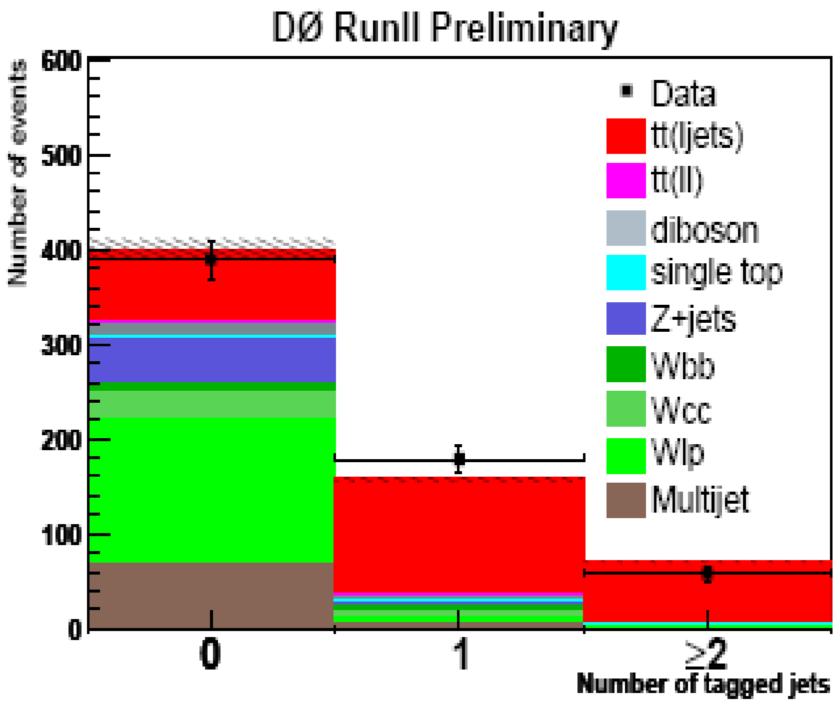
b/c/tau Tagging at the Tevatron - State of the Art

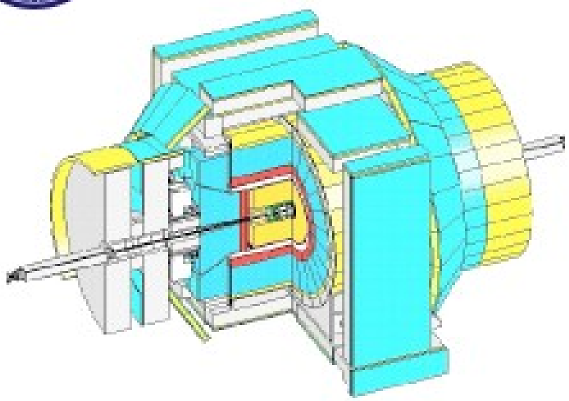
Andy Haas*
Columbia U. / Nevis Labs

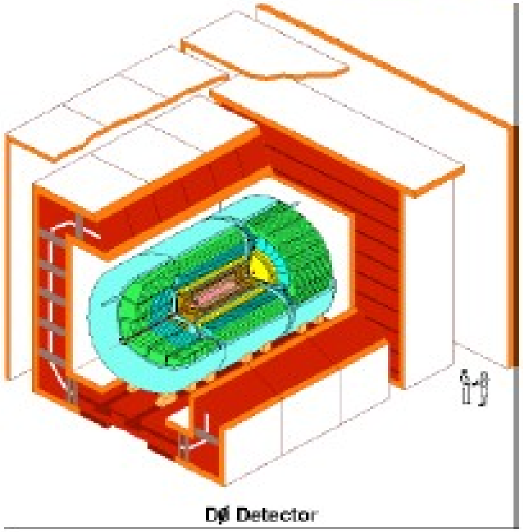
ALCPG 2007
Fermilab
October 24, 2007



**Lots of slides stolen from:* Lorenzo Feligioni (CPPM)

Introduction





B-Tagging is critical to physics at the Tevatron

Top (t->Wb, *single-top*)

SUSY (stop, sbottom, ...)

Higgs (h->bb)

We've learned a lot...

Smart (multivariate) b/c reconstruction

Accurate measurement of efficiency / fake rate

Optimal *use* of info in analyses

Disclaimers and Qualifications

I'm on D0

- Biased towards our methods
- But CDF is similar, and we do share techniques

b-ID convener, for past 2 years

Now Higgs convener

- Single-top has slightly different needs...

I like searches!

- Other issues may dominate for precision measurements (top cross-section, etc.)

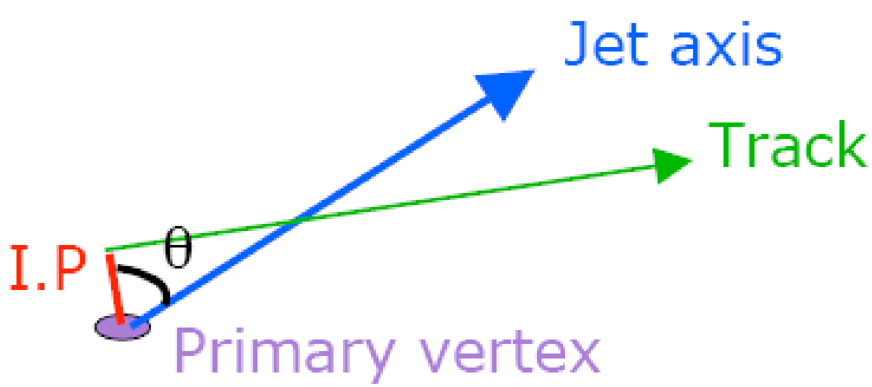
Tevatron (p-pbar) is an extremely difficult environment

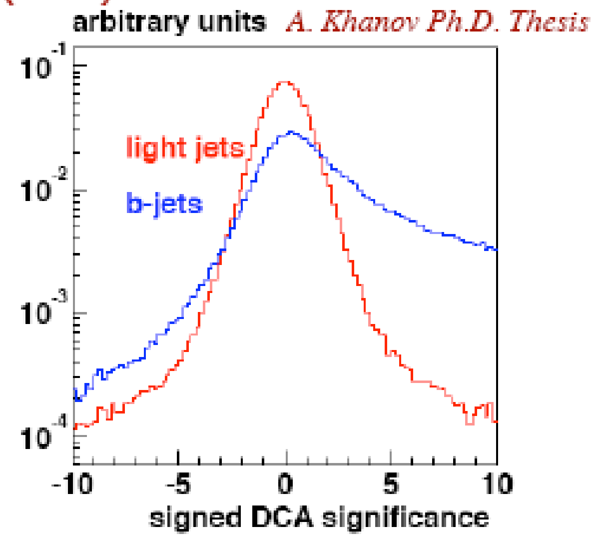
- QCD background!
  sigma(Z->bb)/sigma(bb)=~0.05 (under peak)!
- Personpower-limited
  ~150 FTE / experiment!

b/c Tagging Techniques

Tagging Algorithms

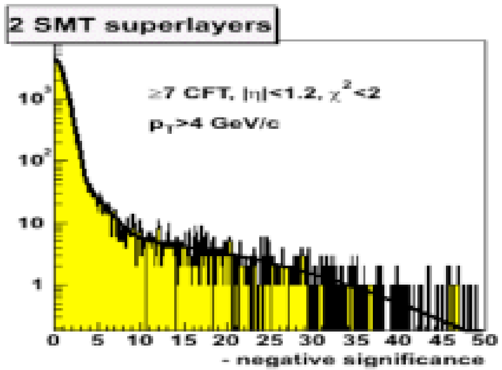
- DØ b-tagging group: several groups worked on development and evaluation of different algorithms:
  1. Counting Signed Impact Parameter:
     - simple algorithm and robust design
     - Based on Impact Parameter Significance: S(IP) = IP/s(IP)
     - IP is a signed quantity w.r.t to jet axis (so is S(IP)) :
       - positive if $\theta < \pi/2$, negative if $\theta > \pi/2$
     - **Requirements to tag a jet positively (negatively):**
       - at least 2 tracks with S(IP) > 3 (<-3)
       - or at least 3 tracks with S(IP) > 2 (<-2)





