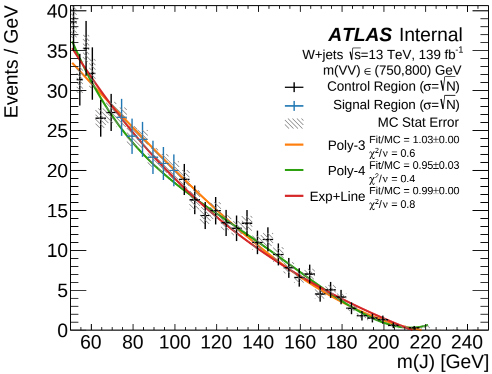
- Common method: functional fits
  1. 3rd order polynomial
  2. 4th order polynomial
  3. Exponential + linear
- Many problems:
  - Arbitrary choice of function leads to bias!
  - In this example, 8% difference between poly-3 and poly-4
  - Functional form can drastically change between bins



$m_T(VV) \in [650, 700]$ GeV

An Alternative Approach

- No one function is the best
  - What if we average over these 3?
  - Biases may average out to 0
  - But why these 3 and not others?
- Average over **every possible** fit function
- But how to do that?
  - Can't fit an infinite number of functions
  - Prioritize "good-fitting" functions
  - Prioritize "smooth"/simple functions



An Alternative Approach

- No one function is the best
  - What if we average over these 3?
  - Biases may average out to 0
  - But why these 3 and not others?
- Average over **every possible** fit function
- But how to do that?
  - Can't fit an infinite number of functions ⟵ Posterior $p(f|\mathbf{y})$
  - Prioritize "good-fitting" functions ⟶ Likelihood $p(\mathbf{y}|f)$
  - Prioritize "smooth"/simple functions ⟶ Prior $p(f)$

- Use Bayesian statistics!

$$\underbrace{p(f|\mathbf{y})}_{\text{posterior}} = \underbrace{p(\mathbf{y}|f)}_{\text{likelihood}} \underbrace{p(f)}_{\text{prior}} \; / \; \underbrace{p(\mathbf{y})}_{\substack{\text{marginal}\\ \text{likelihood}}}$$

- f = function, $\mathbf{y}$ = data

The Prior: Gaussian Process

$$\underbrace{p(f|\mathbf{y})}_{\text{posterior}} = \underbrace{p(\mathbf{y}|f)}_{\text{likelihood}} \underbrace{p(f)}_{\text{prior}} \;/\; \underbrace{p(\mathbf{y})}_{\text{marginal likelihood}}$$

- Hard part: defining the prior
  - How to define probabilities over functions?
  - How to prioritize "smooth" functions?
- Use formalism called Gaussian process
  - Generalization of multivariate Gaussian to ∞ dimensions

The Prior: Gaussian Process

$$\underbrace{p(f|\mathbf{y})}_{\text{posterior}} = \underbrace{p(\mathbf{y}|f)}_{\text{likelihood}} \underbrace{p(f)}_{\text{prior}} \; / \; \underbrace{p(\mathbf{y})}_{\text{marginal likelihood}}$$

- Hard part: defining the prior
  - How to define probabilities over functions?
  - How to prioritize “smooth” functions?
- Use formalism called Gaussian process
  - Generalization of multivariate Gaussian to ∞ dimensions

- Can use to define probabilities over functions
  - Approximate functions as ∞-dimensional vectors
  - ∞-dimensional Gaussian $\equiv$ PDF over functions
- Like normal Gaussians, fully defined by its mean and covariance
  - Covariance is called a **kernel**, controls the smoothness

The Prior: Gaussian Process

$$\underbrace{p(f|\mathbf{y})}_{\text{posterior}} = \underbrace{p(\mathbf{y}|f)}_{\text{likelihood}}\ \underbrace{p(f)}_{\text{prior}}\ /\ \underbrace{p(\mathbf{y})}_{\text{marginal likelihood}}$$

- Hard part: defining the prior
  - How to define probabilities over functions?
  - How to pr functions?
- Use formalism called Gaussian process
  - Generalization of multivariate Gaussian to ∞ dimensions

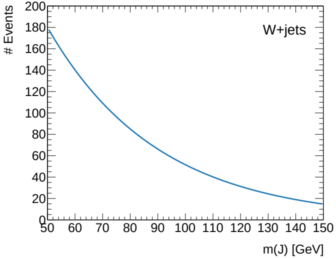
- Can use to define probabilities over functions
  - s ∞-dimensional
  - $\equiv$ PDF over functions
- Like normal Gaussians, fully defined by its mean and covariance
  - Covariance is called a **kernel**, controls the smoothness

Choosing a prior $\implies$ Choosing a kernel

Understanding the Kernel

$$\underbrace{p(f|\mathbf{y})}_{\text{posterior}} = \underbrace{p(\mathbf{y}|f)}_{\text{likelihood}} \underbrace{p(f)}_{\text{prior}} \; / \; \underbrace{p(\mathbf{y})}_{\text{marginal likelihood}}$$

- What is the kernel?
  - Arbitrary function $k(x_1, x_2)$
  - Covariance between $f(x_1)$ and $f(x_2)$
- How correlated should neighboring points be?

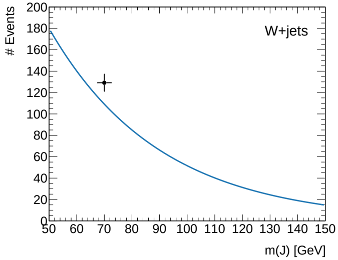


Prior
(no data)

Understanding the Kernel

$$\underbrace{p(f|\mathbf{y})}_{\text{posterior}} = \underbrace{p(\mathbf{y}|f)}_{\text{likelihood}} \underbrace{p(f)}_{\text{prior}} \Big/ \underbrace{p(\mathbf{y})}_{\text{marginal likelihood}}$$

- What is the kernel?
  - Arbitrary function $k(x_1, x_2)$
  - Covariance between $f(x_1)$ and $f(x_2)$
- How correlated should neighboring points be?

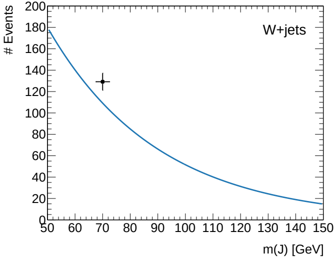


Prior
+ datapoint

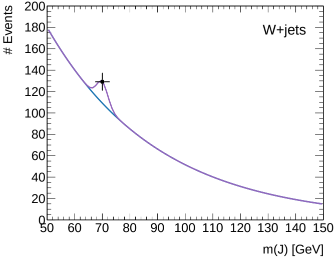
Understanding the Kernel

$$\underbrace{p(f|\mathbf{y})}_{\text{posterior}} = \underbrace{p(\mathbf{y}|f)}_{\text{likelihood}} \underbrace{p(f)}_{\text{prior}} \Big/ \underbrace{p(\mathbf{y})}_{\text{marginal likelihood}}$$

- What is the kernel?
  - Arbitrary function $k(x_1, x_2)$
  - Covariance between $f(x_1)$ and $f(x_2)$
- How correlated should neighboring points be?



Prior
+ datapoint

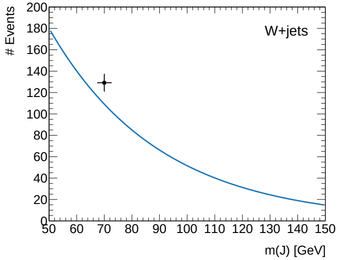


Posterior
small correlations
e.g. $k(70, 80) = 0$

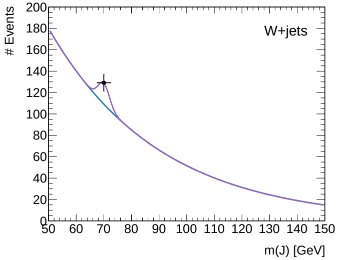
Understanding the Kernel

$$\underbrace{p(f|\mathbf{y})}_{\text{posterior}} = \underbrace{p(\mathbf{y}|f)}_{\text{likelihood}}\ \underbrace{p(f)}_{\text{prior}}\ /\ \underbrace{p(\mathbf{y})}_{\text{marginal likelihood}}$$

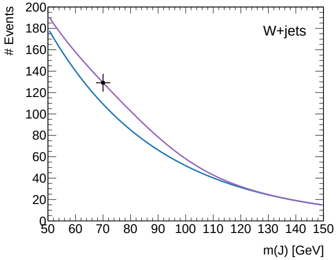
- What is the kernel?
  - Arbitrary function $k(x_1, x_2)$
  - Covariance between $f(x_1)$ and $f(x_2)$
- How correlated should neighboring points be?



