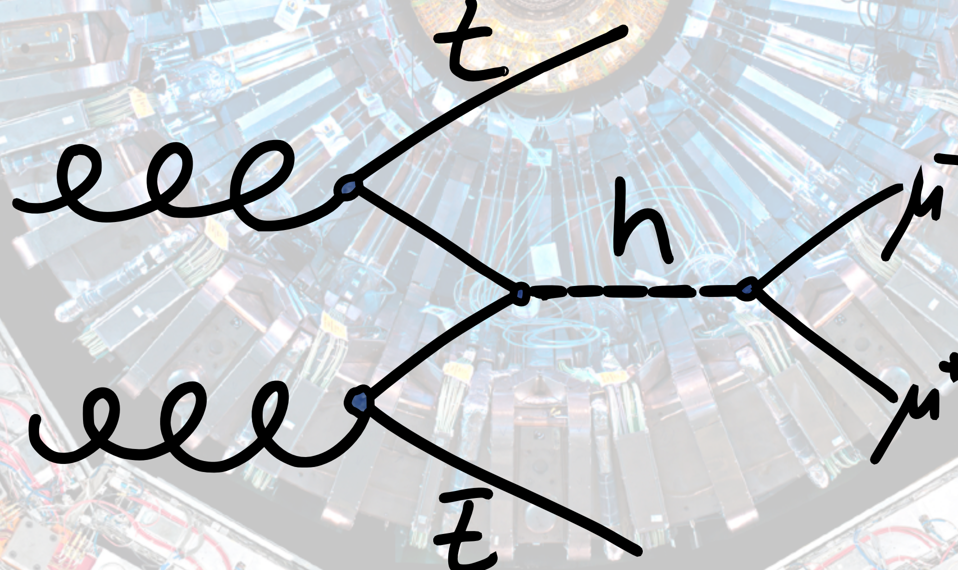
Search for SM Higgs Decays to Muons with the CMS Experiment on full Run2 data

-

Higgs in the associated Production with a Pair of Top Quarks



Tobias Kramer, Torben Lange, **Oliver Rieger**, Peter Schleper

Quantum Universe H1+H2 meeting, 17.12.2019

Outline

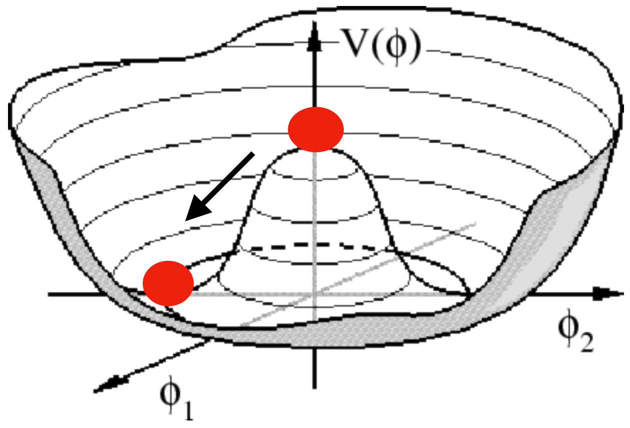
1. **Motivation & Introduction**
2. **Analysis Strategy for ttH → μμ**
   A. **Preselection**
   B. **Category definition -
ttH leptonic & ttH hadronic**
   C. **BDT optimization & categorization**
   D. **Limit setting**
3. **Conclusion - Estimate on sensitivity**

Motivation

Higgs mechanism

- Higgs observed $m_H = 125.09$ GeV
- Introduces SU(2) doublet
- Non zero VEV
- → Particles acquire mass

- Higgs coupling proportional to particle mass



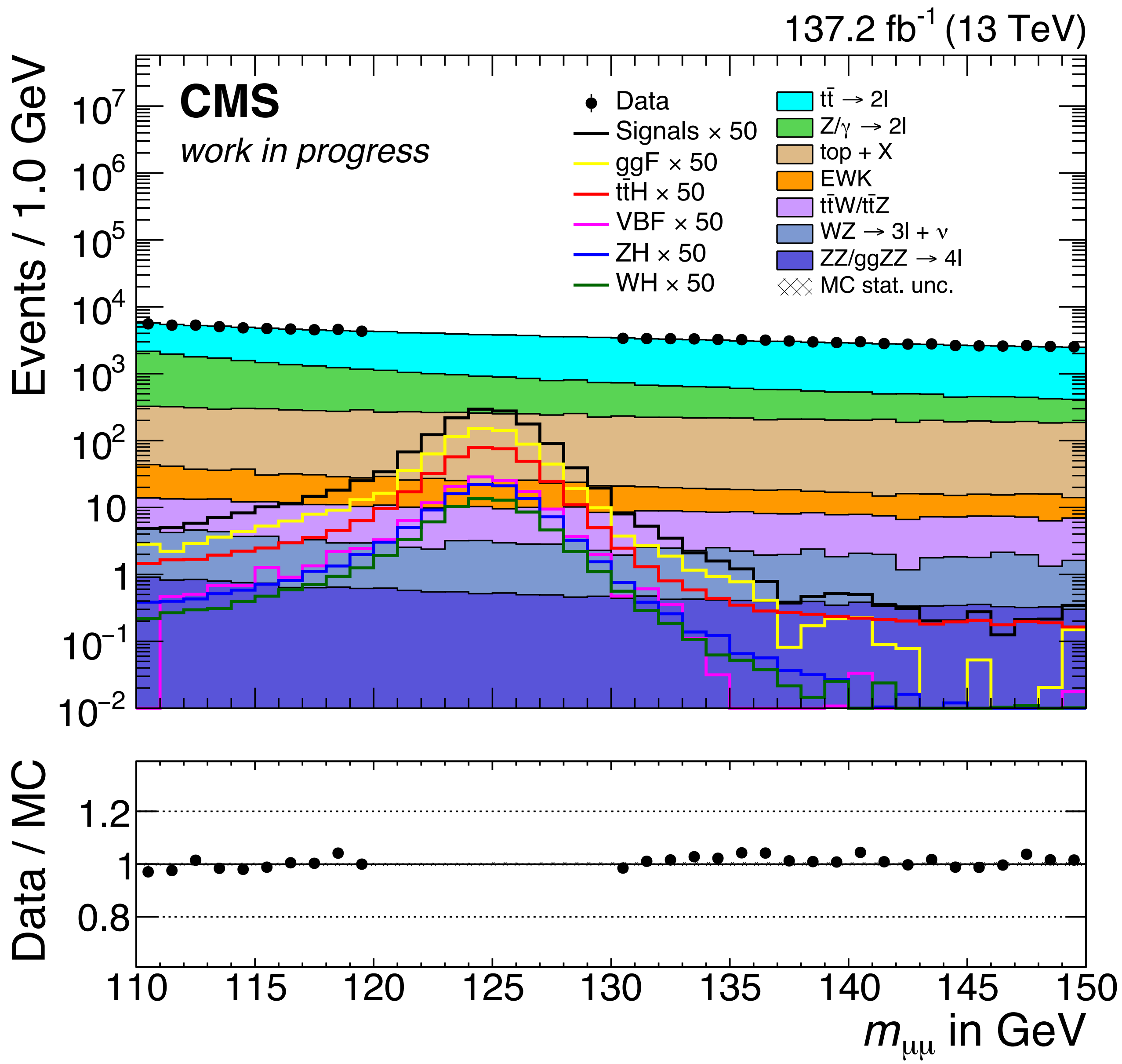
Higgs coupling to 2nd generation

- Higgs coupling to 2nd generation not observed yet
- Differences compared to SM predictions can point to physics beyond the SM

Introduction

Main challenges

- ttH has the smallest production rate among all considered Higgs channels
- Large backgrounds from dileptonic top quark pair production and Drell Yan with fake or non-prompt leptons or ISR jets
- Irreducible background from ttZ events