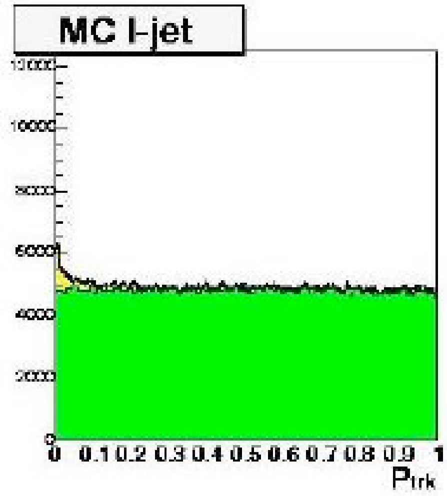
Tagging Algorithms

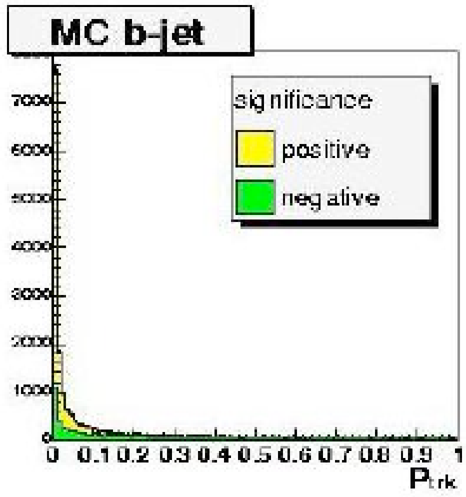
- DØ b-tagging group: several groups worked on development and evaluation of different algorithms:
  1. Counting Signed Impact Parameter:
  2. Jet Lifetime Impact Parameter:
     - Fit a resolution function (R) on the negative S(IP) distribution. Assume these tracks to **originate from primary vertex**.
       - Several track categories defined
     - Use then R to define a **probability** for a track to originate from the primary vertex



S. Greder Ph.D. Thesis

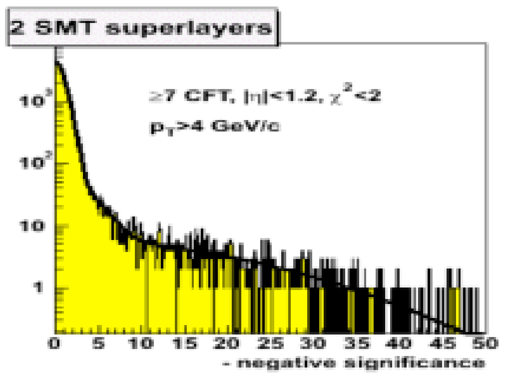
$$\mathcal{P}_{trace}(\mathcal{S}_{IP}) = \frac{\int_{-50}^{-|\mathcal{S}_{IP}|} \mathcal{R}(s)ds}{\int_{-50}^{0} \mathcal{R}(s)ds}. \qquad (4.6)$$





Tagging Algorithms

- DØ b-tagging group: several groups worked on development and evaluation of different algorithms:
  1. Counting Signed Impact Parameter:
  2. Jet Lifetime Impact Parameter:
     - Fit a resolution function (R) on the negative S(IP) distribution. Assume these tracks to **originate from primary vertex.**
       - Several track categories defined
     - Use then R to define a **probability** for a track to originate from the primary vertex
     - Combine then each track's probability to compute the probability for N tracks to originate from primary vertex : P_{jet} the jet lifetime probability



S. Greder Ph.D. Thesis

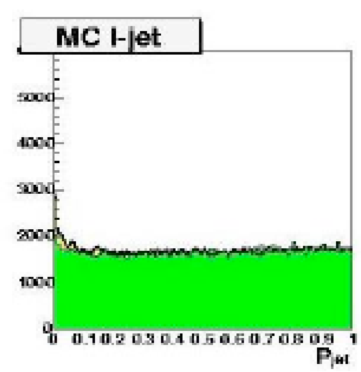
$$\mathcal{P}^{\pm}_{jet} = \Pi^{\pm} \times \sum_{j=0}^{N^{\pm}_{traces}-1} \frac{(-\log \Pi^{\pm})^j}{j!} \quad \text{où :} \quad \Pi^{\pm} = \prod_{i=1}^{N^{\pm}_{traces}} \mathcal{P}_{trace}(S^{IP>0}_{IP<0}) \qquad (4.7)$$

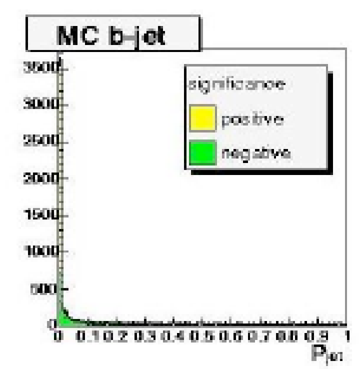
Requirements to tag a jet positively :

- P_{jet} (+) < cut

Requirements to tag a jet negatively :

- P_{jet} (-) < cut

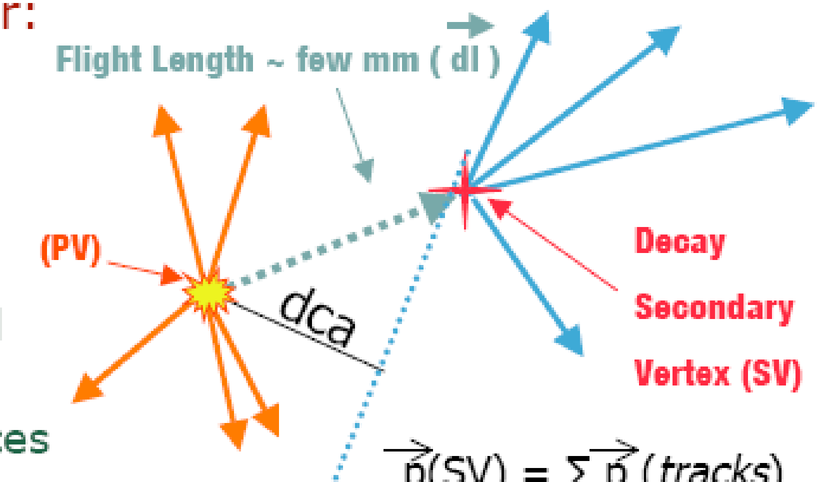




Tagging Algorithms

- DØ b-tagging group: several groups worked on development and evaluation of different algorithms:
  1. Counting Signed Impact Parameter:
  2. Jet Lifetime Impact Parameter:
  3. Secondary Vertex Tagger:
     - Build up track-based jets and fit their tracks to a secondary vertex
     - Select tracks with **high IP** to build secondary vertices
     - Tracks with high IP will form vertices with high **decay length significance** :

       $S(Lxy) = Lxy/s(Lxy)$.
     - Jet are then tagged by requiring a dR < 0.4 match between the jet axis and the SV.



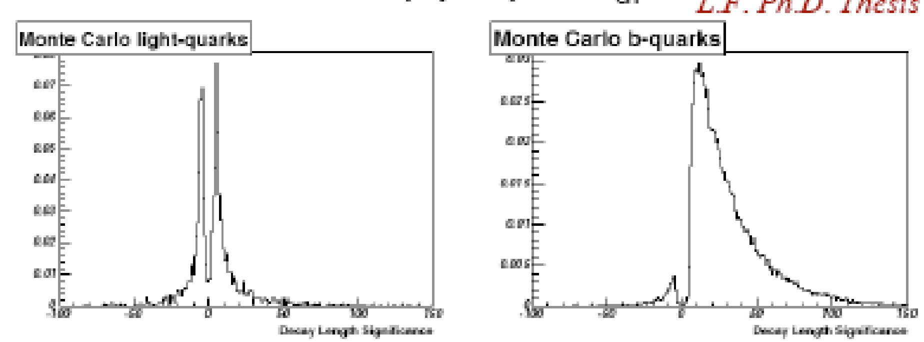
$$\vec{p}(SV) = \Sigma\, \vec{p}\,(tracks)$$

$$dls = \frac{\vec{dl} \cdot \vec{p}(SV) \cdot |\vec{dl}|}{|\vec{dl} \cdot \vec{p}(SV)| \cdot \sigma_{dl}}$$

L.F. Ph.D. Thesis

Requirements to tag a jet positively :
- match a secondary vertex with S(Lxy) > cut

Requirements to tag a jet negatively :
- match a secondary vertex with S(Lxy) < -cut