Prior
+ datapoint



Posterior
small correlations
e.g. $k(70, 80) = 0$

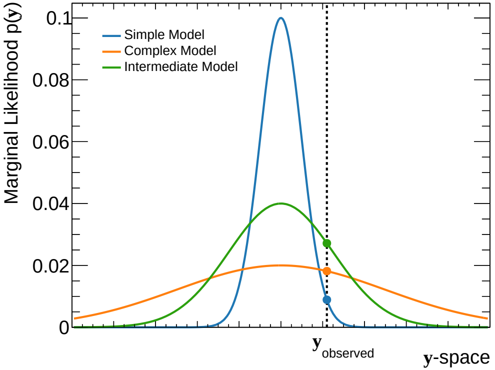


Posterior
large correlations
e.g. $k(70, 80) > 0$

Choosing a Kernel

$$\underbrace{p(f|\mathbf{y})}_{\text{posterior}} = \underbrace{p(\mathbf{y}|f)}_{\text{likelihood}} \underbrace{p(f)}_{\text{prior}} \Big/ \underbrace{p(\mathbf{y})}_{\text{marginal likelihood}}$$

- How to choose the kernel?
  - Maximize the marginal likelihood $p(\mathbf{y} \mid \text{kernel})$
  - Pick the one that best matches the sideband data
- Automatic trade-off between complexity and accuracy
  - Simple model can't model data correctly
  - Complex model adapts readily, but has small max probability
  - Intermediate model achieves best balance

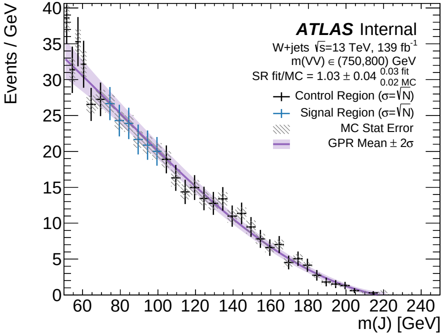


Adapted from Rasmussen and Williams

Bayesian Regression in Action

$$\underbrace{p(f|\mathbf{y})}_{\text{posterior}} = \underbrace{p(\mathbf{y}|f)}_{\text{likelihood}}\ \underbrace{p(f)}_{\text{prior}}\ /\ \underbrace{p(\mathbf{y})}_{\text{marginal likelihood}}$$

- With kernel chosen, we now have a prior!
- Plug and chug Bayes' rule to get posterior
  - Remember, output is a probability distribution for EVERY possible fit function
  - Use max $p(f)$ as predicted value (solid line)
  - Use 1σ CI as uncertainty (shaded band)
- Works great!
  - Smaller errors than functional fits
  - Manifest errors from PDF
  - No arbitrary choice of function anymore
  - No headaches from non-convergence

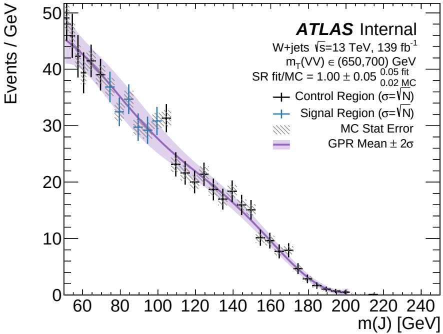


$m(VV) \in [750, 800]$ GeV

Bayesian Regression in Action

$$\underbrace{p(f|\mathbf{y})}_{\text{posterior}} = \underbrace{p(\mathbf{y}|f)}_{\text{likelihood}} \underbrace{p(f)}_{\text{prior}} \Big/ \underbrace{p(\mathbf{y})}_{\text{marginal likelihood}}$$

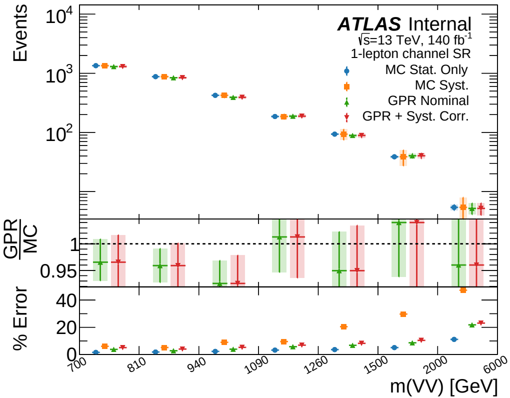
- With kernel chosen, we now have a prior!
- Plug and chug Bayes' rule to get posterior
  - Remember, output is a probability distribution for EVERY possible fit function
  - Use max $p(f)$ as predicted value (solid line)
  - Use 1σ CI as uncertainty (shaded band)
- Works great!
  - Smaller errors than functional fits
  - Manifest errors from PDF
  - No arbitrary choice of function anymore
  - No headaches from non-convergence
  - Adapts readily to changing shapes bin-to-bin



$m_T(VV) \in [650, 700]$ GeV

Closure Tests

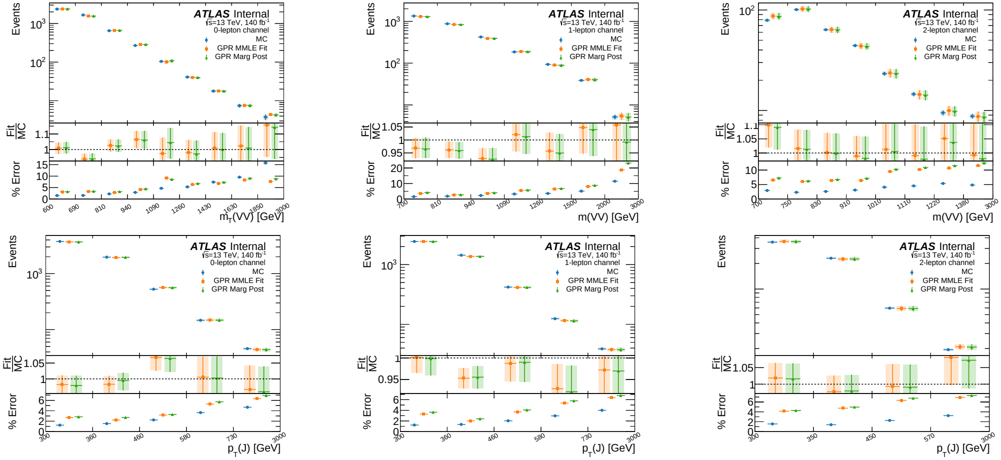
- Repeat for every bin in $m(VV)$
  - Closures within uncertainties
  - Much lower errors compared to Monte Carlo simulations
- Some subtleties not addressed:
  - Uncertainties on kernel choice
  - Dealing with signal contamination and other backgrounds in control region
  - Tests on 2D fits



Compare GPR (red) with MC (orange)

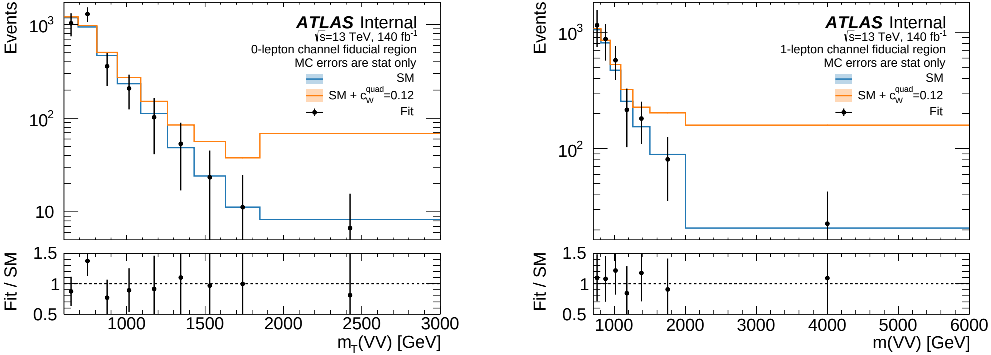
Many Closure Tests

Repeat for bins of $p_T(J)$, and repeat for every lepton channel...



Final Results?

- Money plots: measured diboson distributions (unfolded)
- But first: tests of full procedure on asimov data (fake data == expected events from SM)
- Good closure with SM, large sensitivity to BSM effect (c_W set to 2σ global limit)



0-lepton channel 1-lepton channel

Final Results!

- Still blinded...awaiting EB request!

