Search for $t\bar{t}(H\rightarrow b\bar{b})$ using the Matrix Element Method with ATLAS at 7 TeV

Olaf Nackenhorst

supervised by K. Kröninger, A. Quadt, E. Shabalina

University of Göttingen

Motivation, Strategy & Introduction

Discrimination Studies

NN Studies & First Glance at Limits

Summary & Outlook

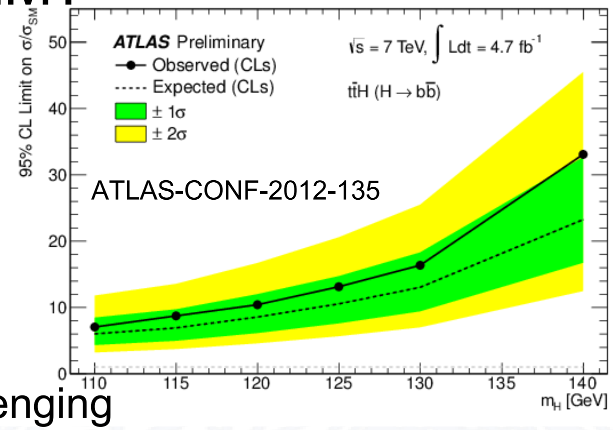
GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

PHYSICS AT THE TERA SCALE
Helmholtz-Alliance

FSP 101 ATLAS

Motivation

- Search for a Higgs in $t\bar{t}H \rightarrow$ l+jets+$b\bar{b}$ with MEM?
  - In $t\bar{t}H$ channel no sign yet for a Higgs
  - Direct top-Higgs Yukawa coupling measurement
    - Any deviations from SM hint for new physics
  - Small signal (σ ~ 90 fb), large background
  - $t\bar{t}b\bar{b}$ similar kinematic signature after selection
    - Event reconstruction / jet-to-parton assignment challenging
    - Not very powerful discriminating variables for MVA



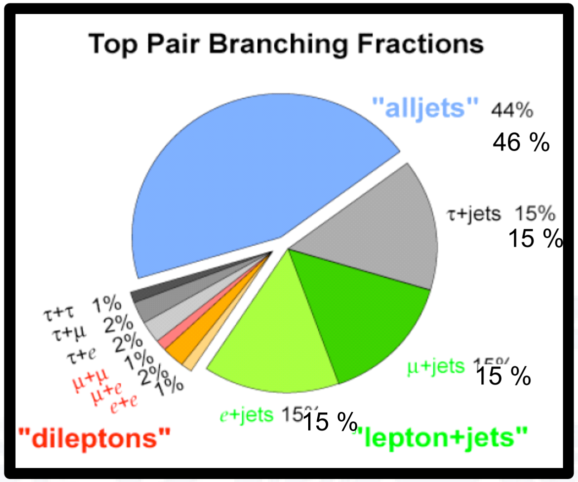
- Goal:
  - Use MEM in order to construct a **Higgs mass dependent discriminator**
  - Likelihood ratio most powerful discriminant for separating BG from signal:
    - Neyman-Pearson:

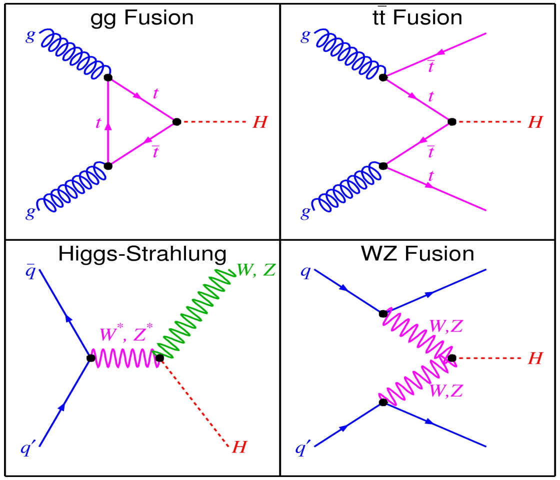
$$r_{\text{sig}}\left(\vec{x}|\vec{a}\right) = \frac{P_{\text{sig}}\left(\vec{x}|\vec{a}\right)}{\sum_{\text{bkg}} f_{\text{bkg}} P_{\text{bkg}}\left(\vec{x}|\vec{a}\right)}$$

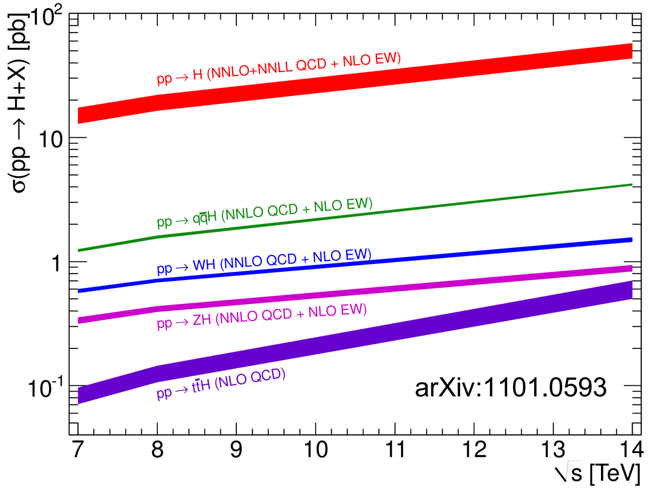
  - Discriminator can then be **further used in current MVA or standalone**

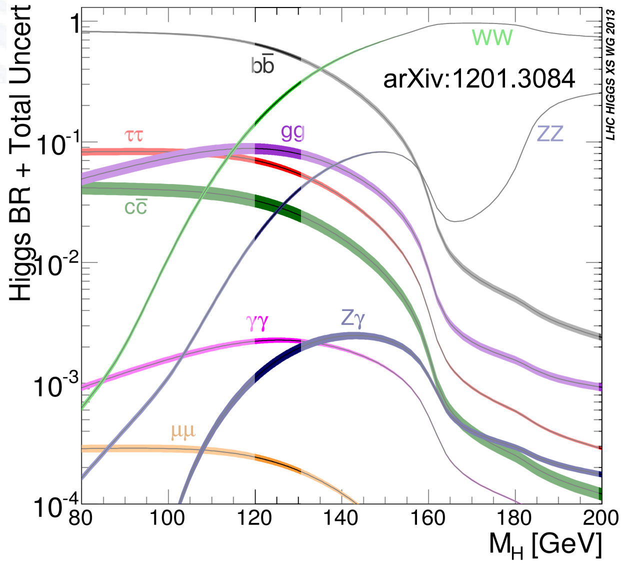
Top and Higgs Physics

- Higgs boson production at the LHC
  - Four production mechanism
  - Dominated by gluon fusion, ttH suppressed
- Higgs boson decay
  - Dominated by $b\bar{b}$ (not yet observed) and WW
- Top quark decay
  - Almost 100 % to b-quark and W-boson
  - final states classified according to W decay







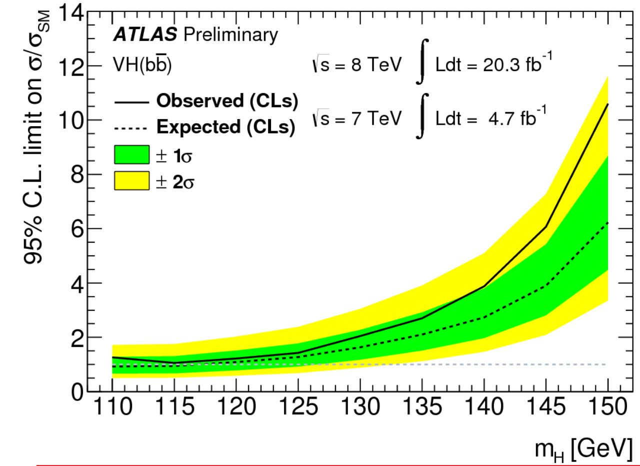


Search for H→$b\bar{b}$ in VH

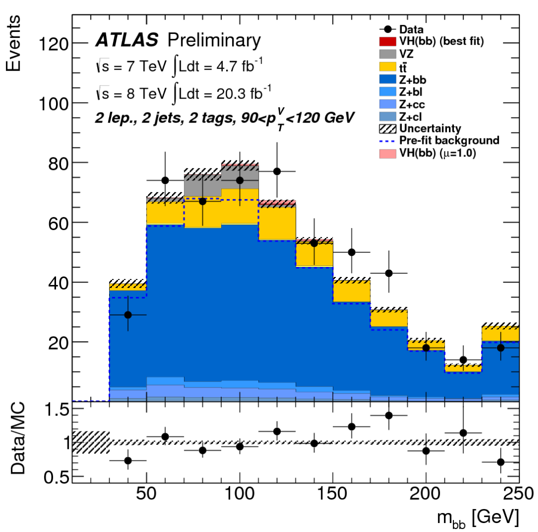
Analysis Outline

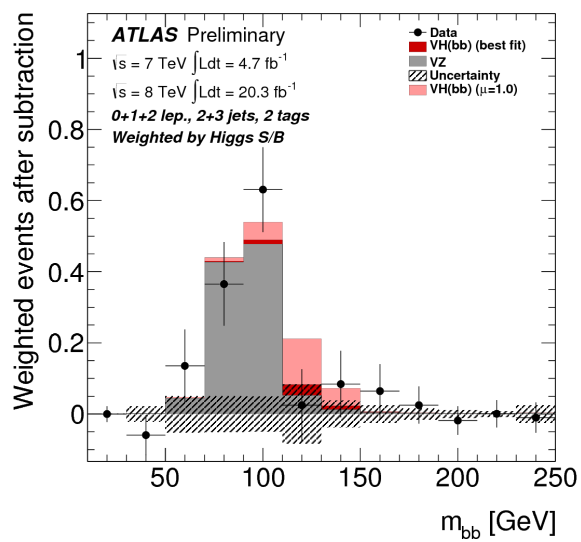
ATLAS-CONF-2013-079

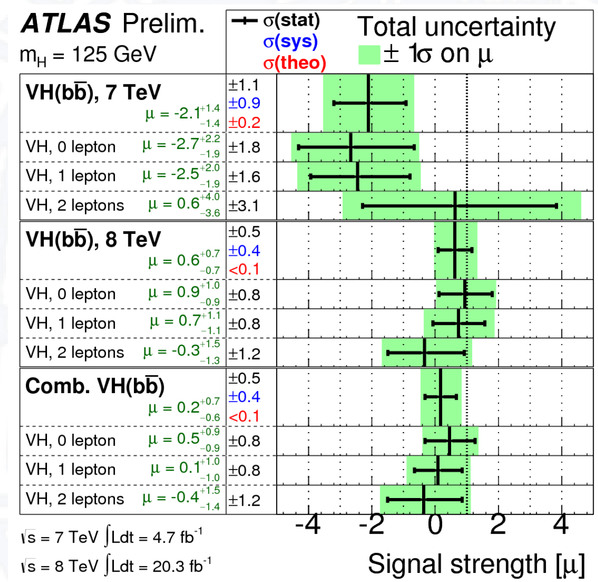
- Full 7 and 8 TeV data (4.7 fb^{-1} / 20.3 fb^{-1})
- Leptonic W/Z decays: 0/1/2 leptons in final state
- Signal (2 b-tags) and control regions (0/1 b-tags)
- Split in sub channels: # lepton, # jets, p_T^V interval
- Calculate di-jet mass: b-tags and highest p_T
- Perform simultaneous profile likelihood fit
- Main uncertainties: $t\bar{t}$ modeling, c-tagging, multijet