Tagging Algorithms

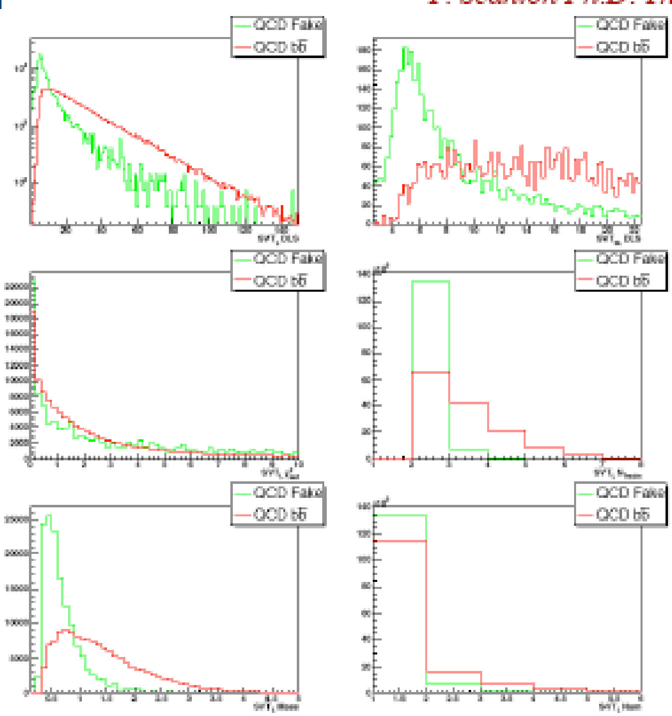
- DØ b-tagging group: several groups worked on development and evaluation of different algorithms:
  1. Counting Signed Impact Parameter:
  2. Jet Lifetime Impact Parameter:
  3. Secondary Vertex Tagger:
- A NN was developed using these different taggers to enhance b-tag performances

Start loose

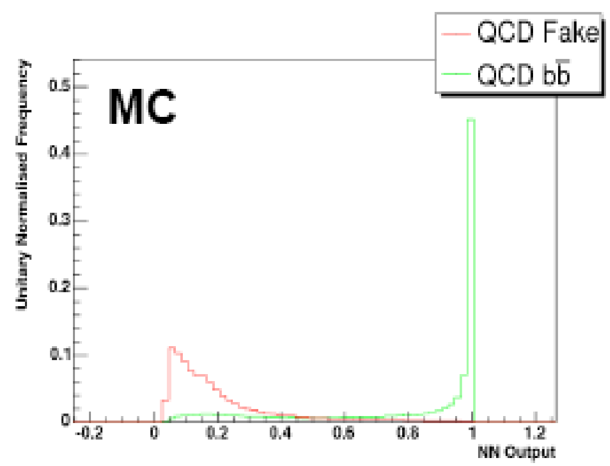
T. Scanlon Ph.D. Thesis

NN variables

Variable
SVT_{SL} DLS
CSIP Comb
JLIP Prob
SVT_{SL} χ^2_{dof}
SVT_L N_{Tracks}
SVT_{SL} Mass
SVT_{SL} Num
JLIP $Prob_{red}$
SVT_{SL} dR

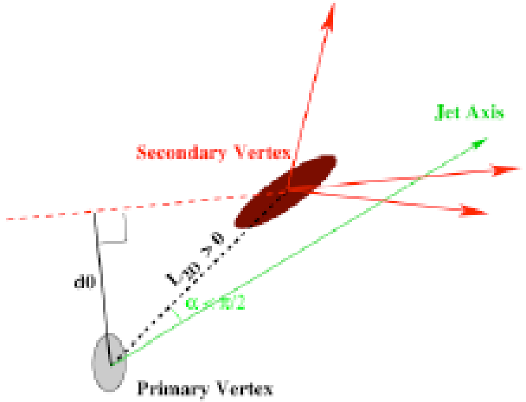


Large separation between b and 'loosely tagged' fake-jets



Tagging Algorithms

- Used for 1st observation of top quark
- Two successive attempts:
  - here shown TIGHT SecVtx configuration:

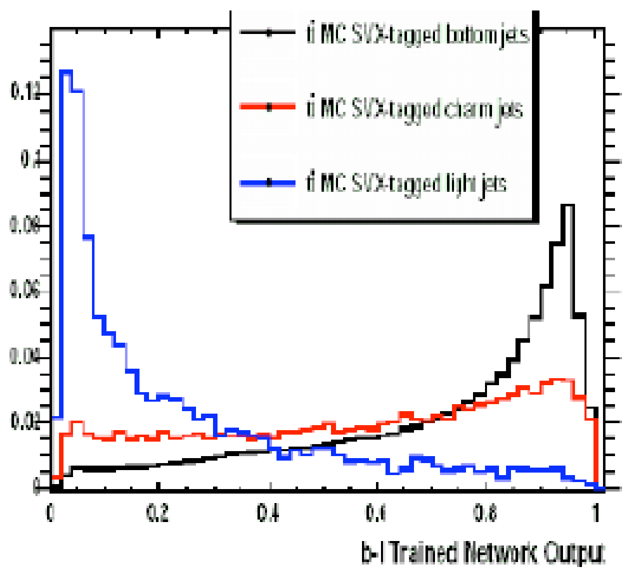


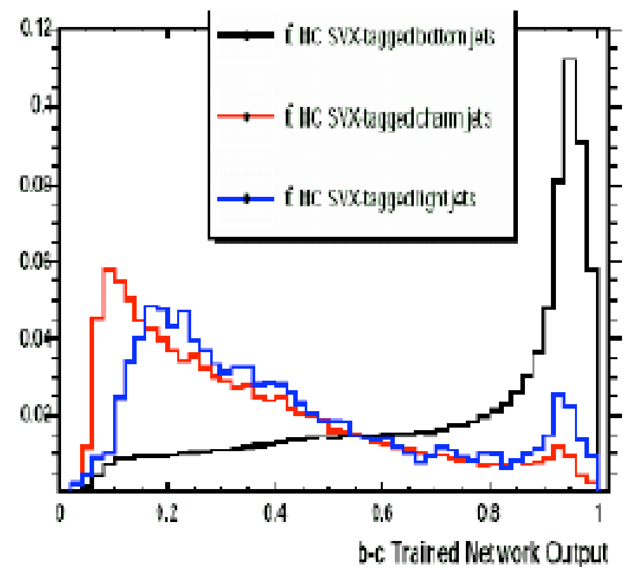
- Pass 1:
  - Loose track selection (pT>0.5 GeV)
  - Require >2 displaced tracks (|S(d0)|>2)
  - Attempt 3D Secondary Vertex reconstruction from displaced tracks (prune incompatible tracks iteratively until converge)
- Pass 2:
  - Tighter track selection but require only 2 displaced tracks (S(d0)>3.5, one with pT>1.5 GeV, other with pT>1 GeV)
  - Attempt Secondary Vertex reconstrucion from two tracks
- Tag if found a SV. and L2D/sigma(L2D) > 7.5

Tagging Algorithms

- Start with a SecVtx tag and filter out c and light jets with NN
- 16 variables (8 relative to SecVtx):

SECVTX variable	SECVTX independent variable
• Number of tracks in SECVTX	• Number of good tracks
• Fit χ^2	• Jet Probability (JetProb)
• Transverse decay length (L_{xy})	• Reconstructed mass of pass 1 tracks
• L_{xy} significance ($L_{xy}/\sigma_{L_{xy}}$)	• Reconstructed mass of pass 2 tracks
• Pseudo-$c\tau$ ($L_{xy} \times M_{SECVTX}/p_T^{SECVTX}$)	• Number of pass 1 tracks
• Vertex Mass ($\sqrt{(\sum \lvert p_{vtx} \rvert)^2 - (\sum p_{vtx})^2}$)	• Number of pass 2 tracks
• $p_T^{vtx}/(\sum_{good\ tracks} p_T)$	• $\sum_{Pass1\ track} p_T/p_T^{jet}$
• Vertex pass number (pass 1 or 2)	• $\sum_{Pass2\ track} p_T/p_T^{jet}$

