Combinatorial likelihood discriminant method for ttH(H->bb) search with the ATLAS detector at the LHC

Yulia Rodina

Supervisors:
Laurent Vacavant, CPPM, Aix-Marseille Université and CNRS/IN2P3, Marseille
and
Aurelio Juste Rozas, IFAE, Universitat Autònoma de Barcelona and ICREA, Barcelona

CPPM PhD Day
October 2016

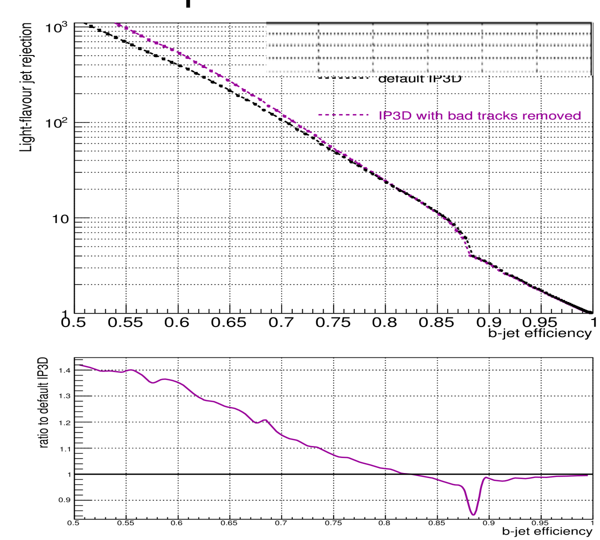
Introduction

- My PhD is a joint project between Aix-Marseille Université (supervisor: Laurent Vacavant) and Universitat Autònoma de Barcelona (supervisor: Aurelio Juste Rozas)
- The project is funded by A*MIDEX

- For the first the first year of my PhD I was working on b-tagging (identifying jets stemming from hadronisation of b-quarks)
- b-tagging is important for ttH(H->bb) channel as there are 4 b-jets in final state → have to identify them
- I was involved into work on optimization of IP2D/IP3D b-tagging algorithms for LHC Run 2. The key goal of this optimization was to take advantage of the new layer of ATLAS inner detector - Insertable B-layer (IBL). I was developing the new track grading which allowed to improve b-tagging performance for Run 2 and working on further algorithm optimization

Track grading for Run 2: ~15-20% improvement with respect to Run 1

Further optimization of IP3D algorithm: ~15% improvement

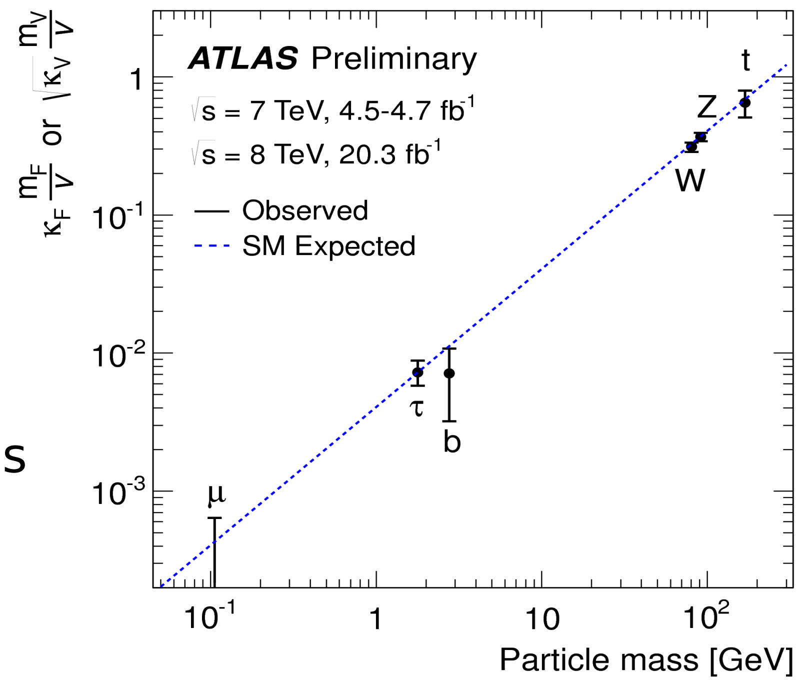


Introduction: ttH(H->bb)

LHC Run 1 (2009-2013): Standard Model like Higgs boson was discovered

$m_H = 125.09 \pm 0.21(stat.) \pm 0.11(syst.)$ GeV

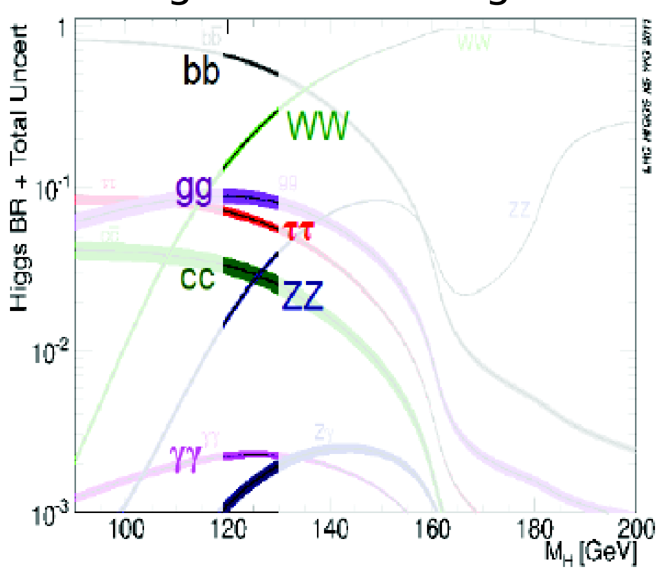
But not yet observed in $t\bar{t}H$ channel!

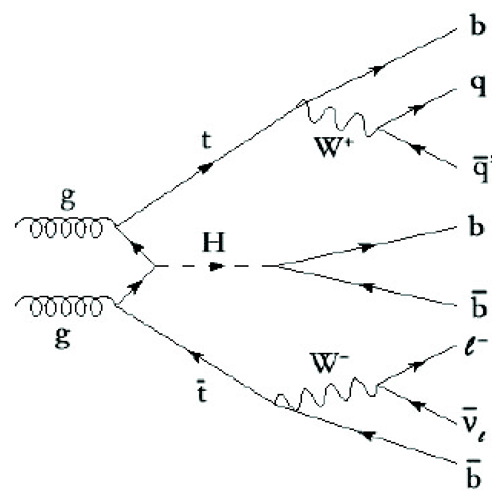


- $t\bar{t}H$ has the lowest cross section of Higgs production at LHC
- with higher energies the impact is becoming larger:
  - Run 1: √s=8 TeV -> σ(ttH)=130 fb
  - Run 2: √s=14 TeV -> σ(ttH)=611 fb – ~5 times larger!

Why important to study $t\bar{t}H$:
gives direct access to the top Yukawa coupling

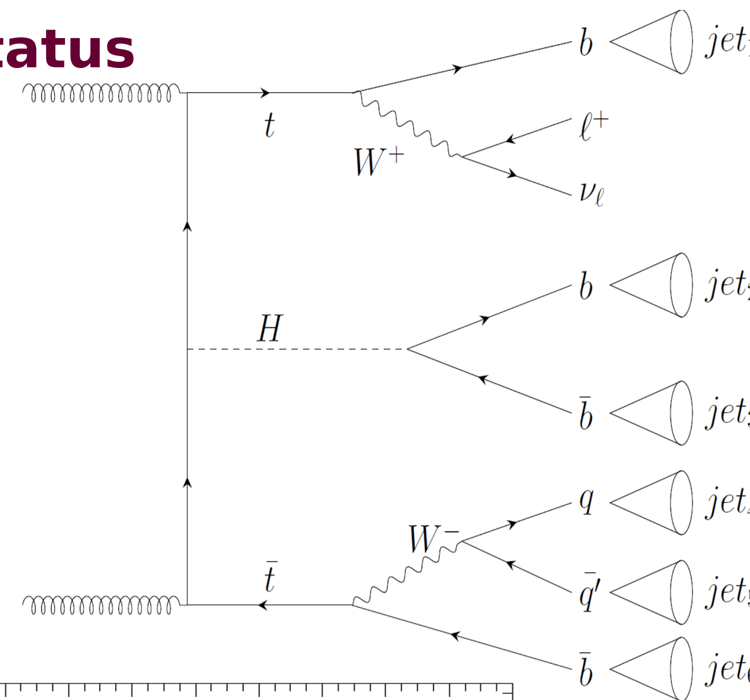
H→bb branching ratio - the largest in the SM:





4 b-jets in the final state → important to identify them

ttH(bb): analysis status



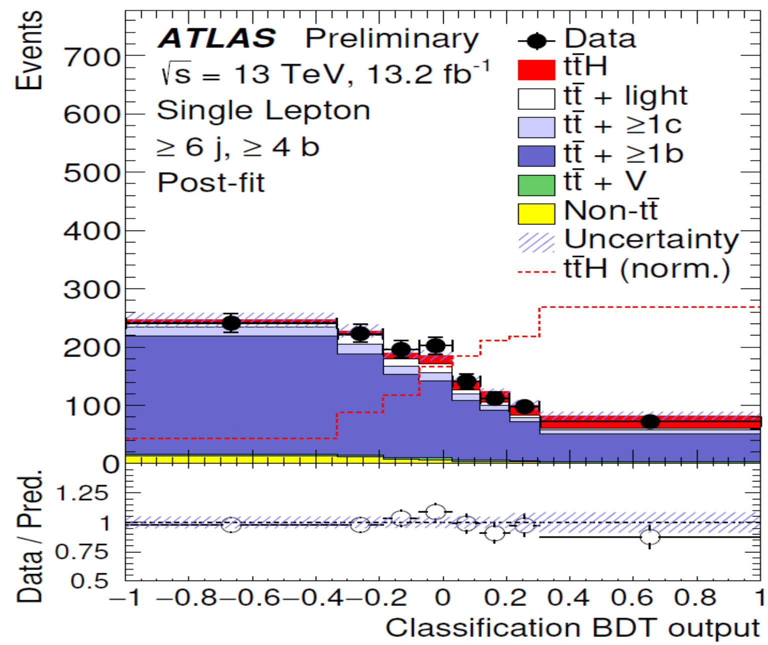
Signal regions:

Reconstruction BDT: matches the reconstructed jets to the partons from top and Higgs decays

