Machine Learning in High Energy Physics

SPMS 2018

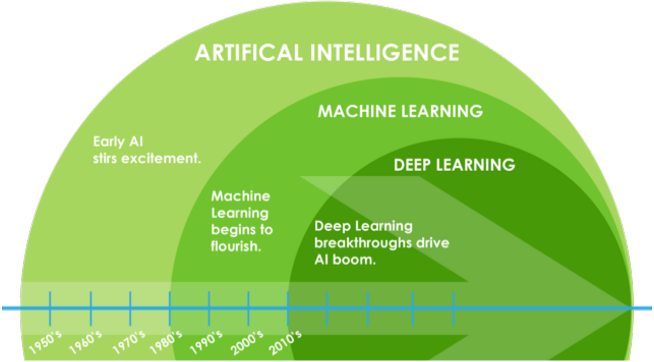
Jiří Franc

June 21, 2018

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

Outline

- Short summary of MVA and ML history in HEP (personal view).
- Current status of ML and MVA in HEP.
- MVA and ML example in one analysis, where we contributed.
- Challenges and future of ML in HEP.



Difference Between AI, ML and DL [source].

Short History

... Multivariate Pattern Recognition

1763 Bayes's theorem

1805 Least squares method

1936 Linear discriminant analysis (Fisher-Discriminants)

1950 Turing's machine

1951 First neural network

1967 Nearest Neighbor

1995 Support vector machines
Random forest

21stC Boom of ML in applications: IBM's Watson, Google Brains and AlphaGO, Chatbot passes the Turing Test, Facebook DeepFace, AI poker bot Libratus, etc.

Examples of application of MVA and ML in HEP experiments

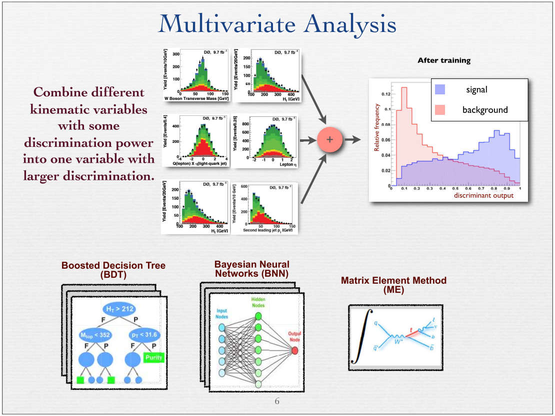
- Linear Decision Boundaries and Naive Bayesian classifiers in τ particle identification and studies:
  - MARK III at SLAC (1980s),
  - LEP collaborations ALEPH and OPAL (1990s).
- Artificial Neural Networks in jet identification and tracking at CDF and D0 (1992).
- Boosted Decision trees (BDT) - MiniBooNE, an experiment at Fermilab searching for neutrino oscillations (2005).
- TMVA - Toolkit for Multivariate Data Analysis (2007).
  The "era of hard cuts" was gradually ending.

TMVA Toolkit for Multivariate Data Analysis with ROOT [source].

Examples of application of MVA and ML in HEP experiments

- Combination of BDT, BNN and ME - Observation of Single Top-Quark Production (2009).
- 49 input variables -> 3 discriminants -> one final discriminant

This approach was reused in 2013 measurement again:

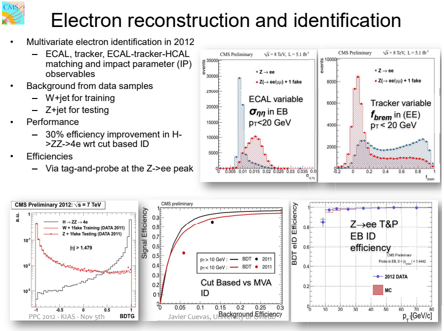


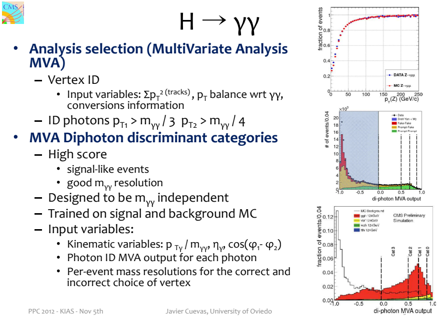
Friday, May 17, 2013

Jyoti Joshi [ADM meeting 17.5.2013]

Examples of application of MVA and ML in HEP experiments

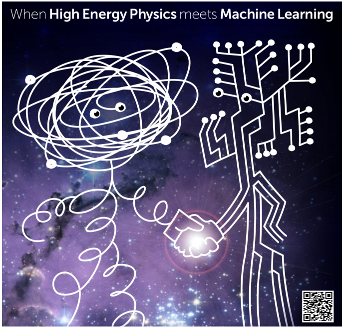
- Observation of Higgs Boson by **CMS** and ATLAS collaborations (2012).





Example of MVA usage by Cuevas in "CMS SM Higgs searches".

HEP meets ML: Kaggle HiggsML challenge



- More than 2000 teams
- Largest competition at the time.
- Winners presented solutions worldwide (DeepMind, OpenAI, ...).
- XGboost was the core of winning solution.
- GPU's were used in many solutions.