Preselection with 2 muons and b-tagged jets

Full signal + sideband region
110 GeV < m(μμ) < 150 GeV

Dileptonic top quark pair production

Drell Yan

Top + X: tZq, semileptonic TTbar, SingleTop production, MultiTop production

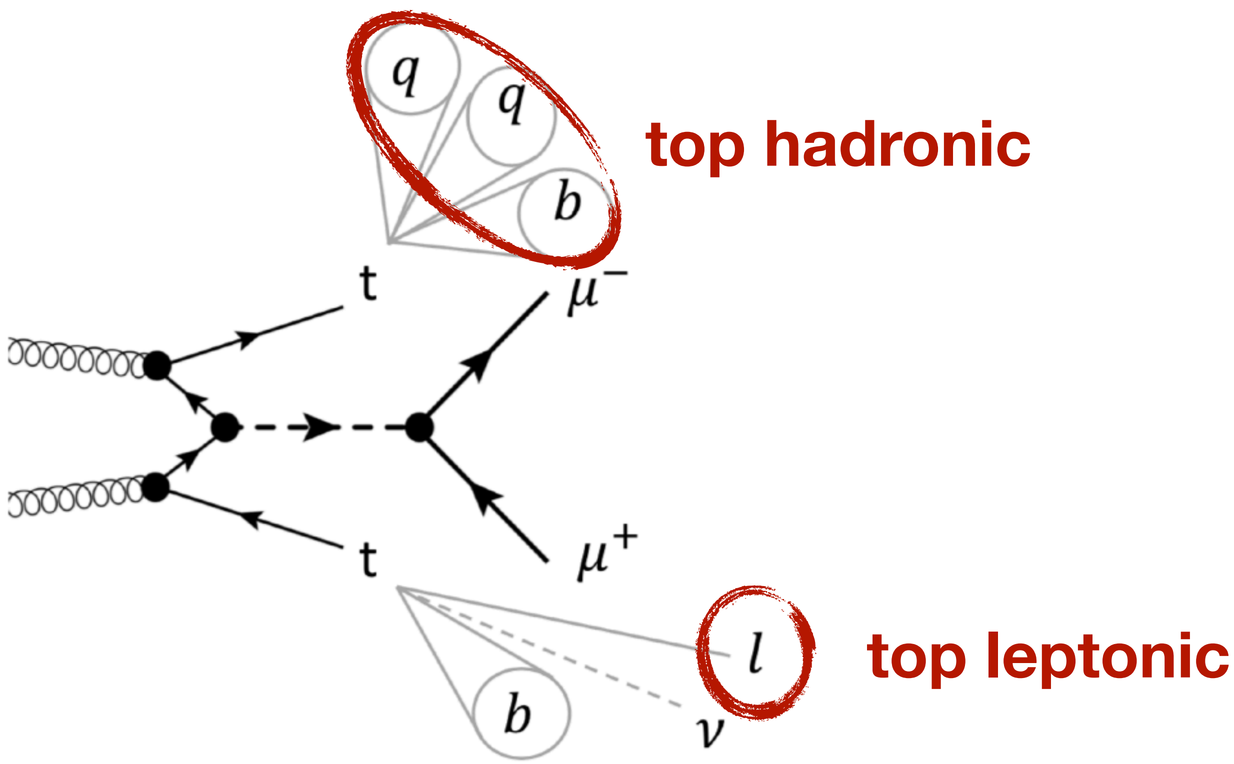
EWK: MultiBoson VVV, VV

ttV: ttZ, ttW, ttWW

Preselection and Categorization

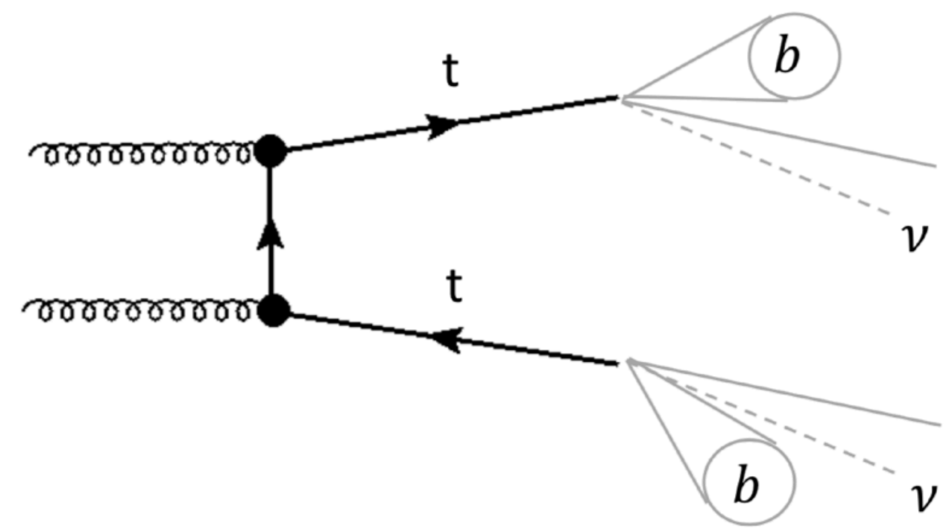
Signal

- Muon pair from Higgs decay
- Top quark additional source of jets/leptons



Dominant Background (reducible)

- Less expected jets in final state
- No prompt 3rd lepton

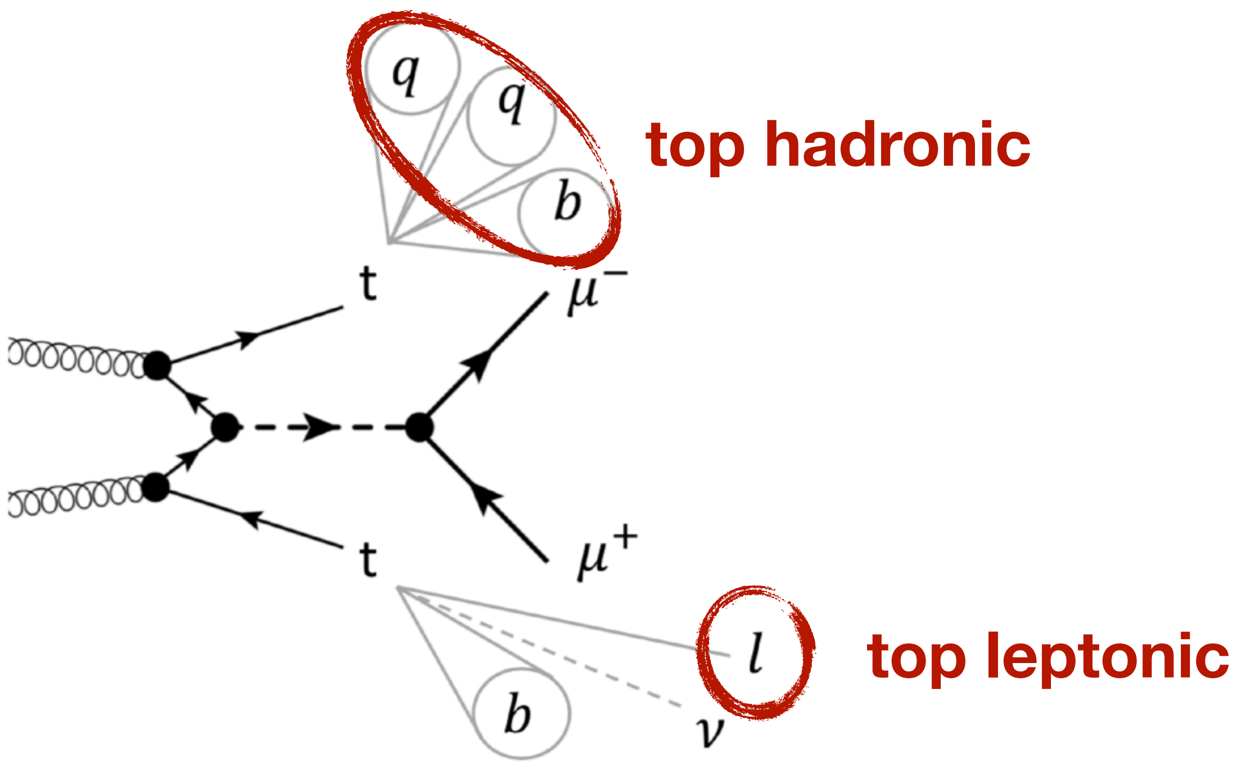


Objectselection: Muons, Electrons with pT > 20 GeV
Mini-isolation (increase acceptance)
MVA lepton ID (improve suppression of non-prompt lepton background)
Jets with pT > 30 GeV, η < 2.5

Preselection and Categorization

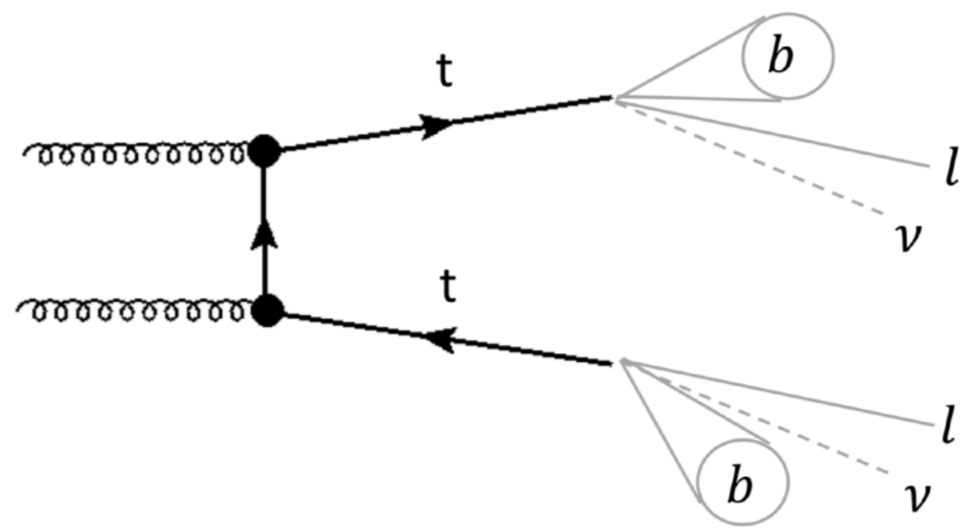
Signal

- Muon pair from Higgs decay
- Top quark additional source of jets/leptons



Dominant Background (reducible)

