# Back to the Formula
# -
# LHC Edition

A. Butter, T. Plehn, N. Soybelman and J. Brehmer
arXiv: 2109.10414

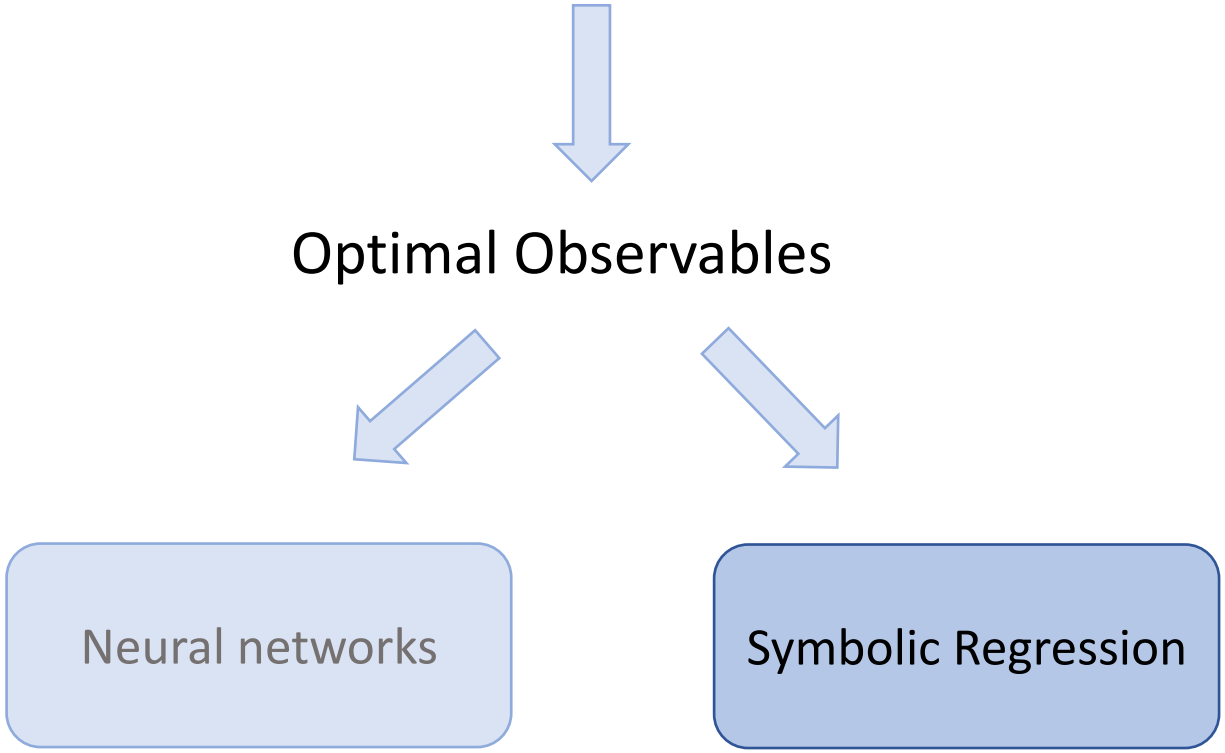
Nathalie Soybelman
ITP, Universität Heidelberg