Conclusion

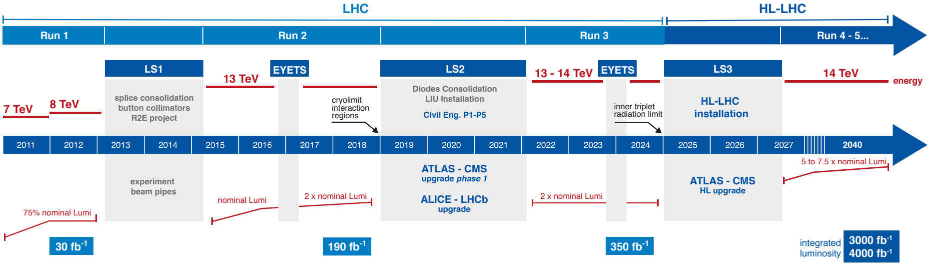
- Optimized Hough transforms for tracking in the trigger
- Introduced Gaussian process regression for background modeling in the VV Semileptonic analysis
- Results pending!

Thank you for your attention!

Backup

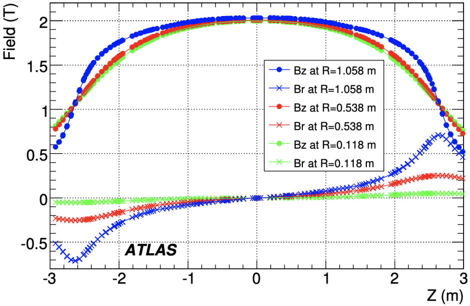
HL-LHC

https://hilumilhc.web.cern.ch/content/hl-lhc-project

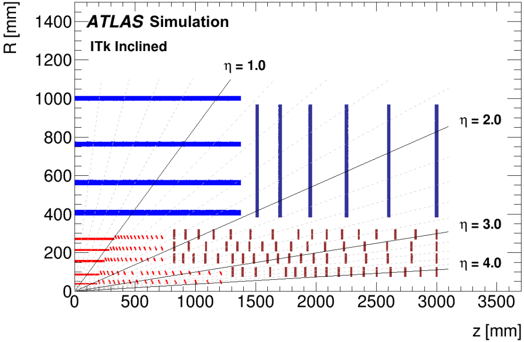


- In 2027 the LHC will be upgraded to the High Luminosity LHC
- Average inelastic collisions per bunch up to $\langle\mu\rangle \sim 200$
- ATLAS will need upgrades to handle this extra pileup, to be installed 2025-2027

ITK Layout



J. Phys.: Conf. Ser. 110 (2008) 092018



https://cds.cern.ch/record/2257755

Hough Setup and Configuration Variables

Layer choice Which detector layers to input. Default 8 layers: Pixel 4, Strip 0 inner side, Strip 1-3 both sides.

Threshold Minimum number of hit layers to create a track candidate.

Bin sizes Width of ϕ_t and q/p_T bins. Finer bins = less fake tracks but also lower efficiency.

Padding Add extra bins to nominal region of interest account for boundary effects.

z-slicing See next slide.

Smearing Fill extra phi bins on each side to account for resolution effects and d_0.

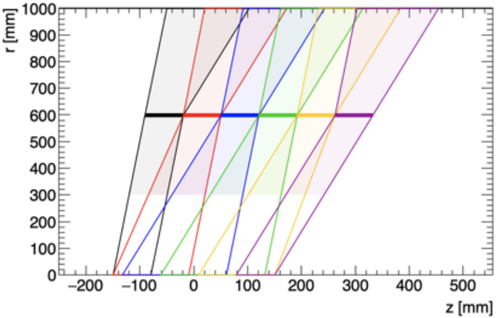
Duplicate removal Many algorithms. Window algorithm: for each bin, look at $n \times n$ neighbors. If any neighbor

1. Has more hit layers
2. Has equal hit layers, but more hits
3. Has equal hit layers and hits, but is in a lower-left bin

do not create road from this bin.

Key-layer Slicing

- Define slices by splitting a key layer along z, on module boundaries
  - For each slice, select tracks that go through the respective modules on the key layer
  - Add all other modules these tracks hit to the slice
- Best key layer appears to be the outer short strip layer (for current layer choice)



(Elliot)

<clusters> per slice:

Layer	6 Slice	18 Slice
0	102	70
1	86	51
2	45	13
3	66	35
4	71	43
5	72	44
6	224	85
7	278	98
Ave	118	55

Single muons flat in:

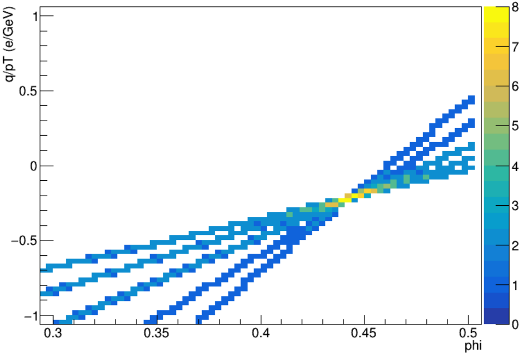
$\phi \in [0.3, 0.5]$

$d_0 \in [-2, 2]$ mm

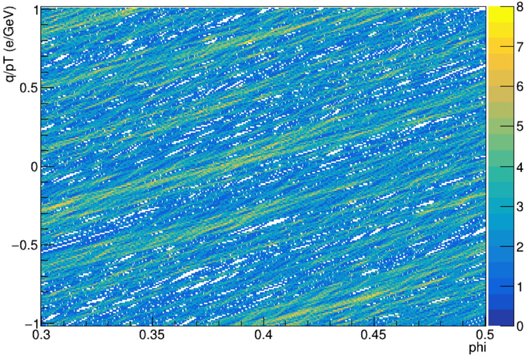
$z_0 \in [-150, 150]$ mm

$1/p_T \in [-1, 1]$ GeV

Sample Images

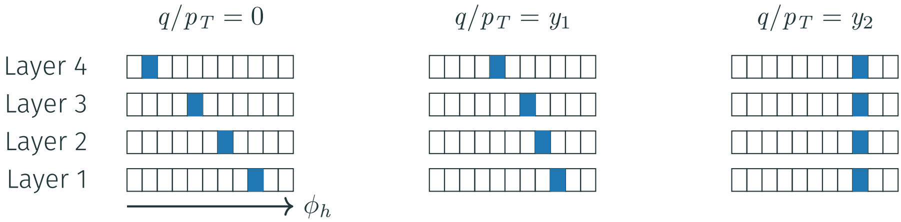


Single track with no background hits



Single track with $\mu = 200$ background hits

1D Hough Transform



- Instead of binning in q/p_T and ϕ_t, bin in ϕ_h (per layer)
- Assume hits from same layer all have the same r
- Iterate different $q/p_T \implies$ bitshift the ϕ_h bins to ϕ_t

$$\phi_t - \phi_h = \frac{r_i B}{2} \cdot \frac{q}{p_T}$$

- Multiple layers have hit at same bin $\implies$ intersection at common ϕ_t for the given $q/p_T \implies$ track!

Task Force Decision

- Task forces created December 2020 to evaluate new EF baseline
- Reports submitted June 2021, reviewed by a trigger committee
- Committee concluded to not pursue custom hardware, but keep investigating commodity solutions
- TDR amendment published: ATLAS-TDR-029-ADD-1

TF	Algorithm	Hardware	Pros	Cons
Custom	Pattern matching	AM chip	Low power	Costly and difficult
Hetero	Hough transform	FPGA	Cheap, ok power	Lots of fakes
Software	ITk fast tracking	CPU	Best at tracking	Power intensive

Comparison - Cost, Power, and Effort

Architecture	Cost (MCHF)	Effort (FTE)	Power (MW)
Custom	4.55	140	0.59
Heterogeneous commodity	2.0	80[1]	0.91
Software-only	3.28	14[1]	1.82

- HTT cost is 17.45 MCHF, the EF farm is 3.4 MCHF, and the Run-3 farm is 2.85 MCHF
- 1 MW $\times$ duty cycle = 3 GWh/year $\approx$ 200 homes
- Point-1 expected limit of 2.5 MW

[1] Review committee notes that these seem underestimated

W/Z Branching Ratios

Mode	Branching Ratio (%)
$e\nu$	10.71 ± 0.16
$\mu\nu$	10.63 ± 0.15
$\tau\nu$	11.38 ± 0.21
hadrons	67.41 ± 0.27