Was it pure success or HEP community gave it up to be state of the art in MVA and ML?

- Can we understand to MVA and ML and how can we code it in our systems.
- Before TMVA we used simple cuts, now we have tools which bring significant improvement, but is it enough?
- We can not use "black boxes", we need to understand what we are doing.
- etc.

Current status of MVA and ML in HEP

Huge change in ML and HEP communities in last 5 years.

- ML community found out that HEP problems are interesting and unique.
- HEP community discovered ML tools out of ROOT and TMVA.

To follow what ML community is doing in HEP read and see

- Machine Learning Community White Paper [pdf file]
- Inter-experimental LHC Machine Learning Working Group in CERN
  - The community increasingly uses non CERN standard libaries like keras, scikit-learn, XGBoost, either as standalone libraries or through interfaces with ROOT.
  - 1st ILM Machine Learning Workshop
  - 2nd ILM Machine Learning Workshop
- Fermilab Machine Learning Group

Current status of MVA and ML in HEP

	HEP (Circa 2013)	HEP (Circa 2018)	Industry
Data Layer	ROOT Files	ROOT Files	DB / HDFS etc.
Loading Layer	Ad hoc ROOT ETL logic	Numpy / HDF5 Converters / Loaders	Numpy / HDF5 Converters / Loaders
Training Layer	TMVA	Keras, TensorFlow, PyTorch, XGBoost, scikit-learn, ...	Keras, TensorFlow, PyTorch, XGBoost, scikit-learn, ...
Serving Layer	Deployment Target (TMVA)	Deployment Target (lwtnn, TensorFlow, TMVA wrappers)	Deployment Target (TensorFlow Serving, SageMaker, etc.)

source: Luke de Oliveira talk at IML2

Experiment Management: Make people training ML models more productive.
Universal Serving Layer: Make people using ML models more productive.

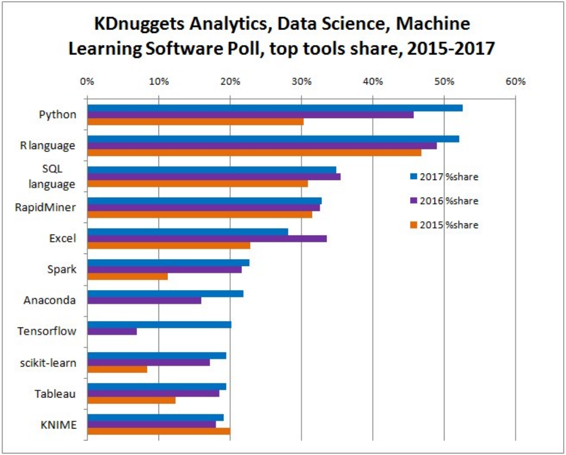
Current status of MVA and ML in HEP

	TMVA	TensorFlow	Theano	Scikit Learn	R	Spark ML	VW	libFM
ROOT [**T, C**]	✓							
CSV [**F**]		✓	✓	✓	✓	✓	×	×
libSVM [**M**]							×	✓
VW [**M**]							✓	
RGF [**M**]								
NumPy [**R**]		✓	✓	✓	✓	✓	×	×
Avro [**S, R**]					✓	✓		
Parquet [**S, C**]					✓	✓		
HDF5 [**S**]								
R df [**R**]					✓			

Where **T** Trees, **F** at tables, **M** sparse matrices, **R** row-wise arrays, **C** column-wise arrays, **S** static data structures. Source: ML white papers.

Python and R in HEP

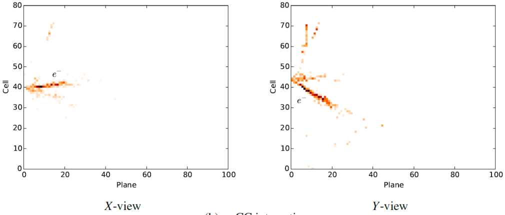
- Free and runs on any standard computing OS.
- Frequent releases, active development, very active user community.
- ROOT/TMVA Keras interface.
- TensorFlow is growing much faster and gains more support (Thanks Google :-).



Examples of application of MVA and ML in HEP experiments

Application of Deep Neural Networks (ten years later than Google)

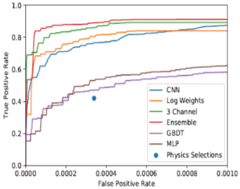
- Identification of neutrino interactions at NOvA by CNN (2016)

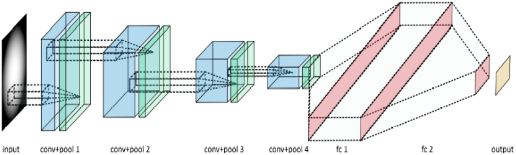


(b) ν_e CC interaction.

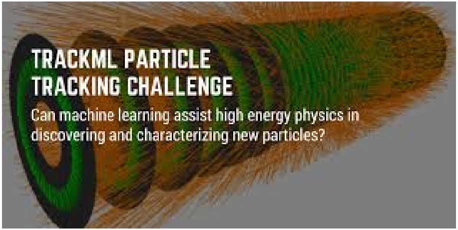
See talks of Petr Bour and Miroslav Kubu for more details.

