Measurement of the inclusive $t\bar{t}$ production cross section in $p\bar{p}$ collisions at D0 and extraction of the top quark mass

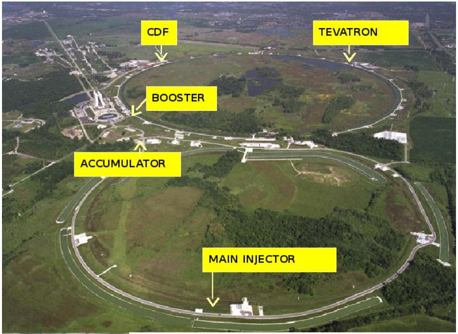
Les Rencontres de Physique de la Vallée d'Aoste 2015

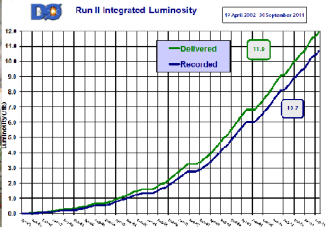
Jiří Franc
and
DØ Collaboration

March 05, 2015

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

DØ experiment



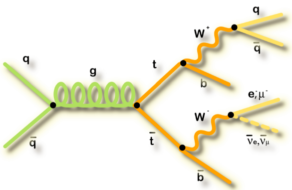


Tevatron - $p\bar{p}$ collider

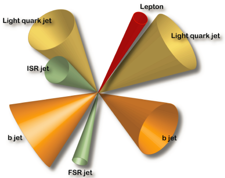
- **Tevatron:** $p\bar{p}$ circular particle accelerator (6.86 km).
- **Unique data set:** worlds largest $p\bar{p}$ data set for a long time.
- **Center Mass energy:** $\sqrt{s} = 1.96$ TeV.
- **Experiments:** CDF and DØ with well understood detectors.
- **Run II:** begun in 2001 and each experiment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



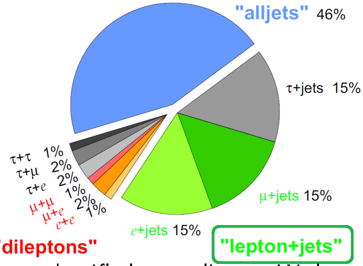
Feynman diagram of the $t\bar{t}$ production.



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

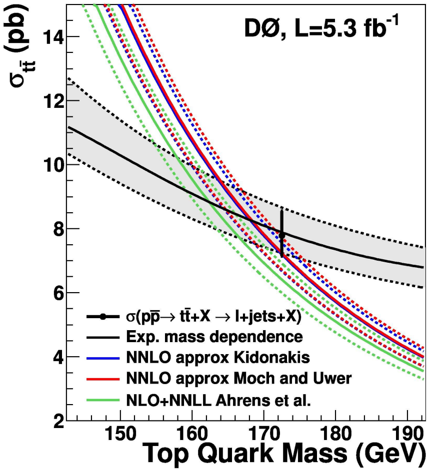


Samples are classified according to W-decay: $l + jets$ channel is under concern in this analysis.

Top pair production cross section measurement in lepton+jets channel

- **6 analysis channels:** Electron: 2 Jets, 3 Jets, 4+ Jets.
  Muon: 2 Jets, 3 Jets, 4+ Jets.
- **Data sample:** Full Data Set ($9.7 fb^{-1}$) with selection: Phys.Rev.D 90,092006 (2014)

