Measurement of the inclusive $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV and determination of the top quark pole mass

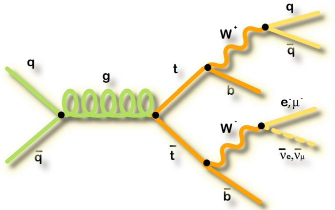
Top Workshop 2016

Jiří Franc
and
DØ Collaboration

September 20, 2016

Czech Technical University in Prague
Faculty of Nuclear Sciences and Physical Engineering

Top pair production via strong interaction - The goal of this analysis



Top Quark & Tevatron:

- Worlds largest $p\bar{p}$ data set for a long time.
- Center Mass energy $\sqrt{s} = 1.96$ TeV.
- **Production at Tevatron:**
  $\approx 85\%$ by $q\bar{q}$ annihilation
  $\approx 15\%$ by gg fusion
- **Top decay:** $t \rightarrow W + b \approx 100\%$

Samples are classified according to W-decay: $\ell +$jets and $\ell\ell$ channels are analyzed.

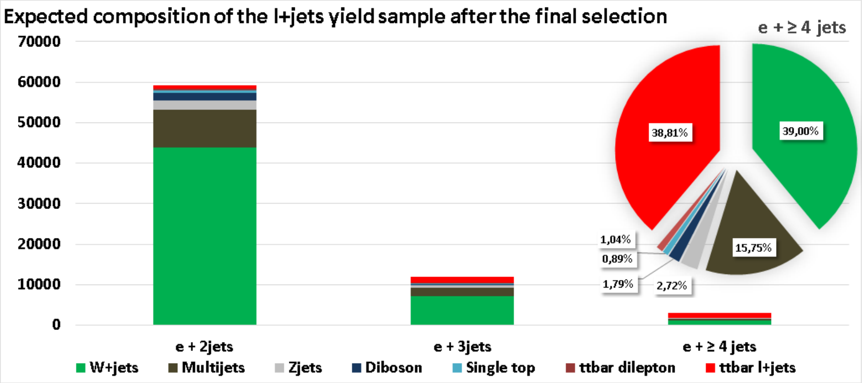
This analysis is based on **full Tevatron data sample** recorded by the D0 detector and corresponding to an integrated luminosity of **9.7fb^{-1}**.

The Goal:

- Measurement of the $t\bar{t}$ cross section using MVA methods.
- Extract the top quark mass.

$\ell +$ jets Yield table:

6 analysis channels in $\ell +$ jets : to the lepton type and the number of jets (2, 3, $\geq$ 4)

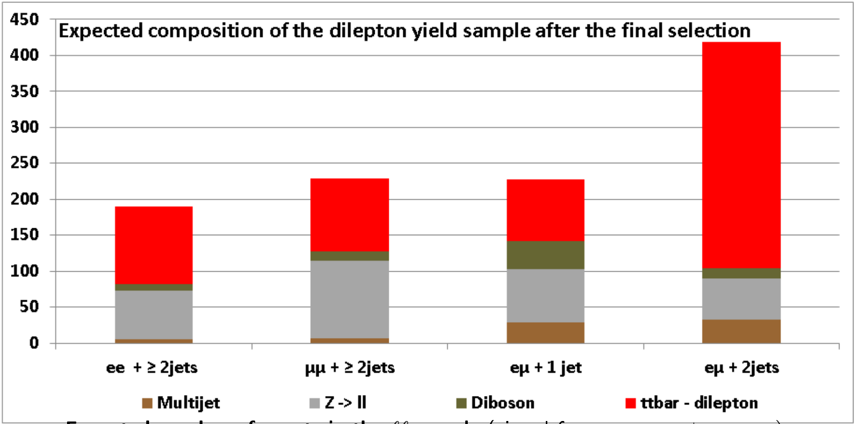


Expected number of events in the $\ell +$ jets sample(signal from ALPGEN+PYTHIA):

Process	$e + 2$ jets	$e + 3$ jets	$e + \geq 4$ jets	$\mu + 2$ jets	$\mu + 3$ jets	$\mu + \geq 4$ jets
$\sum$ bknd	58479±2900	10465±650	1834±140	44381±1650	8123±350	1402±80
$t\bar{t}$, $\ell +$ jets	669± 30	1460± 70	1177± 60	393± 20	1002± 50	909±50
$\sum$ MC	59148	11925	3011	44773	9125	2310
Data	59122	11905	3007	44736	9098	2325

$\ell\ell$ yield table:

4 analysis channels in dilepton decay channel:

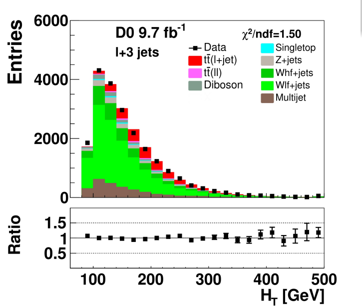


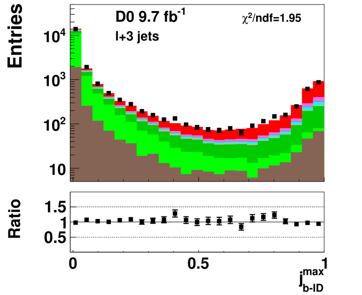
Expected number of events in the $\ell\ell$ sample (signal from ALPGEN+PYTHIA):

Process	$ee+\geq 2$ jets	$\mu\mu+\geq 2$ jets	$e\mu+1$ jets	$e\mu+\geq 2$ jets
$\sum$ background	82.2±18	172.2±22	141.4±18	104.7±15
$t\bar{t}$, $\ell\ell$	107.7±15	101.5±12	86.5±11	313.7±38
$\sum$ (sig + bg)	190	229	228	418
Data	215	242	236	465

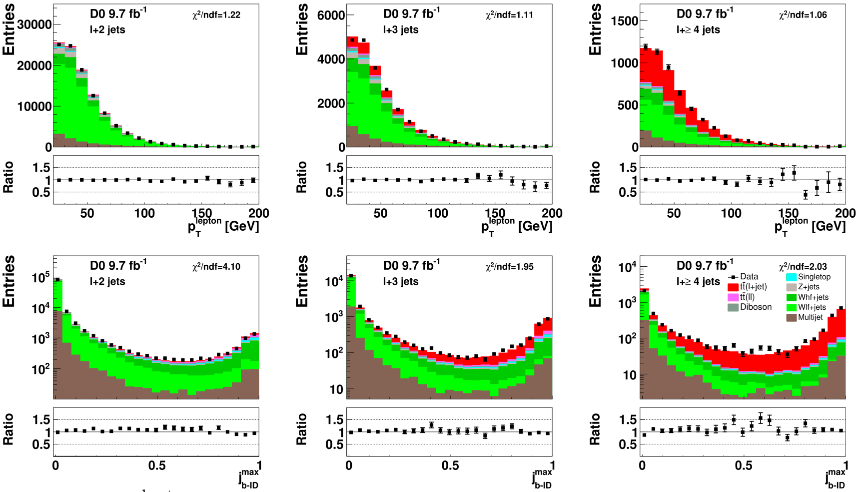
Variables selection in $\ell+$jets decay channel

- 46 kinematical and topological variables have been investigated (e.g. H_T, p_T^{lepton}, Centrality, Sphericity, $M_T(jets)$, $p_T^{j_i}$, , ...).
- 25 variables with good MC vs. Data agreement and good separation power between signal and background have been selected.
- Selection has been done by combination of different rankings: (Hypothesis tests Kolmogorov-Smirnov, Anderson-Darling, χ^2, and TMVA ranking).
- No b-ID information is used in the selection, but the TMVA includes the maximal MVA b-ID value of all jets, $j_{\mathrm{b-ID}}^{\mathrm{max}}$.





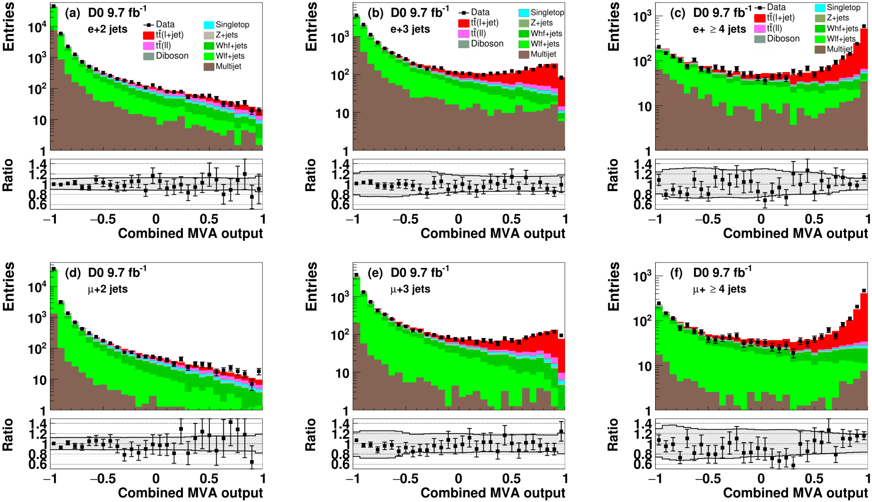
Control plots in $\ell +$jets channel



Distributions of p_T^{lepton} (upper row) and $j_{\text{b-ID}}^{\text{max}}$ (lower row), for events with a lepton and 2, 3 or 4 or more jets.