- Identification of objects or particular particle types on ATLAS/CMS (2017)

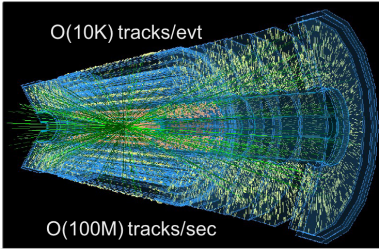




HEP meets ML again: Kaggle TrackML challenge



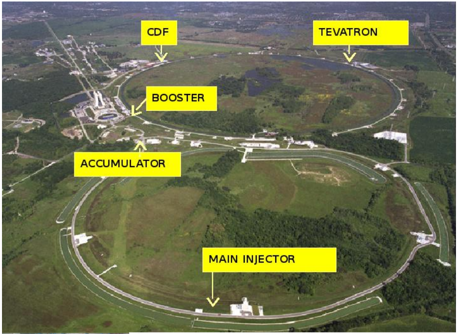
- $25,000 Prize Money
- Now 441 teams
- still 2 months to go until merger deadline

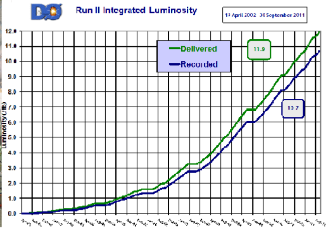


Challenge Datasets

- Accurate simulation engine to produce realistic dataset
- One file with list 3D points
- Typical events with 200 parasitic collisions (10.000 tracks/event)
- Large training sample 100k events, 10 billion tracks 100 GB.

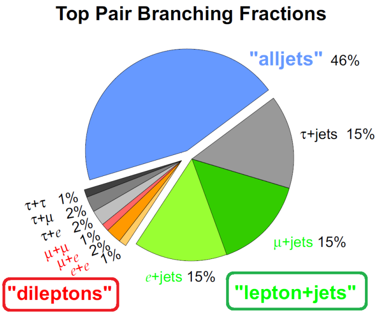
DØ experiment





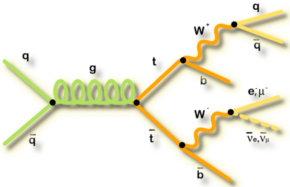
- **Tevatron** - $p\bar{p}$ **collider:**
  - circular accelerator (6.86 km).
  - $p\bar{p}$ collisions at 1.7 MHz.
  - most energetic collider until the turn on of the LHC.
- **Unique data set:** worlds largest $p\bar{p}$ data set for a long time.
- **Center Mass energy:** $\sqrt{s} = 1.96$ TeV.
- **Experiments:** Two multi-purpose detectors CDF and DØ with well understood detectors.
- **Run II:** begun in 2001 and each experiment recorded $\approx 10 fb^{-1}$ until September 2011.
- Presenting measurement has been done with the **full dataset 9.7fb^{-1}**.

Strong interaction: Top pair production

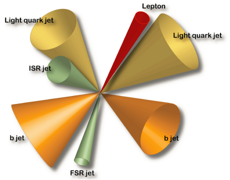


Top Quark at Tevatron:

- **Mass:** $m_t = 174.34 \pm 0.64\,GeV$
- **Lifetime:** $t \approx 5 \times 10^{-25} s << \Gamma_{QCD}$
- **Production:**
  $\approx 85\%$ by $q\bar{q}$ annihilation
  $\approx 15\%$ by gg fusion
- **Top decay:** $\mathrm{BR}(t \to W + b) \approx 100\%$

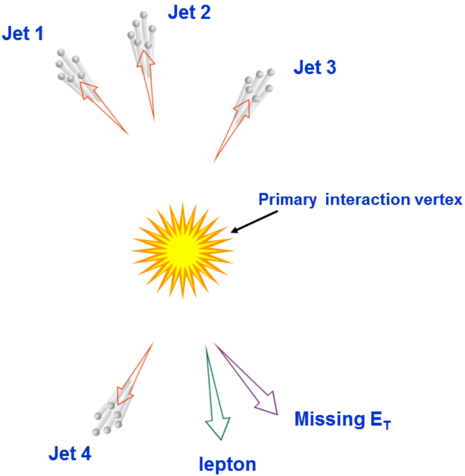


Situation in detector (+ missing transverse energy)

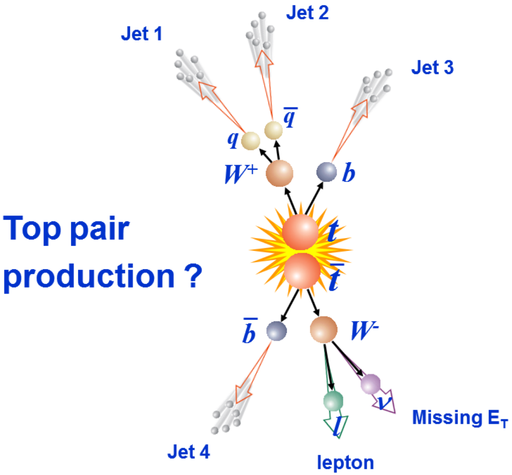


Samples are classified according to W-decay: $\ell+$jets and $\ell\ell$ channels are under concern and **full dataset 9.7fb^{-1}** is used.

