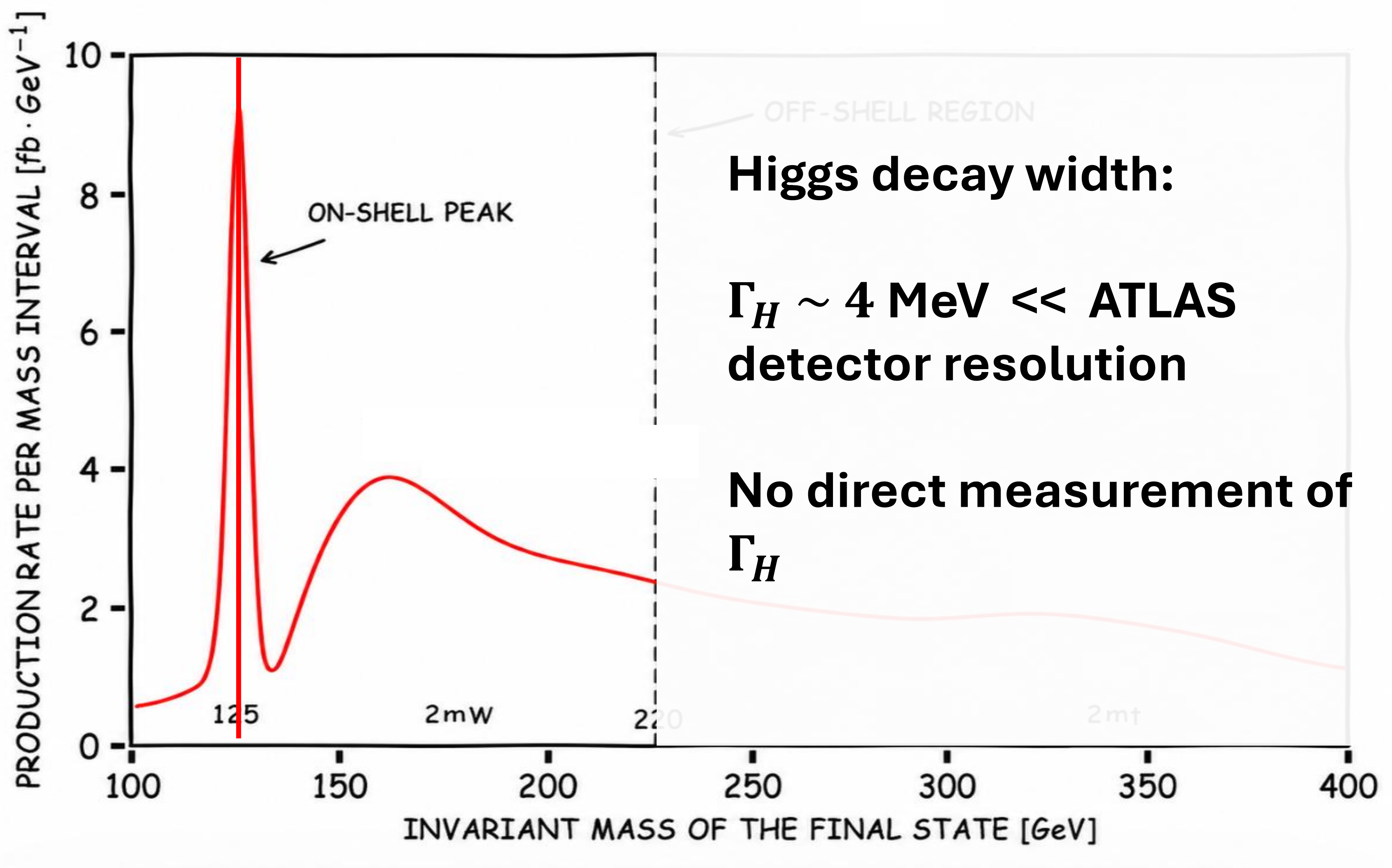
Off-Shell Higgs Production measurement in the $H \to ZZ^{(*)} \to 4l$ and $H \to ZZ^{(*)} \to 2l2v$ decay channels with ATLAS Run 3 data using Neural Simulation Based Inference

