Multivariative analysis of $H \to b\bar{b}$ in $t\bar{t}H$ associated production mode in ATLAS with fast simulation

Sergey Kotov

MPI für Physik

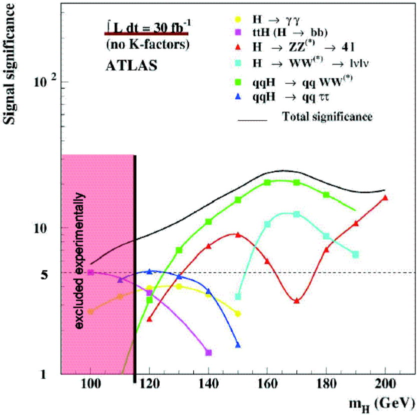
DPG meeting, Berlin, Mar 3-10, 2005

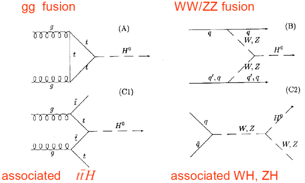
Max
Planck
Institut für Physik

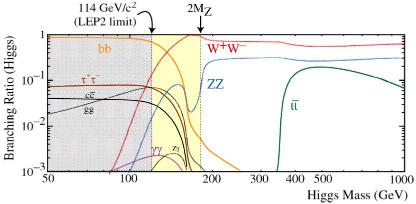
Low mass SM Higgs boson overview

LEP2 experimental bounds on Higgs mass

- precision measurements of EW observables: $m_H = 117^{+67}_{-45}$ GeV
- direct searches: $m_H > 114$ GeV







Signature channels for low mass SM Higgs

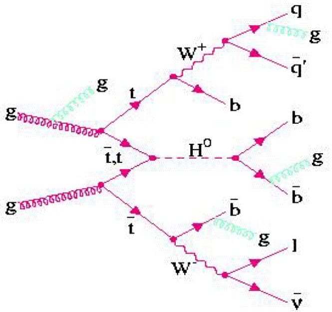
- $H \to \tau^+\tau^-$ in VBF production mode
- $H \to WW^* \to l\nu l\nu$ in VBF production mode
- $H \to \gamma\gamma$ in gluon fusion production mode
- $H \to b\bar{b}$ in $t\bar{t}H$ associated production mode

Channel description

Features of $t\bar{t}H$, $H \to b\bar{b}$ channel

- Complex final state
  - 6 jets: 4 b-jets and 2 light jets
  - 1 high-p_t lepton (trigger)
  - missing energy E_t^{miss} from neutrino
  - additional jets from ISR/FSR
- Large backgrounds
  - combinatorial from mis-pairing of b-jets in signal events
  - irreducible from $t\bar{t}b\bar{b}$ events
  - reducible from $t\bar{t}jj$ events

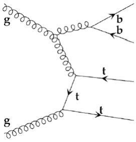
full reconstruction of event and very good b-jet tagging are needed to suppress backgrounds



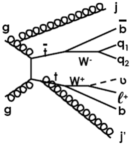
$t\bar{t}H$ signal process diagram

Signal and background cross sections

Process	$t\bar{t}H$ (m_H=120 GeV)	$t\bar{t}b\bar{b}$	$t\bar{t}jj$ ($\geq$ 6 jets)
σ, pb	0.5	9.2	72

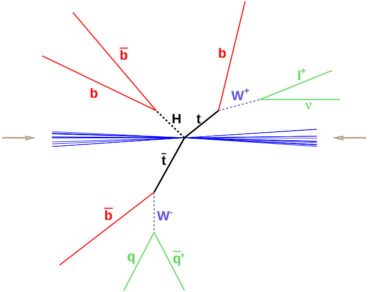


$t\bar{t}b\bar{b}$ background



$t\bar{t}jj$ background

Event topology, preselection and data samples



PYTHIA 6.2 and AcerMC 2.3 programs are used for event generation
ATLFAST package is used for fast simulation of ATLAS detector response

Preselection: event topology cuts

- 1 isolated lepton with p_t > 20(25) GeV for $\mu(e)$ and η < 3
- $\geq$ 6 jets with p_t > 20 GeV and η < 5
- $\geq$ 4 of jets identified as b-jets

Data samples

Process	$t\bar{t}H$	$t\bar{t}b\bar{b}$	$t\bar{t}jj$
Generated events, M	0.6	1	1
Preselection efficiency, %	2.1	0.8	0.02

ATLFAST underestimates the b-tagging efficiency in events with multiple b-jets (ϵ_b=0.42 compared to ϵ_b=0.60-0.65 in full simulation), so we use an efficiency scale factor of 3 in our final calculations

Event reconstruction

Main reconstruction challenge

How to reconstruct $t\bar{t}$-pair from all possible $bl\nu - bjj$ combinations properly with correct assignment of b-jets?