- Less expected jets in final state
- No prompt 3rd lepton



→ Two independent ttH categories aiming at the decay products of the top quark

ttH leptonic (semi-leptonic ttbar decays)

- B-tagged event
- OS dimuon pair
- exactly 1 additional μ or e

ttH hadronic (fully hadronic ttbar decays)

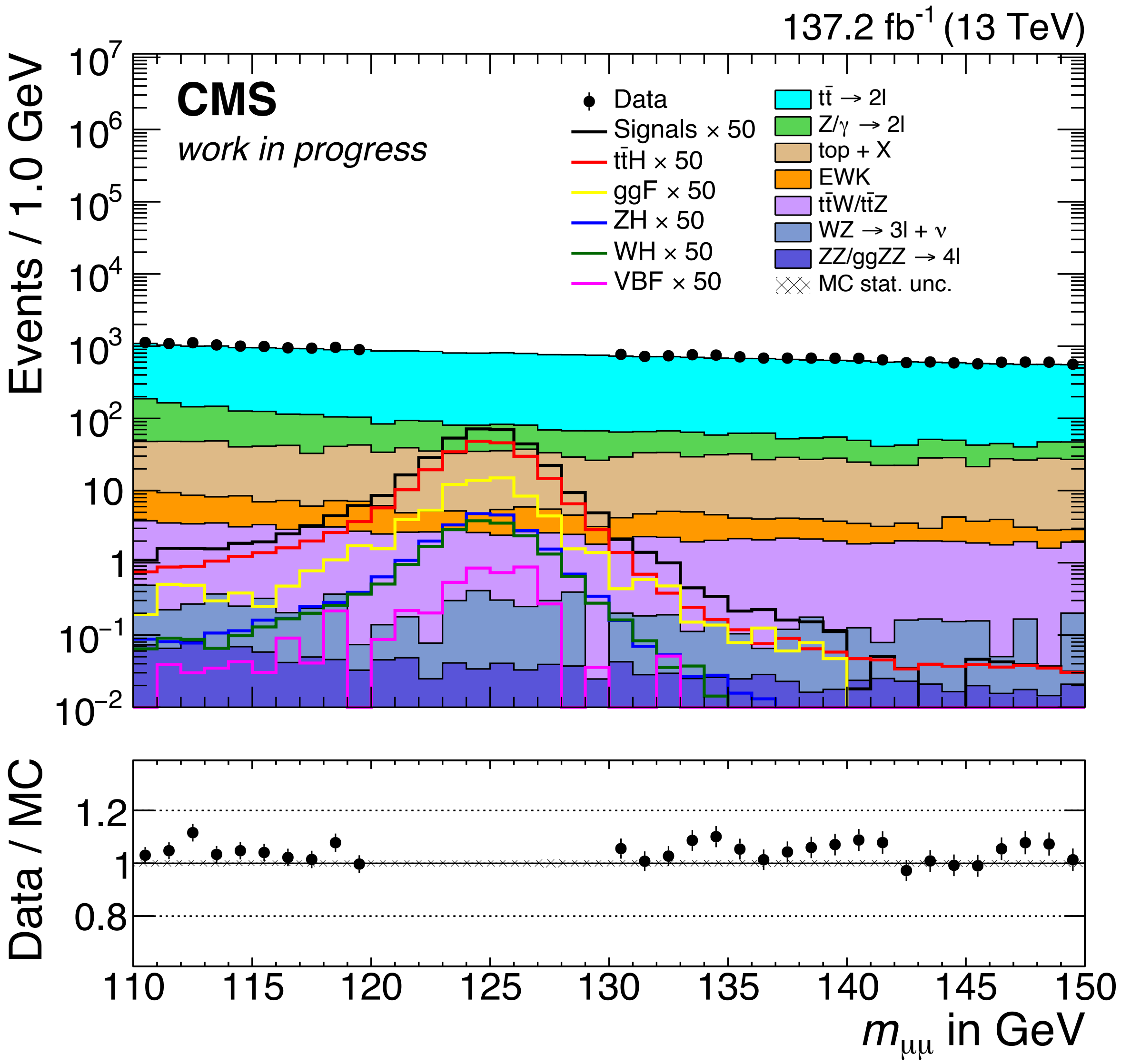
- B-tagged event
- OS dimuon pair
- veto additional μ or e
- Tag jet-triplet (Top tagger)

Optimization - ttH hadronic

ttH hadronic

Category Preselection

- B-tagged event
- OS dimuon pair
- veto additional μ or e
- At least 3 jets in 100 GeV < m(jjj) < 300 GeV

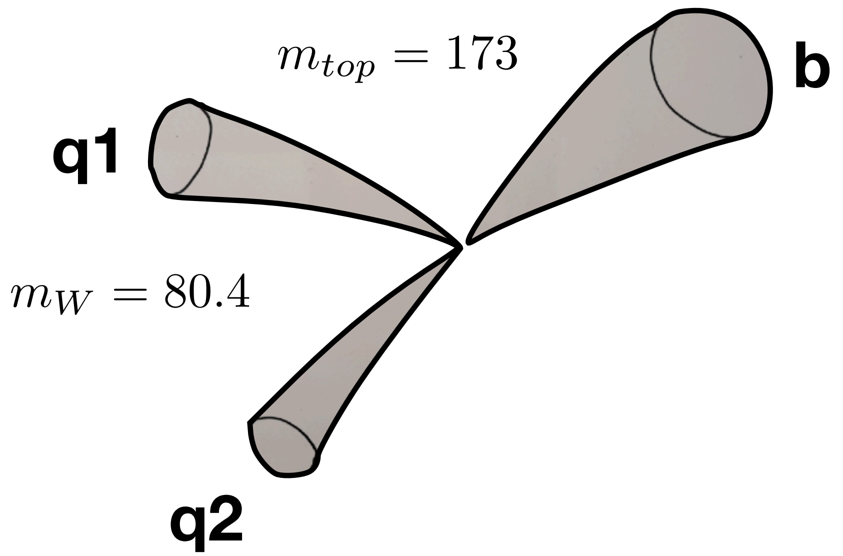


Signal strongly dominated by O(10^3) dileptonic ttbar and Drell Yan

→ Train BDT against dominant backgrounds

1) Resolved Top tagger
2) Mass independent BDT

Resolved Hadronic Top Tagging



Resolved Hadronic Top Tagger

Kinematic Fit → Boosted Decision Tree

Particle-level BDT based on a kinematic fit of jet-triplets for identification of low to moderately boosted top quark decays

Boosted Decision Tree setup:

- use only fit variables & properties of the jet-triplets
- no global event information contained

- Inputs
  single jet: DeepCSV, QGL, pT
  jet triplet: angular distributions, m_W, m_{top}, P^t_{top}
- Training
  Signal: gen-matched jet triplets from ttH
  Background: random jet combinations from ttbar, DY

Jet permutations e.g. Njets = 3

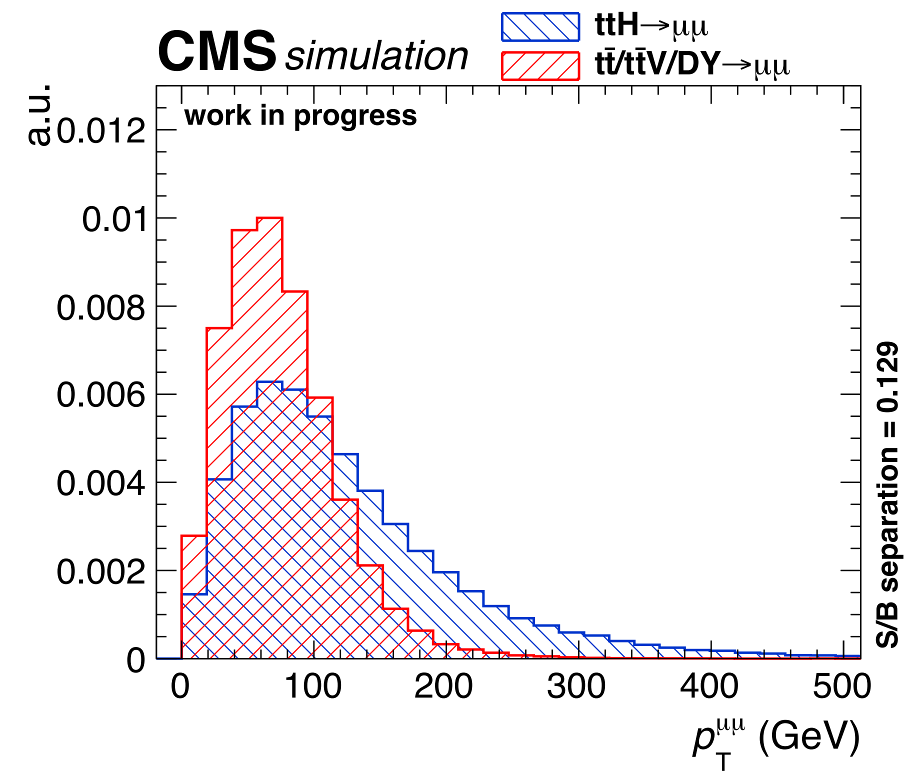
b	q1	q2
Jet1	Jet2	Jet3
Jet2	Jet1	Jet3
Jet3	Jet1	Jet2