$\mu = 0.2 \pm 0.5$ (stat) ± 0.4 (syst)

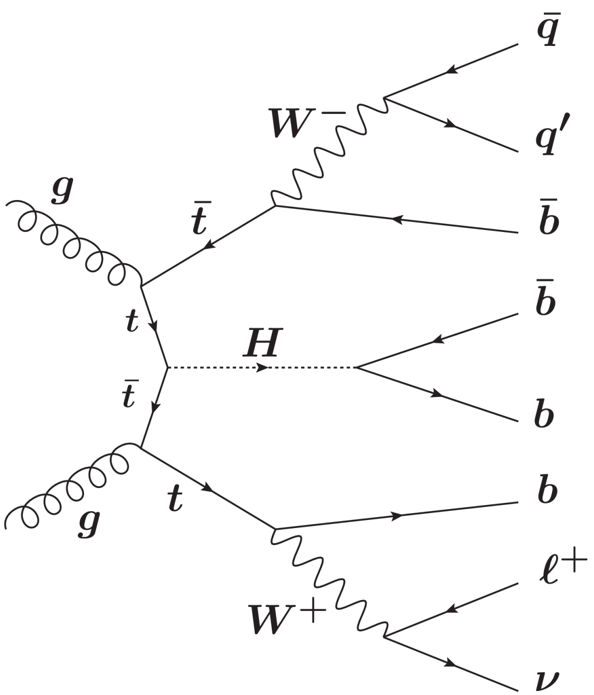


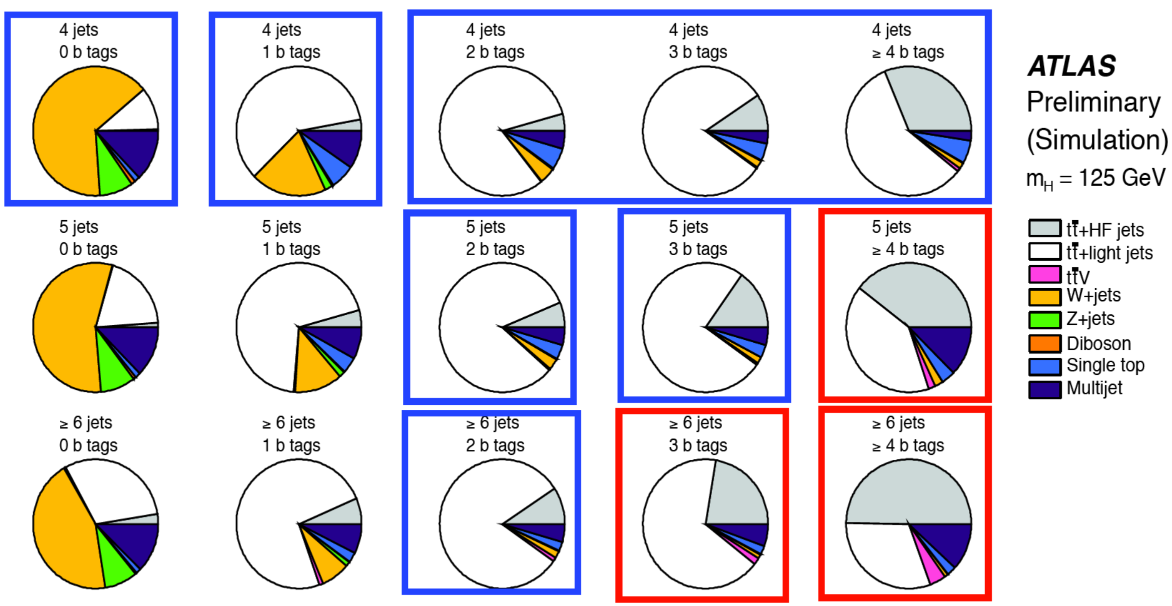




Overview tt(H→bb) l+jets

- Basic Strategy for 7 and 8 TeV
  - Selection: similar to standard $t\bar{t}$ selection (symmetric in e/mu)
  - Split in sub channels dependent on jet and b-tag multiplicity
  - Simultaneous profile likelihood fit in all channels
    - H_T in BG regions to constrain systematics
    - NN output in signal regions to extract signal

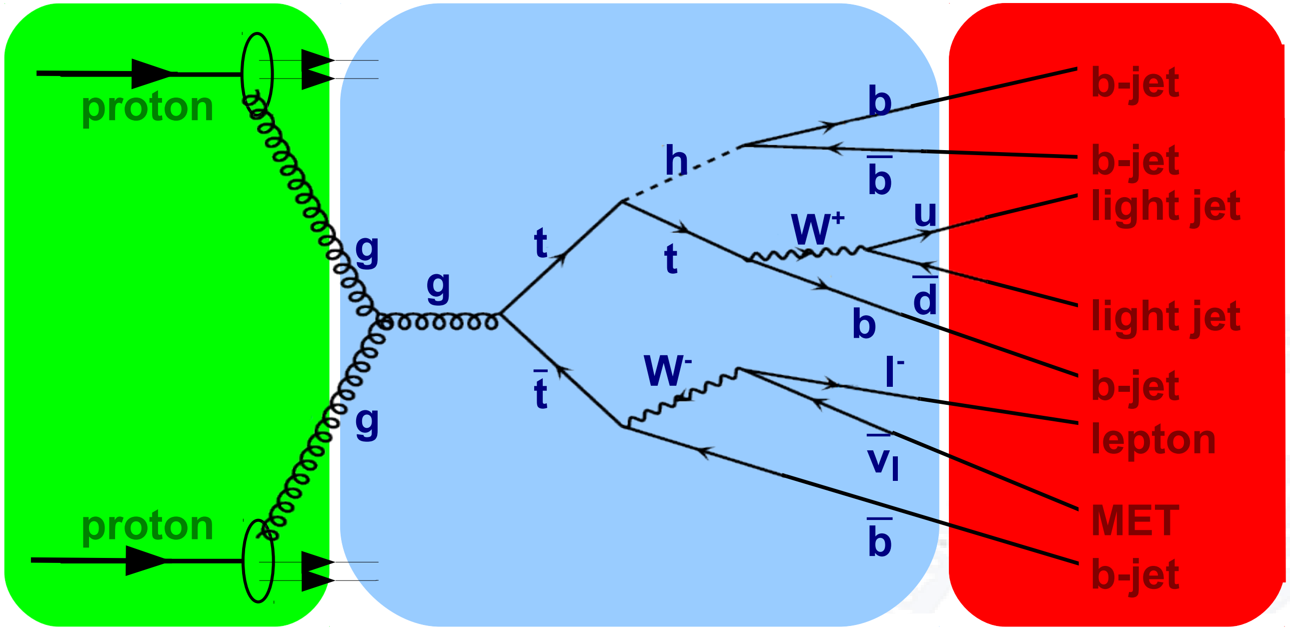




S = signal regions B = background regions

ATLAS-CONF-2012-135

Matrix Element Method

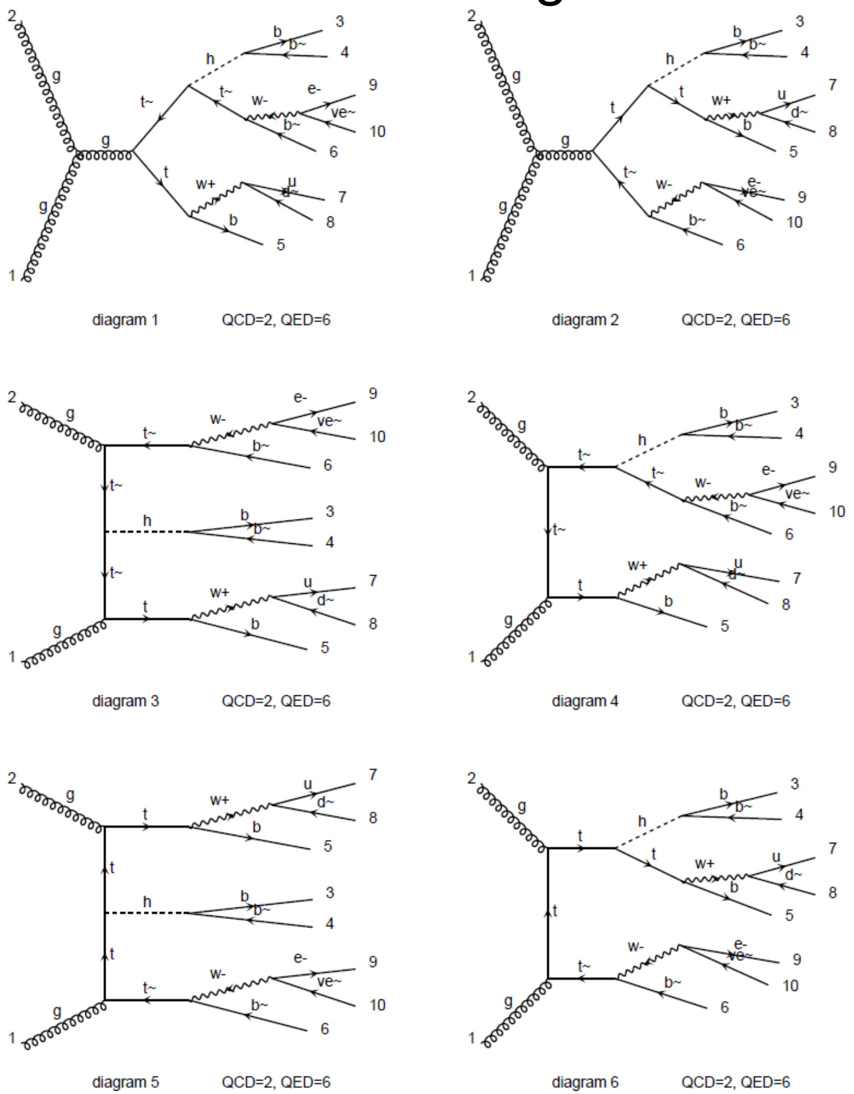


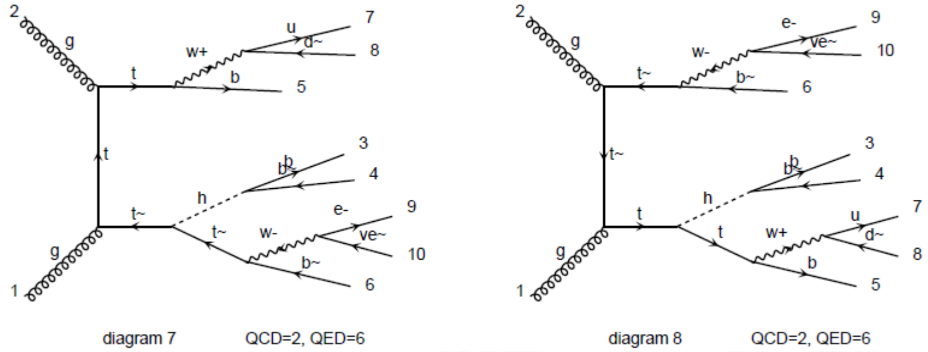
$$\underbrace{P_{t\bar{t}H}\left(\vec{x}_{\text{Detector}}, m_H\right)}_{\text{probability}} = \underbrace{\frac{1}{\sigma_{t\bar{t}H}\left(m_H\right)}}_{\text{normalization}} \int \underbrace{\mathrm{d}p_{g1}\mathrm{d}p_{g2} f\left(p_{g1}\right) f\left(p_{g2}\right)}_{\text{parton density function}} \underbrace{\mathrm{d}\sigma_{t\bar{t}H}\left(\tilde{x}_{\text{Parton}}, m_H\right)}_{\text{differential cross section}} \underbrace{W\left(\vec{x}_{\text{Parton}}, \vec{x}_{\text{Detector}}\right)}_{\text{transfer functions}}$$

- Obtain event probability to be a $t\bar{t}H$ decay
  - Normalized by total xsec (considering efficiency & acceptance)
  - PDFs account for production mechanism
  - Differential cross section proportional to $|M|^2$, consider only LO ME
  - Transfer functions map detector response to parton level

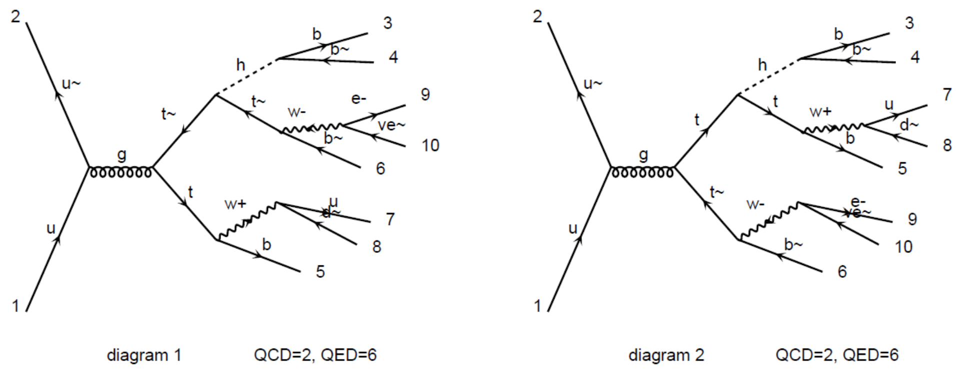
Looking at Signal Diagrams

Gluon Fusion: 8 Diagrams





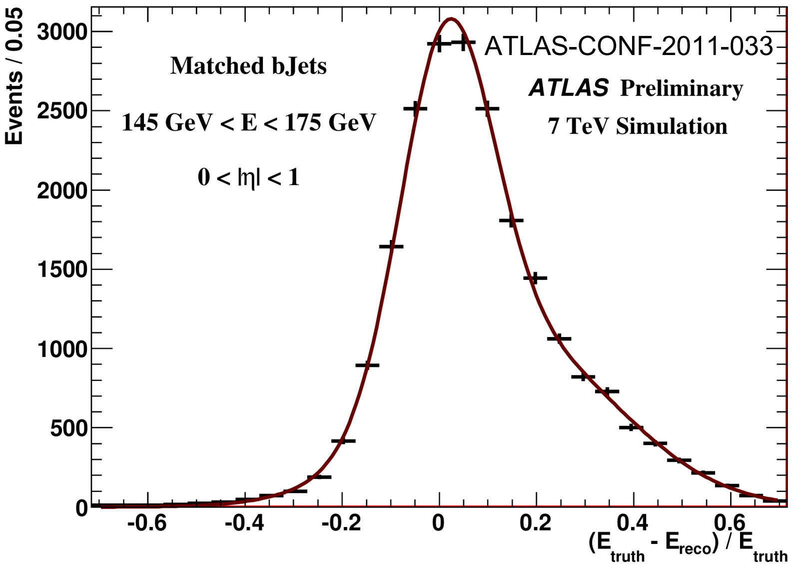
$q\bar{q}$-Annihilation: 2 Diagrams



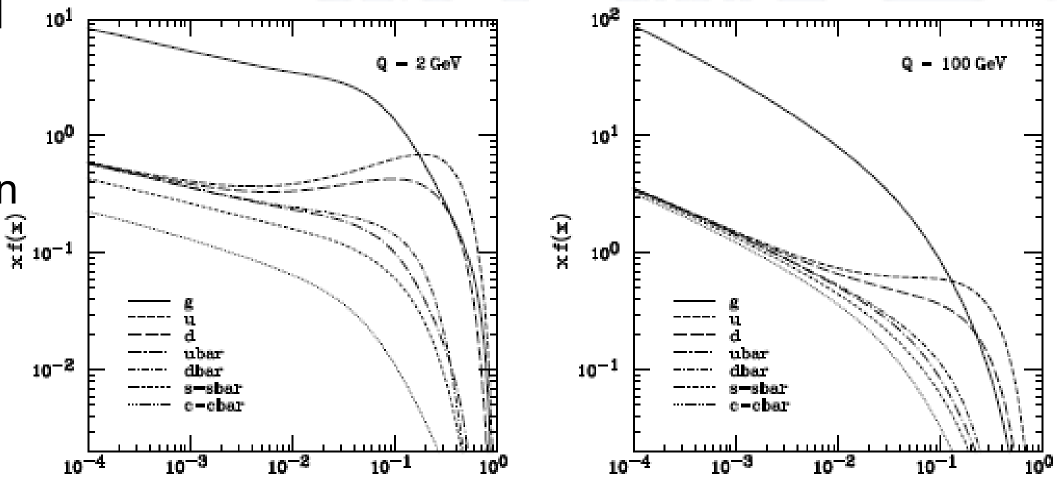
in $t\bar{t}H$: $gg/q\bar{q}$ production ~ 50/50!

Transfer Functions and PDFs

- Transfer Functions W (E_{det} | E_{parton})
  - map detector response to parton level object
  - double-Gaussian parameterization in
    - different η regions (detector geometry)
    - E intervals
  - → take asymmetric tails into account
  - Separately for e, μ, light and b-jets

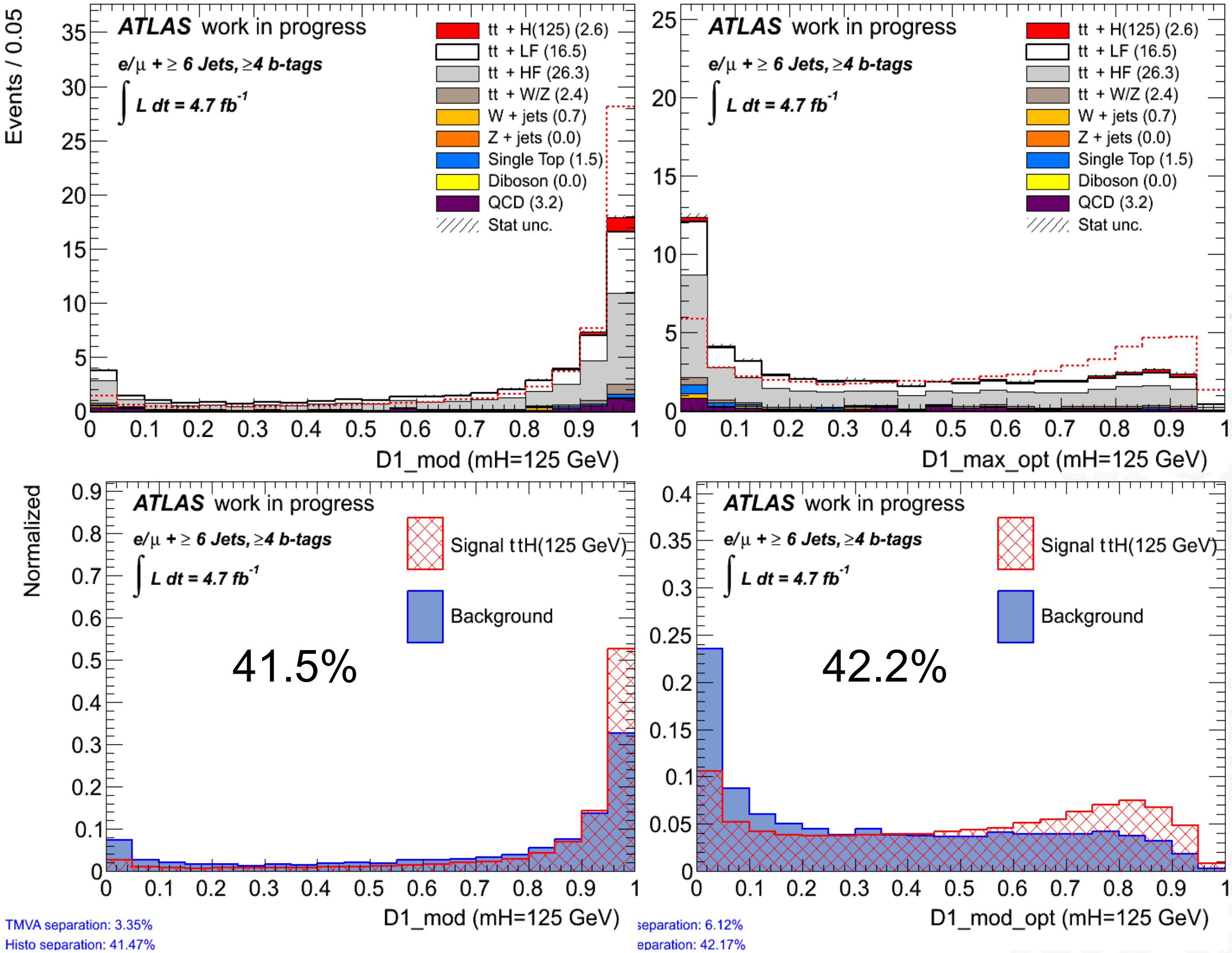


- Parton Density Function CTEQ6M
  - Energy distribution of hadronic content of proton
  - Probability density to find a parton
    - at momentum fraction x
    - at energy scale Q
  - small x → gluons dominate
  - large x → light quarks dominate



arXiv:hep-ph/0201195

Neyman-Pearson Likelihood ratio for different normalizations



$$\mathcal{D}_i^1 = \frac{L\left(t\bar{t}H|y_i\right)}{L\left(t\bar{t}H|y_i\right) + \alpha L\left(t\bar{t}b\bar{b}|y_i\right)}$$

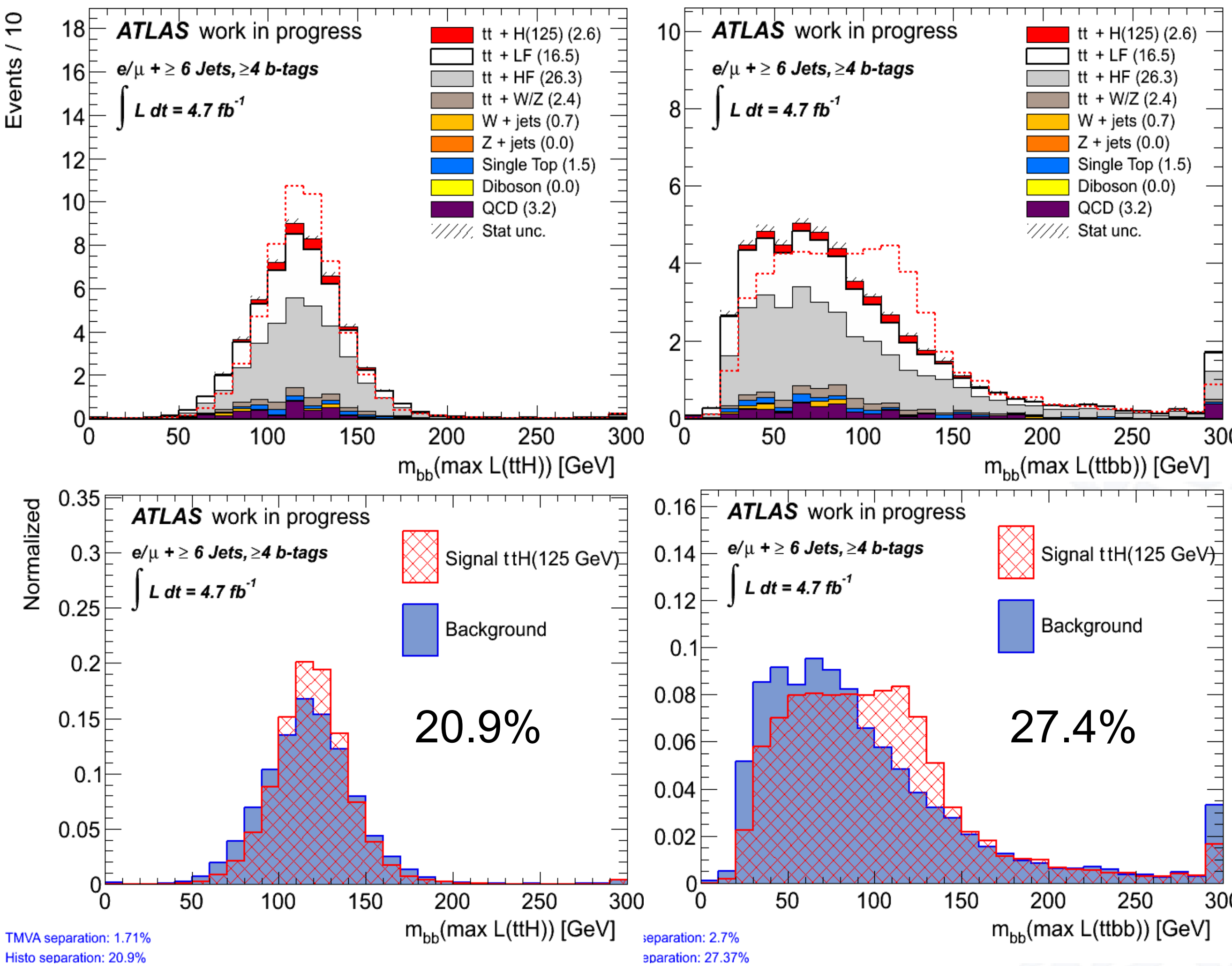
MC composition

- Xsec norm (left)
  - Most $t\bar{t}H$ in signal-like bin ~1
  - $t\bar{t}/V$ looks like signal
- Max sep. norm (right)
  - Better ID of BG
  - Substructure

Separation

- Similar and great discrimination in both

Invariant $m_{b\bar{b}}$ of non $t\bar{t}$ b-jets in best permutation



Max-L ($t\bar{t}H$)

- Forced to find two jets which mimic H mass
- Signal has slightly sharper peak
- Correlation to D1

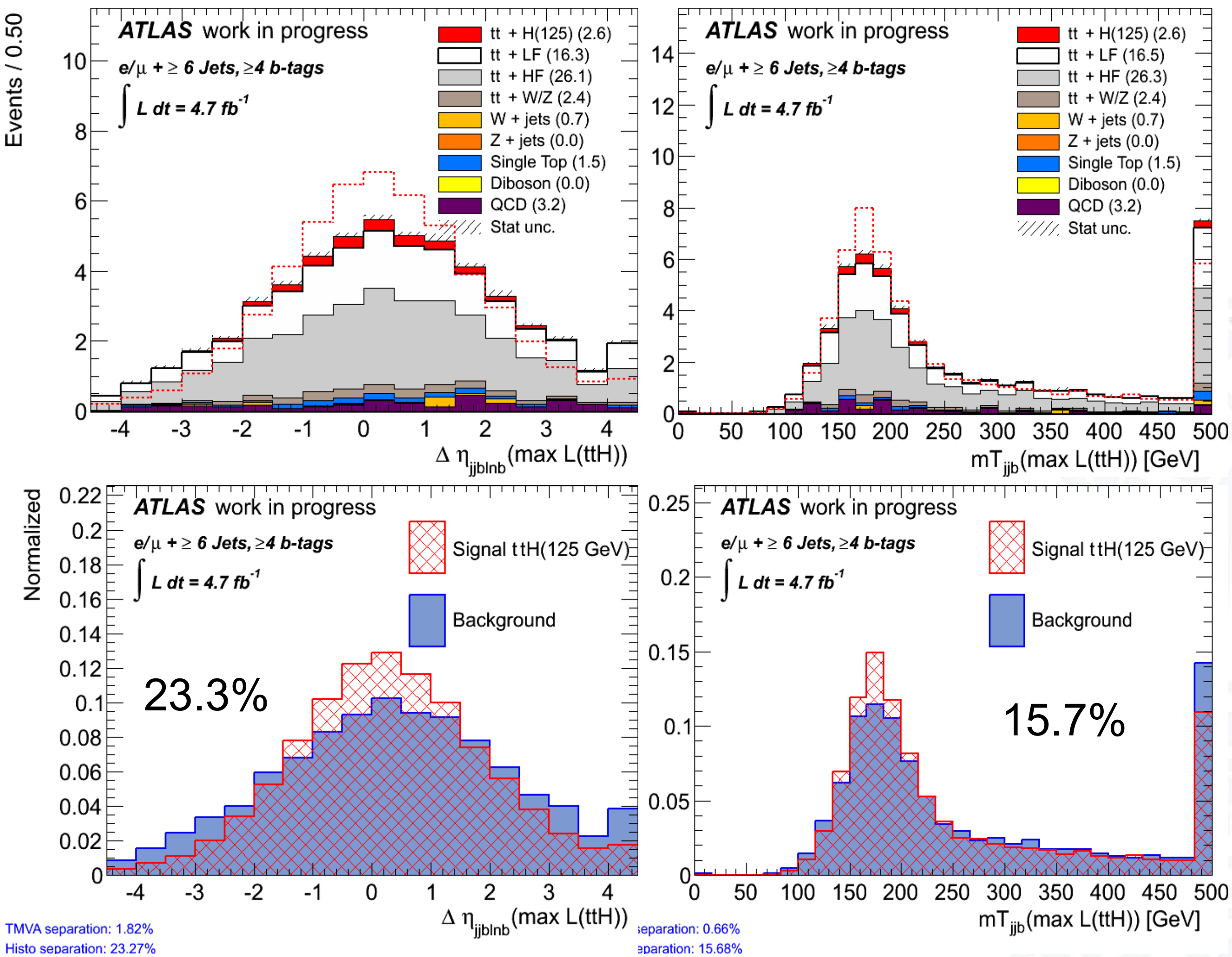
Max-L ($t\bar{t}b\bar{b}$)

- In signal Higgs peak visible
- Broad combinatorial shoulder
- Good discrimination power

→ Use BG hypothesis!

Other Discriminants

Other discriminating variables for best permutation



$\Delta\eta$ between tops

- Maximized for $t\bar{t}H$
- Full event reconstr.
- Additional boost by radiated Higgs

Hadronic top mass

- Maximized for $t\bar{t}H$
- Good mass reconstruction
- Better for $t\bar{t}H$ due to additional constraint

Ranking of Variables MVA+MEM

Table 1: Variables used as input to NN ordered by their importance

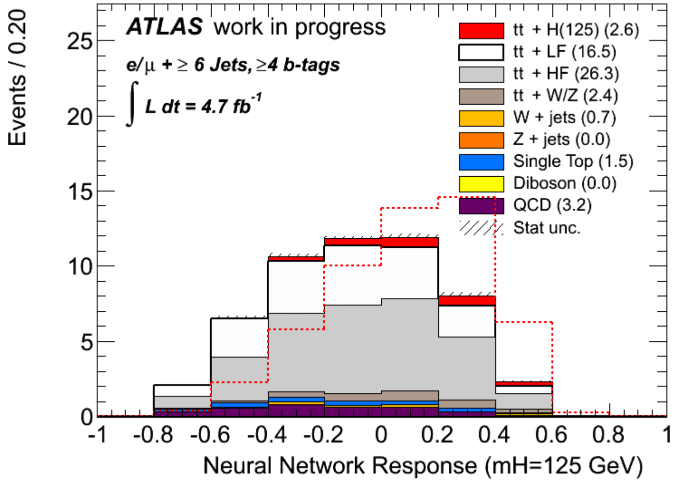
Variable	Significance(σ)
D1_mod (mH=125 GeV)	28.72
H_{T} (jets)	17.69
Aplanarity	19.89
mW_{jj}(max L(ttH)) [GeV]	9.1
Number of Jets (p_{T}>40 GeV)	7.7
#Delta #eta_{jjblnb}(max L(ttH))	7.12
m (bj^{max p_{T}})	6.56
#Delta #eta_{jjb}(max L(ttH))	5.14
#Delta #eta_{bb}(max L(ttH))	5.23
#Delta #eta_{jj}(max L(ttH))	3.46

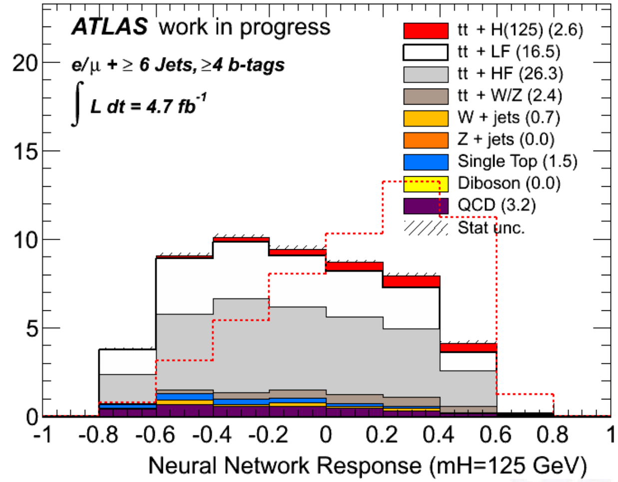
ATLAS work in progress

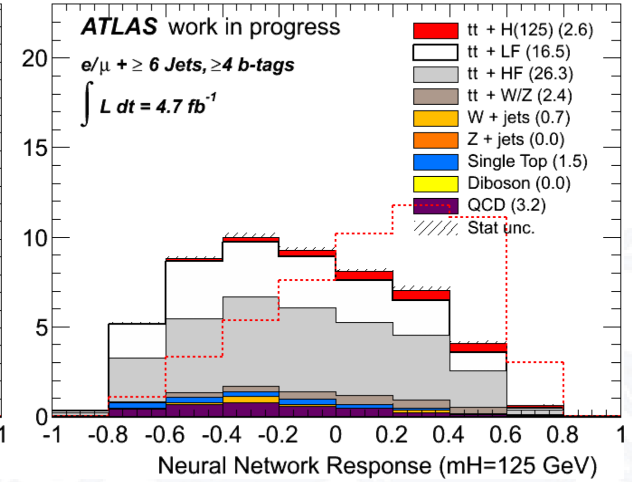
Total correlation to the target: 37.15% Total significance: 42.84

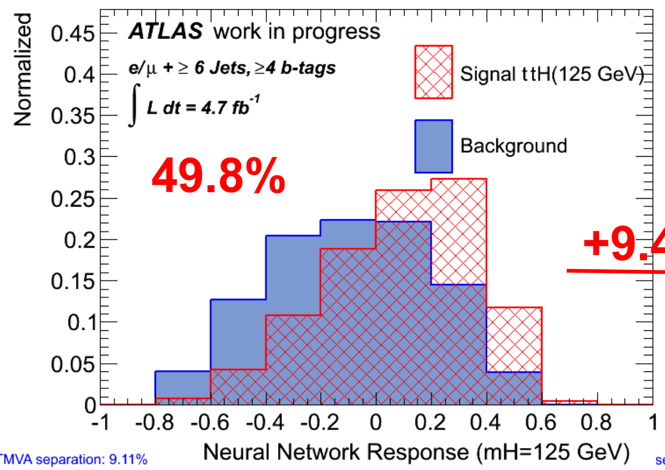
NN Output

NN output 5MVA → 4MVA+1MEM → 4MVA+6MEM



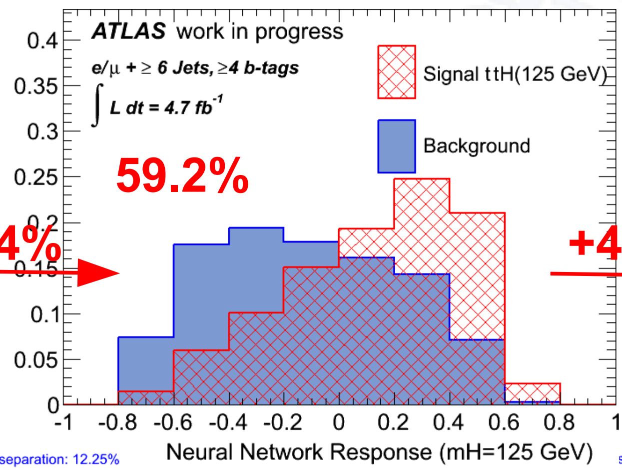






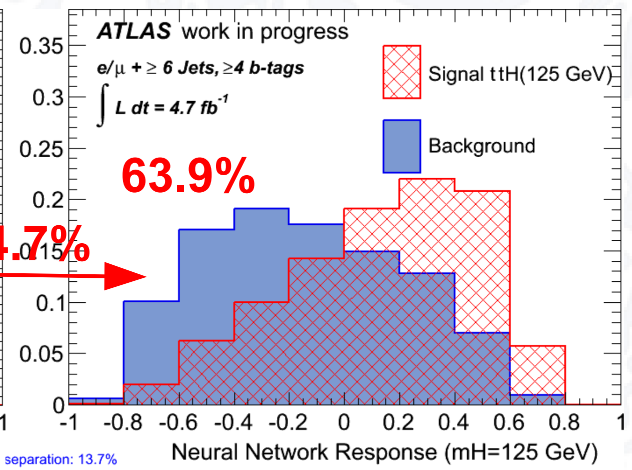
TMVA separation: 9.11%
Histo separation: 49.8%

+9.4%



separation: 12.25%
eparation: 59.14%

+4.7%



separation: 13.7%
eparation: 63.87%

Statistical Limits

- Statistical limits (95% CL on σ/σ_{SM}) set in 6ji/4bi using best ranked variables
  - 5 MVA variables (only MVA)
  - 4 MVA variables + 1 MEM (5 Var)
  - 4 MVA variables + 6 MEM (10 Var)
  - 10 MEM variables (only MEM)
- Including MEM variables improves the exp. limit and degrades the uncertainties
- MEM only can compete with highly optimized MVA only limit
- Improvements in the limit including systematic uncertainties are significant

ATLAS work in progress

	5MVA	4MVA1MEM	4MVA6MEM	10MEM
expected	6.00	5.67	5.59	5.92
+- 1 sigma	[4.32, 8.52]	[4.08, 8.08]	[4.03, 7.96]	[4.27, 8.42]
+- 2 sigma	[3.22, 11.85]	[3.04, 11.29]	[3.00, 11.14	[3.18, 11.71]

Summary & Outlook

Summary

- MEM is powerful tool for discriminating $t\bar{t}H$ from BG
  - First application of MEM at LHC to search (complex channel)
  - Good balance between precision & reasonable CPU time needed (4 m/evt)
  - Development of large computing framework was needed
    - Multi-user submission, automated monitoring, etc
  - Complex method with high potential for good discrimination
  - Neyman-Pearson likelihood ratio shows impressive separation power
  - MEM variables in NN improves sensitivity
    - 20-30% relative improvement in discrimination
    - Expected statistical only limit degrades

Outlook

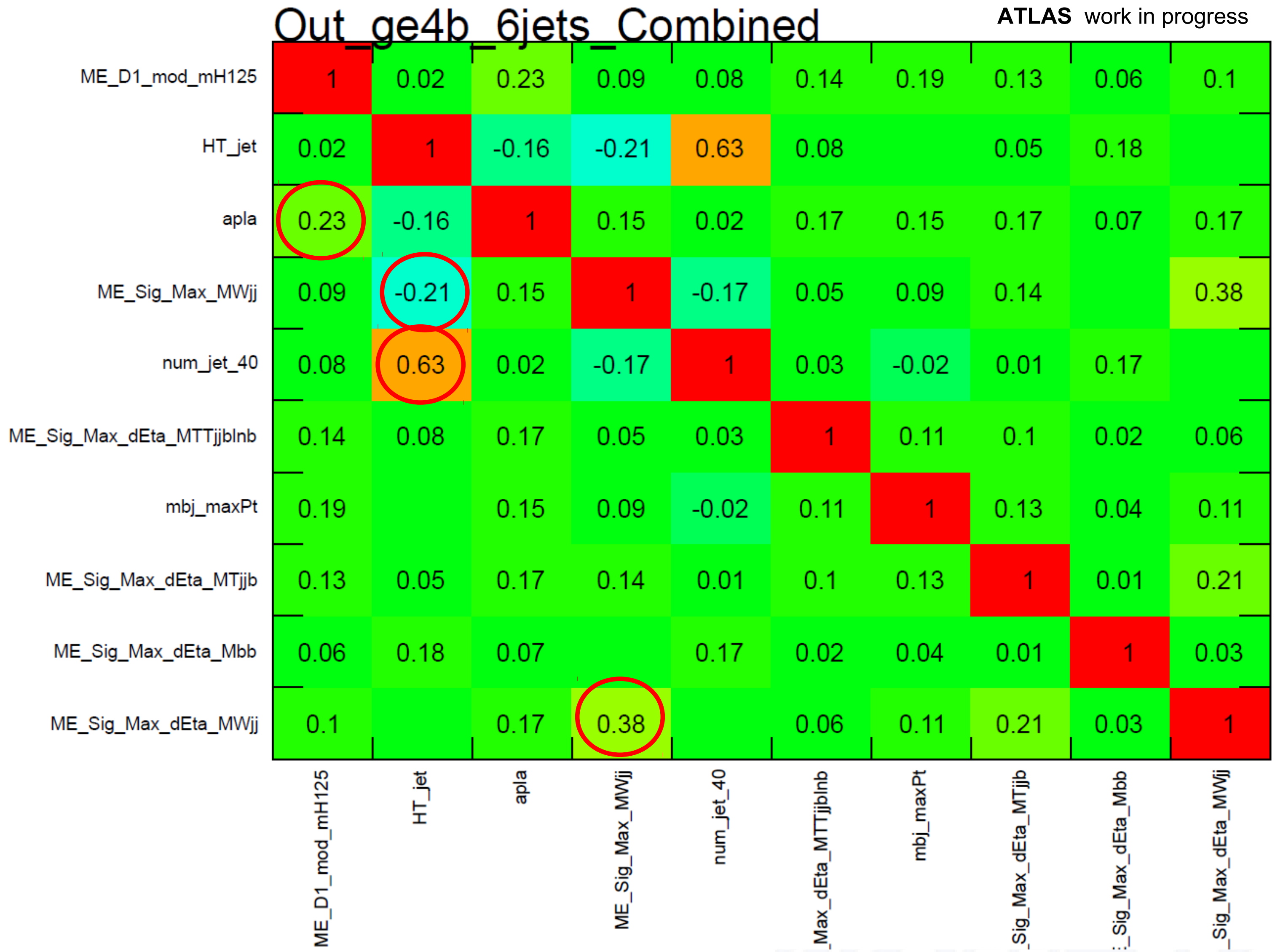
- Significant gain in sensitivity including systematics (7TeV)
- Run on 8 TeV with current description and full systematic
  - Possibility of using lookup tables for systematic breakdowns
  - Use permutation ranking to simultaneously run on 0/3 b-tag

BACKUP

$t\bar{t}H$ with Matrix Element Method

- Realization with MEMTool in ≥ 6 jets, ≥ 4 b-tags channel
  - Model:
    - LO PDFs: CTEQ6L1
    - Use *optimized* MadGraph LO ME ($t\bar{t}H$ / $t\bar{t}b\bar{b}$ diagrams)
    - KLFitter TF
  - Default options:
    - Veto b-tag method
    - Selection: 4 jets b-tagged ranked, 2 jets in pt order (light jets)
    - Kinematic transformation to m_W and m_H :
      - No lep int., NWA for lept. M_W, no NWA for hadr. m_W and m_H
  - Test signal and background hypothesis
    - Max[L($t\bar{t}H$) | m_H] – permutation which maximizes $t\bar{t}H$ likelihood
    - Max[L($t\bar{t}bb$)] – permutation which maximizes $t\bar{t}b\bar{b}$ likelihood
  - Goal of MEM:
    - Neyman-Pearson likelihood ratio as m_H dependent discriminator
    - Event Reconstruction

Correlations MVA+MEM



Ranking of Variables only MEM

Table 1: Variables used as input to NN ordered by their importance

Variable	Significance(σ)
D1_mod (mH=125 GeV)	28.72
HT_{all}(max L(ttH)) [GeV]	14.07
#Delta #eta_{jj}(max L(ttH))	10.45
#Delta #eta_{jjblnb}(max L(ttH))	9.43
#Delta #eta_{bb}(max L(ttH))	8.18
#Delta #eta_{jjb}(max L(ttH))	7.33
mW_{jj}(max L(ttH)) [GeV]	5.05
#Delta #eta_{bb}(max L(ttbb))	4.7
cos((j_{1}j_{2})#angle(b_{1}b_{2}))(max L(ttbb))	4.53
#Delta cos(#phi)_{lnu}(max L(ttH))	3.41

ATLAS work in progress

Total correlation to the target: 32.69% Total significance: 37.7

Correlations only MEM

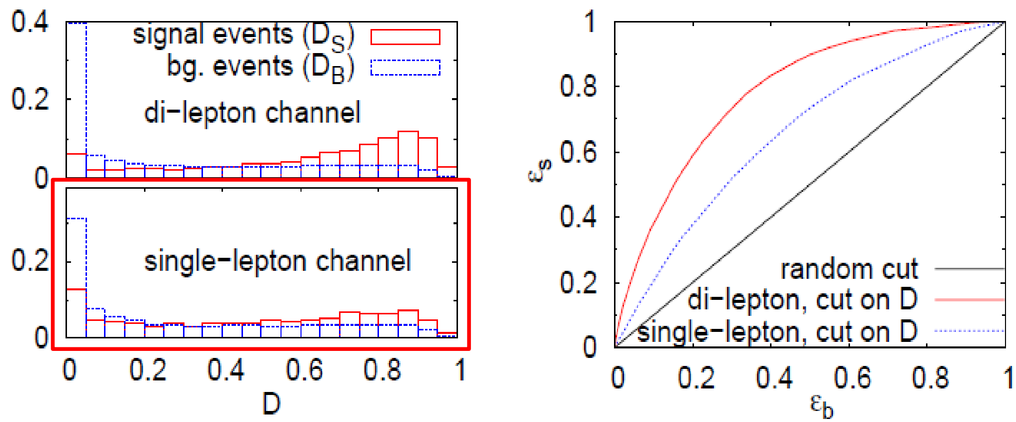
Out_ge4b_6jets_Combined

ATLAS work in progress

	ME_D1_mod_mH125	ME_Sig_Max_EtTTbb	Sig_Max_dEta_MWjj	Max_dEta_MTTjjblnb	_Sig_Max_dEta_Mbb	Sig_Max_dEta_MTjjb	ME_Sig_Max_MWjj	Bkg_Max_dEta_Mbb	kg_Max_CP_bb_mbb	Max_dCosPhi_MWlnu
ME_D1_mod_mH125	1	0.04	0.1	0.14	0.06	0.13	0.09	0.04	0.05	0.04
ME_Sig_Max_EtTTbb	0.04	1	0.01	0.13	0.19	0.07	-0.18	0.04	-0.01	0.23
ME_Sig_Max_dEta_MWjj	0.1	0.01	1	0.06	0.03	0.21	0.38			0.02
ME_Sig_Max_dEta_MTTjjblnb	0.14	0.13	0.06	1	0.02	0.1	0.05	0.05	0.02	0.1
ME_Sig_Max_dEta_Mbb	0.06	0.19	0.03	0.02	1	0.01		0.27	0.03	0.02
ME_Sig_Max_dEta_MTjjb	0.13	0.07	0.21	0.1	0.01	1	0.14	0.06	0.01	0.03
ME_Sig_Max_MWjj	0.09	-0.18	0.38	0.05		0.14	1	-0.04		
ME_Bkg_Max_dEta_Mbb	0.04	0.04		0.05	0.27	0.06	-0.04	1	0.02	0.03
ME_Bkg_Max_CP_bb_mbb	0.05	-0.01		0.02	0.03	0.01		0.02	1	
ΛE_Sig_Max_dCosPhi_MWlnu	0.04	0.23	0.02	0.1	0.02	0.03		0.03		1

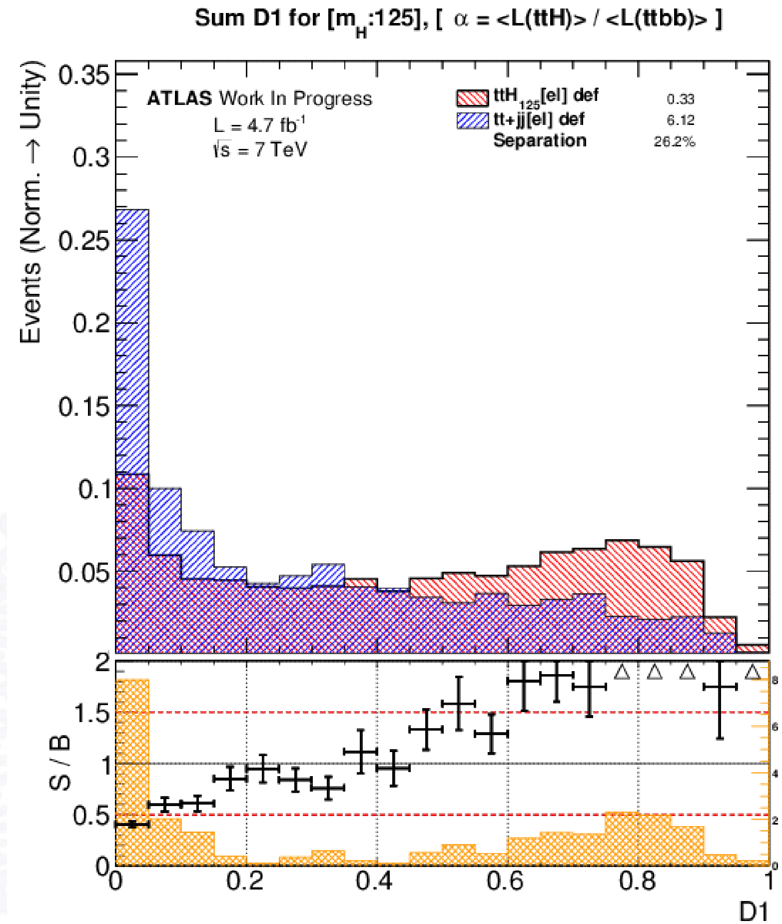
Theory vs Reality

- Paper by MadWeight authors:
  - Discussing analyses at 14 TeV
  - Analysis simplified by theorist, important experimental affects not taken into account, MadEvents
  - Not optimized in terms of CPU power
    - → fewer approximations!
- MEMTool
  - I might be biased, but I think we are better...



Unravelling ttH via MEM (arXiv:1304.6414)

FIG. 1: Left: Normalized distributions of events with respect to the MEM-based observable D for the di-lepton (top) and single-lepton (bottom) channels. Right: Efficiency of selecting signal vs. background using a $D > D_{min}$ cut.



Matrix Element Method (1)

- Method:
  - Probability of observing an event in the detector to be consistent with model:

$$P_{ev}(\boldsymbol{x}|\boldsymbol{a})=\sum_i f_i P_i(\boldsymbol{x}|\boldsymbol{a})$$

    - Observed quantity **x**
    - Model parameters **a** (theoretical & instrumental)
    - fractions f_i of different process / production channel (non interfering)
  - Calculate probability P_i to measure **x** with given **a** for each channel *i*
  - Combine all event probabilities into one likelihood function for n events:

$$L(\boldsymbol{a})=\prod_{j=1}^{n} P_{ev}(\boldsymbol{x}|\boldsymbol{a})$$

  - Extract model parameters **a** by maximizing the likelihood function
    - Model parameters might be, e.g.: m_H, JES, bJES, signal fraction f_{ttH}
    - → Allows to further reduce the systematical uncertainties

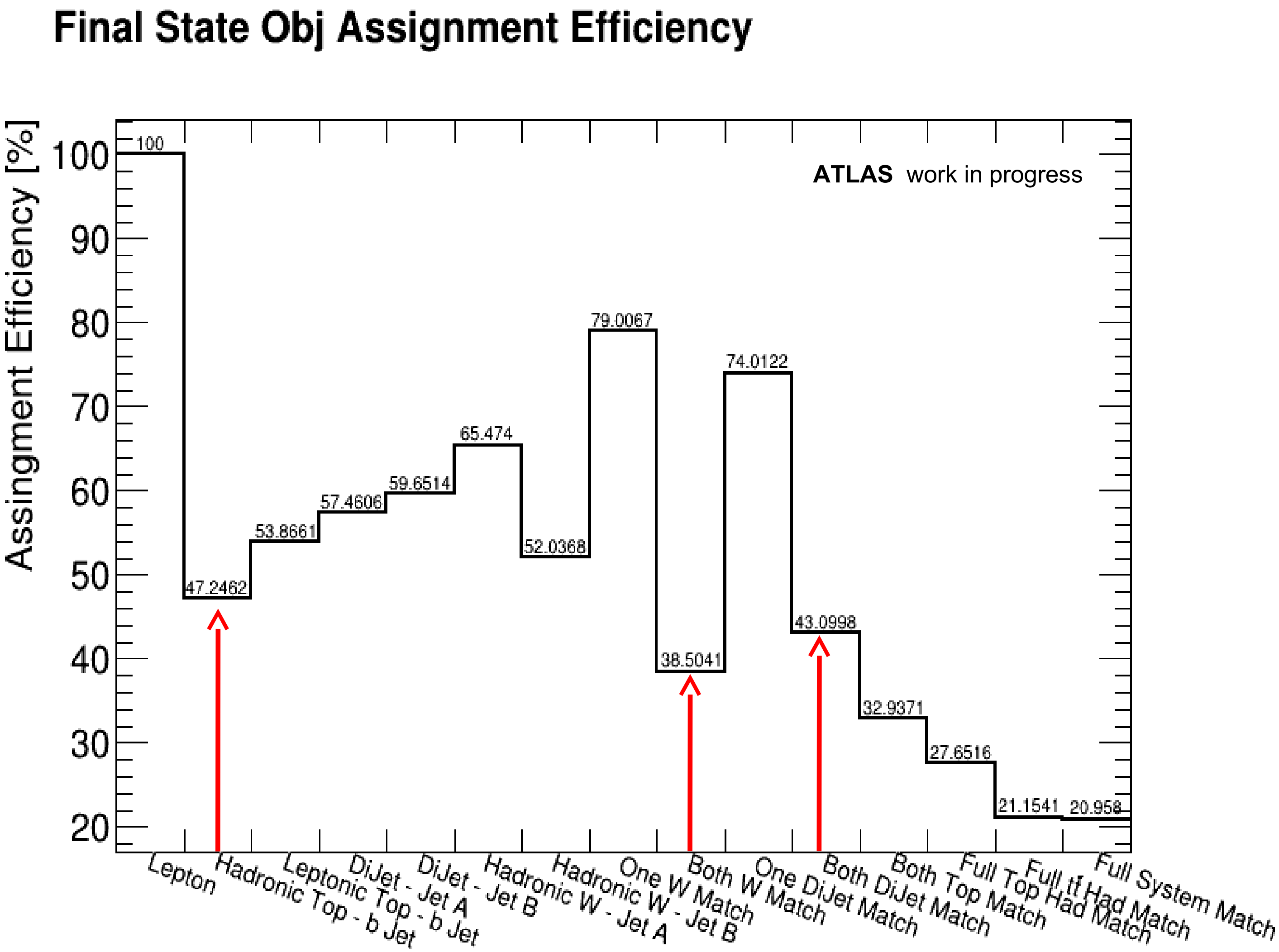
Reconstruction Max L(perm|ttH)

Assignment Efficiencies and Migrations for:

- Max L(P|ttH) where P is in '12 "b-Tag Veto" Perms'

- Events with good matches on all Truth partons

- Event with 6 n jet **inclusive**



* Jet "**A**" is always the leading pt parton, and "**B**" is always the sub-leading parton.

Reconstruction Max L(perm|ttH)

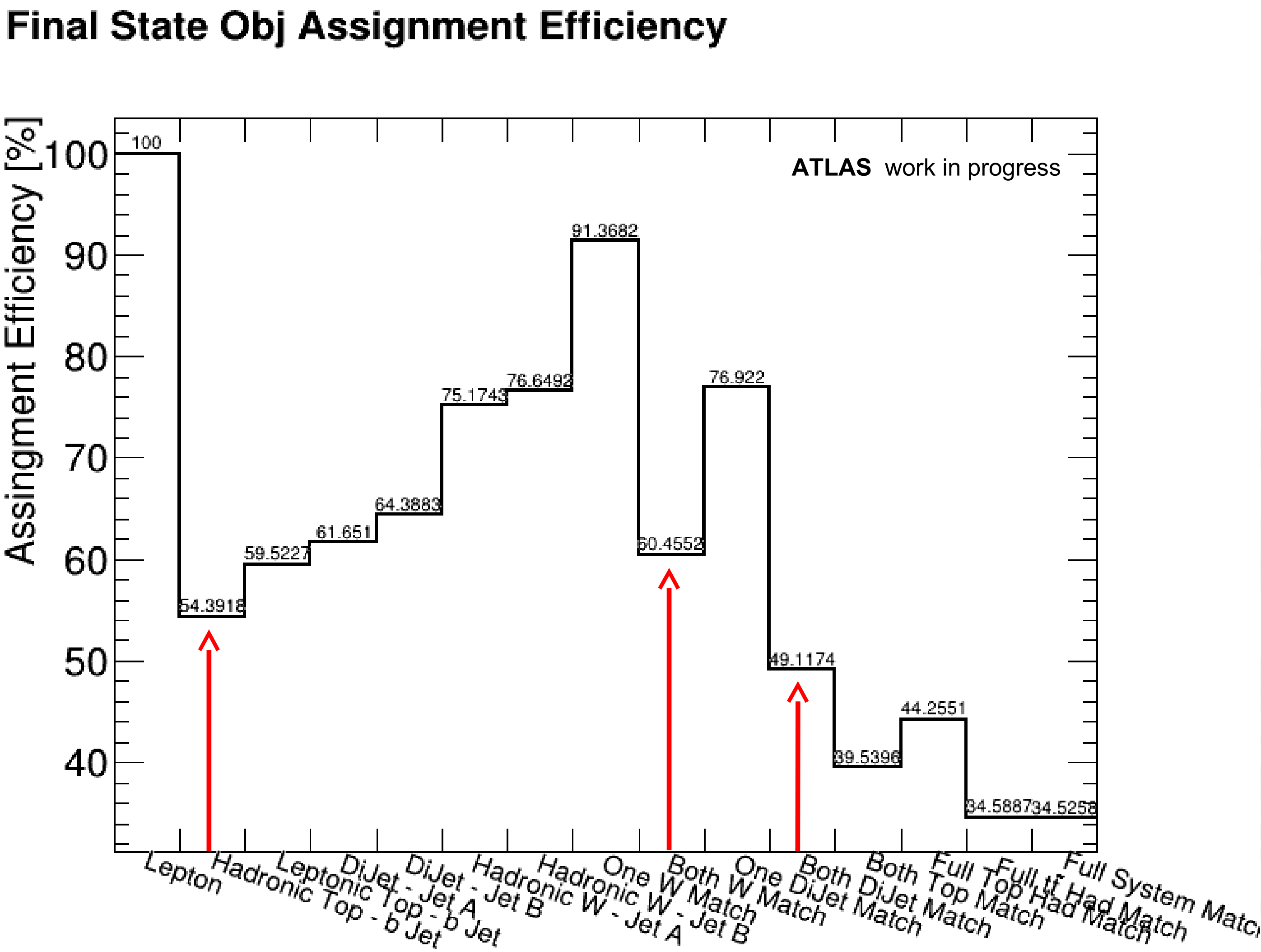
Assignment Efficiencies and Migrations for:

- Max L(P|ttH) where P is in '12 "b-Tag Veto" Perms'

- Events with good matches on all Truth partons

- Event with 6 n jet **exclusive**

* Jet "**A**" is always the leading pt parton, and "**B**" is always the sub-leading parton.



Reconstruction Max L(perm|ttbb)

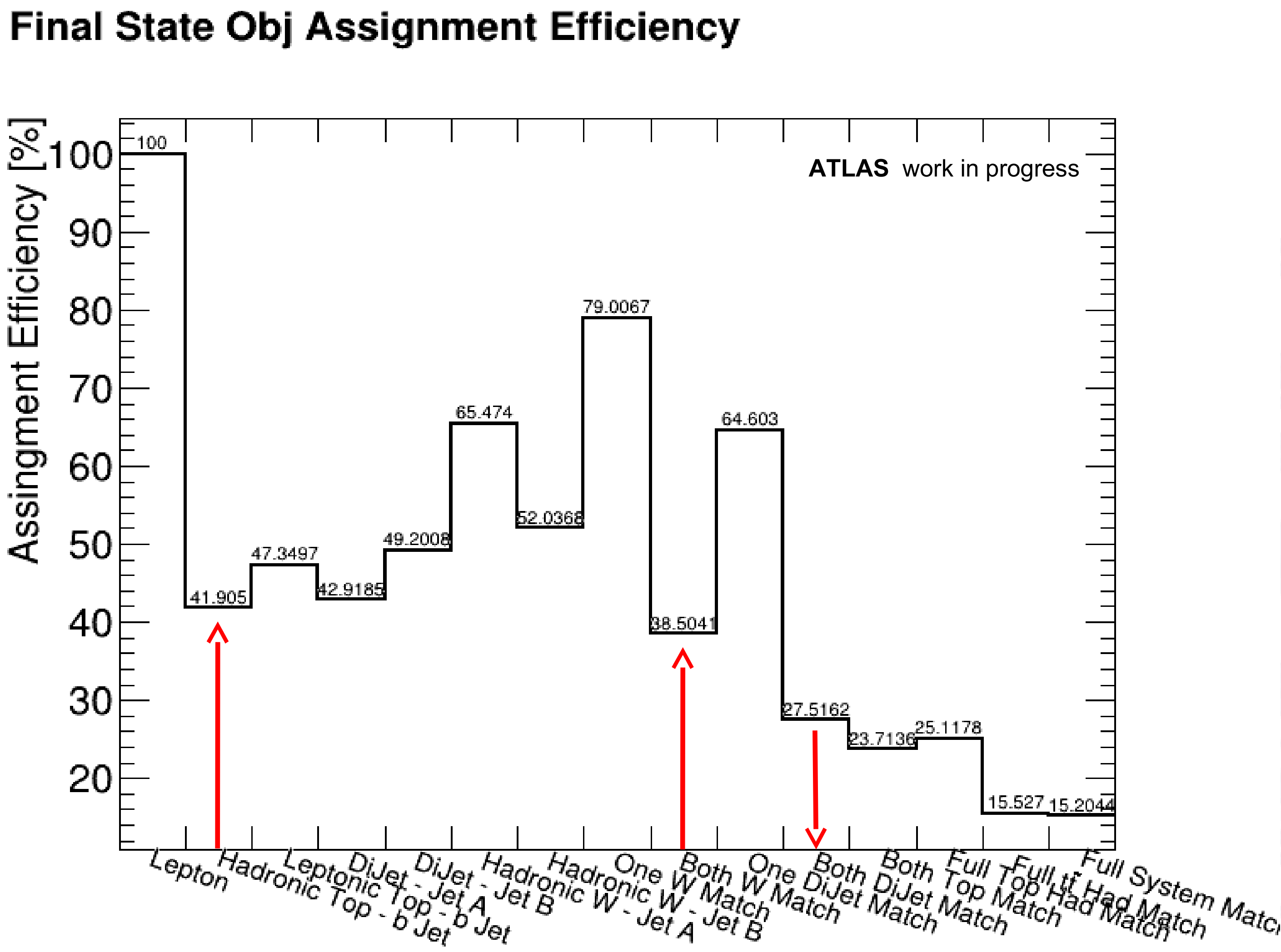
Assignment Efficiencies and Migrations for:

- Max L(P|ttbb) where P is in '12 "b-Tag Veto" Perms'

- Events with good matches on all Truth partons

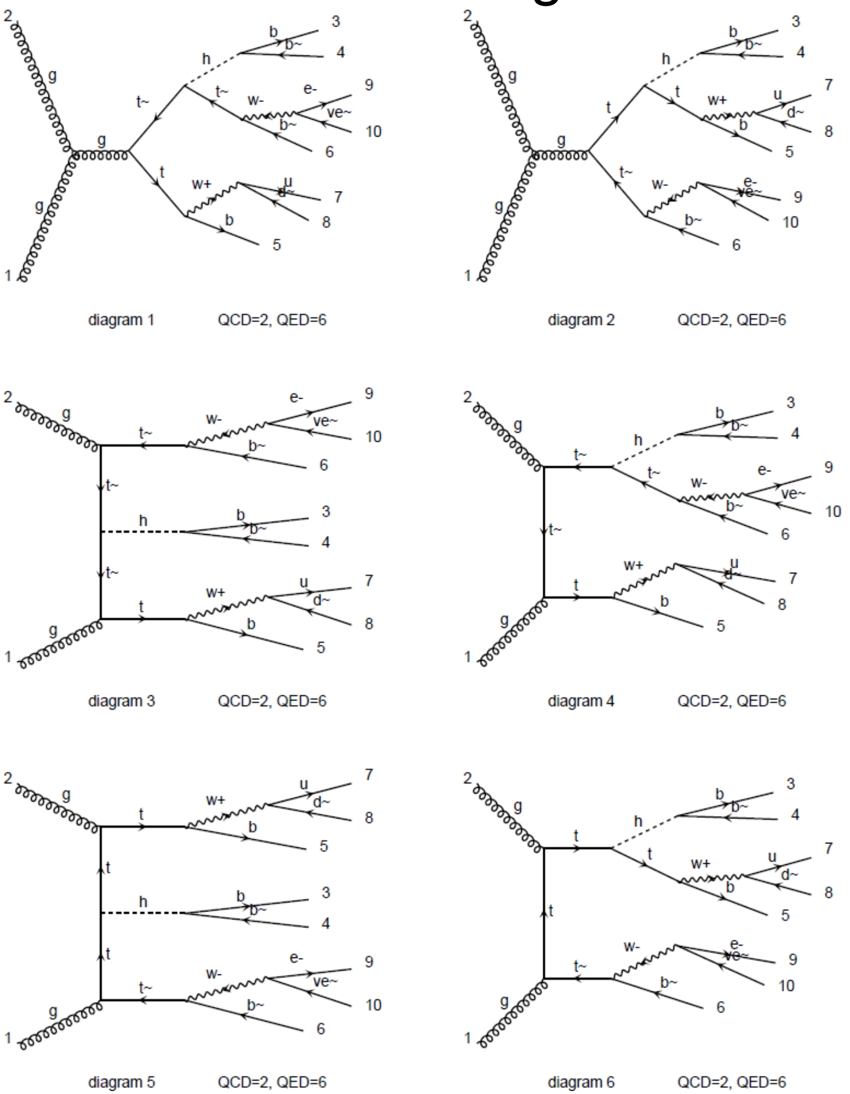
- Event with 6 n jet **inclusive**

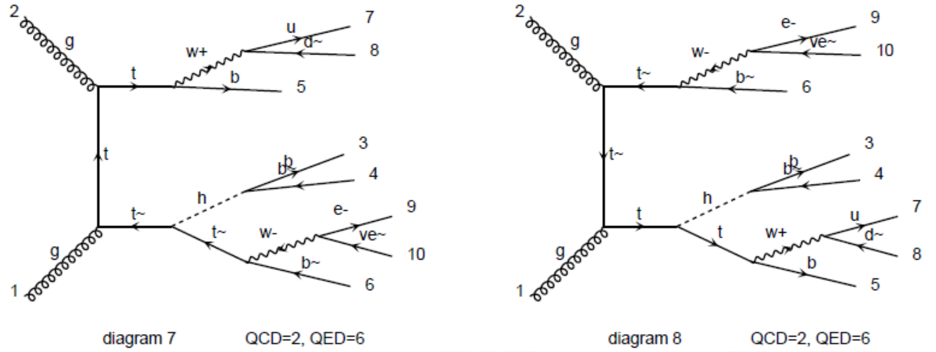
* Jet "**A**" is always the leading pt parton, and "**B**" is always the sub-leading parton.



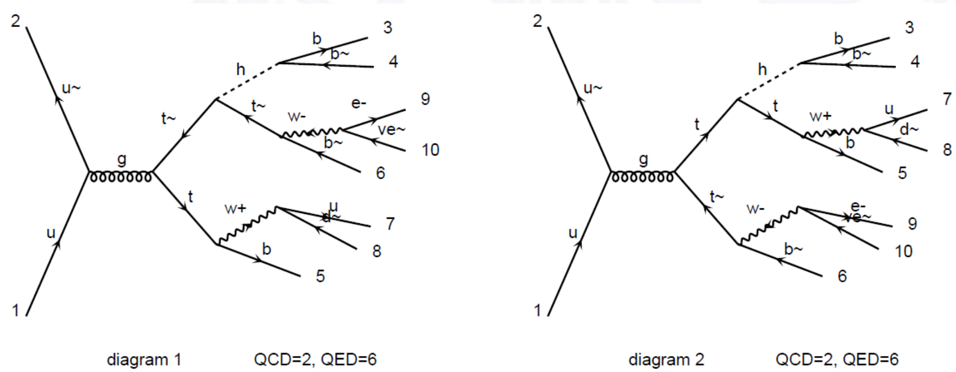
Looking at Signal Diagrams

Gluon Fusion: 8 Diagrams





$q\bar{q}$-Annihilation: 2 Diagrams



in ttH: gg / $q\bar{q}$ production ~ 50 / 50!