Other kinematic variables e.g. sum of jet pT, average deltaR between two b-jets

Classification BDT: provides final output to separate signal from background

More about BDT in
Royer Ticse Torres' PhD defense talk



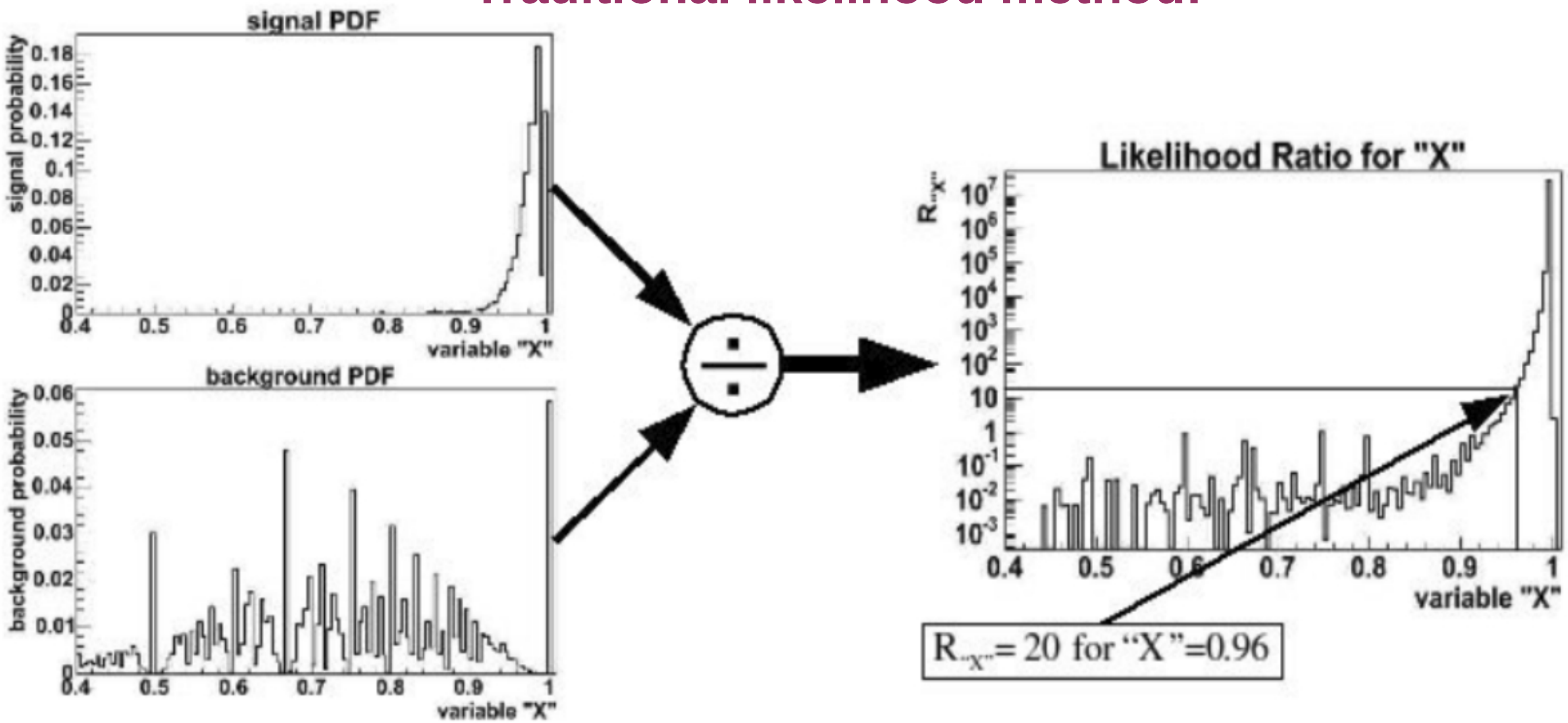
Recent result (presented at ICHEP conference, August 2016): ratio of the measured ttH cross section to the Standard Model prediction: $\mu = 2.1^{+1.0}_{-0.9}$

Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in *pp* collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector", ATLAS-CONF-2016-080

Likelihood discriminant method: introduction

For the second year of my PhD I was working on applying a new method of separating signal from background for ttH(bb) analysis: combinatorial likelihood method

Traditional likelihood method:



- The presented method is based on combinatorial likelihood approach: it extends traditional likelihood method by taking into account the combinatorics of the complex final state

- Was developed and successfully applied for flavor-changing neutral current (FCNC) analysis: t->Hq(H->bb) (see JHEP 12 (2015) 061, arXiv:1509.06047 [hep-ex])

- Being tested in the context of the ttH(bb) analysis

Likelihood discriminant method

- Discriminant between signal and background:

$$D(x) = \frac{P^{sig}(x)}{P^{sig}(x) + P^{bkg}(x)}$$

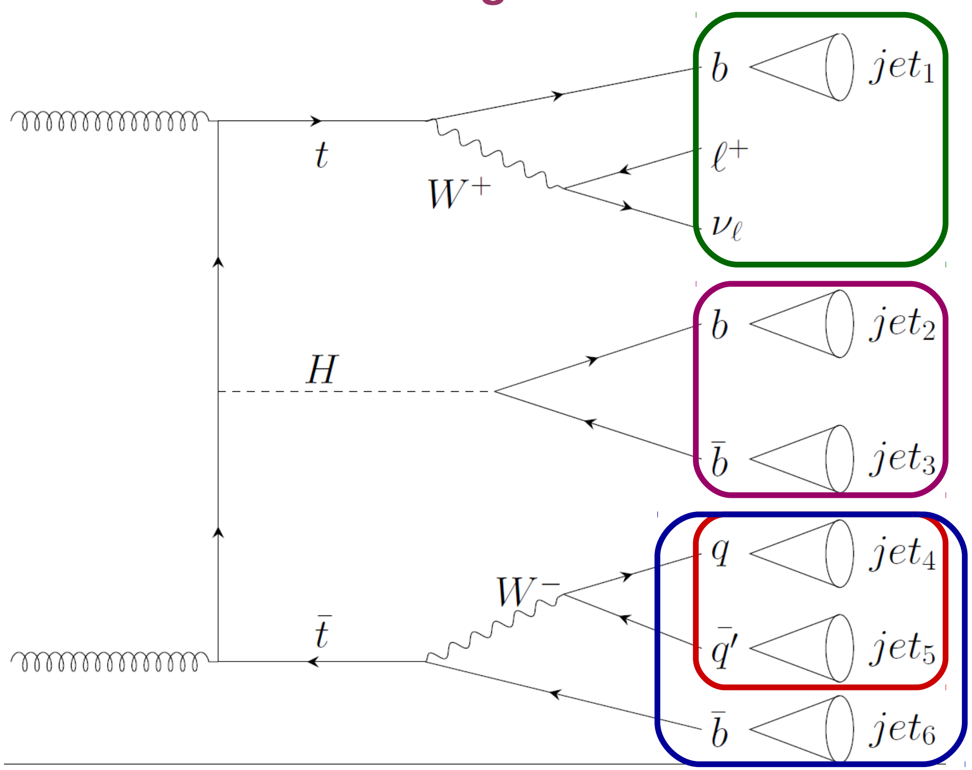
$P^{sig}(x), P^{bkg}(x)$ - PDF of a given event under signal or background hypothesis

X – the 4-momentum vectors of all final state objects at reconstruction level: lepton, neutrino, jets

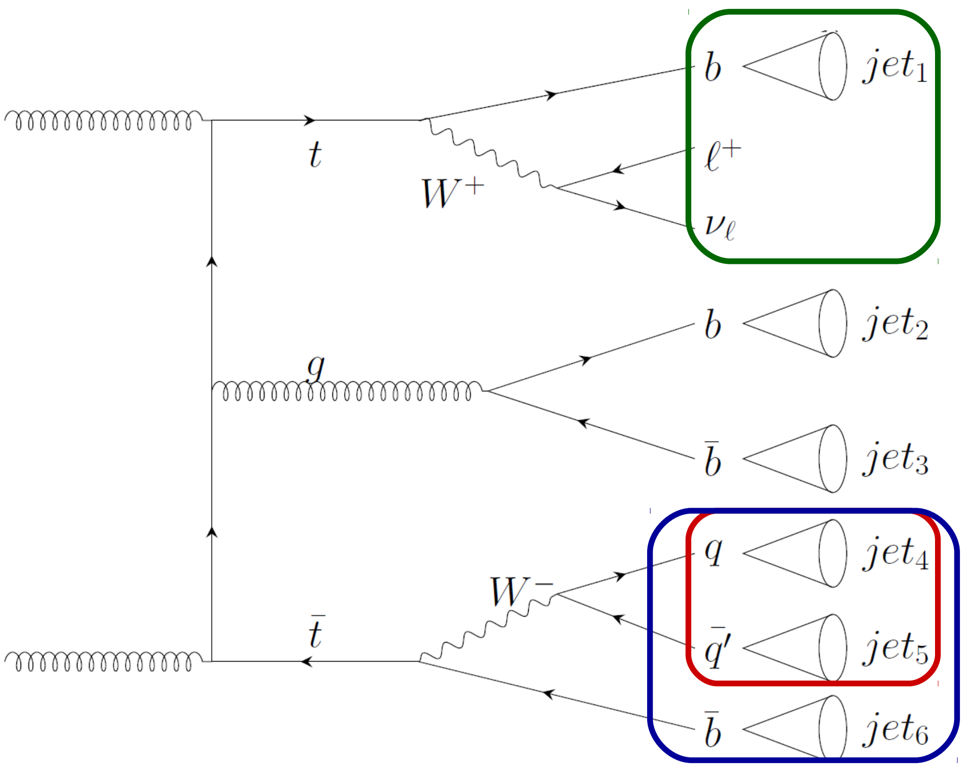
Main observables exploited in this method - various invariant mass resonances: leptonic top, Higgs, hadronic top, hadronic W

Restricted (for now) to two processes:

Signal: ttH



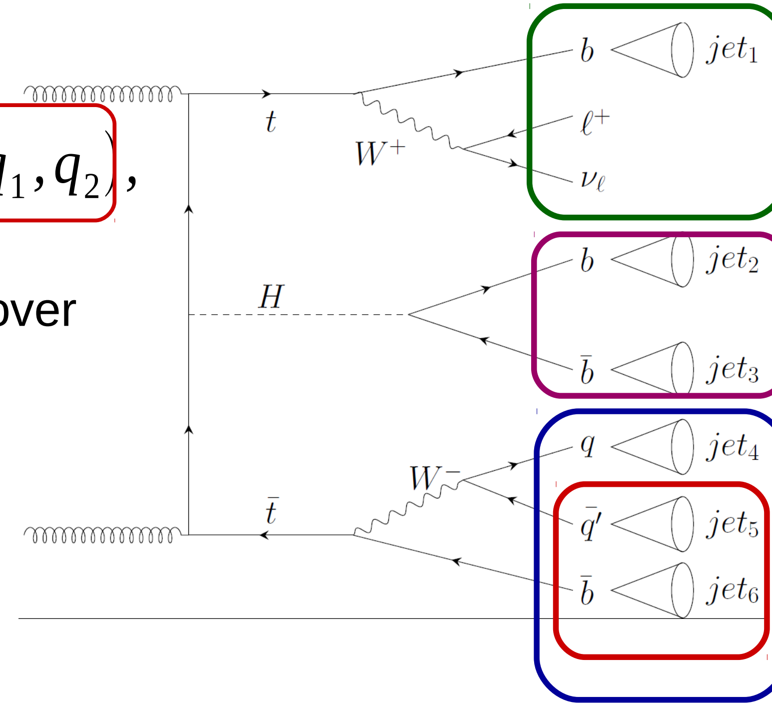
Main background: ttbar+b



Signal probability

If we know origin of each jet, intuitively signal probability is

$$P^{sig}(x)=\boxed{P^{sig}_{M_{l\nu b_l}}(l,\nu,b_l)}\boxed{P^{sig}_{M_{b_1 b_2}}(b_1,b_2)}\boxed{P^{sig}_{M_{q_1 q_2 b_h}}(q_1,q_2,b_h)}\boxed{P^{sig}_{M_{q_1 q_2}}(q_1,q_2)},$$



But we don't know the jets' origin, so we sum over all possible N_p permutations, weighting them with b-tagging terms:

$$P^{sig}(x)=\frac{\sum_{k=1}^{N_p} P^{sig}_{kin}(x^k)\,P^{sig}_{btag}(x^k)}{\sum_{k=1}^{N_p} P^{sig}_{btag}(x^k)}$$

$P^{sig}_{kin}(x)$ - kinematic probability term, using PDFs of the reconstructed invariant masses in the event