- all combinations are tested
- redundant combinations avoided (q1 ~ q2 do not permutate)

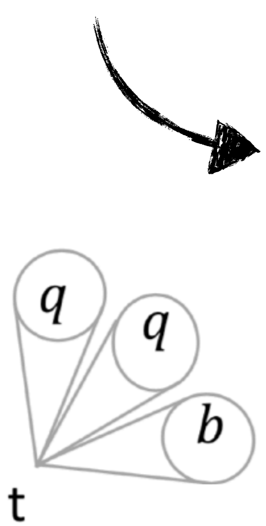
ttH hadronic BDT

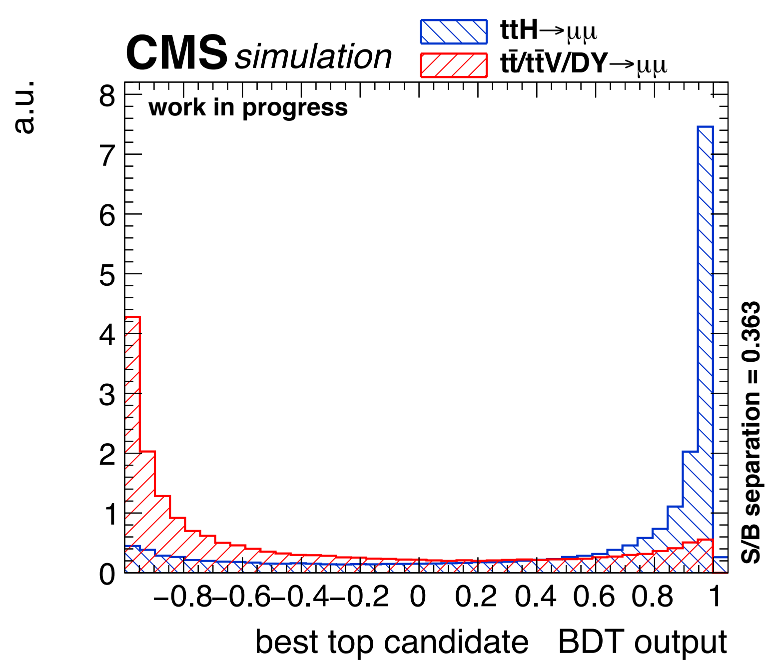
BDT training:

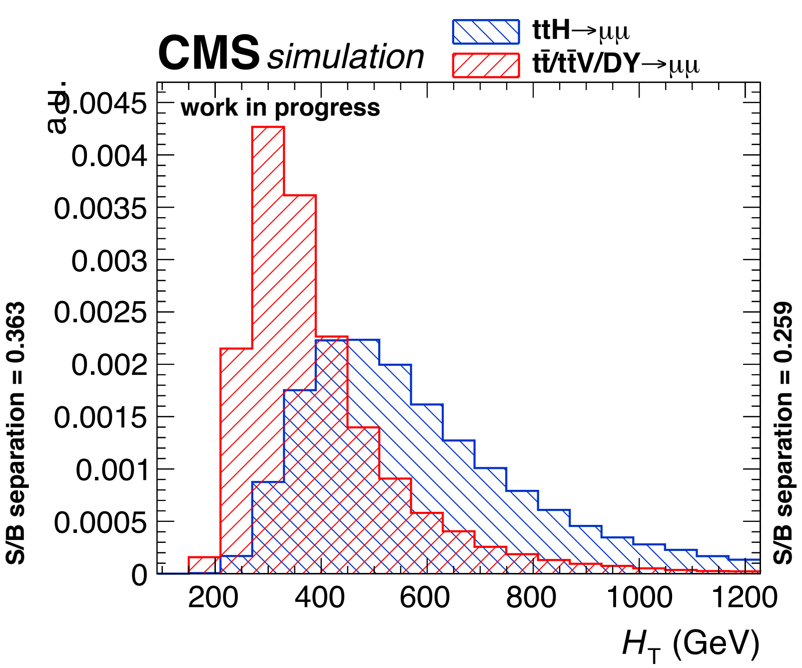
- In full signal region 110 GeV < m(μμ) < 150 GeV
- Signal: ttH (m(μμ)=120,125,130)
- Background: dileptonic ttbar, Drell Yan
- No dimuon mass information included
- Observables:
  Dimuon system (pT(μμ), y(μμ), cos(Θ˟), ϕ˟)
  Hadronic top quark decay & high hadronic event activity (best top tagger BDT output, HT, MET)
  (Total: 14)