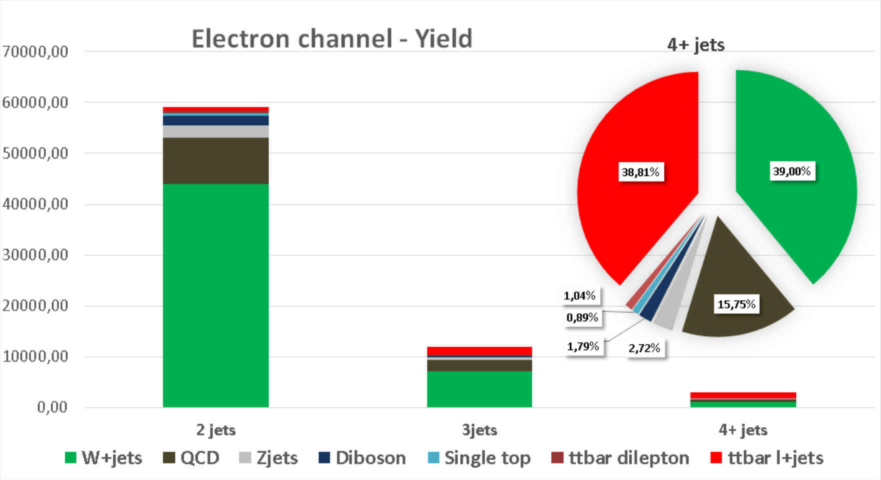
variable	kinematic range
lepton $\eta(e)$	$\|\eta(e)\| < 1.1$
lepton $\eta(\mu)$	$\|\eta(\mu)\| < 2.0$
lepton $p_T(l)$	$p_T(l) > 20$ GeV
$\not{E}_T$	$\not{E}_T > 20$ GeV



The Goal:

- The inclusive $t\bar{t}$ cross section measurement using MVA methods in l+jets channel.
- Improve the precision compare to previous analyses by reducing systematical uncertainties.
- Extract the mass dependence of the $t\bar{t}$ cross section by applying the same trained MVA to $t\bar{t}$ samples generated at varous top mass values.

Yield table:



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

	2jb	3jb	4+jb
Electron	1.17%	12.24%	38.81%
Muon	0.88%	11.01%	39.01%

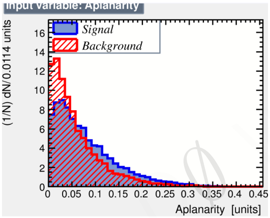
Variables selection - Visual control plots

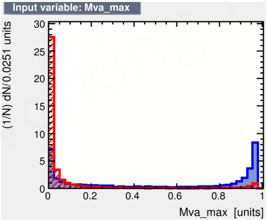
- 46 kinematical and topological variables were analyzed (e.g.Aplanarity, Sphericity, H_T^l, $M_T(jets)$, lepton p_T, ...) + mva_max variable.

MVA b-ID output distribution has been included

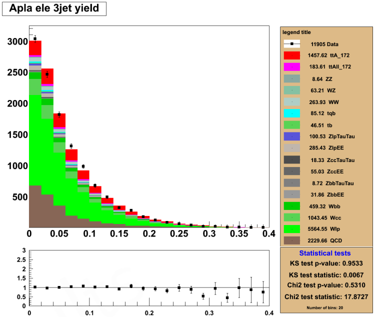
$$\text{mva\_max} = \max_{i \in \forall jets} (\text{bID MVA output of } i\text{th jet}).$$

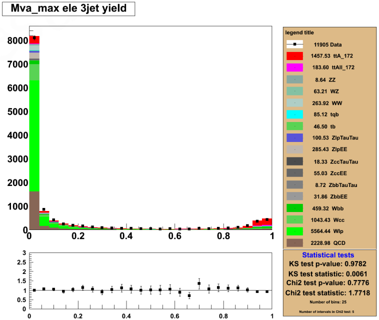
- Task: select variables with good MC vs. Data agreement and good separation power between signal and background.



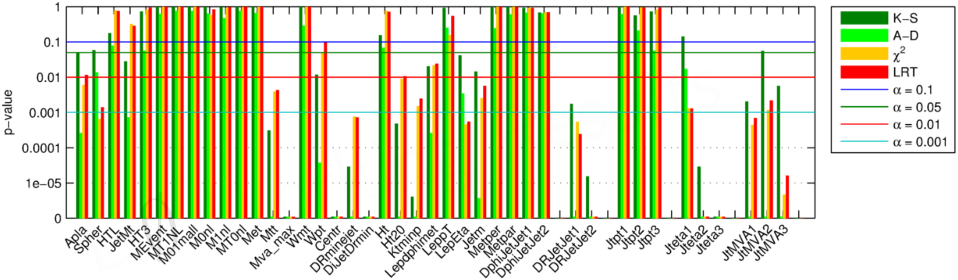


Example for Electron 3jb channel:





Variables selection - Statistical approach



ROOT hypothesis testing

- Kolmogorov-Smirnov test (TH1),
- Goodness-of-fit tests: Chi2Test (TH1).