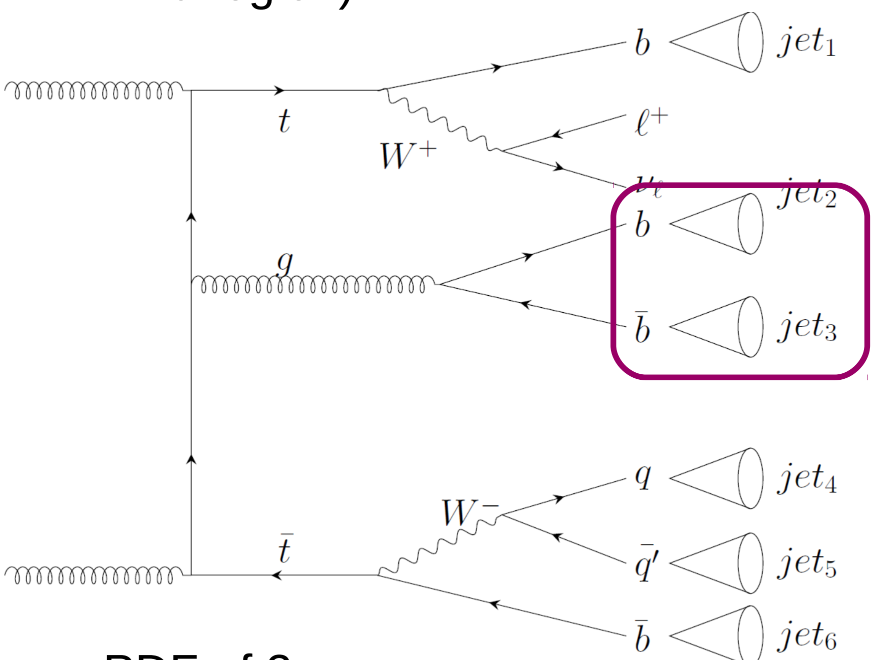
$P^{sig}_{btag}(x)$ - b-tagging probability term, used the suppress the combinatorial background

Background probability

Similar to signal, but instead of Higgs – extra jets

bb

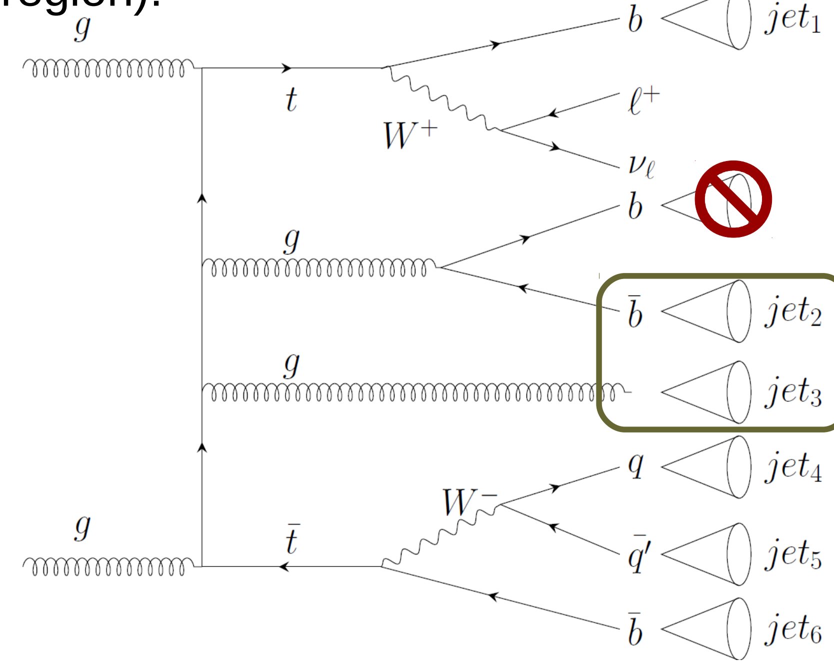
>=2 truth extra b (~78% in >=6 j >= 4 b region):



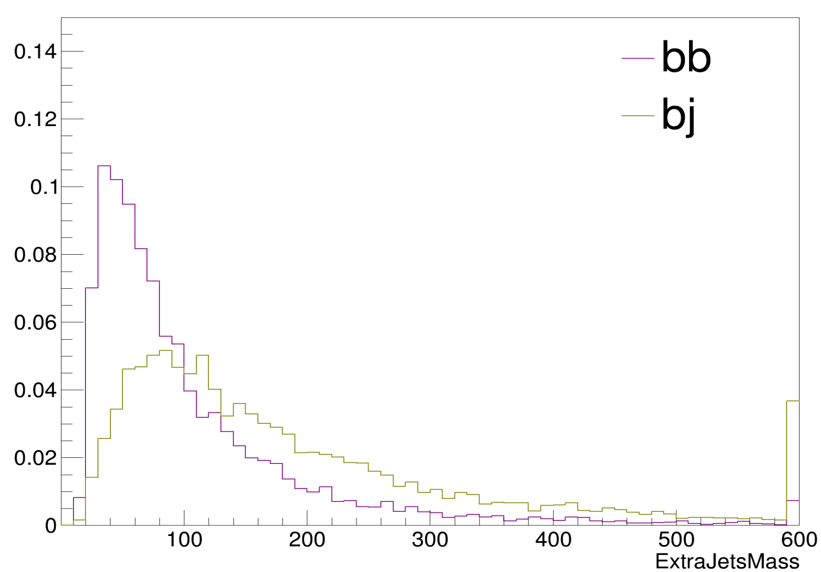
- mass PDF of 2 leading in pT extra b-jets

bj

1 truth extra b: (~22% in >=6 j >= 4 b region):



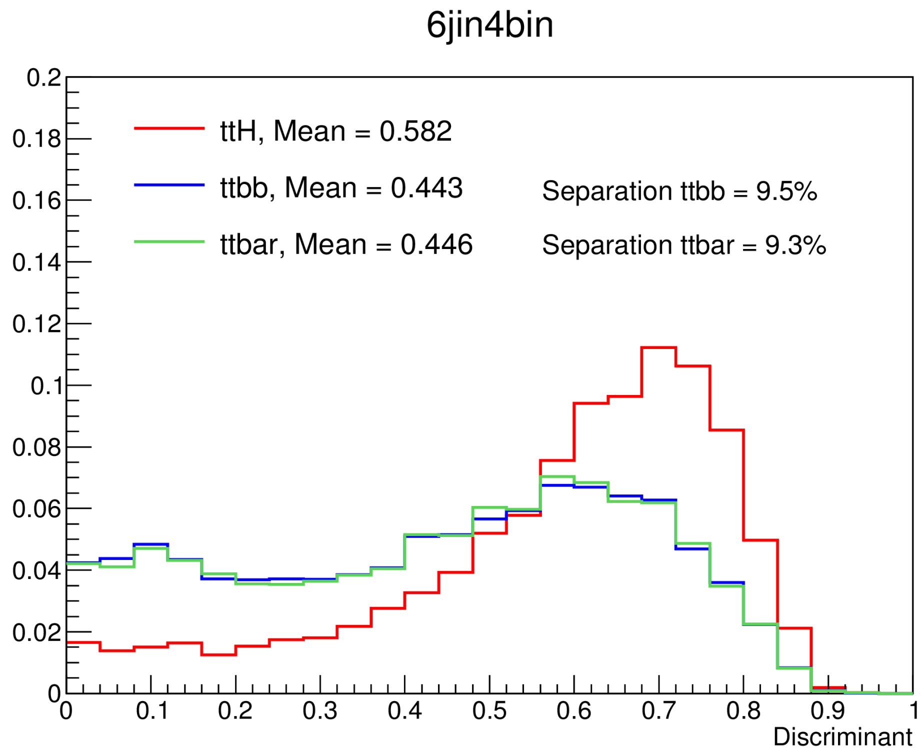
- mass of 1 extra b-jet and leading in pT non-b-jet



Results, summary and plans

Discriminant between signal and background:

$$D(x) = \frac{P^{sig}(x)}{P^{sig}(x) + P^{bkg}(x)}$$



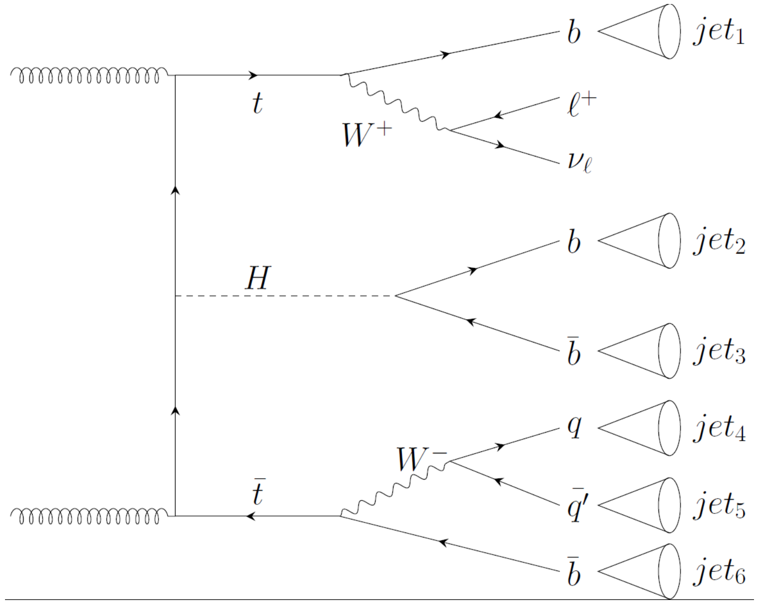
- Status of work:
  - The Likelihood discriminant method is being tested in the context of ttH(bb) analysis
  - Preliminary performance promising, comparable to alternative methods
  - Work is ongoing on optimizations such as adding additional kinematic information, in particular some angular variables (e.g. equivalent to DeltaR(b,b) for extra bb system), background probability, hypothesis of missing parton
- Further plans:
  - The milestone is including the method in the upcoming ttH(bb) publication in summer 2017

Backup material

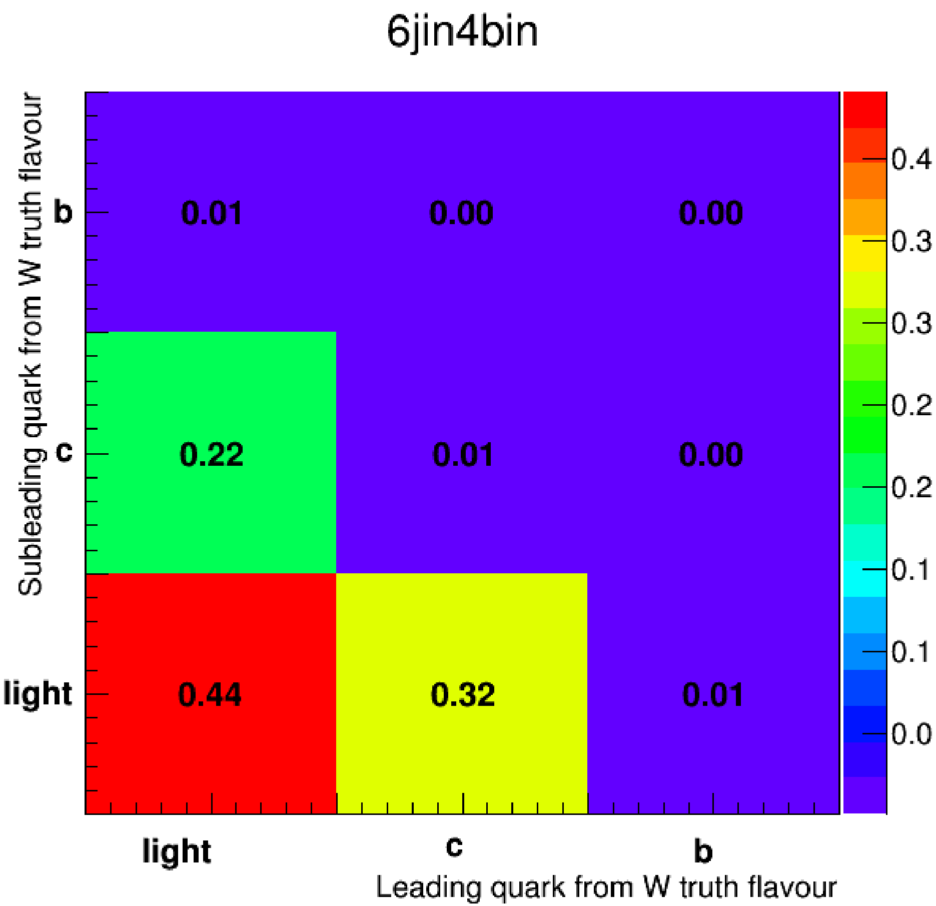
b-tagging probability

$$P^{sig}_{btag}(x) = P_b(j1)P_b(j2)P_b(j3)[f_{light-light}P_{light}(j4)P_{light}(j5) + f_{c-light}P_{light}(j4)P_c(j5)]P_b(j6)$$