Looking at BG Diagrams

Gluon Fusion: 16 Diagrams (1-12)

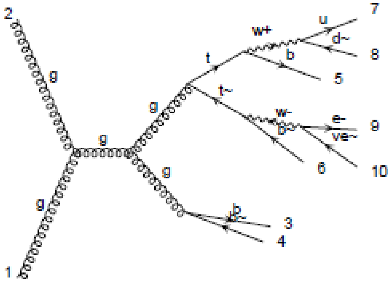


diagram 1 QCD=4, QED=4

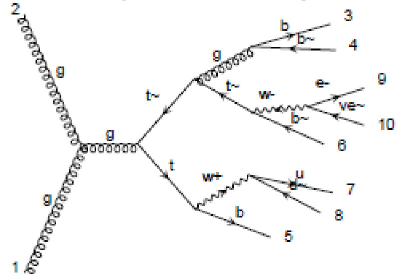


diagram 2 QCD=4, QED=4

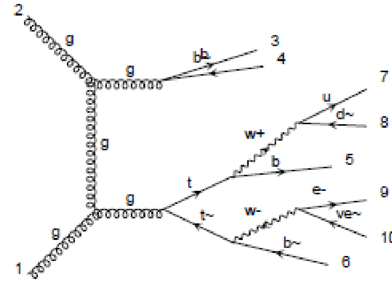


diagram 7 QCD=4, QED=4

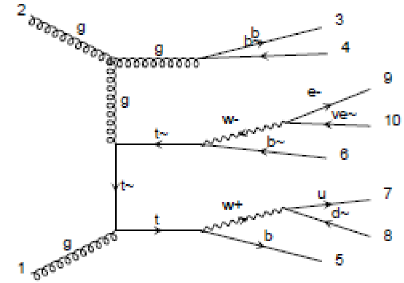


diagram 8 QCD=4, QED=4

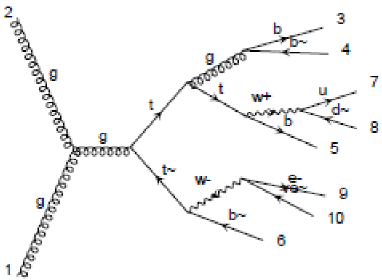


diagram 3 QCD=4, QED=4

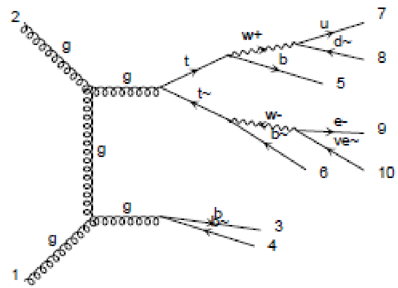


diagram 4 QCD=4, QED=4

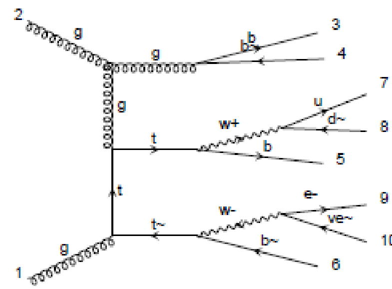


diagram 9 QCD=4, QED=4

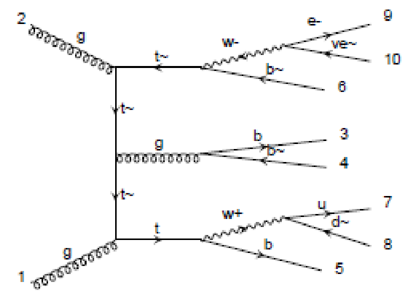


diagram 10 QCD=4, QED=4

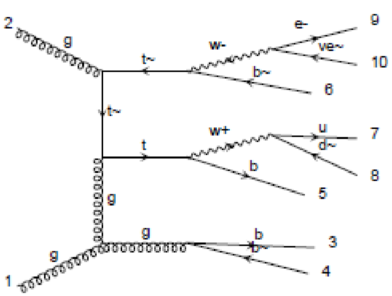


diagram 5 QCD=4, QED=4

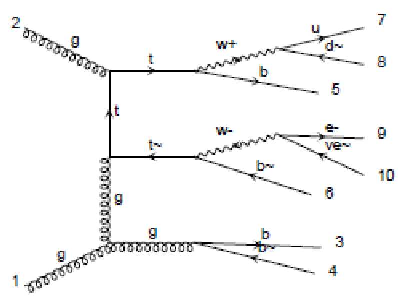


diagram 6 QCD=4, QED=4

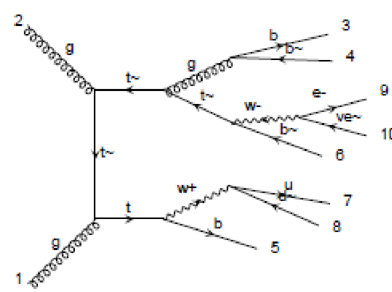


diagram 11 QCD=4, QED=4

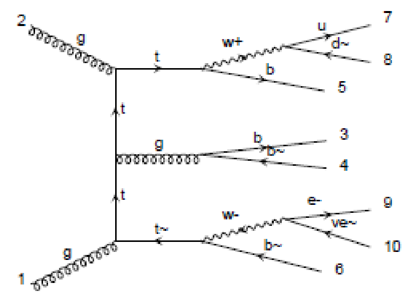
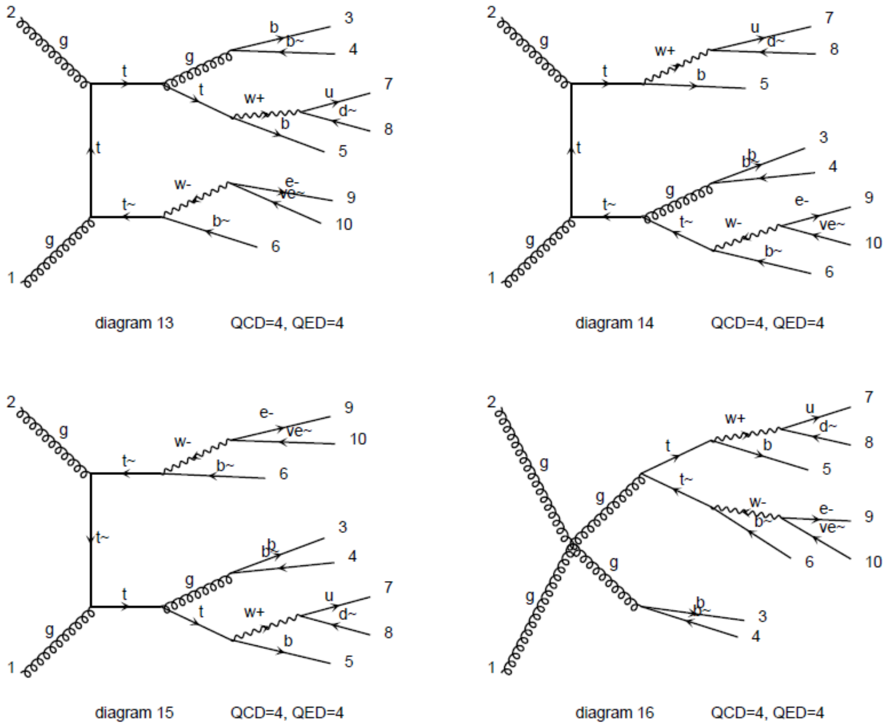


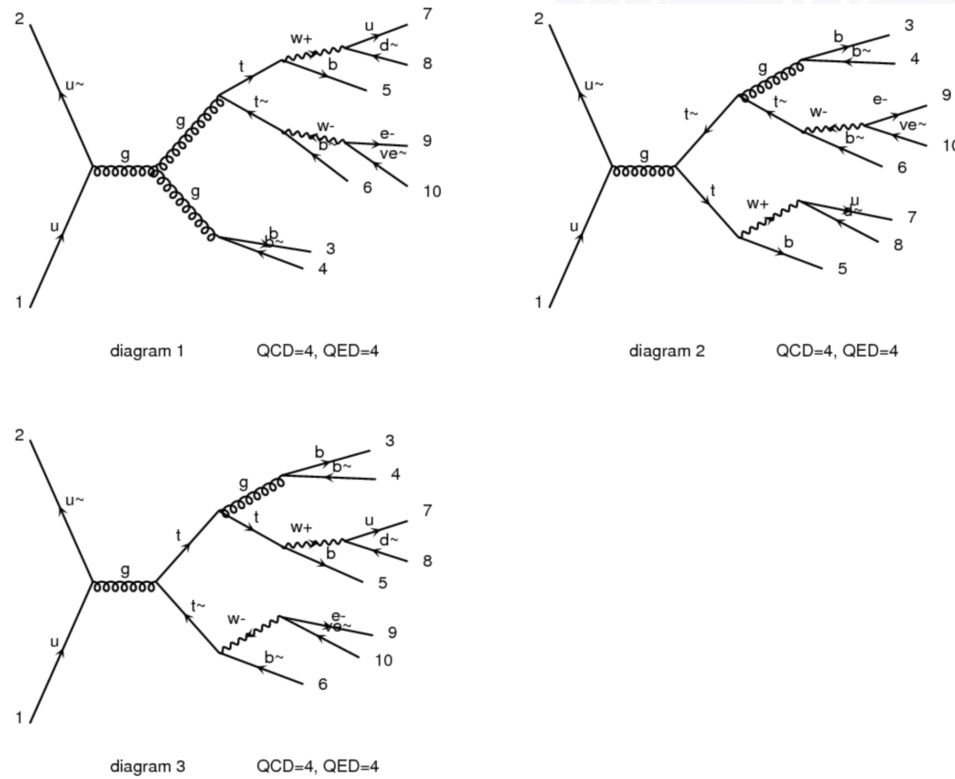
diagram 12 QCD=4, QED=4

Looking at BG Diagrams

Gluon Fusion: 16 Diagrams (13-16)



$q\bar{q}$-Anhilitaion: 3 Diagrams



MEMTool

- MEMTool
  - Tool for ME calculation, developed in Göttingen for ttbar → ljets
  - Modular designed
    - Can handle all kinds of ME, including c++ generated MadGraph ME
    - All kinds of phase space transformations
    - A lot of different input interfaces exist
    - A set of PDFs in a library available (LHAPDF)
    - Lots of different integration techniques included (via GSL,default is VEGAS)
    - Different sets of transfer functions, including latest KLFitter TFs
  - Generalization of tool and lots of development
    - Can handle all kinds of event topologies (model particle, perm. tables, etc.)
    - Optimizations on ME calculation (helicity states): 20 times faster!
    - MadGraph ME can be automatically implemented including optimizations
    - Optimization of the integration (break up conditions, PS volume, etc.)
    - A lot of nice tools (plotting, grid / batch submission, etc.)
  - Several other groups are interested in the tool