- equidistant bins and can't use weighted EDF.

TMVA ranking

- variables ranking
- TMVA method ranking

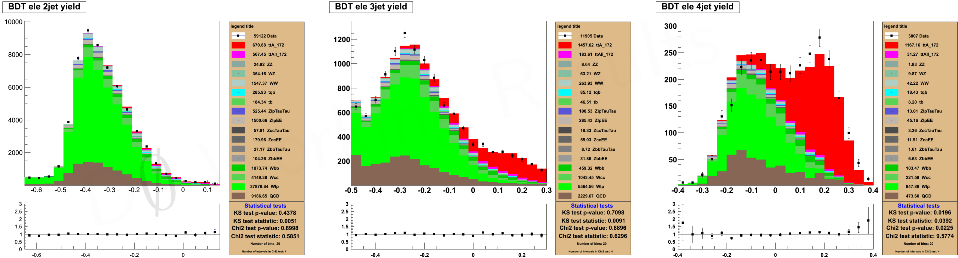
Correct statistical hypothesis testing

- Kolmogorov-Smirnov test.
- Anderson - Darling test.
- χ^2 Goodness of Fit test.
- Likelihood ratio test.
- Φ divergences comparison(L1 norm, Rényi divergence, ...)

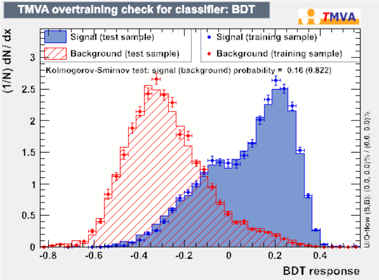
- used weighted EDF, quantile bins.

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

Discrimination by TMVA

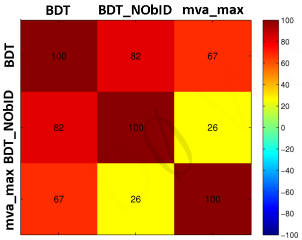


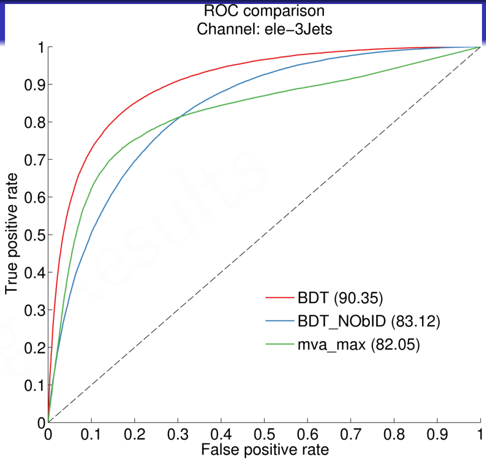
- MVA methods combine different variables with small discrimination power into one final variable with larger discrimination.
- TMVA BDT with adaptive boost has been trained on 30 variables + mva_max variable
- Attempts without mva_max variable and with another MVA methods has been tried.



Discrimination - Method comparison

- The area under the ROC curve is around 0.9 for all 6 channels when using TMVA discrimination.
- The improvement of the separation in the combination of both methods is about 10%.
- MVA b-ID variable has behave better in "strict threshold area", while pure topological BDT behave better in '"lax treshold area"





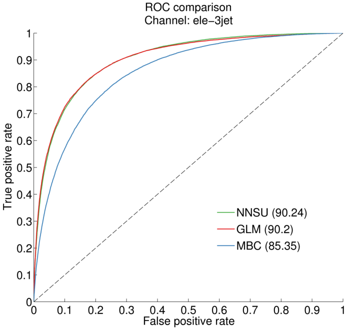
- Discriminants are not perfectly linear correlated.
- Receiver operating characteristic (ROC) curve describes the performance of a binary classifier plotting the *Sensitivity* against *1-Specificity* at various thresholds.

