

UPPSALA  
UNIVERSITET

Unfolding of multivariate tools and statistical analysis for Higgs boson pair production searches in the ATLAS detector at the Large Hadron Collider

Christina Dimitriadi  
*Partikeldagarna 2019*  
*Linköping*



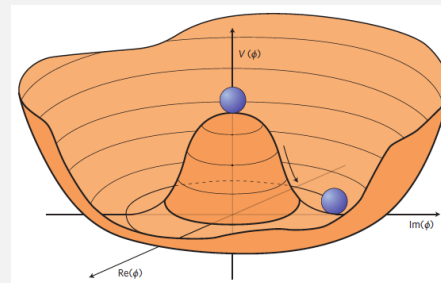
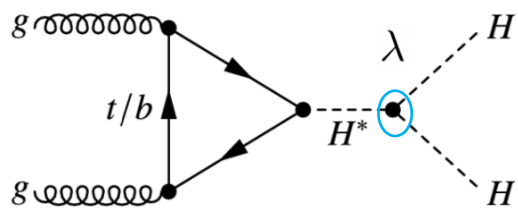
# Motivation

## Why di-Higgs?

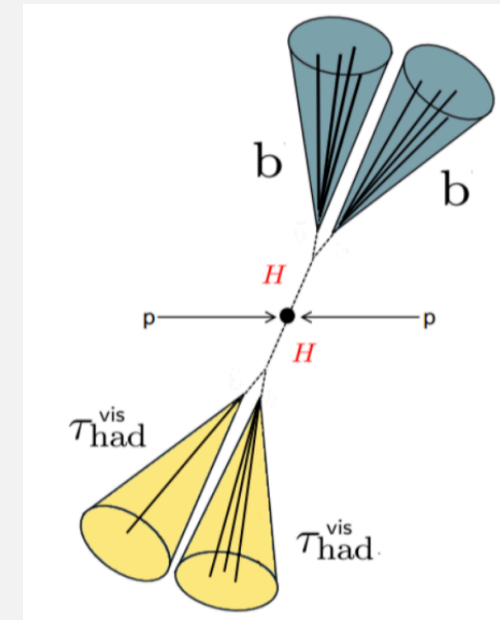


- ❖ The Higgs boson has the unique ability to couple to itself, resulting in the simultaneous production of two Higgs bosons
- ❖ The Higgs self-coupling is promptly linked to the Higgs potential, whose actual shape has not been measured experimentally yet

$$V = V_0 + \lambda v^2 H^2 + \lambda v H^3 + \frac{\lambda}{4} H^4$$



- This study uses the results of the search for SM pair production of Higgs bosons in the  $b\bar{b}\tau^+\tau^-$  final state, where both  $\tau$ -leptons decay in a hadronic mode ( $\tau_{had}\tau_{had}$ ).

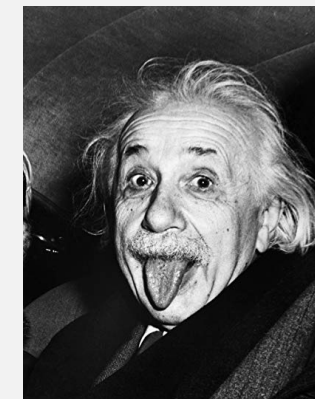
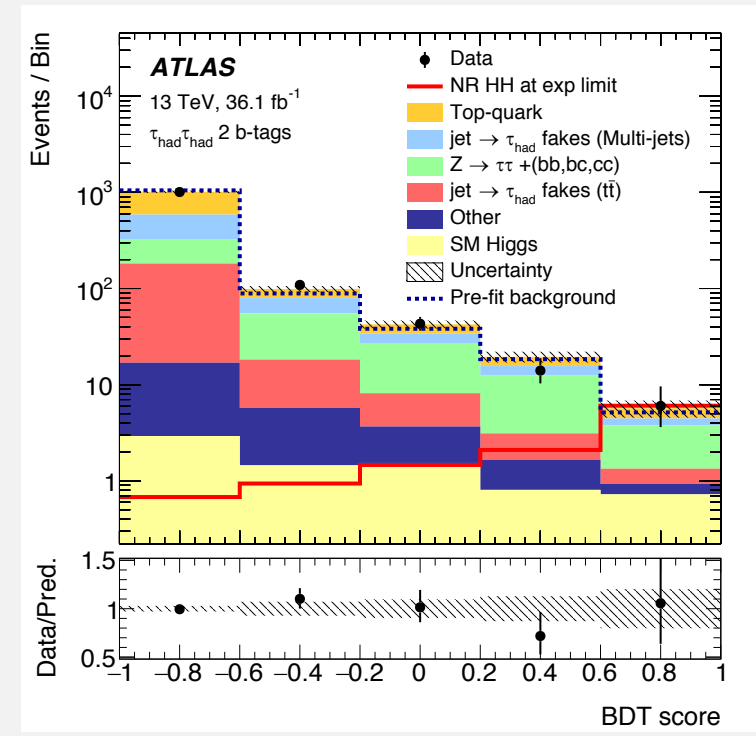