$P_f(j_i)$ - probability that a jet i with its given b-tagging weight to be originating of parton with flavour f



- For jets from W there are two flavour hypotheses: light-light (u-d) and c-light (c-s)
- They are considered with corresponding weights extracted from quark from W truth flavour PDF: $f_{light-light} \approx 0.45, f_{c-light} \approx 0.55$



Eventually the weights are planned to be calculated using the b-tagging efficiency maps

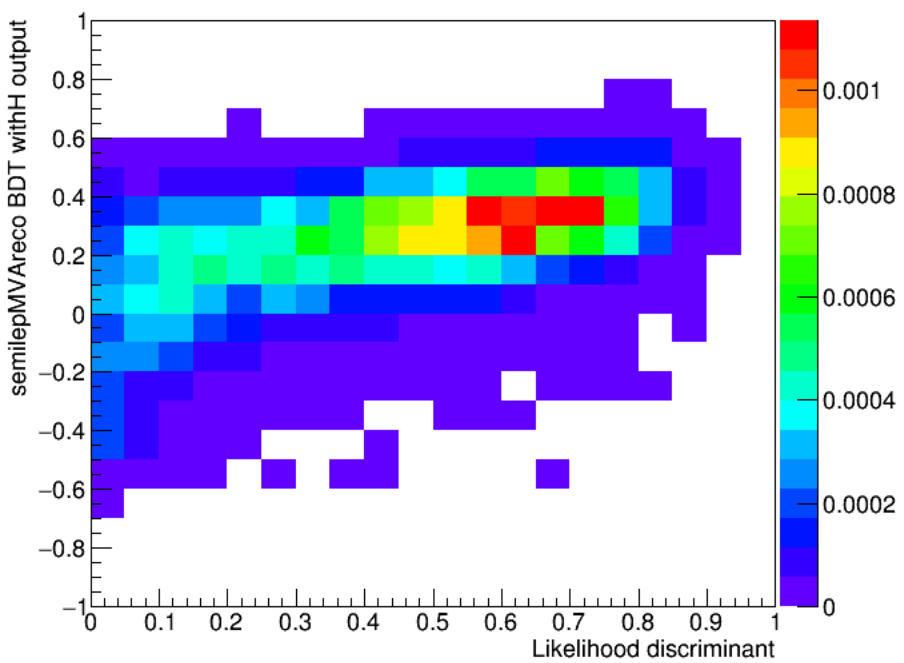
For now using flat numbers:

$P_f(j_i) = eff_f(w_{jet_i})$ at 70% OP if jet is tagged

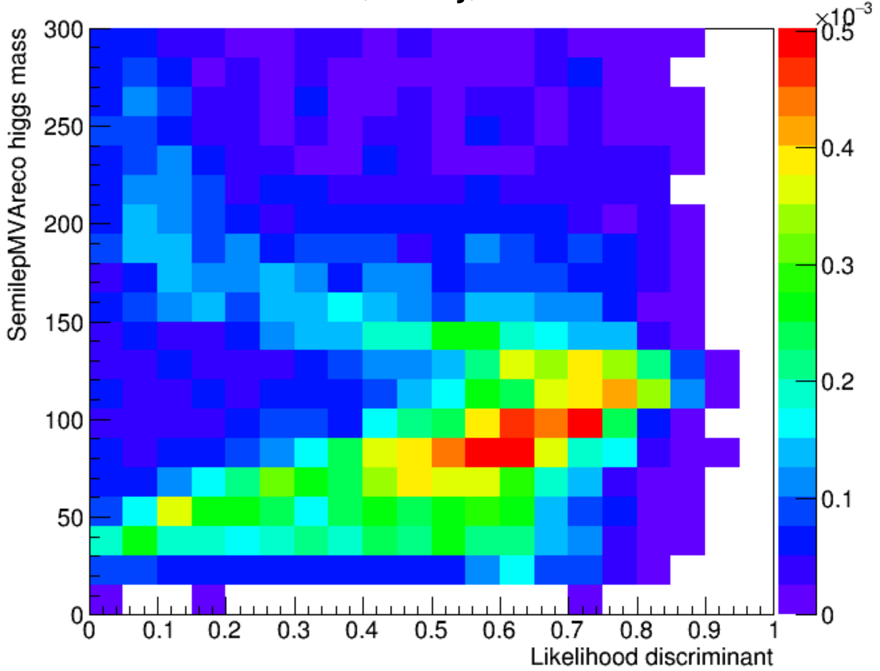
$P_f(j_i) = 1 - eff_f(w_{jet_i})$ if not tagged

Results and outlook

Likelihood discriminant vs recoBDT with Higgs output and Higgs mass:



ttbb, >=6j, >=4b



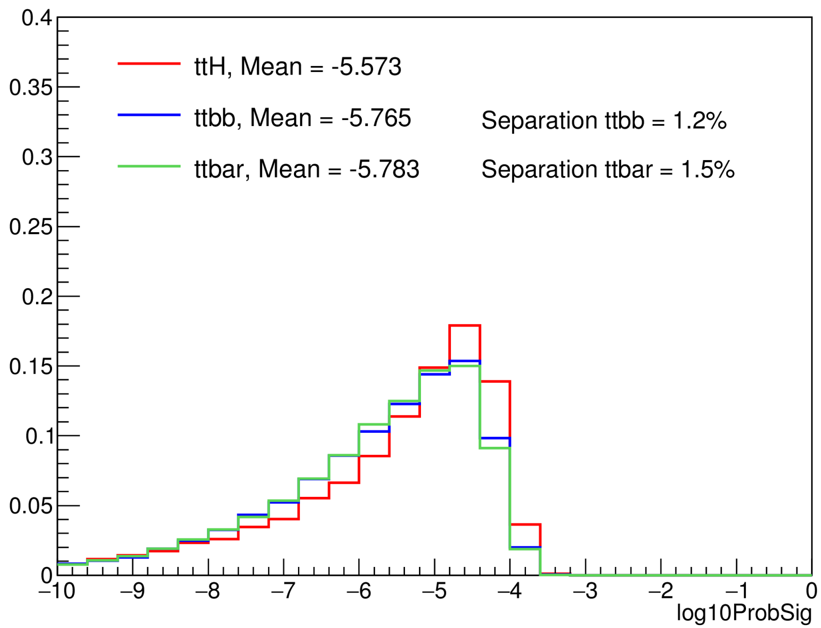
- The method was migrated into the TTHbbLeptonic framework

- Planned optimizations:
  - adding additional kinematic information, in particular some angular variables (e.g. equivalent to DeltaR(b,b) for extra bb system)
  - background probability
  - hypothesis of missing parton
  - better understanding of b-tagging probability term and try to find unified strategy to achieve optimal performance in all selection regions

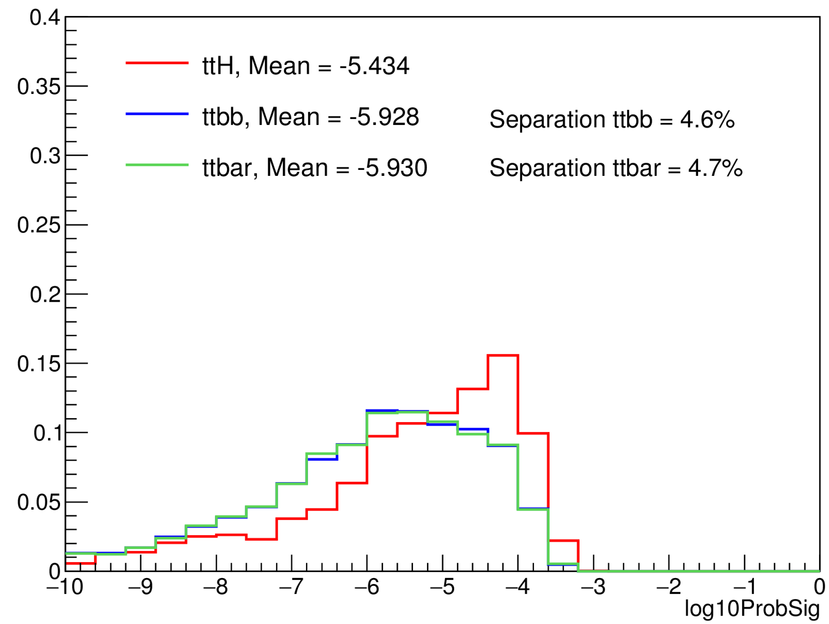
Signal probability

The signal probabilities distributions are as expected: the signal is more likely to be signal than background