W boson

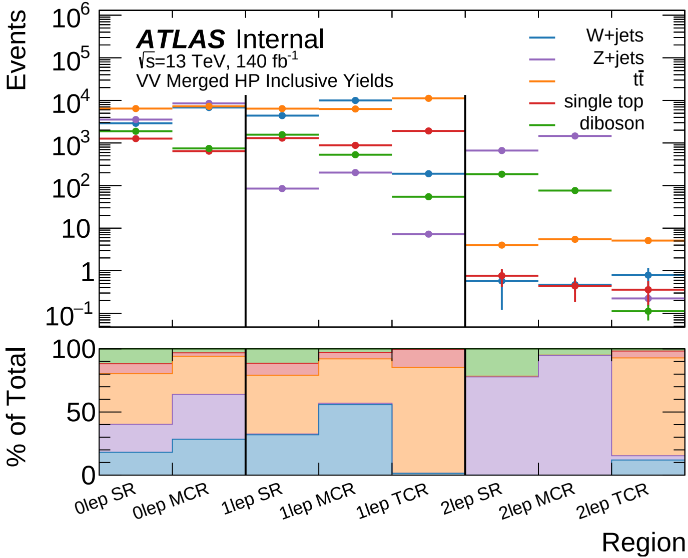
Mode	Branching Ratio (%)
e^+e^-	3.3632 ± 0.0042
$\mu^+\mu^-$	3.3662 ± 0.0066
$\tau^+\tau^-$	3.3696 ± 0.0083
invisible	20.000 ± 0.055
hadrons	69.911 ± 0.056

Z boson

PDG 2022

Event Selection

0-lepton	1-lepton	2-lepton
Lepton Selection		
0 “Loose” leptons	1 “Tight” lepton and 0 “Loose” leptons	2 opp.-sign same-flavor “Loose” leptons
$E_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV $E_{\mathrm{T}}^{\mathrm{miss}}$ sig. > 10	$p_T(\ell) > 30$ GeV $E_{\mathrm{T}}^{\mathrm{miss}} > 100$ GeV $p_T(W) > 200$ GeV	Leading ℓ: $p_T > 27$ GeV Subleading ℓ: $p_T > 25$ GeV $m(\ell\ell)$: $83 < m_{ee} < 99$ GeV $m_{\mu\mu} > 85.6 - 0.0117 p_T(\ell\ell)$ GeV $m_{\mu\mu} < 94.0 + 0.0185 p_T(\ell\ell)$ GeV
Jet Selection		
Leading large-R jet		
$p_T(J) > 300$ GeV 72 GeV $< m(J) <$ 102 GeV Pass W/Z substructure tagger		
Topology Requirements		
$\Delta\phi(\text{jets}, E_{\mathrm{T}}^{\mathrm{miss}}) > \frac{\pi}{9}$	$E_{\mathrm{T}}^{\mathrm{miss}}/p_T(W) > 0.2$ (electron-only)	-
-	No additional b-jets	-
-	$p_T(V)/m(VV) > 0.35$	


ATLAS Internal
√s=13 TeV, 140 fb^{-1}
VV Merged HP Inclusive Yields
W+jets
Z+jets
$t\bar{t}$
single top
diboson
Events
% of Total
0lep SR
0lep MCR
1lep SR
1lep MCR
1lep TCR
2lep SR
2lep MCR
2lep TCR
Region

Recent ATLAS/CMS Diboson Results

	Process	Note	$\mathcal{L}$ [fb^{-1}]	Ref.	ATLAS?
ATLAS	$pp \to 4\ell$	SF+OS x2	139	↗	
	$WZ \to 3\ell$	Polarizations, no EFT fit	139	↗	
	$WW \to e\mu$	No EFT fit	140	↗	
	$WW + j \to e\mu$	Interference enhanced	139	↗	
	$Z + 2j \to 2\ell$	VBF	139	↗	
	$Z + \gamma\gamma \to 2\ell$	Dim-8 aQGC limits	139	↗	
	$ZZ + 2j \to 4\ell$	Dim-8 aQGC limits, VBS	139	↗	
	$W^\pm W^\pm + 2j \to 2\ell$	Dim-8 aQGC limits, VBS	139	↗	
	$W\gamma + 2j \to \ell$	Dim-8 aQGC limits, VBS	139	↗	
	$VV + 2j \to \ell\ell jj$	Dim-8 aQGC limits, VBS	139	↗	
CMS	$WW \to 2\ell$		36	↗	✓
	$WZ \to 3\ell$	Polarization angles	138	↗	✓
	$W\gamma \to \ell$	Interference resurrection	138	↗	7 TeV ; Run-2 in progress?
	$Z\gamma + 2j \to 2\ell$	Dim-8 aQGC limits, VBS	36	↗	No EFT limits
	$W\gamma + 2j \to \ell$	Dim-8 aQGC limits, VBS	138	↗	✓
	$W^\pm W^\pm + 2j \to 2\ell$	Dim-8 aQGC limits, VBS	137	↗	✓
	$WZ + 2j \to 3\ell$	Dim-8 aQGC limits, VBS	137	↗	
	$ZZ + 2j \to 4\ell$	Dim-8 aQGC limits, VBS	137	↗	✓
	$VV + 2j \to \ell\ell jj$	Dim-8 aQGC limits, VBS	36	↗	✓

Disclaimer: This list is probably far from comprehensive.

Effective Field Theory

- Framework to organize indirect measurements
- No assumptions about new particles or decay modes! Only major assumption:

$$\underbrace{\text{Energy BSM physics}}_{\Lambda} \gg \underbrace{\text{Energy SM physics}}_{E}$$

- Effectively, Taylor expansion around SM in terms of $1/\Lambda$

$$\text{Taylor:} \quad f(x) = e^x \approx 1 + x + \frac{1}{2!}x^2 + \ldots$$

$$\text{EFT:} \quad \mathcal{L}_{\text{BSM}} \approx \mathcal{L}_{\text{SM}} + \frac{1}{\Lambda^2}\sum_i c_i \mathcal{O}_i^{(6)} + \frac{1}{\Lambda^4}\sum_i c_i \mathcal{O}_i^{(8)} + \ldots$$

  - 59 operators $\mathcal{O}_i^{(6)}$ at leading order
  - Experiments fit the coefficients c_i/Λ^2

Effective Field Theory

- Framework to organize indirect measurements
- No assumptions about new particles or decay modes! Only major assumption:

$$\text{Energy BSM physics} \gg \text{Energy SM physics}$$

Targets of indirect measurements:

1. Unfolded cross section
2. EFT coefficient limits

- Effectively, Taylor

Taylor: $f(x) \ldots \frac{1}{2!}x^2 + \ldots$

$$\text{EFT:} \quad \mathcal{L}_{\text{BSM}} \approx \mathcal{L}_{\text{SM}} + \frac{1}{\Lambda^2}\sum_i c_i \mathcal{O}_i^{(6)} + \frac{1}{\Lambda^4}\sum_i c_i \mathcal{O}_i^{(8)} + \ldots$$

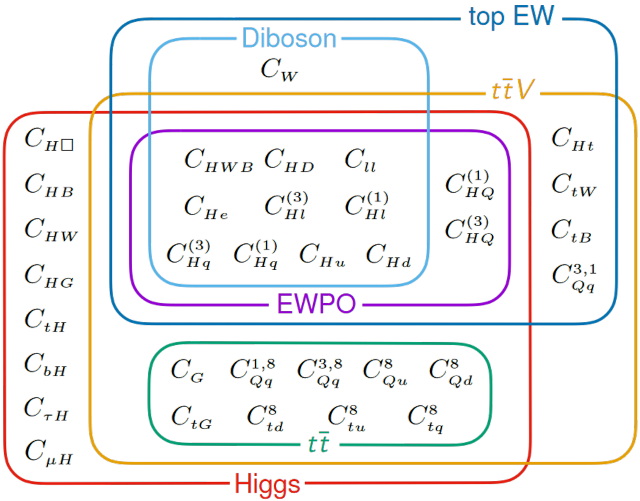
  - 59 operators $\mathcal{O}_i^{(6)}$ at leading order
  - Experiments fit the coefficients c_i/Λ^2

EFT Formalism

$$\sigma \propto |\mathcal{M}_{\text{SMEFT}}|^2 = |\mathcal{M}_{\text{SM}}|^2 + \underbrace{\frac{1}{\Lambda^2}\mathcal{M}^*_{\text{SM}}\mathcal{M}^{(6)}_{\text{EFT}}}_{\text{linear term}} + \underbrace{\frac{1}{\Lambda^4}\left|\mathcal{M}^{(6)}_{\text{EFT}}\right|^2}_{\text{quadratic term}} + \underbrace{\frac{1}{\Lambda^4}\mathcal{M}^*_{\text{SM}}\mathcal{M}^{(8)}_{\text{EFT}}}_{\mathcal{O}(8)\text{ interference}}$$

- First order operators are dimension 6
  - Odd-dimension terms violate lepton/baryon number
  - 59 operators
- Linear interference term is first order effect (Λ^{-2})
- Quadratic term is second order effect (Λ^{-4})
  - Same order as un-modeled dim-8 interference
  - Taken as an uncertainty
- Two fits: just linear, and linear + quadratic