ATLAS EXPERIMENT

MAX-PLANCK-INSTITUT FÜR PHYSIK

Off-Shell Higgs

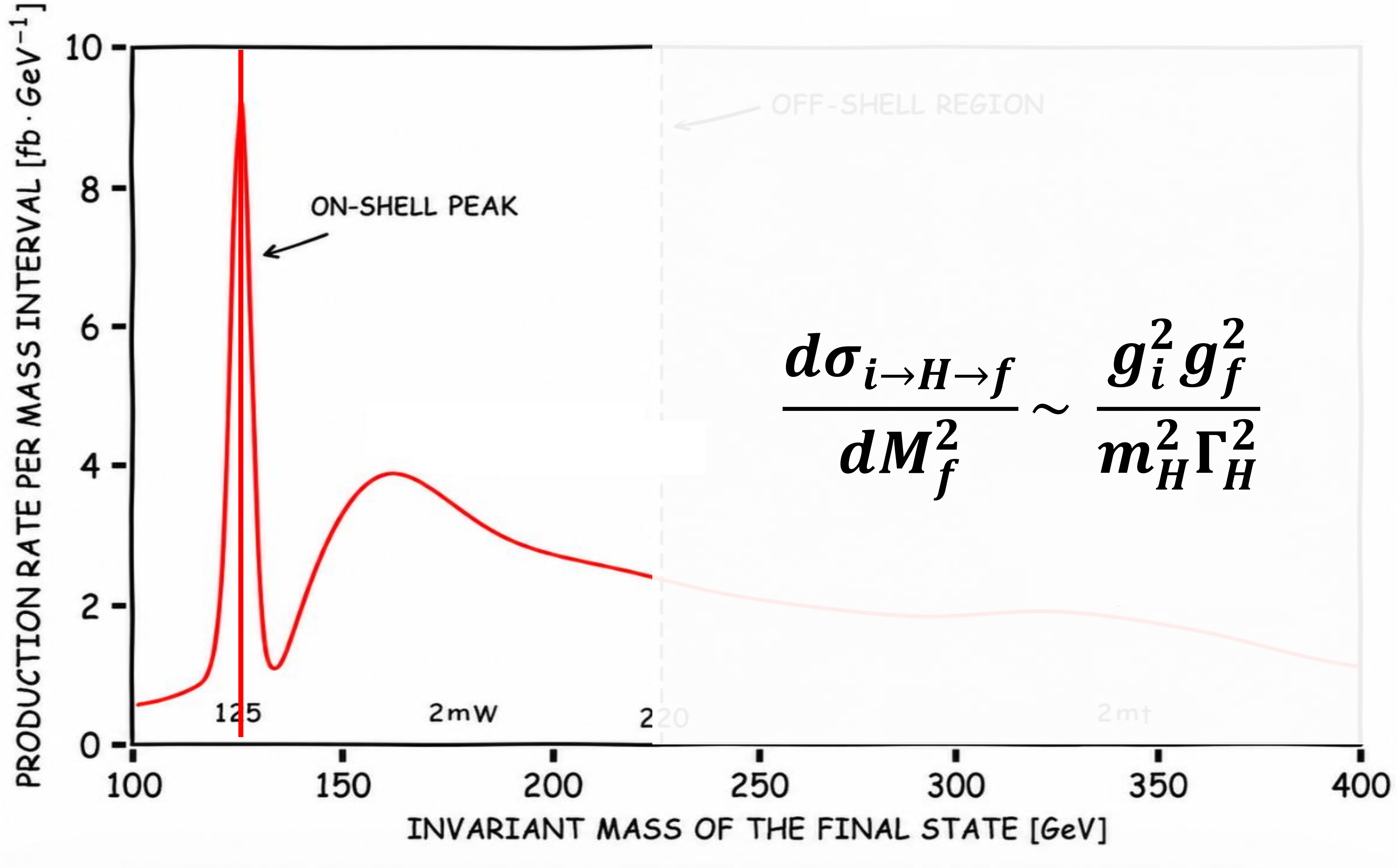


Higgs decay width:

$\Gamma_H \sim 4$ MeV << ATLAS detector resolution

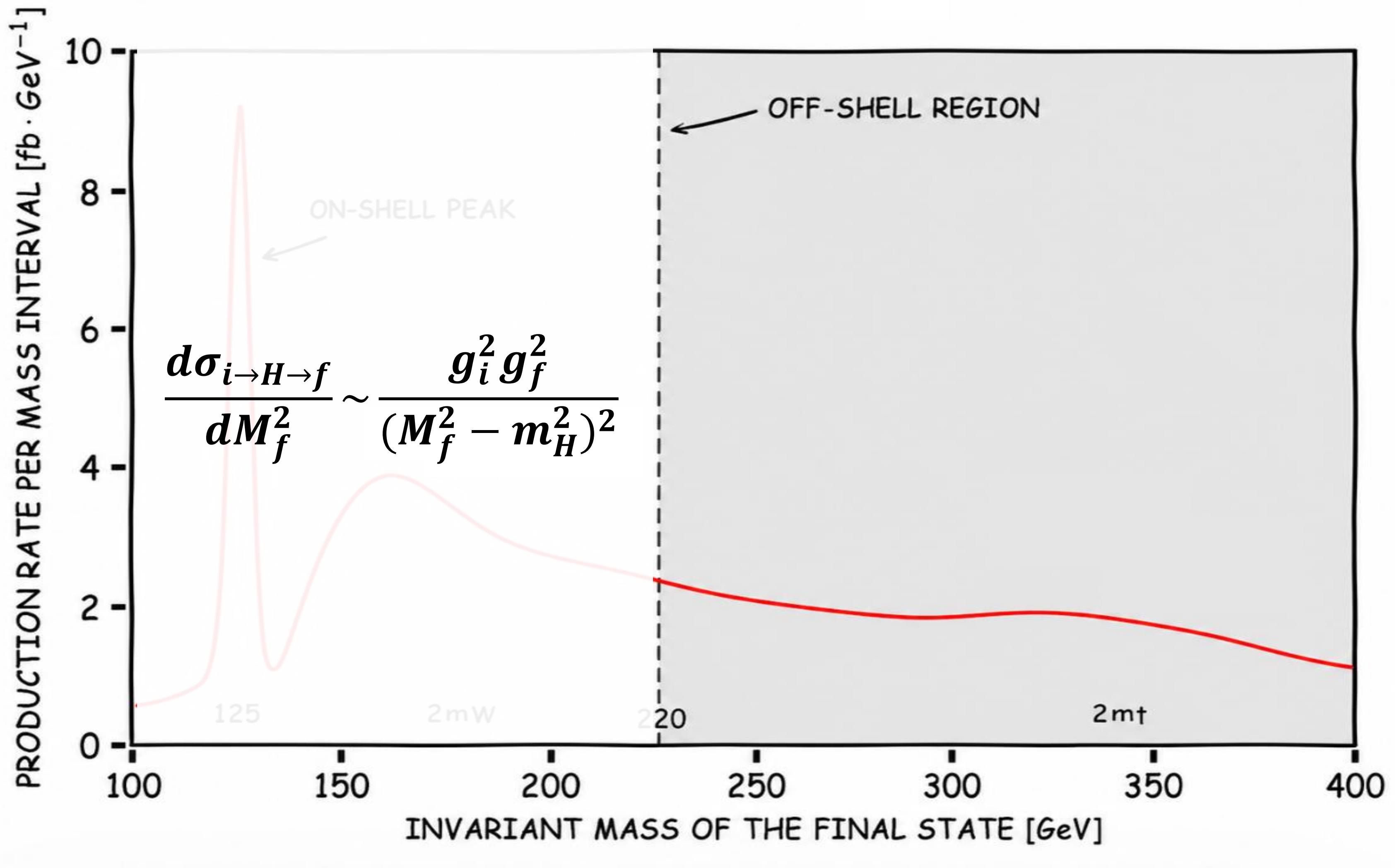
No direct measurement of Γ_H

Off-Shell Higgs



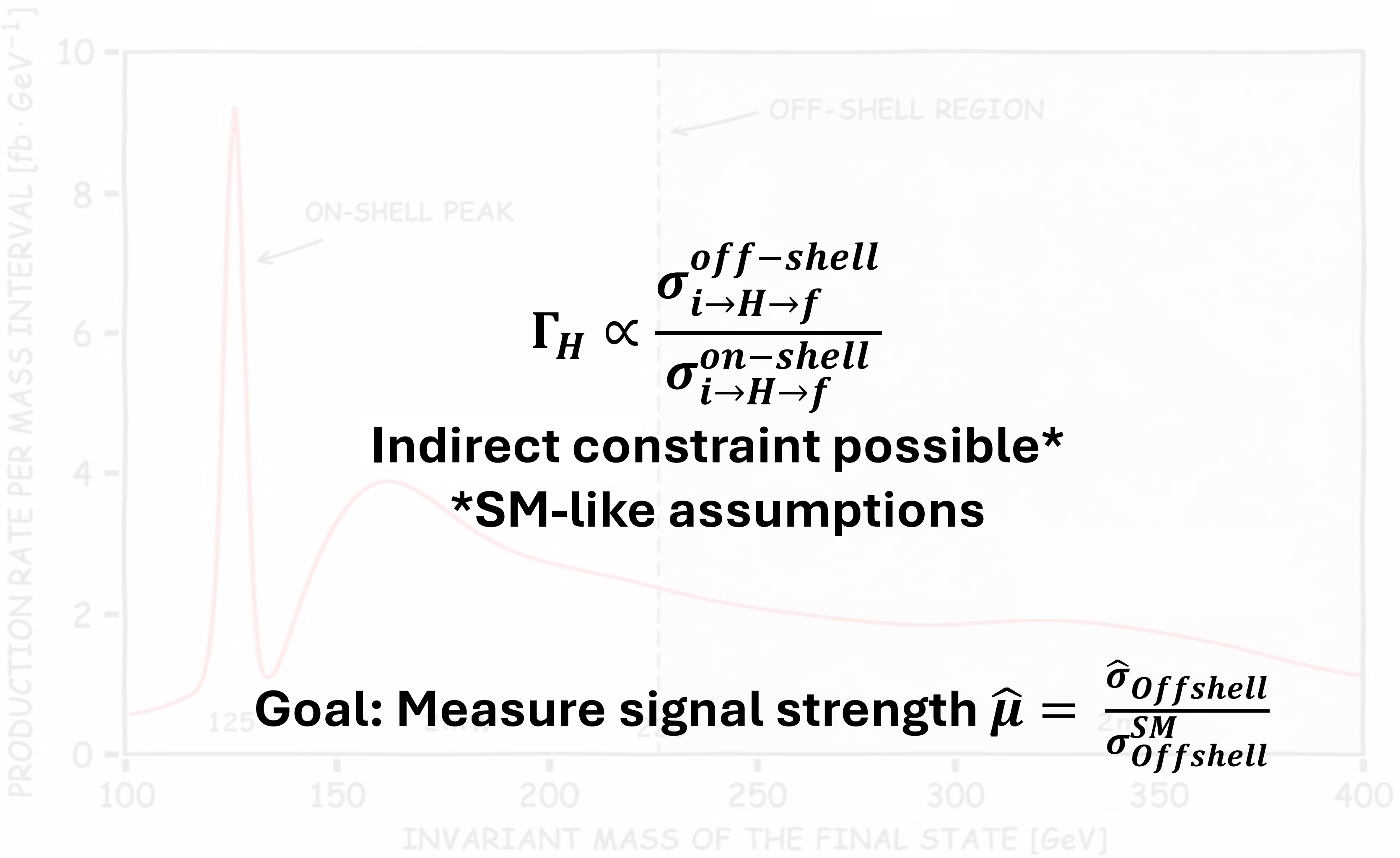
$$\frac{d\sigma_{i\to H\to f}}{dM_f^2} \sim \frac{g_i^2 g_f^2}{m_H^2 \Gamma_H^2}$$

