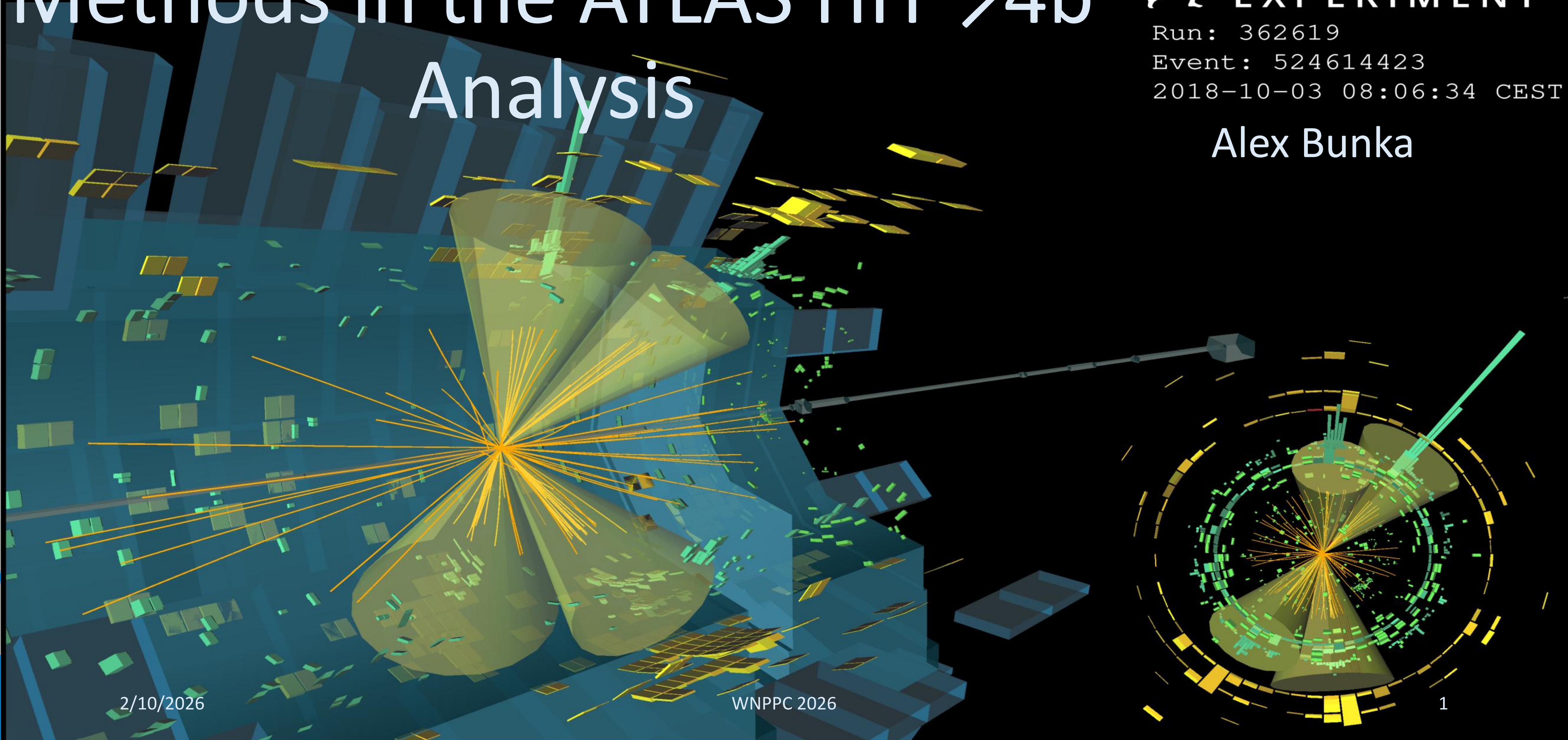
New Background Uncertainty Methods in the ATLAS HH→4b Analysis

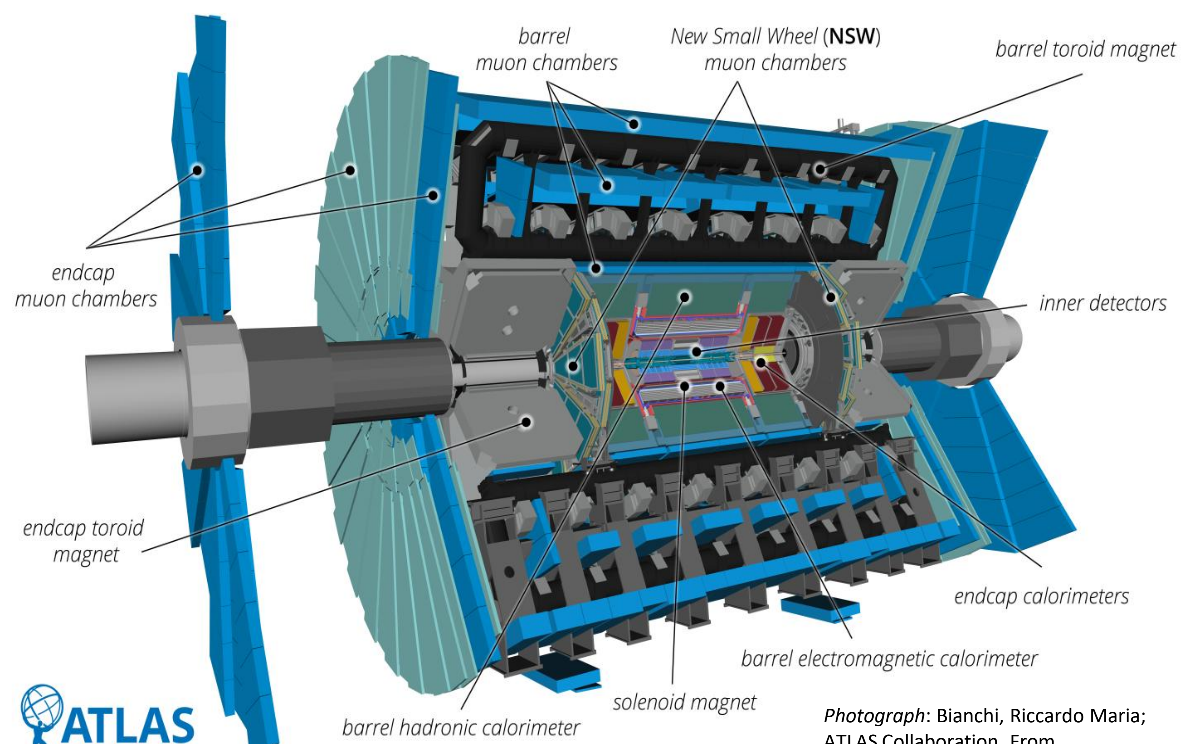
Image Credit



Alex Bunka

ATLAS and the LHC

- LHC is a 27km proton collider
  - Located near Geneva, Switzerland
  - Highest energy particle collisions ever achieved
- ATLAS is the largest general purpose particle detector ever built
  - 25m diameter, 44m length
  - Measures position, momentum and energy of billions of particles every second!



Photograph: Bianchi, Riccardo Maria; ATLAS Collaboration. From https://cds.cern.ch/record/2837191

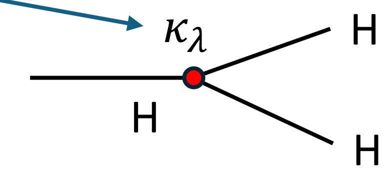
ATLAS EXPERIMENT

THE UNIVERSITY OF BRITISH COLUMBIA

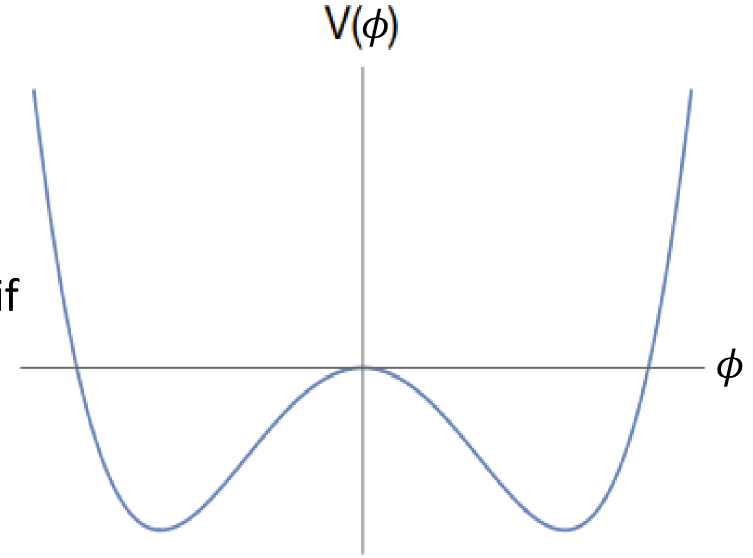
The Higgs Boson

$$V = V_0 + \frac{1}{2} m_h^2 h^2 + \frac{m_h^2}{2\nu} h^3 + \ldots$$

$$\lambda_{hhh}^{SM} = \frac{m_h^2}{2\nu}, \kappa_\lambda = \frac{\lambda_{hhh}}{\lambda_{hhh}^{SM}}$$



This is the general form after electroweak symmetry breaking, but if $\kappa_\lambda \neq 1$, it's different!



- First theorized in 1964, elementary particles gain mass by interacting with the Higgs field
  - Higgs potential includes Higgs boson mass, Higgs field, h^3 term predicts 3 Higgs coupled together
- Discovered by ATLAS and CMS (another LHC experiment) in 2012
  - Many accurate measurements of different properties, not enough data to extricate trilinear Higgs coupling

ATLAS EXPERIMENT

https://doi.org/10.1103/PhysRevLett.13.508 Higgs theory paper!
https://doi.org/10.1016/j.physletb.2012.08.020 Higgs discovery paper (ATLAS)!

THE UNIVERSITY OF BRITISH COLUMBIA

Motivation: HH->4b

- Triangle diagram has dependence on κ_λ, the HHH coupling strength.
- Box and triangle diagrams interfere destructively – harder to isolate!
- Insights to Electroweak Symmetry Breaking and BSM physics!