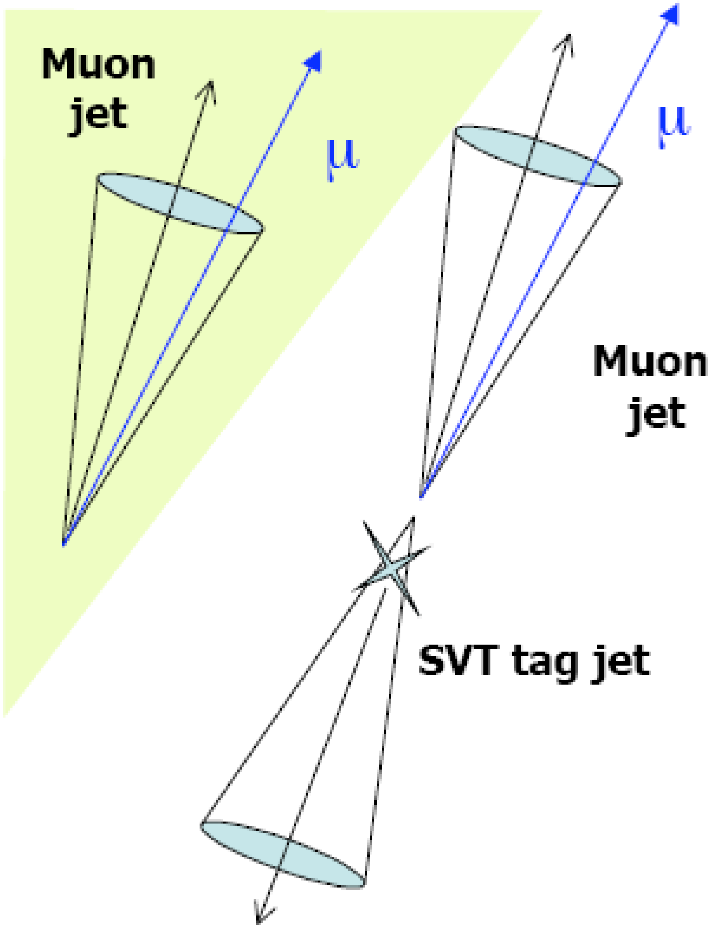




Measuring b/c Tagging Efficiencies *in data*

b-tagging efficiency from data

- System 8: System of **8 equation** and **8 unkowns**
  - 2 samples
    - Muon Jets (n)
    - MJ + opposite lifetime tag Jet (p)
  - 2 different b and light content
    - $n_b \neq p_b$
  - 2 uncorrelated tagging algoritms
    - Track based
    - Muon Tag ($p_T^{rel} > 0.8$ GeV)
- 2 taggers are applied separately and together to the two different samples
- Correlation among the two taggers is derived from simulation
  - only MC input
- **System 8 is analytically solvable!**
- High statistics allow to solve the system in bins of jet p_T and η
- DØ only b-tagging efficiency estimation algorithm

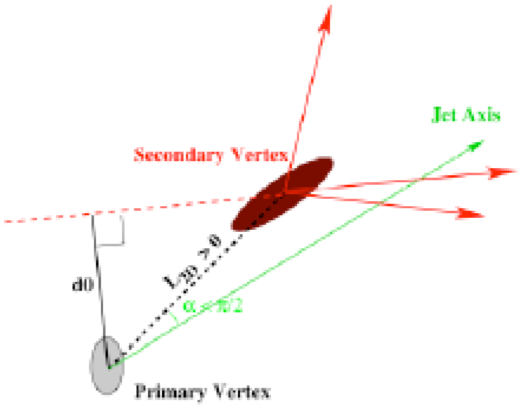


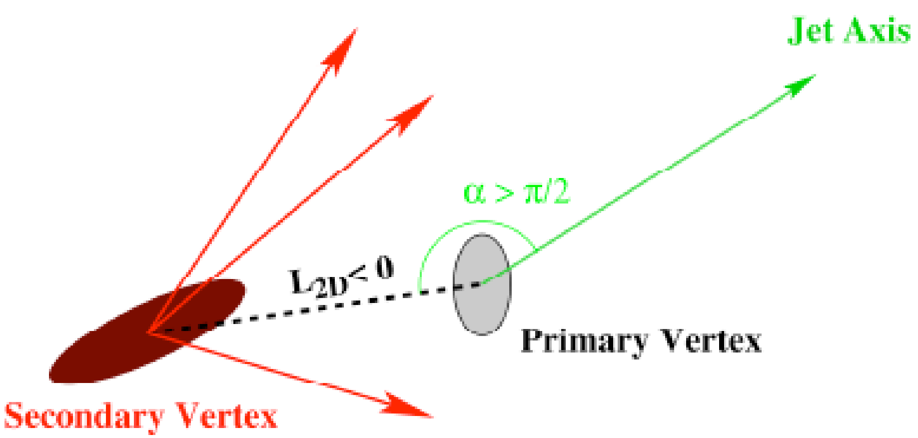
Efficiency measurement

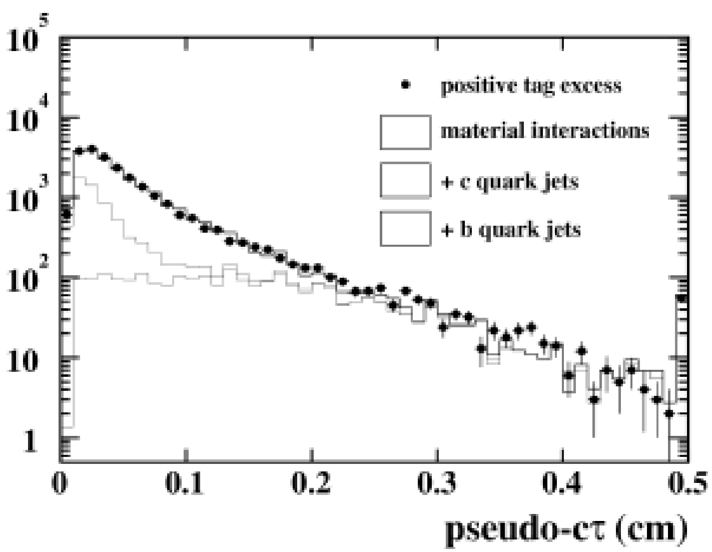
- Efficiency measured in data
  - «Scale Factor» = eff(data)/eff(MC) is a constant
  - MC describes well enough non-uniformity (ET, eta, phi, # tracks) -> taken as systematic error
- Use di-jet sample triggered by 8 GeV electron or muon
- Opposite jet b-tag used to purify lepton-jet HF content
  - Electron: use conversion to estimate fraction of light jets
  - Muon: use pT(relative to jet axis)
- Efficiency in data is ~ 10% too low. Estimated at 7% level. Scale Factor = 0.909±0.06 (as of last year)
- Scale Factor NOT efficiency applied for physics analysis

Mistag rate from data

- «Negative» tag rate parametrized in several jet samples, as a function of:
- Jet ET, eta, phi, #tracks in jet, Sum(ET)
- Correct for asymmetry between negative and positive tag rate
  - decay in flight plus material interaction
- Estimate this asymmetry with c-tau fit in jet samples. Pos fake = 1.37±0.15· Neg Fake

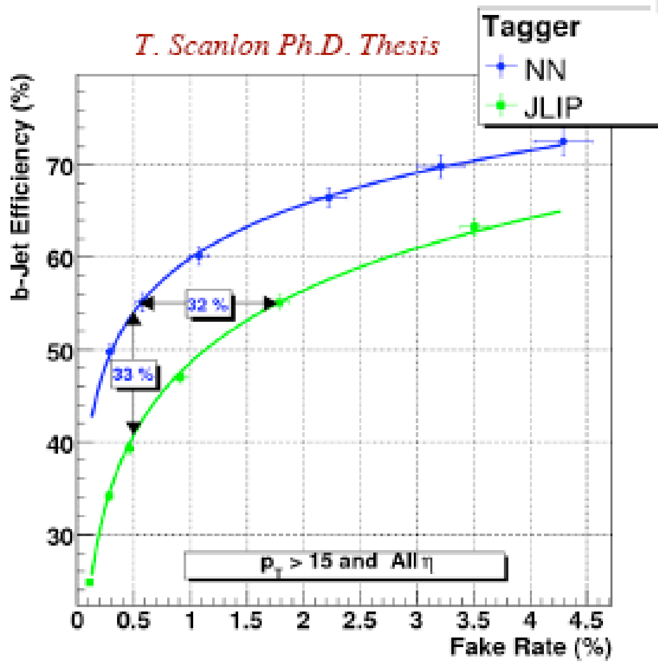




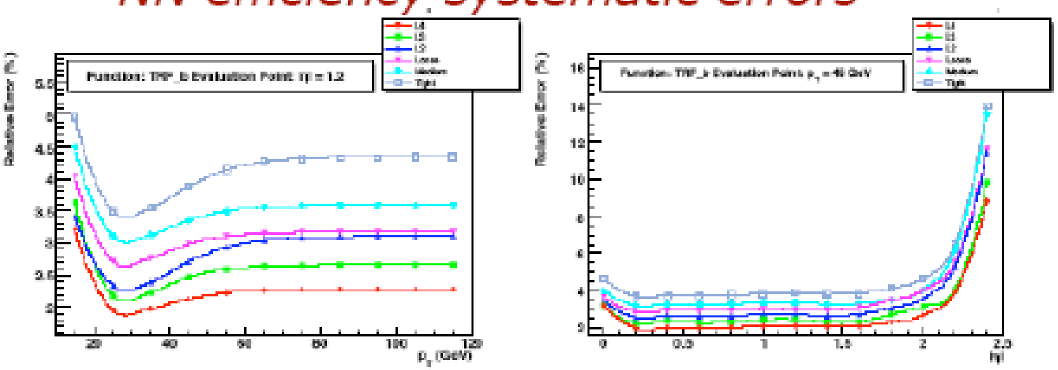


Tag Rate Functions

- Taggers are optimized for several working points
  - Efficiencies, scale factors, systematic uncertainties calculated only for these settings
- Different sources of systematic errors:
  - System 8 correlation functions
    - efficiency is recalculated after varying correlation factors
  - Parameterization error
    - correlation between p_T and η TRF parameterizations
  - MC sample dependencies