Off-Shell Higgs



$$\frac{d\sigma_{i \to H \to f}}{dM_f^2} \sim \frac{g_i^2 g_f^2}{(M_f^2 - m_H^2)^2}$$

Off-Shell Higgs



$$\Gamma_H \propto \frac{\sigma_{i\to H\to f}^{off-shell}}{\sigma_{i\to H\to f}^{on-shell}}$$

Indirect constraint possible*

***SM-like assumptions**

Goal: Measure signal strength $\hat{\mu} = \frac{\hat{\sigma}_{Offshell}}{\sigma_{Offshell}^{SM}}$

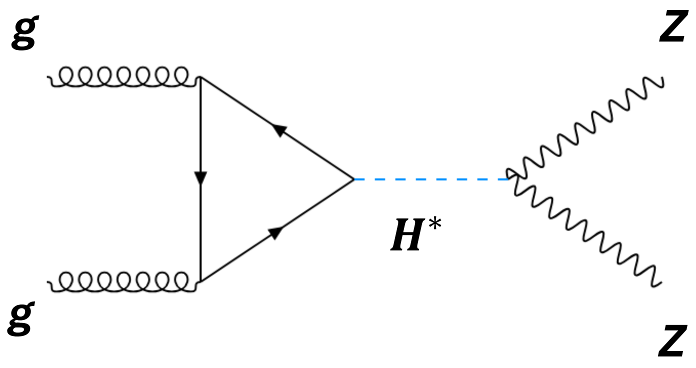
Off-Shell $H \to ZZ^*$ Production

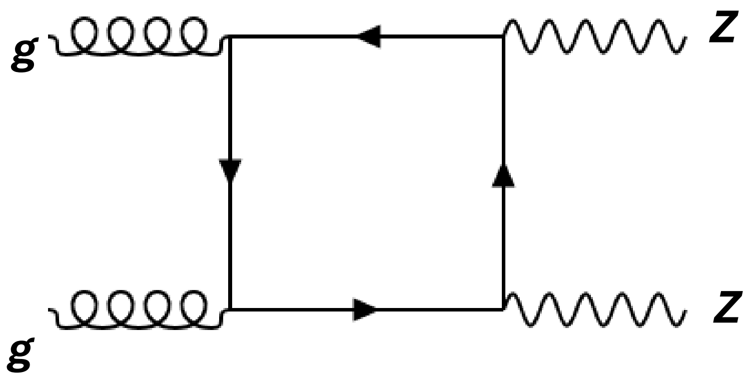
$H \to ZZ^{(*)} \to 4l$

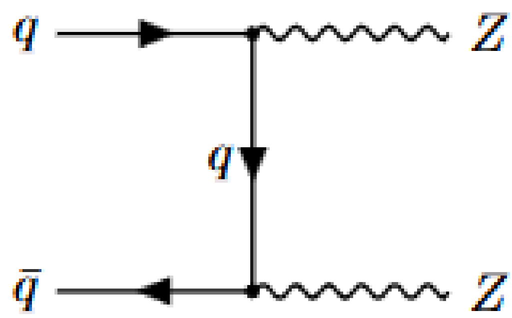
- Backgrounds: *qq→ZZ* (dominant), *gg → ZZ, VVV*
- Fully reconstructable final state
- Small branching ratio limits statistical power

$H \to ZZ^{(*)} \to 2l2v$

- Backgrounds: *qq→ZZ, ggZZ, WZ,* $t\bar{t}$*, Z+jets*
- Final state reconstruction challenging due to escaping neutrinos
- Much larger branching ratio



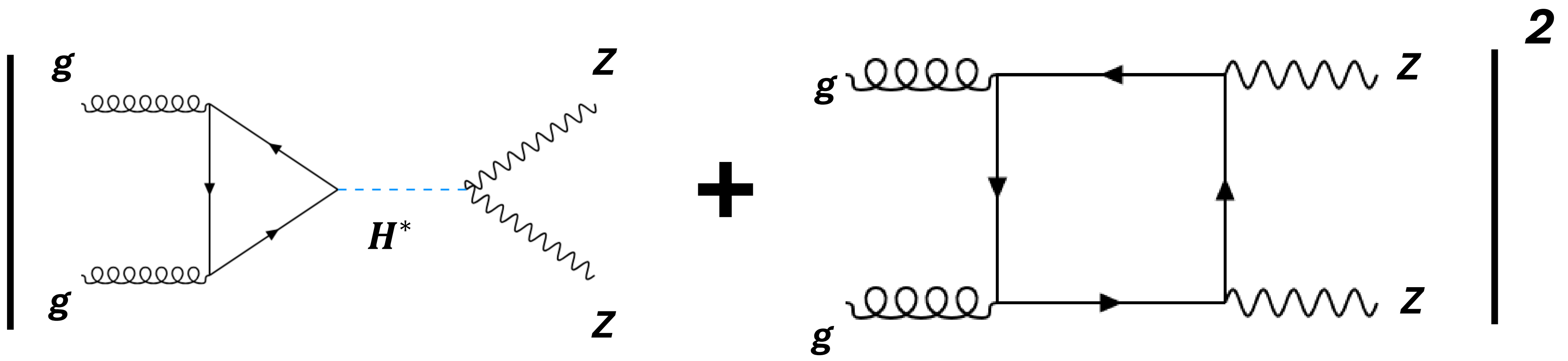




Off-Shell $H \rightarrow ZZ^*$ Production

Substantial interference between *gg→H→ZZ* Signal and *gg→ZZ* background

Signal can not be measured independently from background



Probability Density Model

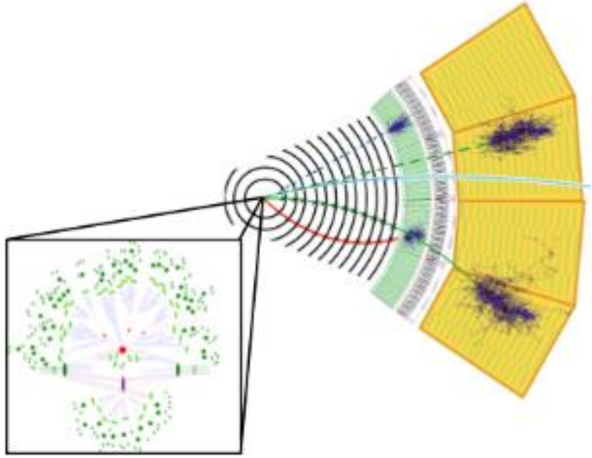
Estimate density p(x) → Deconstruct by contributing template processes

$$p(x|\mu) = \frac{1}{\nu(\mu)}\left[\mu \cdot \underbrace{v_S \cdot p_S(x)}_{\text{Signal}} + \sqrt{\mu} \cdot \underbrace{v_I \cdot p_I(x)}_{\text{Interference}} + \underbrace{v_B \cdot p_B(x)}_{\text{Background}} + \underbrace{v_{NI} \cdot p_{NI}(x)}_{\text{Non-Interfering Background}}\right]$$

