Heavy flavor properties of jets produced in $p\bar{p}$ interactions at $s^{1/2}$=1.8 TeV

A small step forward in resolving the long-standing (alleged) discrepancies between the measured and predicted b-cross sections at the Tevatron

hep-ex/0311051

P. Giromini
INFN/Frascati

Jets with heavy flavor

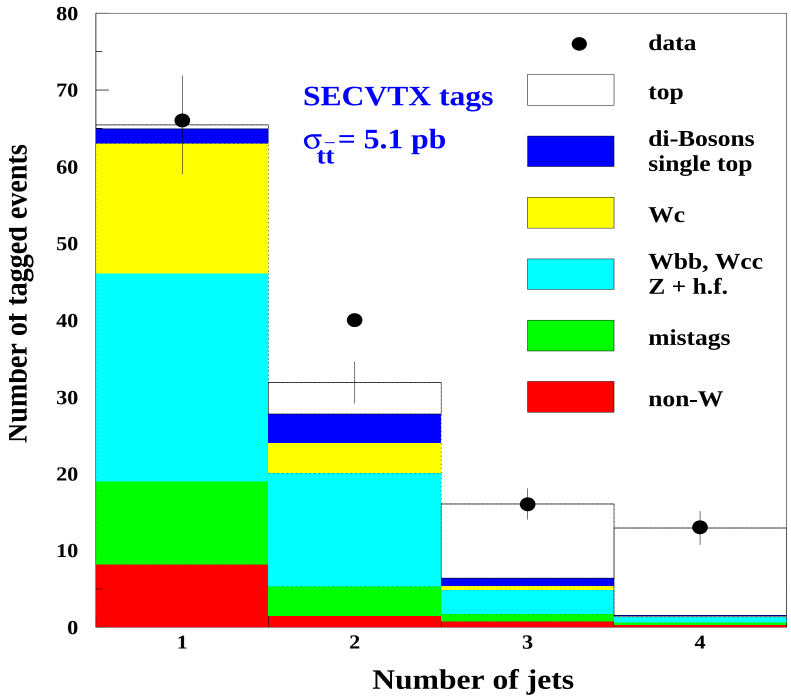
- Over several years CDF has been comparing the fraction of jets with heavy flavor (b and c quarks) to a simulation based upon the Herwig and CLEO (QQ) Monte Carlo generators
- Heavy flavor-identification : Efficiency
-

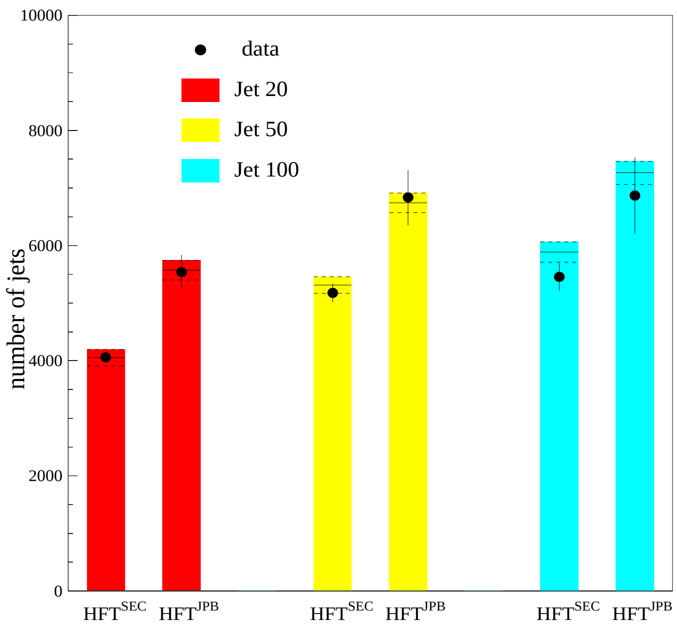
		b quark	c quark
✓ SECondary VerTeX	(SECVTX)	43%	9%
✓ Jet-ProBability	(JPB)	43%	30%

- Data sets : *W*+ jet events, generic-jet data (JET20, JET50, and JET 100), di-jet events with one jet containing a lepton (lepton-triggered sample)

Jets with heavy flavor

- We have used the lepton-triggered sample to calibrate the data-to-simulation scale factors for the SECVTX and JPB tagging algorithms
- We have used generic-jet data to tune the parton-level cross sections evaluated in Herwig