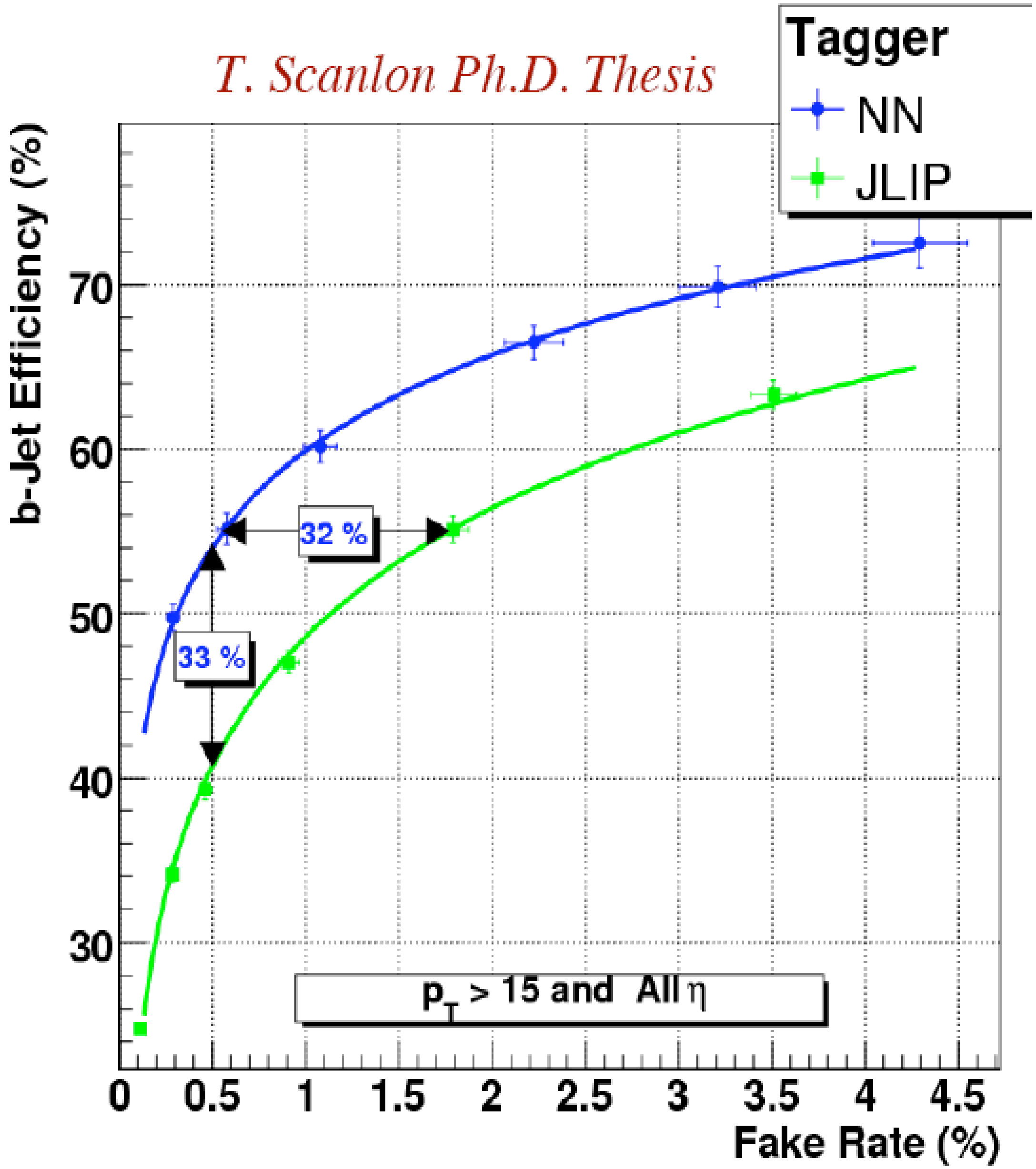
NN efficiency systematic errors

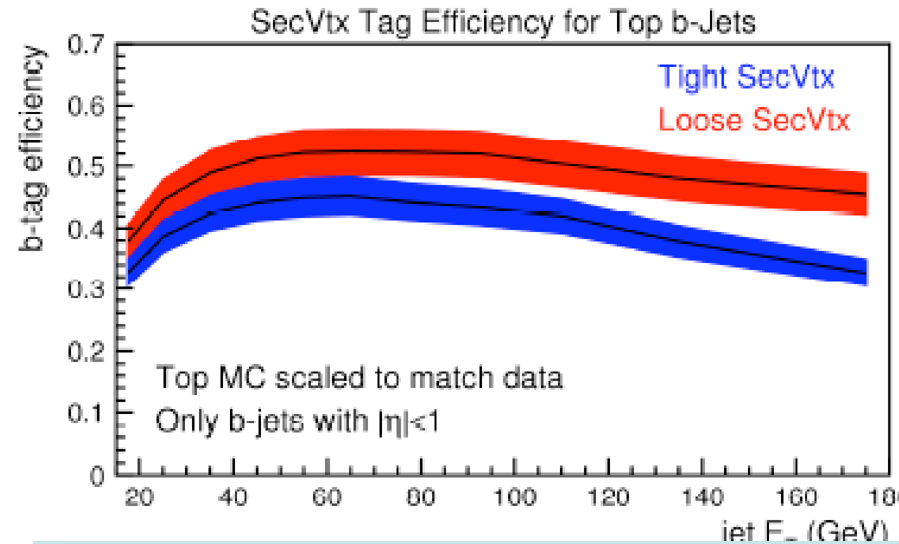


Analysis Strategies

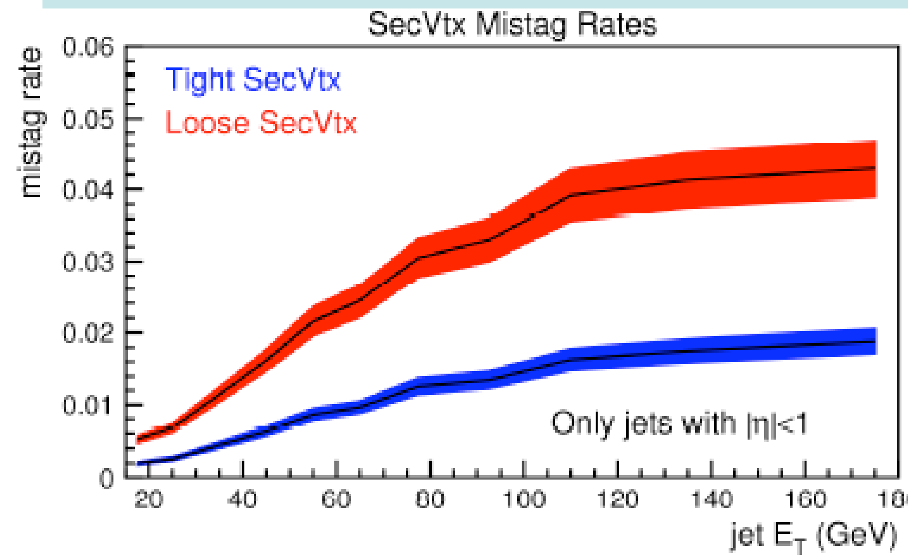
- Use Tag Rat Functions
  - derived from data
  - known systematics
- Use Scale Factors
  - b-tagging on MC
  - allows MC to predict unkown source of systematics
    - I.e. hidded taggability dependency

Performance





Eff Scale Factor = 0.909±0.06 for the tight



Use of b-tagging in physics analyses

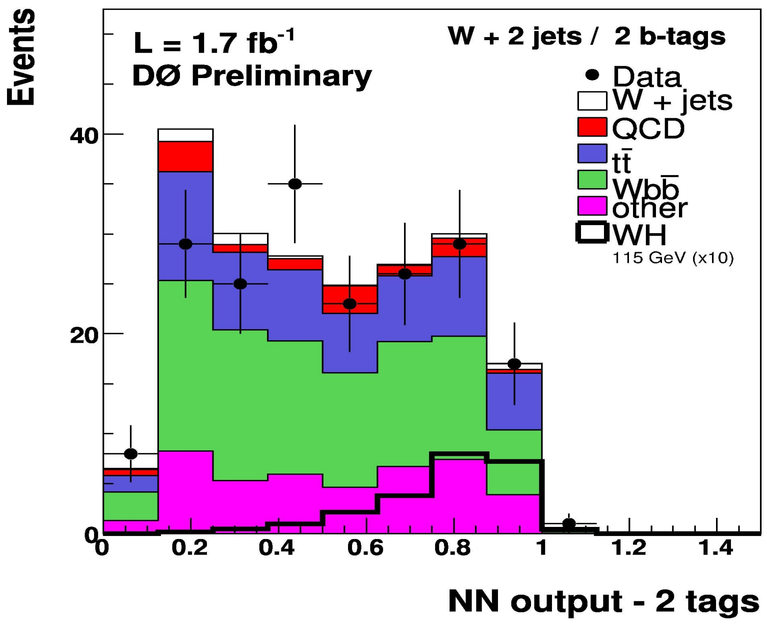
Single-tight / Double-loose

WH(->bb) as an example

- *Note: 2 b's in most final states!*

Base:

Find optimal double b-tag working point (usually very loose)



Single-tight / Double-loose

WH(->bb) as an example

- *Note: 2 b's in most final states!*

Base:

Find optimal double b-tag working point (usually very loose)

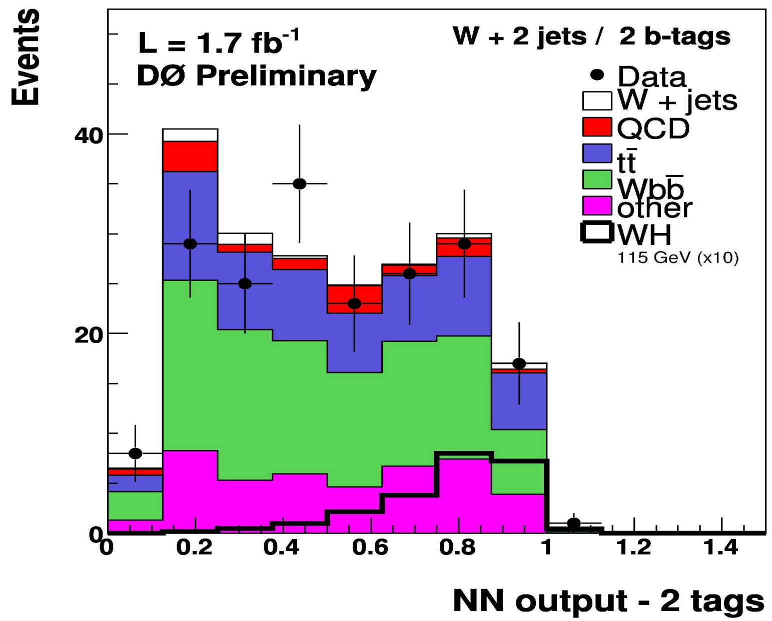
Better:

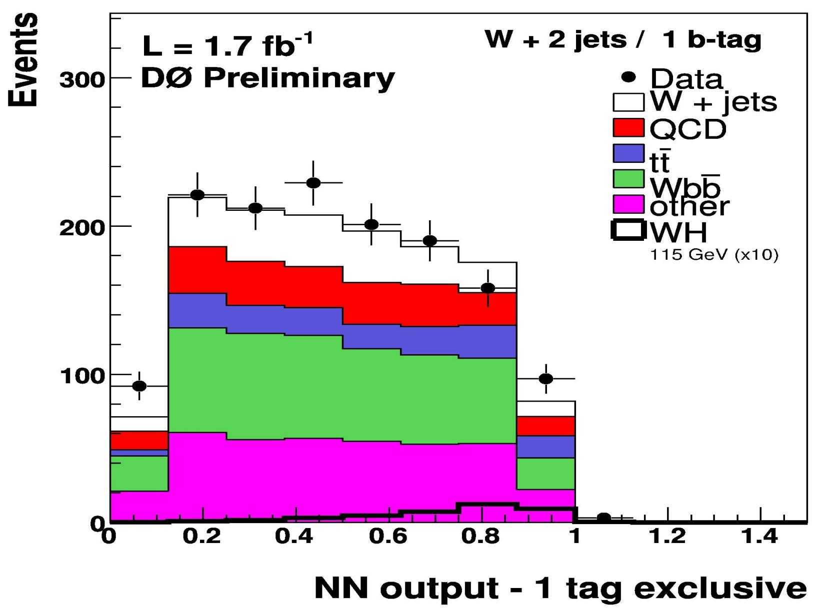
Add optimal (orthogonal) single b-tag working point

- Usually much tighter
- Careful: adding single tag can change the optimal double tag working point!

Gain in performance:

~10% in sensitivity compared to one working point





Two Working Points

Better still:

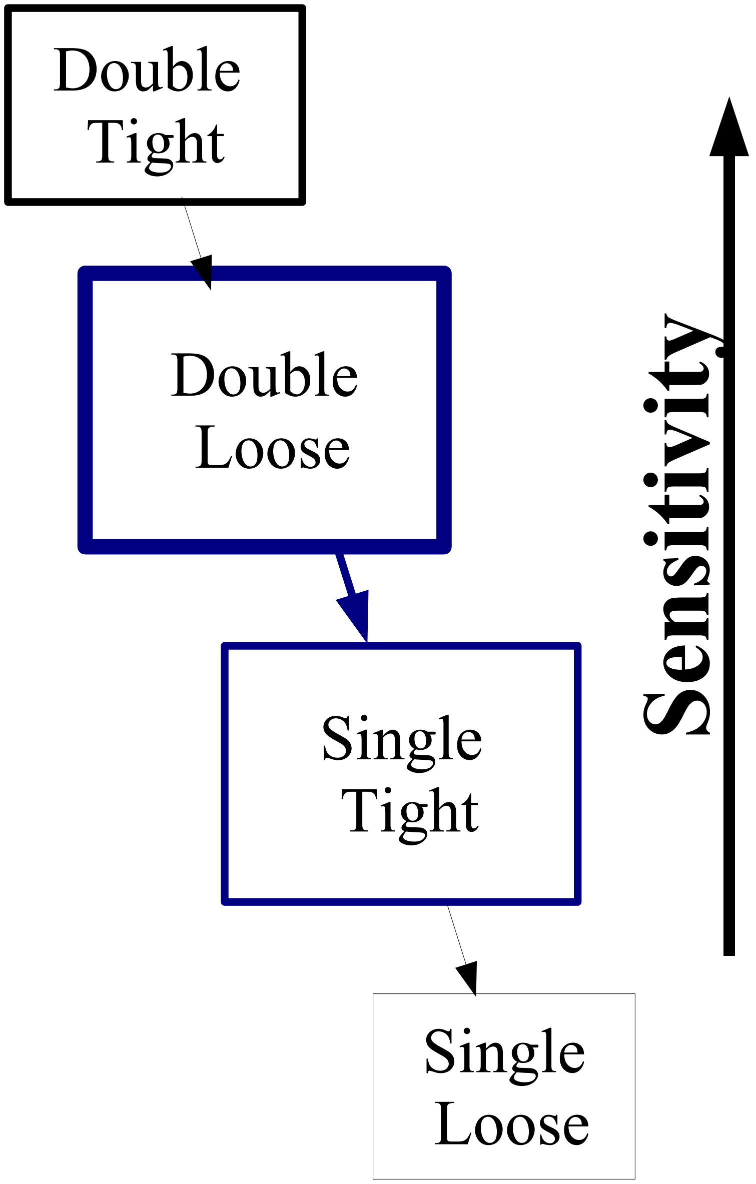
Add double-tight working point

And single-loose working point

(Keep selections orthogonal!)