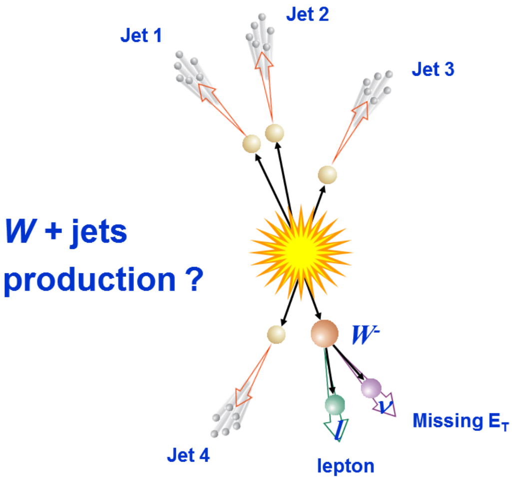
Primary interaction vertex in Top pair production



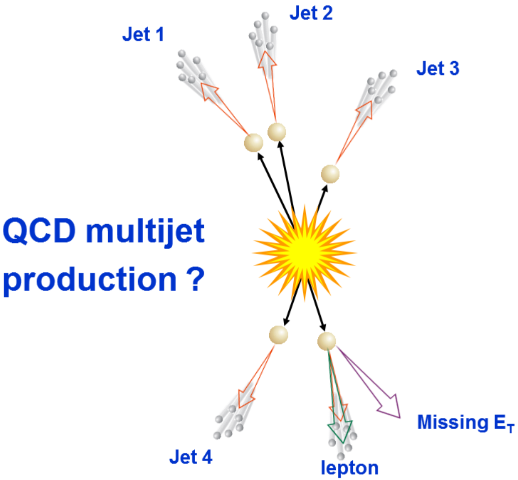
Primary interaction vertex in Top pair production



Primary interaction vertex in Top pair production

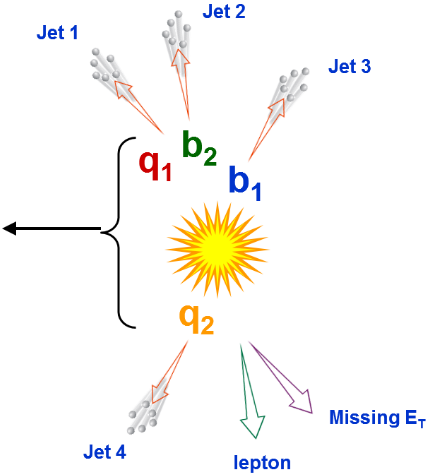


Primary interaction vertex in Top pair production



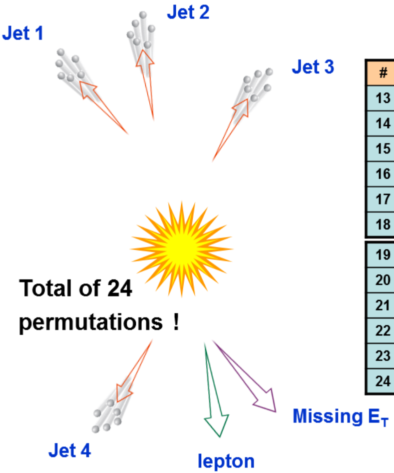
Primary interaction vertex in Top pair production

#	J1	J2	J3	J4
1	q_1	q_2	b_1	b_2
2	q_1	q_2	b_2	b_1
3	q_1	b_1	q_2	b_2
4	q_1	b_1	b_2	q_2
5	q_1	b_2	q_2	b_1
6	q_1	b_2	b_1	q_2



Primary interaction vertex in Top pair production

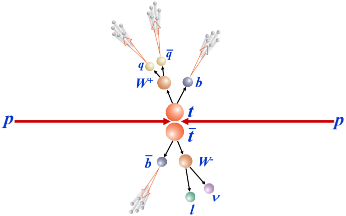
#	J1	J2	J3	J4
1	q_1	q_2	b_1	b_2
2	q_1	q_2	b_2	b_1
3	q_1	b_1	q_2	b_2
4	q_1	b_1	b_2	q_2
5	q_1	b_2	q_2	b_1
6	q_1	b_2	b_1	q_2
7	q_2	q_1	b_1	b_2
8	q_2	q_1	b_2	b_1
9	q_2	b_1	q_1	b_2
10	q_2	b_1	b_2	q_1
11	q_2	b_2	q_1	b_1
12	q_2	b_2	b_1	q_1

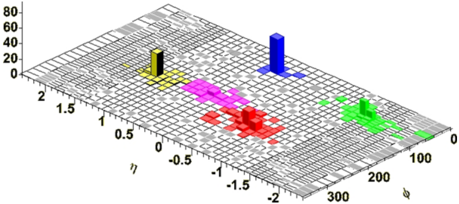


Total of 24 permutations !