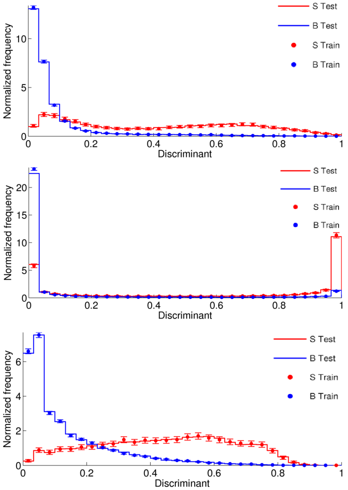
Discrimination by non TMVA methods

Three separation technique out of the ROOT has been tried and compared with TMVA methods.

- Generalized Linear Models (GLM);
- Model based clustering (MBC);
- Neural networks with switching units (NNSU);

They have different shapes of probability density function, but final Xsec measurement gives consistent results.





Output discriminants are not perfectly correlated and they can be used as new input variables for the eventual second step separation.

Inclusive Xsec calculation - precisious measurement

More than 30 different flat and shape dependent systematic uncertainty has been included. Error is dominated by systematic uncertainties and the most influenced are:

- Hadronization uncertainties (Alpgen+Herwig vs. Alpgen+Pythia).
- Higher orders signal model (MC@NLO+Herwig vs. Alpgen+Herwig).
- ISR-FSR (initial state radiation vs. final state radiation).
- The uncertainty on the Xsec due to the uncertainty on PDF.

Inclusive Xsec calculation - precisious measurement

More than 30 different flat and shape dependent systematic uncertainty has been included. Error is dominated by systematic uncertainties and the most influenced are:

- Hadronization uncertainties (Alpgen+Herwig vs. Alpgen+Pythia).
- Higher orders signal model (MC@NLO+Herwig vs. Alpgen+Herwig).
- ISR-FSR (initial state radiation vs. final state radiation).
- The uncertainty on the Xsec due to the uncertainty on PDF.

Comparison of previous D0 results ($5.3 fb^{-1}$) and new expected results ($9.7 fb^{-1}$) obtained by BDT combining MVA b-ID variables and other topological variables.

	lepton+jets	dilepton	Combination
$5.3 fb^{-1}$	$\approx 9.1\%$	$\approx 11.5\%$	$\approx 8\%$
$9.7 fb^{-1}$	$\approx 7\%$	$\approx 10\%$	$\approx 6\%$

$$\sigma_{tt} = \frac{N^{signal}}{\epsilon \cdot \mathcal{L} \cdot B}$$

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

The $t\bar{t}$ Xsec measurement and nuisance parameter fit of MC to Data performed with Collie (A Confidence Level Limit Evaluator, DØ note 5595).

Summary and The End of The Talk

- The combination of $\ell + jets$ and $\ell\ell$ channels using the MVA method for the measurement of the inclusive cross section yields for a top quark mass of 172.5 GeV has been done.
- Final numbers and the measured mass dependency of the top quark pair production cross section will be ready soon.

Main result

- 25% improvement in comparison with a previous DØ result ($5.3 fb^{-1}$).
- The expected precision of the the inclusive $t\bar{t}$ production cross section in $p\bar{p}$ collisions is $\approx$ 6% !!!

Note: The current Tevatron cross section combination has an uncertainty of 5.4%.

Summary and The End of The Talk

- The combination of $\ell + jets$ and $\ell\ell$ channels using the MVA method for the measurement of the inclusive cross section yields for a top quark mass of 172.5 GeV has been done.
- Final numbers and the measured mass dependency of the top quark pair production cross section will be ready soon.

Main result

- 25% improvement in comparison with a previous DØ result ($5.3 fb^{-1}$).
- The expected precision of the the inclusive $t\bar{t}$ production cross section in $p\bar{p}$ collisions is $\approx$ 6% !!!

