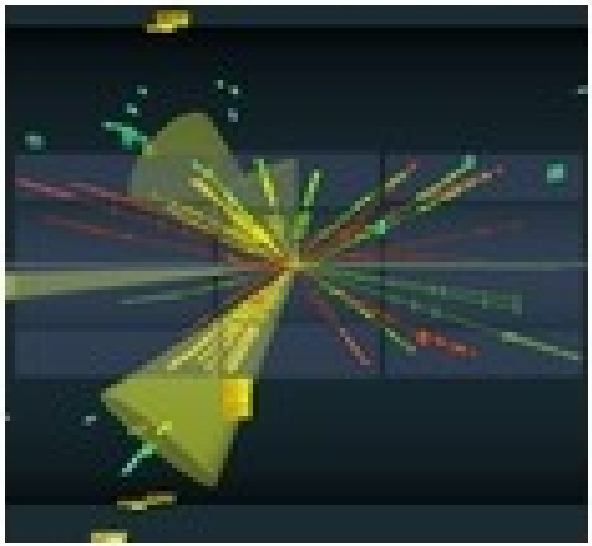
Dileptonic ttH events in CMS events in CMS

Analysis of the ttH events with 2 leptons and 4 b-jets in the final state



N.Bartosik, **J.Garay García**, J.Hauk

Dileptonic ttH events in CMS events in CMS

Benasque, September 2013

Thesis directed by A. Meyer and J. Haller

HELMHOLTZ ASSOCIATION

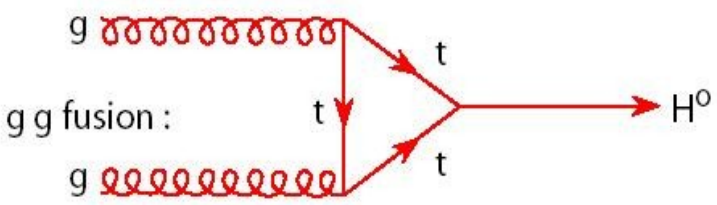
DESY

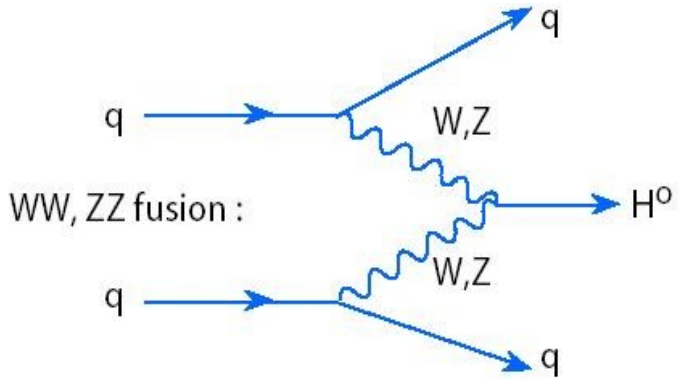
Outline

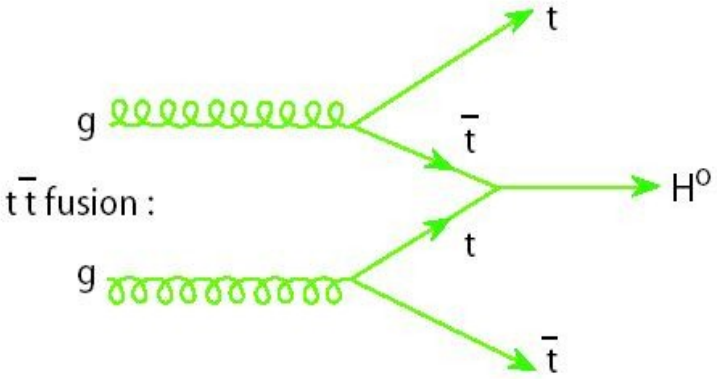
- Higgs and Top production at the LHC
- The analysis
  - Boosted Decision Trees
  - Types of jet combinations
  - MVA input variables
- Results
  - Separation power
  - Dijet mass
- Summary and conclusions

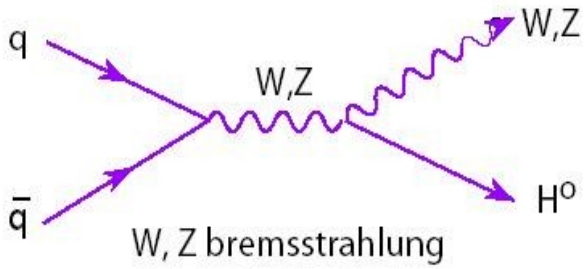
Higgs production at LHC

> Higgs production at LHC comes from very different processes:



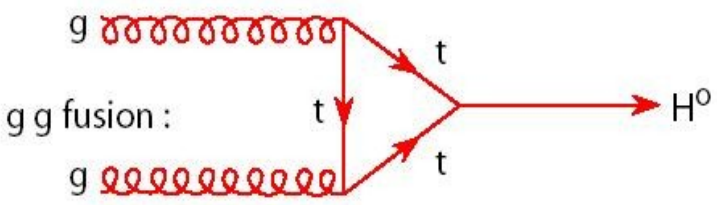


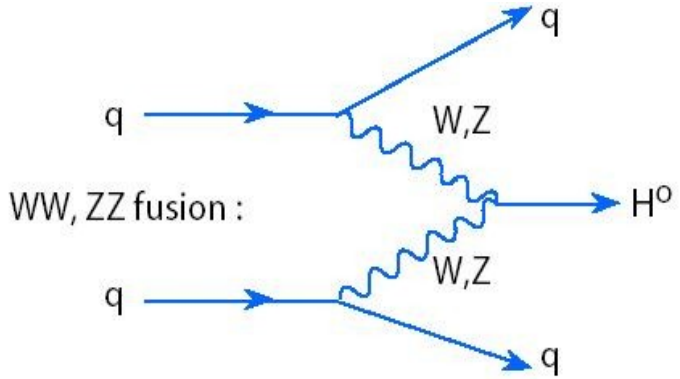


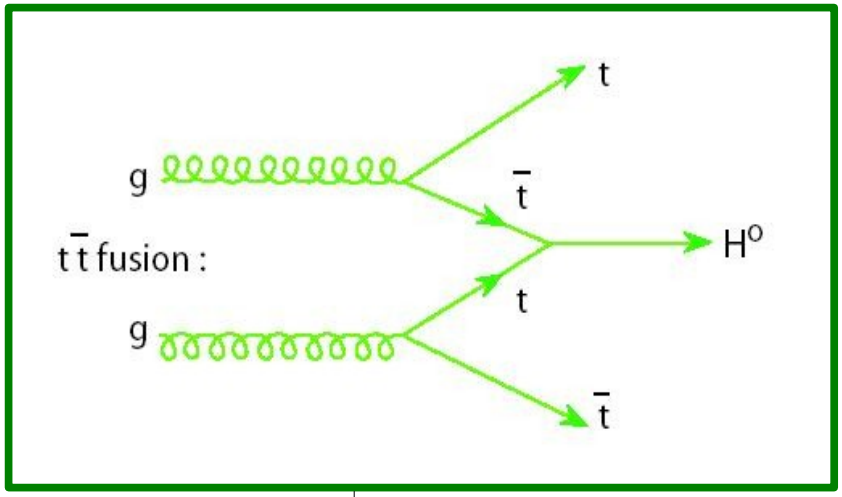


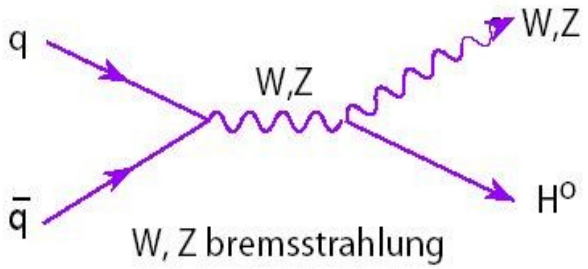
Higgs production at LHC

> Higgs production at LHC comes from very different processes:





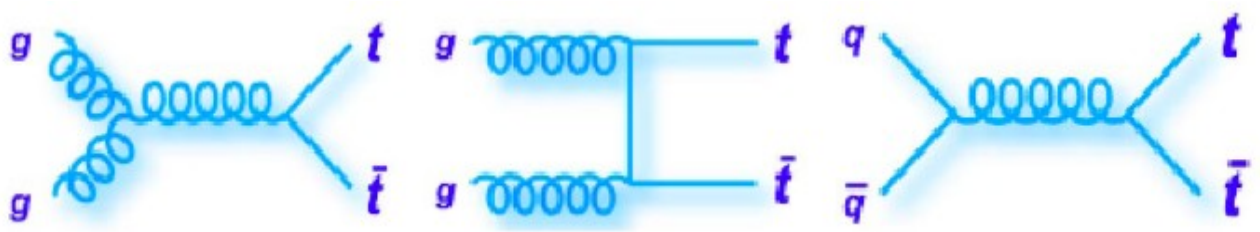




This is what we will talk about!

Top quark production at LHC

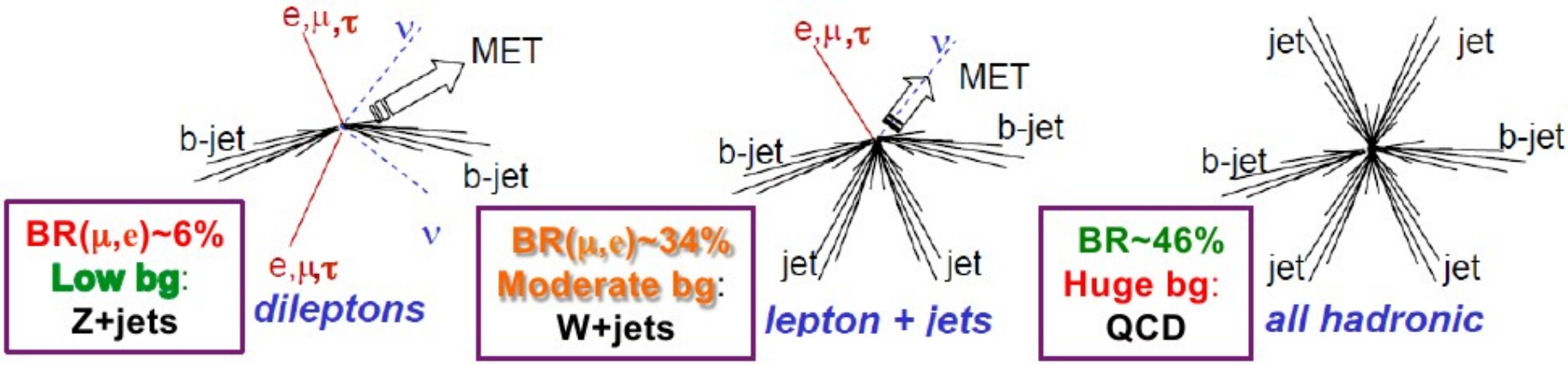
> Top pairs mainly produced by gluon fusion at LHC