#	J1	J2	J3	J4
13	b_1	q_1	q_2	b_2
14	b_1	q_1	b_2	q_2
15	b_1	q_2	q_1	b_2
16	b_1	q_2	b_2	q_1
17	b_1	b_2	q_1	q_2
18	b_1	b_2	q_2	q_1
19	b_2	q_1	q_2	b_1
20	b_2	q_1	b_1	q_2
21	b_2	q_2	q_1	b_1
22	b_2	q_2	b_1	q_1
23	b_2	b_1	q_1	q_2
24	b_2	b_1	q_2	q_1

Selection of $t\bar{t}$ Candidates





Main selection cuts in $\ell +$ jets channel:

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

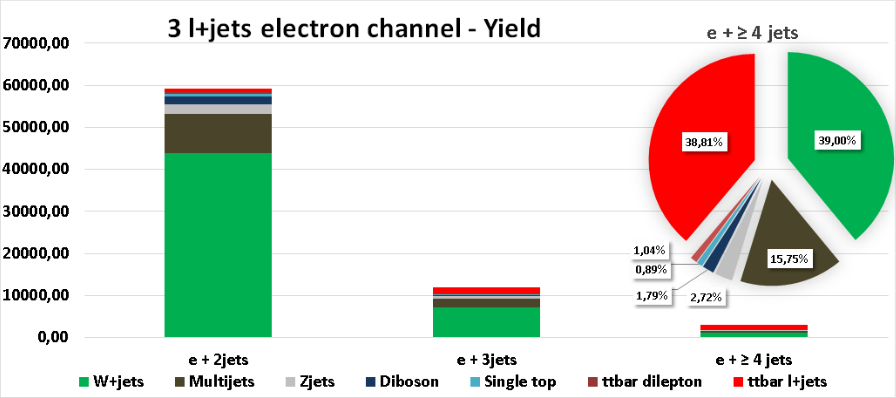
+ additional cuts

The measurements in both decay channels employ the b-tagging discriminant output distribution as provided by the b-ID MVA.

Data sample: Full Data Set ($9.7 fb^{-1}$) with selection: Phys.Rev.D 90,092006 (2014)
The main goal: Measurement of the inclusive $t\bar{t}$ cross section using MVA and b-ID methods in $\ell +$ jets and $\ell\ell$ channels and compute pole mass.

$\ell +$jets Yield table:

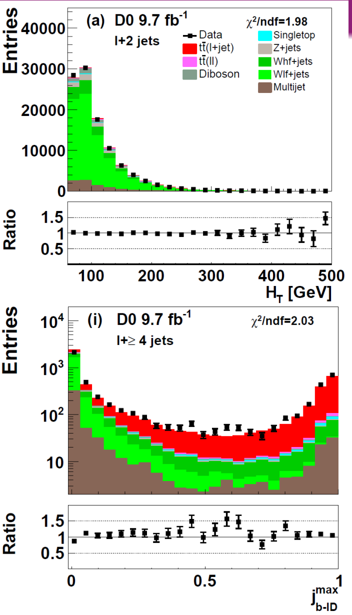
6 analysis channels in $\ell +$jets :
to the lepton type (electron, muon) and the number of jets (2, 3, $\geq$ 4)



Signal rate in MC(Signal from Alpgen+Pythia):

	2 jets	3 jets	$\geq$ 4 jets
e	1.17%	12.24%	38.81%
μ	0.88%	11.01%	39.01%

Does the MC describe the data?

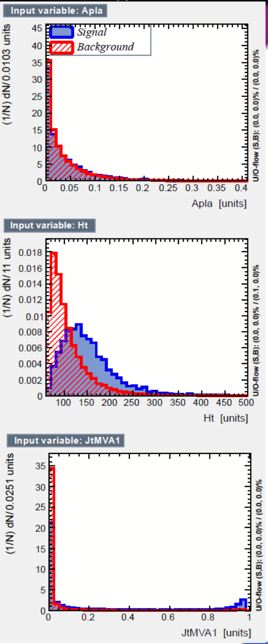


- **Task:** Check the MC and data agreement for all used kinematical and topological variables.
- **Tools:** Control plots and statistical hypothesis testing.
- **Weighted homogeneity tests:** More information in J. Trusina and A. Novotny talks - stay tuned!

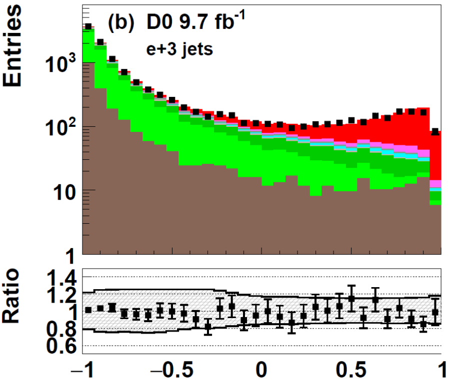
The data are compared to the sum of predicted contributions from signal and background processes, using the theoretical value of $\sigma_{t\bar{t}} = 7.48$ pb and $m_t = 172.5$ GeV.