Note: The current Tevatron cross section combination has an uncertainty of 5.4%.

This is the end => Thank you for your attention

and special acknowledgments to Organizers and Regional Government of valle d'Aosta

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_{mva}^{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 $\not{E}_T$.

GLM: Generalized Linear Models

Applied on former selection and only 20 variables has been used

Response has binomial distribution from exponential family and measure the relationship between a categorical (binary) dependent variable and given independent variables.

Model:

$$E(Y_i) = \mu = g^{-1}(\mathbf{X}\beta),$$

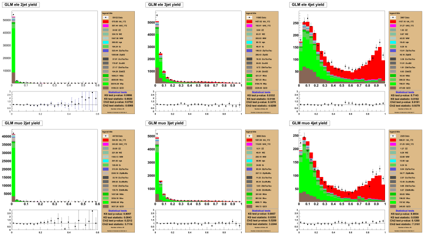
where random variable Y_i are *iid* and g is given link function.

The aim is to estimate parameter β from training sample and than compute the probability that given event from yield sample is signal.

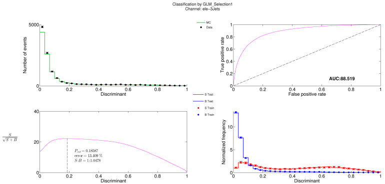
$$\text{Logistic link function:} \quad g(\mu) = \ln\left(\frac{\mu}{1-\mu}\right)$$

$$\text{Loglog ling function:} \quad g(\mu) = \ln(-\ln(1-\mu))$$

$$\text{Probit ling function:} \quad g(\mu) = \mathbf{\Phi}^{-1}(\mu)$$



Control plots of GLM discriminants with **quasibinomial family** (dispersion parameter is not fixed at one) and **probit link**.



MBC: Model-Based Clustering

MBC method try to describe both signal and background by different distribution mixtures.

Distribution mixture:

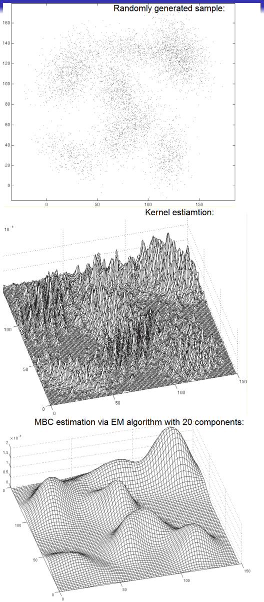
$$p(x|\Theta) = \sum_{j\in M} \alpha_j p_j(x|\theta_j), \quad \sum_{j\in M} \alpha_j = 1, \quad \alpha_j \geq 0,$$

α_j is weight of the jth component,
$p_j(x|\theta_j)$ is Gaussian density function of the jth component.

Parameters of GMM is computed by iterative EM algorithm:

- monotone convergence guaranteed (local maximum)
- strong dependence on initialization

Final evaluation is done by the help of the Bayes formula.

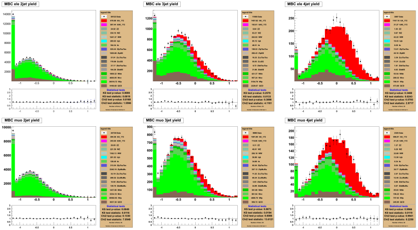


MBC: Model-Based Clustering

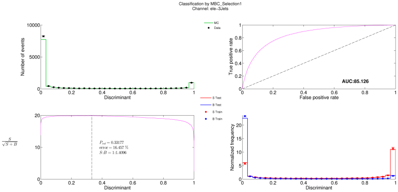
Applied on former selection and only 20 variables has been used

Tasks:

- Find optimal number of components – the risk of overfitting, degeneration of the model.
- Stability: diagonal/full covariance matrix.
- Event Transformation:
  - Gaussianisation (Via inverse error function, Box-Cox transformation).
  - Decorrelation.
- Best fit on elliptical clusters.
- Better results for less variables with unimodal distribution (appreciated for combination).
- Results improved by genetic optimization.
- More information in:
  - Phys. Conf. Ser. 490 012225 (2014)
  - Phys. Conf. Ser. 574 012150 (2015)



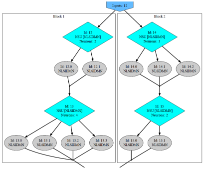
Control plots of MBC discriminants (log transformation applied) with different number of componnents for signal and background (depend on channel).