Gain in performance:

~2% in sensitivity compared to two working points



Many Working Points

Even Better Still:

Use 3,4,... <u>12</u> working points

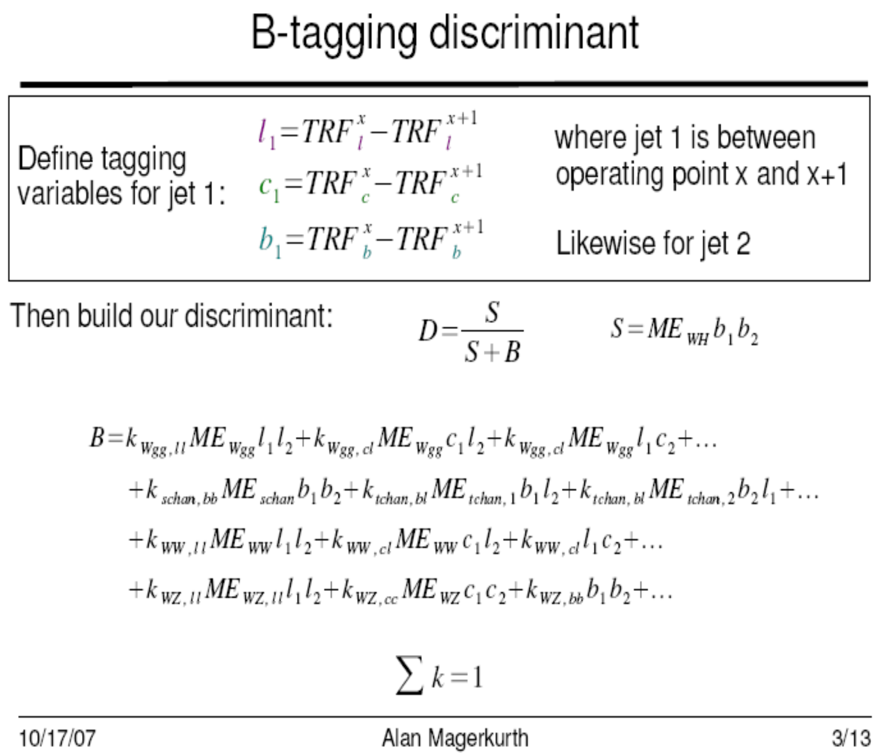
Including *untagged* there are 169 boxes (1 and 2 tags)

Order in terms of sensitivity

(and keep orthogonal)

Gain in performance:

~10% in sensitivity compared to two working points



Infinite Working Points

Use the output of the NN tagger

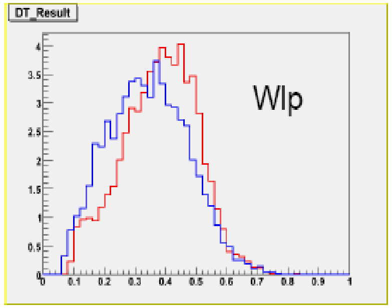
Danger: how to account for data/MC scale factors?

Under development

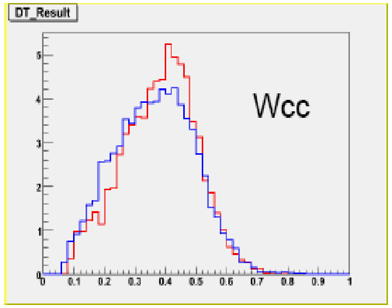
Gain in performance:

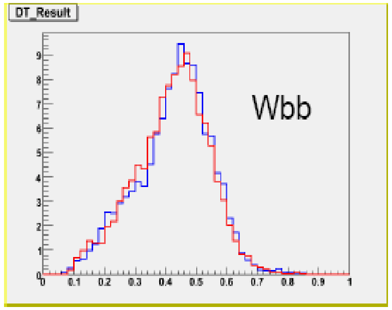
~0% in sensitivity compared to 12 working points

But can be added much more simply as an input to another NN...



- Red is DT output with binary NNTag variable
- Blue is DT output with continuous NN output





John BackusMayes

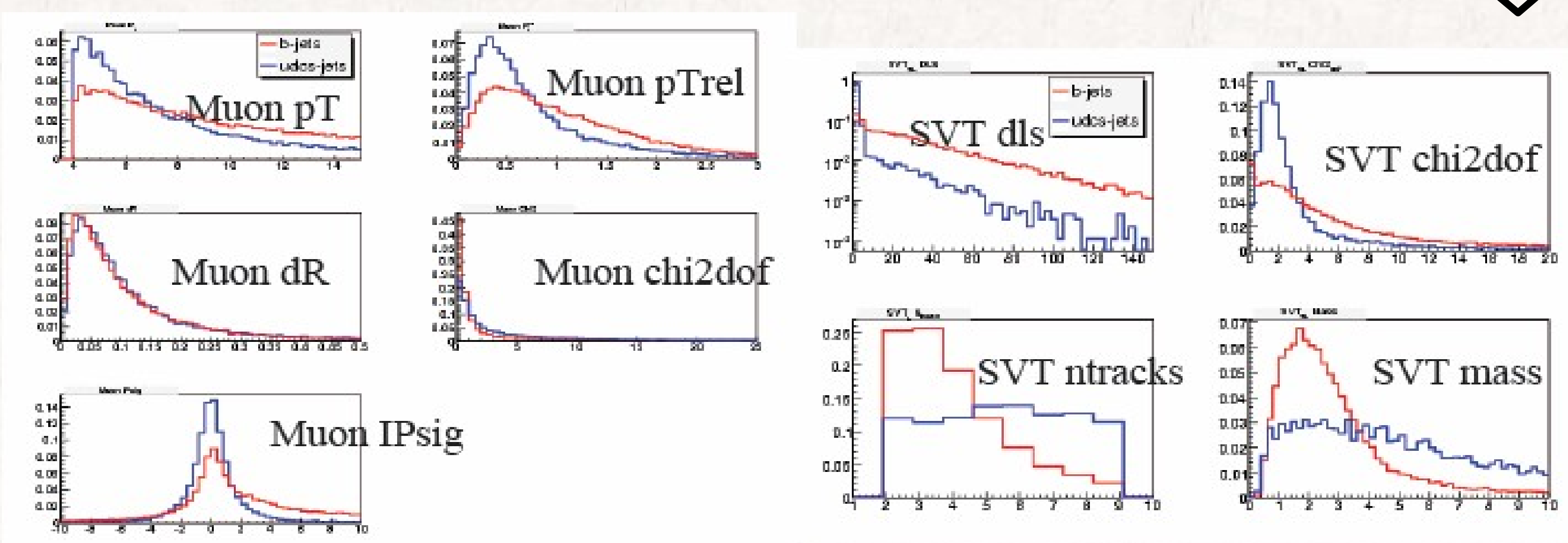
Soft-lepton Tagging

BROWN UNIVERSITY

Neural Network

- Input variables
  - muon: P_t^{rel}, IP significance, ΔR, pT, χ^2_{dof}
  - SVT: $SVT_{SL}DLS$, $SVT_{SL}\chi^2_{dof}$, $SVT_L N_{tracks}$, $SVT_{SL}Mass$
  - SVT variables involved with muon track only

Muon-only so far!



09/27/07 b-id Meeting *Hwidong Yoo* *Slide 6*

tau Tagging

Use a tau NN

- Input track ip as a variable?

No use of tracking for tau identification at Tevatron, so far

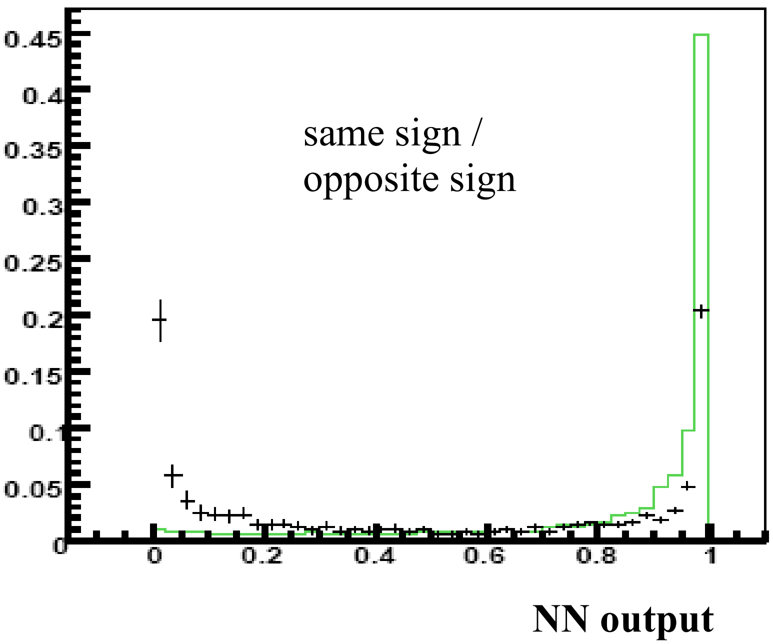
- (at D0, and I'm pretty sure at CDF)

Issues:

- Short tau lifetime
- Small tau mass
- Backgrounds from b/c jets
- Data/MC disagreements

Could be used if there were *excellent* tracking and different backgrounds

- ILC?



Conclusions / Summary

Good b/c tagging starts with good tracking!