Preform Likelihood Fit on this to extract $\hat{\mu}$

Histogram based approach to likelihood profiling

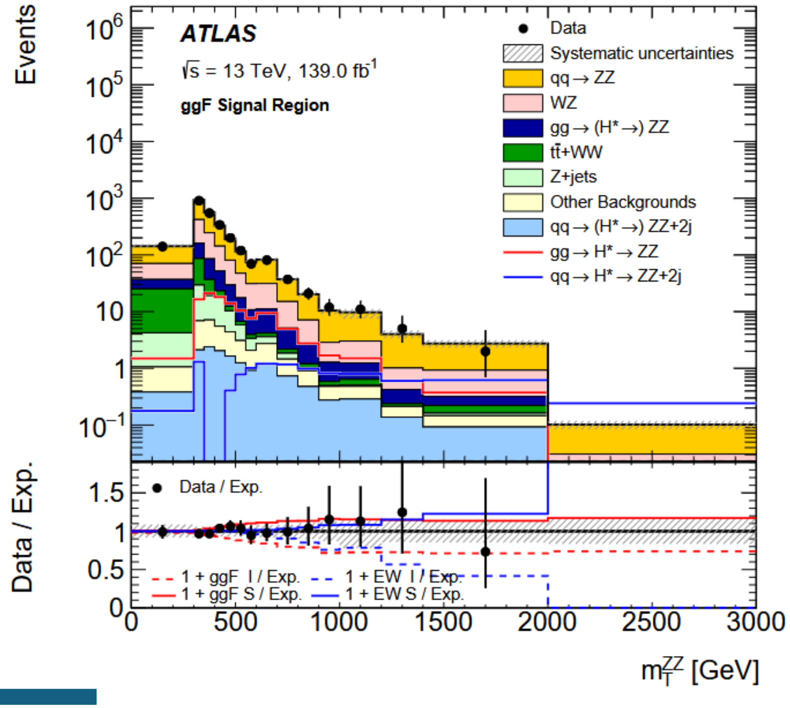
High dimensional Data



Low dimensional summary variable

$$m_T^{ZZ}$$

Histogram of summary variable



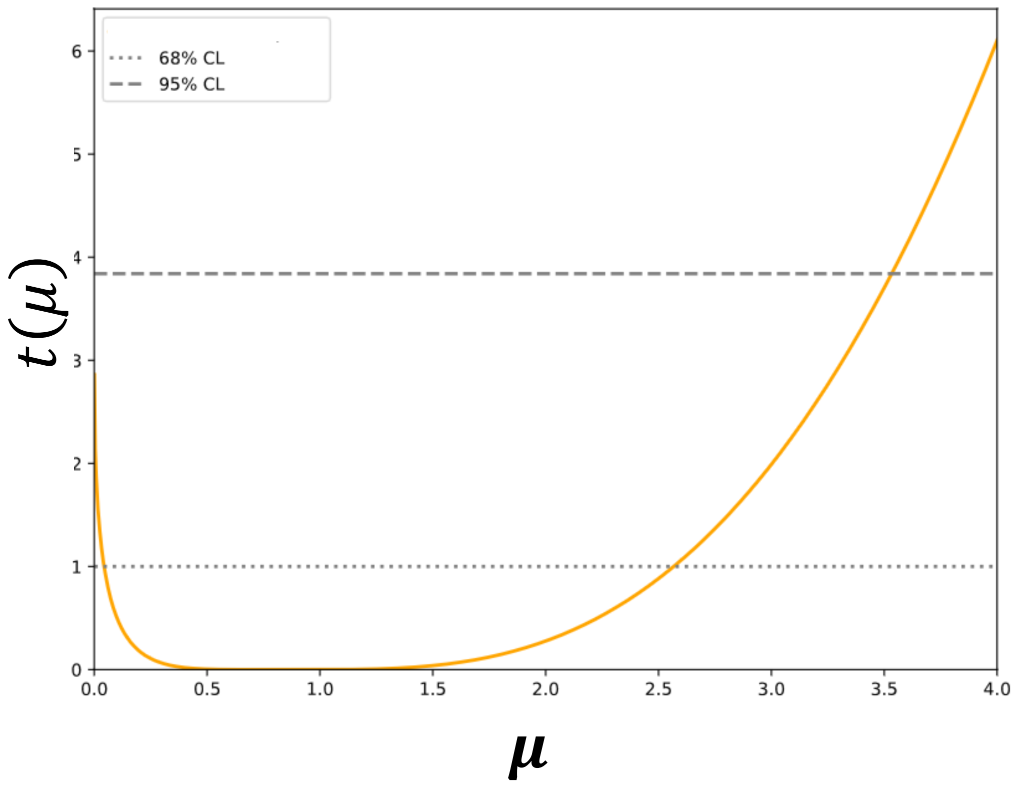
Poisson-based Likelihood

$$L(\mu) = \prod_{bins} Poisson\left(N_{obs} \middle| N_{exp}(\mu)\right)$$

Fit to data

$$\hat{\mu}$$

Test statistic scan



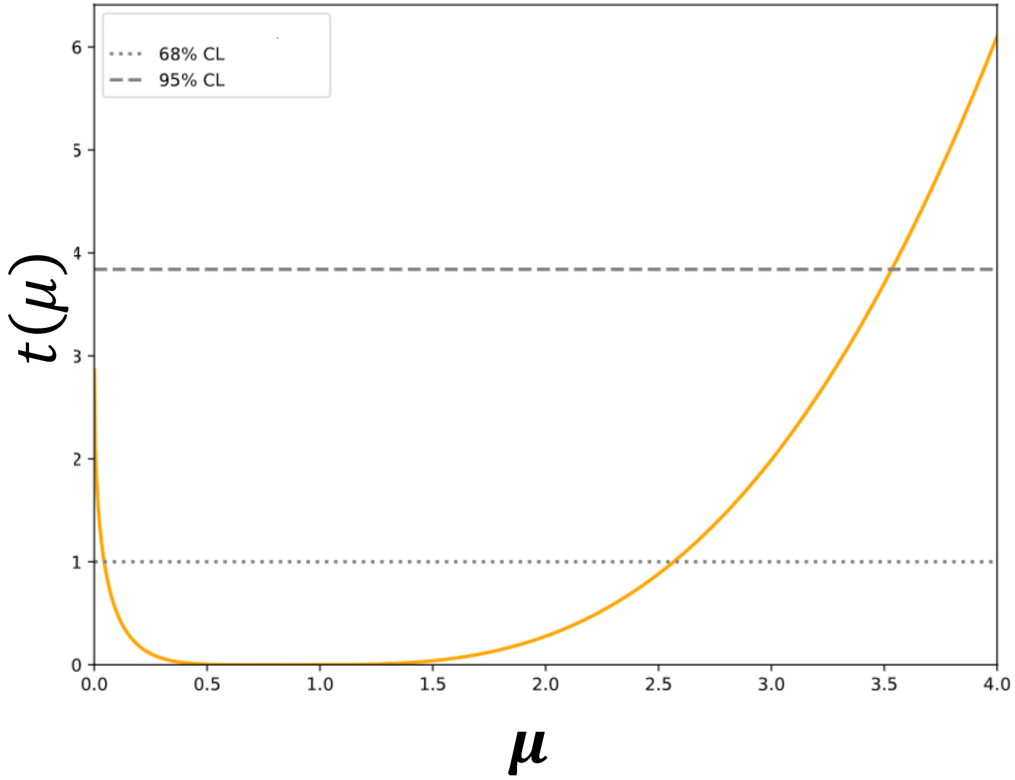
How "different" from the minimum is the Likelihood at other values of μ?

Test statistic

$$t(\mu) = -logL(\mu) - (-logL(\hat{\mu}))$$

"Sharper" $t(\mu)$ curve indicates tighter confidence interval in the end

Test statistic scan



Caveat of interference effect:

No single observable can offer optimal sensitivity at all values of the parameter of interest μ

Using only one summary variable histogram causes loss of sensitivity!

Optimally sensitive observables for every μ?

Neyman-Pearson-Lemma:

Probability densitiy ratio is the optimal observable to compare two hypothesis

Idea:

Reformulate probability density problem to probability density ratio problem

$$p(x|\mu) \rightarrow \frac{p(x|\mu)}{p_{ref}(x)}$$

„Arbitrary“ reference probability distribution

Probability Density Ratio Estimation

$$\frac{p(x|\mu)}{p_{ref}(x)} = \frac{1}{\nu(\mu)}[\underbrace{\mu \cdot v_S \cdot \frac{p_S(x)}{p_{ref}(x)}}_{\text{Signal}} + \underbrace{\sqrt{\mu} \cdot v_I \cdot \frac{p_I(x)}{p_{ref}(x)}}_{\text{Interference}} + \underbrace{v_B \cdot \frac{p_B(x)}{p_{ref}(x)}}_{\text{Background}} + \underbrace{v_{NI} \cdot \frac{p_{NI}(x)}{p_{ref}(x)}}_{\text{Non-Interfering Background}}]$$

$$L(x|\mu) = Pois(N^{Obs}|N^{Exp}(\mu)) \prod_i^N \frac{p(x_i|\mu)}{p_{ref}(x_i)}$$

Optimal observable construction event-by-event!

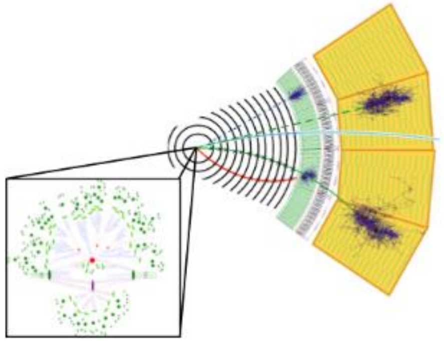
Total event count term

Event-By-Event Shape term

Probability Density Ratio Construction

Efficient probability density ratio estimators:
Neural Network classifiers

Monte Carlo Simulation



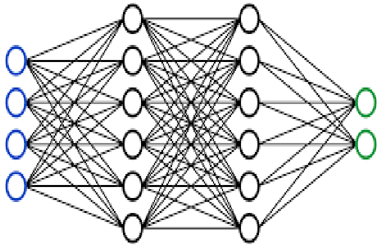
$\boldsymbol{p_i(x)}$

$\boldsymbol{p_{ref}(x)}$

Training Sample

$$\hat{x}$$

Neural Network Classifier

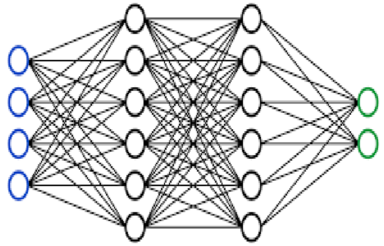


Neural Network Prediction

$$\hat{s} \approx \frac{p_i(x)}{p_i(x) + p_{ref}(x)}$$

Neural Simulation Based Inference

"Likelihood Ratio Trick"

$\hat{x}$ →  → $$\hat{s} \approx \frac{p_i(x)}{p_i(x) + p_{ref}(x)}$$ → $$\frac{p_i(x)}{p_{ref}(x)} = \frac{s}{1-s}$$

Easy direct construction of probability density ratios from trained Neural Network classifier!

„Neural Simulation-Based Inference" (NSBI)

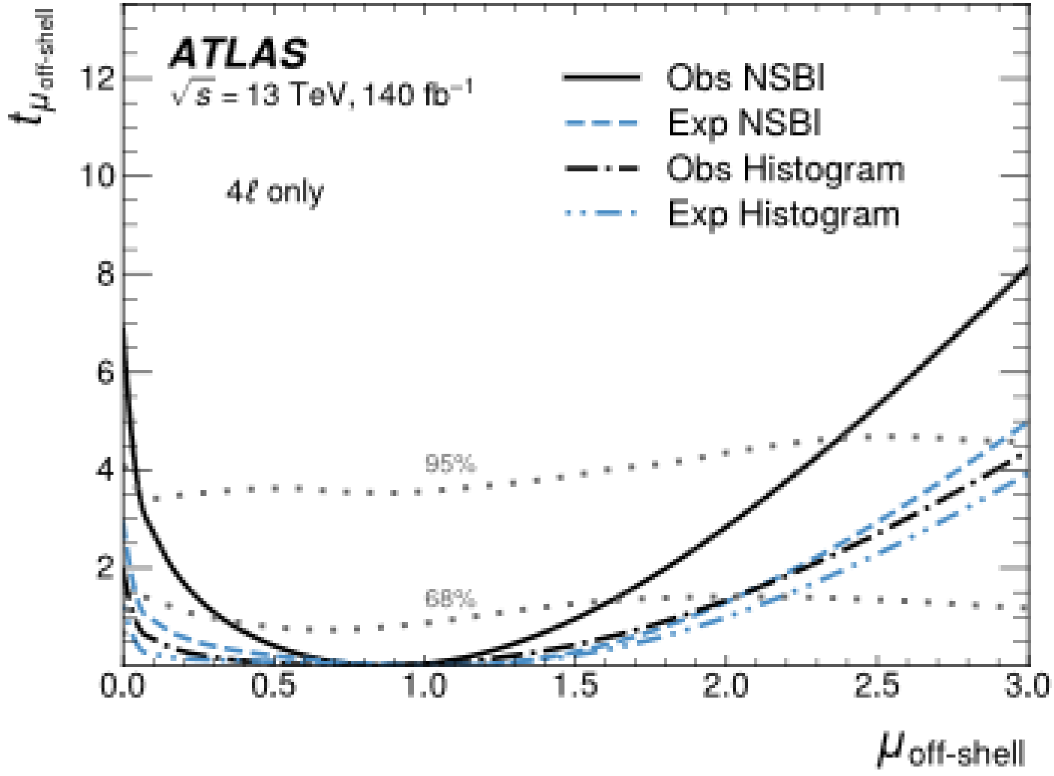
NSBI test statistic scan

$$t(\mu) = \boxed{-2log\left(\frac{P\left(n_k^{obs}|n_k^{exp}(\mu)\right)}{P\left(n_k^{obs}|n_k^{exp}(\hat{\mu})\right)}\right)}$$

Poisson Term

$$\boxed{-2\sum_{i=i}^{n_{obs}} log\frac{p(x_i|\mu)/p_{ref}(x_i)}{p(x_i|\hat{\mu})/p_{ref}(x_i)}}$$

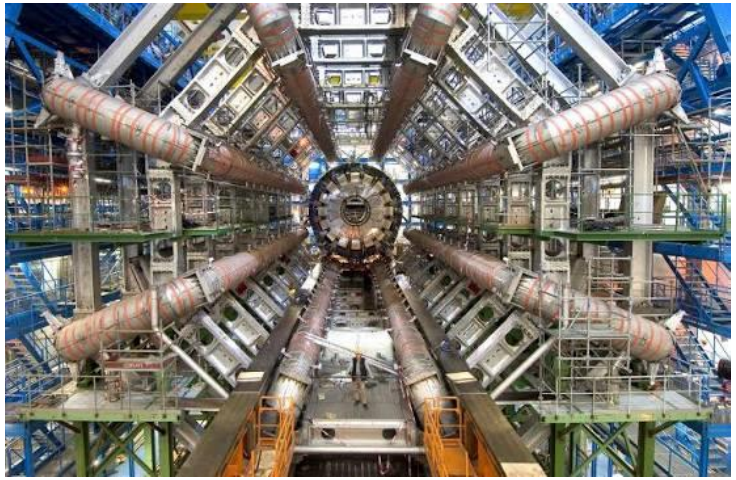
Event-By-Event Likelihood Ratios



https://iopscience.iop.org/article/10.1088/1361-6633/adcd9a/pdf

About 30% tighter confidence interval with NSBI compared to traditional histogram approach!

Run 3 ATLAS Off-Shell NSBI strategy



Full ATLAS Run 2 dataset at 13 TeV (2015 – 2018):

- $L \approx 140\ fb^{-1}$

Current ATLAS Run 3 dataset at 13.6 TeV (2022-2026):

- $L > 350\ fb^{-1}$

NSBI-based Off-Shell Higgs Production measurement with Run 3 data

Previously: NSBI only used in $H \to 4l$ channel

New: Extension to $H \to 2l2v$ channel for combined NSBI result

Studies on further improvements on the NSBI workflow to increase efficiency of the procedure

Conclusion

- **Neural Simulation-Based Inference offers a clean way to reformulate probability density estimation problems into probability density ratio estimation problems**

- **Processes with non-trivial probability density models, like Off-shell Higgs Boson production can benefit from this change in problem definition**

- **A combined NSBI-based Off-Shell Higgs production measurement in the $H \rightarrow ZZ^{*} \rightarrow 4l$ and $H \rightarrow ZZ^{*} \rightarrow 2l2v$ channels is being worked on for the ATLAS Run 3 dataset**