	LHC (7TeV)	Tevatron
gg	~80%	~15%
qq̄	~20%	~85%

Courtesy of M.Aldaya

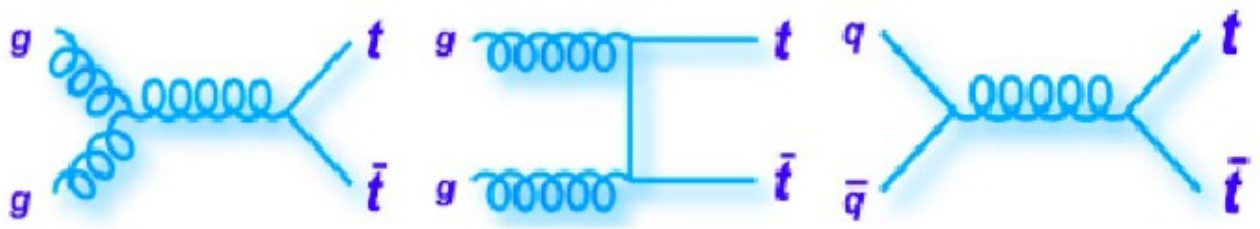
> The possible final states of the top decays are three: **dileptonic,** lepton+jets and full hadronic



Courtesy of M.Aldaya

Top quark production at LHC

> Top pairs mainly produced by gluon fusion at LHC

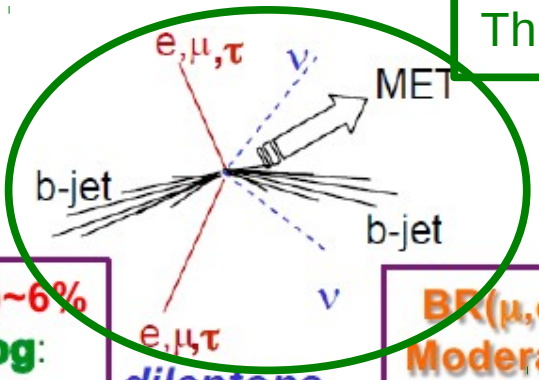


	LHC (7TeV)	Tevatron
gg	~80%	~15%
qq̄	~20%	~85%

Courtesy of M.Aldaya

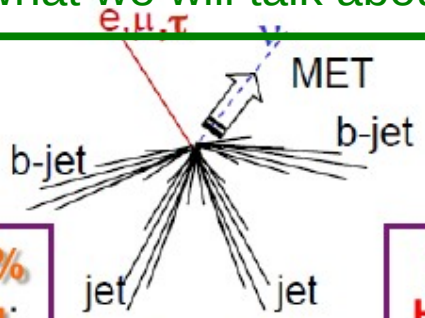
> The possible final states of the top decays are three: **dileptonic,** lepton+jets and full hadronic

This is what we will talk about!



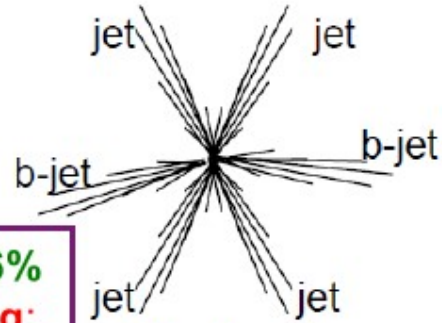
BR(μ,e)~6%
Low bg:
Z+jets

dileptons



BR(μ,e)~34%
Moderate bg:
W+jets

lepton + jets



BR~46%
Huge bg:
QCD

all hadronic

Courtesy of M.Aldaya

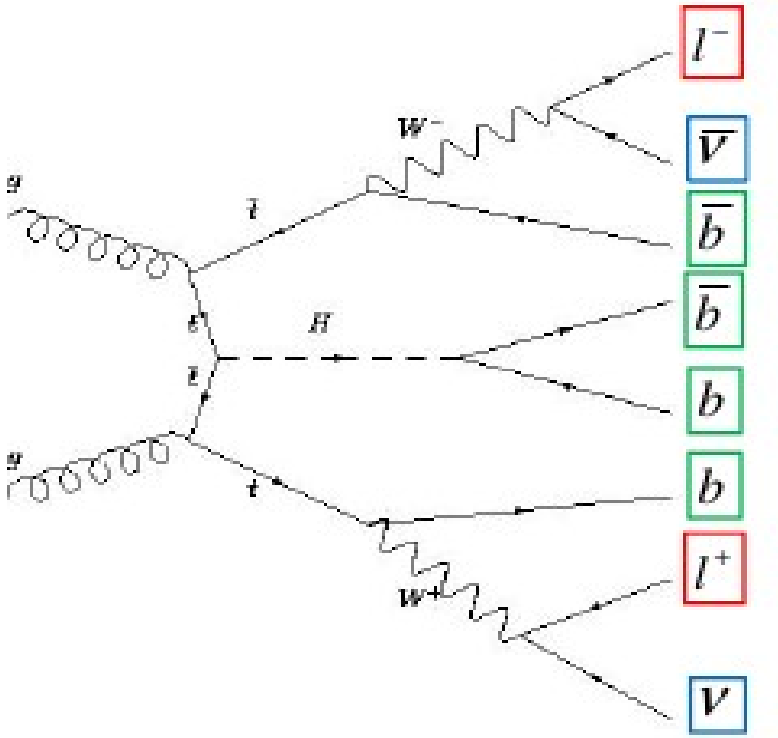
Event Topology

> Analysis of associated top-pair Higgs production

- Dileptonic final state of top system
  - → 2 oppositely charged isolated leptons (e,μ)
  - → MET mostly from two neutrinos
  - → Two b-jets
- Higgs decay into pair of b-jets
  - → Two b-jets