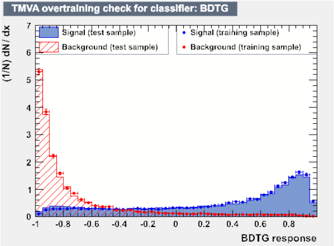
$\ell +$ jets discrimination by TMVA - Boosted Decision Trees with Gradients

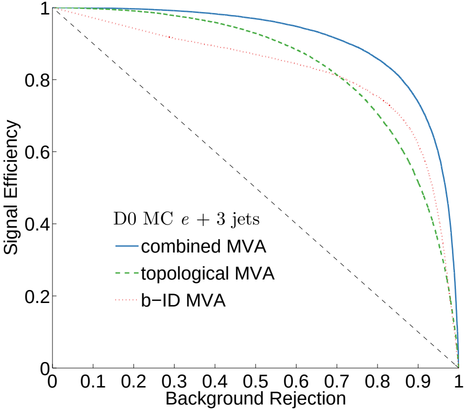
The MVA BDTG output distributions using a theoretical $\sigma_{t\bar{t}} = 7.48$ pb and $m_t = 172.5$ GeV



$\ell+$jets discrimination by TMVA - Boosted Decision Trees with Gradients

- TMVA BDTG with gradient boost trained on 25 types of variables + $j^{\text{lead}}_{\text{b-ID mva}}$.
- Each individual background contribution was used in the training and verified that there is no bias due to overtraining of the method.

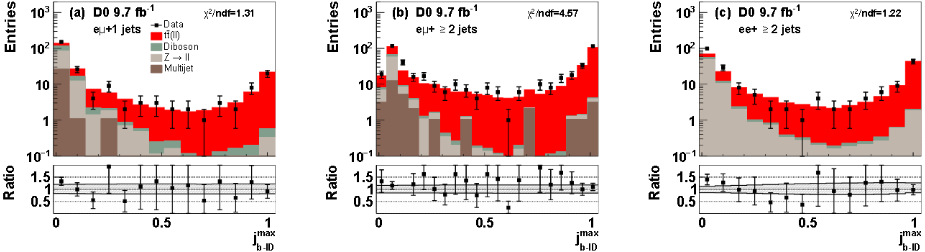




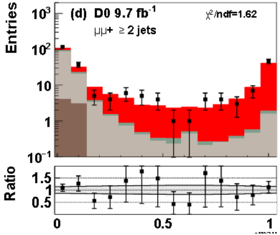
- The combination of pure topological and MVA b-ID method improved the separation by 10%.
- The area under the ROC curve is around 0.9 for all 6 analysis channels when using BDTG method.

$\ell\ell$ decay channel discrimination

The pre-fit *b*-ID MVA output distributions in the dilepton decay channel.



- In the $\ell\ell$ channel discriminantion is solely based on the output distribution of the MVA employed to identify *b*-tagged jets: $j^{\mathrm{max}}_{\mathrm{b-ID}}$.
- The shape of the $j^{\mathrm{max}}_{\mathrm{b-ID}}$ variable allows one to distinguish between $t\bar{t}$ events located at high output values and the most dominant $Z/\gamma^{\star}$ +jetsbackground contribution typically located at low output values.



Sources of grouped postfit systematic uncertainties for the $\sigma_{t\bar{t}}$ measurement

Each source of systematic uncertainty yields a modified discriminant distribution.

A log-likelihood profile fit of MC templates to data using a nuisance parameter for every source of systematic uncertainty has been performed:

$$\mathcal{L}(\vec{s},\vec{b}|\vec{n},\vec{\theta})\times\pi(\vec{\theta}) = \prod_{i=1}^{N_C}\prod_{j=1}^{N_{\text{bins}}} \mu_{ij}^{n_{ij}} \frac{e^{-\mu_{ij}}}{n_{ij}!} \times \prod_{k=1}^{n_{\text{sys}}} e^{-\theta_k^2/2}$$

50 different systematic uncertainties has been taken into account (list in backup).

Source of uncertainty	$\delta_{\ell+\text{jets}}$, pb	$\delta_{\ell\ell}$, pb	$\delta_{\text{comb},}$, pb
Signal modeling			
Signal generator	±0.21	±0.05	±0.17
Hadronization	±0.26	±0.33	±0.25
Color reconnection	±0.08	±0.05	±0.09
ISR/FSR variation	±0.08	±0.04	±0.06
PDF	±0.04	±0.03	±0.02
Detector modeling			
Jet modeling & ID	±0.11	±0.08	±0.04
b-jet modeling & ID	±0.27	±0.26	±0.23
Lepton modeling & ID	±0.20	±0.26	±0.17
Trigger efficiency	±0.32	±0.08	±0.16
Luminosity	±0.30	±0.30	±0.27
Sample Composition			
MC cross sections	±0.07	±0.13	±0.09
Multijet contribution	±0.11	±0.02	±0.10
W+jets scale factor	±0.21	±0.01	±0.15
Z/γ^*+jets scale factor	±0.07	±0.11	±0.12
MC statistics	±0.01	±0.01	±0.02
Total systematic uncertainty	±0.67	±0.73	±0.55