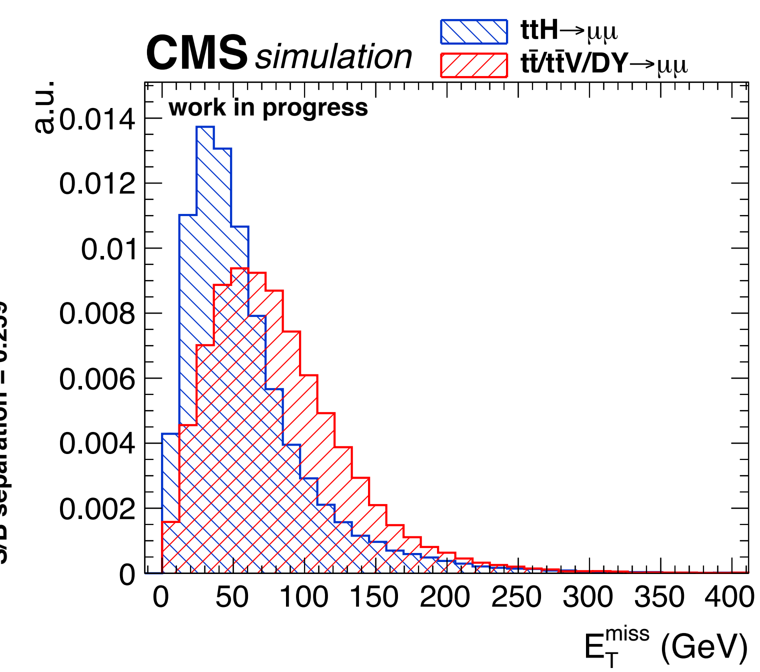


Additional information from hadronic top quark decay





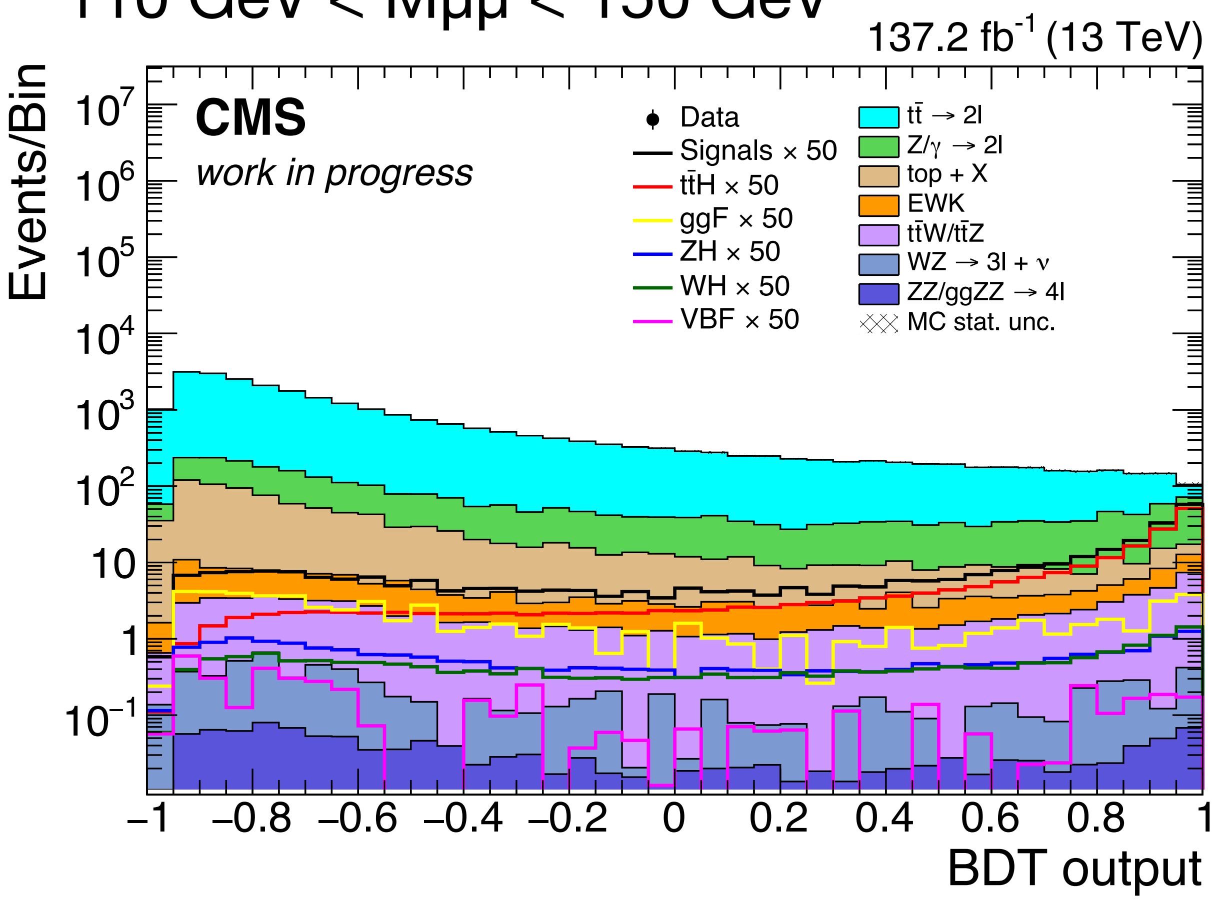




ttH hadronic BDT

Fully blinded signal and sideband region

110 GeV < Mμμ < 150 GeV



BDT Performance

Good separation between signal and backgrounds

Suppress background by several orders of magnitude in the high BDT score region

Optimization - ttH leptonic

ttH leptonic

Event selection

- B-tagged event
- OS dimuon pair
- exactly 1 additional μ or e

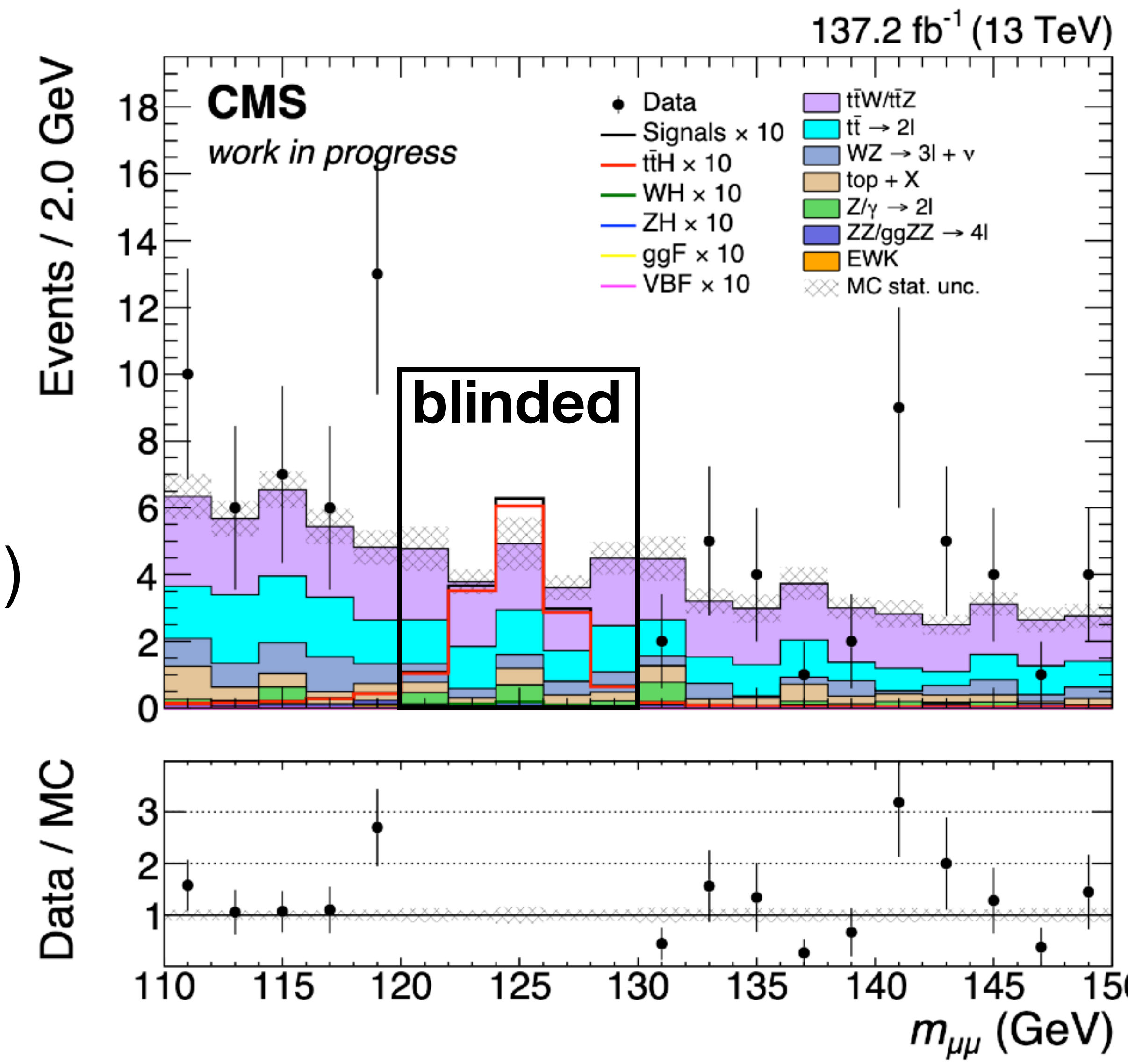
1. Optimize dimuon pair finding & reduce combinatorial background (ttV, VV, VVV)

→ choose 110 GeV < m(μμ) < 150 GeV
if several possibilities highest Pt(μμ)

→ cut on mass window around Z peak
(veto event if any OS combination of muons in 81 GeV < m(μμ) < 101 GeV)

2. Separation between prompt and non-prompt leptons

→ Use multivariate lepton identification trained against non-prompt leptons



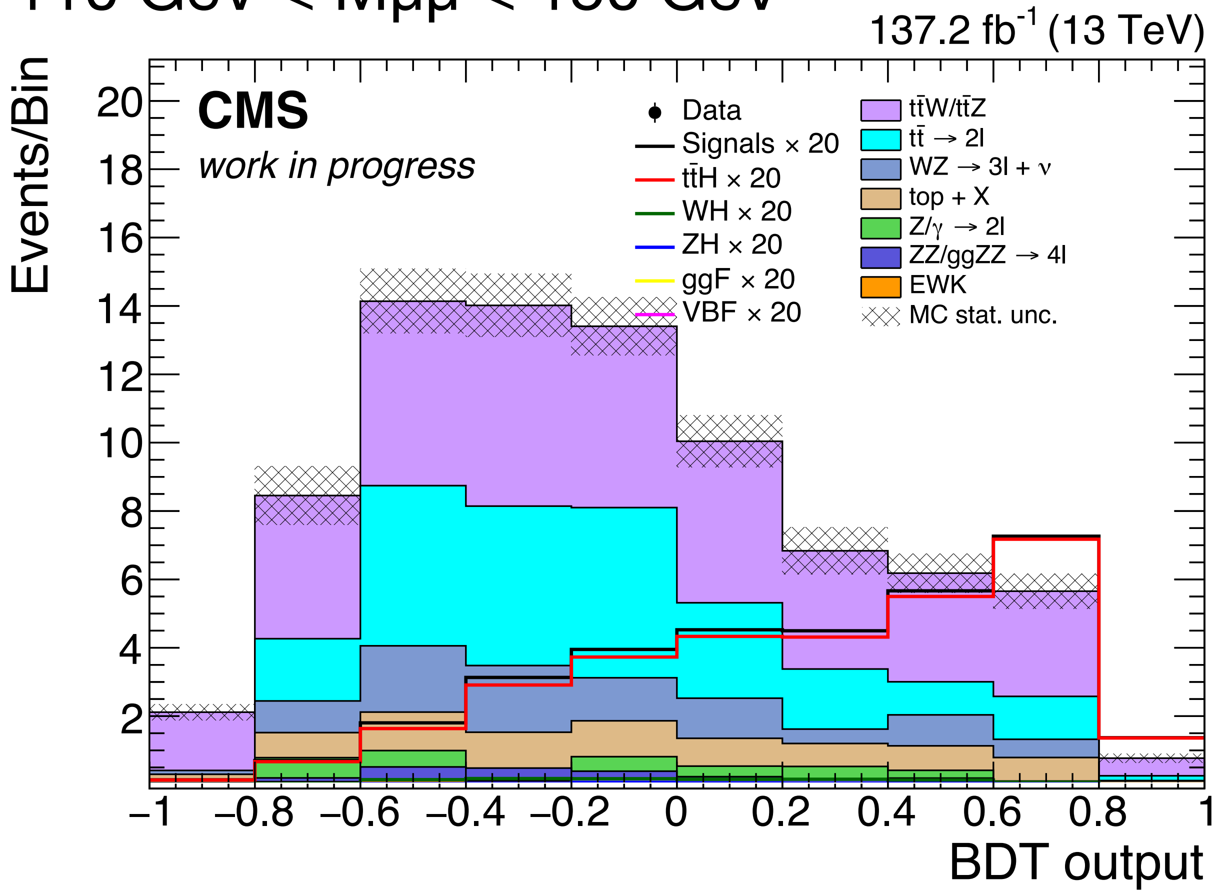
- Reduction of fake lepton background achieved due to LeptonMVA
- Dominant irreducible background is ttW/ttZ