# Purpose of study



- Rescoping of the multivariate analysis (MVA) ([Phys. Rev. Lett. 121, 191801 \(2018\)](#)) by introducing a cut-based analysis.
- MVA techniques are widely used in High Energy Physics, where there is an ever-increasing production of large data, which typically involve multiple variables including their correlations.
- But is MVA **always** the best option? What is the gain by using MVA? Can a cut-based analysis perform equally well?





Looking forward to answering your questions in the  
Poster Session later 😊



**Unfolding of multivariate tools and statistical analysis for Higgs boson pair production searches in the ATLAS detector at the Large Hadron Collider**  
Ch. Dimitriadi, A. Ferrari, P. Bokan

**Motivation**  
Searches for Higgs boson pair production provide insight into the Higgs boson self-coupling, promptly linked to the shape of the Higgs potential in and beyond the Standard Model (SM).

**METHODS**

- Look at the distributions of the BDT input variables
- Get expected upper limits on the  $HH$  production cross section (without systematic uncertainties)
- Fit the BDT score  $\Rightarrow$  exp.:  $15.2 \times \sigma_{SM}$  @ 95% CL
- Fit  $m_{HH}$  in signal (SR) and control regions (CR) based on the last 2 bins of the BDT score  $\Rightarrow$  exp.:  $21.7 \times \sigma_{SM}$  @ 95% CL

**SM Higgs boson pair production at the LHC**

**Higgs boson pair decay processes**  
Decay channel of interest for this study:  
 $HH \rightarrow b\bar{b}\tau^+\tau^- \rightarrow b\bar{b}\tau_{had}\tau_{had}$

- relatively clean final state
- relatively large branching ratio (7.4%)
- both  $\tau$ -leptons decay hadronically

**Purpose of study**  
Rescoping of multivariate analysis (MVA) by introducing a cut-based analysis, so that the gain of using the MVA can be estimated.

**MVA - Boosted decision tree (BDT)**  
 $\Rightarrow$  to separate signal from background

**Cut-based analysis**  
Definition of variables associating the mass and  $\Delta R$  of di- $\tau$  and di- $b$  objects

$$X_{m_{bb}, m_{\tau\tau}} = \sqrt{\left(\frac{m_{bb} - 116 \text{ GeV}}{0.1 \times m_{bb}}\right)^2 + \left(\frac{m_{\tau\tau} - 112 \text{ GeV}}{0.1 \times m_{\tau\tau}}\right)^2}$$

$$X_{\Delta R_{bb}, \Delta R_{\tau\tau}} = \sqrt{\left(\frac{\Delta R_{bb} - 1.229}{0.1 \times \Delta R_{bb}}\right)^2 + \left(\frac{\Delta R_{\tau\tau} - 1.227}{0.1 \times \Delta R_{\tau\tau}}\right)^2}$$

**Optimal cuts on these variables are found, so that most of the background is removed while a large fraction of the signal survives. The aim is to find the combination of cuts giving the highest significance.**

**Expected upper limits on the  $HH$  production cross section (without systematic uncertainties) @ 95% CL, when fitting  $m_{HH}$  in SRs and CRs based on the optimal combination of cuts**

Cut 1	$\sigma(HH \rightarrow b\bar{b}\tau\tau)$ [pb]	-2 $\sigma$	-1 $\sigma$	Expected	+1 $\sigma$	+2 $\sigma$
		0.42	0.56	0.78	1.08	1.45
	$\sigma/\sigma_{SM}$	12.5	16.8	<b>23.4</b>	32.4	43.4

**Conclusion**  
Neither using the BDT nor fitting its score distribution worsens the sensitivity to  $HH \rightarrow b\bar{b}\tau\tau$  by 50%. This is a logical outcome since the BDT is by construction the best final discriminant and has the best separation. However, it was of great importance to suggest a cut-based analysis that could be easier re-interpreted by theorists compared to the BDT analysis.

Phys. Rev. Lett. 121, 191801 (2018)

Partikeldagarna 2019  
Linköping