The $\sigma_{t\bar{t}}$ measurement and nuisance parameter fit of MC to Data are performed using the software package **Collie** (A Confidence Level Limit Evaluator[1])

[1] CDF & D0 Collaborations, V. M. Abazov et al., Higgs Boson Studies at the Tevatron, Phys. Rev. D 88, 052014).

Inclusive Cross Section Results

For a top quark mass of 172.5 GeV we measure
in the $\ell +$jets decay channel:

$$\sigma_{t\bar{t}} = 7.33 \pm 0.14\,(\text{stat.})\,^{+0.71}_{-0.61}\,(\text{syst.}),$$

with a relative total uncertainty of 9.2%.
in the $\ell\ell$ decay channel:

$$\sigma_{t\bar{t}} = 7.58 \pm 0.35\,(\text{stat.})\,^{+0.69}_{-0.58}\,(\text{syst.}),$$

with a relative total uncertainty of 9.6%.

in the **combination** of $\ell +$jets and $\ell\ell$ channels :

$$\boldsymbol{\sigma_{t\bar{t}} = 7.26 \pm 0.13\,(\text{stat.})\,^{+0.57}_{-0.50}\,(\text{syst.})},$$

with a relative total **uncertainty of 7.6%**.

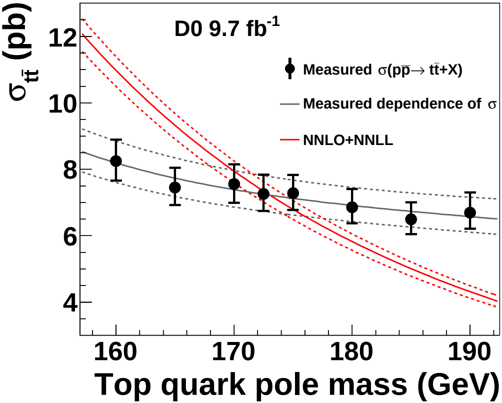
The combined $\ell +$jets and $\ell\ell$ $t\bar{t}$ cross section does not coincide with the weighted average of the individual results, which is attributed to the effect of correlations.

Top Quark Pole Mass Determination - from the inclusive $t\bar{t}$ xsec

- The measured $t\bar{t}$ production cross section dependence on the top quark mass parametrized by a quartic function and compared to the dependence provided by the NNLO pQCD calculation TOP++.

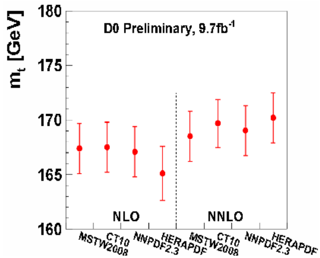
Pole mass of the Top Quark:

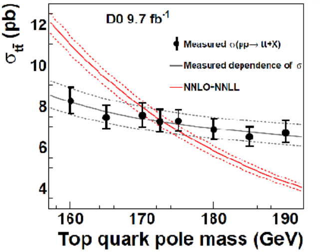
$$m_t = 172.8 \pm 1.1(theo.)^{+3.3}_{-3.1}(exp.) GeV$$

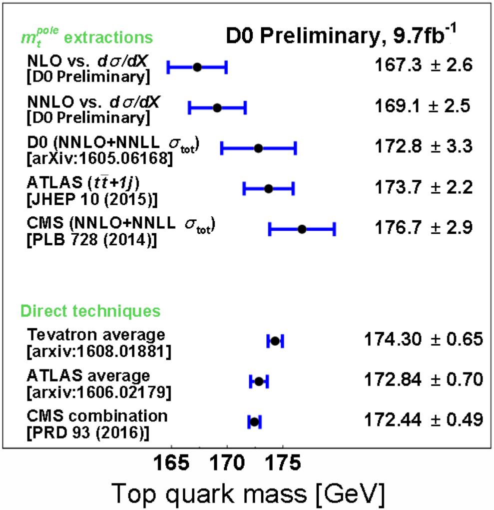


- The precision is **1.9%** and represents the **most precise** determination of the top quark pole mass from the **inclusive** $t\bar{t}$ cross section **at the Tevatron**.

NEW: Top Quark Pole Mass Determination from Differential Cross Section







Comparison of pole mass from inclusive (1.9% precison) and differential cross section (1.5 % precison).