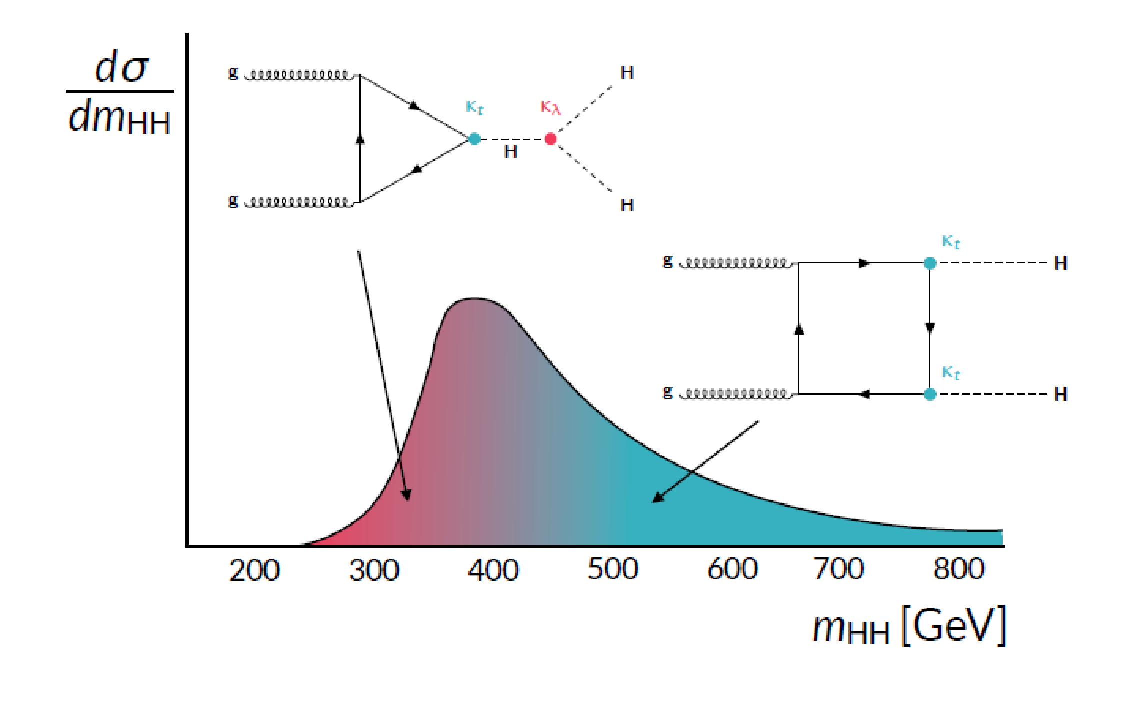


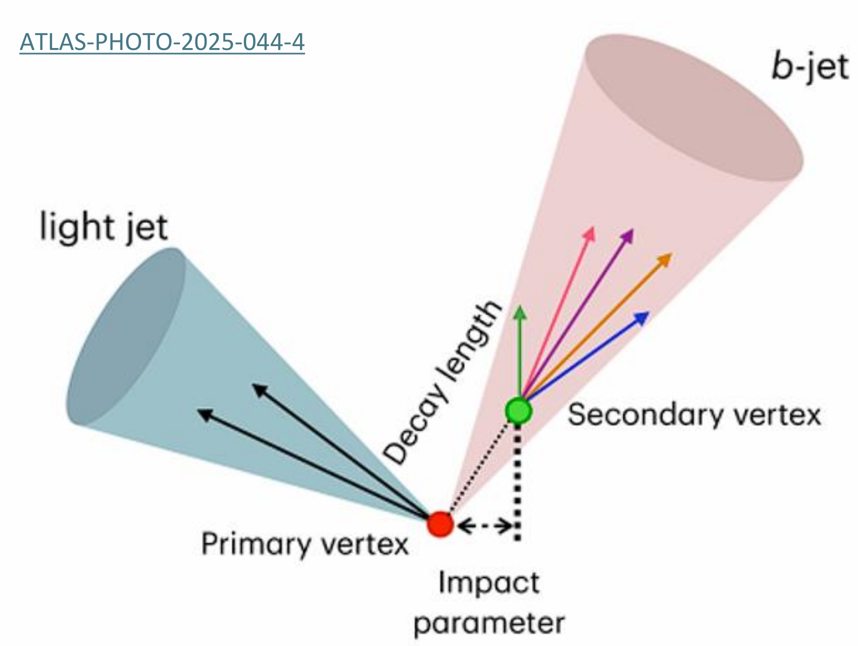
Figure from:
The ATLAS Collaboration, Search for nonresonant pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state in pp collisions at $\sqrt{s}$ =13 TeV with the ATLAS detector, Phys. Rev. D 108, 052003 (2023)
https://journals.aps.org/prd/abstract/10.1103/PhysRevD.108.052003

ATLAS EXPERIMENT

THE UNIVERSITY OF BRITISH COLUMBIA

Branching Ratios and B Tagging

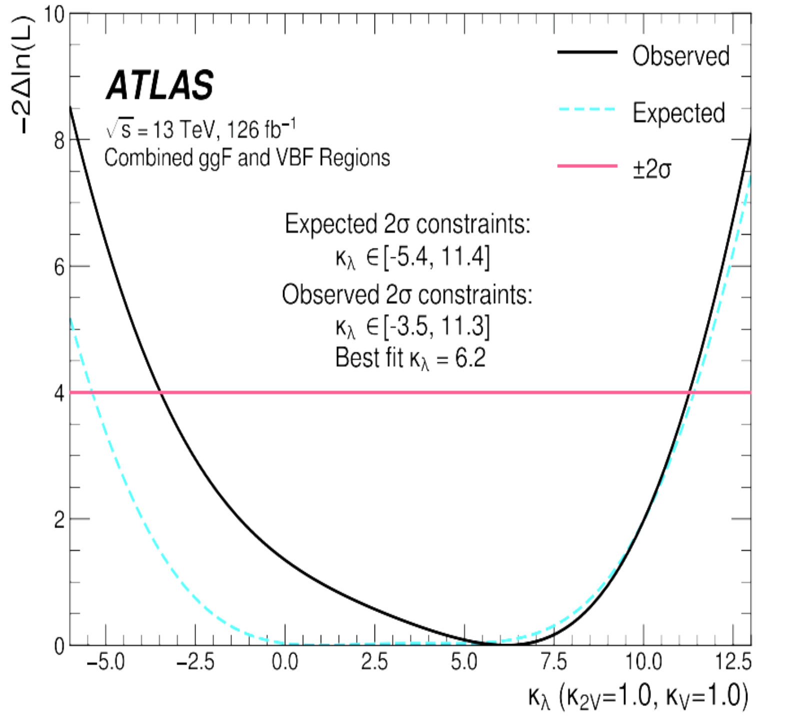
- Higgs Bosons decay quickly, with many different decay channels
  - The most common decay is to a pair of bottom quarks, $H \rightarrow bb$, at 58%. ~1/3 of HH pairs decay to 4b
- 4b final state is the most common! But it has the most background
  - B quarks show up as particle jets, and the ATLAS detector sees a lot of jets from many different sources
- Data is sorted by the number of b quark jets. Jets are b tagged based on different features, primarily the long lifetime of b quark hadrons



ATLAS EXPERIMENT

THE UNIVERSITY OF BRITISH COLUMBIA

Previous Results



(a)

Source of Uncertainty	$\Delta\mu/\mu$
Theory uncertainties	
Theory uncertainty in signal cross-section	−9.0%
All other theory uncertainties	−1.4%
Background modeling uncertainties	
Bootstrap uncertainty	−7.1%
CR to SR extrapolation uncertainty	−7.5%
3b1f nonclosure uncertainty	−2.0%

- Run 2 analysis, 2023:
  - Broad κ_λ constraints, -3 $\longrightarrow$11!
  - No evidence of signal
- Total background uncertainty was the largest in the analysis, which greatly hinders the final results.

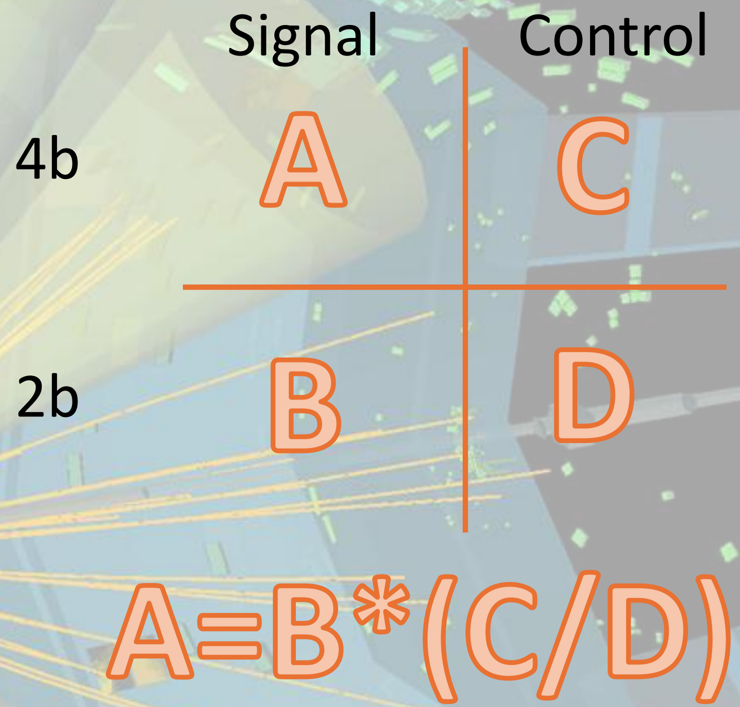
Improve uncertainty,
Improve limit!