$\ell +$ jets variables selection

- **Analyzed:** 46 kinematical and topological variables (e.g. H_T, Aplanarity, Sphericity, $M_T(jets)$, lepton p_T, ...) + $j^{\text{lead}}_{\text{b-ID mva}}$.
- **Task:** Select variables with good MC vs. Data agreement and good separation power between signal and background.
- **Tools:** Statistical hypothesis testing (KS test, AD test, χ^2 test, LR test, . . .) and TMVA ranking.
- **Selection:** 25 types of good modeled variables with best separation power.
- *b*-**tagging:** MVA *b*-ID output distribution $j^{max}_{b-\text{ID}}$ has been included

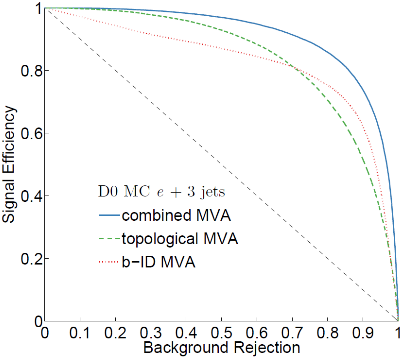


$\ell +$ jets discrimination by TMVA

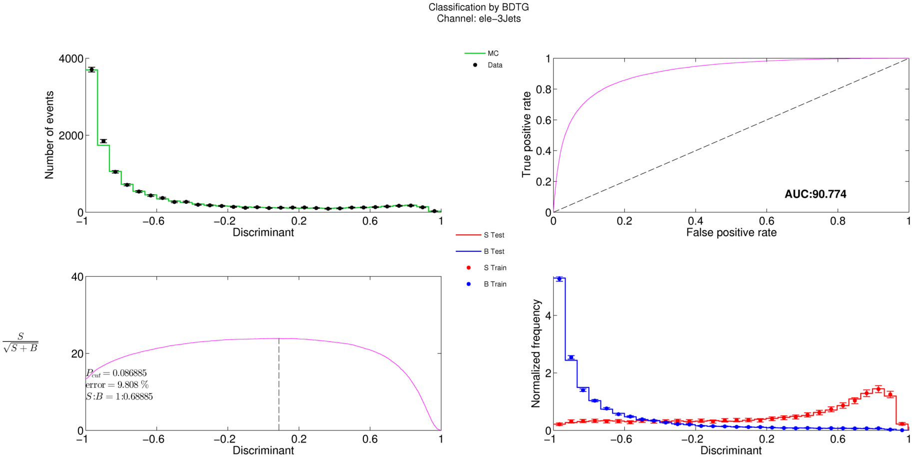


- The combination of pure topological and MVA b-ID method improved the separation by 10%.
- The area under the ROC curve is around 0.9 for all 6 analysis channels when using TMVA discr.

- Different MVA methods has been tried.
- 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.



Discrimination by BDTG - optimal cutting:



Signal rate in MC - Improvement:

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

Summary and The End of The Talk

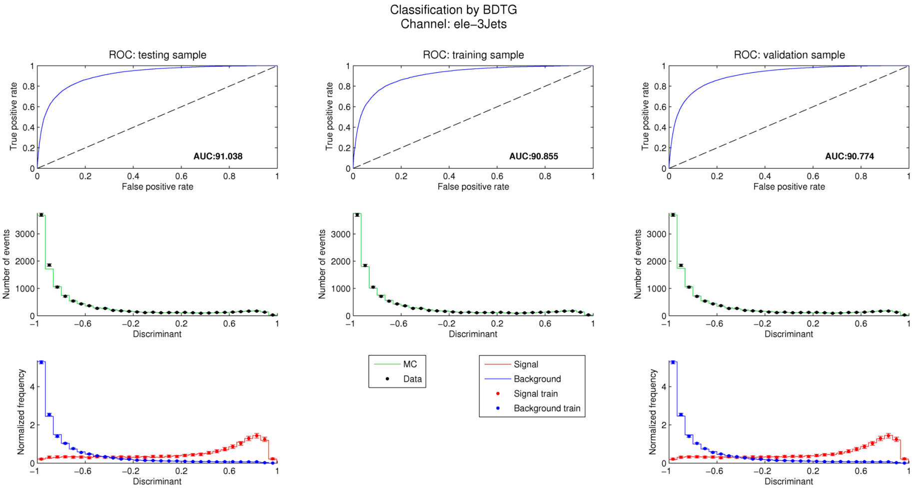
This is the end => Thank you for your attention.

Backup

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 $\not{E}_T$.

Discrimination by BDTG - overtraing check:



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

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).

Yiled table: l+jets

TABLE I. Expected number of events in the ℓ+jets channel with 2, 3 or ≥ 4 jets. The sum of signal and background agrees well with the number of data events by construction; uncertainties are statistical and systematic added in quadrature (see Sec. IX A 4 for details). Events from $t\bar{t}$ dilepton decays are treated as background and denoted as "$t\bar{t}$ $\ell\ell$".