Summary and The End of The Talk

Final result of the cross section measurement at $D\emptyset$

- We perform a simultaneous loglikelihood fit to profile systematic uncertainties and, for a top quark mass of 172.5 GeV, we measure a combined $t\bar{t}$ cross section of

$$\sigma_{t\bar{t}} = 7.26 \pm 0.13\,(\mathrm{stat.})\,^{+0.57}_{-0.50}\,(\mathrm{syst.}),$$

which corresponds to a relative uncertainty of 7.6%.

- The Top Quark pole mass has been computed:

$$m_t = 172.8^{+3.4}_{-3.2}\,(\mathrm{tot.})\,\mathrm{GeV}$$

.

Analysis and results are available: arXiv:1605.06168.

Summary and The End of The Talk

Final result of the cross section measurement at $D\emptyset$

- We perform a simultaneous loglikelihood fit to profile systematic uncertainties and, for a top quark mass of 172.5 GeV, we measure a combined $t\bar{t}$ cross section of

$$\sigma_{t\bar{t}} = 7.26 \pm 0.13\,(\mathrm{stat.})\,^{+0.57}_{-0.50}\,(\mathrm{syst.}),$$

which corresponds to a relative uncertainty of 7.6%.

- The Top Quark pole mass has been computed:

$$m_t = 172.8^{+3.4}_{-3.2}\,(\mathrm{tot.})\,\mathrm{GeV}$$

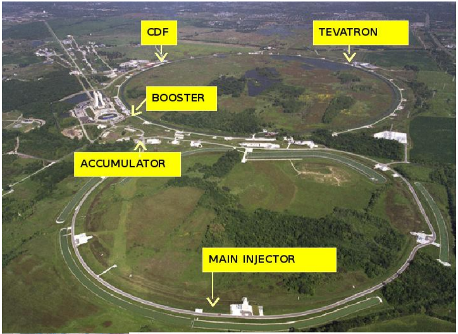
.

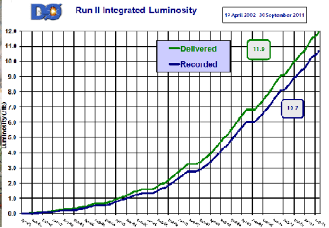
Analysis and results are available: arXiv:1605.06168.

This is the end => Thank you for your attention.

Backup

DØ experiment



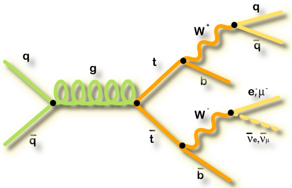


Tevatron - $p\bar{p}$ collider

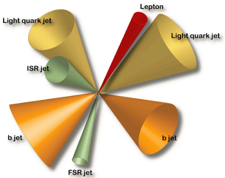
- **Tevatron:** $p\bar{p}$ circular particle accelerator (6.86 km).
- **Unique data set:** worlds largest $p\bar{p}$ data set for a long time.
- **Center Mass energy:** $\sqrt{s} = 1.96$ TeV.
- **Experiments:** CDF and DØ with well understood detectors.
- **Run II:** begun in 2001 and each experiment recorded $\approx 10fb^{-1}$ until September 2011.
- Presenting measurement has been done with the **full dataset 9.7fb^{-1}**.

Strong interaction: Top pair production

Top is the heaviest fundamental particle discovered so far!



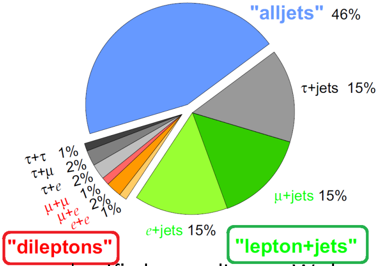
Feynman diagram of the $t\bar{t}$ production.



Situation in the detector
(+ missing transverse energy)

- **Top Quark:**
  - **Mass:** $m_t = 173.2 \pm 0.87\, GeV$
  - **Lifetime:** $t \approx 5 \times 10^{-25} s << \Gamma_{QCD}$
  - **Production at Tevatron:**
    - $\approx 85\%$ by $q\bar{q}$ annihilation
    - $\approx 15\%$ by gg fusion
  - **Top decay:** $t \rightarrow W + b \approx 100\%$
  - **Different decay channels:**

Top Pair Branching Fractions



Samples are classified according to W-decay:
$l + jets$ channel is under concern in this analysis.