Figures and results from:
The ATLAS Collaboration, Search for nonresonant pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state in pp collisions at √s =13 TeV with the ATLAS detector, Phys. Rev. D 108, 052003 (2023)
https://journals.aps.org/prd/abstract/10.1103/PhysRevD.108.052003

THE UNIVERSITY OF BRITISH COLUMBIA

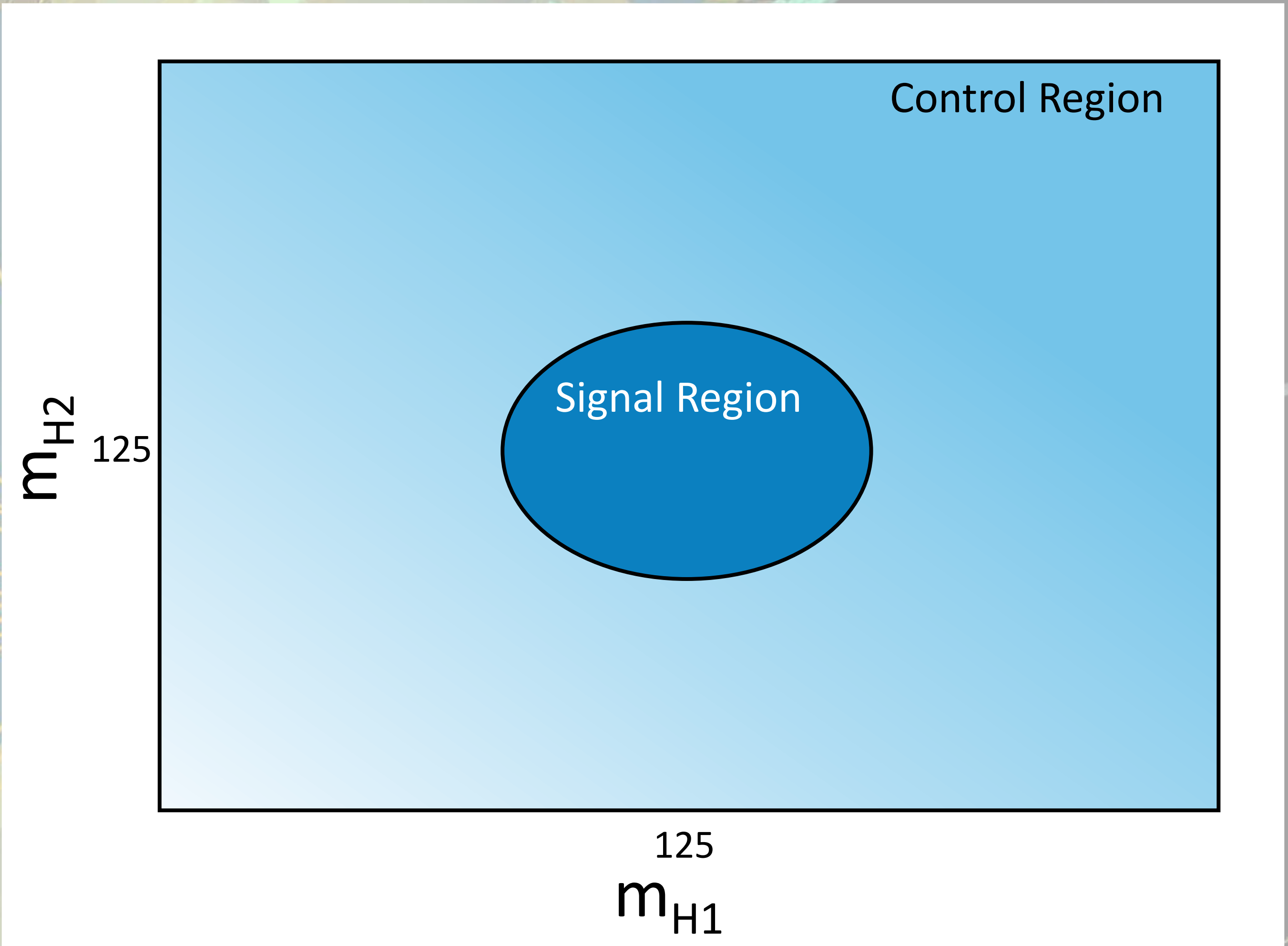
Past Methods

- The background is primarily multijet events
  - Very complicated QCD calculations are difficult to model accurately
- Previously: 2b to 4b ABCD
  - Learn the relationship between 2b and 4b backgrounds in a control region
  - Apply to the 2b signal region to estimate the 4b SR
- Sources of background uncertainties
  - Uncertainty – stdev of 100 reweighting functions, statistical uncertainty in 2b, 3b-4b non closure…
- What can we do to improve this?



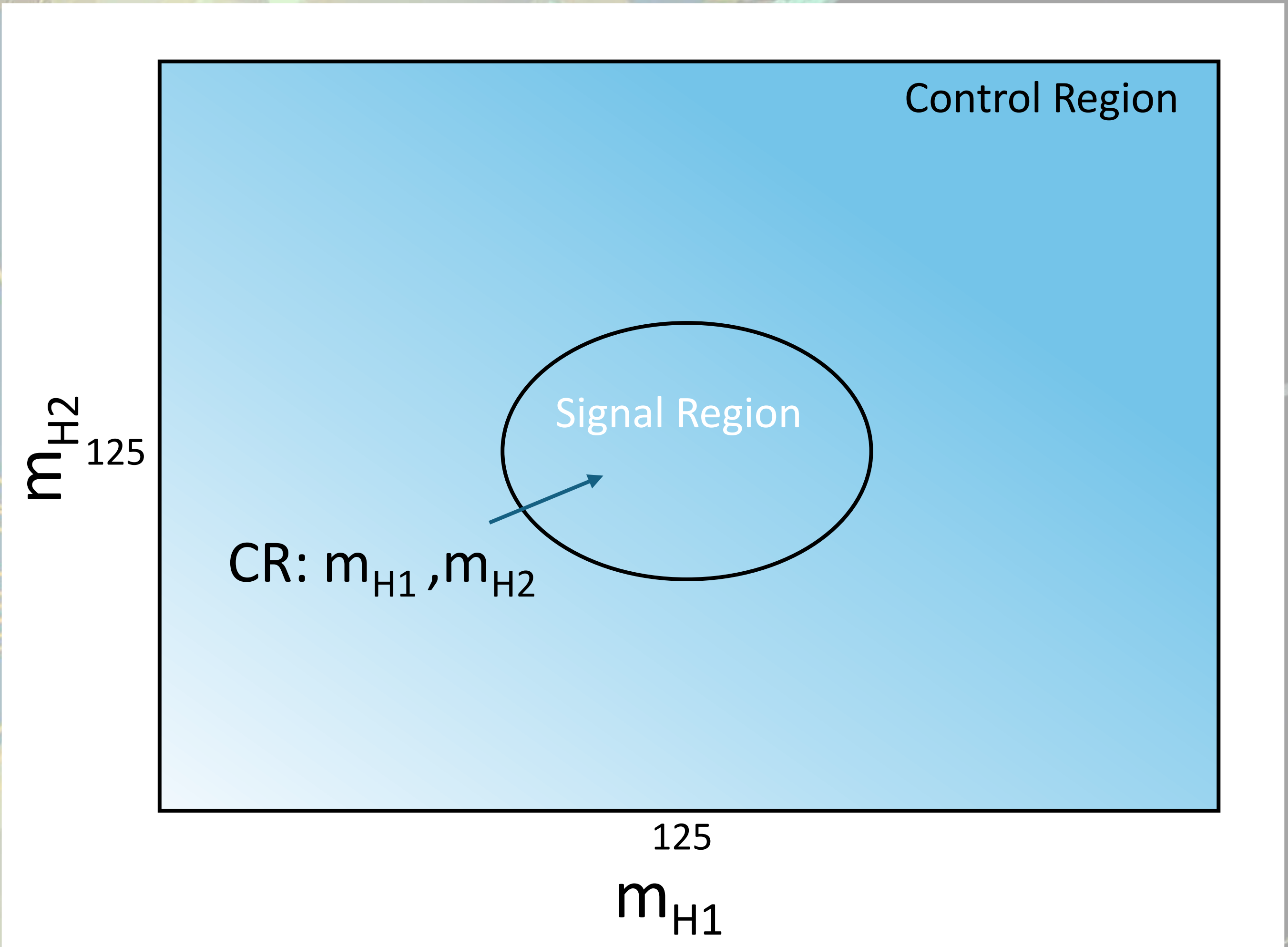
New Background Methods

- **We know the control region, but the signal region is blinded.**
- Higgs mass is 125 GeV, so the signal region is an ellipse centered around (x, y) = (125, 125).
- Outside of this ellipse is the control region.



New Background Methods

- **Fit the 2d histogram of masses to smoothly fill in the signal region, which gives m_{H1} & m_{H2} for the signal.**
- Learn how to "move" between different m_{H1}, m_{H2} points in the control region
- Interpolate from the control region to the signal region