→ Mass-independent BDT against ttW/ttZ

ttH leptonic BDT

Fully blinded signal and sideband region

110 GeV < Mμμ < 150 GeV



BDT Performance

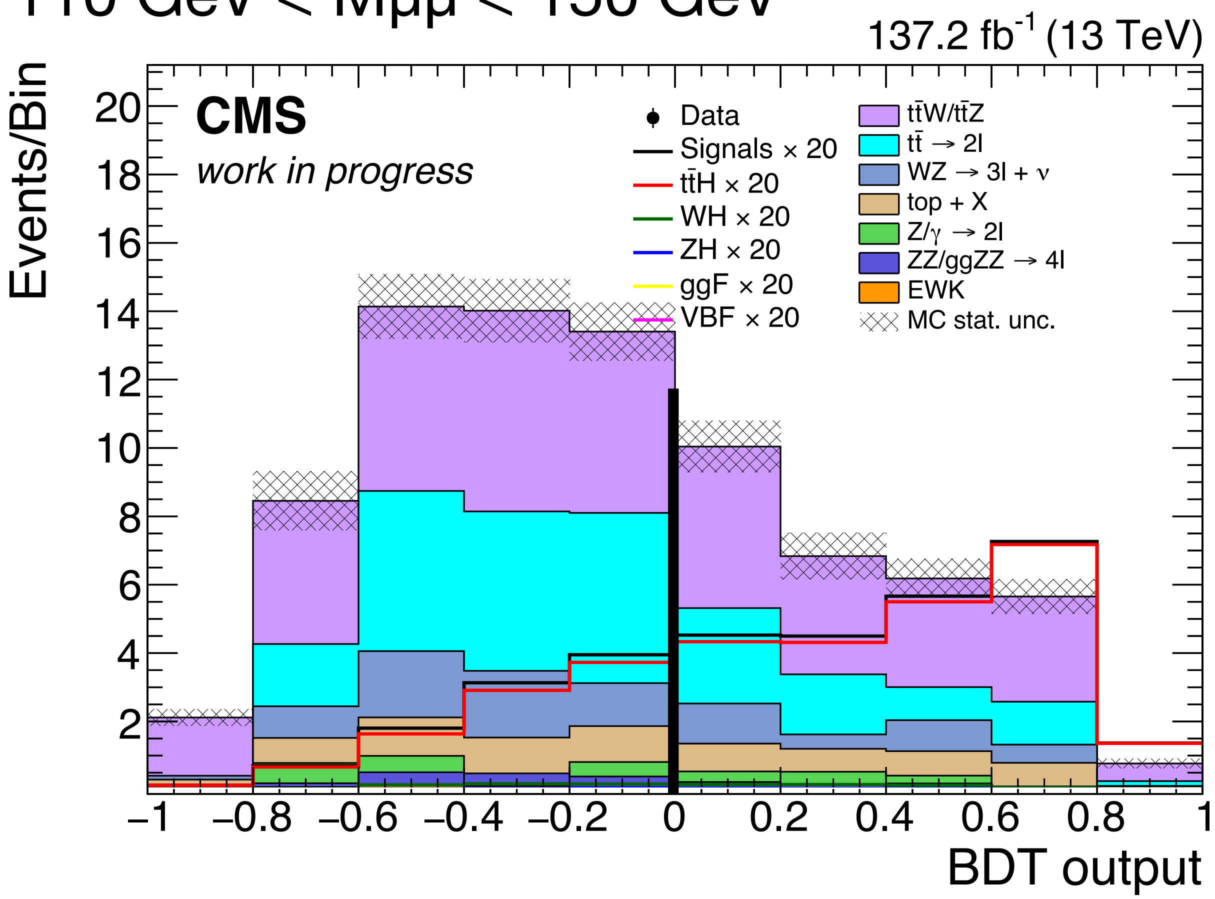
Good separation between signal and backgrounds