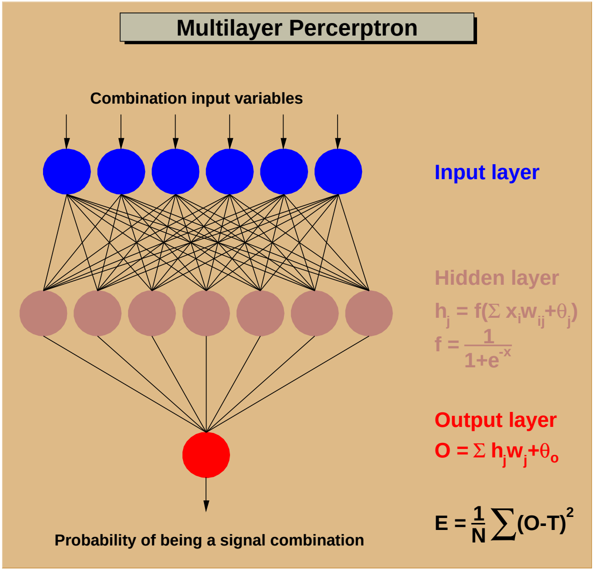
Various approaches to $t\bar{t}$-pair reconstruction

- ATLAS Technical Design Report: select combination with minimal $\Delta^2 = (m_{bl\nu} - m_t)^2 + (m_{bjj} - m_t)^2$
- a recent improved approach which uses likelihood techniques for reconstruction of $t\bar{t}$-pair
- this analysis uses neural network technique for $t\bar{t}$-pair reconstruction

Reconstruction strategy with Neural Network

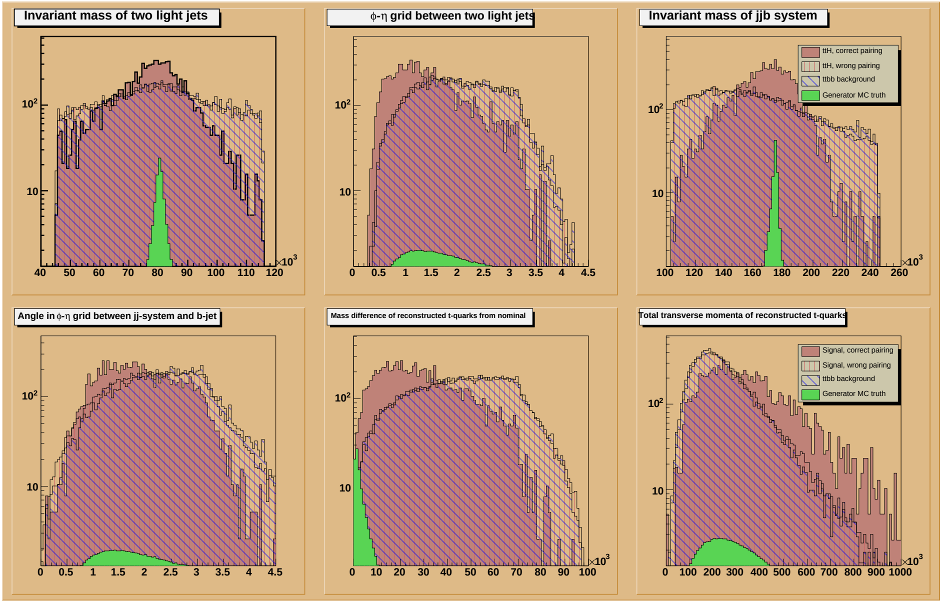
- use events which pass preselection criteria (1 lepton, 4 b-jets, 2 light jets)
- determine p_ν from p_l and p_{miss} using m_W constraint (if fails, use approximation $p_\nu^z = p_l^z$)
- select all possible reconstructed combinations of lepton, neutrino, 2 light jets and 2 b-jets for which the reconstructed invariant masses m_{jj}, $m_{bl\nu}$, and m_{bjj} fit inside some mass windows of W boson and t-quark (30 GeV and 70 GeV respectively)
- feed parameters of these combinations through a neural network (which was trained beforehand on a sample of combinations matched and non-matched to MC generator truth table) and select combination with the highest NN output value
- assign the remaining 2 b-jets to the Higgs boson and plot their invariant mass m_{bb}