SMEFT Sensitive Operators



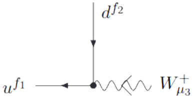
x2012.02779

SMEFT Operators: other than 4 fermion

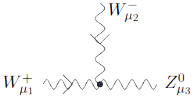
X^3		φ^6 and $\varphi^4 D^2$		$\psi^2\varphi^3$	
Q_G	$f^{ABC} G_\mu^{A\nu} G_\nu^{B\rho} G_\rho^{C\mu}$	Q_φ	$(\varphi^\dagger\varphi)^3$	$Q_{e\varphi}$	$(\varphi^\dagger\varphi)(\bar{l}_p e_r \varphi)$
$Q_{\widetilde{G}}$	$f^{ABC} \widetilde{G}_\mu^{A\nu} G_\nu^{B\rho} G_\rho^{C\mu}$	$Q_{\varphi\Box}$	$(\varphi^\dagger\varphi)\Box(\varphi^\dagger\varphi)$	$Q_{u\varphi}$	$(\varphi^\dagger\varphi)(\bar{q}_p u_r \widetilde{\varphi})$
Q_W	$\varepsilon^{IJK} W_\mu^{I\nu} W_\nu^{J\rho} W_\rho^{K\mu}$	$Q_{\varphi D}$	$\left(\varphi^\dagger D^\mu\varphi\right)^\star \left(\varphi^\dagger D_\mu\varphi\right)$	$Q_{d\varphi}$	$(\varphi^\dagger\varphi)(\bar{q}_p d_r \varphi)$
$Q_{\widetilde{W}}$	$\varepsilon^{IJK} \widetilde{W}_\mu^{I\nu} W_\nu^{J\rho} W_\rho^{K\mu}$				

$X^2\varphi^2$		$\psi^2 X\varphi$		$\psi^2\varphi^2 D$	
$Q_{\varphi G}$	$\varphi^\dagger\varphi\, G_{\mu\nu}^A G^{A\mu\nu}$	Q_{eW}	$(\bar{l}_p\sigma^{\mu\nu} e_r)\tau^I\varphi W_{\mu\nu}^I$	$Q_{\varphi l}^{(1)}$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu\,\varphi)(\bar{l}_p\gamma^\mu l_r)$
$Q_{\varphi\widetilde{G}}$	$\varphi^\dagger\varphi\, \widetilde{G}_{\mu\nu}^A G^{A\mu\nu}$	Q_{eB}	$(\bar{l}_p\sigma^{\mu\nu} e_r)\varphi B_{\mu\nu}$	$Q_{\varphi l}^{(3)}$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu^I\,\varphi)(\bar{l}_p\tau^I\gamma^\mu l_r)$
$Q_{\varphi W}$	$\varphi^\dagger\varphi\, W_{\mu\nu}^I W^{I\mu\nu}$	Q_{uG}	$(\bar{q}_p\sigma^{\mu\nu} T^A u_r)\widetilde{\varphi}\, G_{\mu\nu}^A$	$Q_{\varphi e}$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu\,\varphi)(\bar{e}_p\gamma^\mu e_r)$
$Q_{\varphi\widetilde{W}}$	$\varphi^\dagger\varphi\, \widetilde{W}_{\mu\nu}^I W^{I\mu\nu}$	Q_{uW}	$(\bar{q}_p\sigma^{\mu\nu} u_r)\tau^I\widetilde{\varphi}\, W_{\mu\nu}^I$	$Q_{\varphi q}^{(1)}$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu\,\varphi)(\bar{q}_p\gamma^\mu q_r)$
$Q_{\varphi B}$	$\varphi^\dagger\varphi\, B_{\mu\nu} B^{\mu\nu}$	Q_{uB}	$(\bar{q}_p\sigma^{\mu\nu} u_r)\widetilde{\varphi}\, B_{\mu\nu}$	$Q_{\varphi q}^{(3)}$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu^I\,\varphi)(\bar{q}_p\tau^I\gamma^\mu q_r)$
$Q_{\varphi\widetilde{B}}$	$\varphi^\dagger\varphi\, \widetilde{B}_{\mu\nu} B^{\mu\nu}$	Q_{dG}	$(\bar{q}_p\sigma^{\mu\nu} T^A d_r)\varphi\, G_{\mu\nu}^A$	$Q_{\varphi u}$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu\,\varphi)(\bar{u}_p\gamma^\mu u_r)$
$Q_{\varphi WB}$	$\varphi^\dagger\tau^I\varphi\, W_{\mu\nu}^I B^{\mu\nu}$	Q_{dW}	$(\bar{q}_p\sigma^{\mu\nu} d_r)\tau^I\varphi\, W_{\mu\nu}^I$	$Q_{\varphi d}$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu\,\varphi)(\bar{d}_p\gamma^\mu d_r)$
$Q_{\varphi\widetilde{W}B}$	$\varphi^\dagger\tau^I\varphi\, \widetilde{W}_{\mu\nu}^I B^{\mu\nu}$	Q_{dB}	$(\bar{q}_p\sigma^{\mu\nu} d_r)\varphi\, B_{\mu\nu}$	$Q_{\varphi ud}$	$i(\widetilde{\varphi}^\dagger D_\mu\varphi)(\bar{u}_p\gamma^\mu d_r)$

x1008.4884



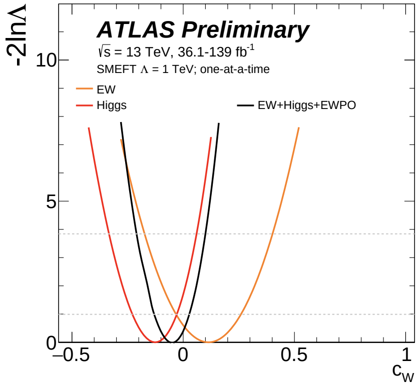
$$-\frac{i\bar{g}}{\sqrt{2}}K_{f_1f_2}\gamma^{\mu_3}P_L - 2vp_3^{\nu}K_{f_1g_1}C^{dW}_{g_1f_2}\sigma^{\mu_3\nu}P_R - \frac{i\bar{g}v^2}{\sqrt{2}}K_{f_1g_1}C^{\varphi q3}_{g_1f_2}\gamma^{\mu_3}P_L$$
$$-\frac{i\bar{g}v^2}{2\sqrt{2}}C^{\varphi ud}_{f_1f_2}\gamma^{\mu_3}P_R - 2vp_3^{\nu}K_{g_1f_2}\sigma^{\mu_3\nu}P_L C^{uW*}_{g_1f_1}$$



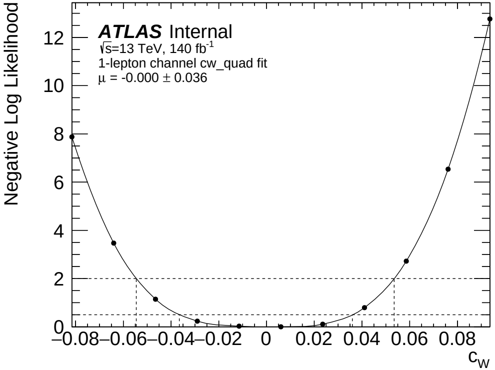
$$+\frac{i\bar{g}^2}{\sqrt{\bar{g}^2+\bar{g}'^2}}\left(\eta_{\mu_1\mu_2}(p_1-p_2)^{\mu_3} + \eta_{\mu_2\mu_3}(p_2-p_3)^{\mu_1} + \eta_{\mu_3\mu_1}(p_3-p_1)^{\mu_2}\right)$$
$$-\frac{6i\bar{g}}{\sqrt{\bar{g}^2+\bar{g}'^2}}C^{W}\left(p_3^{\mu_1}p_1^{\mu_2}p_2^{\mu_3} - p_2^{\mu_1}p_3^{\mu_2}p_1^{\mu_3} + \eta_{\mu_1\mu_2}\left(p_1^{\mu_3}p_2\cdot p_3 - p_2^{\mu_3}p_1\cdot p_3\right)\right.$$
$$\left.+\eta_{\mu_2\mu_3}\left(p_2^{\mu_1}p_1\cdot p_3 - p_3^{\mu_1}p_1\cdot p_2\right) + \eta_{\mu_3\mu_1}\left(p_3^{\mu_2}p_1\cdot p_2 - p_1^{\mu_2}p_2\cdot p_3\right)\right)$$
$$+\frac{i\bar{g}\bar{g}'v^2}{\left(\bar{g}^2+\bar{g}'^2\right)^{3/2}}C^{\varphi WB}\left(\eta_{\mu_1\mu_2}(\bar{g}'^2p_1^{\mu_3} - \bar{g}'^2p_2^{\mu_3}) + \eta_{\mu_2\mu_3}(\bar{g}'^2p_2^{\mu_1} + \bar{g}^2p_3^{\mu_1})\right.$$
$$\left.+\eta_{\mu_3\mu_1}(-\bar{g}^2p_3^{\mu_2} - \bar{g}'^2p_1^{\mu_2})\right)$$
$$-\frac{2i\bar{g}}{\sqrt{\bar{g}^2+\bar{g}'^2}}C^{\widetilde{W}}\left(\epsilon_{\mu_1\mu_2\mu_3\alpha_1}\left(p_1^{\alpha_1}p_2\cdot p_3 + p_2^{\alpha_1}p_1\cdot p_3 + p_3^{\alpha_1}p_1\cdot p_2\right)\right.$$
$$+\epsilon_{\mu_1\mu_2\alpha_1\beta_1}(p_1-p_2)^{\mu_3}p_1^{\alpha_1}p_2^{\beta_1} + \epsilon_{\mu_2\mu_3\alpha_1\beta_1}(p_2-p_3)^{\mu_1}p_2^{\alpha_1}p_3^{\beta_1}$$
$$\left.+\epsilon_{\mu_3\mu_1\alpha_1\beta_1}(p_3-p_1)^{\mu_2}p_3^{\alpha_1}p_1^{\beta_1}\right)$$
$$-\frac{i\bar{g}\bar{g}'v^2}{\sqrt{\bar{g}^2+\bar{g}'^2}}C^{\varphi\widetilde{W}B}\epsilon_{\mu_1\mu_2\mu_3\alpha_1}p_3^{\alpha_1}$$