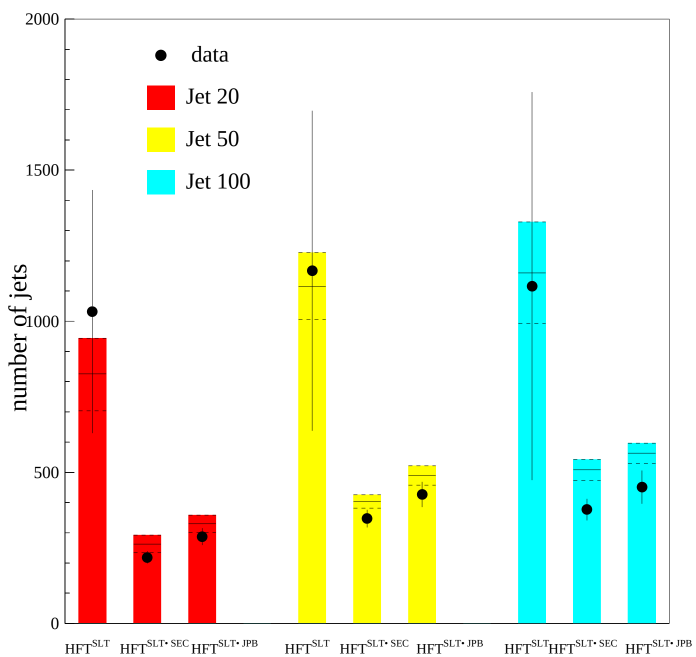
W+jet events

P~50%

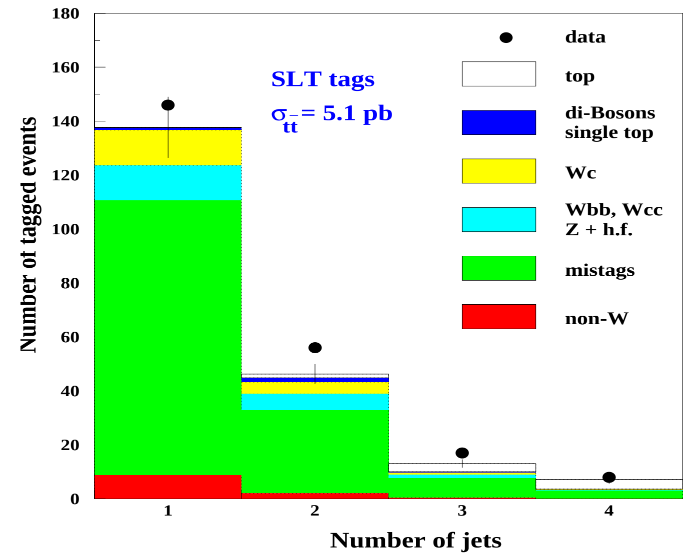
- PRD 64, 032002 (2001)

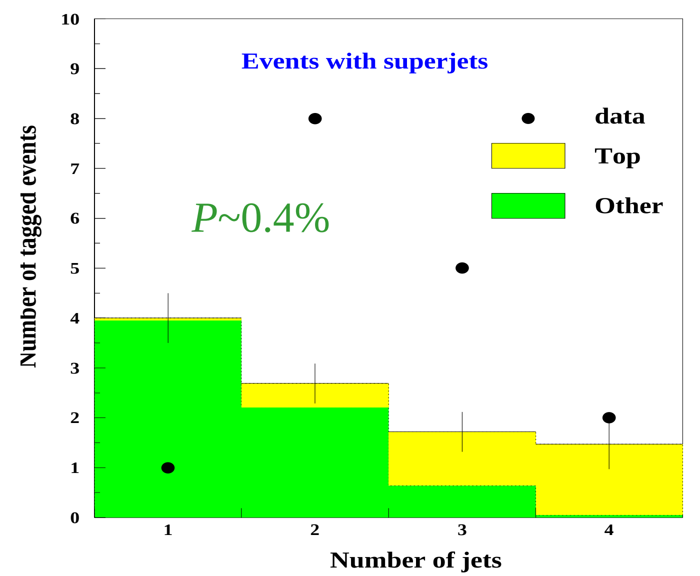
Jets with heavy flavor

- We also identify heavy flavors by searching jets for semileptonic decays (SoftLeptonTagging)
  efficiency 6.4 % (b) and 4.6% (c)
- PRD 65, 052007 (2002)