BDT Categorisation

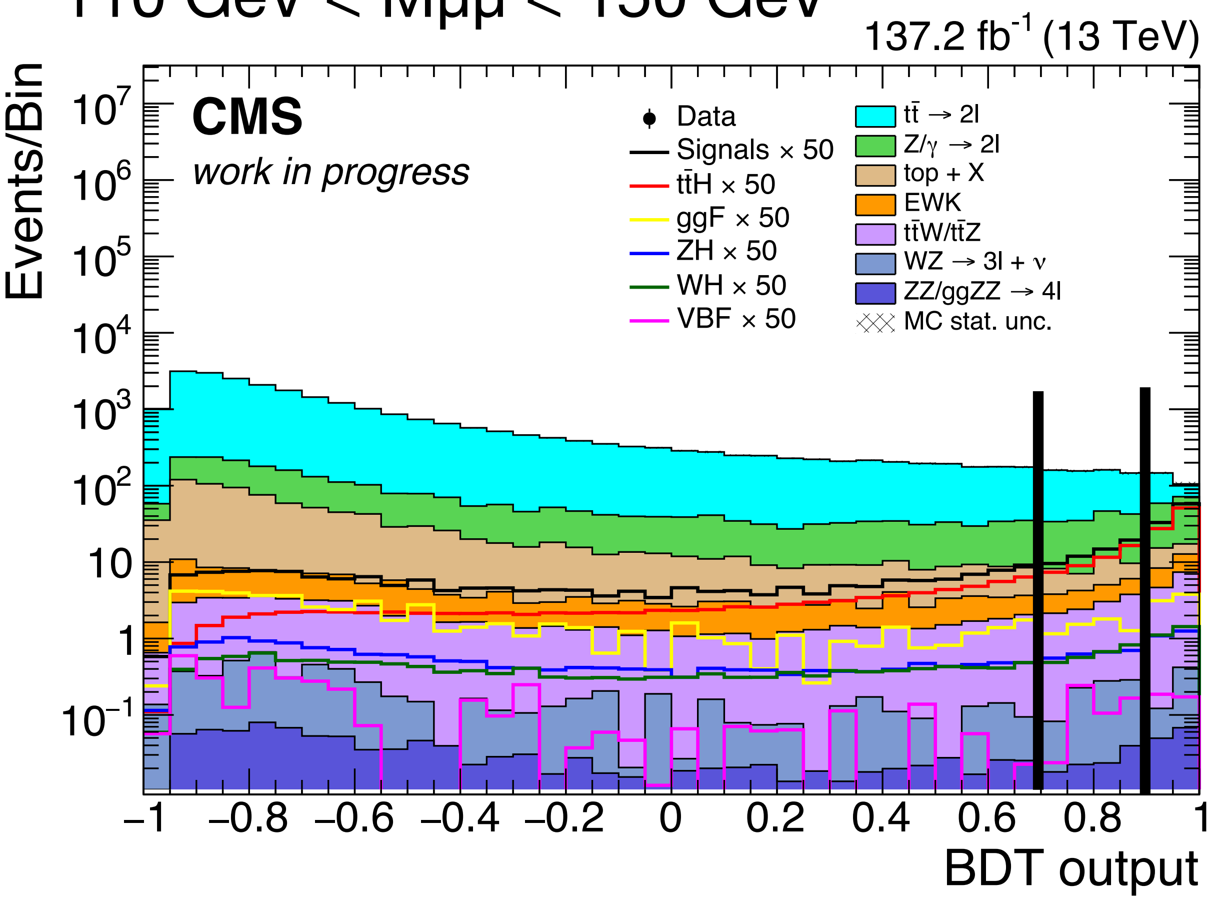
BDT categorisation

Fully blinded signal and sideband region

110 GeV < Mμμ < 150 GeV



110 GeV < Mμμ < 150 GeV

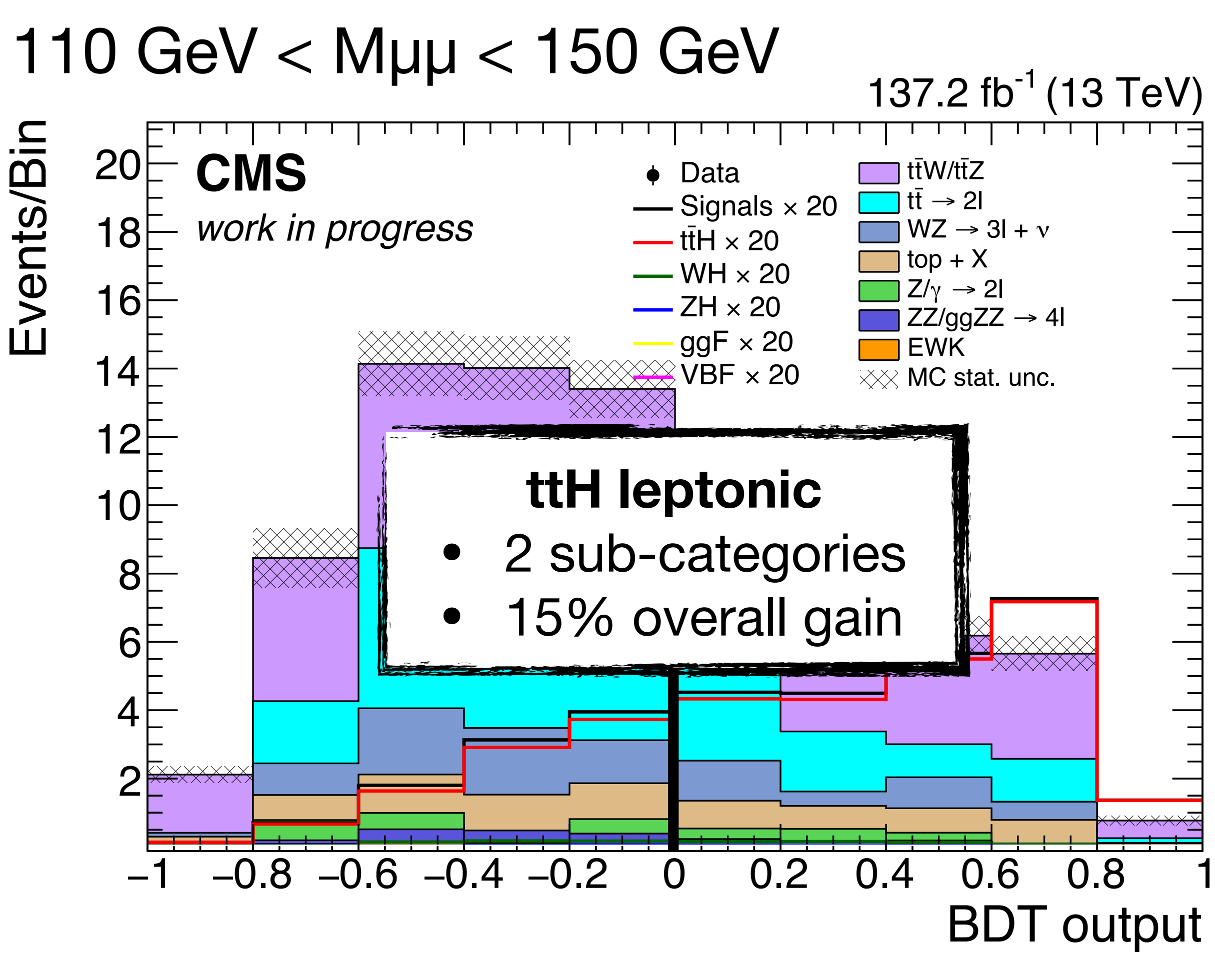


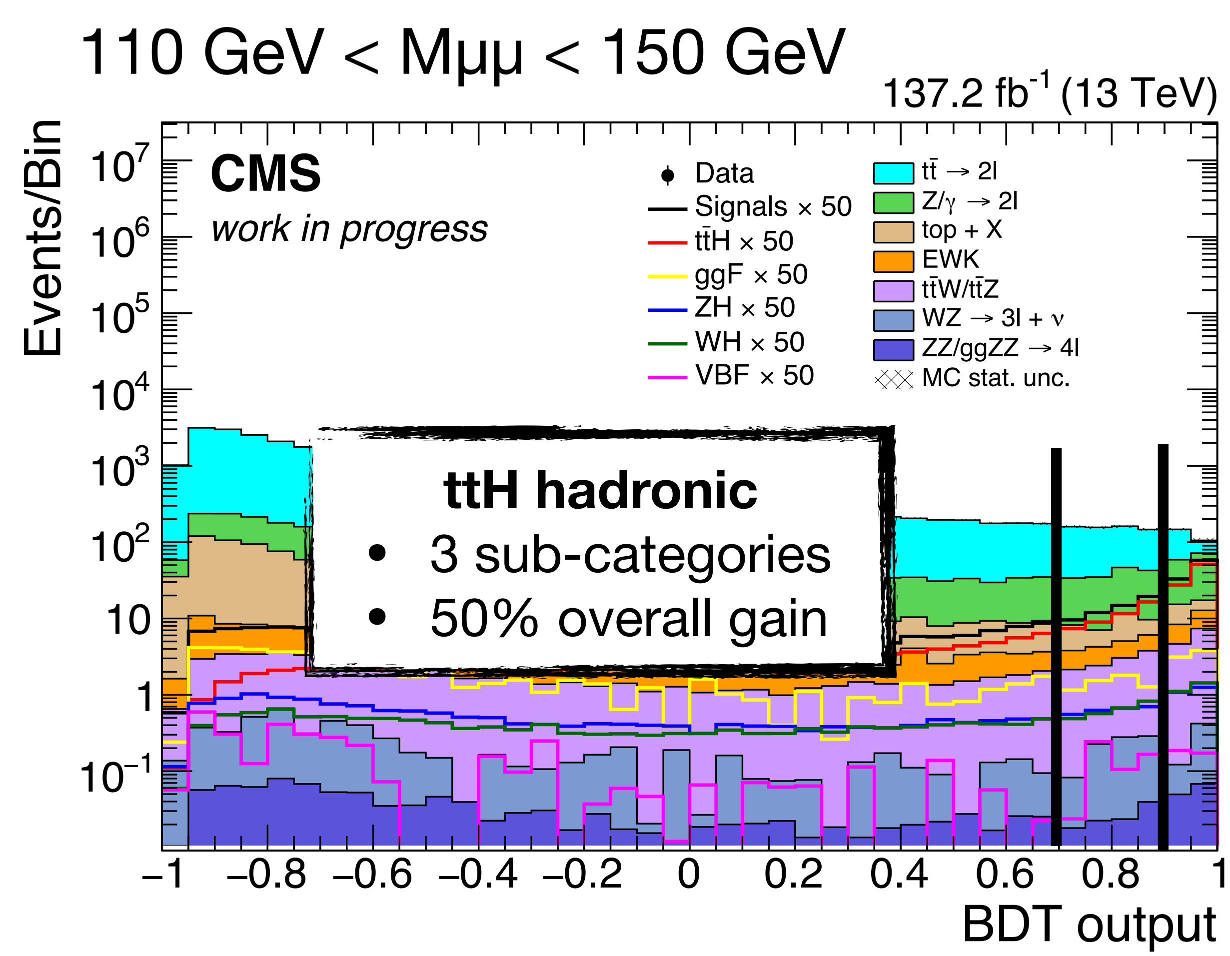
BDT splitting

- Determine the optimal boundaries for n subcategories with an iterative procedure to have the best expected limit
- Stop iteration when gain less than O(2%) or statistics too small