https://arxiv.org/pdf/1704.03888.pdf

EFT Fit



Current ATLAS global limits



Our 1-lep fit (asimov, no systematics)

Posterior Form

$$p(\mathbf{y}_*|X,\mathbf{y},X_*) \sim \mathcal{N}(\mu,\Sigma)$$

$$\begin{aligned}\boldsymbol{\mu} &= \boldsymbol{\mu}_0(X_*) + K_{X_*X}K^{-1}(\mathbf{y} - \boldsymbol{\mu}_0(X)),\\ \Sigma &= K_{X_*X_*} - K_{X_*X}K^{-1}K_{XX_*}\\ K &\equiv K_{XX} + \text{diag}(\boldsymbol{\sigma}^2)\end{aligned}$$

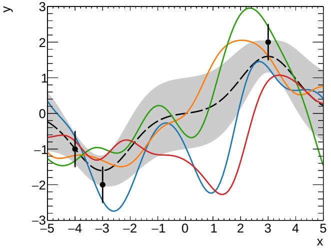
X	training inputs	X_*	test inputs
$\mathbf{y}$	training targets	$\mathbf{y}_*$	predicted values
$\boldsymbol{\sigma}^2$	noise on $\mathbf{y}$		
$\boldsymbol{\mu}_0$	mean of the prior	$K_{XX'}$	covariance of the prior, evaluated between X and X'

$$p(\mathbf{y}) = \frac{1}{\sqrt{(2\pi)^n|K|}}\exp\left[-\frac{1}{2}\mathbf{y}^TK^{-1}\mathbf{y}\right]$$

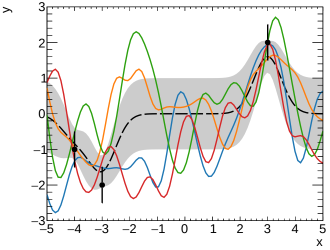
Example with toy data, radial basis function $\implies$ 2 hyperparameters

$$k(x,x') = \gamma^2 \times \exp\left[-\frac{(x-x')^2}{2\ell^2}\right]$$

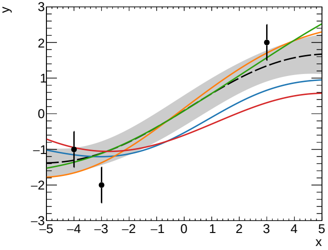
- γ overall variance
- ℓ length scale



$\ell = 1, \gamma = 1$



$\ell = 0.5, \gamma = 1$

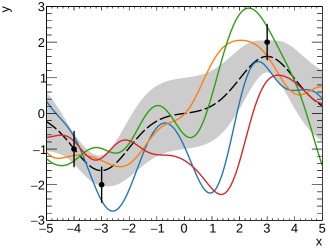


$\ell = 5, \gamma = 1$

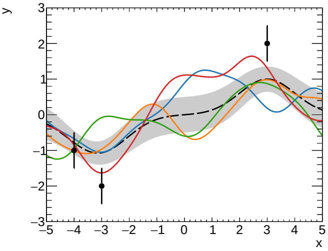
Example with toy data, radial basis function $\implies$ 2 hyperparameters

$$k(x, x') = \gamma^2 \times \exp\left[-\frac{(x - x')^2}{2\ell^2}\right]$$

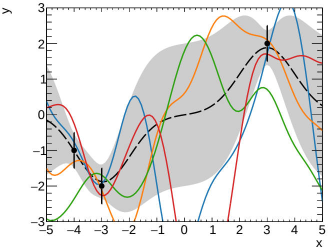
- γ overall variance
- ℓ length scale



$\ell = 1, \gamma = 1$



$\ell = 1, \gamma = 0.5$



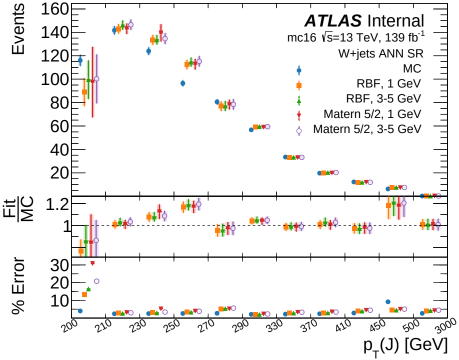
$\ell = 1, \gamma = 2$

- Isotropic kernels (functions of $r = |\mathbf{x}_1 - \mathbf{x}_2|$) behave very similarly
- RBF:

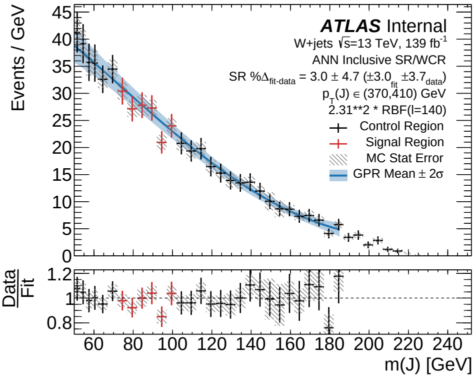
$$k(r) = \gamma^2 \times \exp\left[-\frac{r^2}{2\ell^2}\right]$$

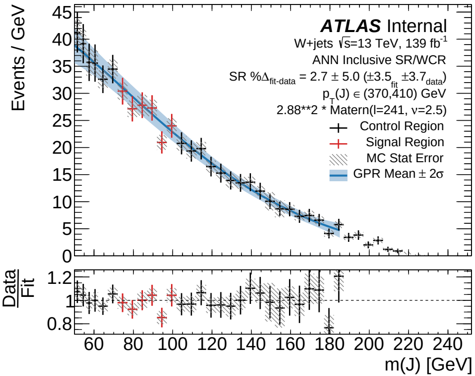
- Matern 5/2:

$$k(r) = \gamma^2 \left(1 + \frac{\sqrt{5}r}{\ell} + \frac{5r^2}{3\ell^2}\right) \exp\left(-\frac{\sqrt{5}r}{\ell}\right)$$



Alternative Kernels





RBF Kernel

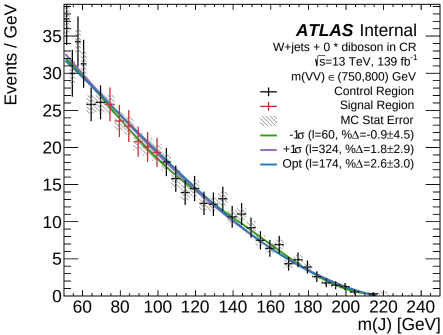
Matern 5/2 Kernel

$$k(r) = \gamma^2 \times \exp\left[-\frac{r^2}{2\ell^2}\right]$$

$$k(r) = \gamma^2 \left(1 + \frac{\sqrt{5}r}{\ell} + \frac{5r^2}{3\ell^2}\right) \exp\left(-\frac{\sqrt{5}r}{\ell}\right)$$

Systematic Errors

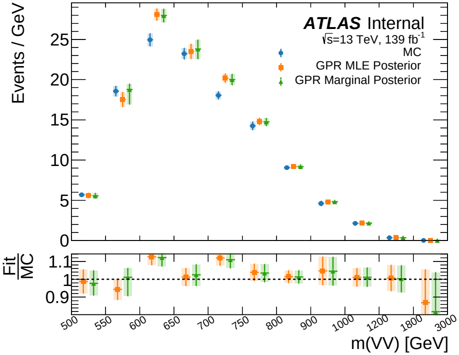
- How dependent is the output on the choice of kernel?
  - A little, $\approx 2\%$ in example on right
  - Which to choose? Error on choice?
- A familiar idea:
  - Don't choose a single kernel
  - Average over all kernels / hyperparameters
  - Weight by how well each kernel fits the data
- Another round of Bayes' rule!
  - Jargon: marginalizing the kernel choice
  - Net effect: slightly larger errors