	ℓ+jets decay channel					
Process	$e + 2$ jets	$e + 3$ jets	$e+ \geq 4$ jets	$\mu + 2$ jets	$\mu + 3$ jets	$\mu+ \geq 4$ jets
Multijet	9160 ± 2350	2266 ± 550	464 ± 120	1546 ± 630	418 ± 170	99 ± 40
Single top	471 ± 60	129 ± 20	27 ± 5	331 ± 40	92 ± 10	20 ± 3
Wlp + jets	$37937 \pm^{1350}_{700}$	$5544 \pm^{200}_{100}$	$850 \pm^{30}_{20}$	$32701 \pm^{1150}_{600}$	$5313 \pm^{200}_{100}$	$835 \pm^{30}_{15}$
$(Wc\bar{c} + Wb\bar{b})$+jets	$6020 \pm^{1000}_{1400}$	$1502 \pm^{250}_{350}$	$329 \pm^{60}_{80}$	$4998 \pm^{850}_{1150}$	$1391 \pm^{250}_{300}$	$315 \pm^{50}_{70}$
$Z/\gamma^* lp$+jets	2031 ± 400	390 ± 80	57 ± 10	2557 ± 500	422 ± 80	49 ± 10
$(Z/\gamma^* c\bar{c} + Z/\gamma^* b\bar{b})$+jets	369 ± 70	114 ± 20	24 ± 5	485 ± 100	120 ± 20	21 ± 5
Diboson	1926 ± 140	338 ± 20	52 ± 5	1417 ± 100	249 ± 20	40 ± 5
$t\bar{t}$, $\ell\ell$	566 ± 30	182 ± 10	31 ± 5	345 ± 20	118 ± 10	22 ± 5
$\sum$ bknd	58479 ± 2900	10465 ± 650	1834 ± 140	44381 ± 1650	8123 ± 350	1402 ± 80
$t\bar{t}$, ℓ+jets	669 ± 30	1460 ± 70	1177 ± 60	393 ± 20	1002 ± 50	909 ± 50
$\sum$ (sig + bknd)	59148 ± 2900	11925 ± 650	3011 ± 140	44773 ± 1650	9125 ± 350	2310 ± 80
Data	59122	11905	3007	44736	9098	2325

Yiled table: dilepton

TABLE II. Expected number of events in the $ee+\geq 2$ jets, $\mu\mu+\geq 2$ jets, $e\mu+1$ jets and $e\mu+\geq 2$ jets channels due to each process; uncertainties are statistical and systematic added in quadrature (see Sec. IX A 4 for details).

Process	dilepton decay channel $ee+\geq 2$ jets	$\mu\mu+\geq 2$ jets	$e\mu+1$ jets	$e\mu+\geq 2$ jets
Multijet	$5.7\pm{}^{0.9}_{0.9}$	$7.0\pm{}^{3.3}_{2.6}$	$28.3\pm{}^{6.6}_{6.6}$	$32.5\pm{}^{7.4}_{7.4}$
$Z/\gamma^*\rightarrow\ell\ell$+jets	$66.6\pm{}^{17.9}_{17.2}$	$107.6\pm{}^{22.1}_{22.0}$	$74.6\pm{}^{15.8}_{15.8}$	$57.5\pm{}^{13.8}_{13.4}$
Diboson	$9.9\pm{}^{2.4}_{2.2}$	$12.6\pm{}^{2.8}_{3.0}$	$38.5\pm{}^{4.6}_{4.2}$	$14.7\pm{}^{3.7}_{3.5}$
$\sum$ bknd	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 + bknd)	190 ± 23	229 ± 25	228 ± 21	418 ± 42
Data	215	242	236	465

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**

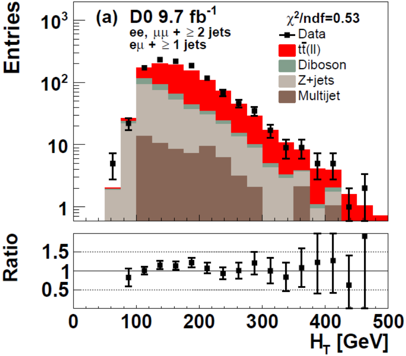
Xsec calculation in l+jets - Sources of systematic uncertainties

- Each source of systematic uncertainty yields a modified discriminant distribution of the MVA topological (or MVA b-ID) method.
- 50 different systematic uncertainties has been taken into account (list in backup).
- The pre-fit uncertainty in percent from each source on the inclusive cross section is given for the $\ell +$jets and $\ell\ell$ channel.

Source of uncertainty	$\delta_{\ell+\text{jets}}$, %	S/N	$\delta_{\ell\ell}$, %	S / N
Modeling of signal				
Alternative signal model	±10	S	±4	S
Hadronization	±8	S	±4	S
Color reconnection	±2	S	±2	S
ISR/FSR variation	±2	S	±2	S
PDF	±7	N	±1	N
Modeling of detector				
Jet modeling & ID	±8	S	±3	S
b-jet modeling & ID	±5	S	±12	S
Lepton modeling & ID	±3.5	S	±6	N
Trigger efficiency	±5	N	±2	N
Luminosity	±4.7	N	±4.3	N
Sample Composition				
MC cross sections & BRs	±0.9	N	±1.3	N
Z/W p_T reweighting	±1.5	S	±4	S
Multijet contribution	±23	S/N	±15	S/N
Z/γ^*+jets scale factor	±25	S/N	±2	S/N
W+jets heavy flavor SF	$^{+17}_{-23}$	S/N	n.a.	n.a.
W+jets light parton SF	$^{+3.5}_{1.8}$	S/N	n.a.	n.a.
MC statistics	n.a.	n.a.	±3	S/N

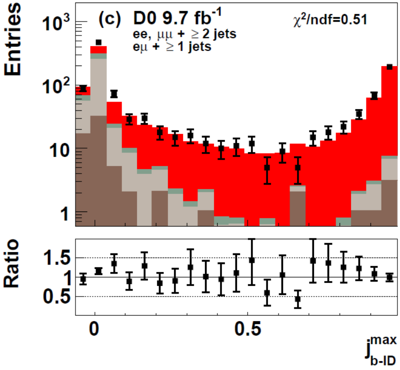
The $\sigma_{t\bar{t}}$ measurement and nuisance parameter fit of MC to Data are performed using the software package **Collie** (A Confidence Level Limit Evaluator, DØ note 5595).