BDT categorisation

Fully blinded signal and sideband region



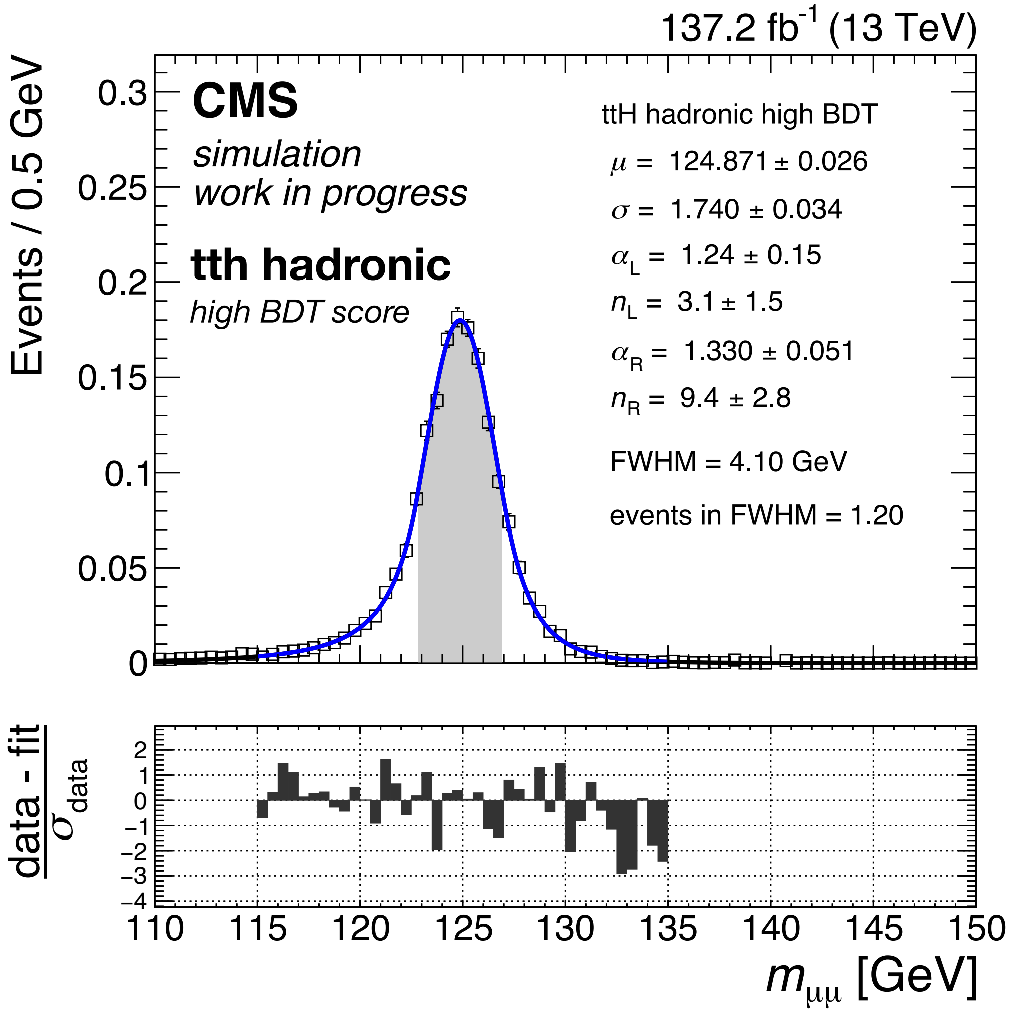


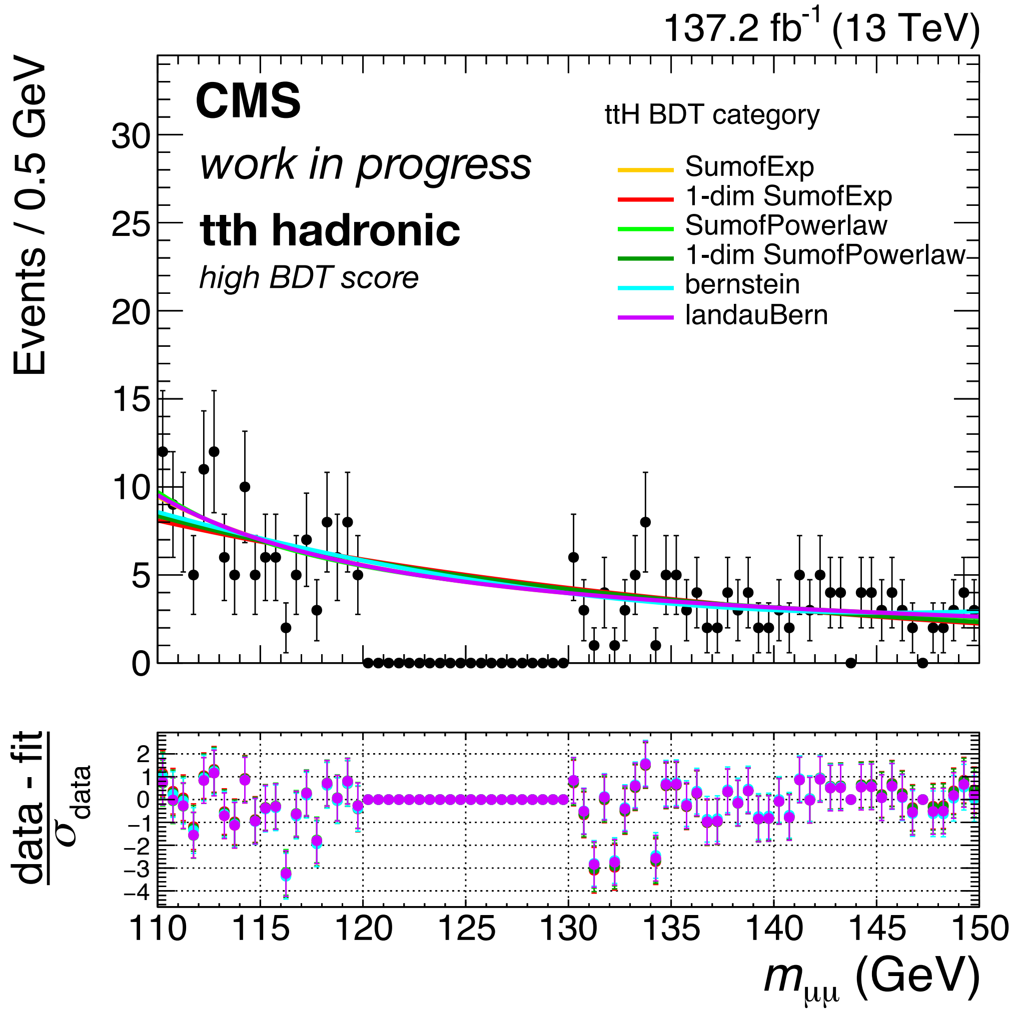
BDT splitting

- Determine the optimal boundaries for n subcategories with an iterative procedure to have the best expected limit
- Stop iteration when gain less than O(2%) or statistics too small

Limit Setting

Background Estimation & Limit Setting





- **Signal model**: double-sided Crystal Ball
- **Background**: Fully data-driven background estimation from sideband in dimuon mass
- Use a set of agnostic functions: exponential, powerlaw & polynomials with DoF = 1, 2
- Bias studies performed to check systematic uncertainty from the choice of background function
- Sum of exponentials is unbiased on all subcategories and chosen to model the background
- Systematics evaluated on signal

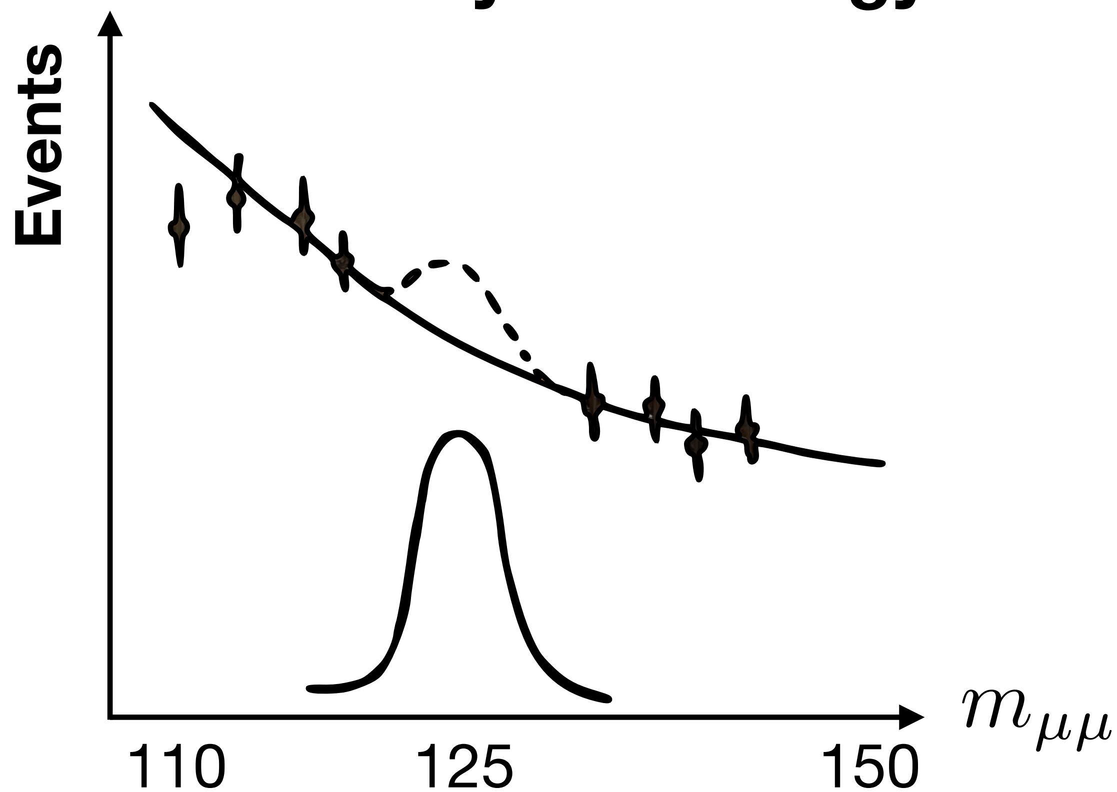
Results & Summary

- Development of a search strategy for SM ttH→μμ in progress
- Orthogonal phase-space selection using b-tagged jets
- Dividing into a leptonic and hadronic category
- BDT optimization and sub-categorization in 5 BDT categories
- Parametrisation of signal and background using analytical functions fitted to dimuon mass
- First estimates of the sensitivity are comparable to the results of the CMS full Run1 Higgs to muons analysis

BackUp

Introduction to SM H→μμ

General Analysis Strategy



- SM Higgs decays to muons appear as sharp peak in the dimuon mass spectrum at 125 GeV

1. BDT independent of m(μμ) separates signal and background
2. Define categories according to the output of the BDT
3. Fully data-driven background model from sideband fit to data

Latest ATLAS and CMS results

ATLAS: H → μμ full Run2 results

[ATLAS-CONF-2019-028]

- Preliminary results public
- Signal strength limits: 1.3 exp. / 1.7 obs.

CMS: H → μμ RunI + 2016 results

[Phys. Rev. Lett. 122, 021801]

- Inclusive approach using all dimuon events independent from the production channel
- Sensitivity driven by
  1) 1-jet events (GGF enriched, high μμ p_t)
  2) 2-jet events (VBF enriched)
- Signal strength limits: 2.16 exp. / 2.92 obs.