Posterior modes with 3 different hyperparameter settings

Systematic Errors

- How dependent is the output on the choice of kernel?
  - A little, $\approx 2\%$ in example on right
  - Which to choose? Error on choice?
- A familiar idea:
  - Don't choose a single kernel
  - Average over all kernels / hyperparameters
  - Weight by how well each kernel fits the data
- Another round of Bayes' rule!
  - Jargon: marginalizing the kernel choice
  - Net effect: slightly larger errors



Summary of yields from multiple $m(VV)$ bins

Marginalize the Posterior

- Bayesian is awesome! Marginalize hyperparameters out from posterior
- $\boldsymbol{\theta} = (\ell, \gamma)$ list of hyperparameters
- Add explicit $\boldsymbol{\theta}$ dependence on previous formulas

$$\underbrace{p(f|\mathbf{y})}_{\text{marginal posterior}} = \int \mathrm{d}\boldsymbol{\theta} \underbrace{p(f|\mathbf{y},\boldsymbol{\theta})}_{\text{original posterior}} \times \underbrace{p(\boldsymbol{\theta}|\mathbf{y})}_{\text{hyperposterior}}$$

$$\underbrace{p(\boldsymbol{\theta}|\mathbf{y})}_{\text{hyperposterior}} = \frac{\overbrace{p(\mathbf{y}|\boldsymbol{\theta})}^{\text{marginal likelihood}} \times \overbrace{p(\boldsymbol{\theta})}^{\text{hyperprior}}}{\underbrace{p(\mathbf{y})}_{\text{normalizing constant}}}$$

Example: Marginal Likelihood

$$p(f|\mathbf{y}) \propto \int \mathrm{d}\boldsymbol{\theta}\, p(f|\mathbf{y},\boldsymbol{\theta})p(\mathbf{y}|\boldsymbol{\theta})$$

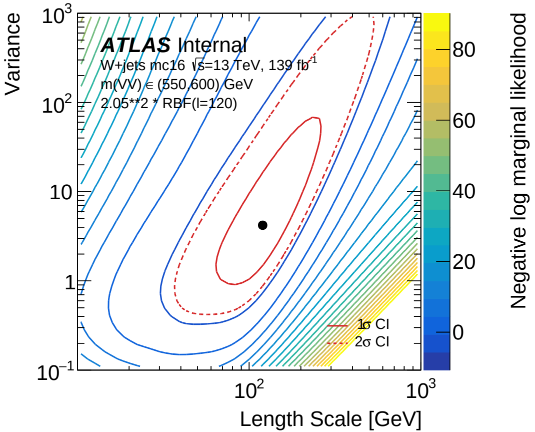
- RBF kernel, so two hyperparameters γ and ℓ

$$k(\mathbf{x}_1,\mathbf{x}_2) = \gamma^2 \cdot \exp\left[-\frac{|\mathbf{x}_1-\mathbf{x}_2|^2}{2\ell^2}\right]$$

- Black dot is maximum (MMLE)
- Technically, we should be looking at

$$\underbrace{p(\boldsymbol{\theta}|\mathbf{y})}_{\text{hyperposterior}} \propto \underbrace{p(\mathbf{y}|\boldsymbol{\theta})}_{\text{marginal likelihood}} \times \underbrace{p(\boldsymbol{\theta})}_{\text{hyperprior}}$$

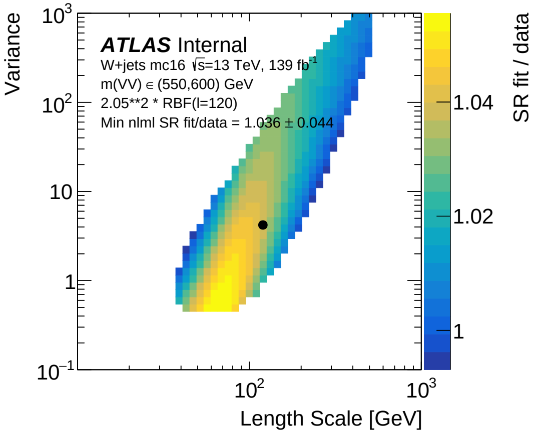
- But assuming $p(\boldsymbol{\theta})$ ~log-constant in grid, reduces to marginal likelihood



Example: Posterior

$$p(f|\mathbf{y}) \propto \int \mathrm{d}\boldsymbol{\theta}\, p(f|\mathbf{y},\boldsymbol{\theta})\, p(\mathbf{y}|\boldsymbol{\theta})$$

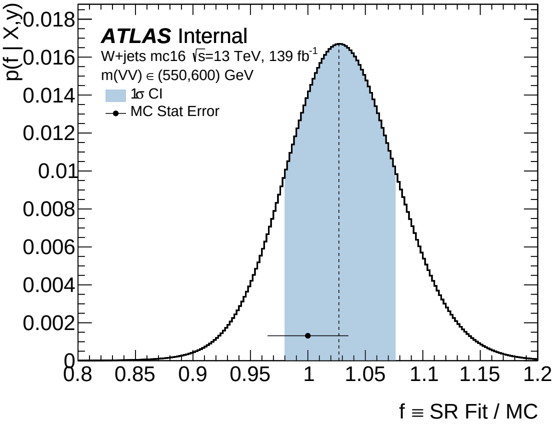
- This is the usual output of GPR, evaluated at each $\boldsymbol{\theta}$
- Keep in mind the output is actually a Gaussian distribution
- Here we show just the mean scaled by MC event count
- For simplicity, only evaluate within 1σ or 2σ CI



Example: Marginal Posterior

$$p(f|\mathbf{y}) \propto \int \mathrm{d}\boldsymbol{\theta}\, p(f|\mathbf{y},\boldsymbol{\theta})p(\mathbf{y}|\boldsymbol{\theta})$$

- Evaluate integral at multiple f to get full distribution
- Pick mode and 1σ CI as predicted values
- Repeat in each $m(VV)$ bin



GPR Signal Contamination

- Fit: $N_{\text{W+jets}} = N_{\text{data}} - N_{\text{top},t\bar{t}} - \mu_{\text{guess}} N_{\text{diboson}}$
  - Take $\mu_{\text{guess}} = 1$
  - But if $\mu_{\text{guess}} \neq \mu_{\text{truth}}$, signal contamination!
- Test: vary μ_{truth} and fit GPR to

$$N^{\text{MC}}_{\text{W+jets}} + (\mu_{\text{truth}} - 1) N^{\text{MC}}_{\text{diboson}}$$

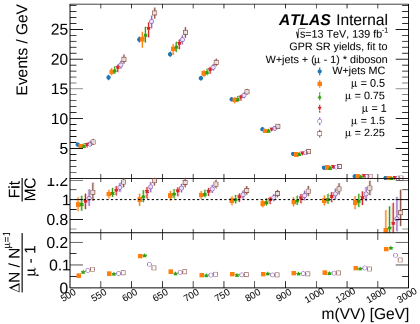
- Idea: keep $\mu_{\text{guess}} = 1$ in GPR, correct in likelihood fit
  - Simply subtract out correlation term:

$$N_{\text{no contam}} = N^{\text{fit yield}}_{\text{contam}} - (\mu - 1)\Delta N$$

  - ΔN a "nuisance histogram" determined from right
  - As fitter varies μ, automatically corrected

GPR yield ≈ linear vs μ_{truth}:

$$N^{\mu}_{\text{GPR}} = (\mu - 1)\Delta N + N^{\mu=1}_{\text{GPR}}$$



	$\mu_{\text{truth}} = 1$	$\mu_{\text{truth}} = 1.5$
No correction	$1.00 \pm .10$	$1.40 \pm .10$
With correction	$1.00 \pm .13$	$1.50 \pm .12$

Unfolding

- Unfolding: undo detector-specific effects $\rightarrow$ "truth" distribution
  - Adjust for detector efficiencies and resolutions
  - So people outside of ATLAS can use measurement
- Likelihood-based approach:
  - Assume linear detector response, determine matrix from simulation

$$N_i^{\text{Detector}} = R_{ij} N_j^{\text{Truth}}$$

  - Treat each truth bin as separate "signal samples" in likelihood fit
  - "Signal strength" == unfolded event counts N_j^{Unfolded}

$$\text{Normal Likelihood:} \quad L(\text{data} \mid \mu) = \prod_i P(\text{data}_i \mid \text{bkgs}_i + \mu \cdot \text{signal}_i)$$