# Content

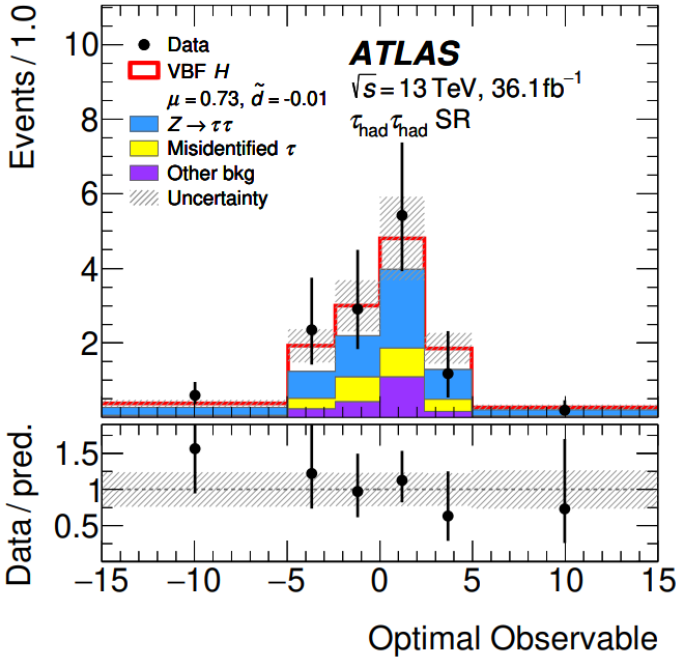
- Motivation
- Optimal observable
- CP violation in WBF
- Results

# Motivation

New techniques for improved LHC analysis



# Optimal observable




- Best observable to measure a parameter
- First introduced in:

  D. Atwood, A Soni: Phys. Rev. D45 (1992) 2405
  M. Diehl, O. Nachtmann: Z. Phys. C 62 (1994) 397-412
  M. Davier et al.: Phy. Lett. B306 (1993) 411

- Traditionally introduced for processes in e+e- colliders
- Used for instance in ATLAS analysis arXiv 2002.05315

# Optimal observable

Traditional optimal observables numericaly involved

→ More general approach by Brehmer et al. arXiv 1805.00020 (MadMiner)

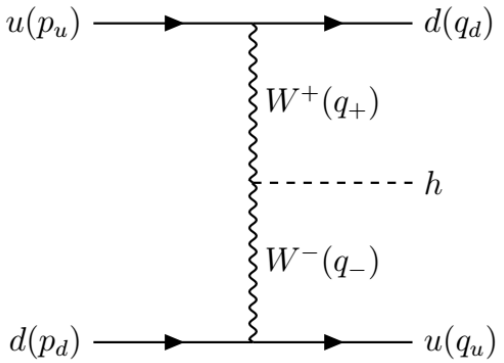
Likelihood $$p(z|\theta) = \frac{1}{\sigma(\theta)} \frac{\mathrm{d}\sigma(z|\theta)}{\mathrm{d}z}$$

**Score** $$t(z|\theta) = \nabla_\theta \log p(z|\theta) = \frac{\nabla_\theta \mathrm{d}\sigma(z|\theta)}{\mathrm{d}\sigma(z|\theta)} - \frac{\nabla_\theta \sigma(\theta)}{\sigma(\theta)} \sim \frac{\nabla_\theta |\mathcal{M}|^2}{|\mathcal{M}|^2}$$

$z$ contains latent variables (parton shower, detector effects, etc.)

Marginalize over latent variables → Optimal observable

# CP violation in weak boson fusion




Hankele et al. [arXiv: hep-ph/0609075]

Brehmer et al. [arXiv: 1712.02350]

Parameter $\theta = f_{W\widetilde{W}}$

Operator $\mathcal{O}_{W\widetilde{W}} = -(\phi^\dagger\phi)\widetilde{W}^k_{\mu\nu}W^{\mu\nu k}$

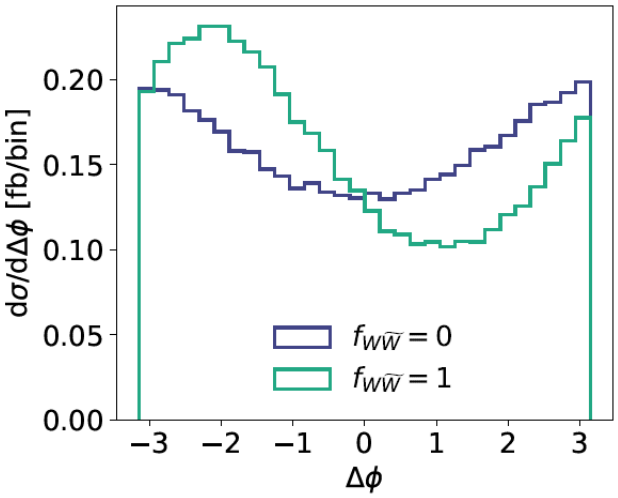
Matrix element $|\mathcal{M}|^2 = p_0 + a\theta + b\theta^2$