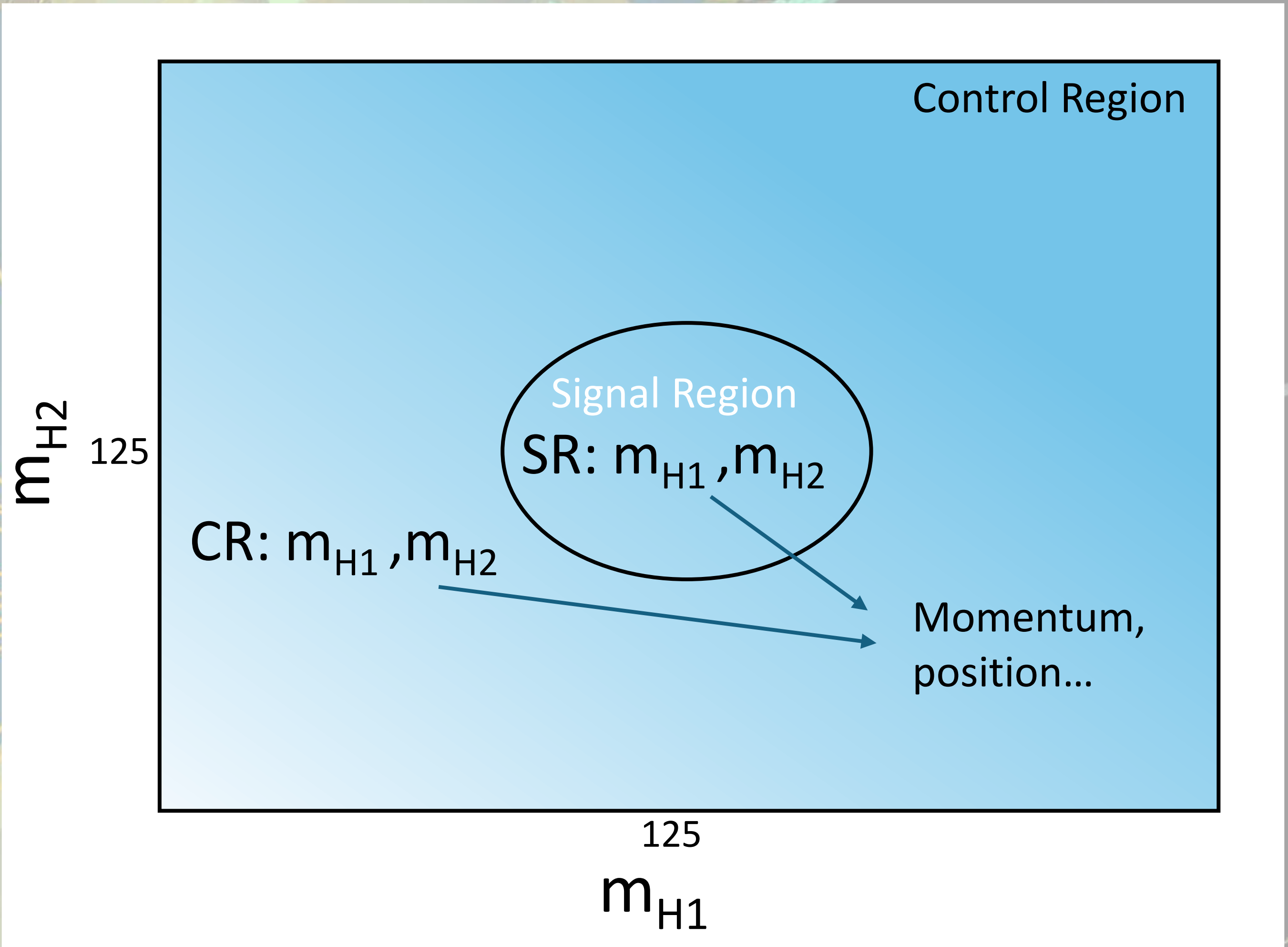


ATLAS EXPERIMENT

THE UNIVERSITY OF BRITISH COLUMBIA

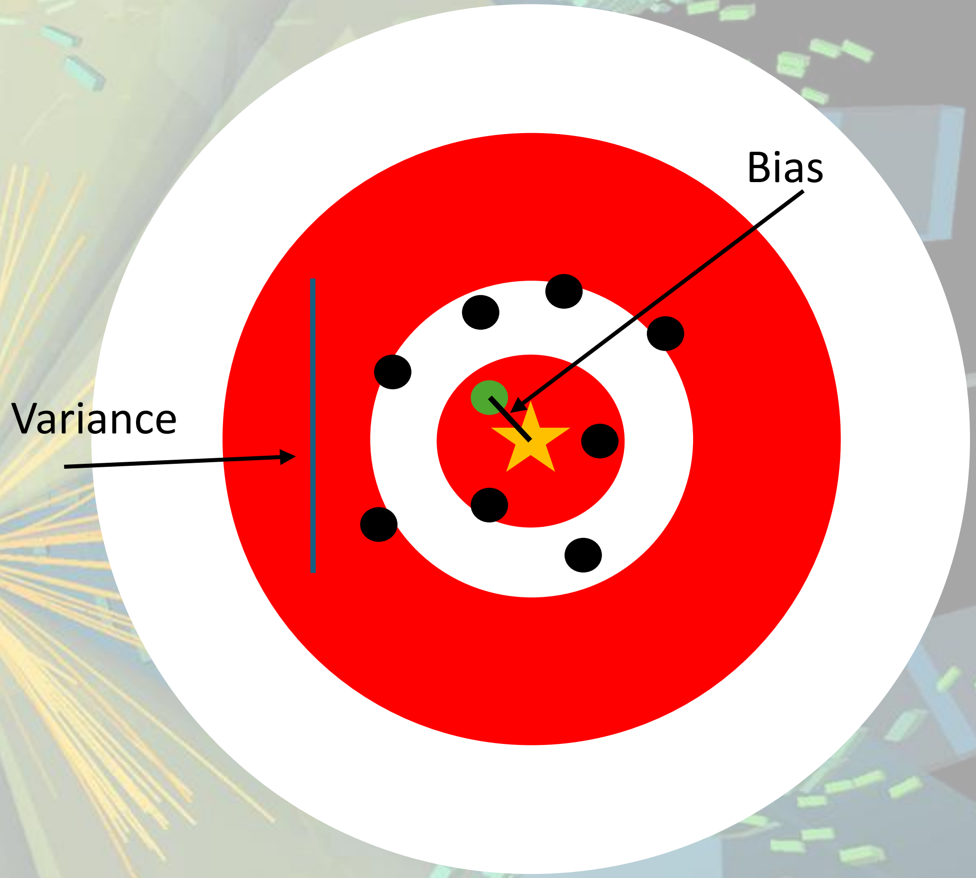
New Background Methods

- **Use a generative model to morph these first simple variables into other physics features**
- Start in the signal region, and find the relationship between the mass plane and other physics distributions
- Apply this relationship to the signal region mass plane from the previous step



Why Use 3b for Uncertainty?

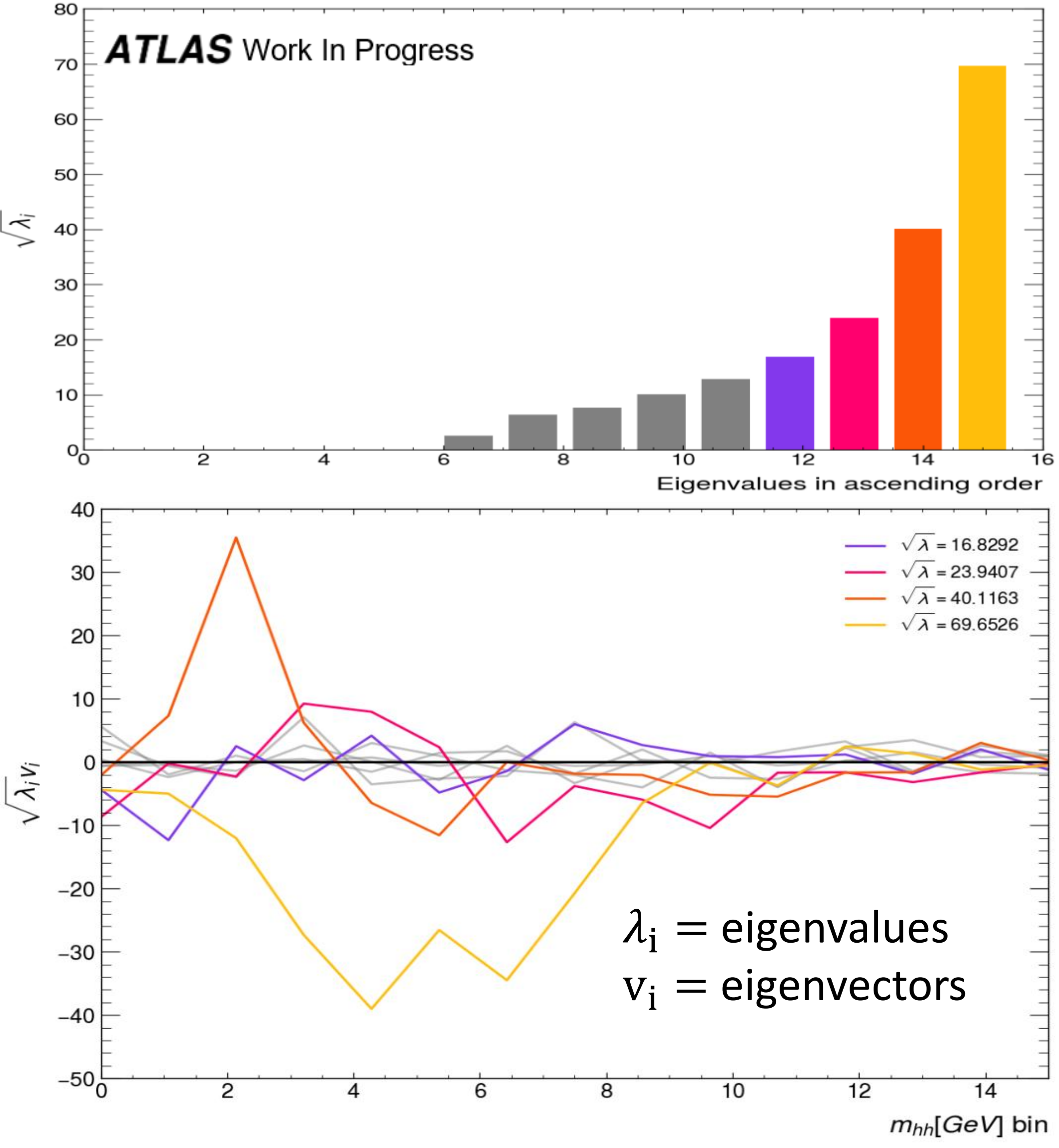
- The statistics of 4b are limited
  - Roughly ten times more data passes 3b requirements than 4b
  - If we randomly downsample the 3b data, we have similar statistical effects with ten versions to show variations
  - Majority of 3b data is 4b that failed one b-tag
- 4b is blinded
  - We cannot look at the 4b signal region data until we have all the background modelling complete
  - Can't risk biasing our analysis work!



ATLAS EXPERIMENT

THE UNIVERSITY OF BRITISH COLUMBIA

Background Uncertainty: 3b Predictions



- Use 3b data for the systematics, $N_{3b} = 10N_{4b}$, create 10 subsets.
- Train a background estimate from each of the subsets (call these replicas).
- Principle Component Analysis on the average of the replicas.
  - Solve an eigenvalue problem:
  - Eigenvectors are variations in shape, eigenvalues are the magnitude of the changes.

THE UNIVERSITY OF BRITISH COLUMBIA

Uncertainty in Two Parts

The Variance Step:

- How much do the individual replicas differ from their average?
- Unconstrained fit of each replica vs average, then iteratively add EVs.
- Continue till pulls are uncorrelated.

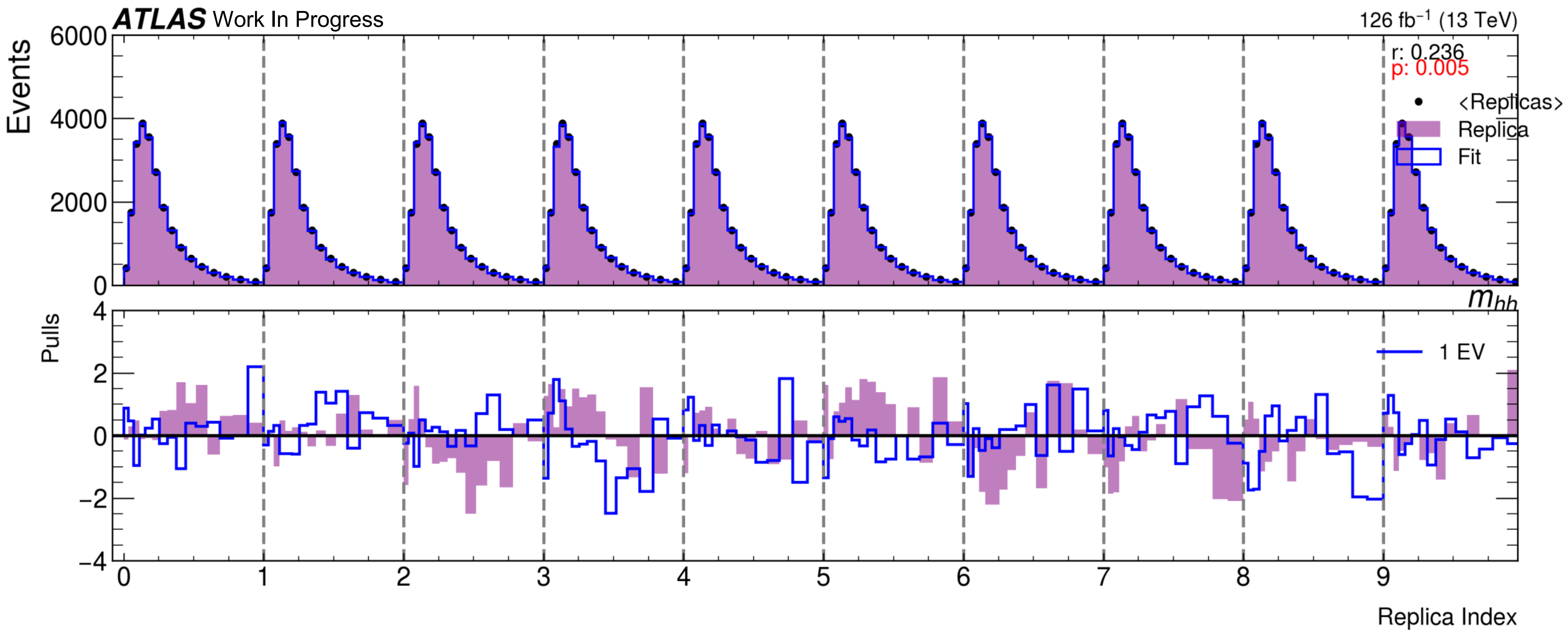
The Bias Step:

- How much does the average differ from data?
- Use EVs as nuisance parameters, constrained from variance step.
- Unconstrain NPs until the fit is good.

ATLAS EXPERIMENT

THE UNIVERSITY OF BRITISH COLUMBIA

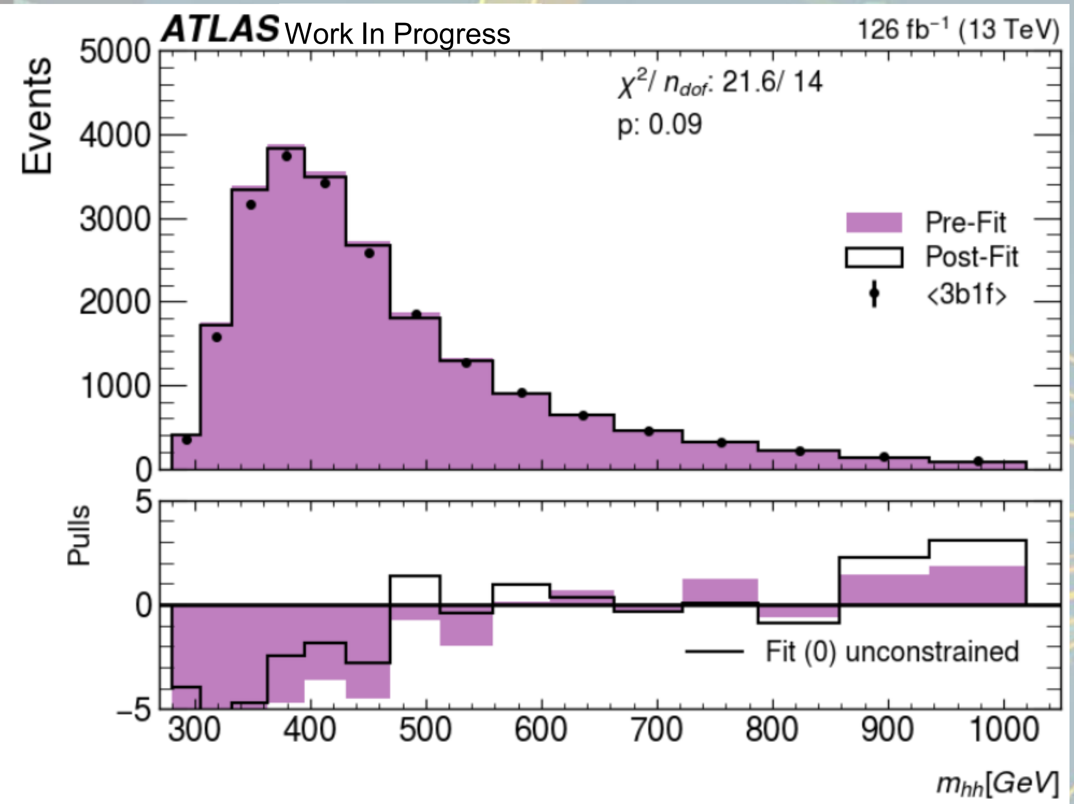
Variance: Fitting Against Average



The Bias Step

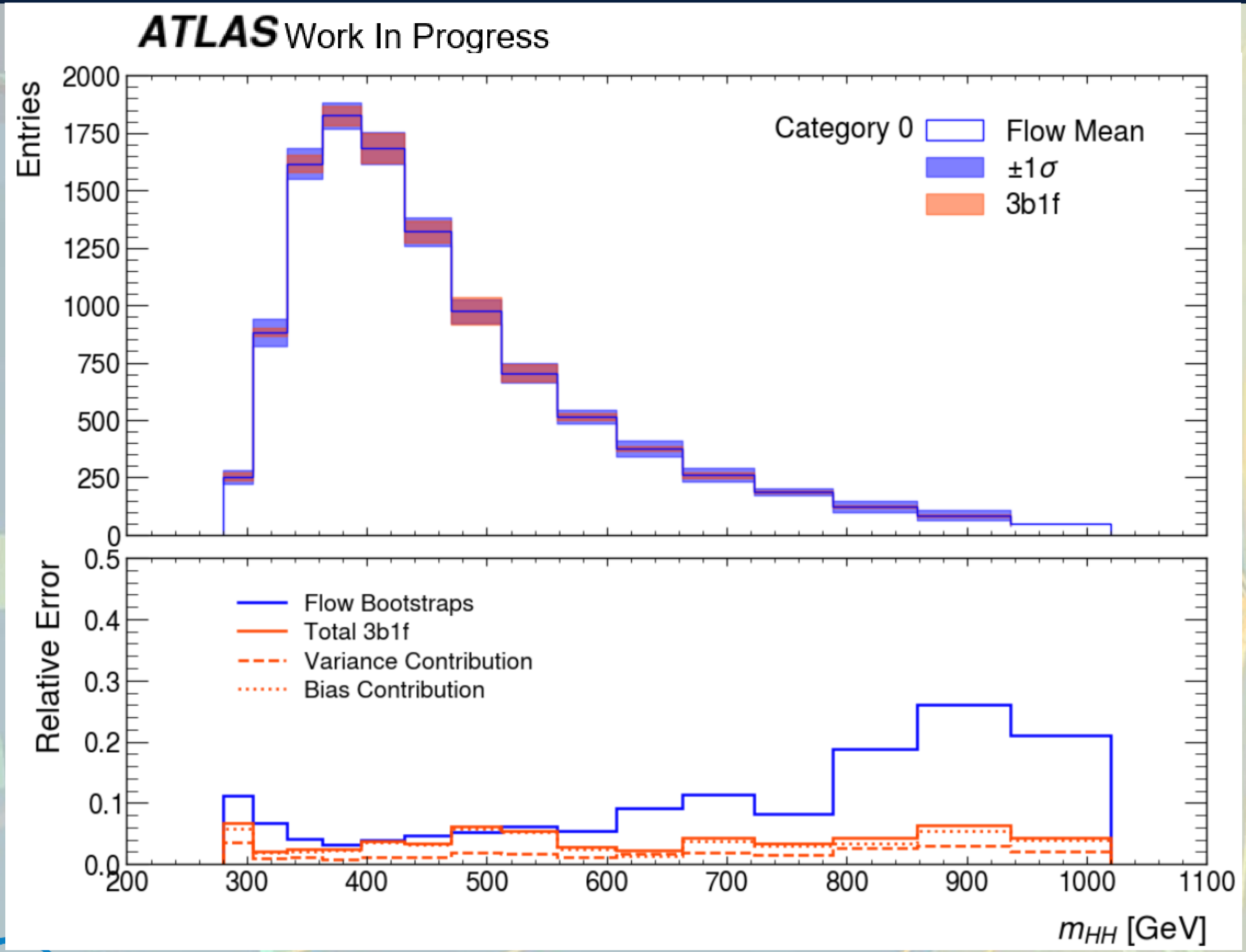
- Compare the average of the replicas to real 3b data
- Use EVs as nuisance parameters, constrained from variance step
  - If the lowest correlation was after 6 fits, we have 6 EVs to use
- Unconstrain NPs until the fit is good
  - Different statistical methods tested: Chi2, F-test…
- Loosened fit value and variance uncertainty (from before) combined for final uncertainty

Bias: Fitting Against Data



- Typically achieve closure (p>0.05) with 0 or 1 NP
  - Does not change with different stats methods
- The bias step contributes more to the final uncertainty than the variance
  - Even though it converges in fewer fit steps, the fit value is much larger

Total Uncertainty



- Red lines are the uncertainties from each step, and combined
- Blue is the stdev of 25 trainings
- Uncertainty less than or equal to that of older methods
- Reduced background uncertainty will greatly improve our analysis power!

Summary

- The HH->4b analysis is important to investigate the behaviour of the Higgs Boson and study BSM physics.
- The power of this analysis can be limited by large uncertainties, and improvements are needed keep up with the increase in data.
- Many updates have been made to the background prediction and uncertainty.
- The variance/bias method works to produce a reasonable background uncertainty.
  - Smaller than earlier methods, looks to cover the required range.
- Ideally, this will be used in both the partial Run 3 and full Run 3 analyses.

ATLAS EXPERIMENT

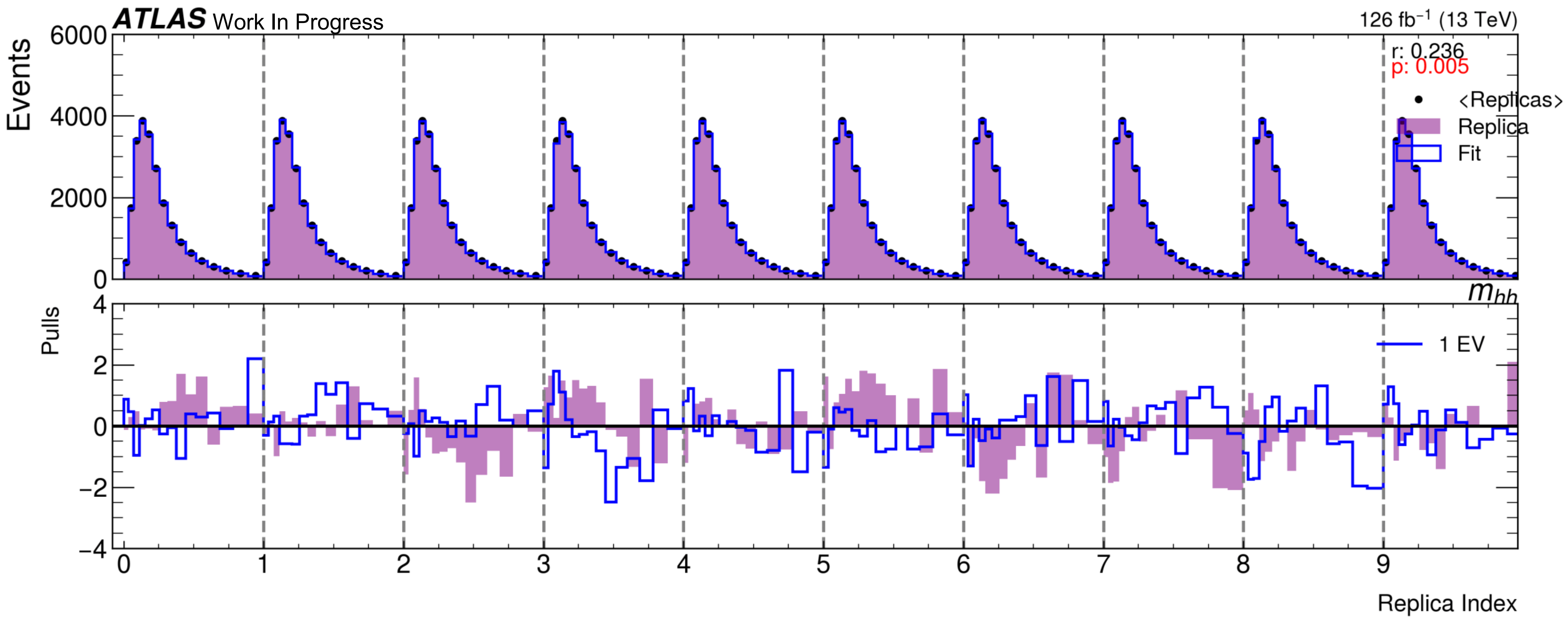
THE UNIVERSITY OF BRITISH COLUMBIA

Thank You!

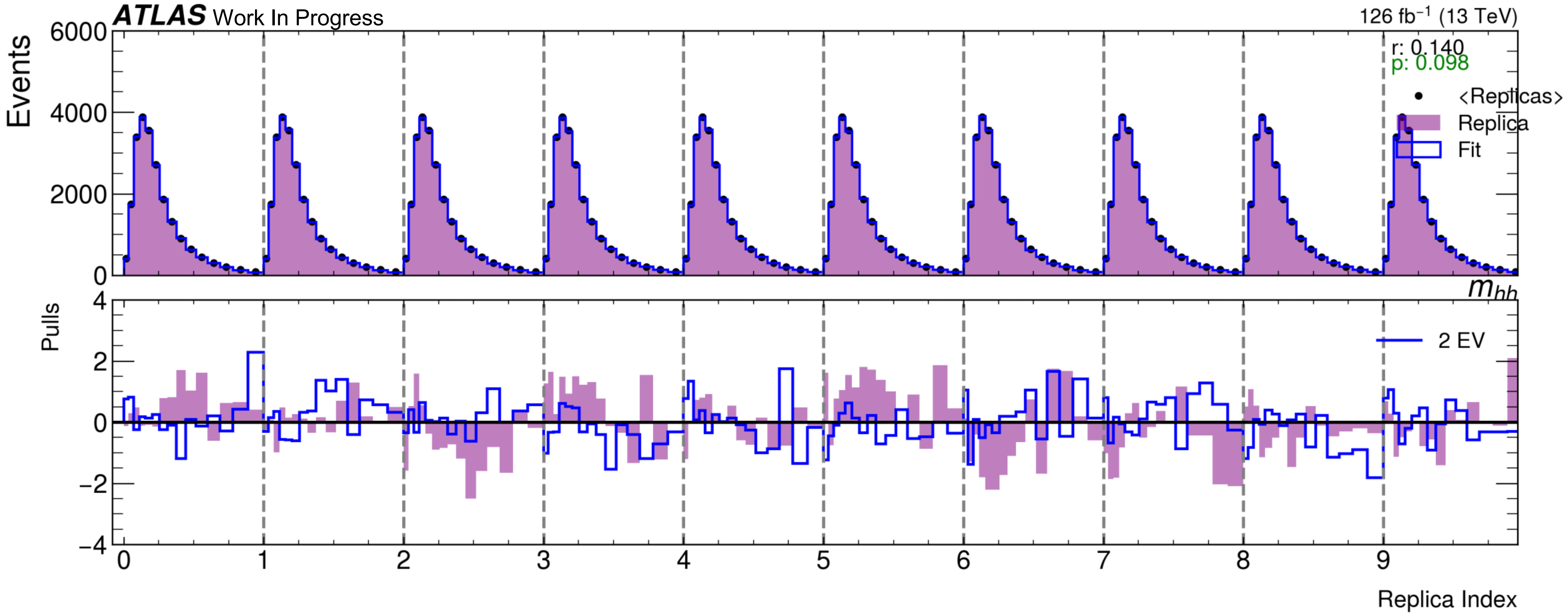
ATLAS EXPERIMENT

THE UNIVERSITY OF BRITISH COLUMBIA

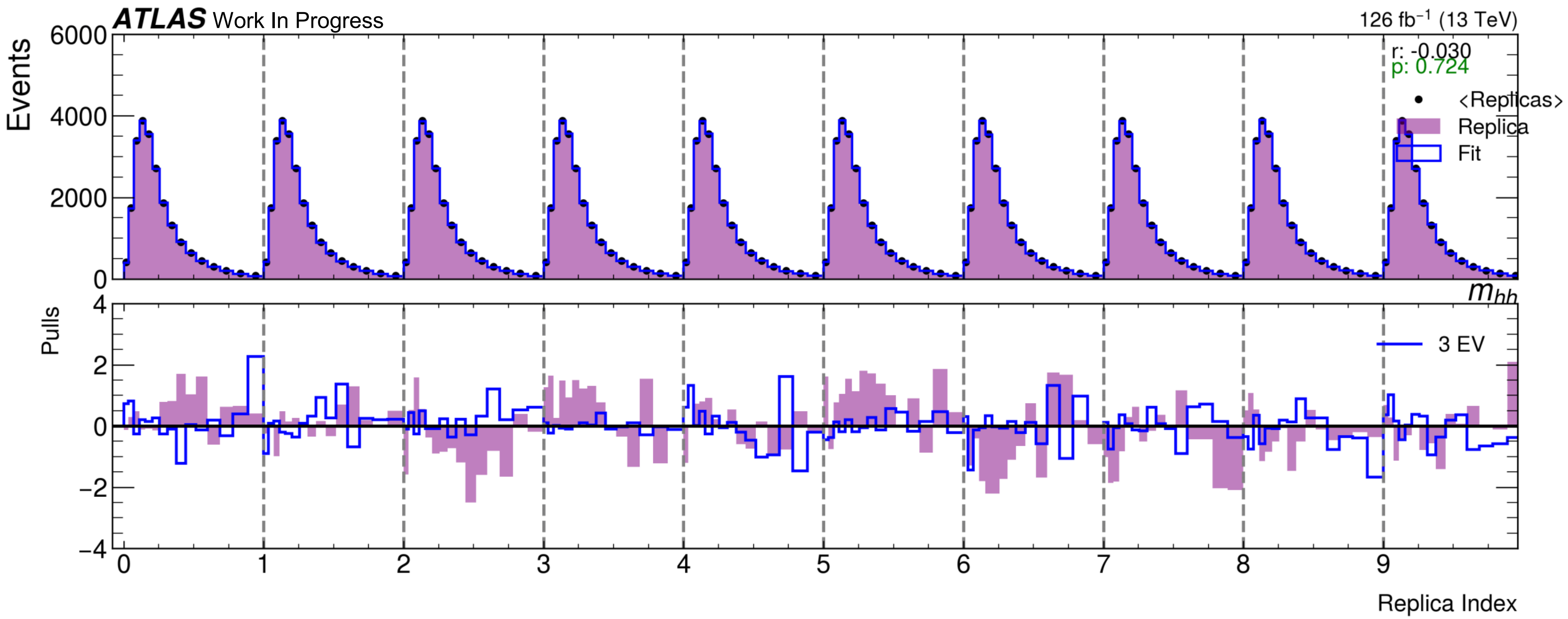
mhh Variance Steps



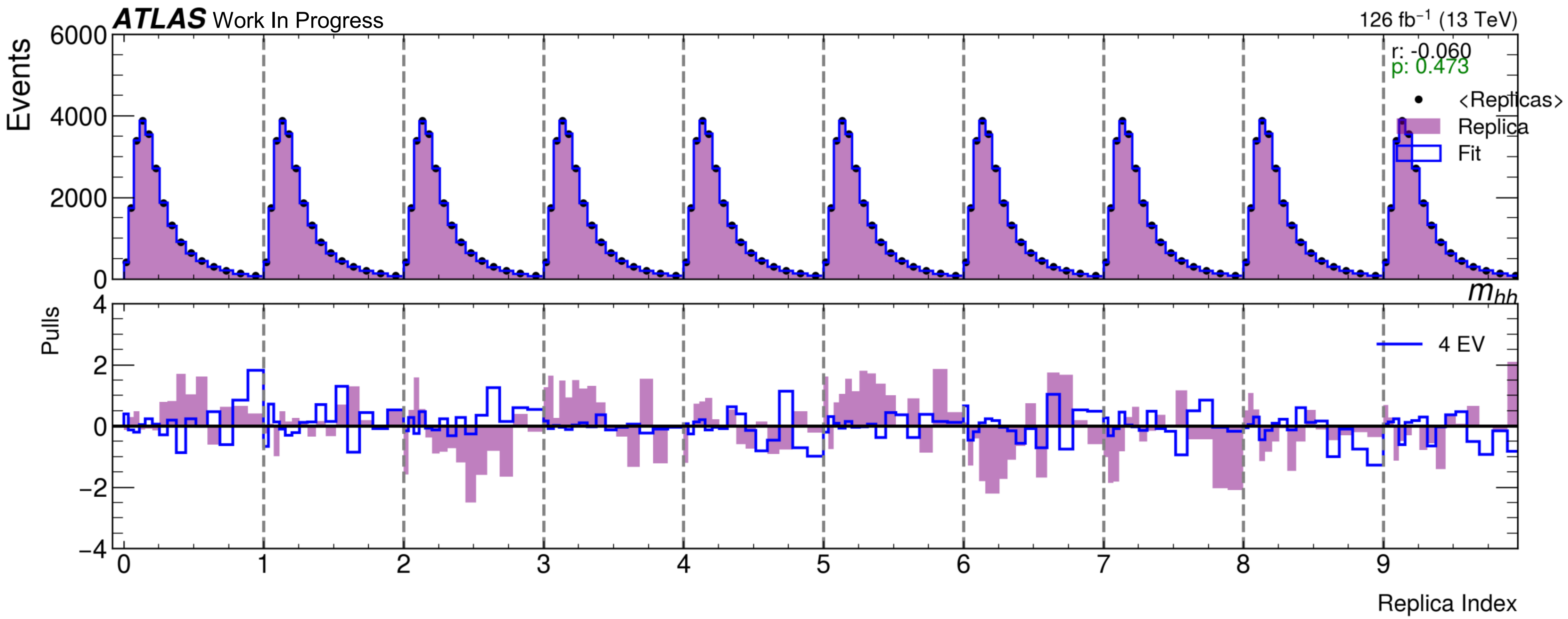
mhh Variance Steps



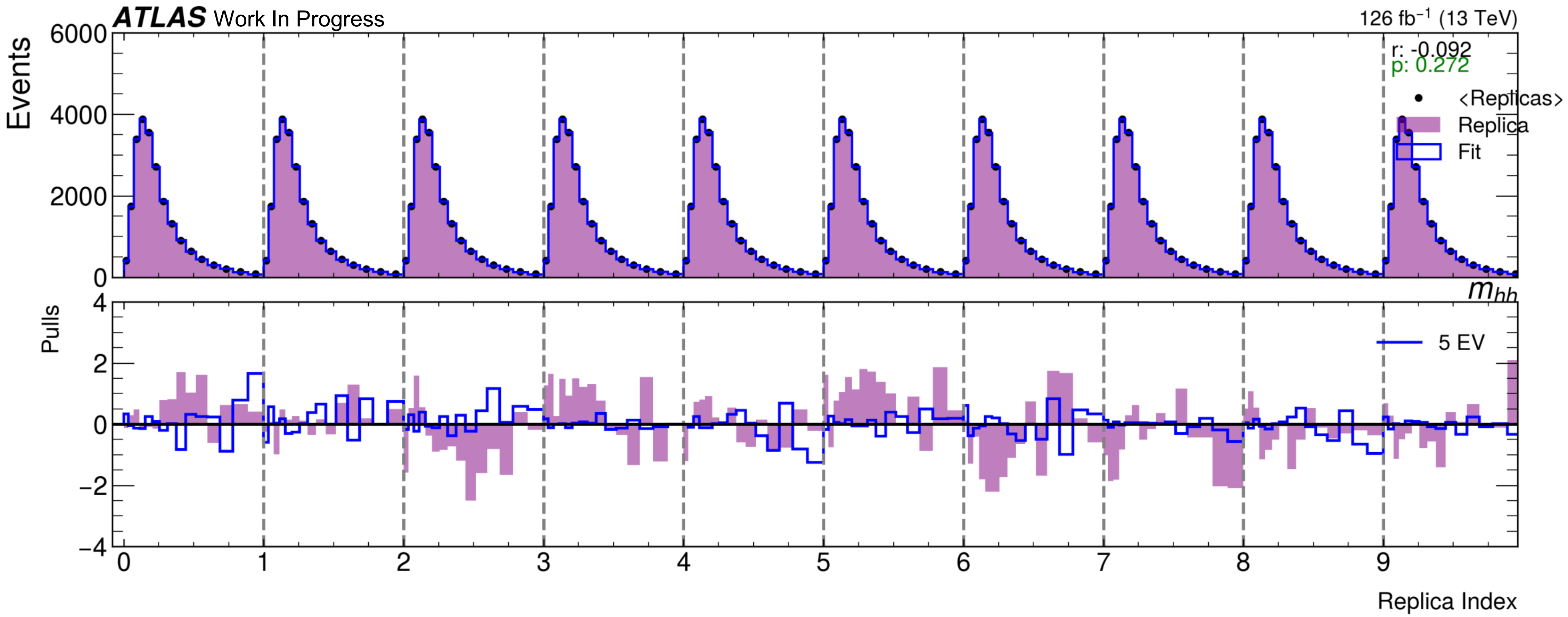
mhh Variance Steps



