



# FAIR Universe HiggsML Uncertainty Challenge

*Wahid Bhimji, Paolo Calafiura, Ragansu Chakkappai, Po-Wen Chang, Yuan-Tang Chou, Sascha Diefenbacher, Jordan Dudley, Steven Farrell, Aishik Ghosh, Isabelle Guyon, Chris Harris, Shih-Chieh Hsu, Elham E Khoda, Benjamin Nachman, Peter Nugent, David Rousseau, Benjamin Thorne, Ihsan Ullah, Yulei Zhang*

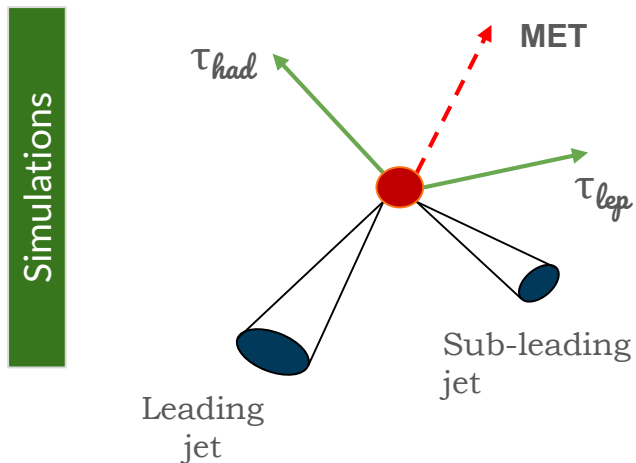
European AI for Fundamental Physics Conference  
2025 - Cagliari, Sardinia Italy

université  
PARIS-SACLAY

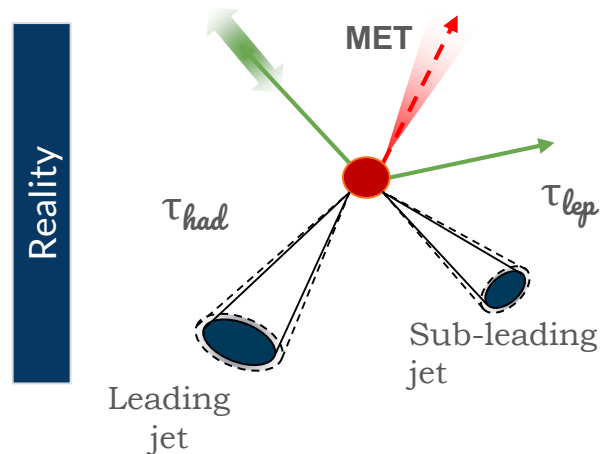


# Bias and uncertainty in Higgs Physics

At the LHC we analyse proton collision with specific final state Eg. evidence for the higgs boson