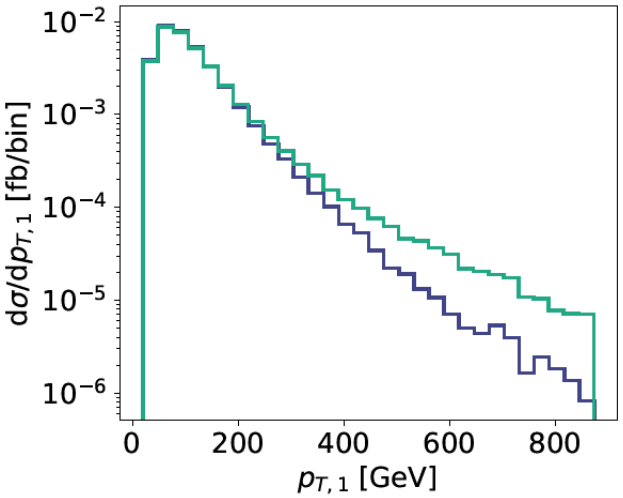
Score $$t(z|\theta) \sim \frac{\nabla_\theta|\mathcal{M}|^2}{|\mathcal{M}|^2} = \frac{a+2b\theta}{p_0+a\theta+b\theta^2} \sim \frac{a}{p_0} + \frac{1}{p_0}\left(2b - \frac{a^2}{p_0}\right)\theta$$

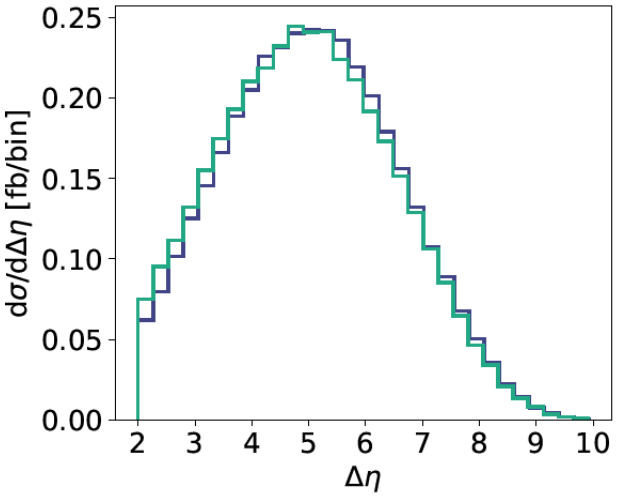
CP observable $\epsilon_{\mu\nu\rho\sigma}p_u^\mu p_d^\nu q_u^\sigma q_d^\rho \sim p_{T1}p_{T2}\sin\Delta\phi$