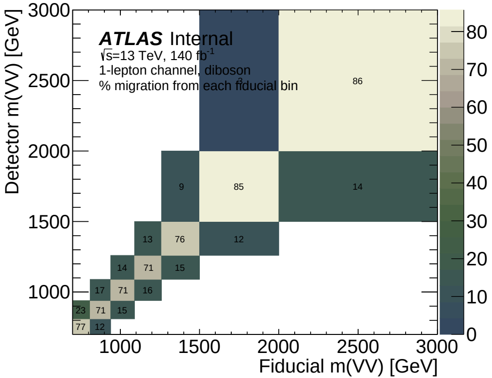
$$\text{Unfolding Likelihood:} \quad L(\text{data} \mid \vec{N}^{\text{Unfolded}}) = \prod_i P(\text{data}_i \mid \text{bkgs}_i + R_{ij} N_j^{\text{Unfolded}})$$

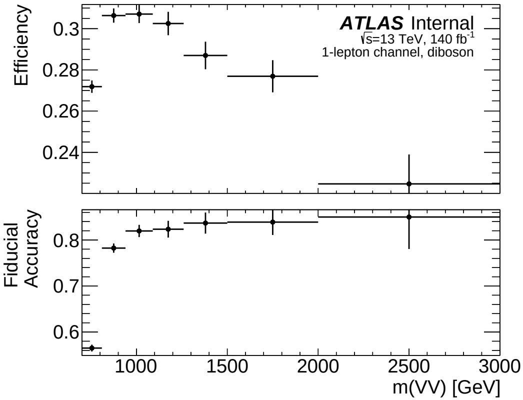
Response Matrix Components

Response matrix	$R_{ij} = \frac{\epsilon_j}{\alpha_i} M_{ij}$
Migration matrix	$M_{ij} = \frac{N_{ij}^{\text{Reco} \bigcap \text{Truth}}}{N_j^{\text{Reco} \bigcap \text{Truth}}}$
Efficiency	$\epsilon_j = \frac{N_j^{\text{Reco} \bigcap \text{Truth}}}{N_j^{\text{Truth}}}$
Acceptance	$\alpha_i = \frac{N_i^{\text{Reco} \bigcap \text{Truth}}}{N_i^{\text{Reco}}}$

i: reco bin j: truth bin

Response Matrix Components





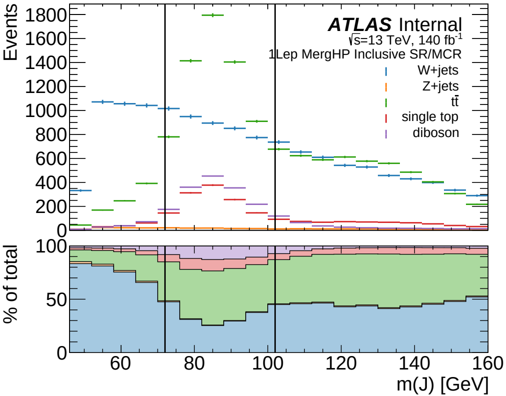
Unfolding Workflow

MCR is not pure V+jets. Method:

1. Fit top and $t\bar{t}$ to TCR (99% pure, use nominal for other samples)
2. Subtract top, $t\bar{t}$, VV from MCR

$$N_{V\text{+jets}} = N_{\text{data}} - \mu_{t\bar{t}} N_{t\bar{t}}^{\text{MC}} - \mu_t N_t^{\text{MC}} - \mu_{VV} N_{VV}^{\text{MC}}$$

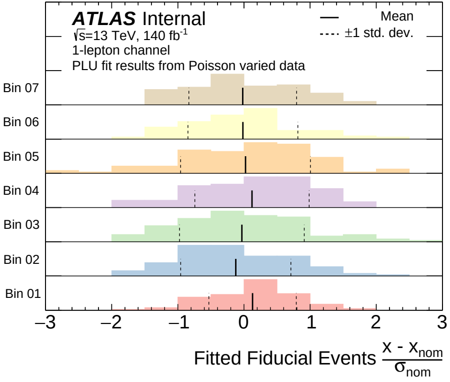
3. Fit GPR to event-subtracted $N_{V\text{+jets}}$ to get SR yield
4. Repeat steps 2–3 for various μ_{VV} – determine signal contamination correction (slides from before)
5. Repeat steps 1–3 for $\pm 1\sigma$ of all systematics and $\mu_{t\bar{t}}$, μ_t
6. Repeat for all bins in $m(VV)$ / other variable



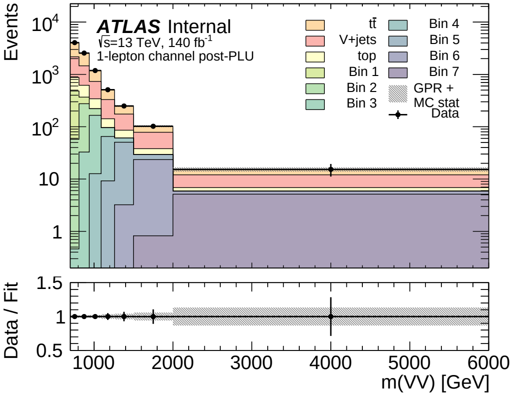
End result: GPR yield histograms in $m(VV)$ for each $\pm 1\sigma$ systematic and $\pm 1\sigma$ on $\mu_{t\bar{t}}$, μ_t

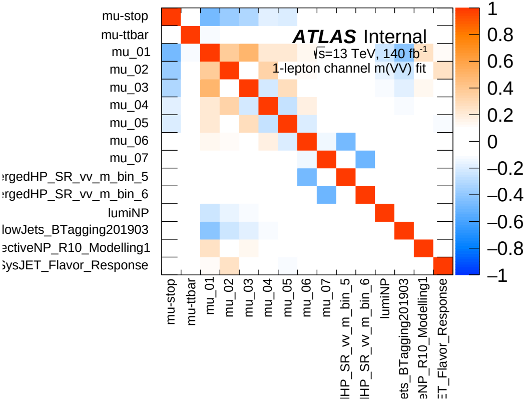
PLU Toy Test

- Vary the data in each bin by Poisson statistics
- Do PLU fit, repeat 100 times
- Plot difference against original yields
- Confirms no bias, accurate uncertainties

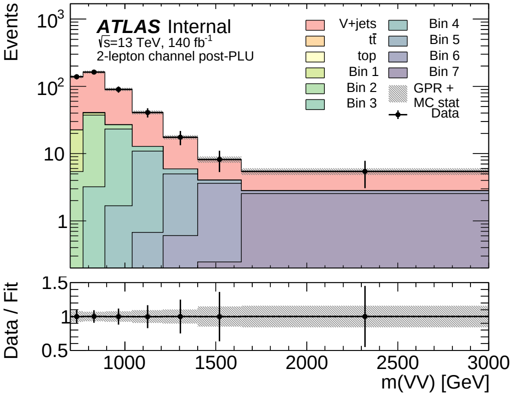


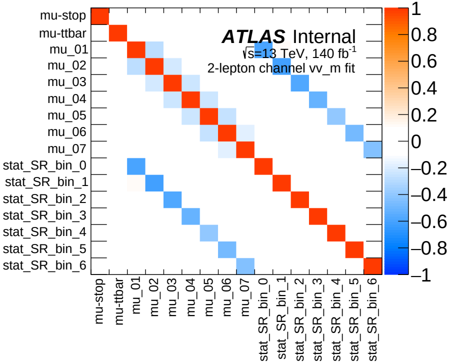
PLU Fit Correlations (1-lep)



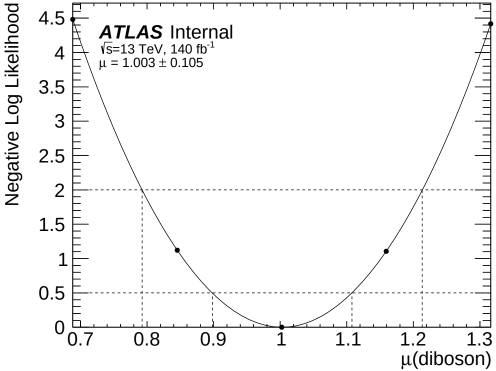


PLU Fit Correlations (2-lep)





SM Diboson Cross Section



Systs.	Channels	μ	σ
✖	0+1+2	0.999 ± 0.105	9.5
	0	1.004 ± 0.132	7.6
	1	1.064 ± 0.156	6.8
	2	1.001 ± 0.212	4.7
✓	1+2	1.038 ± 0.164	6.3
	1	1.058 ± 0.217	4.9
	2	1.012 ± 0.218	4.6