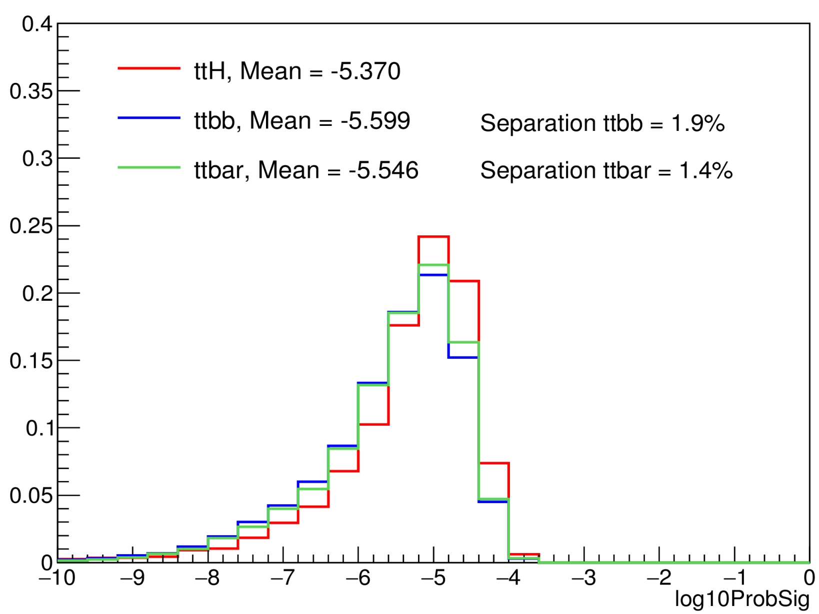
6jex3bex



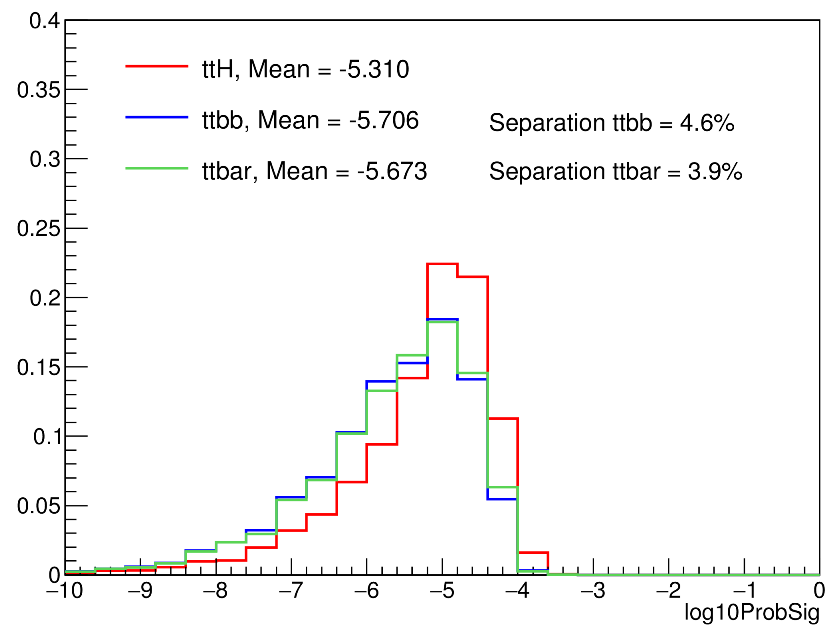
6jex4bin



7jin3bex



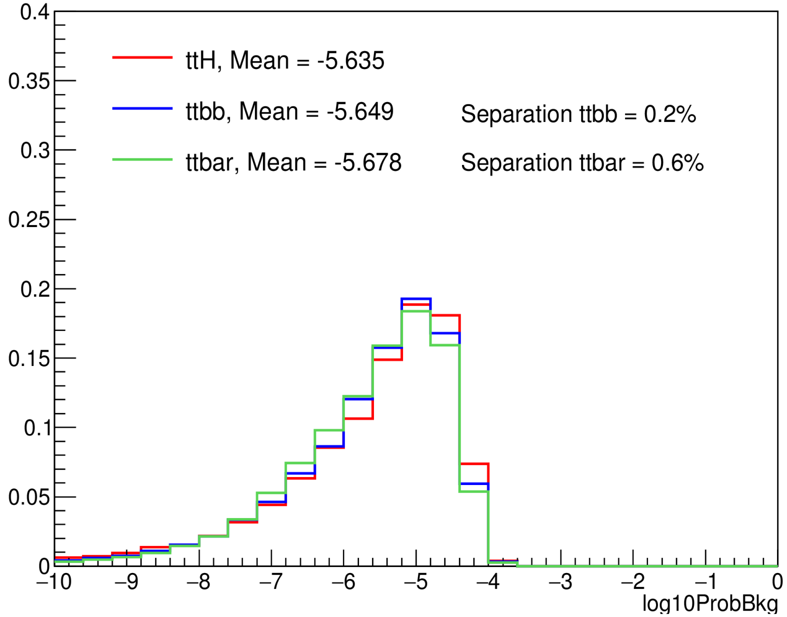
7jin4bin



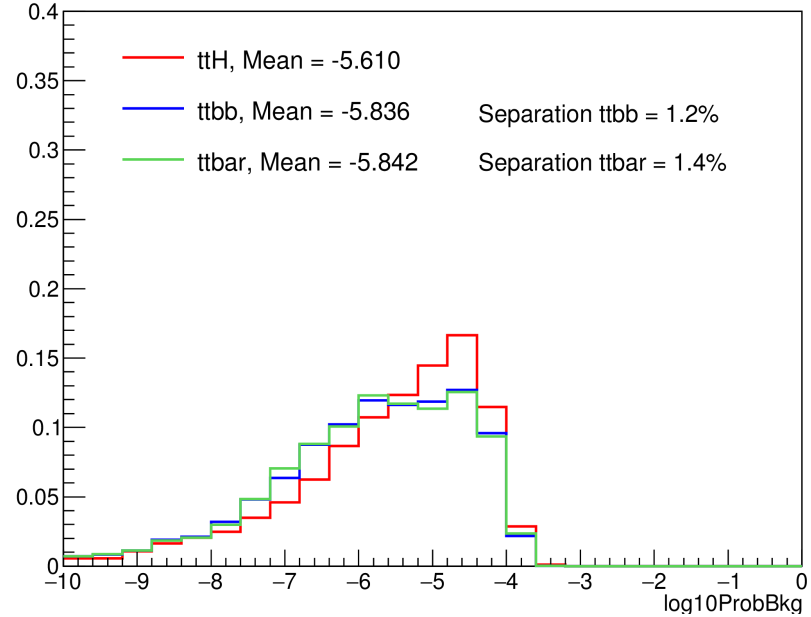
Background probability

The bkg probabilities are not separating signal from bkg correctly: signal is even more likely to be bkg than bkg → smth for further optimisation (probably will add angular variables)

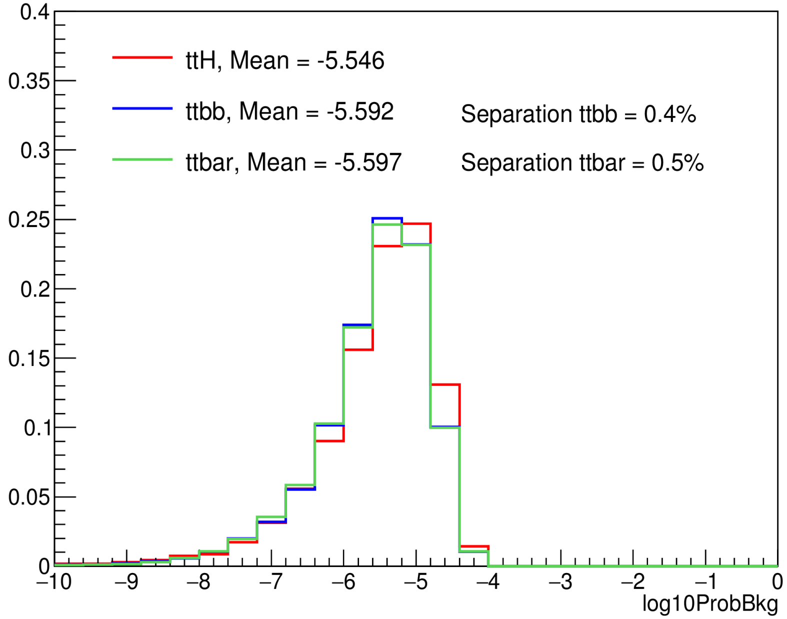
6jex3bex



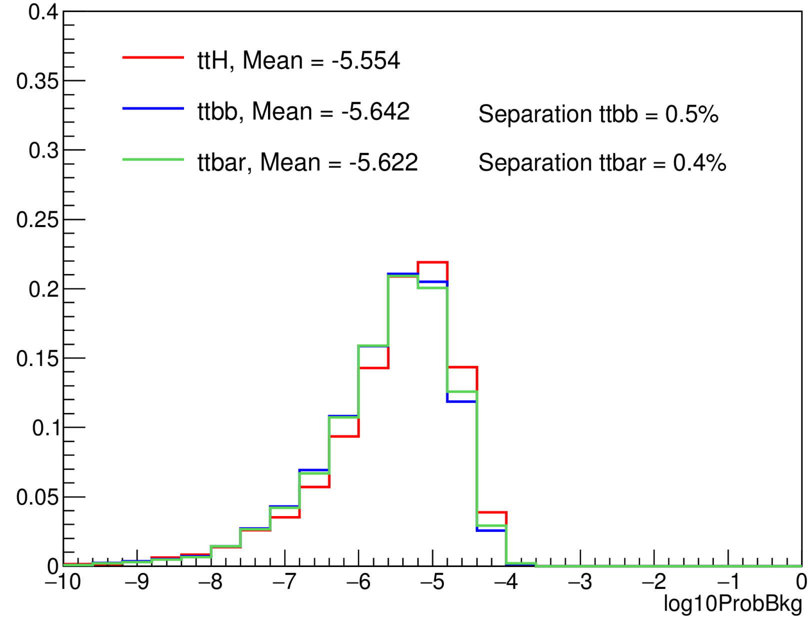
6jex4bin



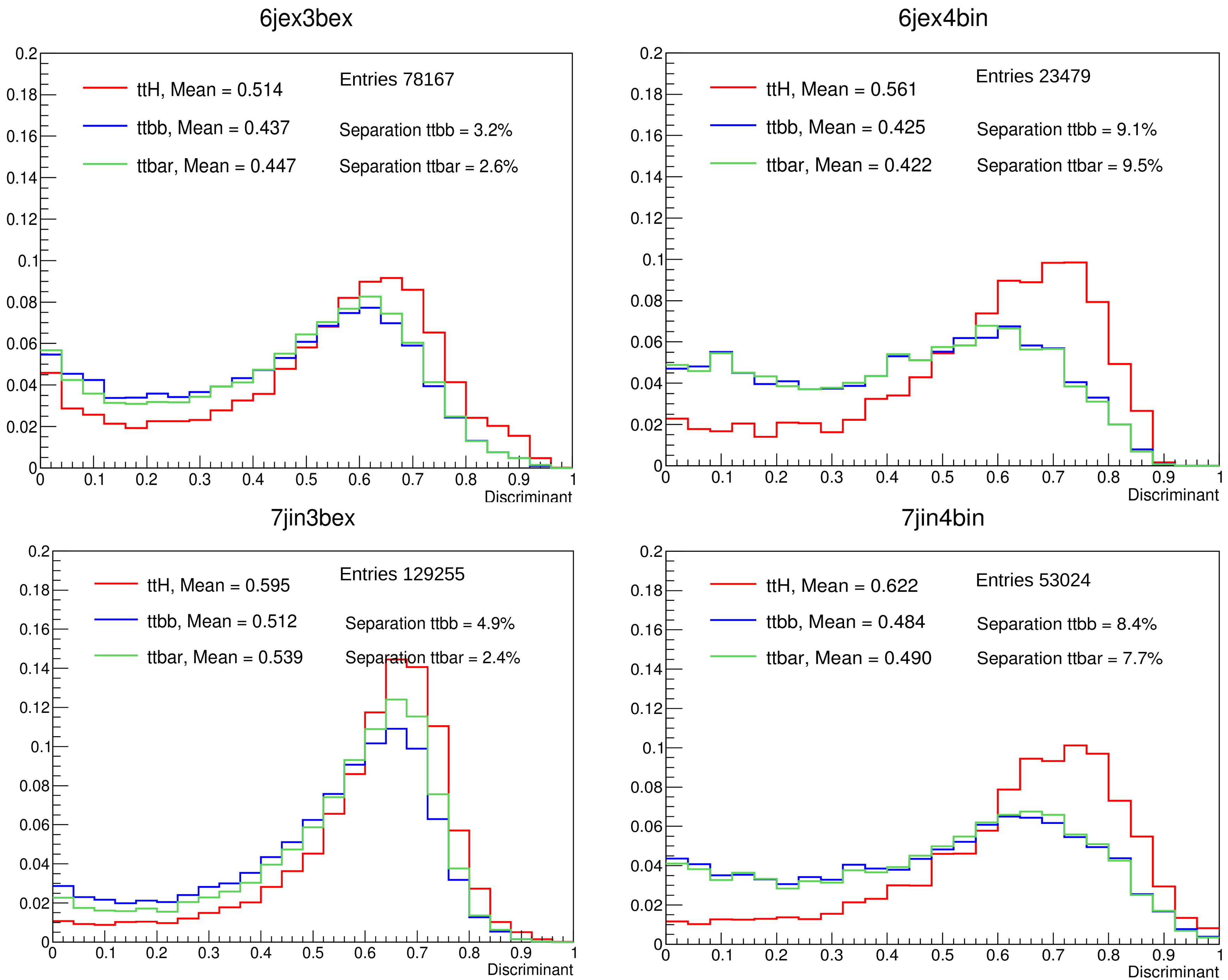
7jin3bex



7jin4bin



Final discriminant



b-tagging

For the first year of my PhD I was working on b-tagging
Goal of b-tagging: identify jets stemming from hadronisation of b-quarks
B-tagging used in many analysis domains:

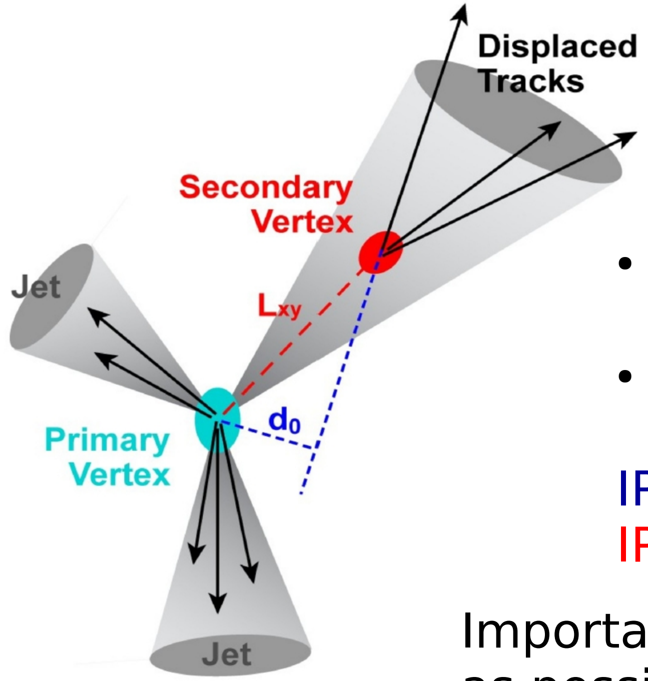
- Top physics
- Higgs physics
- Beyond Standard Model physics

b-tagging algorithms rely on b-hadron properties:

- High mass (~5 GeV)
- Relatively long lifetime (~1.5 ps)

b-tagging algorithms:

- Spatial tagging
  - Impact parameter based: IP2D, IP3D
  - Secondary vertex based: SV0, SV1, JetFitter
  - Multivariate: MV1, MV2
- Soft lepton tagging



Impact parameter based taggers

- Use lifetime signed impact parameter (IP) significance of tracks matched to a jet
- IP - distance of closest approach to the primary vertex (PV)

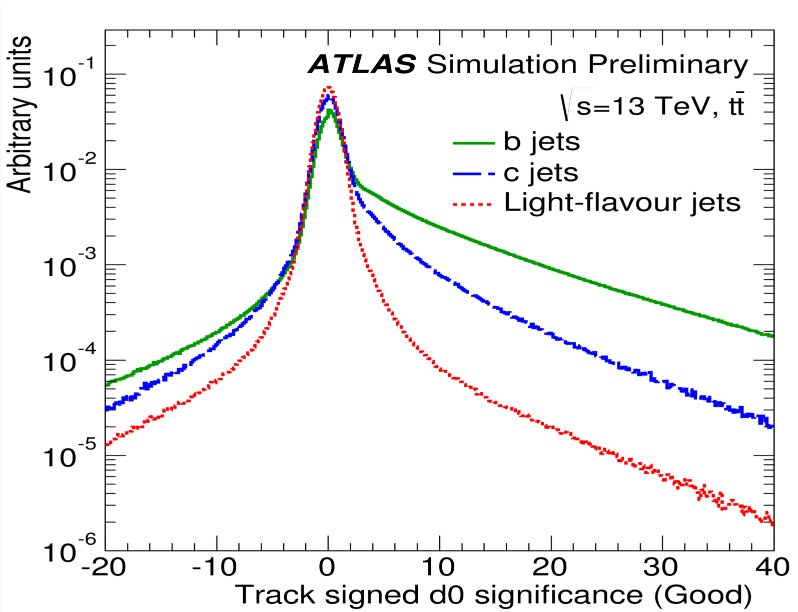
IP2D: only use transverse IP in 1D likelihood
IP3D: combine with longitudinal IP in 2D likelihood

Important to measure particle as close to interaction point as possible

Algorithm optimisation for Run 2

Major ATLAS inner detector upgrade for LHC Run-2: addition of Insertable B-layer (IBL) as fourth innermost pixel layer

The key goal of IP3D algorithm optimization - to take advantage of the new layer

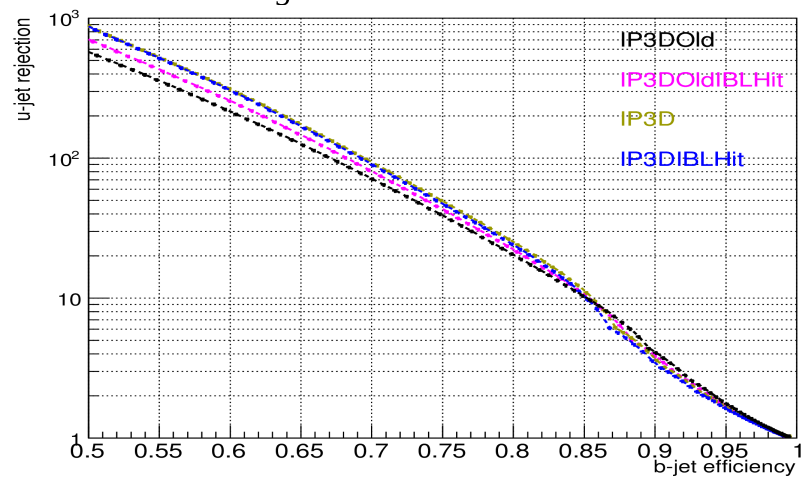


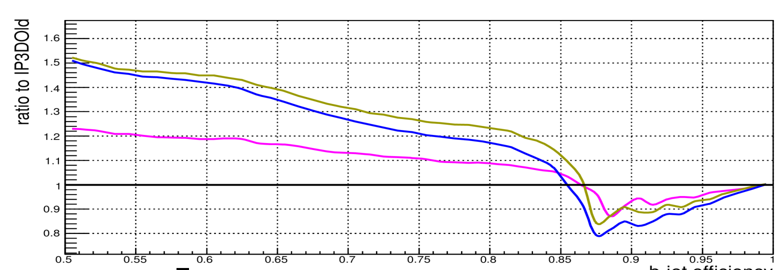
Some tracks give worse resolution:

- those with missing hits in some tracker layers
- with ambiguities in pattern recognition

What we need – divide track into categories and compute IP3D weight for each of them separately:

$$w_{track} = \frac{p_b}{p_{light}}; \qquad w_{jet} = \sum_{tracks} \log\left(w_{track}\right).$$





$t\bar{t}$ events

New track grading for Run-2 conditions was developed, taking into consideration new track variables available: 14 track categories instead of 5 used in Run 1

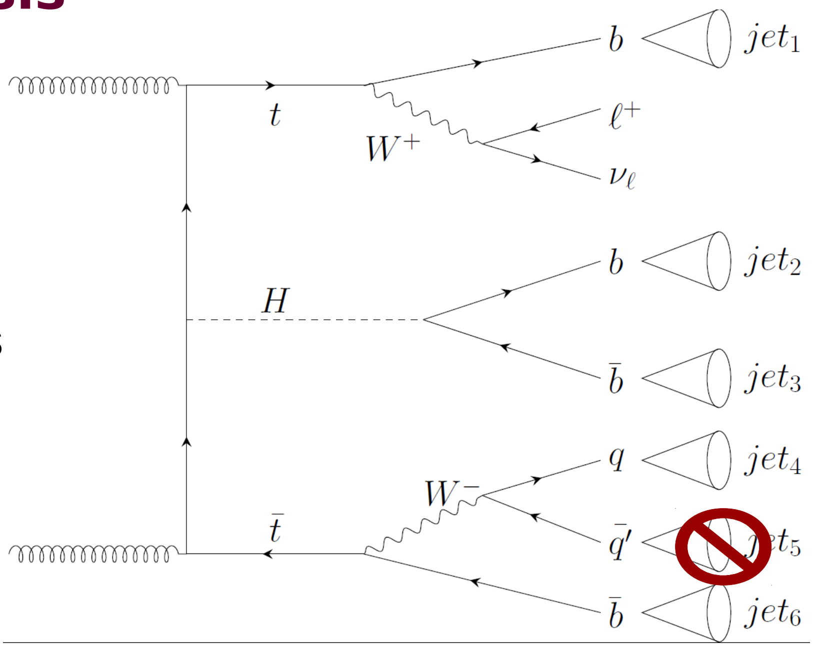
IP3D versions with new grading give better performance than with old grading for up to ~86 efficiency region

Signal hypothesis

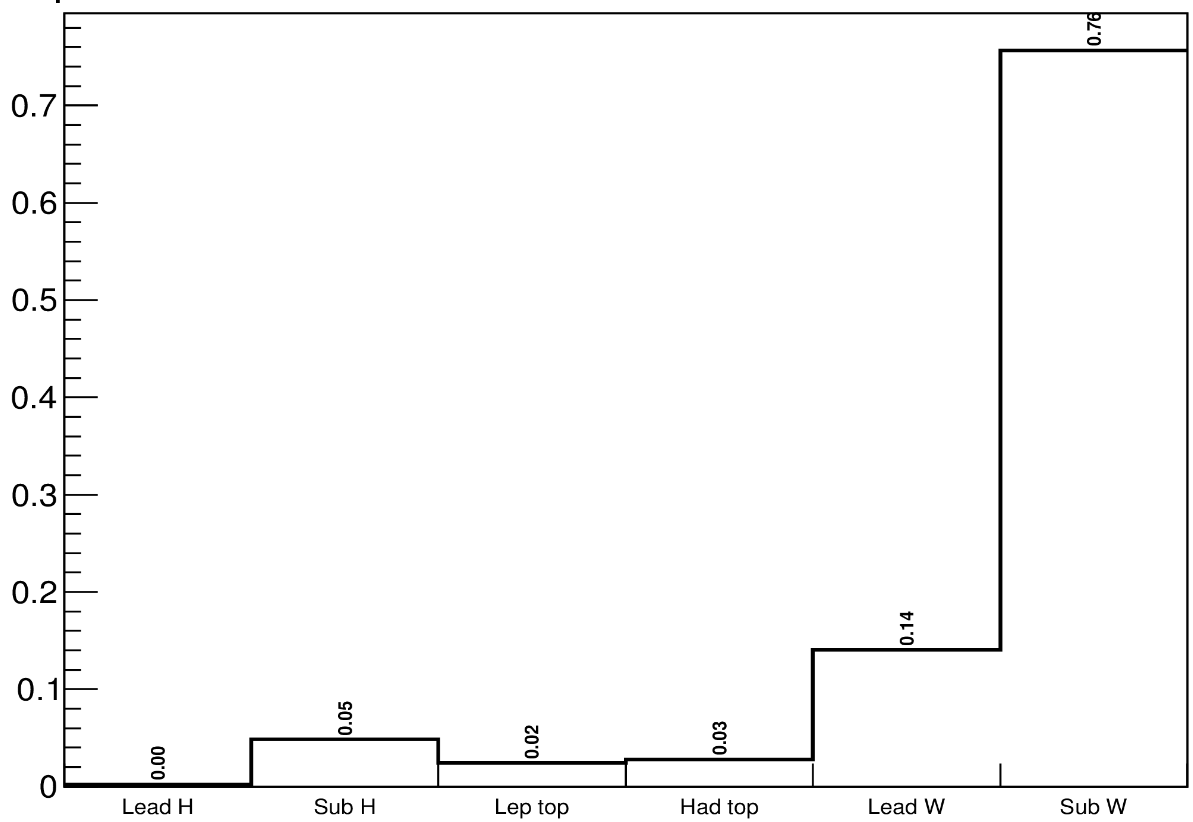
Two hypotheses:
1) all 6 partons are matched to jets
2) 5 partons matched, one parton is missing

- For now all results obtained with only 1)

- In >=6 jets, >= 4 b-jets (main signal region) all partons are matched for 40.3% events, while in 45.8% one jet missing → room for further optimisation



In 6jin4jin with one missing parton in 90% cases it is quark from W:

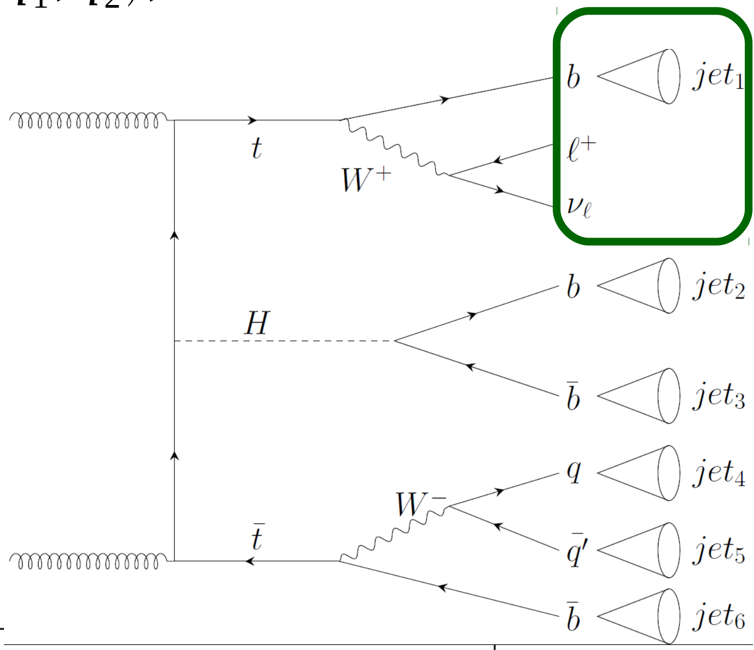


	All matched	Missing parton
6jex4bin	30.9%	52.8%
6jin4bin	40.3%	45.8%
7jin4bin	46.1%	41.5%
6jex3bex	15.5%	47.5%
6jin3bex	22.3%	45.4%
7jin3bex	27.8%	43.7%