Top pair production cross section measurement in $\ell +$ jets channel

Data sample: Full Data Set ($9.7 fb^{-1}$) with selection: Phys.Rev.D 90,092006 (2014)
The general selection criteria applied to both the $\ell +$ jets and $\ell\ell$ decay channels:

1. Accepted events: PV within $|z_{\mathrm{PV}}| < 60$ cm of the center of the detector along the beam axis.
2. The number of tracks associated with the PV is greater or equal three.
3. After correcting the jet energy to the particle level, only jets with a transverse momentum $p_T > 20$ GeV and $|\eta| < 2.5$ are selected.
4. Jets which satisfy the b-tagging requirement are required to have at least two tracks coming from the PV.
5. Identified leptons are required to originate from the PV by demanding $|\Delta z(\ell, \mathrm{PV})| < 1$ cm. These z-values correspond to the point of closest approach to the beam line of these tracks.
6. To ensure that electrons are isolated, an angular separation in ΔR of at least 0.5 between an electron and the closest jet is required.

$\ell +$ jets

variable	kinematic range
lepton $\eta(\ell)$	$\lvert\eta(e)\rvert < 1.1$, $\lvert\eta(\mu)\rvert < 2.0$
lepton $p_T(l)$	$p_T(l) > 20$ GeV
$\not{E}_T$	$\not{E}_T > 20$ GeV
jet $\eta(jet)$	$\lvert\eta(jet)\rvert < 2.5$
jet $p_T(jet)$	$p_T(jet) > 20$ GeV

+ additional cuts

$\ell\ell$

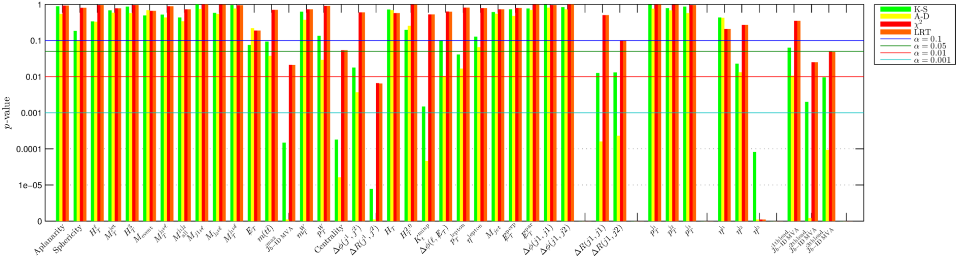
variable	kinematic range
electron	$\lvert\eta(e)\rvert < 2.5$, except $1.1 < \lvert\eta(e)\rvert < 1.$
	$p_T(e) > 20$ GeV
muon	$\lvert\eta(\mu)\rvert < 2.0$
	$p_T(e) > 15$ GeV
	($p_T(e) < 200$ GeV in dimuon)

List of investigated variables

Aplanarity: Diagonalizing the normalized quadratic momentum tensor M yields three eigenvalues λ_i and is defined as $\frac{2}{3}\lambda_3$ and measures the flatness of an event.
Sphericity: Similar to Aplanarit and is defined as $\frac{2}{3}\lambda_2 + \frac{2}{3}\lambda_3$. $t\bar{t}$ events show a more spherical behavior typical for heavy object decays
Centrality: Ratio of the scalar sum of the transverse momentum of all jets to the energy of all jets.
H_T: The scalar sum of the transverse momenta of all jets, the lepton and $\not{E}_T$.
H_T^{ℓ}: The scalar sum of the transverse momenta of all jets and the lepton.
H_T^3: The scalar sum of transverse momenta of jets starting with the 3rd jet multiplicity bin.
$H_T^{2.0}$: The scalar sum of transverse momenta of jets satisfying $|\eta| < 2.0$.
$\not{E}_T$: Missing transverse momentum.
$\not{E}_T{}^{par}$: Missing transverse momentum parallel to.
$\Delta\phi(l,\not{E}_T)$: The separation in azimuth between the lepton and the direction of $\not{E}_T$.
m_{jet}: The invariant mass of the jets.
M_T^{jet}: The transverse mass of the first two leading jets.
M_{event}: The invariant mass of the jets, lepton and the neutrino.

$M_{all}^{j_1 j_2}$: The invariant mass of the system consisting of the leading and second leading jet divided by the total invariant mass of the event.
$M_T^{j_1 \nu l}$: The transverse mass of the system consisting of the leading jet, the neutrino and the lepton.
$M_{j_2 \nu \ell}$: The invariant mass of the system consisting of the second leading jet, the neutrino and the lepton.
$M_T^{j_2 \nu \ell}$: The transverse mass of the system consisting of the second leading jet, the neutrino and the lepton.
$p_T^{j_i}$: The transverse momentum of the individual jets i.
η^{j_i}: The rapidity of the leading jet.
$\Delta\phi(j^1, j^2)$: The separation in azimuth between the leading and second leading jet.
$\Delta R(j_1, j_2)$: The separation in the distance R between the leading and second leading jet.
$j_{\text{b-ID mva}}^{\text{lead}}$: The maximum output value of the MVA b-jet discriminant for the leading jet.
p_T^W: The transverse momentum of the reconstructed W boson, which decays hadronically.
$m(t\bar{t})$: The invariant mass of the $t\bar{t}$ pair.
K_t^{minp}: ΔR_{min} between 2 jets multiplied by minimal p_T and divided by scalar sum of the p_T of the lepton and $\not{E}_T$.