Neural network basics

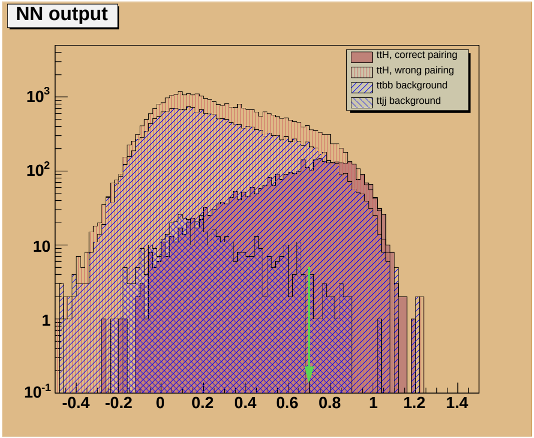


- 14 input variables: 5 invariant mass differences, 8 angular separations in ϕ-η plane, 1 sum of transverse momenta
- TMultiLayerPerceptron ROOT built-in class is used as neural network
- 6000 of matched and 21000 of non-matched combinations were used (with proper weights) to train the neural network

Neural network variables



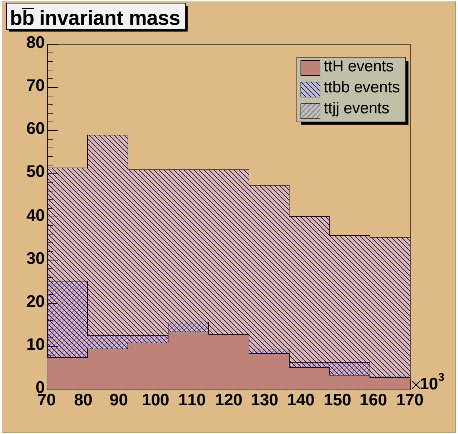
Signal/background separation with neural network



Effectiveness of the neural network for background suppression

Cut	ϵ_{ttH}	ϵ_{cmbn}	ϵ_{ttbb}	ϵ_{ttjj}
0.5	0.768	0.238	0.209	0.163
0.6	0.657	0.157	0.139	0.090
0.7	0.510	0.093	0.080	0.046
0.8	0.330	0.044	0.038	0.023

Analysis results



Expected number of events for $L = 30 fb^{-1}$

ϵ scale factor	3	2	1	Likelihood
ϵ, %	1.62	1.08	0.54	1.32
N_{ttH}	73.4	49.0	24.5	41.9
N_{ttbb}	421.5	281.0	140.5	164.2
N_{ttjj}	97.3	64.9	32.4	54.6
$S/\sqrt{B}$	3.2	2.6	1.9	2.8
Purity	0.33	0.33	0.33	0.29

Purity of reconstructed $t\bar{t}H$ events (fraction of events with all 4 b-jets correctly assigned) is important for finding of the Higgs mass peak

Summary

Conclusions

- a neural network approach was tried for reconstruction of $t\bar{t}$-pair in $t\bar{t}H$ events produced with fast simulation of the ATLAS detector $\Rightarrow$ no considerable improvement with comparison to previously used likelihood approach was achieved
- the obtained signal significance is still rather low, in large due to imperfections of the fast simulation algorithms

Future plans

- ATLAS collaboration now is in the process of mass production of full simulation data for its summer physics workshop in Rome $\Rightarrow$ first results show a HUGE improvement in b-tagging due to development of new combined algorithms
- the neural network approach will be re-tried on this sample of fully simulated events as soon as it will be available
- more elaborate consideration of the background sources is need to improve signal significance (ATLFAST program needs to be tuned for this using full simulation data, because it is impossible to produce the required amount of background events with full simulation of the ATLAS detector)