# Optimal observable for $f_{W\widetilde{W}} = 0$

Event distribution



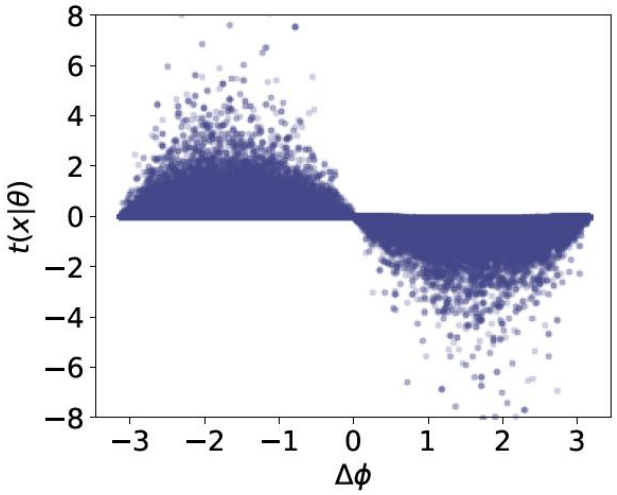



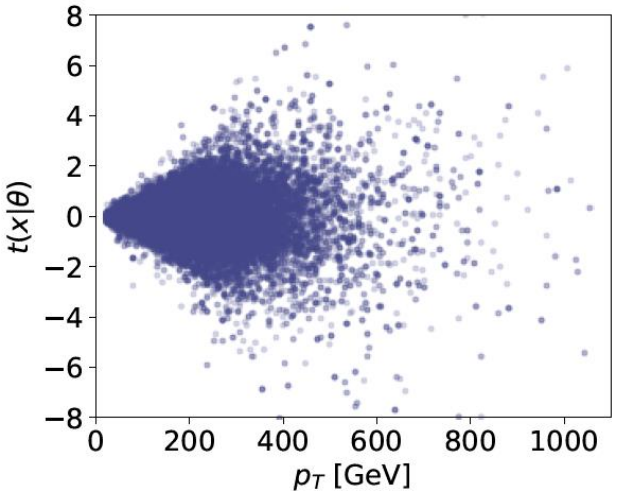

Relevant observables

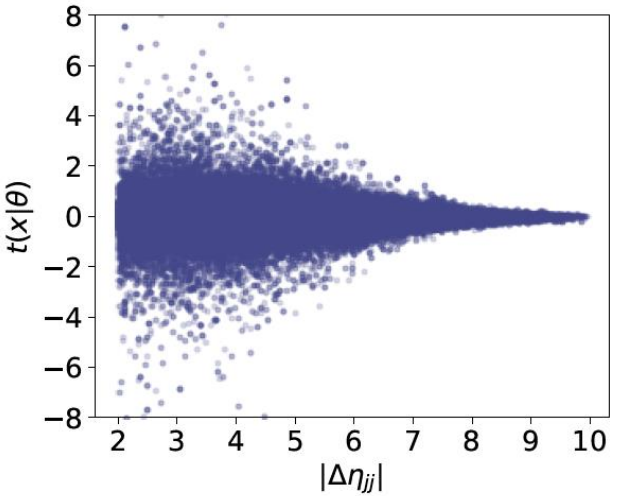
$\Delta\phi,\ p_{T,1},\ p_{T,2},\ \Delta\eta$