> Event fully accessible only in case of successful reconstruction of

- At least two oppositely charged isolated leptons
- At least four jets (or b-jets)



Courtesy of J. Hauk

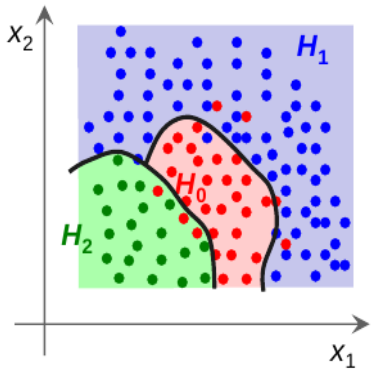
Current analysis goal and strategy

- Try to identify jets from Higgs: large combinatoric background. In each event we have at least four jets.
  - In the ideal case we would have 2 b-jets from top and two b-jets from Higgs.
  - It is crucial being able to distinguish among the jets coming from the different systems.

- **Strategy**: Try to **identify b-jets originating from top system** using Multivariate Analysis (MVA), avoiding introduction of mass bias (towards m_H) for additional jets.
  - Multivariate analysis stands for the observation and analysis of more than one statistical outcome variable at a time.
  - We only choose MVA variables not directly involving information on Higgs jets kinematics.

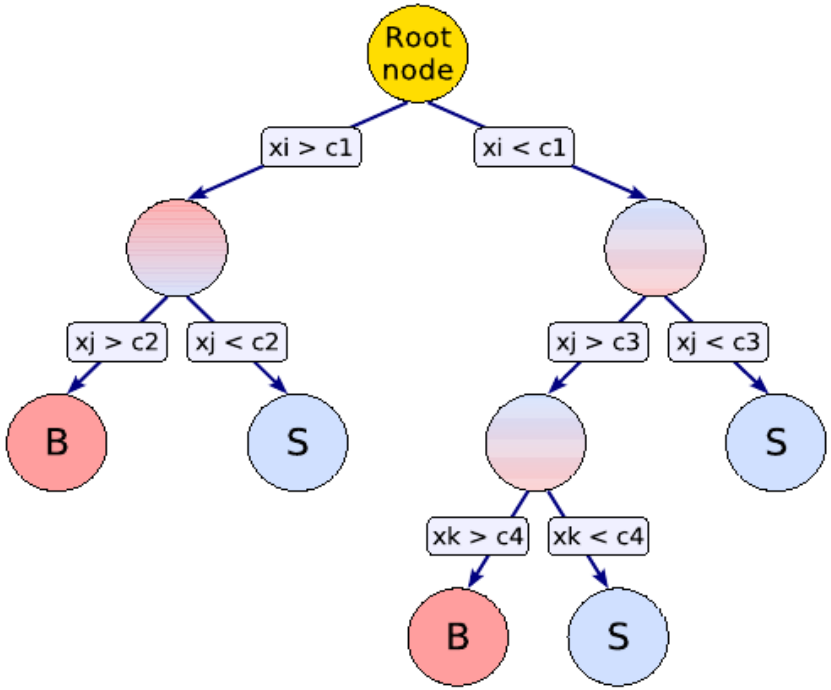
- **Tool**: **Boosted Decision Trees (BDT)**. Successive decision nodes used to categorize events as either signal or background.
  - Several parameters, such as the number of input variables or the number of trees are tuned in order to get the best possible approach.