Search for :  $H \rightarrow \tau^+ (had, \nu) \tau^- (l, \nu, \nu)$



- There are possible biases
- Eg. In calibration of particle energy

Many methods developed, but **NO Common Benchmark**

# Fair Universe: HiggsML Uncertainty Challenge



## FAIR UNIVERSE - HIGGS UNCERTAINTY CHALLENGE

160

PARTICIPANTS

347

SUBMISSIONS



A pool of 4000 USD

<https://www.codabench.org/competitions/2977/>

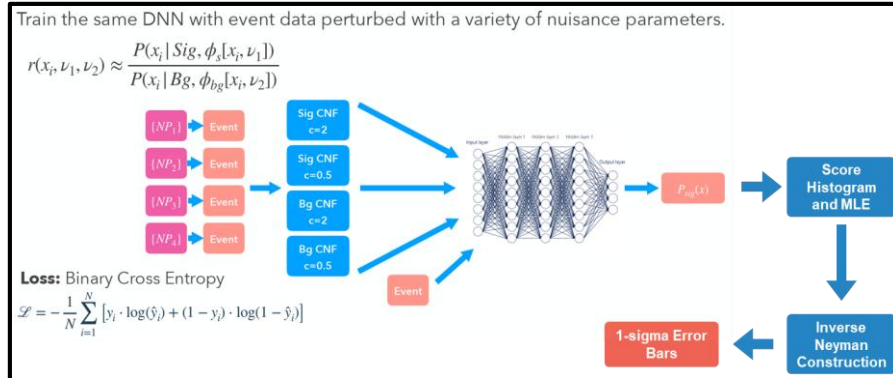
- Ran from September 12 to March 14th
- Accepted as NeurIPS competition 2024
- Dedicated workshop at NeurIPS - 2024
- Final results presented last month at CERN (satellite event of IML - 2025)

- **NEW** Simulated dataset -220M events, 28 features: <https://zenodo.org/records/15131565>
- **6 Parameterized systematics**
- **Metric** : Signal strength ( $\mu$ ) and Confidence Interval
- **Inference** : With pseudo-experiment

# Winners

## Contrastive Normalizing Flow (CNF) - Ibrahim

- Trains CNFs on augmented data
- Trained CNFs to produces bias-free variables
- New Variables Used to Train the DNN



## Unbinned measurements with refinable systematics - HEPHY

- Based on tradition histogram analysis in LHC
- Build ML surrogate for estimating systematics
- Uses these ML surrogates for  $\mu$  estimation

$$d\sigma(x|\mu, \nu) = \mu d\sigma_H(x|\nu_{\text{calib}}) + (1 + \alpha_{\text{bkg}})^{\nu_{\text{bkg}}} d\sigma_Z(x|\nu_{\text{calib}}) + (1 + \alpha_{\text{bkg}})^{\nu_{\text{bkg}}} (1 + \alpha_{\text{tt}})^{\nu_{\text{tt}}} d\sigma_{\text{tt}}(x|\nu_{\text{calib}}) + (1 + \alpha_{\text{bkg}})^{\nu_{\text{bkg}}} (1 + \alpha_{\text{VV}})^{\nu_{\text{VV}}} d\sigma_{\text{VV}}(x|\nu_{\text{calib}})$$

NB1: Still valid measure  
NB2: The additive ansatz facilitates to refinable modeling.

$$d\sigma_H(x|\nu_{\text{calib}}) = \frac{d\sigma_H(x|\nu_{\text{calib}})}{d\sigma(x|1, 0)} d\sigma(x|1, 0) = \frac{d\sigma_H(x|\nu_{\text{calib}})}{d\sigma_H(x|0)} \frac{d\sigma_H(x|0)}{d\sigma(x|1, 0)} d\sigma(x|1, 0) \approx \frac{\hat{S}_H(x|\nu_{\text{calib}})}{\hat{S}_H(x|0)} \hat{g}_H(x) d\sigma(x|1, 0)$$

Form ratio wrt total nominal (a choice)      model systematics per-process (another choice)      ML surrogate for systematics      Likelihood-ratio trick

$$\text{DCR} = \frac{d\sigma(x|\mu, \nu)}{d\sigma(x|1, 0)} \approx \mu \hat{g}_H(x) \hat{S}_H(x|\nu_{\text{calib}}) + (1 + \alpha_{\text{bkg}})^{\nu_{\text{bkg}}} (\hat{g}_Z(x) \hat{S}_Z(x|\nu_{\text{calib}}) + (1 + \alpha_{\text{tt}})^{\nu_{\text{tt}}} \hat{g}_{\text{tt}}(x) \hat{S}_{\text{tt}}(x|\nu_{\text{calib}}) + (1 + \alpha_{\text{VV}})^{\nu_{\text{VV}}} \hat{g}_{\text{VV}}(x) \hat{S}_{\text{VV}}(x|\nu_{\text{calib}})) \equiv \hat{R}(x|\mu, \nu) = \text{surrogate}$$



**Thank you for your  
attention!**

Checkout our at **Poster  
Session A ID 26**