P~50%





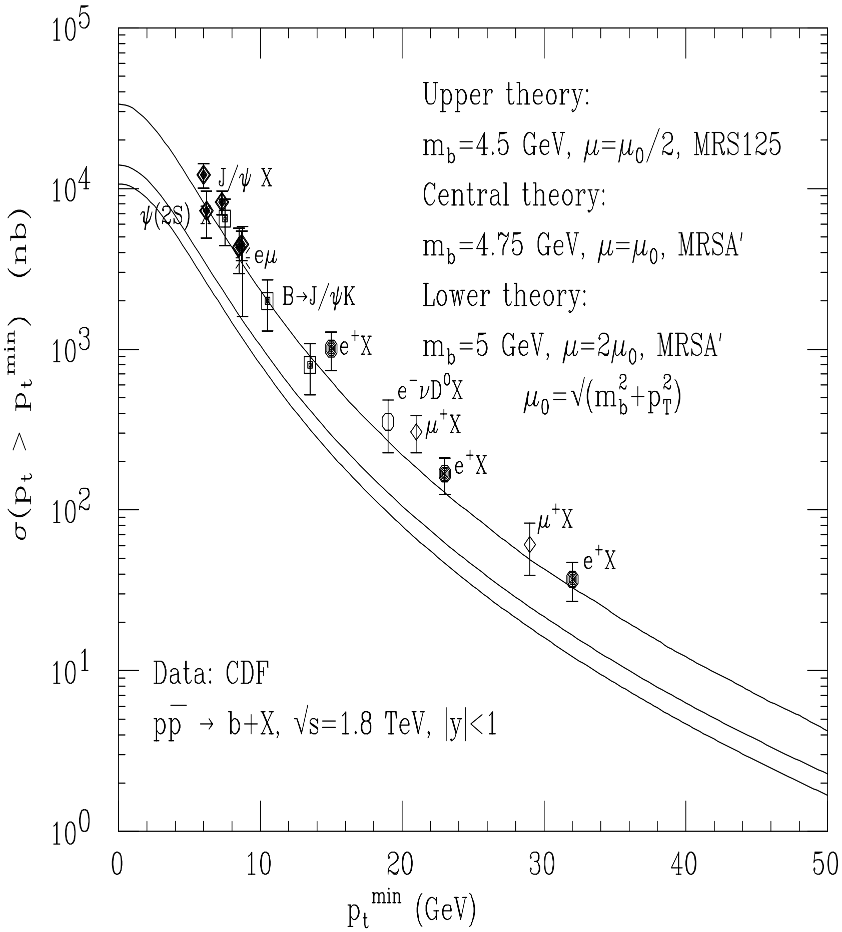
Anomalous W+ 2,3 jet events with a supertag

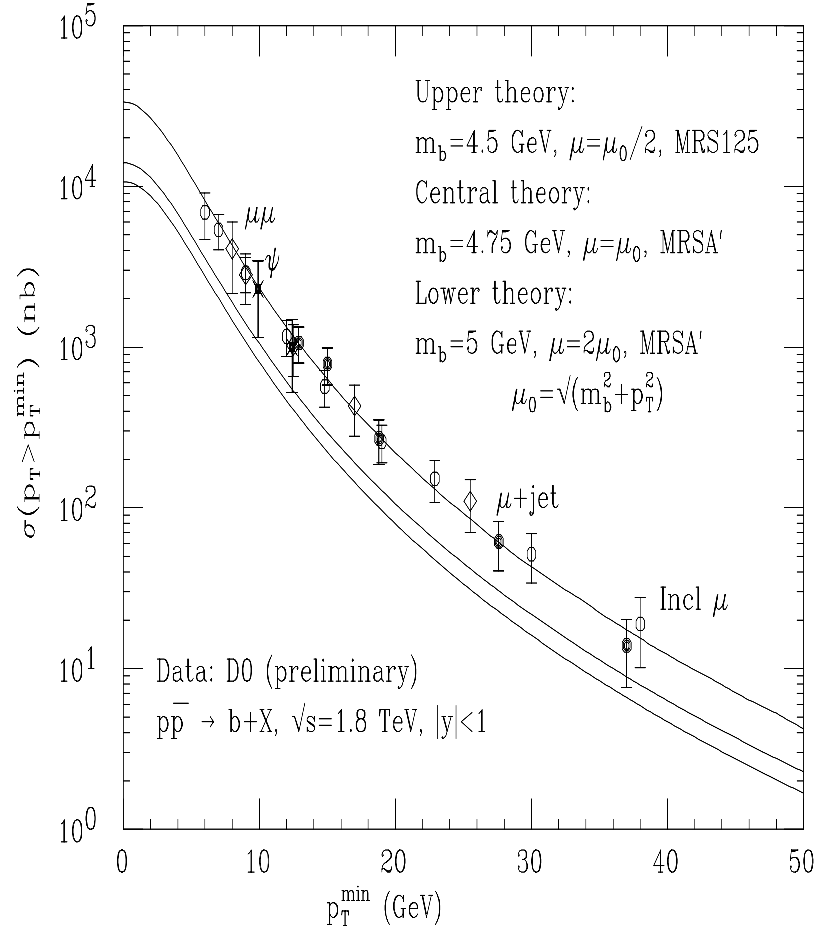
- The kinematics of these events has a 10^{-6} probability of being consistent with the SM simulation [PRD 64, 032004 (2002)]
- hep-ph/0109020 shows that the superjets can be modeled by postulating the existence of a low mass, strong interacting object which decays with a semileptonic branching ratio of the order of 1 and a lifetime of the order of 1 ps
- Since there are no limit to the existence of a charge –1/3 scalar quark with mass smaller than 7 GeV/c^2 [PRL 86, 1963 (2001)], the supersymmetric partner of the bottom quark is a potential candidate

Light sbottom (b_s)

- Lot of recent interest on the subject
- hep-ph/0007318 uses it to resolve the long-standing discrepancy between the measured and predicted value of *R* for $5 < s^{1/2} < 10$ GeV at $e^+ e^-$ colliders
- PRL 86, 4231 (2001) uses it in conjunction with a light gluino which decays to b b_s to explain the difference of a factor of 2 between the measured b-quark production cross section and the NLO prediction
- If light b_s existed, