Higgs measurements using Simulation-Based Inference techniques

Tae Hyoun Park *on behalf of ATLAS and CMS Collaborations*

Standard Model at LHC 2025

ATLAS EXPERIMENT

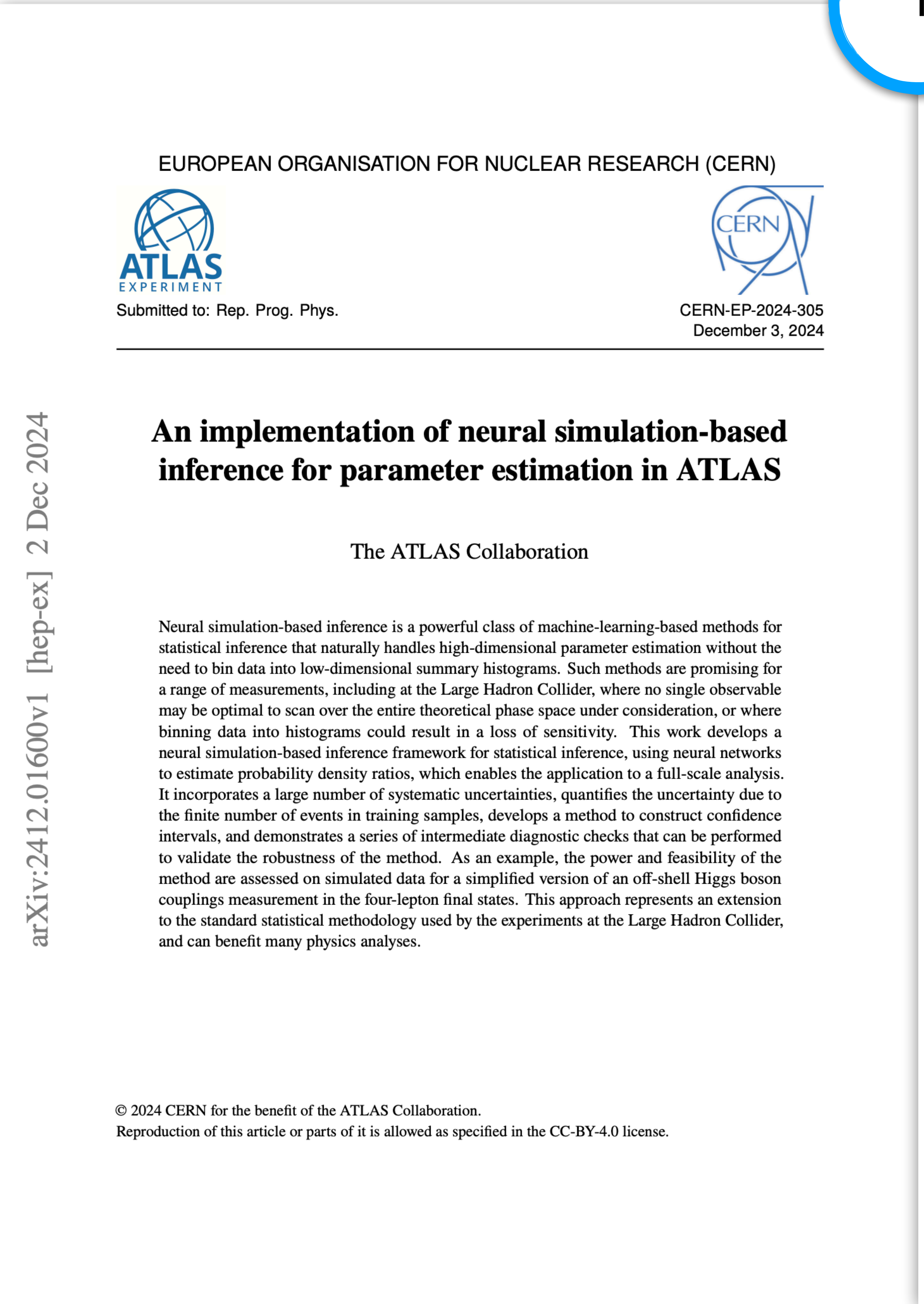
MAX-PLANCK-INSTITUT FÜR PHYSIK

CERN

CMS

Outline

1



EUROPEAN ORGANISATION FOR NUCLEAR RESEARCH (CERN)

Submitted to: Rep. Prog. Phys.

CERN-EP-2024-305
December 3, 2024

arXiv:2412.01600v1 [hep-ex] 2 Dec 2024

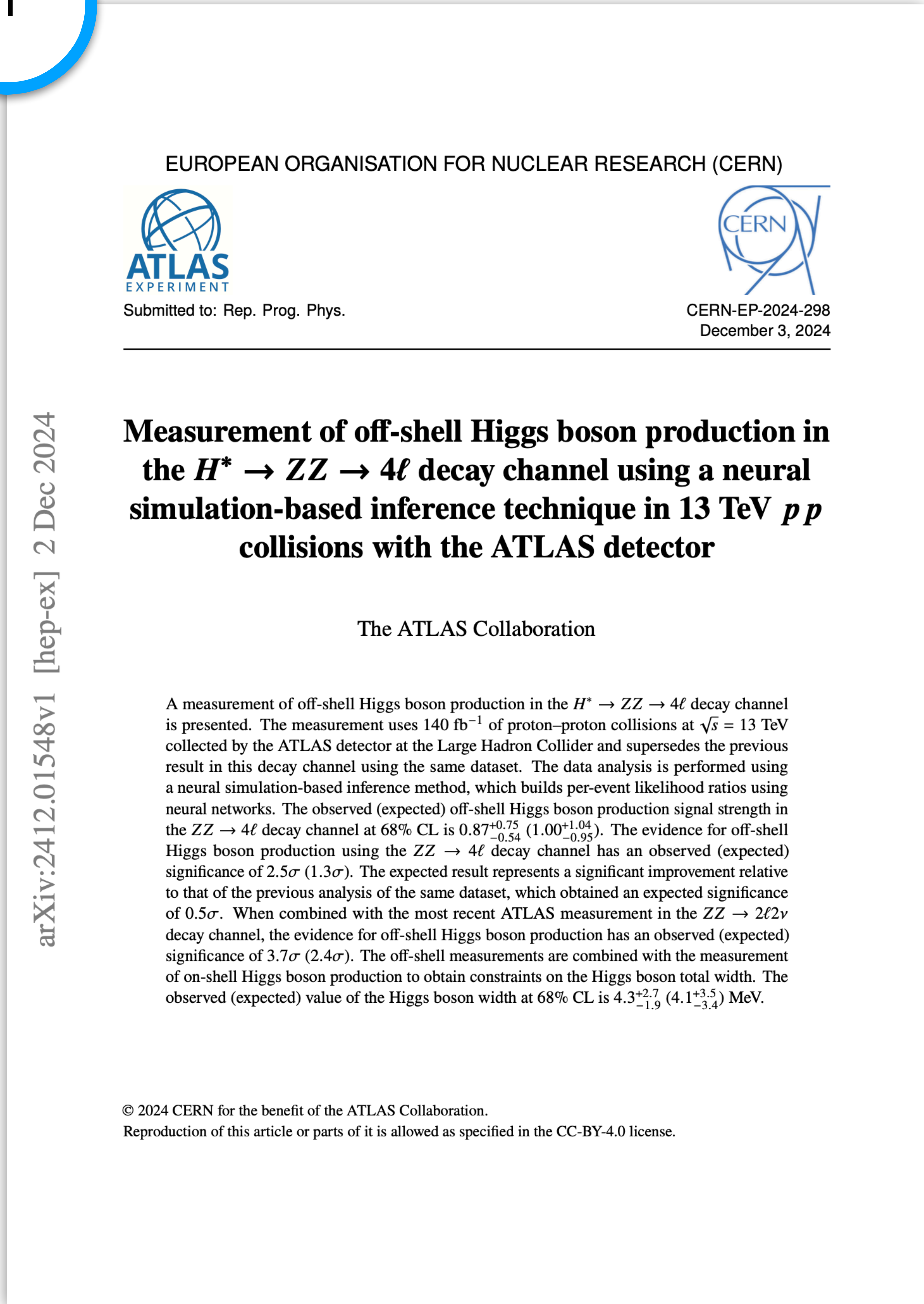
An implementation of neural simulation-based inference for parameter estimation in ATLAS

The ATLAS Collaboration

Neural simulation-based inference is a powerful class of machine-learning-based methods for statistical inference that naturally handles high-dimensional parameter estimation without the need to bin data into low-dimensional summary histograms. Such methods are promising for a range of measurements, including at the Large Hadron Collider, where no single observable may be optimal to scan over the entire theoretical phase space under consideration, or where binning data into histograms could result in a loss of sensitivity. This work develops a neural simulation-based inference framework for statistical inference, using neural networks to estimate probability density ratios, which enables the application to a full-scale analysis. It incorporates a large number of systematic uncertainties, quantifies the uncertainty due to the finite number of events in training samples, develops a method to construct confidence intervals, and demonstrates a series of intermediate diagnostic checks that can be performed to validate the robustness of the method. As an example, the power and feasibility of the method are assessed on simulated data for a simplified version of an off-shell Higgs boson couplings measurement in the four-lepton final states. This approach represents an extension to the standard statistical methodology used by the experiments at the Large Hadron Collider, and can benefit many physics analyses.

© 2024 CERN for the benefit of the ATLAS Collaboration.
Reproduction of this article or parts of it is allowed as specified in the CC-BY-4.0 license.

arXiv:2412.01600



EUROPEAN ORGANISATION FOR NUCLEAR RESEARCH (CERN)

Submitted to: Rep. Prog. Phys.

CERN-EP-2024-298
December 3, 2024

arXiv:2412.01548v1 [hep-ex] 2 Dec 2024

Measurement of off-shell Higgs boson production in the $H^* \to ZZ \to 4\ell$ decay channel using a neural simulation-based inference technique in 13 TeV pp collisions with the ATLAS detector

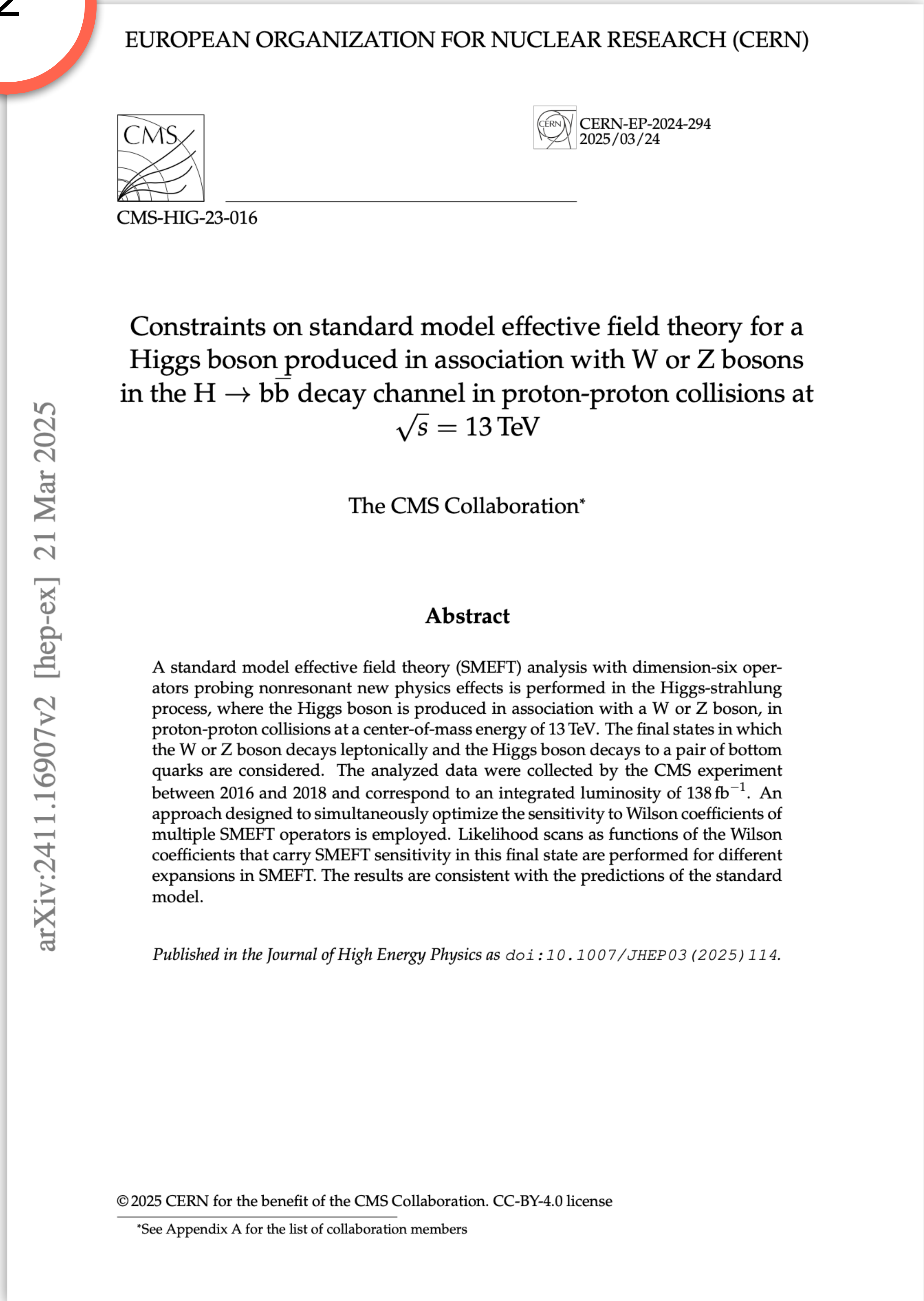
The ATLAS Collaboration

A measurement of off-shell Higgs boson production in the $H^* \to ZZ \to 4\ell$ decay channel is presented. The measurement uses 140 fb^{-1} of proton–proton collisions at $\sqrt{s}$ = 13 TeV collected by the ATLAS detector at the Large Hadron Collider and supersedes the previous result in this decay channel using the same dataset. The data analysis is performed using a neural simulation-based inference method, which builds per-event likelihood ratios using neural networks. The observed (expected) off-shell Higgs boson production signal strength in the $ZZ \to 4\ell$ decay channel at 68% CL is $0.87^{+0.75}_{-0.54}$ ($1.00^{+1.04}_{-0.95}$). The evidence for off-shell Higgs boson production using the $ZZ \to 4\ell$ decay channel has an observed (expected) significance of 2.5σ (1.3σ). The expected result represents a significant improvement relative to that of the previous analysis of the same dataset, which obtained an expected significance of 0.5σ. When combined with the most recent ATLAS measurement in the $ZZ \to 2\ell 2\nu$ decay channel, the evidence for off-shell Higgs boson production has an observed (expected) significance of 3.7σ (2.4σ). The off-shell measurements are combined with the measurement of on-shell Higgs boson production to obtain constraints on the Higgs boson total width. The observed (expected) value of the Higgs boson width at 68% CL is $4.3^{+2.7}_{-1.9}$ ($4.1^{+3.5}_{-3.4}$) MeV.

© 2024 CERN for the benefit of the ATLAS Collaboration.
Reproduction of this article or parts of it is allowed as specified in the CC-BY-4.0 license.

arXiv:2412.01548

2



EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH (CERN)

CMS-HIG-23-016

CERN-EP-2024-294
2025/03/24

arXiv:2411.16907v2 [hep-ex] 21 Mar 2025

Constraints on standard model effective field theory for a Higgs boson produced in association with W or Z bosons in the $H \to b\bar{b}$ decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV

The CMS Collaboration*

Abstract

A standard model effective field theory (SMEFT) analysis with dimension-six operators probing nonresonant new physics effects is performed in the Higgs-strahlung process, where the Higgs boson is produced in association with a W or Z boson, in proton-proton collisions at a center-of-mass energy of 13 TeV. The final states in which the W or Z boson decays leptonically and the Higgs boson decays to a pair of bottom quarks are considered. The analyzed data were collected by the CMS experiment between 2016 and 2018 and correspond to an integrated luminosity of 138 fb^{-1}. An approach designed to simultaneously optimize the sensitivity to Wilson coefficients of multiple SMEFT operators is employed. Likelihood scans as functions of the Wilson coefficients that carry SMEFT sensitivity in this final state are performed for different expansions in SMEFT. The results are consistent with the predictions of the standard model.

Published in the Journal of High Energy Physics as doi:10.1007/JHEP03(2025)114.

©2025 CERN for the benefit of the CMS Collaboration. CC-BY-4.0 license

*See Appendix A for the list of collaboration members

JHEP03(2025)114

Inference @ LHC

Maximum Likelihood: Estimate parameters behind hypothesized probability density, given observed data.

$$\mathcal{L}(\theta|\mathcal{D}) = \mathrm{Pois}(n; \nu(\theta)) \prod_{i \leq n} p\left(x_i \middle| \theta\right)$$

Neyman-Pearson Lemma: Likelihood ratio is the most powerful test statistic.

$$\Lambda(x) = \frac{\mathcal{L}(\theta_1|x)}{\mathcal{L}(\theta_2|x)}$$

Task: Compute the expected probability (ratio).

$$\frac{p(x|\theta_1)}{p(x|\theta_2)} = ?$$

Inference @ LHC: Challenges

Task: Compute hypothesized probability (ratio).

$$\frac{p(x|\theta_1)}{p(x|\theta_2)} = ?$$

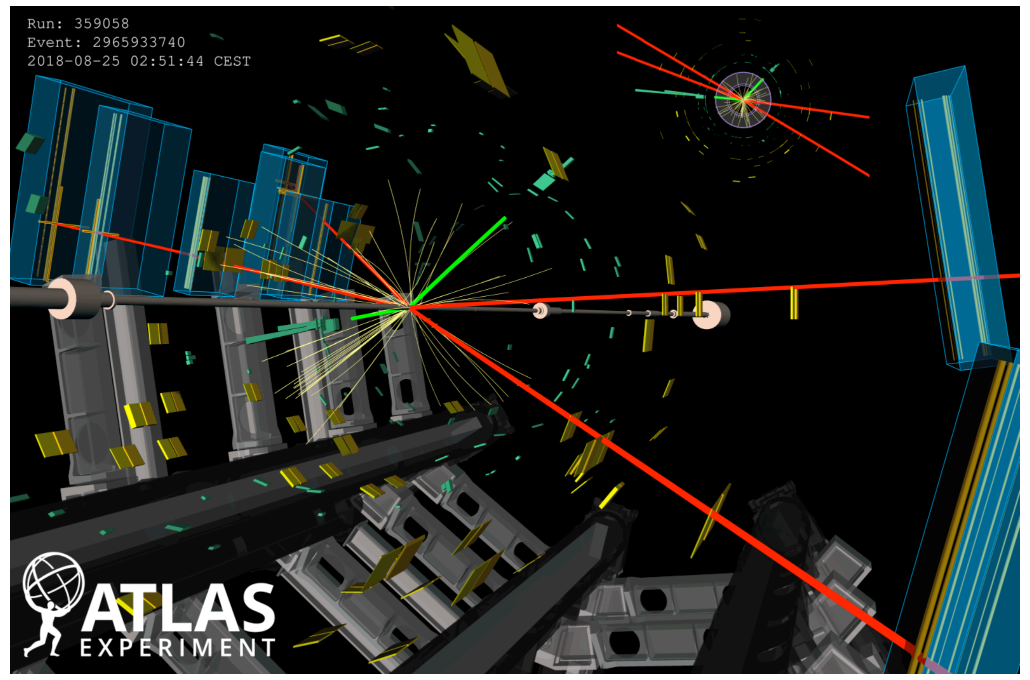
What challenges exist for LHC data?

Example: Consider a “traditional” Higgs signal strength measurement.

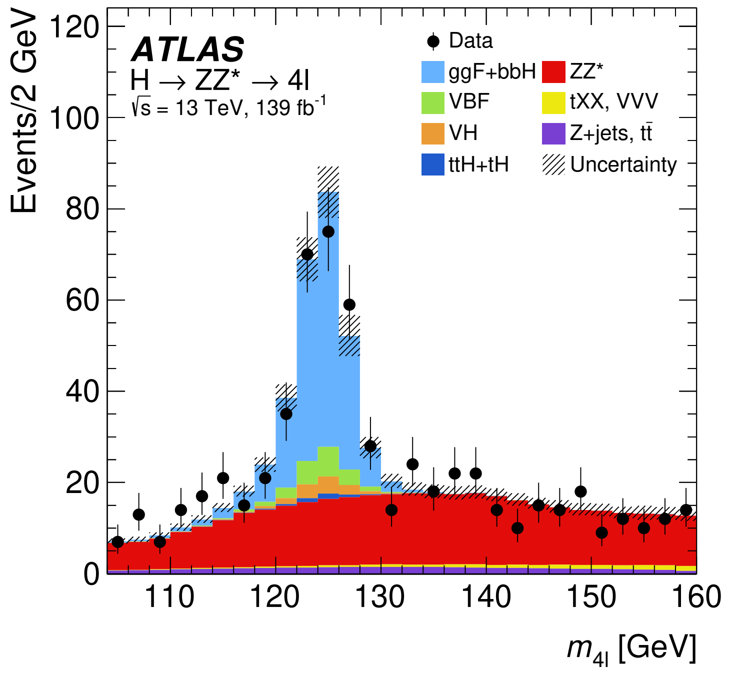
Inference @ LHC: Challenge 1

1. Observable data is high-dimensional.

- Traditional inference: Project data to a low-dimensional summary statistic.



$x \mapsto u(x)$



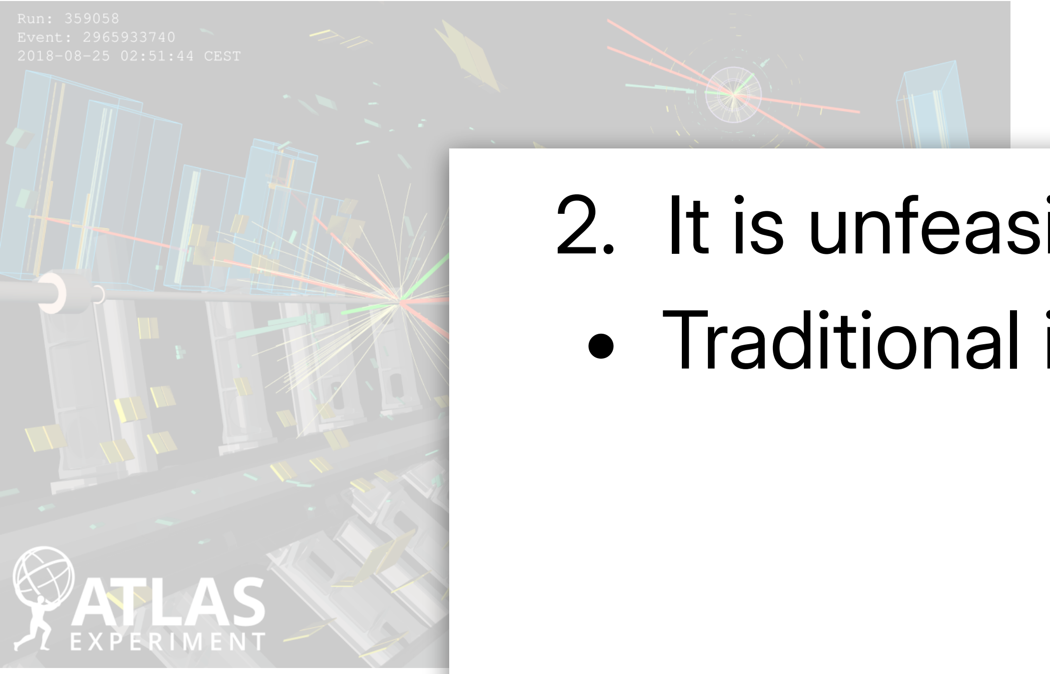
ability (ratio).

LHC data?

nal strength measurement.

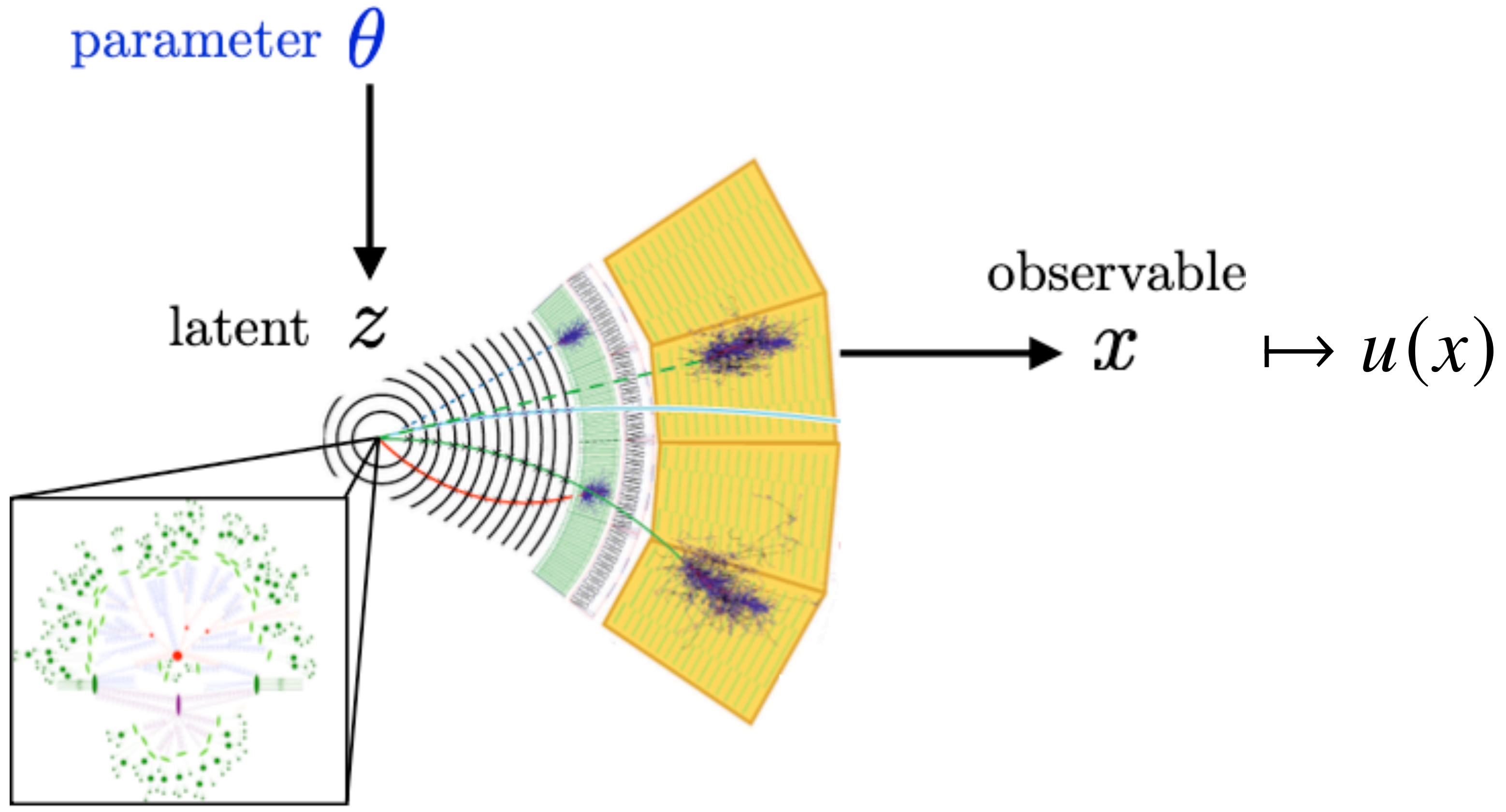
Inference @ LHC: Challenge 2

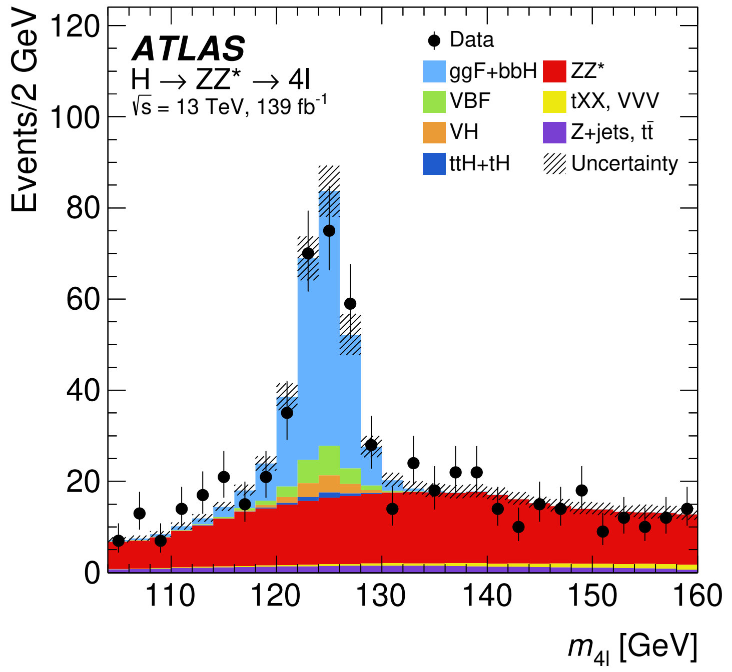
1. Observable data is high-dimensional.
- Traditional inference: Project data to a low-dimensional summary statistic.



obability (ratio).

2. It is unfeasible to evaluate the probability density numerically.
- Traditional inference: Use sampled distributions from simulations as an approximation.





$$p(x|\theta) = \overbrace{\int \mathrm{d}z_{\text{detector}} \int \mathrm{d}z_{\text{shower}} \int \mathrm{d}z_{\text{parton}}\, p(x|z_{\text{detector}})\, p(z_{\text{detector}}|z_{\text{shower}})\, p(z_{\text{detector}}|z_{\text{parton}})}^{\text{Simulation}}\, p(z_{\text{parton}}|\theta)$$

Inference @ LHC: Challenge 3

1. Observable data is high-dimensional.
   - Traditional inference: Project data to a low-dimensional summary statistic.

obability (ratio).

2. It is unfeasible to evaluate the probability density numerically.
   - Traditional inference: Use sampled distributions from simulations as an approximation.

$$p(x|\theta) = \int \mathrm{d}z_{\text{detector}} \int \mathrm{d}z \ldots$$

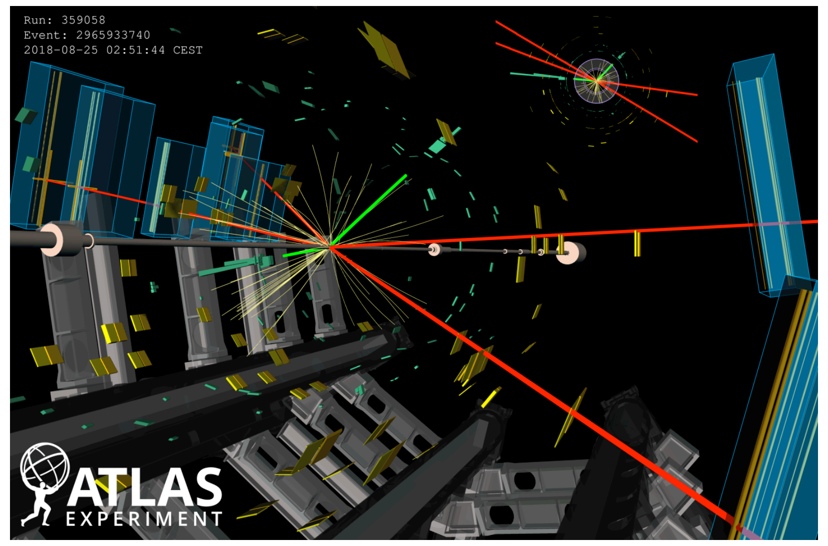
3. Quantum effects can introduce complicated dependence on the parameters of interest.
   - Traditional inference: Keep it simple.

$$\left|\sqrt{\mu}\mathcal{M}_{\mathrm{S}} + \mathcal{M}_{\mathrm{B}}\right|^2 = \mu\left|\mathcal{M}_{\mathrm{S}}\right|^2 + \sqrt{\mu}\,\mathrm{Re}(\mathcal{M}_{\mathrm{S}}^\dagger\mathcal{M}_{\mathrm{B}}) + \left|\mathcal{M}_{\mathrm{B}}\right|^2$$

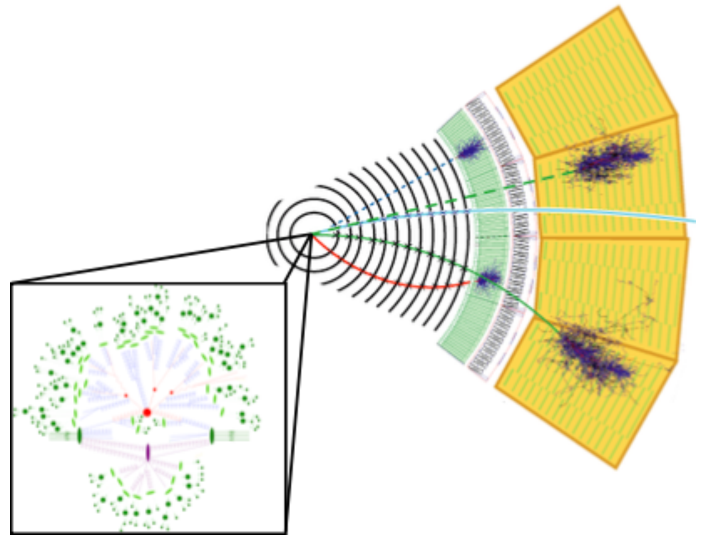
$$\nu(\mu) = \mu\nu_{\mathrm{S}} + \sqrt{\mu}\nu_{\mathrm{I}} + \nu_{\mathrm{B}}$$

   - Can be shown: *one* summary histogram from 1+2 is optimal across *all* μ.

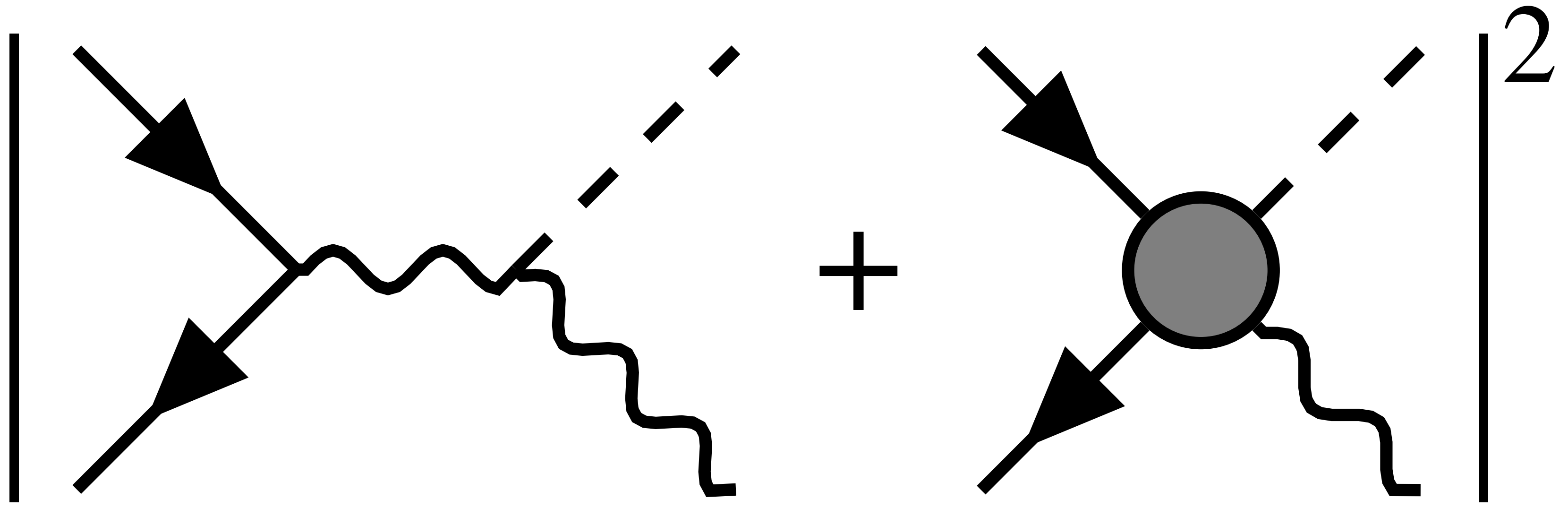
Simulation-based Inference @ LHC?



1. High-dimensionality of data



2. Intractability of the likelihood



3. Quantum interference in probability

$$\{x\} \sim p(x|\theta) = \int \mathrm{d}z\, p_{\mathrm{sim}}(x|z)|\mathcal{M}(z|\theta)|^2$$

What if any/all of these approximations do not hold?

Can simulation-based inference overcome these challenges?

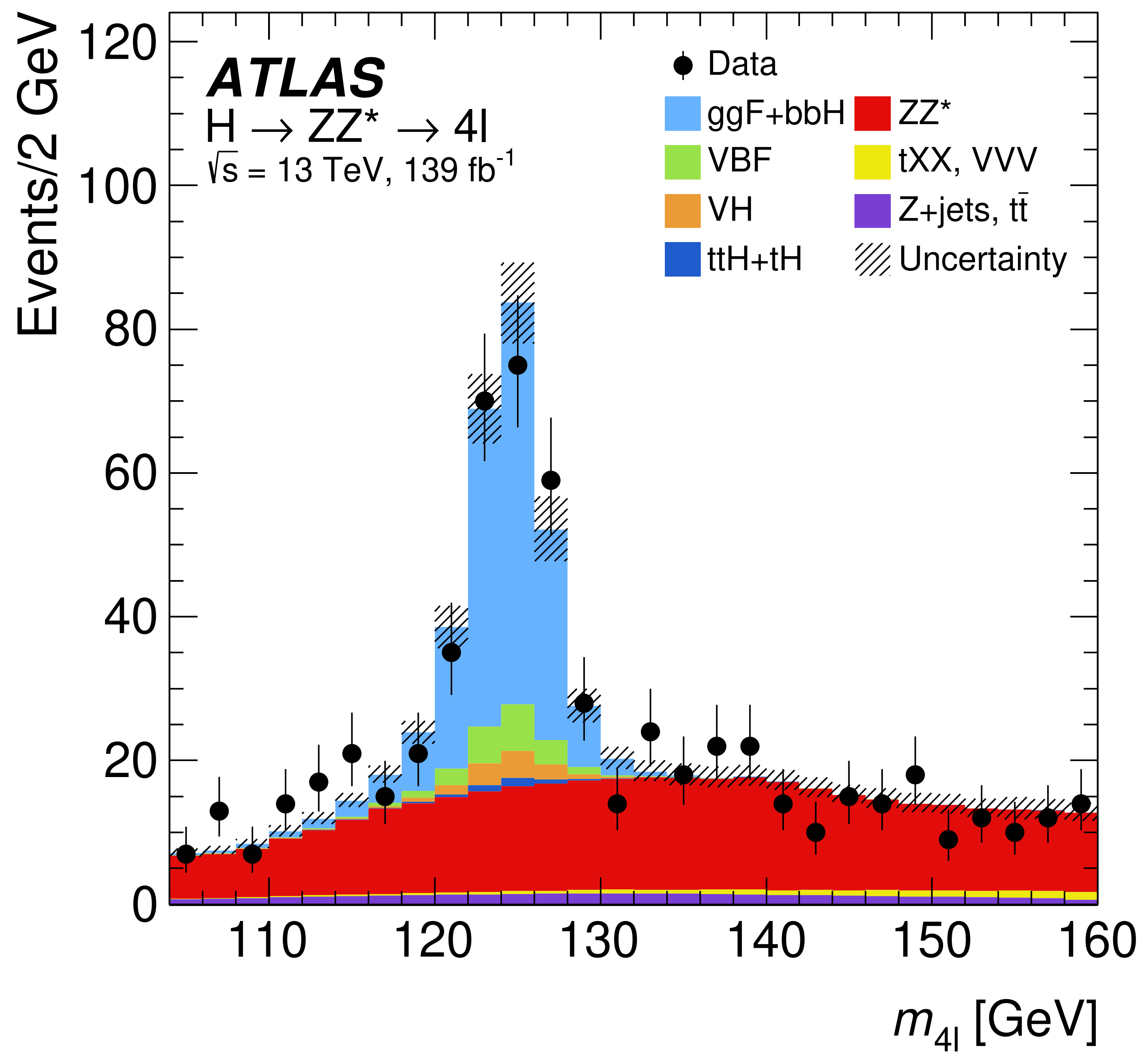
Spoiler: Yes; in fact, these are the ideal conditions under which it excels!

ATLAS $H \to 4\ell$ off-shell measurement

Physics overview

- In/dependence of Higgs production cross-section on Higgs width in on-/off-shell regimes:

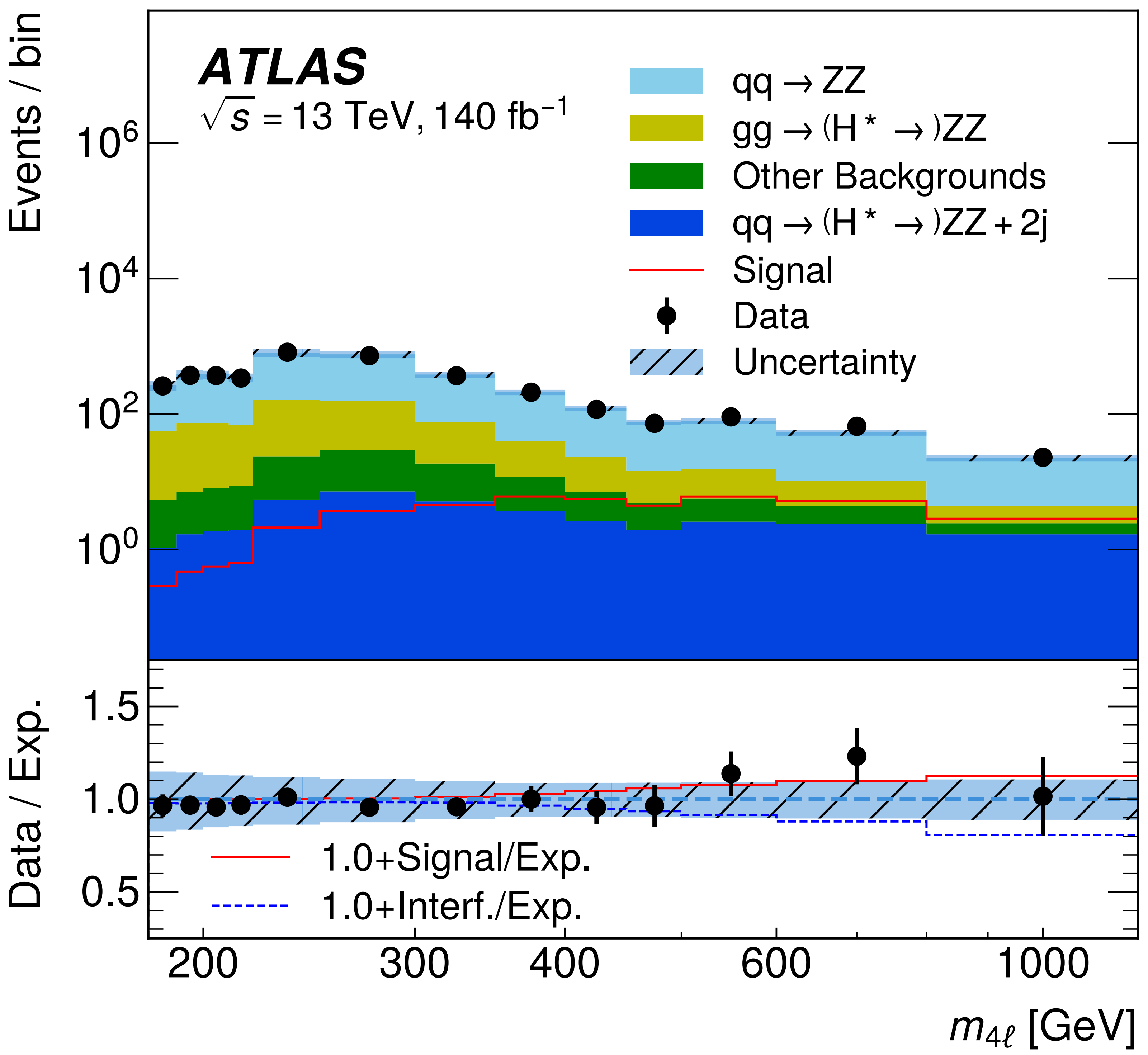
$$\sigma^{H\to VV}_{\text{on-shell}} \propto \frac{g^2_{\text{prod}}(m_H)g^2_{\text{decay}}(m_H)}{\Gamma_H} \longleftarrow \frac{d\sigma^{H\to VV}}{dm^2_{VV}} \propto \frac{g^2_{\text{prod}}(\hat{s})g^2_{\text{decay}}(\hat{s})}{(m^2_{VV}-m^2_H)^2 - m^2_H\Gamma^2_H} \longrightarrow \frac{d\sigma^{H\to VV}_{\text{off-shell}}}{dm^2_{VV}} \propto g^2_{\text{prod}}(\hat{s})g^2_{\text{decay}}(\hat{s})$$



- Negligible interference between signal & background:

$$\nu(\mu) = \mu\nu_S + \nu_B$$

- Comparison of on-/off-shell rates under SM-like assumptions ⇒ indirect measurement of Higgs width.



- Significant interference between signal & background:

$$\nu(\mu) = \mu\nu_S + \sqrt{\mu}\nu_I + \nu_B$$

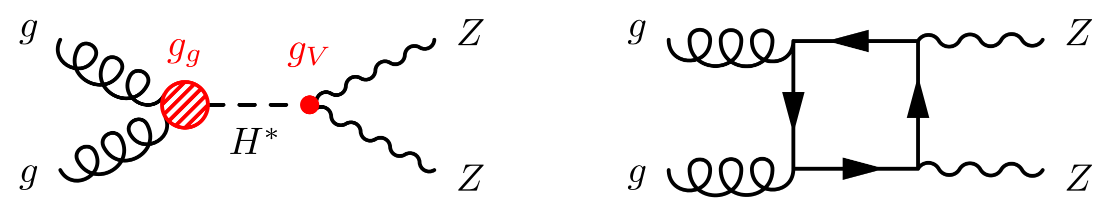
Off-shell Higgs probability density model

$$p(x|\mu^{\text{ggF}}_{\text{off-shell}},\mu^{\text{EW}}_{\text{off-shell}}) = \frac{1}{\nu(\mu^{\text{ggF}}_{\text{off-shell}},\mu^{\text{EW}}_{\text{off-shell}})} \times \Big[\mu^{\text{ggF}}_{\text{off-shell}}\nu^{\text{ggF}}_{\text{S}}p^{\text{ggF}}_{\text{S}}(x) + \sqrt{\mu^{\text{ggF}}_{\text{off-shell}}}\nu^{\text{ggF}}_{\text{I}}p^{\text{ggF}}_{\text{I}}(x) + \nu^{\text{ggF}}_{\text{B}}p^{\text{ggF}}_{\text{B}}(x)$$

$$+ \mu^{\text{EW}}_{\text{off-shell}}\nu^{\text{EW}}_{\text{S}}p^{\text{EW}}_{\text{S}}(x) + \sqrt{\mu^{\text{EW}}_{\text{off-shell}}}\nu^{\text{EW}}_{\text{I}}p^{\text{EW}}_{\text{I}}(x) + \nu^{\text{ggF}}_{\text{B}}p^{\text{ggF}}_{\text{B}}(x)$$

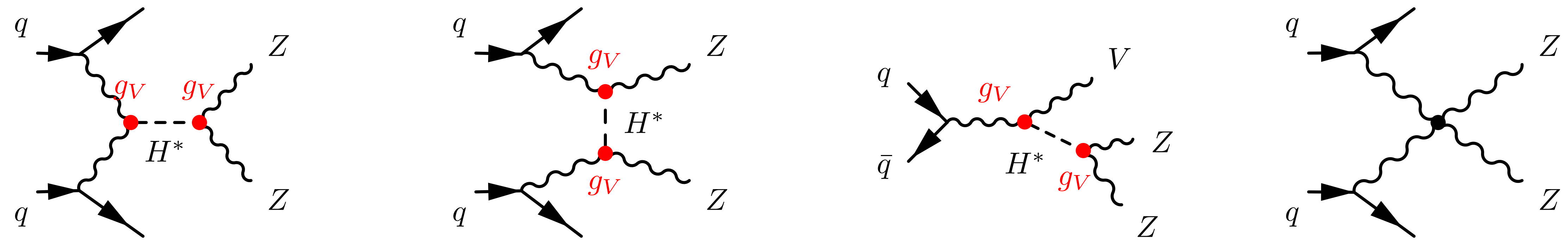
$$+ \nu_{\text{NI}}p_{\text{NI}}(x)\Big]$$

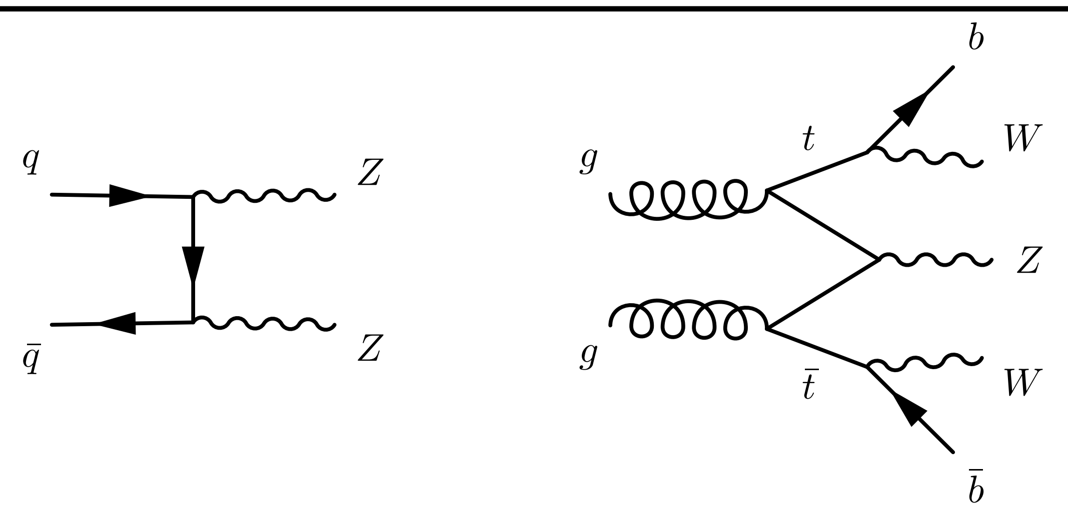
- 3 distinct predictions generated to determine signal+background+interference (SBI) basis:



$\mu_{\text{ggF}} = \kappa_g^2\kappa_V^2$

$\mu_{\text{EW}} = \kappa_V^4$

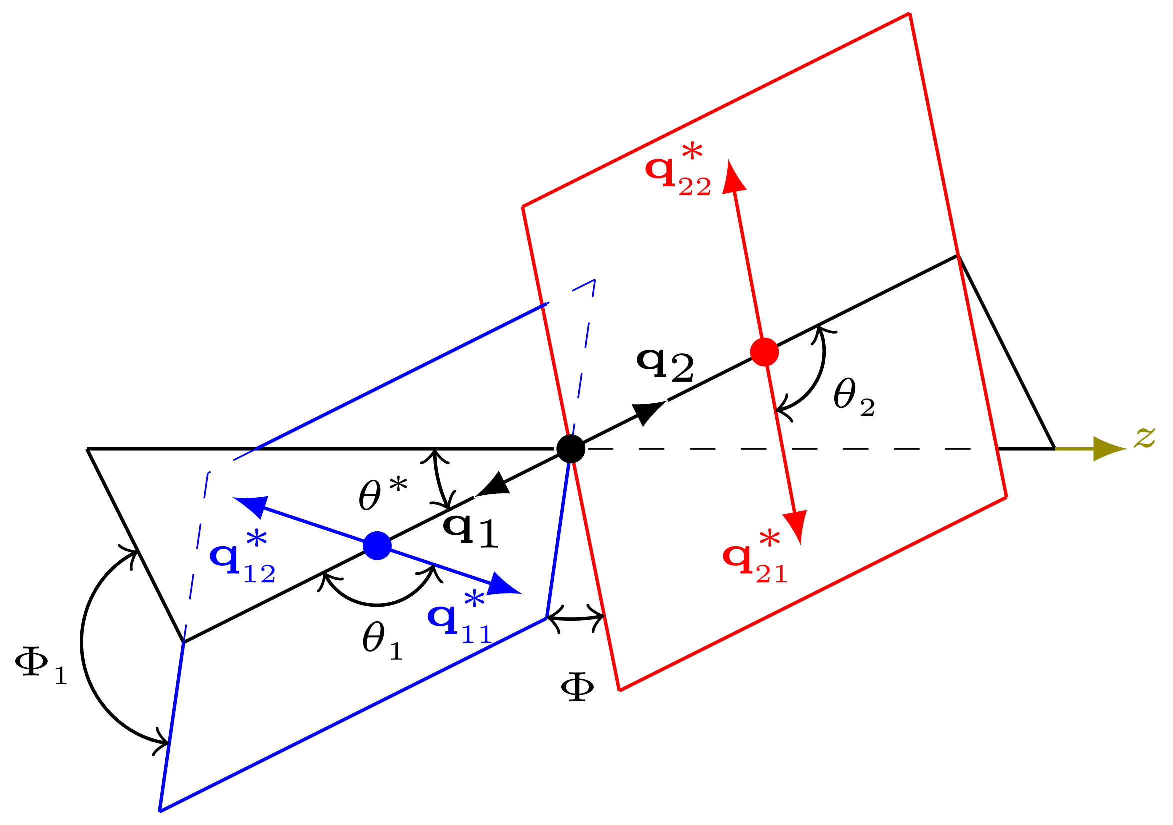




Process	Multipliers	Samples
ggF S	$\mu^{\text{ggF}}_{\text{off-shell}} - \sqrt{\mu^{\text{ggF}}_{\text{off-shell}}}$	$gg \to H^* \to ZZ \to 4\ell$
ggF SBI	$\sqrt{\mu^{\text{ggF}}_{\text{off-shell}}}$	$gg \to (H^* \to)ZZ \to 4\ell$ $(\kappa_V^2 = 1)$
ggF B	$1 - \sqrt{\mu^{\text{ggF}}_{\text{off-shell}}}$	$gg \to ZZ \to 4\ell$ $(\kappa_V^2 = 0)$
EW B	$\frac{(1-\sqrt{10})\mu^{\text{EW}}_{\text{off-shell}} + 9\sqrt{\mu^{\text{EW}}_{\text{off-shell}}} - 10 + \sqrt{10}}{-10+\sqrt{10}}$	EW $qq \to ZZ + 2j \to 4\ell + 2j$ $(\kappa_V^4 = 0)$ $ZZZ \to 4\ell + 2j$
EW SBI$_1$	$\frac{\sqrt{10}\mu^{\text{EW}}_{\text{off-shell}} - 10\sqrt{\mu^{\text{EW}}_{\text{off-shell}}}}{-10+\sqrt{10}}$	EW $qq \to (H^* \to)ZZ + 2j \to 4\ell + 2j$ $(\kappa_V^4 = 1)$ $ZZZ \to 4\ell + 2j$
EW SBI$_{10}$	$\frac{-\mu^{\text{EW}}_{\text{off-shell}} + \sqrt{\mu^{\text{EW}}_{\text{off-shell}}}}{-10+\sqrt{10}}$	EW $qq \to (H^* \to)ZZ + 2j \to 4\ell + 2j$ $(\kappa_V^4 = 10)$ $ZZZ \to 4\ell + 2j$
$q\bar{q}ZZ$ $n_{\text{jets}} = 0$	$\theta^{0j}_{q\bar{q}ZZ}$	$q\bar{q} \to ZZ \to 4\ell$
$q\bar{q}ZZ$ $n_{\text{jets}} = 1$	$\theta^{0j}_{q\bar{q}ZZ}\theta^{1j}_{q\bar{q}ZZ}$	$q\bar{q} \to ZZ \to 4\ell$
$q\bar{q}ZZ$ $n_{\text{jets}} \geq 2$	$\theta^{0j}_{q\bar{q}ZZ}\theta^{1j}_{q\bar{q}ZZ}\theta^{2j}_{q\bar{q}ZZ}$	$q\bar{q} \to ZZ \to 4\ell$
VVV	–	$WWZ \to 4\ell$ $WZZ \to 4\ell$ $t\bar{t}Z \to 4\ell$

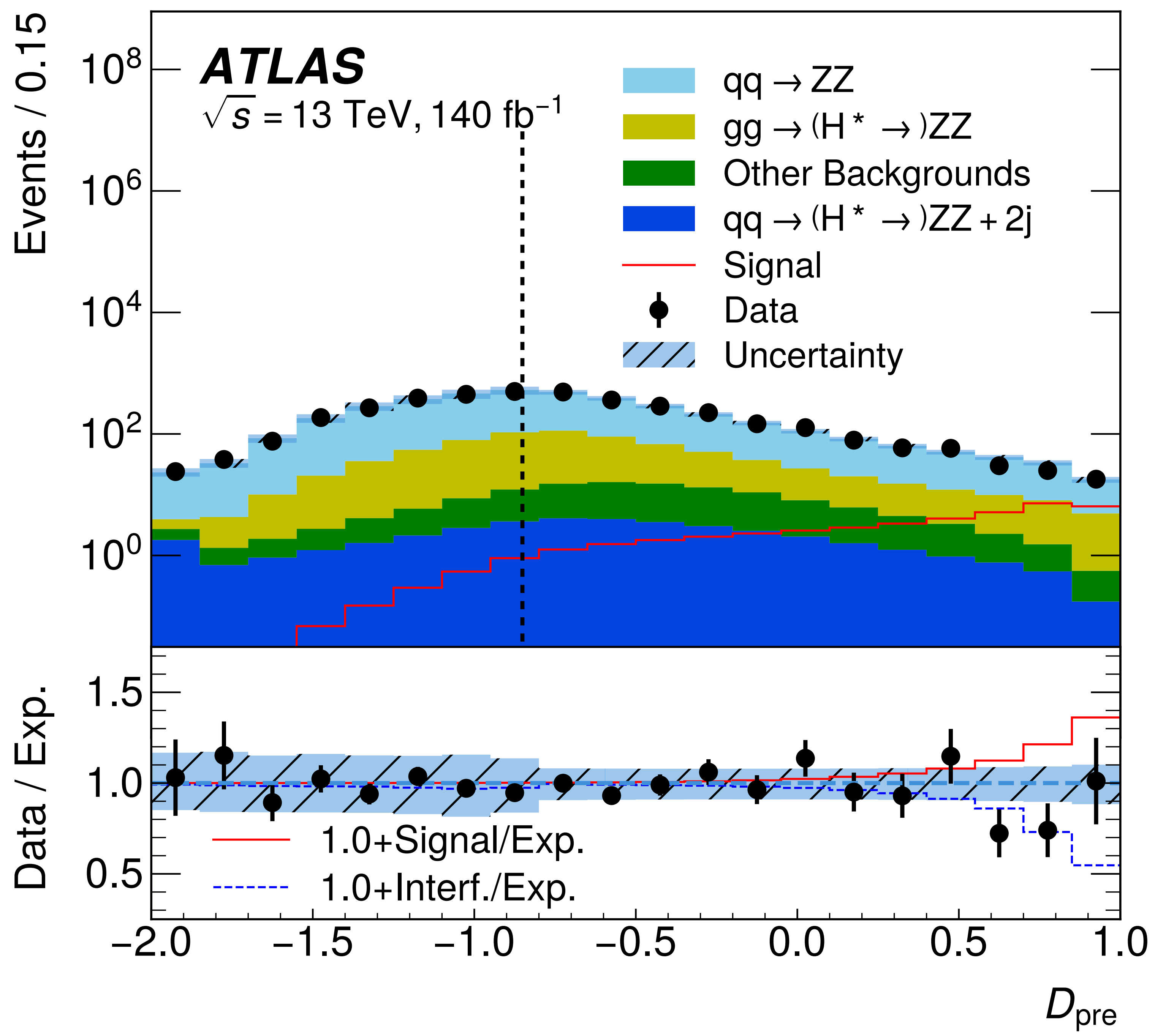
Event selection

- 14 observables to describe kinematics of Higgs(+jets) system.



Variable	Definition
$m_{4\ell}$	quadruplet mass
m_{Z_1}	Z_1 mass
m_{Z_2}	Z_2 mass
$\cos\theta^*$	cosine of the Higgs boson decay angle $[\mathbf{q}_1 \cdot \mathbf{n}_z/\|\mathbf{q}_1\|]$
$\cos\theta_1$	cosine of the Z_1 decay angle $[-(\mathbf{q}_2)\cdot\mathbf{q}_{11}/(\|\mathbf{q}_2\|\cdot\|\mathbf{q}_{11}\|)]$
$\cos\theta_2$	cosine of the Z_2 decay angle $[-(\mathbf{q}_1)\cdot\mathbf{q}_{21}/(\|\mathbf{q}_1\|\cdot\|\mathbf{q}_{21}\|)]$
Φ_1	Z_1 decay plane angle $[\cos^{-1}(\mathbf{n}_1\cdot\mathbf{n}_{sc})\,(\mathbf{q}_1\cdot(\mathbf{n}_1\times\mathbf{n}_{sc})/(\|\mathbf{q}_1\|\cdot\|\mathbf{n}_1\times\mathbf{n}_{sc}\|)]$
Φ	angle between Z_1, Z_2 decay planes $[\cos^{-1}(\mathbf{n}_1\cdot\mathbf{n}_2)\,(\mathbf{q}_1\cdot(\mathbf{n}_1\times\mathbf{n}_2)/(\|\mathbf{q}_1\|\cdot\|\mathbf{n}_1\times\mathbf{n}_2\|)]$
$p_T^{4\ell}$	quadruplet transverse momentum
$y^{4\ell}$	quadruplet rapidity
n_{jets}	number of jets in the event
m_{jj}	leading dijet system mass
$\Delta\eta_{jj}$	leading dijet system pseudorapidity
$\Delta\phi_{jj}$	leading dijet system azimuthal angle difference

$$D_{\text{pre}}(x) = \log \frac{s_{\text{pre}}^{\text{ggF S}}(x) + s_{\text{pre}}^{\text{VBF S}}(x)}{s_{\text{pre}}^{\text{ggF B}}(x) + s_{\text{pre}}^{\text{EW B}}(x) + s_{\text{pre}}^{q\bar{q}ZZ}(x)}$$

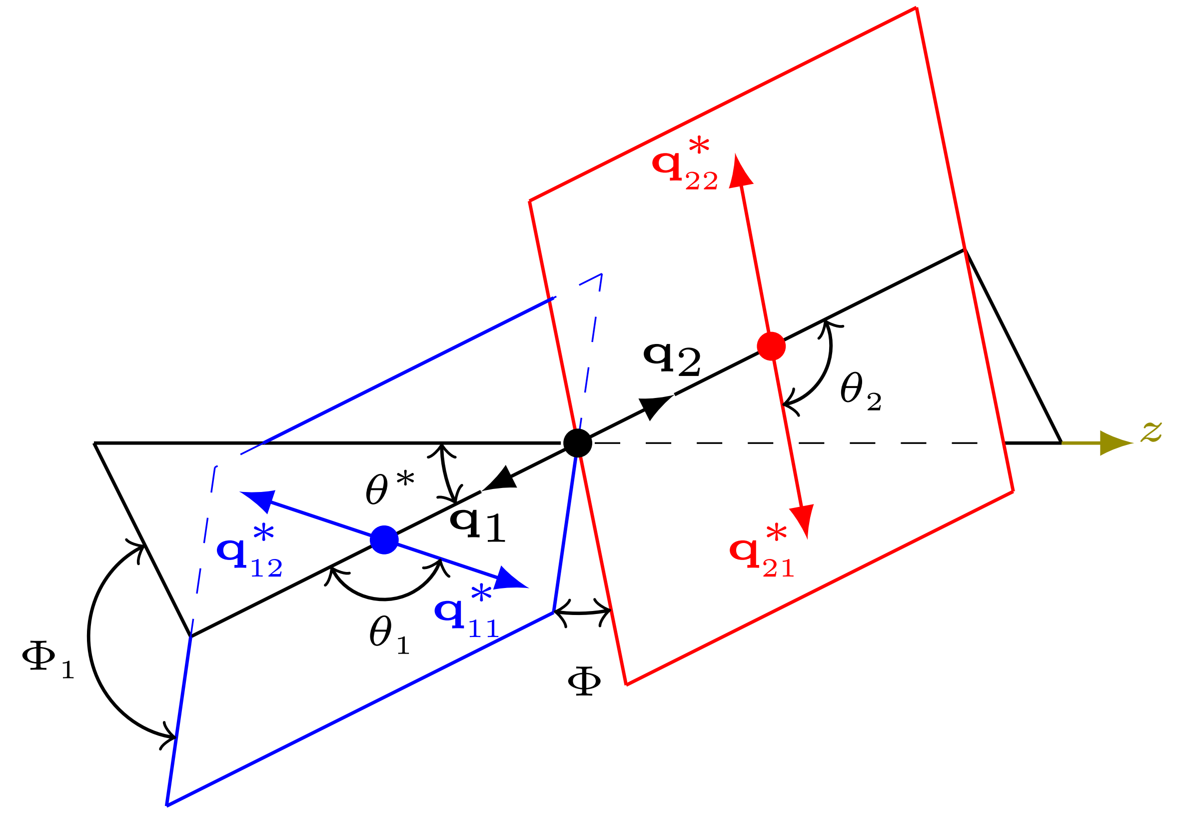


$$\text{SR}: D_{\text{pre}} > -0.85$$

$$\text{CR}: D_{\text{pre}} \leq -0.85$$

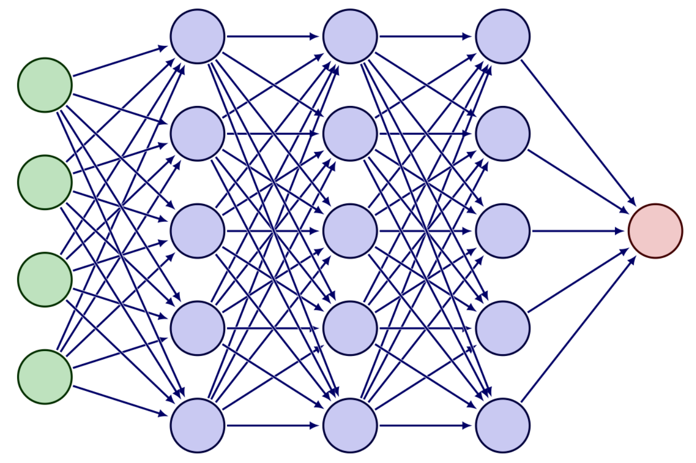
Preselection multi-classifier to define control/signal region.

Neural simulation-based inference (NSBI)



- A *balanced* classifier between hypotheses:

$$\nu(\theta_1) = \nu(\theta_2)$$

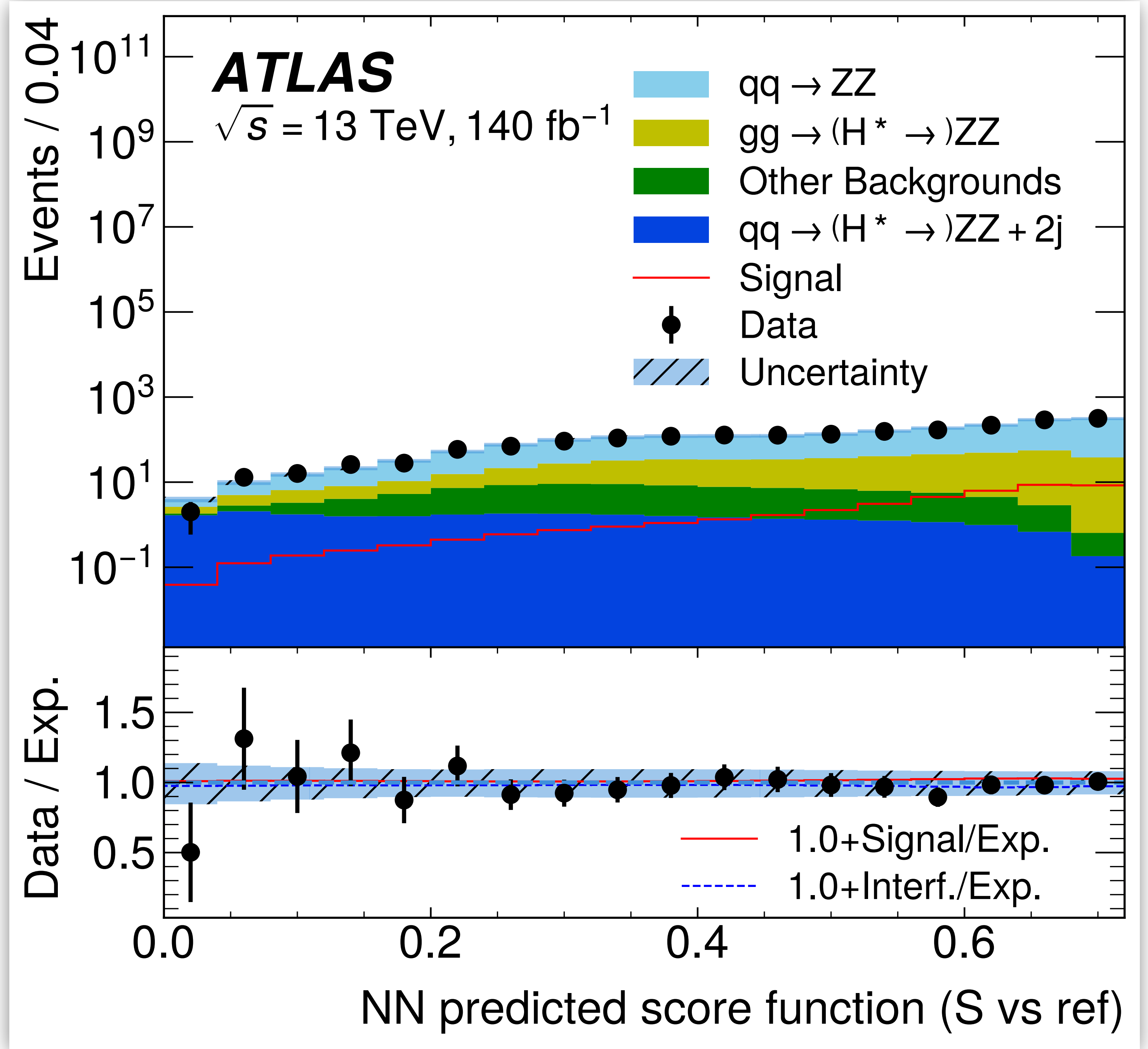


- Implicitly estimates probability ratio:

$$s(x;\theta_1,\theta_2) = \frac{p(x|\theta_1)}{p(x|\theta_1) + p(x|\theta_2)}$$

$$\frac{p(x|\theta_1)}{p(x|\theta_2)} = \frac{s}{1-s}$$

Likelihood ratio trick



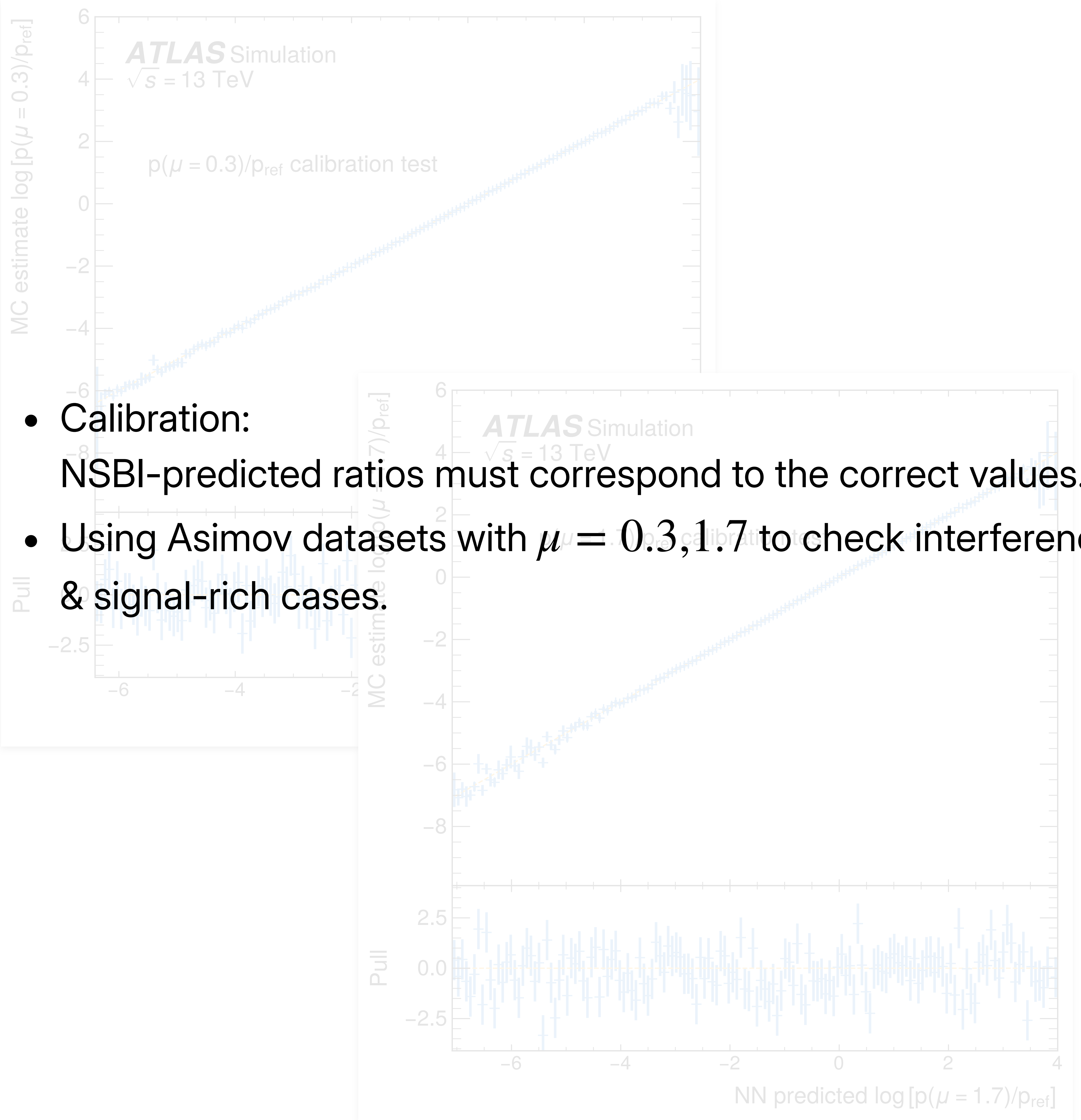
- Estimate each probability ratio against a reference hypothesis:

$$s_X(x) \mapsto \frac{p_X(x)}{p_{\text{ref}}(x)}$$

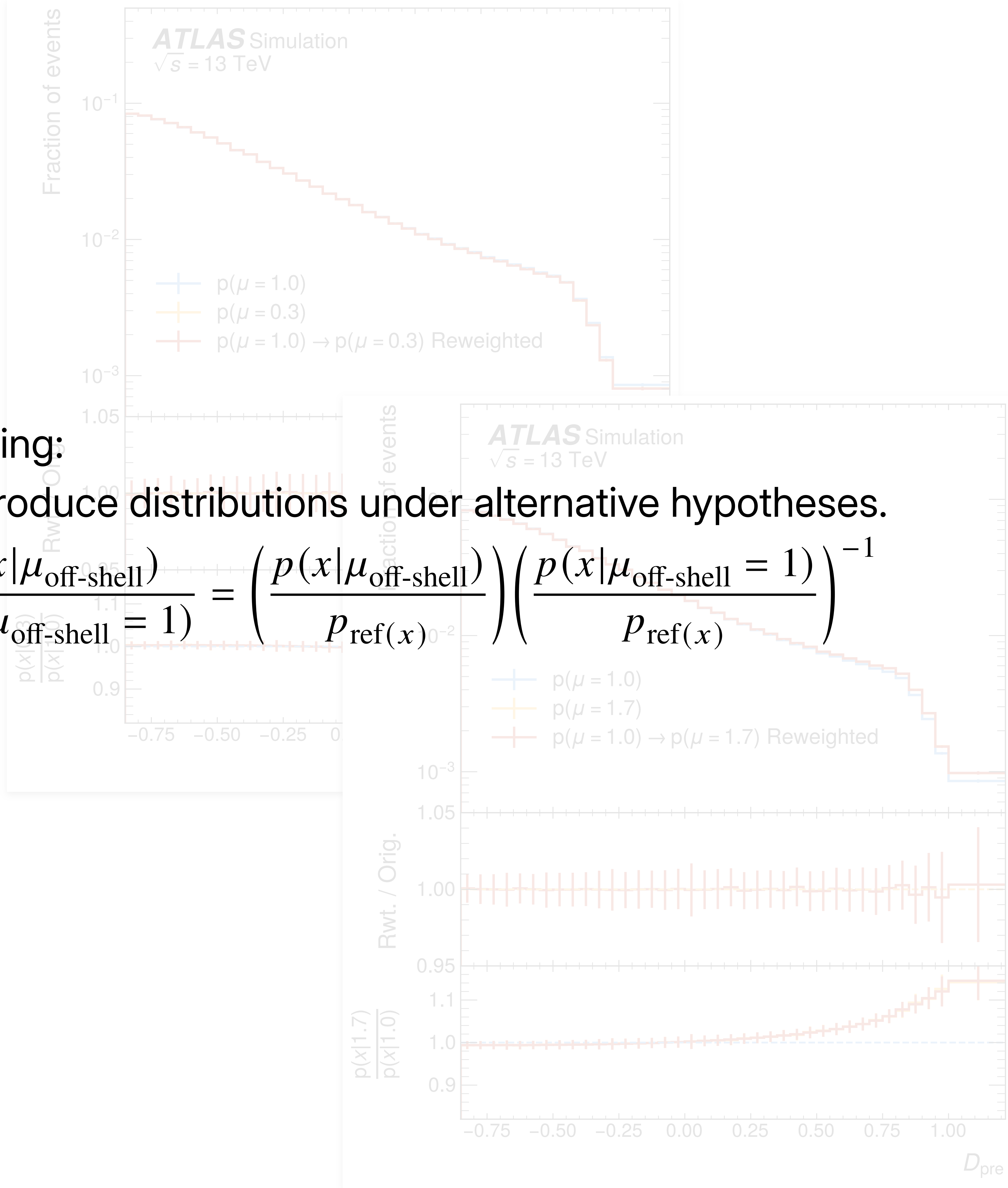
$$\nu_{\text{ref}} p_{\text{ref}} = \nu_S^{\text{ggF}} p_S^{\text{ggF}} + \nu_{\text{SBI}_{10}}^{\text{EW}} p_{\text{SBI}_{10}}^{\text{EW}}$$

- (Left: $X = \text{ggF}$).

NSBI accuracy diagnostics



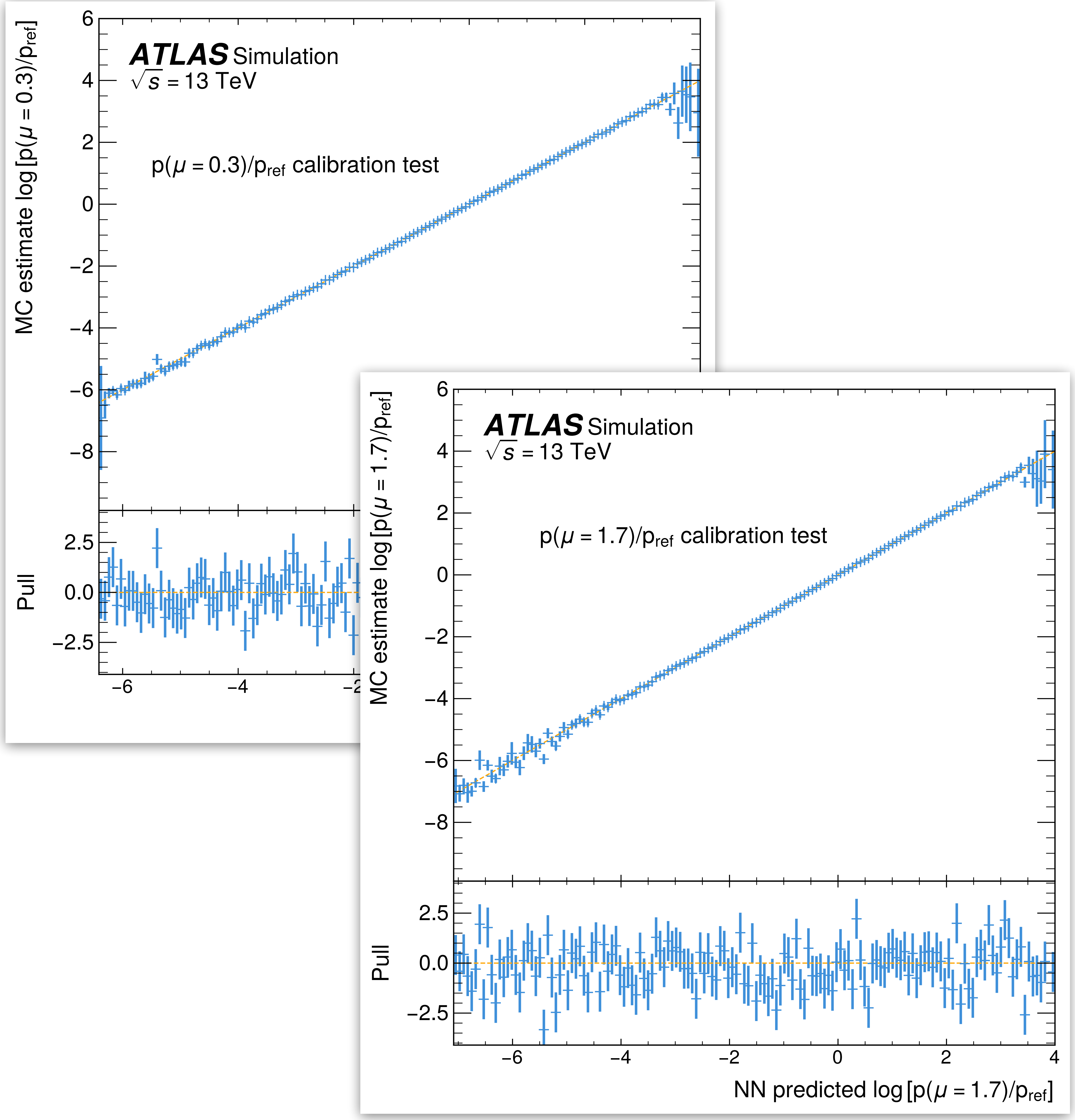
- Calibration:
  NSBI-predicted ratios must correspond to the correct values.
- Using Asimov datasets with $\mu = 0.3, 1.7$ to check interference- & signal-rich cases.

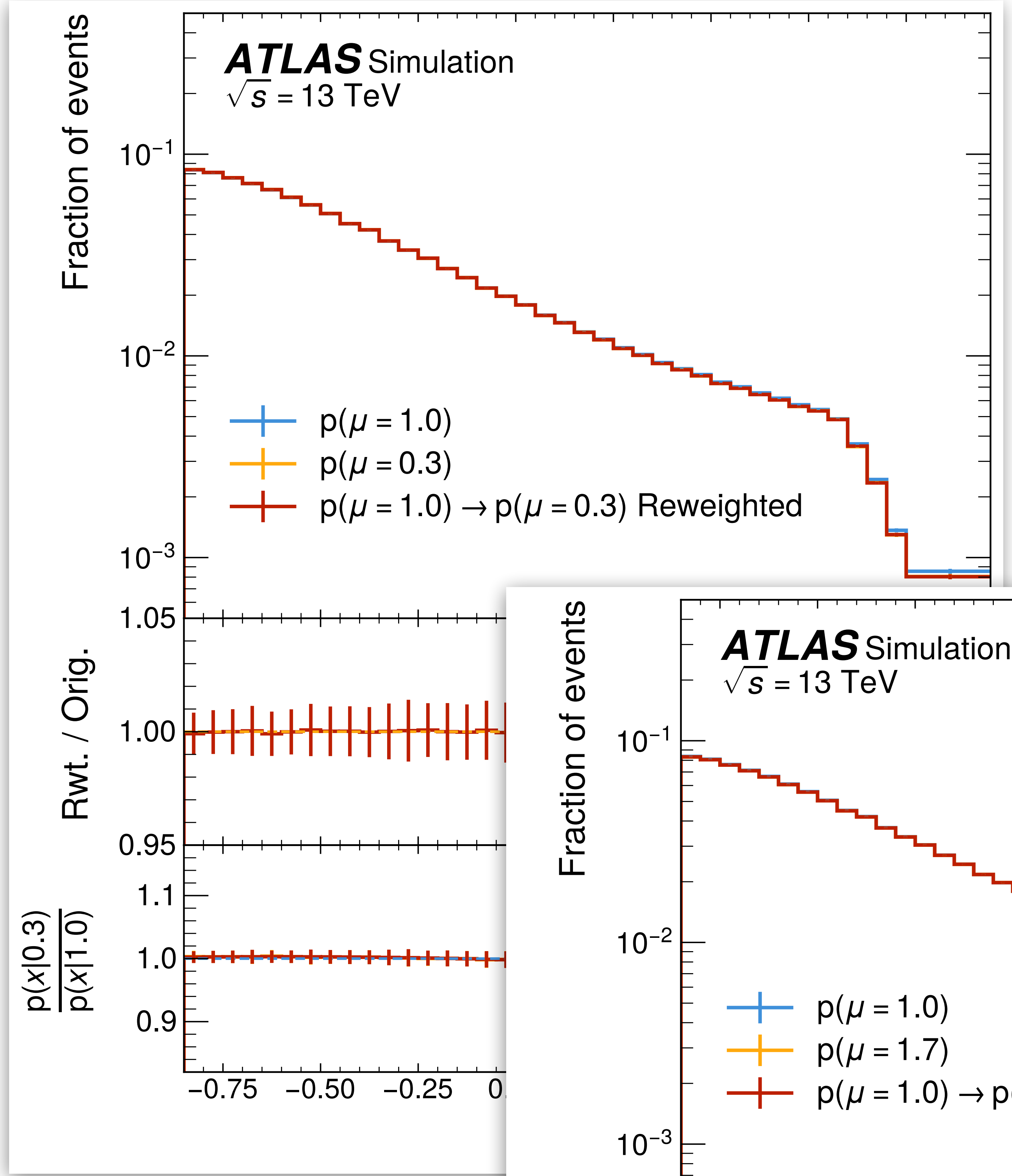


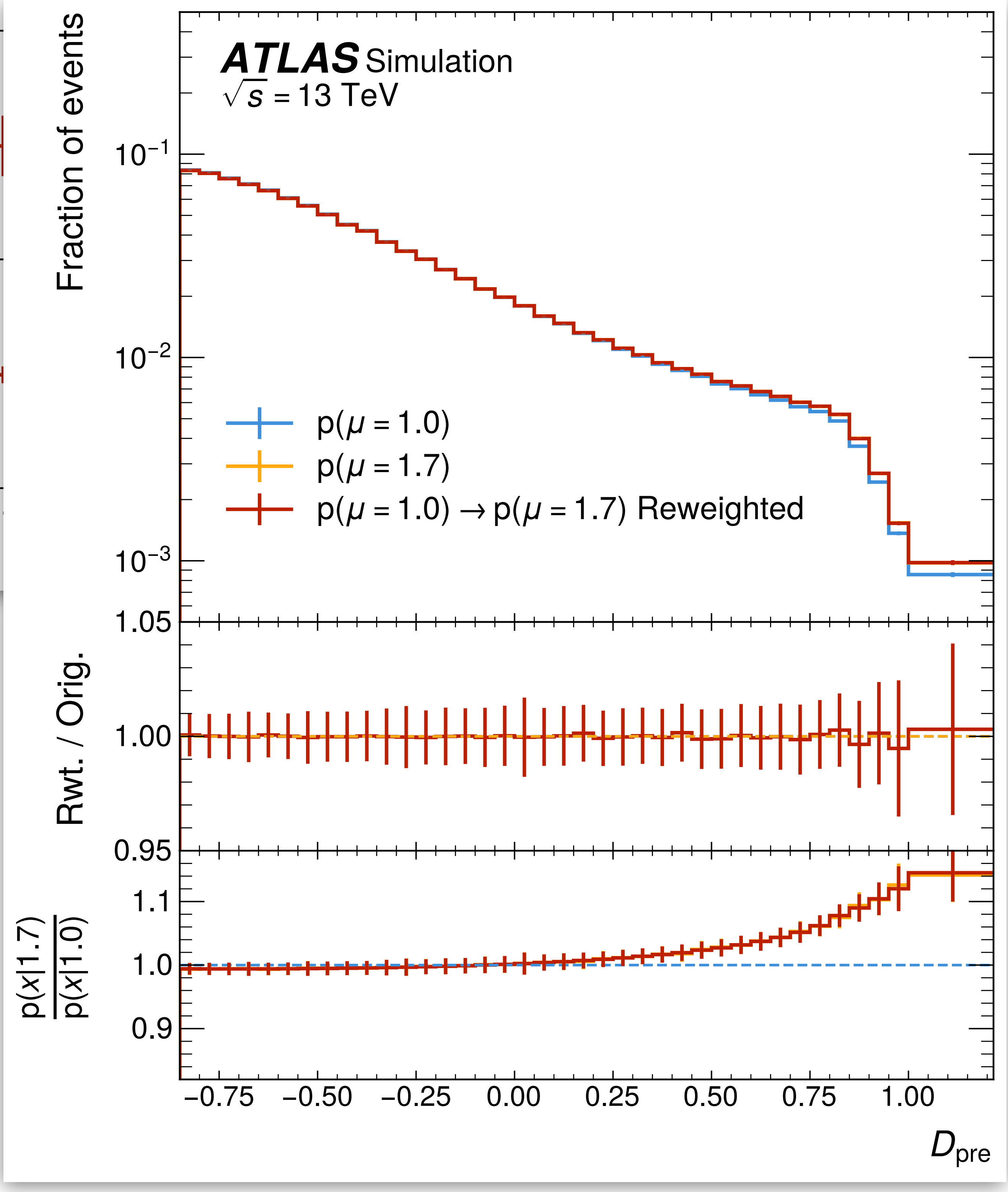
- Reweighting:
  Must reproduce distributions under alternative hypotheses.

$$\frac{p(x|\mu_{\text{off-shell}})}{p(x|\mu_{\text{off-shell}} = 1)} = \left(\frac{p(x|\mu_{\text{off-shell}})}{p_{\text{ref}(x)}}\right)\left(\frac{p(x|\mu_{\text{off-shell}} = 1)}{p_{\text{ref}(x)}}\right)^{-1}$$

NSBI accuracy diagnostics







(Reweighed observable not part of training)

Maximum likelihood with NSBI

1. Maximizing the unconditional likelihood:

$$-2\ln\lambda(\mu,\theta,\alpha) = -2\sum_{\text{regions }(I)} \ln\left[\text{Pois}(N_I|\nu_I(\mu,\theta,\alpha))\right]$$

$$-2\sum_{\text{events }(i)} \ln\left[\frac{p(x_i|\mu,\theta,\alpha)}{p_{\text{ref}}(x_i)}\right] + \sum_{\text{systematics }(m)} (\alpha_m - a_m)^2$$

- The presence of parameter-independent $p_{\text{ref}}(x)$ does not change location of global minimum.

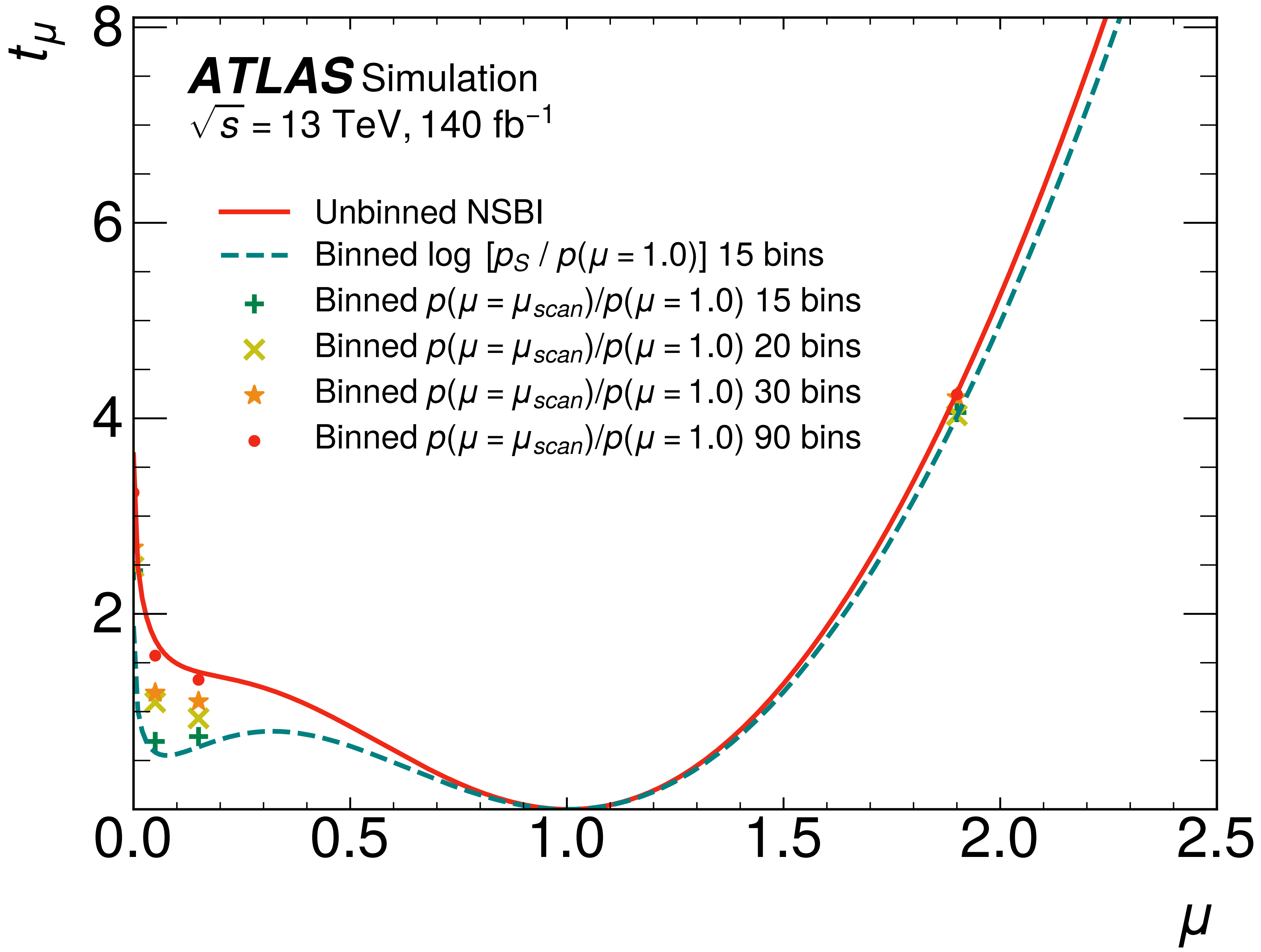
2. Profiling the conditional likelihood:

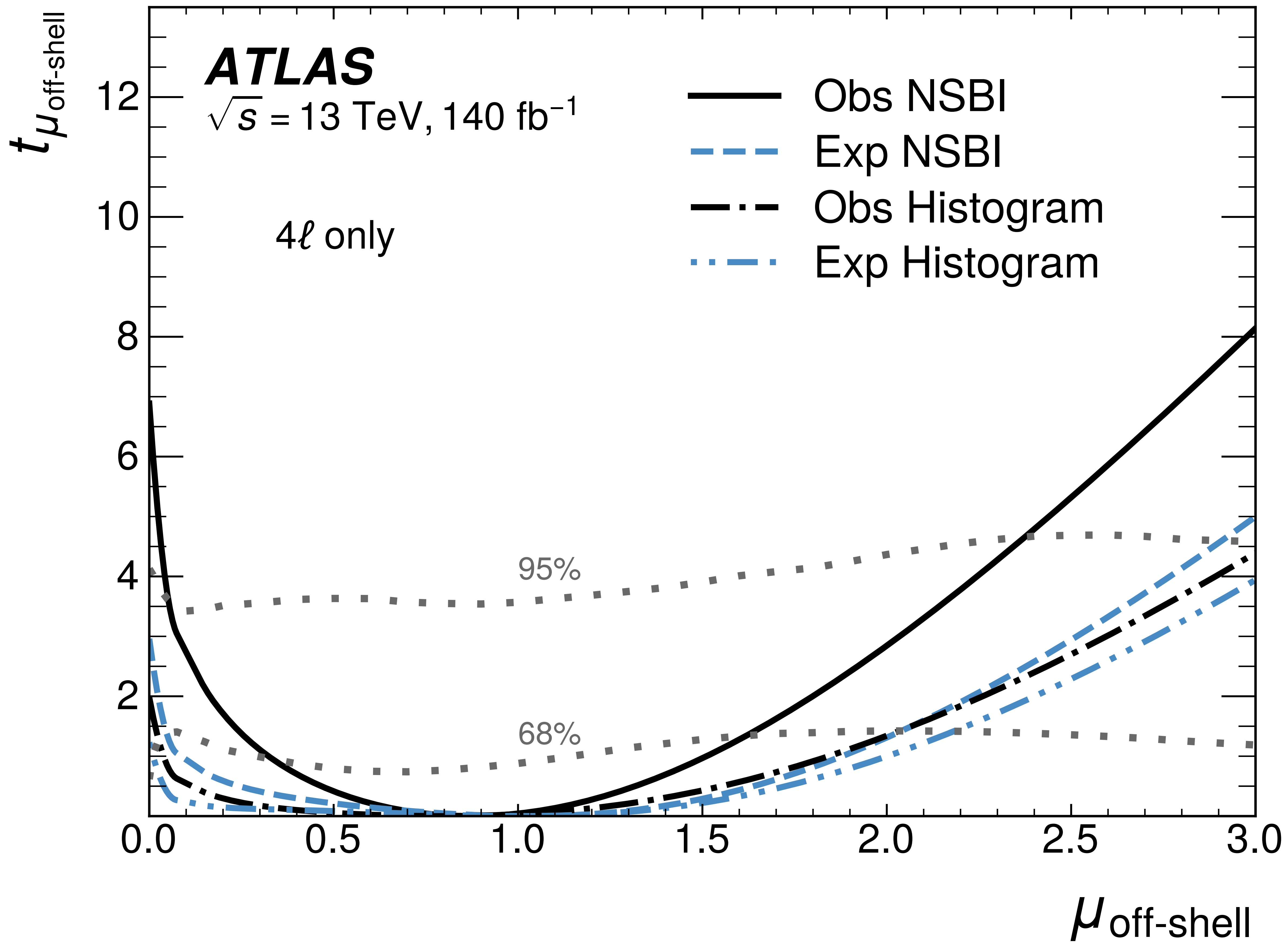
$$t_\mu = -2\ln\frac{\lambda(\mu,\hat{\hat{\alpha}}(\mu))}{\lambda(\hat{\mu},\hat{\alpha})}$$

- The common denominator $p_{\text{ref}}(x)$ cancels out between hypotheses.

$H^* \to 4\ell$ signal strength results

- Expected & observed constraint on off-shell Higgs signal strength both show improved constraint over traditional inference method.
- Improvement understood to be from optimality under Neyman-Pearson Lemma + un-binned nature of analysis:





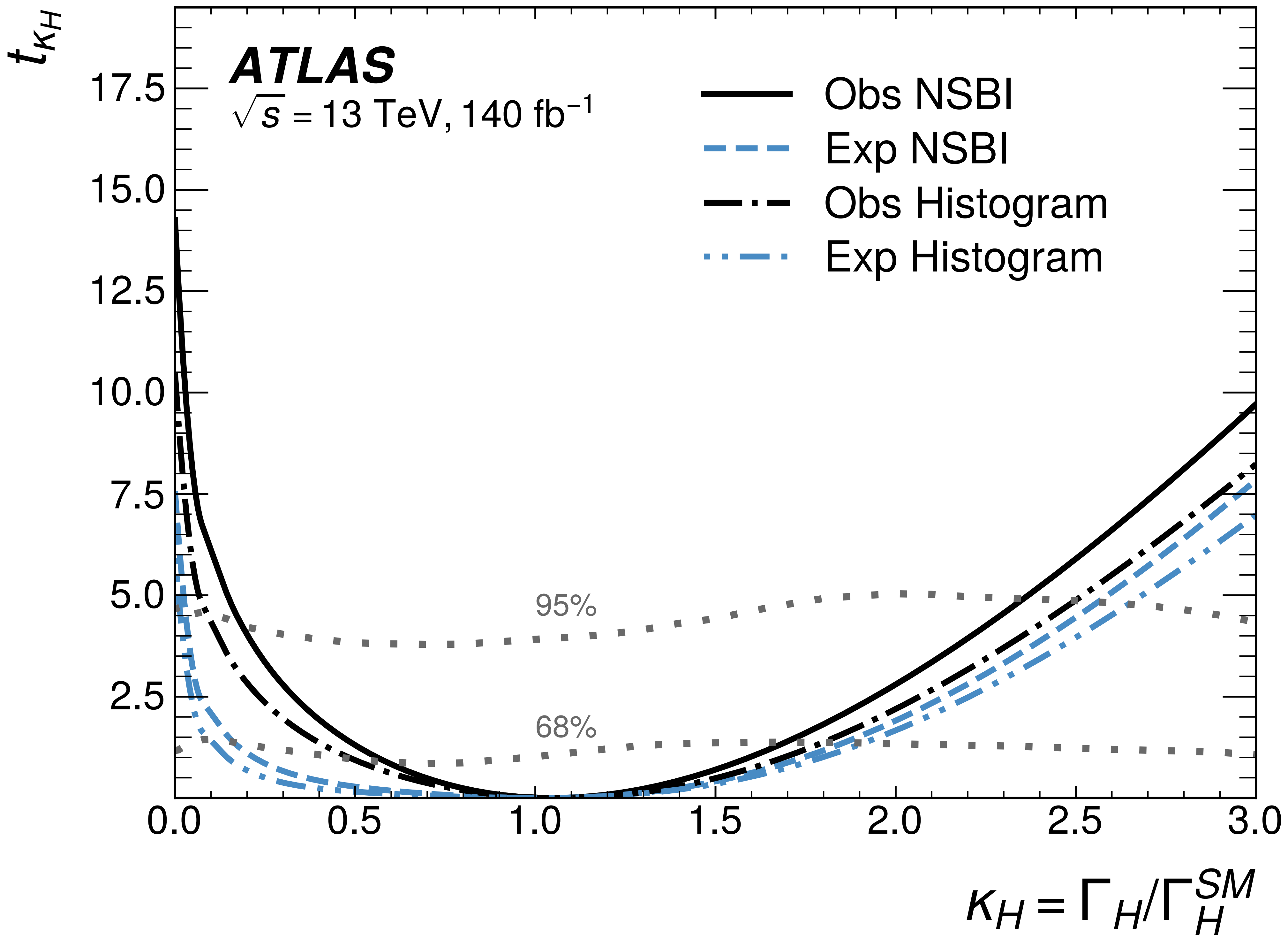
$\mu_{\text{off-shell}} = 0.87^{+0.75}_{-0.54}\ (1.00^{+1.04}_{-0.95})$

Results: combination & Γ_H interpretation

- Statistical combination:
  - NSBI 4ℓ off-shell
  - $2\ell 2\nu$ off-shell
    Phys. Lett. B 846 (2023) 138223
  - 4ℓ on-shell
    Eur. Phys. J. C 80 (2020) 957

$$-2\ln\lambda(\kappa_H,\theta_{HZZ},\theta,\alpha) = -2\sum_{\substack{\text{on-shell } 4\ell\\ \text{regions } (I)}} \ln\left[\text{Pois}(N_I|\nu_I(\theta_{HZZ}/\kappa_H,\alpha))\right] - 2\sum_{\substack{\text{off-shell}\\ \text{regions } (I)}} \ln\left[\text{Pois}(N_I|\nu_I(\theta_{HZZ},\theta,\alpha))\right] - 2\sum_{\substack{\text{off-shell } 4\ell\\ \text{SR events } (i)}} \ln\left[\frac{p(x_i|\theta_{HZZ},\theta,\alpha)}{p_{\text{ref}}(x_i)}\right] + \sum_{\text{systematics } (m)} (\alpha_m - a_m)^2$$

$$\theta_{HZZ} = \kappa^2_{g,\text{on-shell}}\kappa^2_{V,\text{on-shell}} = \kappa^4_{V,\text{on-shell}} = \kappa^2_{g,\text{off-shell}}\kappa^2_{g,\text{off-shell}} = \kappa^4_{V,\text{off-shell}}$$



Significant improvement previous result over same dataset!

New result	Old result
$\Gamma_H = 4.3^{+2.7}_{-1.9}\ (4.1^{+3.5}_{-3.4})$ MeV	$\Gamma_H = 4.4^{+3.1}_{-2.3}\ (4.1^{+3.8}_{-3.8})$ MeV

Different assumptions ⇒ interpretations available in backup.

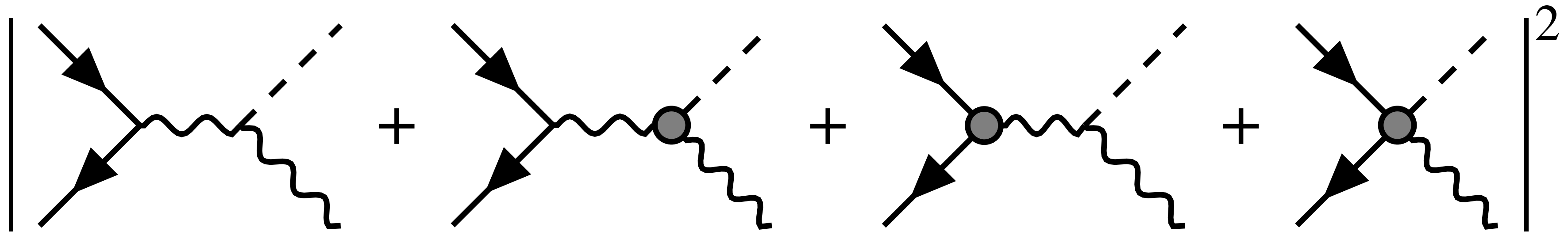
CMS $VHb\bar{b}$ SMEFT constraints

Physics overview

- SMEFT: Interference between SM and EFT operators.

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}}^{(d=4)} + \sum_{d\geq 5}\sum_{i} \frac{c_i^{(d)}}{\Lambda^{d-4}} \mathcal{O}_i^{(d)}$$

$\left|\mathcal{M}_{\text{SMEFT}}\right|^2 =$



- SMEFT operators affecting *VH(bb)* process:

Operator	Definition	Wilson coefficient	Operator	Definition	Wilson coefficient
$\mathcal{O}_{\text{Hq}}^{(1)}$	$i\text{H}^\dagger \overleftrightarrow{D}_\mu \text{H}\bar{\text{q}}_\text{L}\gamma^\mu \text{q}_\text{L}$	$c_{\text{Hq}}^{(1)}$	$\mathcal{O}_{\text{HWB}}$	$\text{H}^\dagger\sigma^a\text{H}\text{W}^a_{\mu\nu}\text{B}^{\mu\nu}$	c_{HWB}
$\mathcal{O}_{\text{Hq}}^{(3)}$	$i\text{H}^\dagger \sigma^a \overleftrightarrow{D}_\mu \text{H}\bar{\text{q}}_\text{L}\sigma^a\gamma^\mu \text{q}_\text{L}$	$c_{\text{Hq}}^{(3)}$	$\mathcal{O}_{\text{H}\widetilde{\text{W}}\text{B}}$	$\text{H}^\dagger\sigma^a\text{H}\text{W}^a_{\mu\nu}\widetilde{\text{B}}^{\mu\nu}$	$c_{\text{H}\widetilde{\text{W}}\text{B}}$
$\mathcal{O}_{\text{Hu}}$	$i\text{H}^\dagger \overleftrightarrow{D}_\mu \text{H}\bar{\text{u}}_\text{R}\gamma^\mu \text{u}_\text{R}$	c_{Hu}	$\mathcal{O}_{\text{HW}}$	$(\text{H}^\dagger\text{H})\,\text{W}^a_{\mu\nu}\text{W}^{a\mu\nu}$	c_{HW}
$\mathcal{O}_{\text{Hd}}$	$i\text{H}^\dagger \overleftrightarrow{D}_\mu \text{H}\bar{\text{d}}_\text{R}\gamma^\mu \text{d}_\text{R}$	c_{Hd}	$\mathcal{O}_{\text{H}\widetilde{\text{W}}}$	$(\text{H}^\dagger\text{H})\,\text{W}^a_{\mu\nu}\widetilde{\text{W}}^{a\mu\nu}$	$c_{\text{H}\widetilde{\text{W}}}$
$\mathcal{O}_{\text{HD}}$	$(\text{H}^\dagger D^\mu\text{H})^*(\text{H}^\dagger D_\mu \text{H})$	c_{HD}	$\mathcal{O}_{\text{HB}}$	$(\text{H}^\dagger\text{H})\,\text{B}_{\mu\nu}\text{B}^{\mu\nu}$	c_{HB}
$\mathcal{O}_{\text{H}\Box}$	$(\text{H}^\dagger\text{H})\Box(\text{H}^\dagger\text{H})$	$c_{\text{H}\Box}$	$\mathcal{O}_{\text{H}\widetilde{\text{B}}}$	$(\text{H}^\dagger\text{H})\,\text{B}_{\mu\nu}\widetilde{\text{B}}^{\mu\nu}$	$c_{\text{H}\widetilde{\text{B}}}$

- WCs rotated to mass eigenstate basis:

$$g_2^{ZZ} = -2\frac{v^2}{\Lambda^2}\left(s_w^2 c_{\text{HB}} + c_w^2 c_{\text{HW}} + s_w c_w c_{\text{HWB}}\right),$$

$$g_2^{Z\gamma} = -2\frac{v^2}{\Lambda^2}\left(s_w c_w (c_{\text{HW}} - c_{\text{HB}}) + \frac{1}{2}(s_w^2 - c_w^2) c_{\text{HWB}}\right),$$

$$g_2^{\gamma\gamma} = -2\frac{v^2}{\Lambda^2}\left(c_w^2 c_{\text{HB}} + s_w^2 c_{\text{HW}} - s_w c_w c_{\text{HWB}}\right),$$

$$g_4^{ZZ} = \tilde{g}_2^{ZZ} = -2\frac{v^2}{\Lambda^2}\left(s_w^2 c_{\text{H}\widetilde{\text{B}}} + c_w^2 c_{\text{H}\widetilde{\text{W}}} + s_w c_w c_{\text{H}\widetilde{\text{W}}\text{B}}\right),$$

$$g_4^{Z\gamma} = \tilde{g}_2^{Z\gamma} = -2\frac{v^2}{\Lambda^2}\left(s_w c_w (c_{\text{H}\widetilde{\text{W}}} - c_{\text{H}\widetilde{\text{B}}}) + \frac{1}{2}(s_w^2 - c_w^2) c_{\text{H}\widetilde{\text{W}}\text{B}}\right),$$

$$g_4^{\gamma\gamma} = \tilde{g}_2^{\gamma\gamma} = -2\frac{v^2}{\Lambda^2}\left(c_w^2 c_{\text{H}\widetilde{\text{B}}} + s_w^2 c_{\text{H}\widetilde{\text{W}}} - s_w c_w c_{\text{H}\widetilde{\text{W}}\text{B}}\right).$$

WCs targeted for *VH(bb)* measurement

Physics overview

$$|\mathcal{M}(\hat{s},\Theta,\theta,\varphi)|^2 = \sum_i a_i(\hat{s}) f_i(\Theta,\theta,\varphi),$$

$$f_1 = f_{LL} = \sin^2\Theta \sin^2\theta$$

$$f_2 = f^1_{TT} = \cos\Theta\cos\theta$$

$$f_3 = f^2_{TT} = (1+\cos^2\Theta)(1+\cos^2\theta)$$

$$f_4 = f^1_{LT} = \cos\varphi\sin\Theta\sin\theta$$

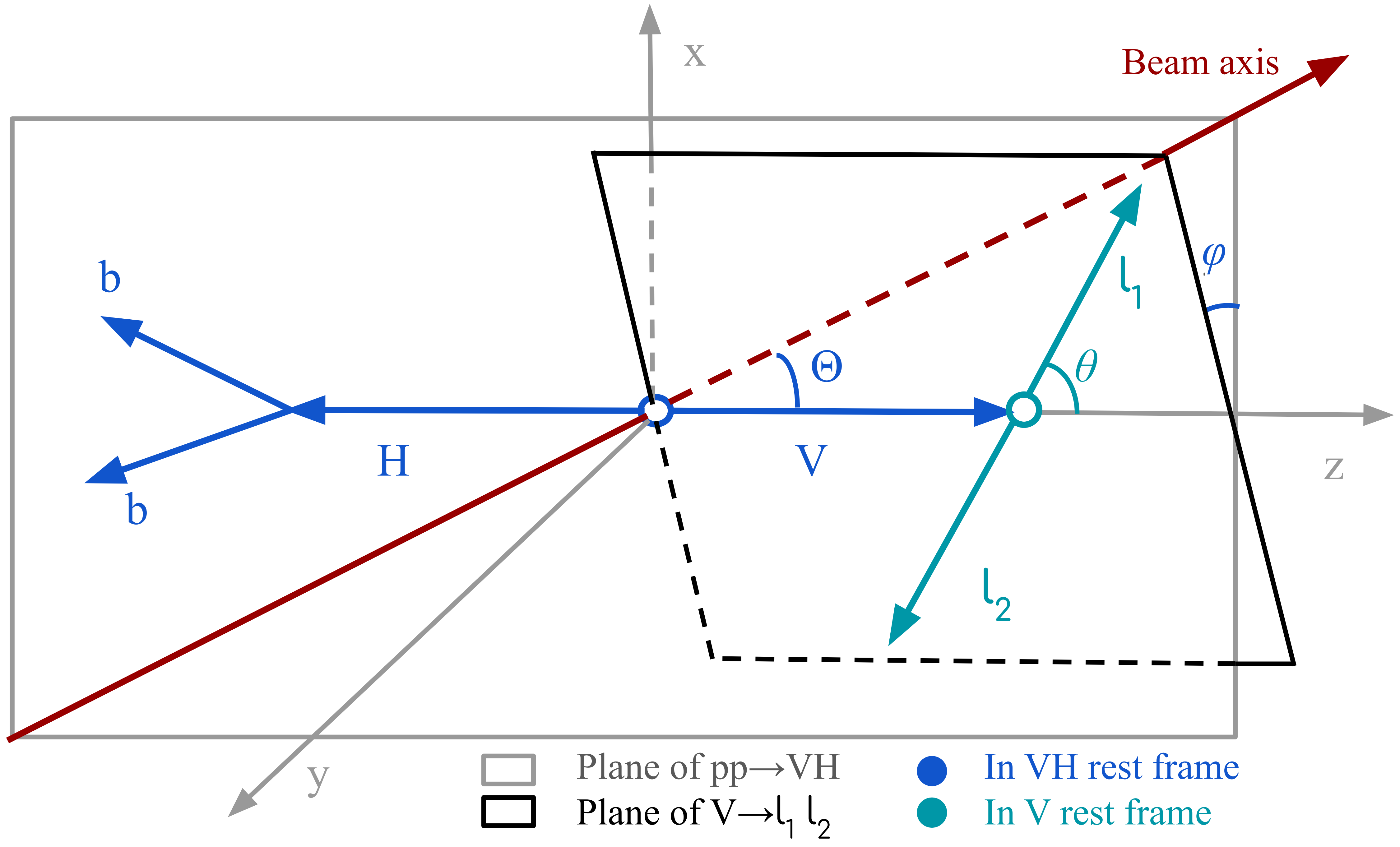
$$f_5 = f^2_{LT} = \cos\varphi\sin\Theta\sin\theta\cos\Theta\cos\theta$$

$$f_6 = \tilde{f}^1_{LT} = \sin\varphi\sin\Theta\sin\theta$$

$$f_7 = \tilde{f}^2_{LT} = \sin\varphi\sin\Theta\sin\theta\cos\Theta\cos\theta$$

$$f_8 = f_{TT'} = \cos^2\varphi\sin^2\Theta\sin^2\theta$$

$$f_9 = \tilde{f}_{TT'} = \sin^2\varphi\sin^2\Theta\sin^2\theta,$$



- EFT operators induces changes in:
  - Inclusive cross-section.
  - Object kinematics + angular distributions.
- Measurement of one = integration over others
  ⇒ Loss of BSM physics information.

→

Instead, use simulation-based inference technique utilizing complete event information to build optimal observable for BSM effects.

Object & event selection

V 0-lepton

Hbb Resolved — DeepJet

Variable	SR	V+HF CR	V+LF CR	$t\bar{t}$ CR
Max (b tag score of b_1 and b_2)	$\geq$ medium	$\geq$ medium	$<$ medium	$\geq$ medium
Min (b tag score of b_1 and b_2)	$\geq$ loose	$\geq$ loose	$\geq$ loose	$\geq$ loose
No. of additional jets	< 2	< 2	< 2	≥ 2
$\Delta\phi\ (\vec{p}_T^{\,trk,miss}, \vec{p}_T^{\,miss})$	< 0.5	< 0.5	< 0.5	—
$M\ (b_1, b_2)$	$\in [90, 150]$	$\in [50, 250] \cup \notin [90, 150]$	$\in [50, 250]$	$\in [50, 250]$

Hbb Boosted — ParticleNet

Variable	SR	V+HF CR	V+LF CR	$t\bar{t}$ CR
H PARTICLENET score	≥ 0.94	≥ 0.94	$\in [0.1, 0.94)$	≥ 0.94
$N_{b\text{-}tagged}$ jets outside H cand.	$= 0$	$= 0$	$= 0$	> 0
$m_{SD}{}^{H}$	$\in [90, 150]$	$\in [50, 250] \cup \notin [90, 150]$	> 50	> 50

V 1-lepton

Hbb Resolved

Variable	SR	V+HF CR	V+LF CR	$t\bar{t}$ CR
Max (b tag score of b_1,b_2)	$\geq$ medium	$\geq$ medium	$\geq$ loose and $<$ medium	$\geq$ tight
Min (b tag score of b_1,b_2)	$\geq$ loose	$\geq$ loose	—	—
No. of additional jets	< 2	< 2	—	≥ 2
$M\ (b_1, b_2)$	$\in [90, 150]$	$\in [50, 250] \cup \notin [90, 150]$	$\in [50, 250]$	$\in [50, 250]$

Hbb Boosted

Variable	SR	V+HF CR	V+LF CR	$t\bar{t}$ CR
H PARTICLENET score	≥ 0.94	≥ 0.94	$\in [0.1, 0.94)$	≥ 0.94
$N_{b\text{-}tagged}$ jets outside H cand.	$= 0$	$= 0$	$= 0$	> 0
$m_{SD}{}^{H}$	$\in [90, 150]$	$\in [50, 250] \cup \notin [90, 150]$	> 50	> 50

V 2-lepton

Hbb Resolved

Variable	SR	V+HF CR	V+LF CR	$t\bar{t}$ CR
Max (b tag score of b_1,b_2)	$\geq$ medium	$\geq$ medium	$<$ loose	$\geq$ tight
Min (b tag score of b_1,b_2)	$\geq$ loose	$\geq$ loose	$<$ loose	$\geq$ loose
m^{V}	$\in [75, 105]$	$\in [85, 97]$	$\in [75, 105]$	$\in [10, 75]$ or ≥ 120
$M\ (b_1, b_2)$	$\in [90, 150]$	$\in [50, 250] \cup \notin [90, 150]$	$\in [90, 150]$	$\in [50, 250]$

Hbb Boosted

Variable	SR	V+HF CR	V+LF CR	$t\bar{t}$ CR
H PARTICLENET score	≥ 0.94	≥ 0.94	< 0.94	≥ 0.94
m^{V}	$\in [75, 105]$	$\in [75, 105]$	$\in [75, 105]$	≤ 75 or ≥ 105
$m_{SD}{}^{H}$	$\in [90, 150]$	$\in [50, 250] \cup \notin [90, 150]$	> 50	> 50

Boosted Information Trees (BITs) as likelihood ratios

- Full-event likelihood:

$$L(\mathcal{D}|\boldsymbol{\theta}) = \frac{e^{-\mathcal{L}\sigma(\boldsymbol{\theta})}}{N!} \prod_{1\leq i\leq N} \mathcal{L}\sigma(\boldsymbol{\theta}) p(\mathbf{x}_i|\boldsymbol{\theta}),$$

- Likelihood ratio between SMEFT hypotheses:

$$q_\theta = \mathcal{L}(\sigma(\boldsymbol{\theta}) - \sigma(\boldsymbol{\theta_0})) - \sum_{1\leq i\leq N} \log R(\mathbf{x}_i|\boldsymbol{\theta}, \boldsymbol{\theta_0})$$

- Probability density ratio intractable at **detector-level**; but can be evaluated at **joint (detector+parton) level**.
- Signal: latent variables (parton-level observables) + regression targets (ratio of EFT weights) extracted from MadGraph5+SMEFTsim.
- Background: no EFT dependence ⇒ always unity.

$$R(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\theta_0}) = \frac{\sigma(\boldsymbol{\theta}) p(\mathbf{x}|\boldsymbol{\theta})}{\sigma(\boldsymbol{\theta_0}) p(\mathbf{x}|\boldsymbol{\theta_0})}$$

$$R(\mathbf{x}, \mathbf{z}|\boldsymbol{\theta}, \boldsymbol{\theta_0}) = \frac{p(\mathbf{x}, \mathbf{z}|\boldsymbol{\theta})}{p(\mathbf{x}, \mathbf{z}|\boldsymbol{\theta_0})} = \frac{p(\mathbf{z}|\boldsymbol{\theta})}{p(\mathbf{z}|\boldsymbol{\theta_0})}$$

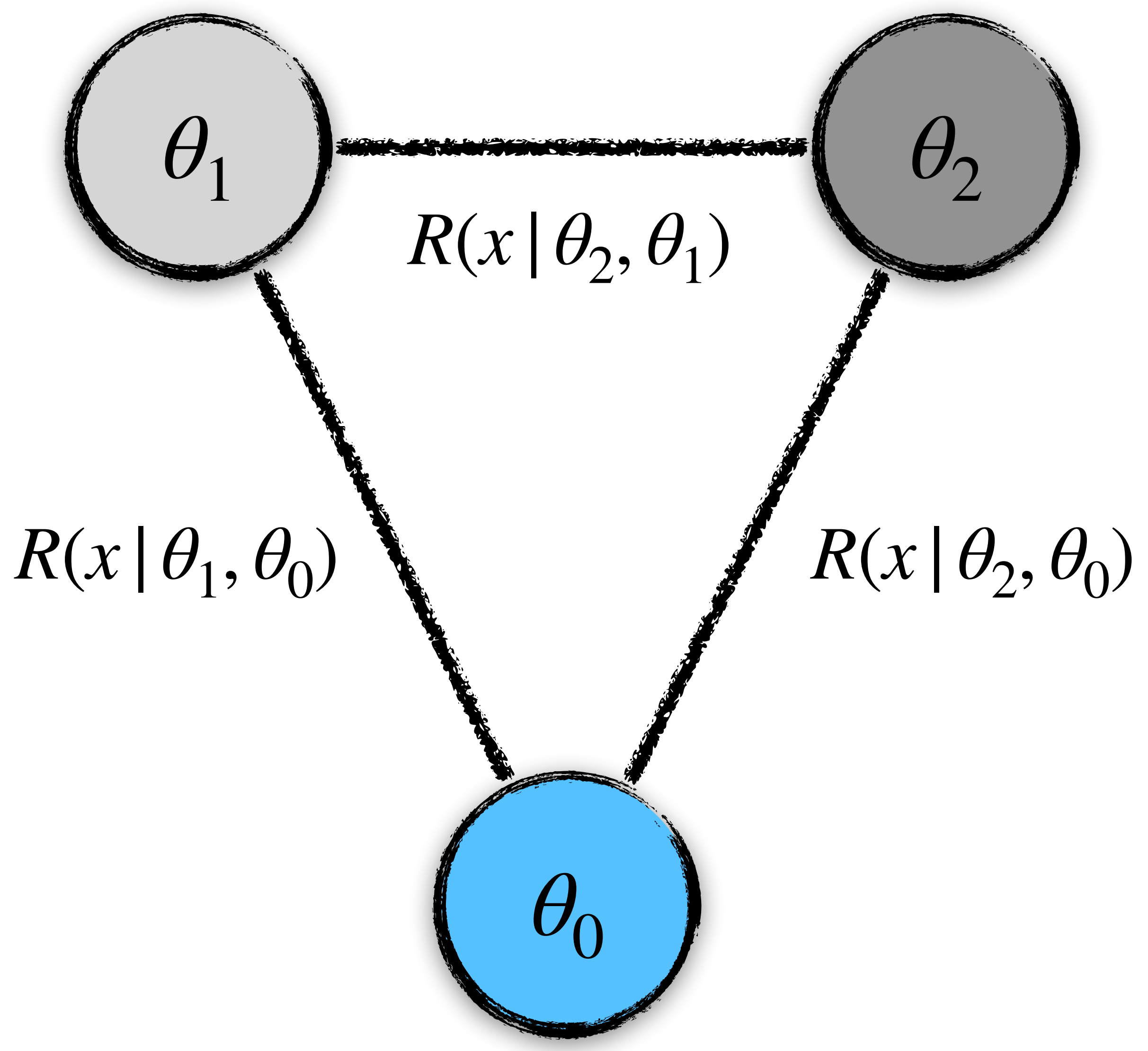
- BDTs trained to estimate the **polynomial expansion** of the likelihood ratio.

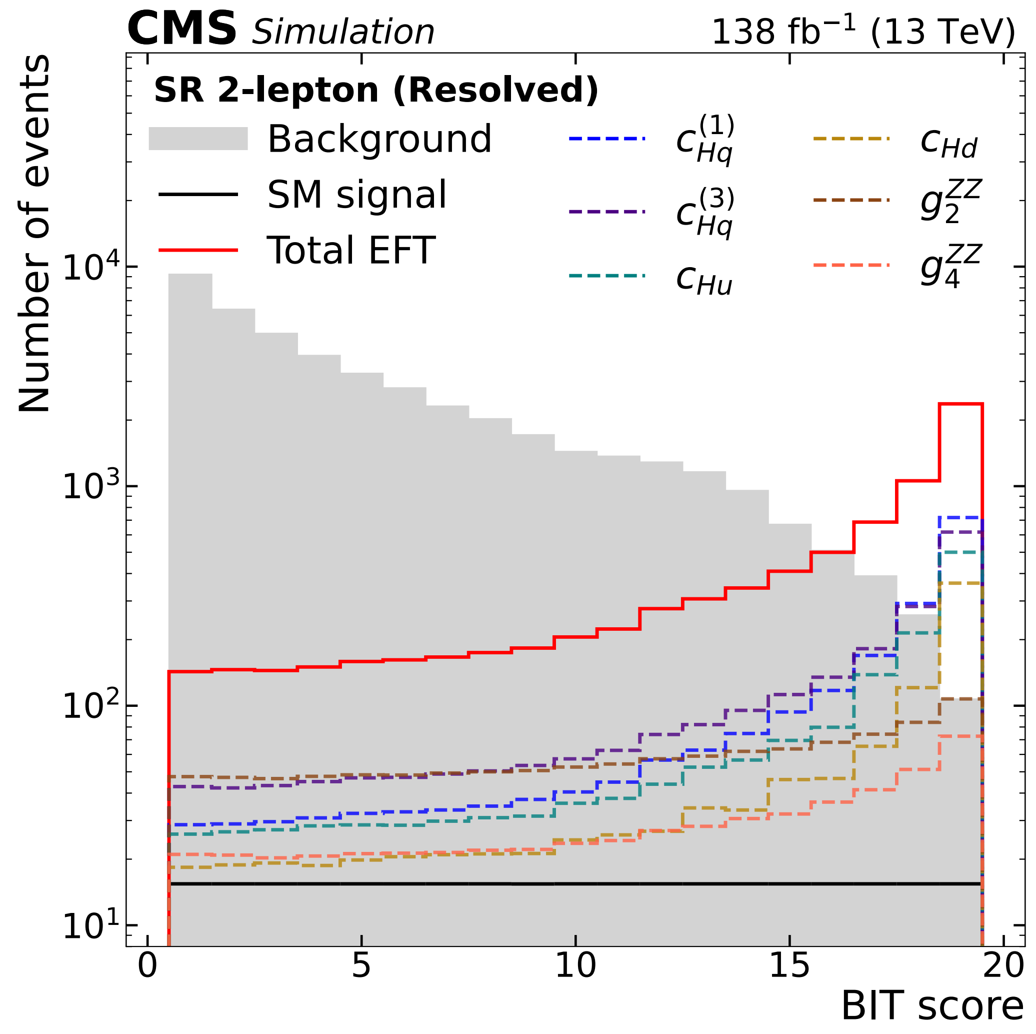
$$R(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\theta}_0) = 1 + \sum_{1\leq i\leq M} (\theta_i - \theta_0) R_i(\mathbf{x}) + \sum_{1\leq i\leq j\leq M} \frac{1}{2} (\theta_i - \theta_0)(\theta_j - \theta_0) R_{i,j}(\mathbf{x})$$

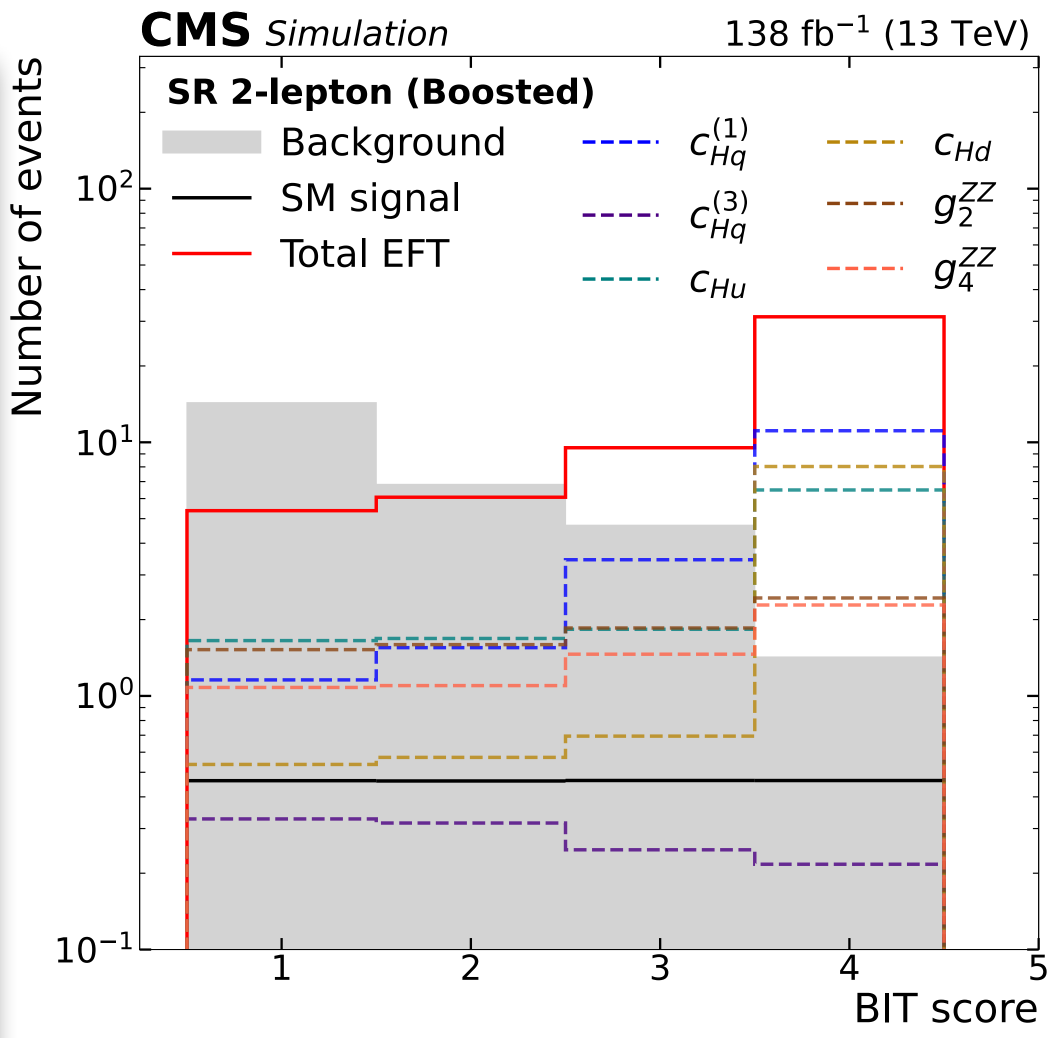
$$R_i(\mathbf{x}) = \frac{\partial}{\partial \theta_i} R(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\theta}_0)\bigg|_{\theta=\theta_0}, \qquad R_{i,j}(\mathbf{x}) = \frac{\partial}{\partial \theta_i}\frac{\partial}{\partial \theta_j} R(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\theta}_0)\bigg|_{\theta=\theta_0}$$

Computer Physics Communications 277 (2022) 108385

Optimization of BIT shapes



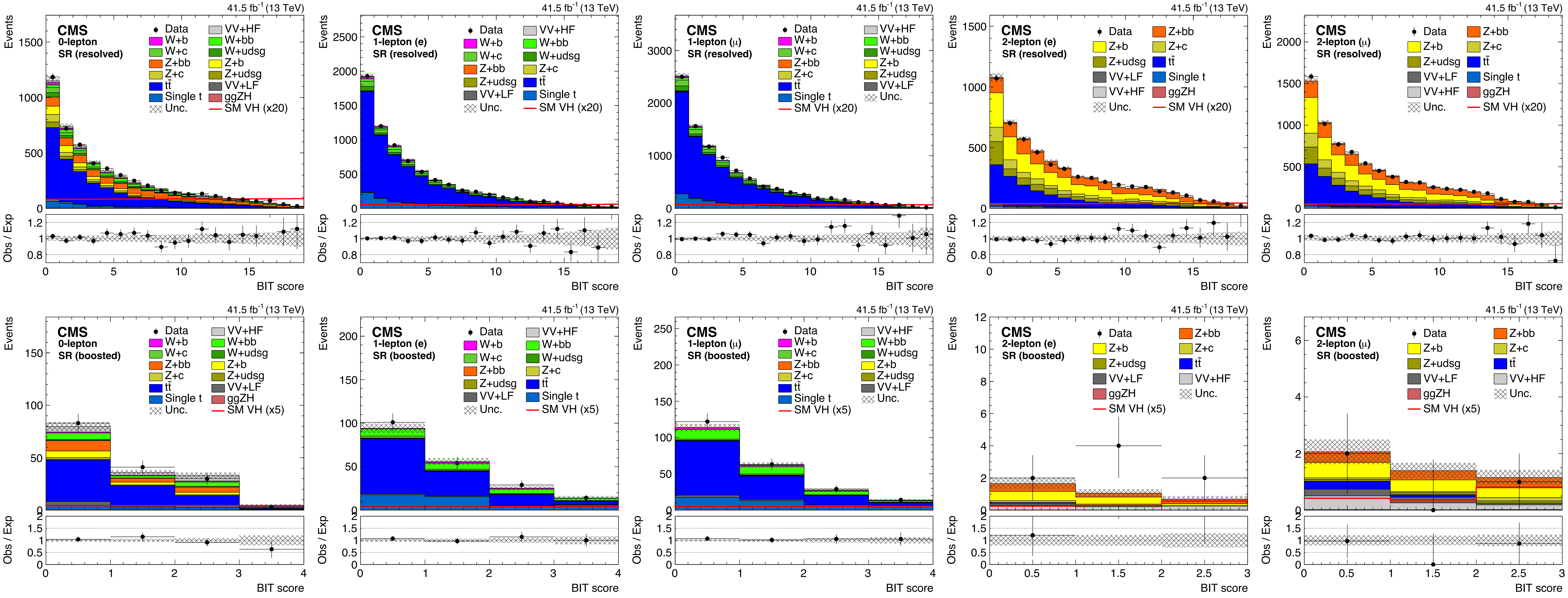




- $R(x\,|\,\theta, \theta_0)$ is the optimal observable to separate between θ and θ_0.
  - It is not in general optimal between (θ', θ).
- Task: select an θ_0 close-to-optimal for all θ_i.
  - Bayesian optimization maximize the expected $95\ \%$ confidence region (volume).
  - Avoid WC space with reduced sensitivity.
- Full WC space still probed in obtaining results.

Results

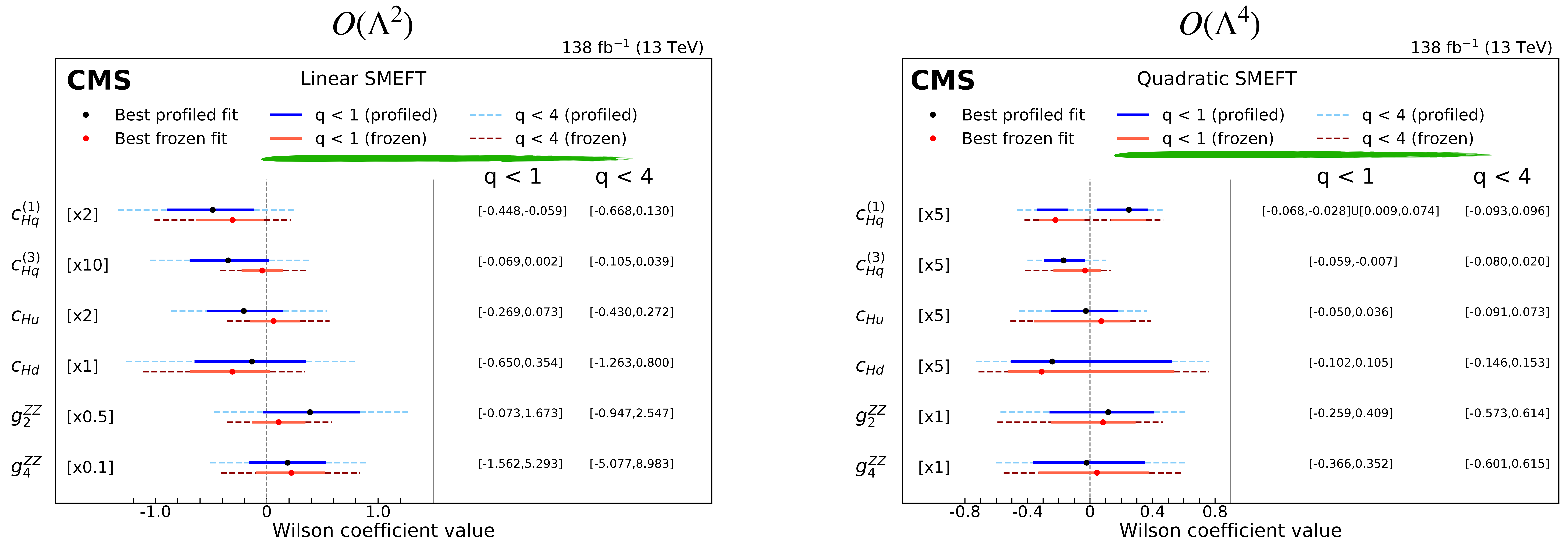
- BIT training inputs: energy-sensitive & $H \to b\bar{b}$ kinematics (0-lepton) + angular variables (1/2-lepton).
- BIT validated with background-free training: agreement between regression outputs & parton-level EFT weight ratios.



- Fit background-only hypothesis fit to data.

Results: SMEFT constraints

Results compatible with SM predictions with p-value of $73\ \%\ (\Lambda^2)$ and $84\ \%\ (\Lambda^4)$.



- Most comprehensive coverage of SMEFT operators achieved in the channel to date.
- Comparable constraints on **simultaneous** vs. **one-at-a-time** achieved.

(Asymptotic approximation not guaranteed)

Additional results: 2-dimensional likelihood scans, upper limit on Λ.

Conclusions

- Traditional inference methodologies to measure HEP processes rely on approximations of their underlying probability densities:
  - Low-dimensional summary of observables + uniqueness of the summarized probability densities on parameters.

- Higgs measurements performed by ATLAS & CMS using novel simulation-based inference techniques presented in this talk.
  - ATLAS: Measurement of the off-shell Higgs production rate & Γ_H interpretation under SM-like assumptions.
  - CMS: Constraints on SMEFT operators affecting $VH, H \to b\bar{b}$ process.
  - Does not rely on approximations associated with traditional inference.

- Simulation-based inference has the potential to significantly improve statistical power of physics measurements at the LHC.

Thank you for your attention!

Backup

Inference @ LHC: Challenges

- Neyman-Pearson Lemma: likelihood ratio is the most powerful test statistic *for **simple** hypotheses*.

$$p(x|\mu) = \frac{\mu \nu_S p_S(x) + \nu_B p_B(x)}{\nu(x)} \qquad s(x) = \frac{p_S(x)}{p_S(x) + p_B(x)} \qquad \frac{p(x_i|\mu)}{p(x_i|\mu = 0)} = \frac{\mu}{\mu \cdot \nu_s + \nu_B}\left(\frac{s(x_i)}{1 - s(x_i)} + \nu_B\right)$$

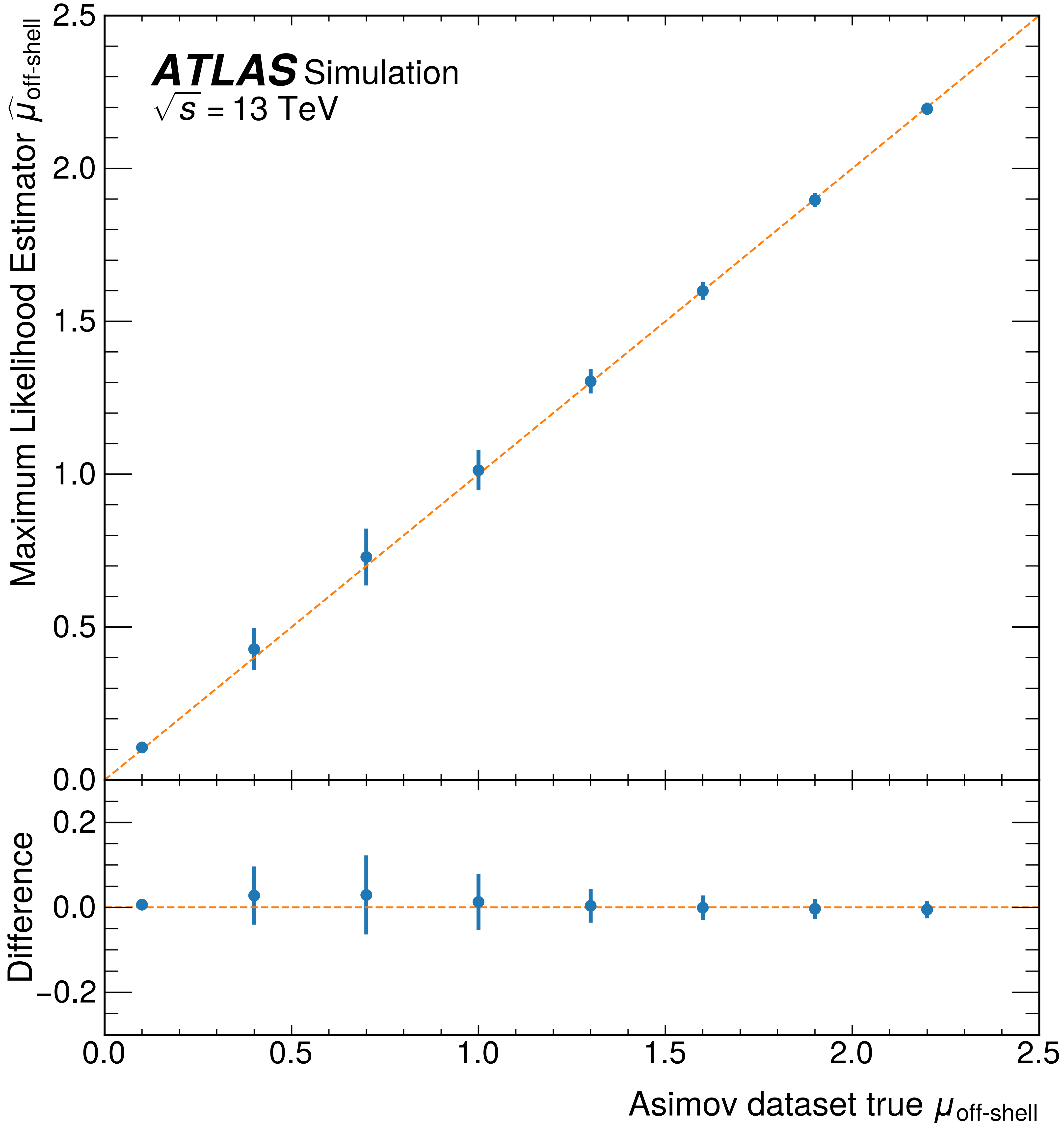
$$\nu(\mu) = \mu \nu_S + \nu_B$$

- Neyman-Pearson Lemma: likelihood ratio is the most powerful test statistic *for **simple** hypotheses*.

ATLAS off-shell: NN training

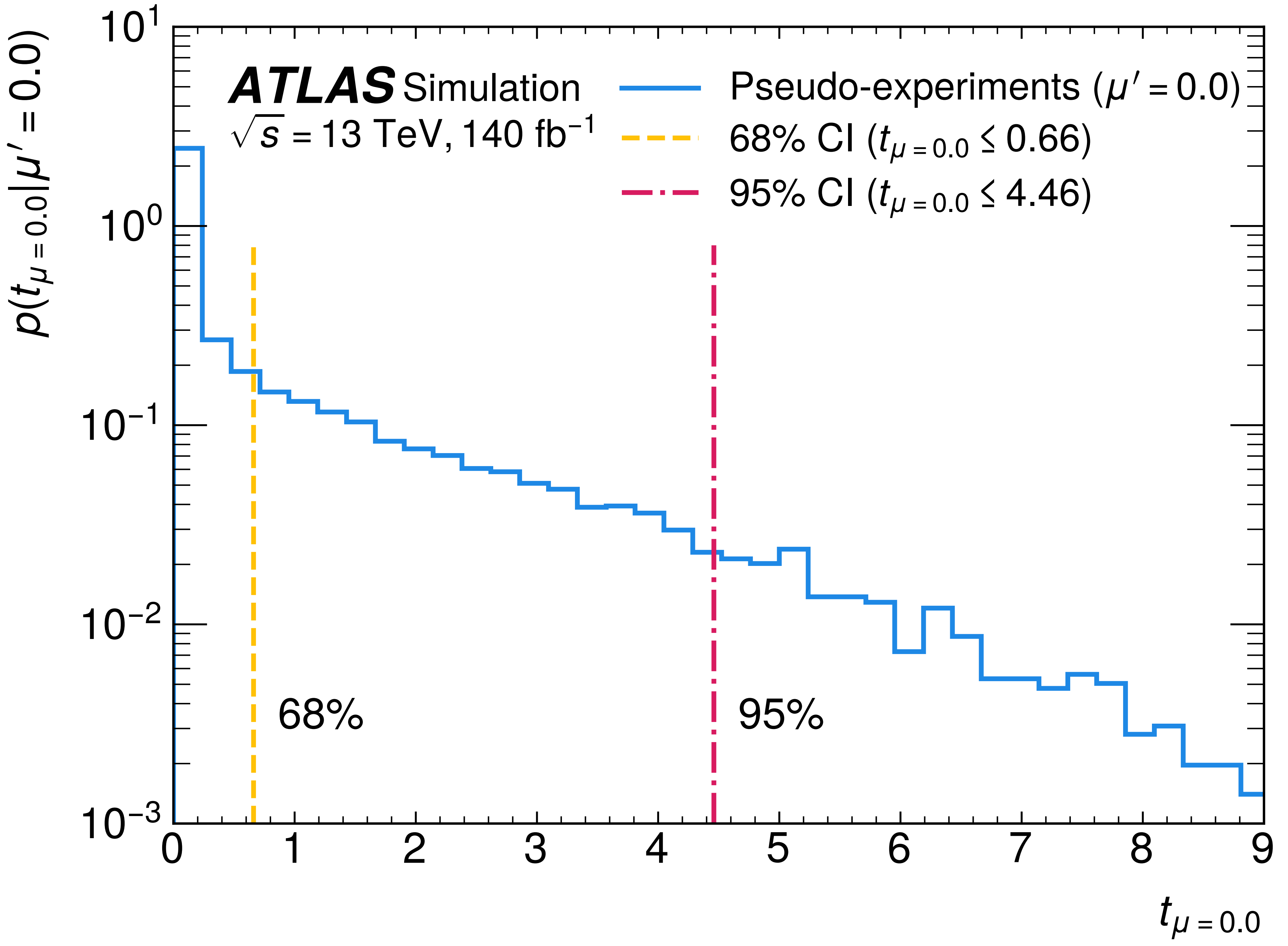
Parameter	**Value**
Number of hidden layers	5
Number of neurons per layer	1 000
Activation function	*swish*
Optimizer (initial learning rate)	*NAdam* (0.1)
Batch Size	256 – 4 096
Training epochs	100
k-fold cross validation	$k = 10$
Ensemble members (total)	100 – 700
Ensemble fraction (sampled without replacement)	80%
Train-validation split	90–10%

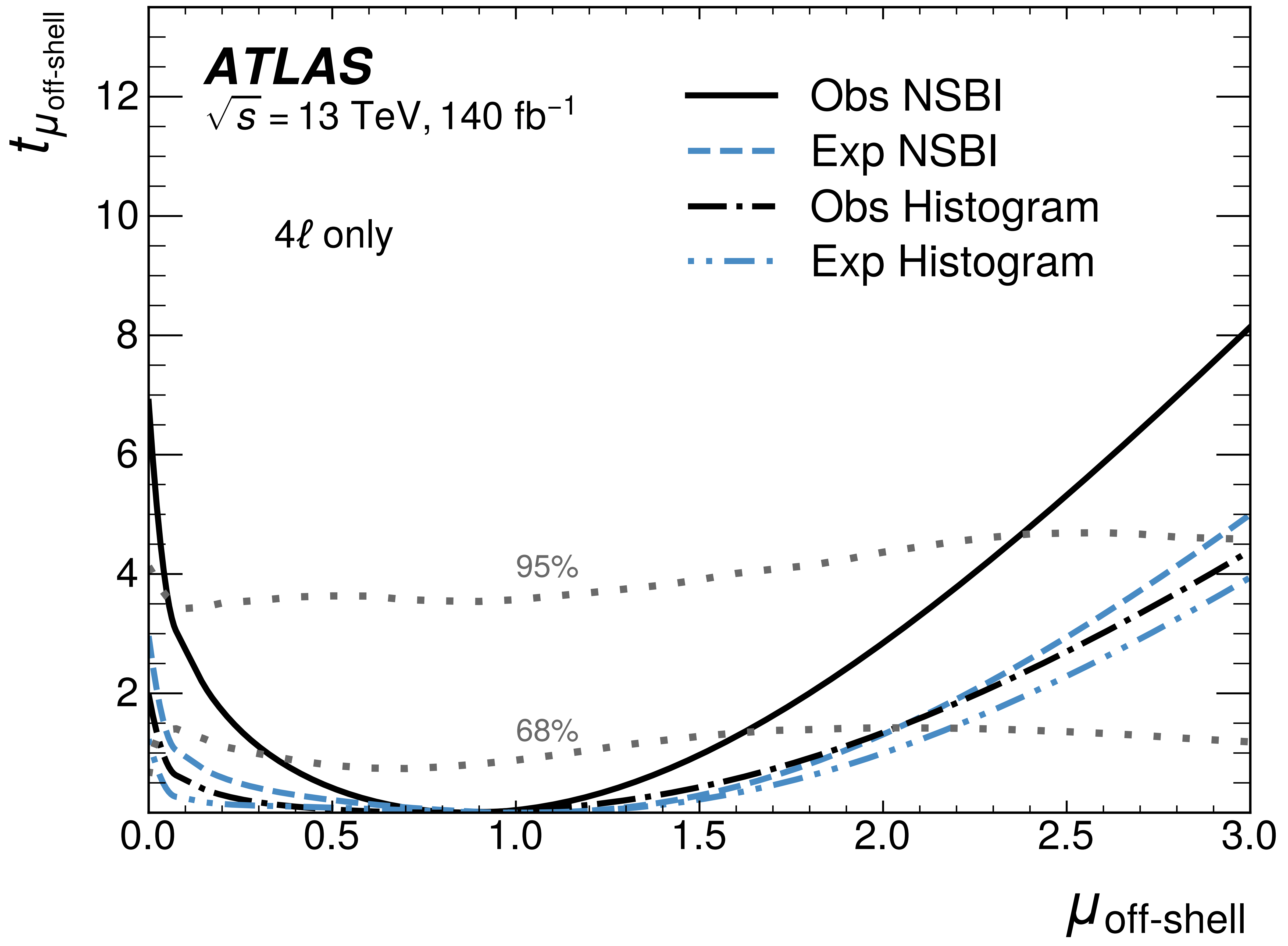
ATLAS off-shell: Asimov dataset closure



$H^* \to 4\ell$ signal strength results

- Expected & observed constraint on off-shell Higgs signal strength both show improved constraint over traditional inference method.
- Full Neyman construction to ensure CI coverage:





$\mu_{\text{off-shell}} = 0.87^{+0.75}_{-0.54}\ (1.00^{+1.04}_{-0.95})$

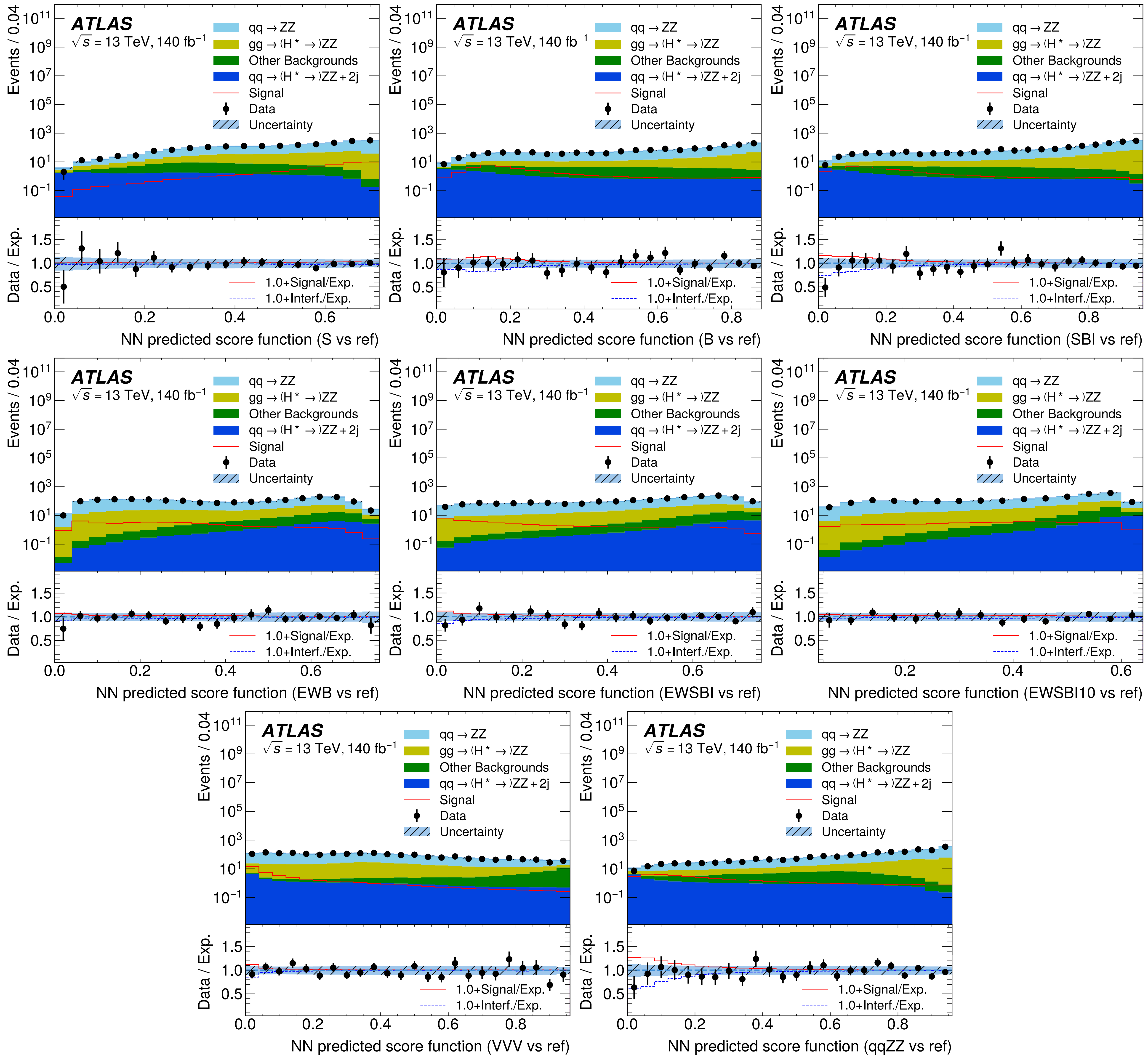
ATLAS off-shell: uncertainties

Uncertainty source	Absolute impact on $\mu_{\text{off-shell}}$	
	Nuisance Parameter	Auxiliary Observable
Electron uncertainties	$(-0.05, +0.06)$	$(-0.05, +0.06)$
Muon uncertainties	$(-0.03, +0.03)$	$(-0.02, +0.03)$
Jet uncertainties	$(-0.10, +0.10)$	$(-0.09, +0.11)$
Luminosity	$(-0.01, +0.01)$	$(-0.01, +0.01)$
Total experimental	$(-0.12, +0.13)$	$(-0.11, +0.12)$
$q\bar{q} \to ZZ$ modeling	$(-0.06, +0.07)$	$(-0.06, +0.07)$
$gg \to ZZ$ modeling	$(-0.08, +0.13)$	$(-0.07, +0.09)$
EW $q\bar{q} \to ZZ + 2j$ modeling	$(-0.01, +0.01)$	$(-0.01, +0.01)$
Total modeling	$(-0.10, +0.15)$	$(-0.09, +0.12)$
Systematic uncertainty	$(-0.16, +0.19)$	$(-0.14, +0.17)$
Statistical uncertainty	$(-0.49, +0.72)$	$(-0.50, +0.73)$
Total uncertainty	$(-0.54, +0.75)$	

ATLAS off-shell: full results

Parameter	Value	68% CL interval Observed	68% CL interval Expected	95% CL interval Observed	95% CL interval Expected
NSBI analysis					
$\mu_{\text{off-shell}}$ (4ℓ only)	0.87	[0.33, 1.62]	[0.05, 2.04]	[0.05, 2.38]	< 2.38
$\mu_{\text{off-shell}}$	1.06	[0.61, 1.67]	[0.17, 1.83]	[0.21, 2.24]	[0.01, 2.42]
Γ_H [MeV] (4ℓ only)	3.43	[1.37, 6.71]	[0.20, 8.25]	[0.18, 9.98]	< 12.09
Γ_H [MeV]	4.29	[2.41, 6.95]	[0.66, 7.61]	[0.76, 9.66]	[0.12, 10.50]
R_{gg}	1.19	[0.53, 2.07]	[0.02, 1.92]	< 2.96	< 2.73
R_{VV}	0.95	[0.61, 1.39]	[0.31, 1.70]	[0.30, 1.86]	[0.06, 2.14]
Histogram-based analysis					
$\mu_{\text{off-shell}}$ (4ℓ only)	0.79	[0.02, 2.00]	< 2.14	< 2.97	< 3.10
$\mu_{\text{off-shell}}$	1.09	[0.54, 1.81]	[0.08, 1.90]	[0.10, 2.41]	[0.01, 2.52]
Γ_H [MeV] (4ℓ only)	3.43	[0.10, 8.42]	< 8.89	< 12.48	< 12.89
Γ_H [MeV]	4.37	[2.13, 7.43]	[0.35, 7.94]	[0.39, 10.14]	< 10.79
R_{gg}	1.23	[0.00, 2.20]	< 1.98	< 3.15	< 2.84
R_{VV}	0.95	[0.60, 1.43]	[0.27, 1.74]	[0.26, 1.90]	[0.02, 2.18]

ATLAS off-shell: NSBI classifier output distributions



CMS VHbb: 2D likelihood scans