# Optimal observable for $f_{W\widetilde{W}} = 0$

Score



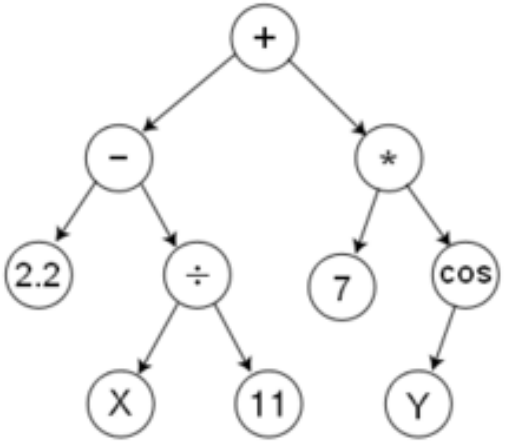




Relevant observables

$\Delta\phi,\ p_{T,1},\ p_{T,2},\ \Delta\eta$

# Symbolic regression




- Genetic algorithm implemented in PySR

  M. Cranmer et al. arXiv 2006.11287

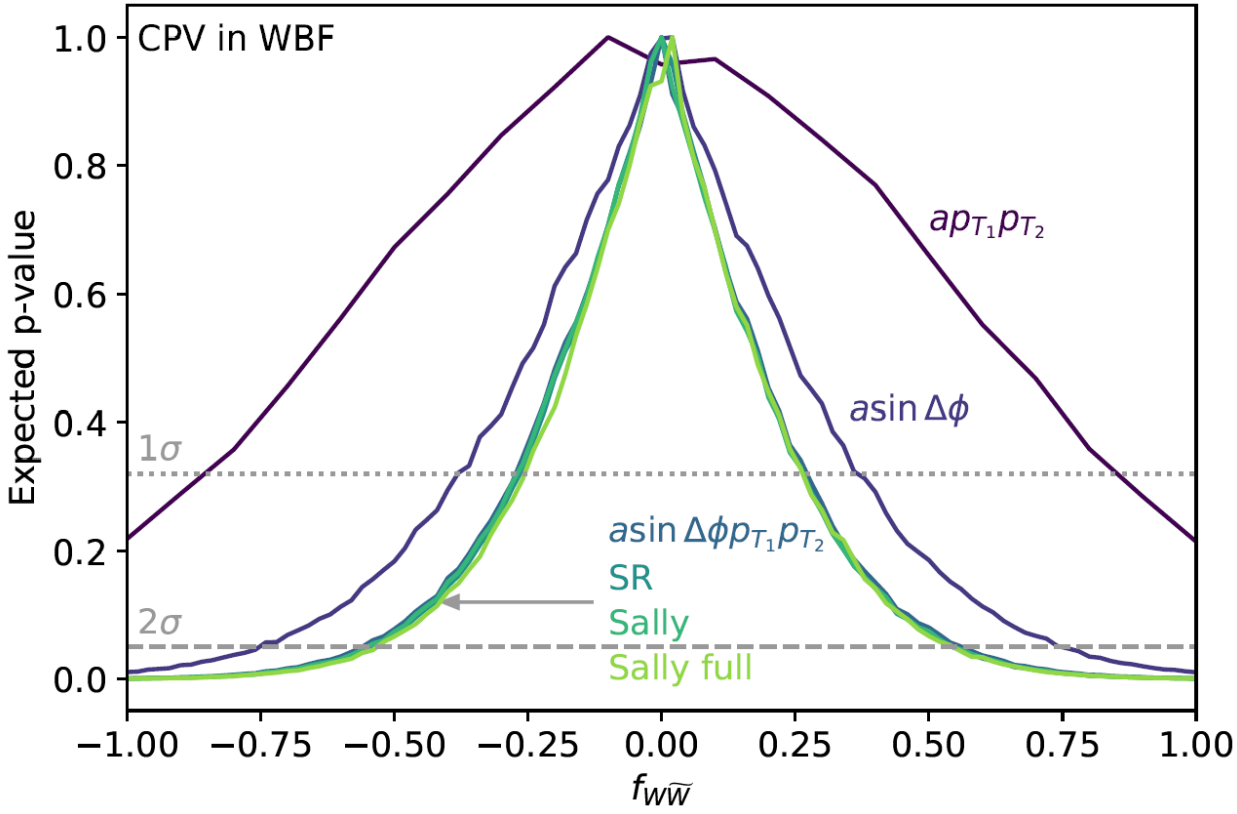
- Input: operators and training data

- Output: hall of fame - best formulas for each complexity

# nodes

| cmpl. | function | MSE |
|---|---|---|
| 4 | $a\sin(\Delta\phi)$ | $1.03 \cdot 10^{-1}$ |
| 8 | $a\sin(\Delta\phi)p_{Tj1}p_{Tj2}$ | $1.49 \cdot 10^{-2}$ |
| 16 | $-p_{Tj1}(a - b\Delta\eta)(p_{Tj2} + c)\sin(\Delta\phi + d)$ | $8.50 \cdot 10^{-3}$ |
| 28 | $(p_{Tj2} + a)(bp_{Tj1}(c - \Delta\phi) - p_{Tj1}(d\Delta\eta + ep_{Tj2} + f)\sin(\Delta\phi + g))$ | $8.18 \cdot 10^{-3}$ |

# Results




Sally: NN using 4 observables

Sally full: NN using full event info

Symbolic regression formula comparable to NN

# Summary

- Optimal observables first introduced for e+e- colliders
- OO evaluation possible numerically or analytically
- Symbolic regression provides interpretable formulas