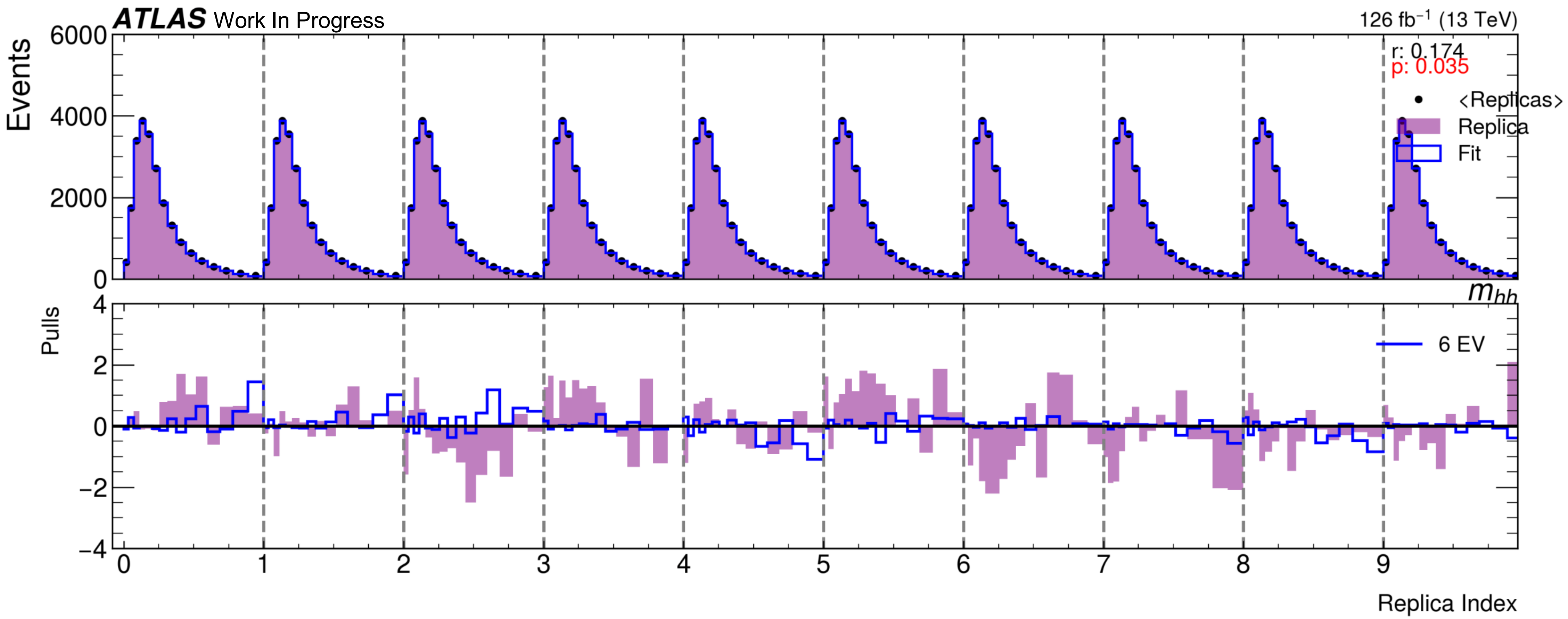
mhh Variance Steps



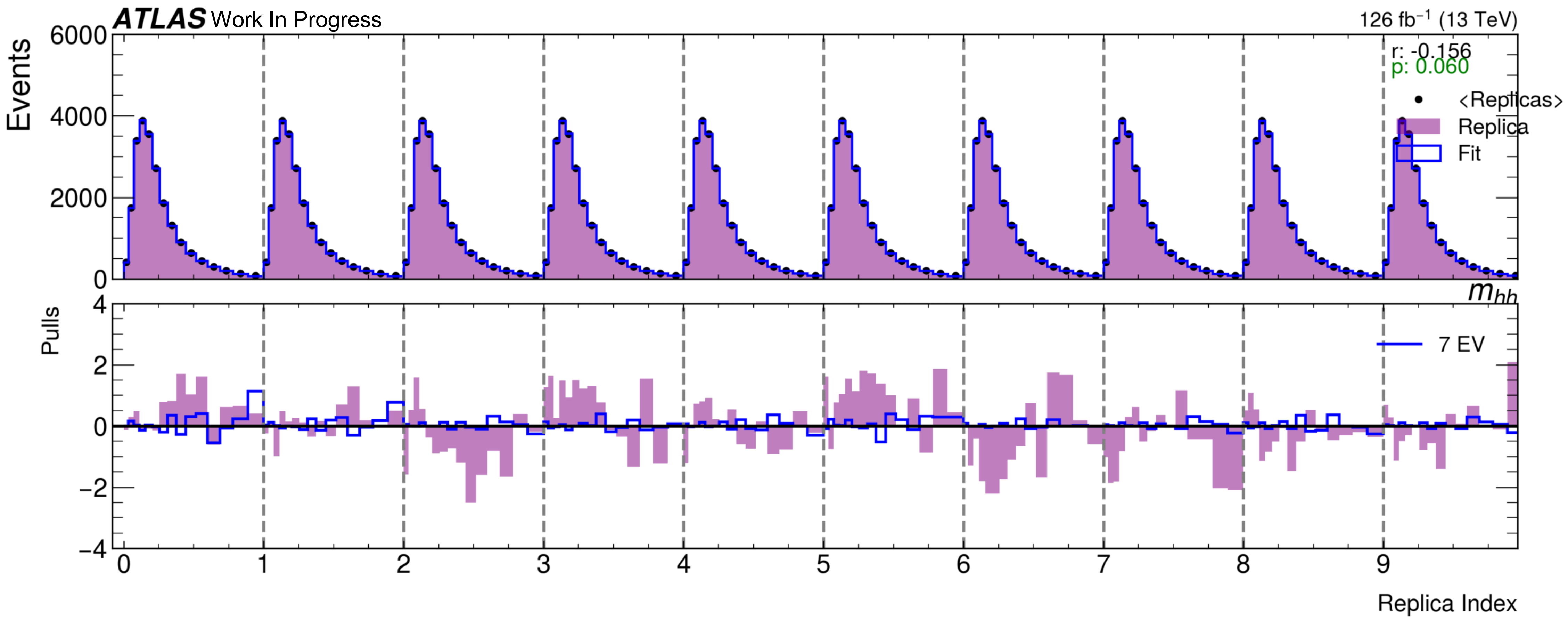
mhh Variance Steps



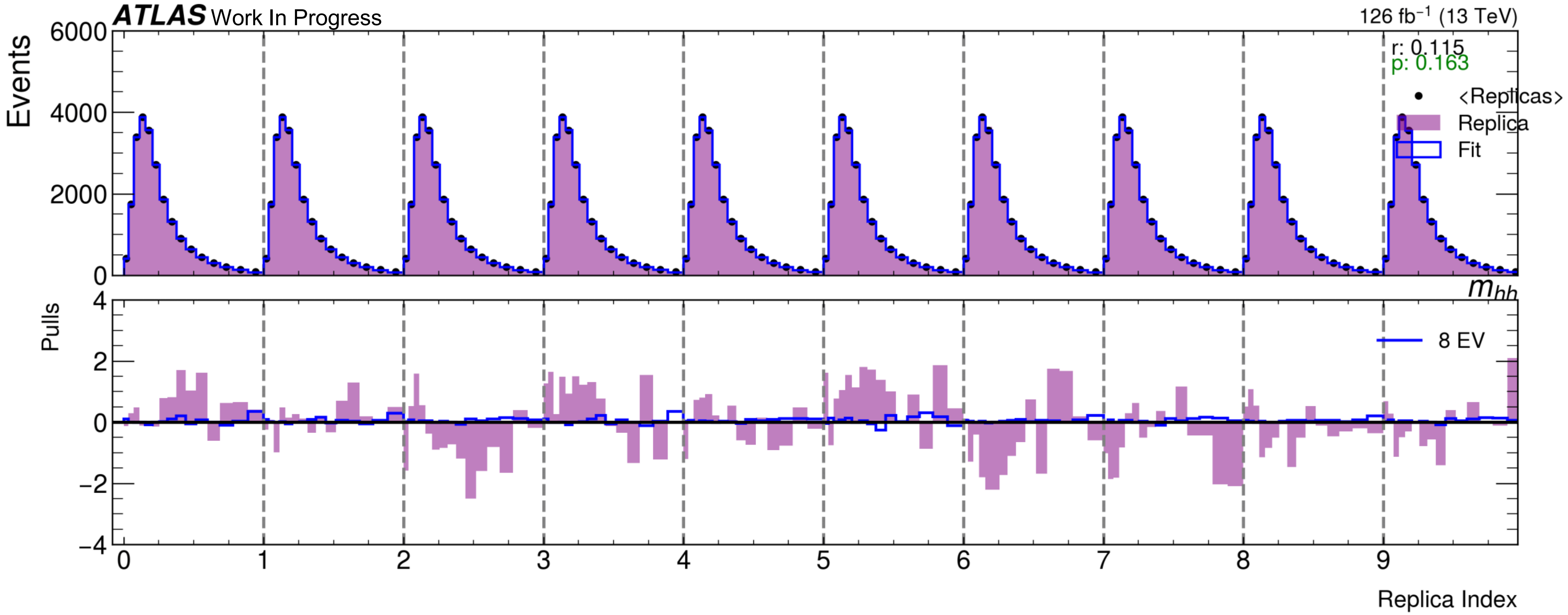
mhh Variance Steps



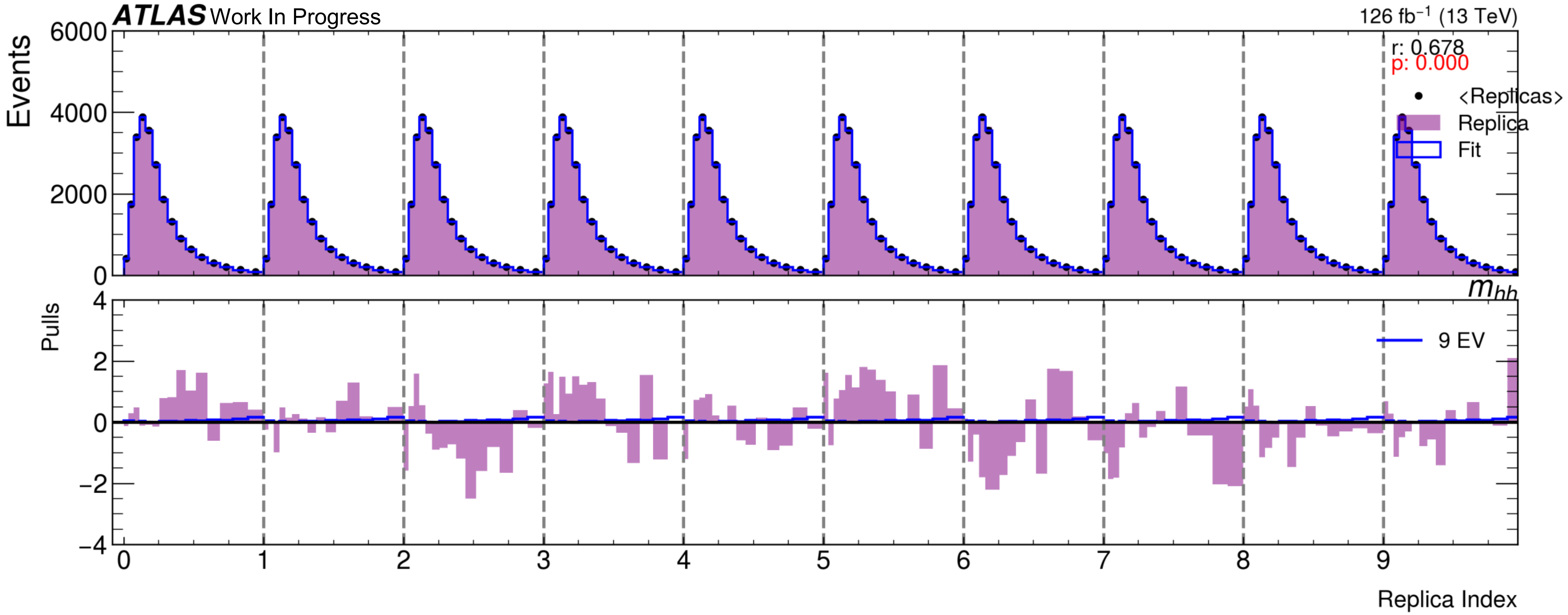
mhh Variance Steps



mhh Variance Steps



mhh Variance Steps



mhh Variance Steps