Boosted Decision Trees

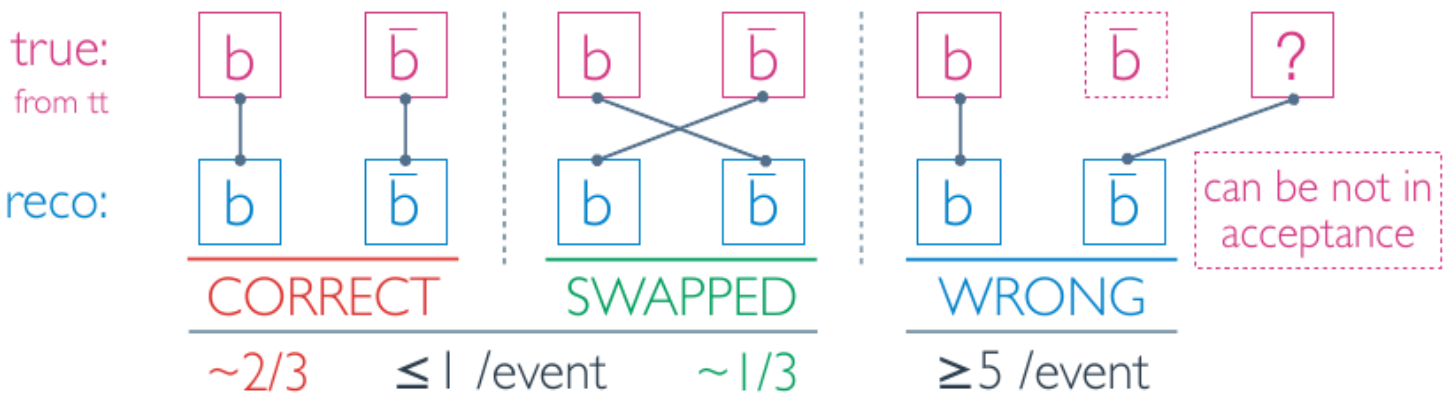


- Non-linearly correlated variables: BDT showed best approach.
- **BDT**: a series of cuts split the initial set into smaller sets, until finally classifying the event as signal or background. It is a training and testing process->provide weights for each event.

- **Adaptative Boosting**: signal events that end up in a background node are given a larger weight.
- Our variables should help us separating the b-jet pair coming from the top system from any other jet pair. So, **we don't classify the event, but the jet pair!**

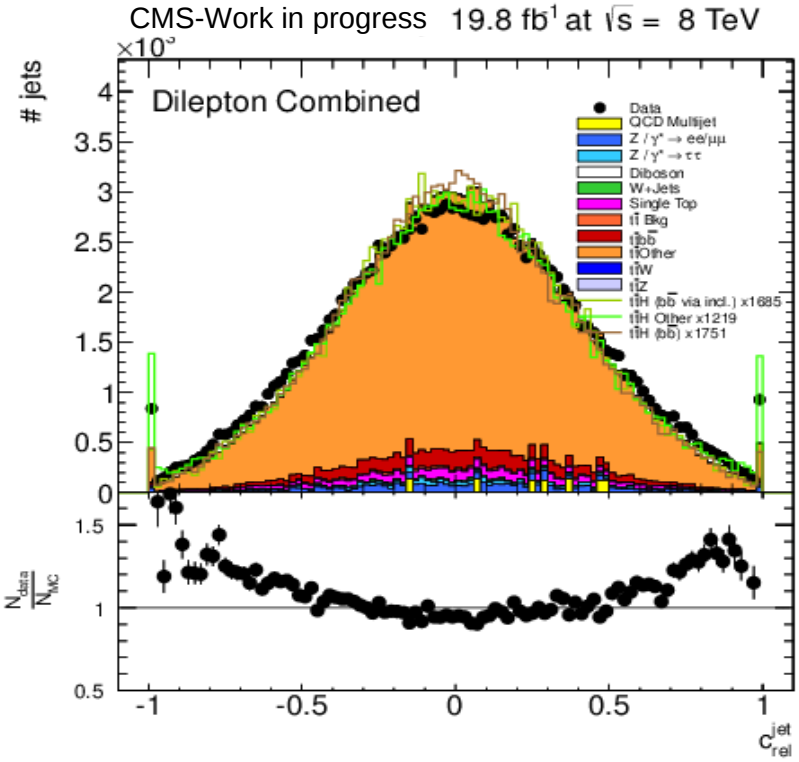


Types of jet combinations



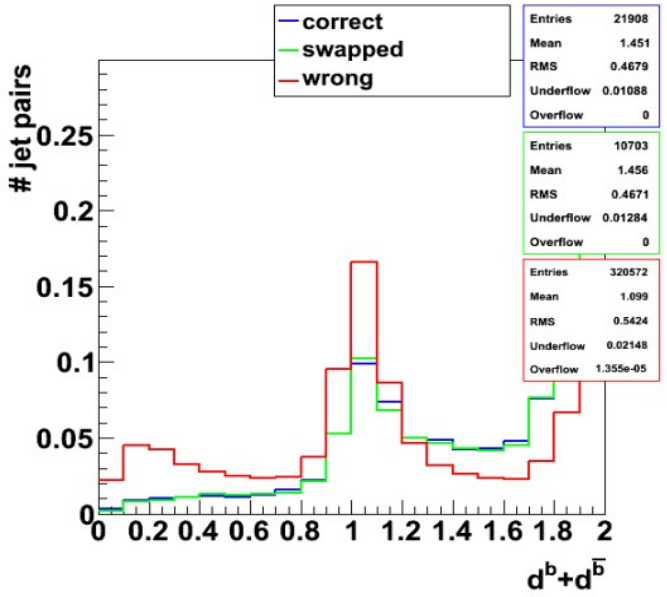
Courtesy of N.Bartosik

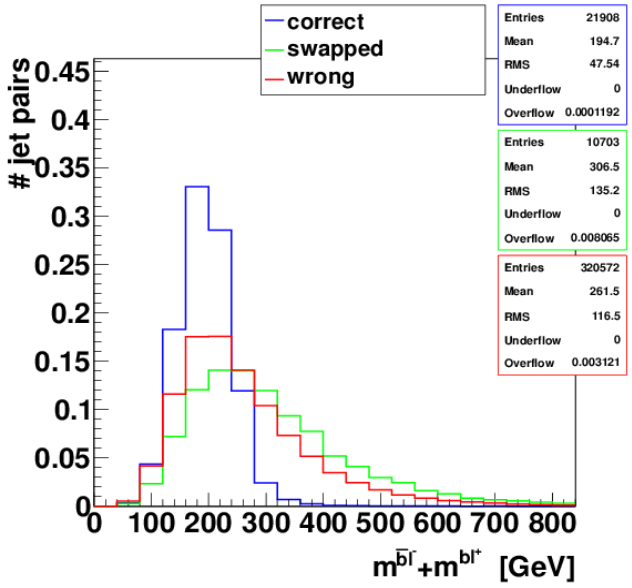
- For correct and swapped pairs, we use jet charge information to identify them:
  - c_{rel} : relative jet charge weighted by p_T with respect to jet axis.
  - In a pair of b-tagged jets:
    - → lower c_{rel}: b jet
    - → upper c_{rel}: anti-b jet