$\ell\ell$ discrimination by b-ID method



- The shape of the $j^{\text{lead}}_{\text{b-ID mva}}$ variable allows to distinguish between $t\bar{t}$ events located at high output values and the most dominant $Z/\gamma^{\star}+$jets background located at low output values.

- $j^{\text{lead}}_{\text{b-ID mva}}$ - the MVA value of the jet with highest MVA output value is used for discrimination.



Xsec calculation in l+jets - Sources of systematic uncertainties

Each source of systematic uncertainty yields a modified discriminant distribution of the MVA topological (or MVA b-ID) method.

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

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

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

Source of uncertainty	$\delta_{\rm comb,}$, pb
Signal modeling	
Signal generator	±0.17
Hadronization	±0.25
Color reconnection	±0.09
ISR/FSR variation	±0.06
PDF	±0.02
Detector modeling	
Jet modeling & ID	±0.04
b-jet modeling & ID	±0.23
Lepton modeling & ID	±0.17
Trigger efficiency	±0.16
Luminosity	±0.27
Sample Composition	
MC cross sections	±0.09
Multijet contribution	±0.10
W+jets scale factor	±0.15
Z/γ^*+jets scale factor	±0.12
MC statistics	±0.02
Total systematic uncertainty (quadratic sum)	±0.60
Total systematic uncertainty (central COLLIE)	±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, DØ note 5595).

Inclusive Xsec Determination and Results

The results of the measurement:

in $\ell+$jets decay channel when using the topological method is

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

with a relative total uncertainty of 9.2%.

in $\ell\ell$ decay channel using the MVA *b*-jet method is

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

with a relative total uncertainty of 9.6%.

In the combination of $\ell+$jets and $\ell\ell$ channels using the topological and MVA *b*-jet method:

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

with a relative total uncertainty of 7.6%.

Theory prediction of σ_{tt} uncertainty is $\approx 3.5\%$ for Tevatron $1.96\,TeV$ (Czakon, et al.).

Xsec calculation for different mass

- Extraction of the top quark mass by using the inclusive $t\bar{t}$ cross section as a function of the top quark MC mass.
- A cubic fit to the individual cross section measurements at various mass points has been provided for the measured dependency.
- Comparison of the top quark mass dependence of the inclusive $t\bar{t}$ cross section with the expected dependency in NNLO pQCD calculation top++.

Top quark mass [GeV]	Cross section $\sigma(t\bar{t})$ [pb]
150	$9.70 \pm 0.16\,(\text{stat.})^{+0.73}_{-0.67}\,(\text{syst.})$
160	$8.25 \pm 0.14\,(\text{stat.})^{+0.63}_{-0.57}\,(\text{syst.})$
165	$7.46 \pm 0.13\,(\text{stat.})^{+0.58}_{-0.51}\,(\text{syst.})$
170	$7.55 \pm 0.13\,(\text{stat.})^{+0.58}_{-0.55}\,(\text{syst.})$
172.5	$7.26 \pm 0.12\,(\text{stat.})^{+0.57}_{-0.50}\,(\text{syst.})$
175	$7.28 \pm 0.12\,(\text{stat.})^{+0.54}_{-0.49}\,(\text{syst.})$
180	$6.86 \pm 0.12\,(\text{stat.})^{+0.53}_{-0.47}\,(\text{syst.})$
185	$6.50 \pm 0.11\,(\text{stat.})^{+0.50}_{-0.43}\,(\text{syst.})$
190	$6.70 \pm 0.11\,(\text{stat.})^{+0.60}_{-0.47}\,(\text{syst.})$

Pole mass of the Top Quark:

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

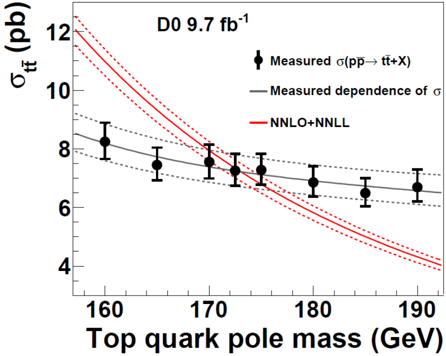
- The precision is 1.9% and represents the most precise determination of the top quark pole mass from the inclusive $t\bar{t}$ cross section at the Tevatron.

Xsec calculation for different mass

- Extraction of the top quark mass by using the inclusive $t\bar{t}$ cross section as a function of the top quark MC mass.
- A cubic fit to the individual cross section measurements at various mass points has been provided for the measured dependency.
- Comparison of the top quark mass dependence of the inclusive $t\bar{t}$ cross section with the expected dependency in 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.