Variables selection - Hypothesis testing (Example of e + 3 jets channel)



3 tools have been used:

- ROOT - KStest, Chi2Test.
- Hypothesis tests based on weighted EDF and Φ divergences.
- TMVA ranking
  - variables ranking, TMVA method ranking

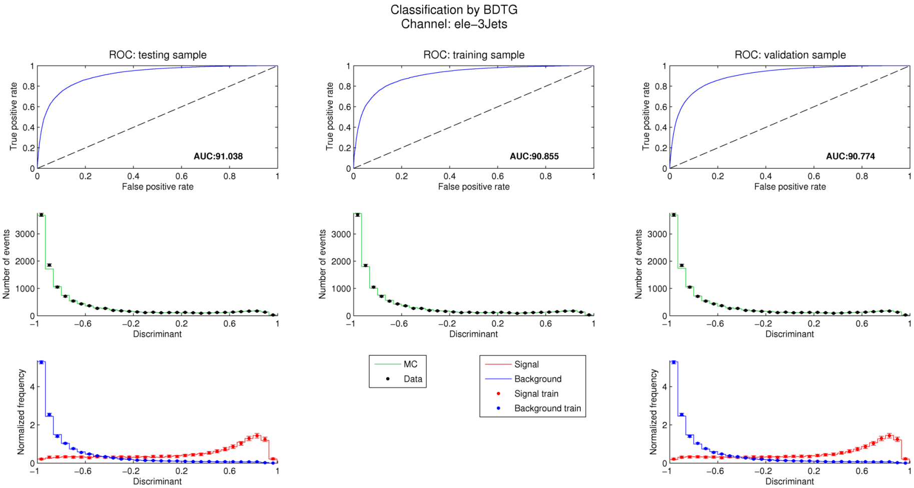
Statistical hypothesis testing

- Kolmogorov-Smirnov test.
- Anderson - Darling test.
- χ^2 Goodness of Fit test.
- Likelihood ratio test.

- weighted empirical distribution function,
- quantile binning.

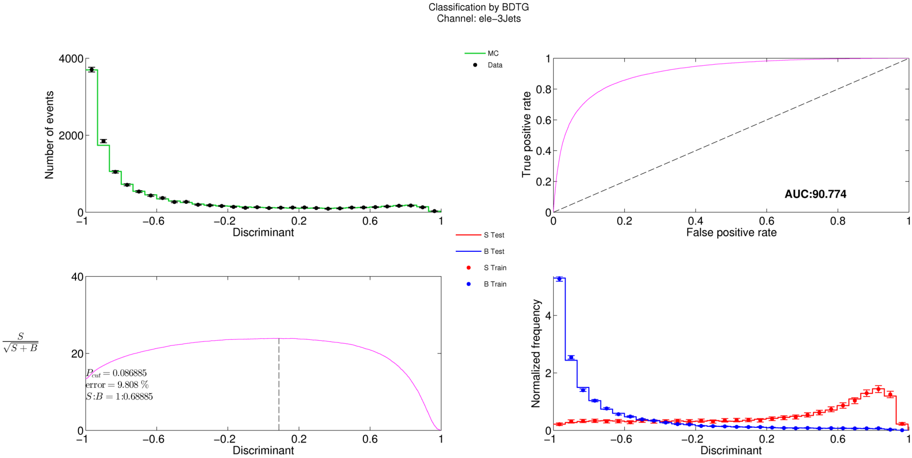
More detailed mathematical description of the tests are in (arXiv:1412.1076)

Discrimination by BDTG - overtraing check:



Similar behave in all $\ell +$ jets analysis channels.

Discrimination by BDTG - optimal cutting:



Signal rate in MC - Improvement:

	Before discrimination	After discrimination with optimal cut
Electron 3jb	12.24%	59.24%

Used Systematics:

Flat systematics:
Diboson Xsec, Single Top Xsec, Z Xesc, Data Quality removes bad events, Fake & signal eff uncertainty, Luminosity, MC BR uncertainty (PDG), Muon ID, Muon isolation, Muon Track, Trigger efficency, Wlp SF, Whf SF.

Shape dependent systematics:
b fragmentation, B-tagging, C-tagging, Jet energy resolution, Jet energy scale, Jet identification, Lepton ID, Lepton Momentum, Light-tagging, PDF, Sample dependent corrs, Taggability, Vertex confirmation, Z & W pT reweighting, Z vertex reweighting.