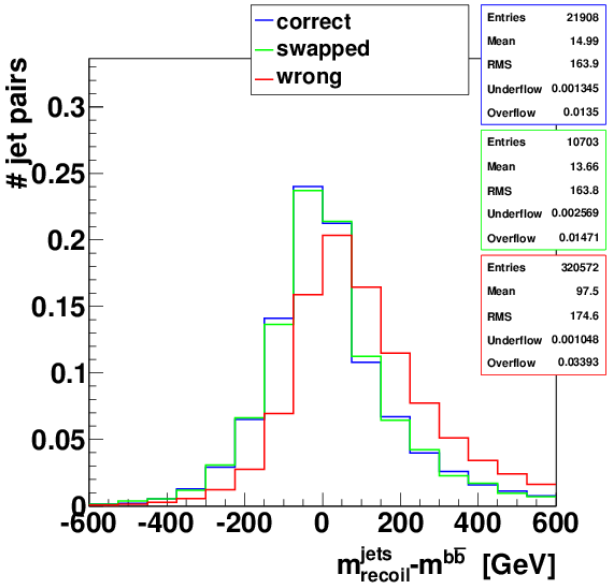


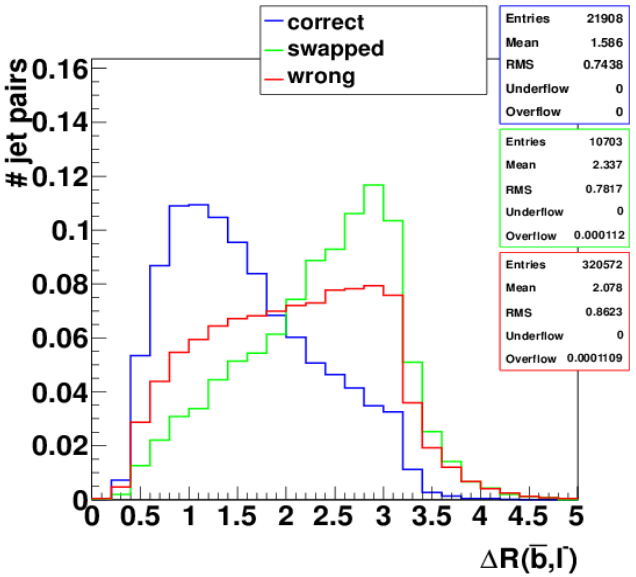
Some MVA input variables

- The training input variables for the MVA are tt system specific (do not depend on kinematics of the Higgs).
- Here we observe the three pairing options. Not always the separation between the three pairings is clear.
- Still searching for new possible variables.