Kinematic probability: leptonic top mass

$$P_{kin}^{sig}(x)=\boxed{P_{M_{l\nu b_l}}^{sig}(l,\nu,b_l)}\,P_{M_{b_1 b_2}}^{sig}(b_1,b_2)\,P_{M_{q_1 q_2 b_h}}^{sig}(q_1,q_2,b_h)\,P_{M_{q_1 q_2}}^{sig}(q_1,q_2),$$

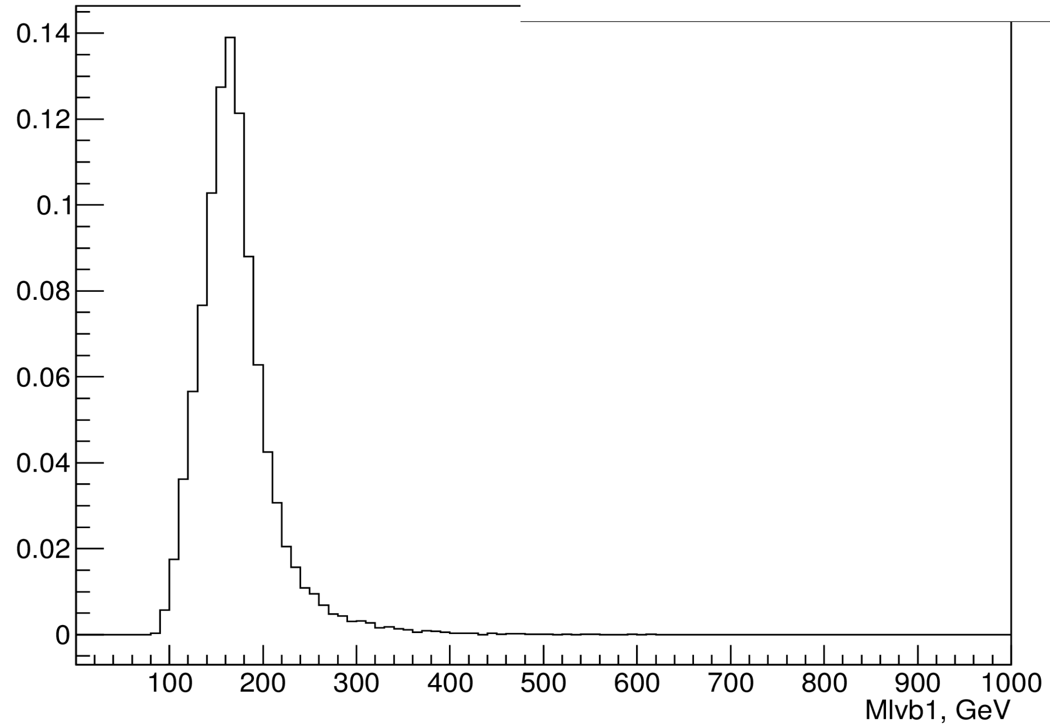
- reconstructed leptonic top mass PDF



- Neutrino 4-vector – solution of the quadratic equation:

$$M_W^2=(p_l+p_\nu)^2$$

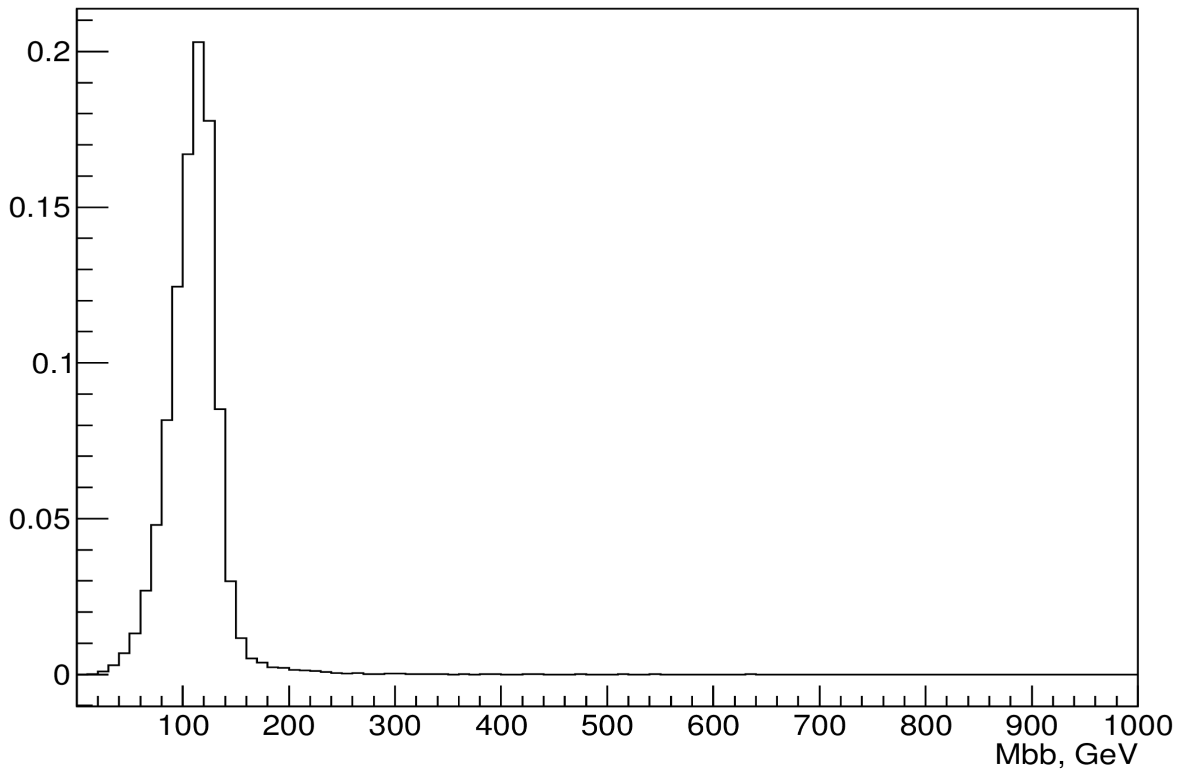
- In the case of two solutions (~66% of events) of the equation both are considered with appropriate weights (fractions of cases when each of the solutions is closer to the truth neutrino)
- In the case of no solution (~34% of events) we vary ETmiss till discriminant of the equation is zero and 1 solution exists

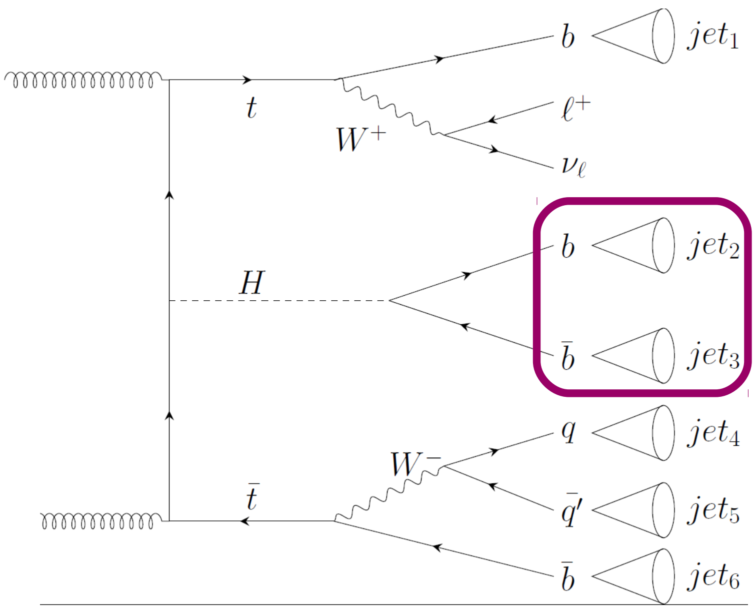


Kinematic probability: Higgs mass

$$P_{kin}^{sig}(x)=P_{M_{l\nu b_l}}^{sig}(l,\nu,b_l)\boxed{P_{M_{b_1 b_2}}^{sig}(b_1,b_2)}P_{M_{q_1 q_2 b_h}}^{sig}(q_1,q_2,b_h)P_{M_{q_1 q_2}}^{sig}(q_1,q_2),$$