Signal related uncertainties:

- ISRFSR (initial state radiation vs. final state radiation).
- Color reconnection (P2011 vs. P2011NOCR).
- Hadronization (Alpgen+Herwig vs. Alpgen+Pythia).
- Higher orders signal model (MC@NLO+Herwig vs. Alpgen+Herwig).

Correlation between l+jets and dilepton (1)

source	Systematic names in l+jets	Systematic names in dilepton	Systematic names in combination N-1 test	Corr
b fragmentation	bfrag_sys	bfrag_sys	b_quark_modeling	1
B-tagging	bTag_sys	bTag_sys	bTag	1
C-tagging	cTag_sys	-------------	-------------	0
Color reconnection	colorRecon_sys	colorRecon_sys	Color_reconnection	1
Data Quality removes bad events	eventdq_sys	eventdq_sys	DQ_event_selection	1
Diboson cross-section	diboson_xsec_sys	diboson_xsec_sys	Diboson_cross_section	1
dZ(lepton, PV)	------	dz_sys	dZ_lepton__PV_	0
Fake & signal eff uncertainty	epsQCDsig_elec_sys	-------------	-------------	0
Fake & signal eff uncertainty	epsQCDsig_muon_sys	-------------	-------------	0
Fake & signal eff uncertainty	-------------	mu_fake_stat_sys	Fit_statistical_error	0
Fake & signal eff uncertainty	-------------	em_fake_stat_sys	Fit_statistical_error	0
Fake & signal eff uncertainty	-------------	mu_fake_rate_sys	Fake_muon_rate	0
Hadronization model	AH_vs_AP_hadro_sys	AH_vs_AP_hadro_sys	AlpgenHerwig	1
ISR/FSR	ttA_isr_fsr_sys	ttA_isr_fsr_sys	ISR_FSR	1
Jet energy resolution	JER_sys	JER_sys	JER	1
Jet energy scale	JES_sys	JES_sys	JES	1
Jet identification	JetID_sys	JetID_sys	JetID	1
Lepton ID	lepid_sys	lepid_sys	Electron_ID__certification_	1
Lepton Momentum	LM_sys	-------------	-------------	0
Lepton Momentum	-------------	muon_res_sys	Muon_momentum_resolution	0
Lepton Momentum	-------------	emes_sys	Electron_energy_scale	0
Lepton Momentum	-------------	emres_sys	Electron_energy_resolution	0
Light-tagging	lTag_sys	lTag_sys	lTag	1
Luminosity	Lumi_sys	Lumi_sys	Luminosity	1

Correlation between l+jets and dilepton (2)

source	Systematic names in l+jets	Systematic names in dilepton	Systematic names in combination N-1 test	Corr
MC BR uncertainty (PDG)	MCBR_sys	MCBR_sys	_Wtoellnu__branching_ratio	1
MC statistics	----	dilepton_mcstat_sys	MC_statistics	**1**
Muon ID	muonid_sys	muonid_sys	Muon_ID	1
Muon isolation	muon_iso_sys	muon_iso_sys	Muon_isolation	1
Muon Track	muon_trk_sys	muon_trk_sys	Muon_track	1
Opposite charge	-------------	emcharge_missid_sys	Electron_charge_missid	0
Opposite charge	-------------	mucharge_missid_sys	Muon_charge_missid	0
PDF	pdf	pdf	pdf	1
Sample dependent corrs	SDC_sys	SDC_sys	SPR	1
Signal modeling	MH_vs_AH_signal_sys	MH_vs_AH_signal_sys	MCatNLOHerwig	1
Single Top cross-section	stop_xsec_sys	-------------	-------------	0
t-quark mass dependence	-------------	-------------	-------------	-
taggability	tagga_sys	tagga_sys	taggability	1
Trigger efficency	TriggEff_sys	-------------	-------------	0
Trigger efficency	-------	trig_diem_sys	Trigger_EMMU	0
Trigger efficency	-------	trig_dimu_sys	Trigger_EMMU	0
Trigger efficency	-------	trig_emmu_sys	Trigger_EMMU	0
Vertex confirmation	vc_sys	vc_sys	vcj	1
Whf SF	Whf_2jb_sys	-------------	-------------	0
Whf SF	Whf_3jb_sys	-------------	-------------	0
Whf SF	Whf_4jb_sys	-------------	-------------	0
Wlp SF	Wlp_2jb_sys	-------------	-------------	0
Wlp SF	Wlp_3jb_sys	-------------	-------------	0
Wlp SF	Wlp_4jb_sys	-------------	-------------	0
Z & W pT reweighting	zwpt_sys	zwpt_sys	Z__p_T__reweighting	1
Z cross-section	z_xsec_sys	z_xsec_sys	Z_cross_section	1
z vertex reweighting	zvtxREW_sys	?	?	**1**