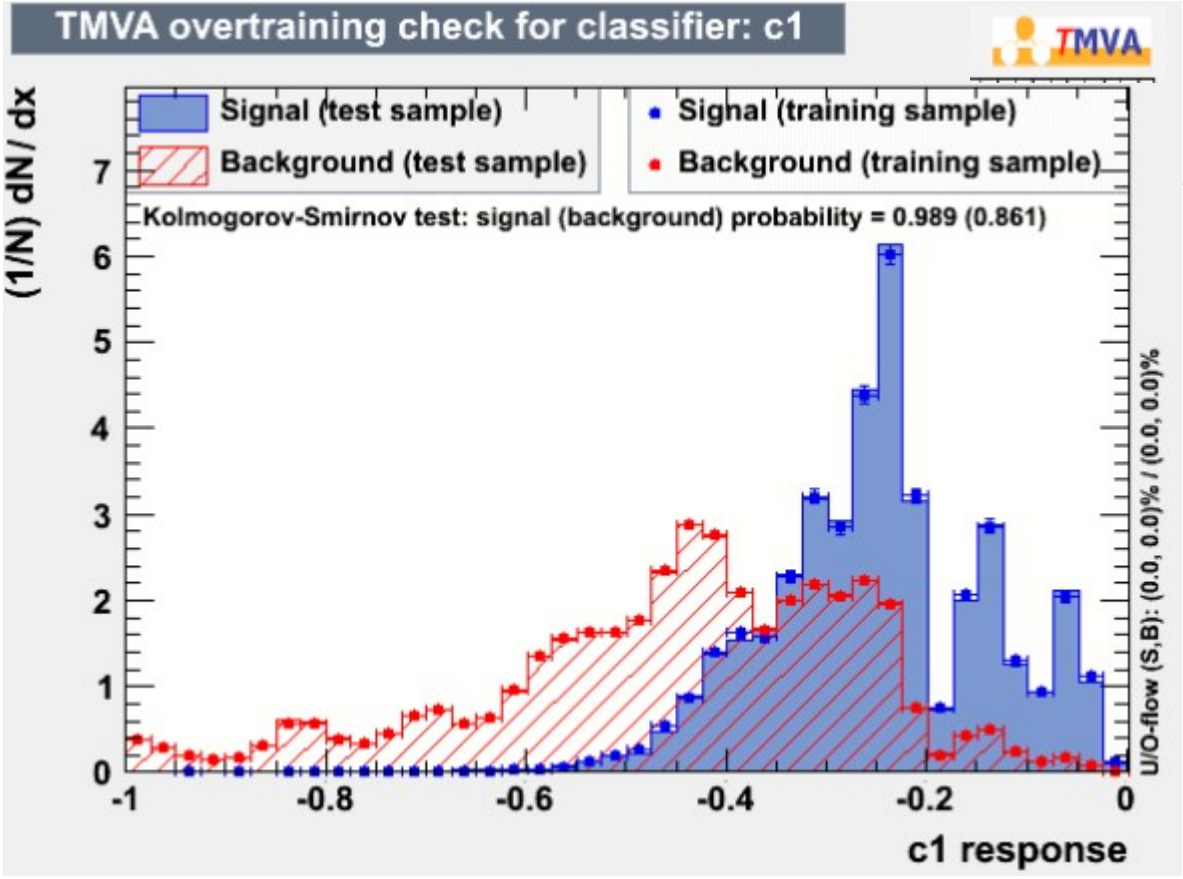


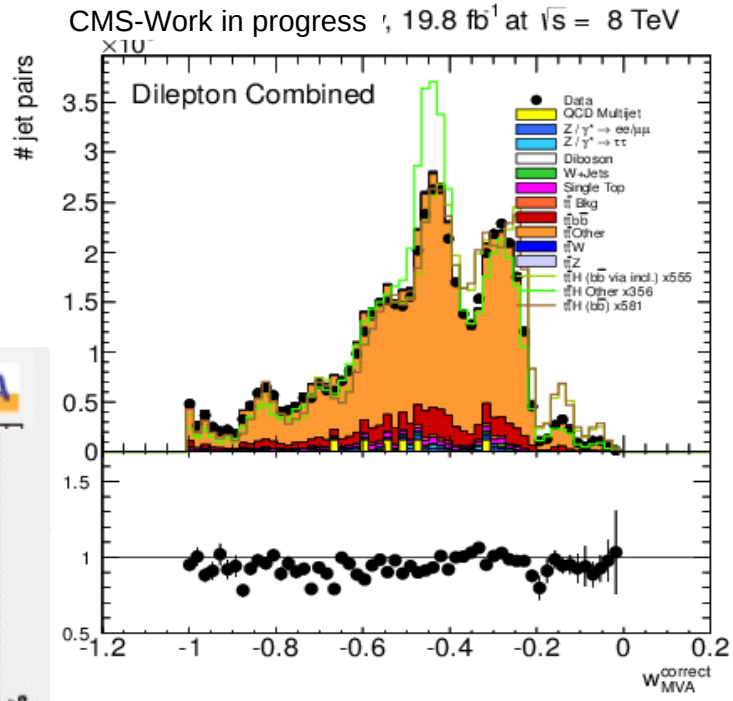




Separation power

- Good agreement for the MVA weights results.
- Highest MVA weight assigned to top pair. From the remaining, the two with the best b-tag are associated to Higgs.

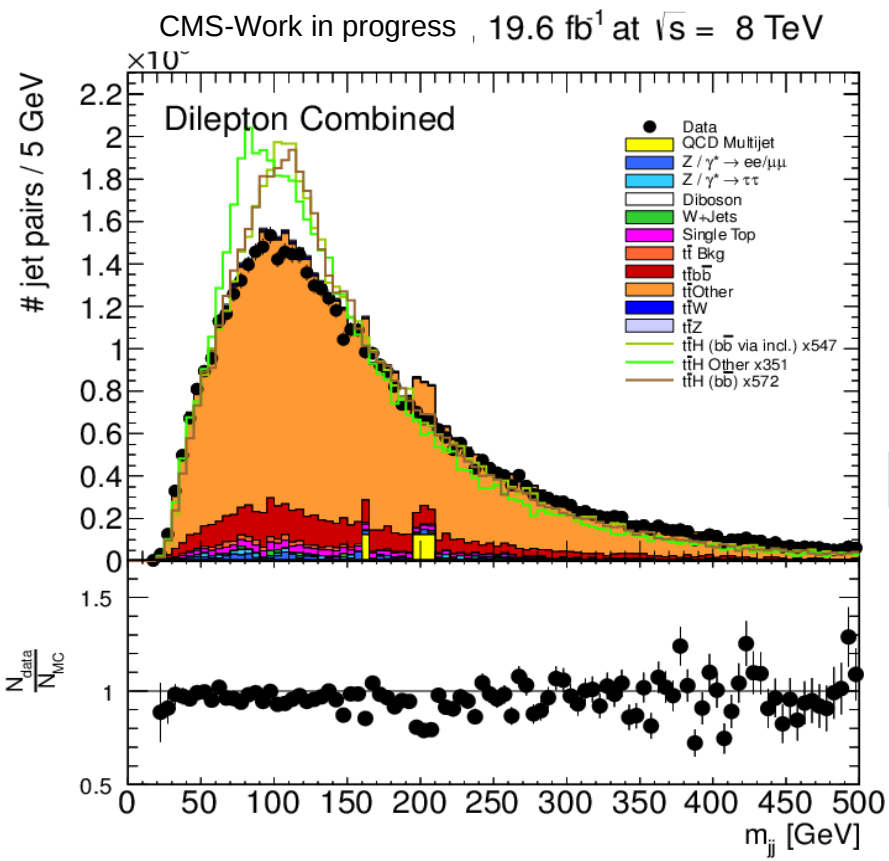




- Almost no overtraining.
- A better separation power performance still needed.