reconstructed Higgs mass PDF

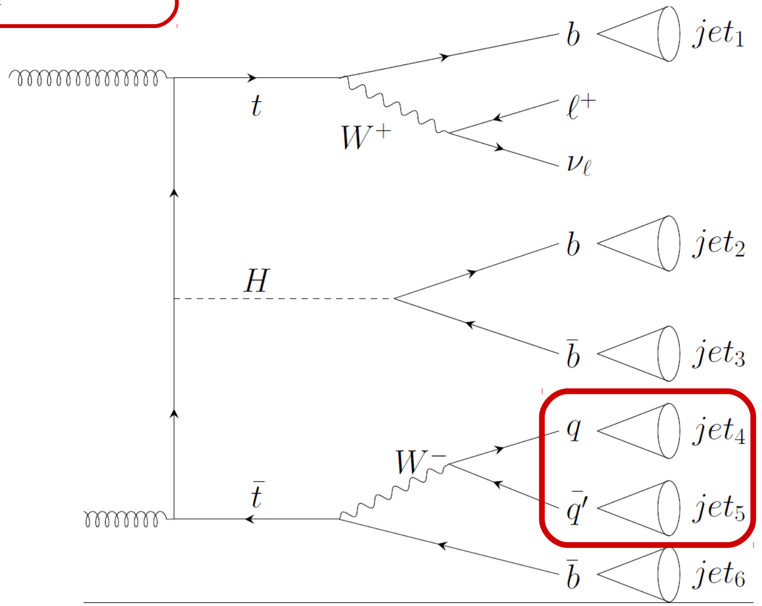


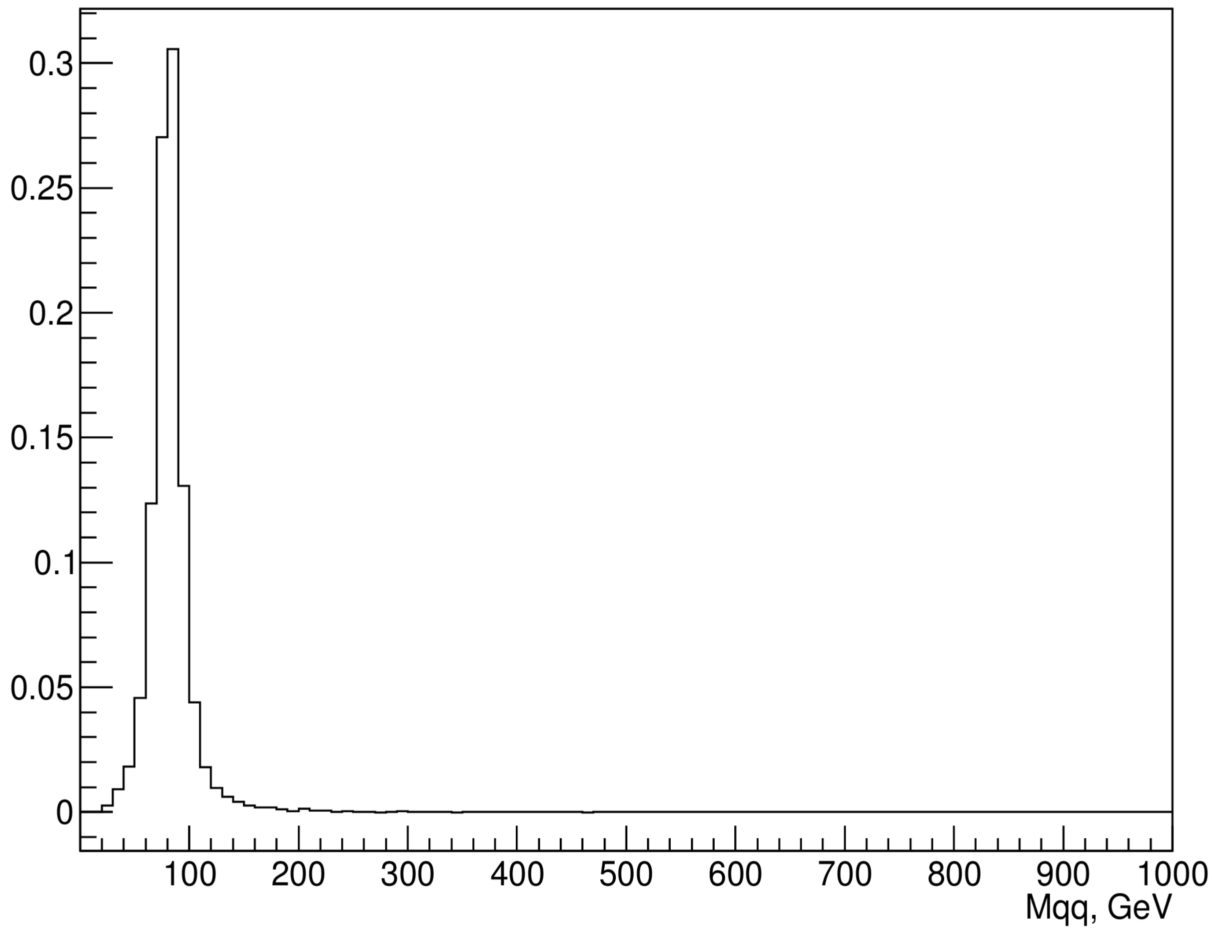


Kinematic probability: hadronic W mass

$$P_{kin}^{sig}(x)=P_{M_{l\nu b_l}}^{sig}(l,\nu,b_l)\,P_{M_{b_1 b_2}}^{sig}(b_1,b_2)\,P_{M_{q_1 q_2 b_h}}^{sig}(q_1,q_2,b_h)\boxed{P_{M_{q_1 q_2}}^{sig}(q_1,q_2)},$$

reconstructed hadronic W mass PDF

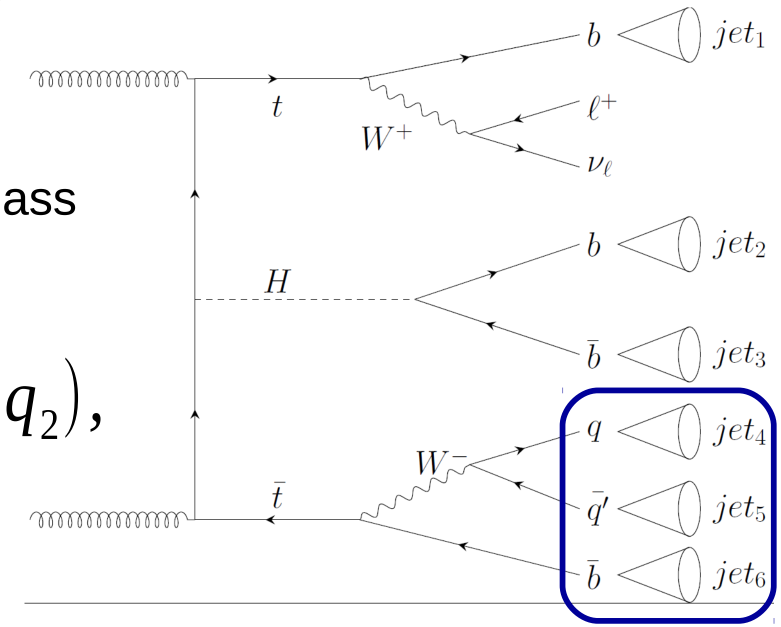




Kinematic probability: hadronic top mass

$$P^{sig}_{kin}(x) = P^{sig}_{M_{l\nu b_l}}(l,\nu,b_l)\, P^{sig}_{M_{b_1 b_2}}(b_1,b_2)\, \boxed{P^{sig}_{M_{q_1 q_2 b_h}}(q_1,q_2,b_h)}\, P^{sig}_{M_{q_1 q_2}}(q_1,q_2),$$

reconstructed hadronic top mass

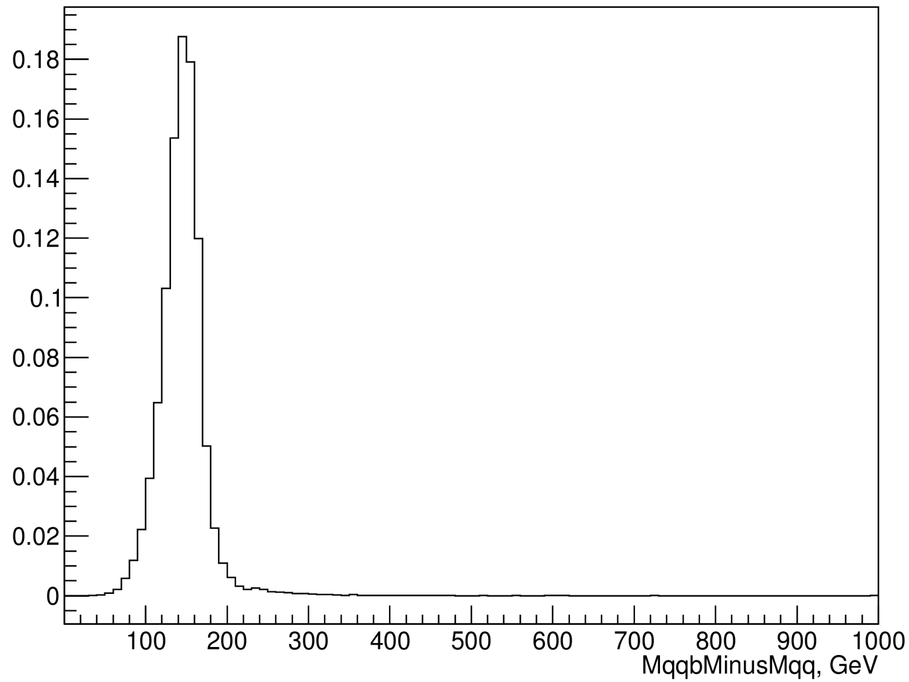


To take into account correlation between hadronic top mass and W mass using instead

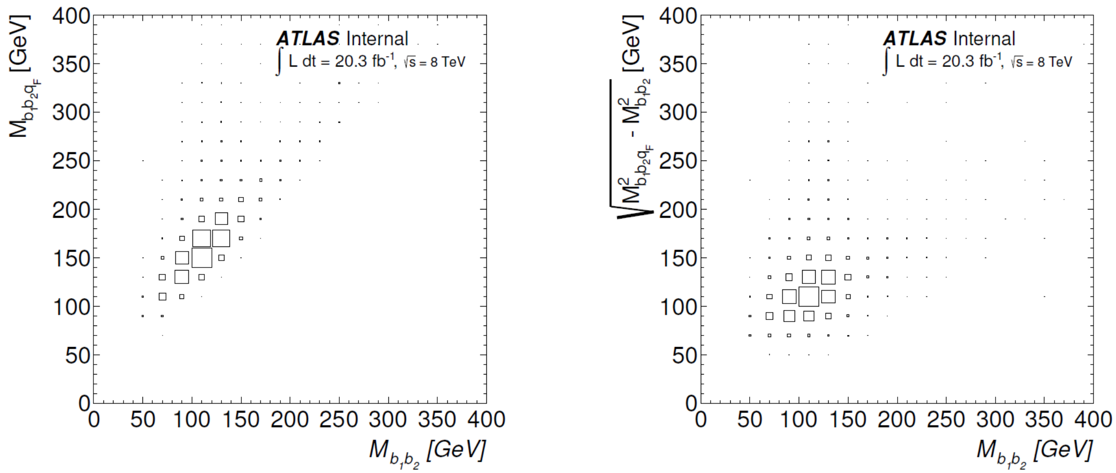
$$P^{sig}_{X_{q_1 q_2 b_h}}(q_1,q_2,b_h),$$

$$X_{q_1 q_2 b_h}(q_1,q_2,b_h) = M_{q_1 q_2 b_h}(q_1,q_2,b_h) \ominus M_{q_1 q_2}(q_1,q_2),$$

where $a \ominus b = \sqrt{a^2 - b^2}$



Correlation between hadronic top mass and H mass and X and H mass in FCNC t->Hq channel



From ATL-COM-PHYS-2014-809, https://cds.cern.ch/record/1742309

Background probability

- Background probability is calculated in similar way to signal probability:

$$P^{bkg}(x)=\frac{\sum_{k=1}^{N_p} P_{kin}^{bkg,bb}(x^k)\,P_{btag}^{bkg,bb}(x^k)+P_{kin}^{bkg,bj}(x^k)\,P_{btag}^{bkg,bj}(x^k)}{\sum_{k=1}^{N_p} P_{btag}^{bkg,bb}(x^k)+P_{btag}^{bkg,bj}(x^k)}$$

These terms are different from Psig

$$P_{kin}^{bkg,bb}(x)=P_{M_{l\nu b_l}}^{bkg}(l,\nu,b_l)\,P_{M_{q_1q_2b_h}}^{bkg}(q_1,q_2,b_h)\,P_{M_{q_1q_2}}^{bkg}(q_1,q_2)\,\boxed{f_{bb}\,P_{M_{b_1b_2}}^{bkg}(b_1,b_2)},$$

$$P_{kin}^{bkg,bj}(x)=P_{M_{l\nu b_l}}^{bkg}(l,\nu,b_l)\,P_{M_{q_1q_2b_h}}^{bkg}(q_1,q_2,b_h)\,P_{M_{q_1q_2}}^{bkg}(q_1,q_2)\,\boxed{f_{bj}\,P_{M_{bj}}^{bkg}(b,j)},$$

$$P_{btag}^{bkg,bb}(x)=P_b(j1)\,P_b(j2)\,P_b(j3)\,P_{light}(j4)\,(f_{light-light}\,P_{light}(j5)+f_{c-light}\,P_c(j5))\,P_b(j6),$$

$$P_{btag}^{bkg,bj}(x)=P_b(j1)\,P_b(j2)\,(f_c\,P_c(j3)+f_{light}\,P_{light}(j3))\,P_{light}(j4)\,(f_{light-light}\,P_{light}(j5)+f_{c-light}\,P_c(j5))\,P_b(j6).$$

Optimization of b-tagging probability calculation

$$P^{sig}_{btag}(x)=P_b(j1)\cdot P_b(j2)\cdot P_b(j3)\cdot P_{light}(j4)\cdot(f_{light-light}P_{light}(j5)+f_{c-light}P_c(j5))\cdot P_b(j6)$$

$P_f(j_i)$ - probability that a jet i with its given b-tagging weight to be originating of parton with flavour f

The problem is combinatorial background, especially for 7jin4bin case

Original approach of b-tagging probability calculation (btag1):

$$P_f(j_i)=eff_f(w_{jet_i})$$ at 70% OP if jet is tagged

$$P_f(j_i)=1-eff_f(w_{jet_i})$$ if not tagged

Alternative approach (btag2):

$$P_b(j_i)=1$$ if jet is tagged

$$P_b(j_i)=0$$ if not tagged

(P_c, P_{light} are not changed)

This strategy is preventing non-b jets being put in p positions and perform better in the extreme case 7jin4bin, with more combinatorial background

Final discriminant separation values for two approaches (ttbb):

	btag1	btag2
6jex4bin	9.1%	8.6%
7jin4bin	8.4%	10.2%

Decided to use btag1 for all regions except 7jin4bin, where btag2 will be used

Analysis regions