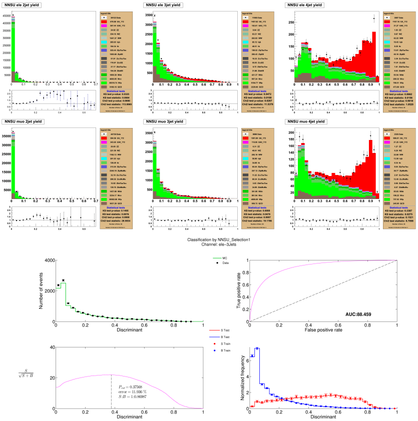


NNSU: Neural nets with switching units

Applied on former selection and only 20 variables has been used

- Switching units with predefined number of clusters.
- Clusters of data are propagated into neural units: logit, probit, quadratic regression.
- Switching and neural units form blocks.
- NNSU is acyclic graph of blocks.
- Genetic algorithm IS used to find optimal net:
  - NNSU connections graph is quickly improved;
  - Representation corresponds in an acceptable way to a directed acyclic graph;
  - Ex. of settings: population of 200 neural networks, 150 generations.
- Fitness function can be defined as: MSE, AUC of R



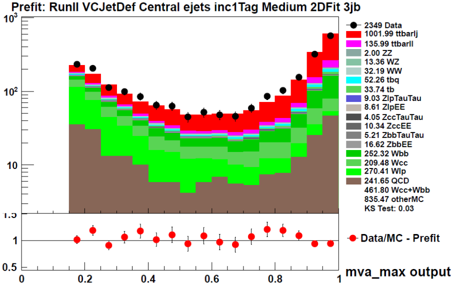


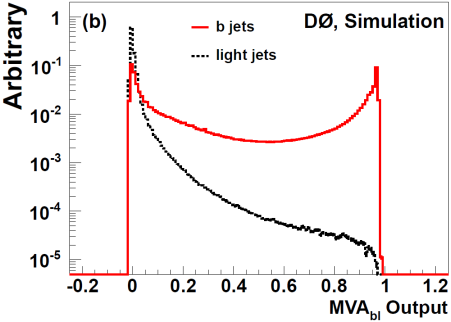
Control plots of NNSU discriminants.

More information about NNSU on http://www.cs.cas.cz/hakl/nnsu-server

Measurement using the b-ID mva technique

MVA output – probability discriminant of b-jet identification technique. b-ID Medium working point is applied - cut at 0.15 in selection.





- *b*-tagging efficiency is approximately 60%, with a light quark misidentification rate of approximately 1.2%.
- Full data and MC templates are splitted into sixteen separate channels with respect to lepton type, amount of jets, and *b*-tags

Full list of systematic uncertainties

Signal systematic uncertainties
Alternative signal model
Color reconnection
Parton showering
ISR/FSR variation

Flat systematic uncertainties
Jet response correction
t-quark mass dependence
Trigger efficiency
Luminosity
Luminosity reweighting
W+jets heavy flavor scale factor
W+jets light parton scale factor
Fake and True lepton rate
Data Quality
MC background cross section
MC signal and background branching ratio
MC statistics

Shape dependent systematic uncertainties
Parton distribution function
Lepton identification
Muon ID, track and isolation
Jet energy scale
Jet energy resolution
Jet identification
Lepton momentum
dZ (lepton, PV) uncertinaty
Z/W pT reweighting
z vertex reweighting and PV scale factor
b-fragmentation
b-tagging uncertainty
c-tag uncertainty
l-tag uncertainty
Taggability uncertainty
Vertex confirmation