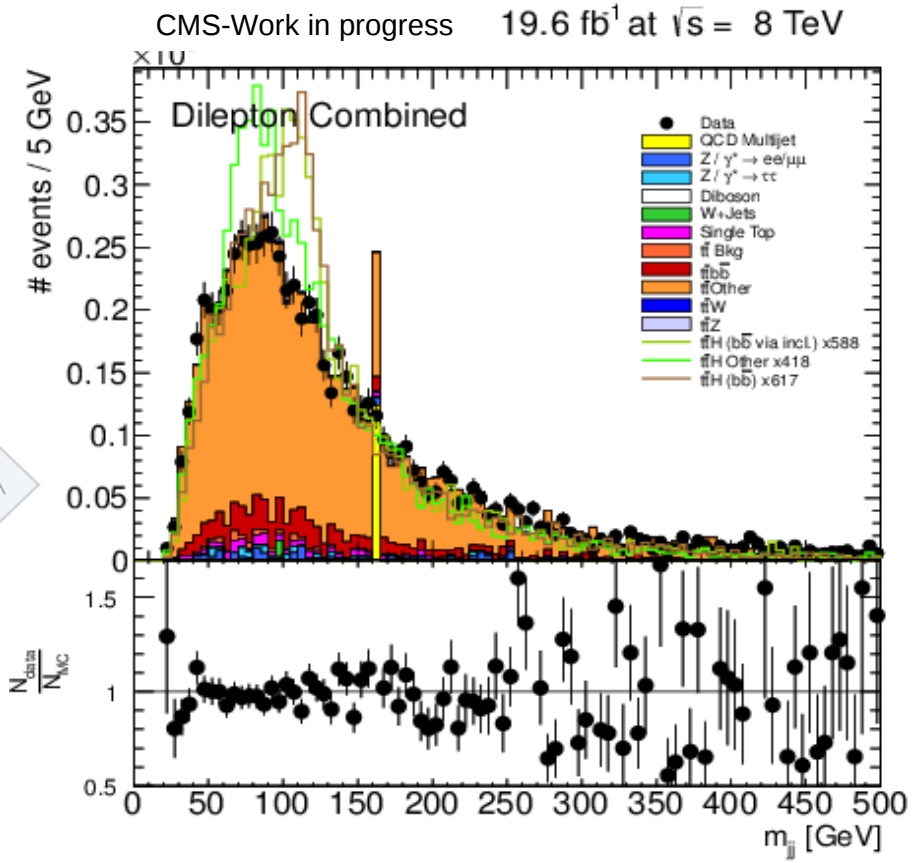
Dijet Mass

All jet combinations (at least 6/event)



MVA

Only correct combination (at most 1/event)



> Some improvement can already be seen.

Summary and conclusions

- The understanding of the top system is our first approach to separate it from the Higgs system.

- MVA seems to start giving some adequate separation power. However, a better understanding of the method and a finer tune of the parameters is needed and foreseen.

- At 8TeV still low statistics, but for higher energies (13-14 GeV) the analysis will be much set up giving rise to clearer results.

- There is a lot of room for many improvements (for example, a new kinematic reconstruction is under study and implementation).

BACKUP SLIDES

Motivation: Why is ttH important?

> Add information about couplings

- Only production mode directly sensitive to top-Higgs
- Add sensitivity to other Higgs coupling measurements

> Search for BSM physics

- Large deviations from SM Higgs couplings not likely
- Some models have enhanced ttH production without changing Higgs BR
- Examples:

  → Vector-like heavy top partner

  → Compositeness, RC, little Higgs

Analysis setup and event selection

- Analysis based on setup for top-pair differential cross-section measurement in dilepton decay channels
  - Common tools, samples, selections, …
- Basic event selection (identical to top analysis)
  - Dilepton triggers (ee, eμ, μμ)
  - 2 oppositely charged isolated leptons with $p_t > 20$ GeV and $|\eta| < 2.4$
  - Dilepton mass $m_{ll} > 20$ GeV (for ee, μμ exclude also Z window 76 GeV < m_{ll} < 106 GeV)
  - For ee, μμ also MET > 40 GeV
  - ≥2 jets with $p_t > 30$ GeV and $|\eta| < 2.4$
  - ≥1 b-tagged jet (CSV loose)
- Further selections
  - Categories in terms of (# jets, # b jets)
  - Background dominated categories (low multiplicities), signal enhanced categories

More about...

- Top cross-section analysis in the dileptonic and semileptonic channel at 8TeV:
  - CMS-PAS TOP-12-027, CMS-PAS TOP-12-028, CMS-PAS TOP-12-042 (Sep'13 results)
  - ATLAS-CONF 2013-099 (Sep'13 results)

M. Aldaya, TOP2013

- TTH (H→bb) analysis in the dileptonic and semileptonic channel:
  - CMS-PAS HIG-13-019
  - ATLAS-CONF 12-135

K. Lannon, TOP2013