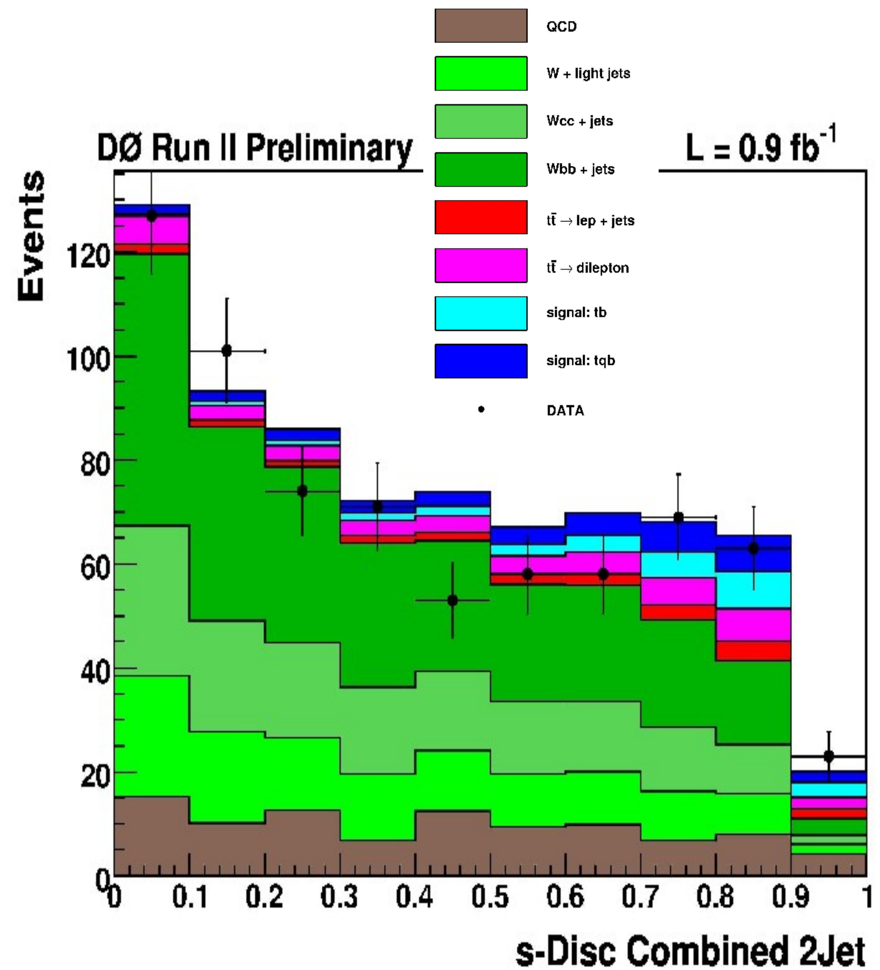
Combine jet lifetime / vertex info into NN's trained on MC to separate b/c/light jets

Measure performance in data using muon-tags / away tags / negative tags -> TRF's

Use *multiple operating points simultaneously* to take advantage of all information from tagging

Soft-lepton tagging (muons, electrons?)

tau-tagging -> for the ILC!



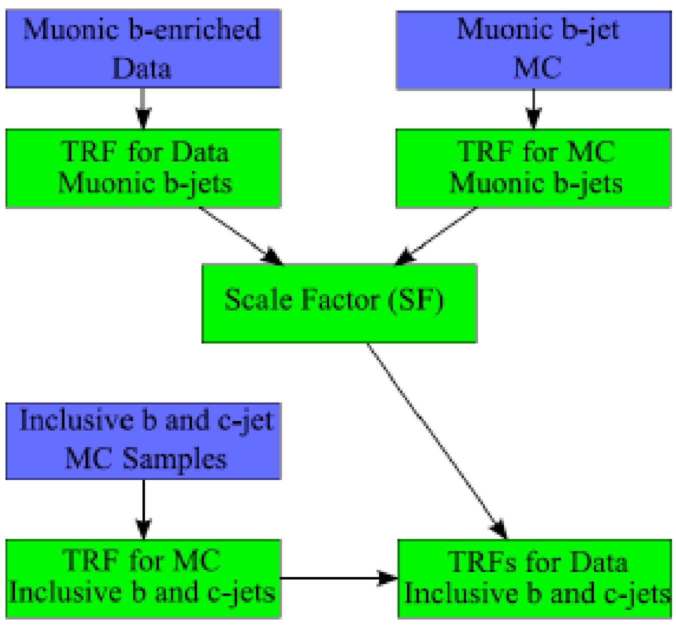
Evidence for single-top!

Backup

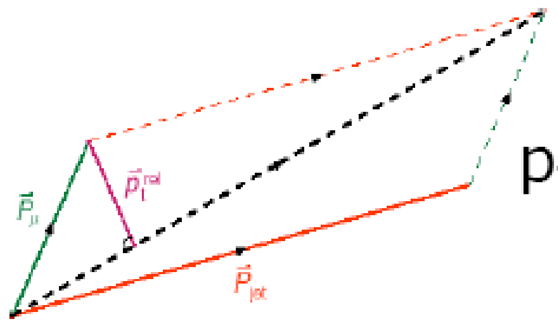
Performance estimation

- For any major software release DØ undergo the process of re-evaluating its b-tagging capability
  - b-tagging efficiency evaluation
    - Measurement from DATA using Muon Jets
      - Muon Jets: jets with an associated muon
    - Scale factors from Monte Carlo
      - Muon Jet b-tagging efficiency
      - c-tagging efficiency
    - Evaluation of systematic errors
    - Tag Rate Functions provided to end users for physics analysis
  - Mistag Rate
    - Negative tagging rate from several samples
      - QCD and EM samples
    - Correction from Monte Carlo
      - asymmetry and heavy flavor content
- Taggability: in order to decouple b-tagging efficiency to effects not simulated in MC (i.e. inefficiency of SMT modules) only jets associated to tracks are used for b-tagging
  - Taggability measured in DATA and used to scale MC

T. Scanlon Ph.D. Thesis

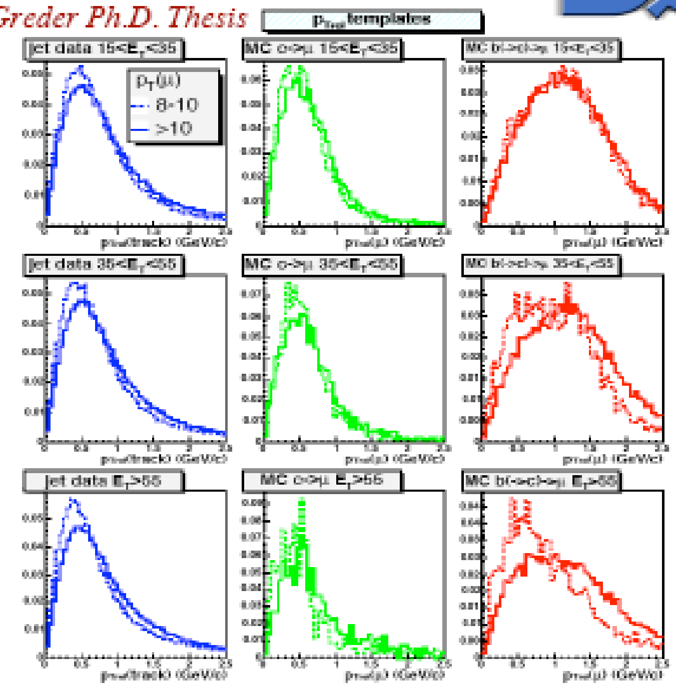


b-tagging efficiency from data



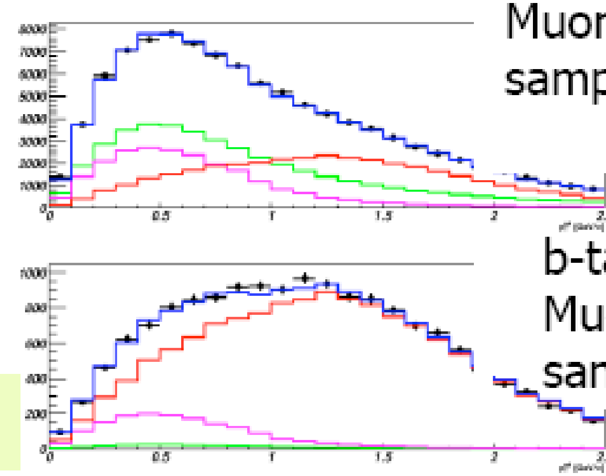
$p_T^{rel}(\mu)$ templates

S. Greder Ph.D. Thesis



- b-tagging efficiency is estimated from Muon Jets sample (high b-quark content)
- b-tagging efficiency is derived from the estimation of b-quark content of the Muon Jet sample before and after applying b-tagging
  - Fit of p_T^{rel} templates
  - p_T^{rel} depends upon jet and muon p_T
    - Several templates used
- Currently used only for counter checks

L.F. Ph.D. Thesis



Muon Jet sample

b-tagged Muon Jet sample

$$\varepsilon_{b\text{-tagging}} = (N_{\text{tagged jets}} \cdot f^{b}_{\text{tagged jets}}) / (N_{\text{jets}} \cdot f^{b}_{\text{jets}})$$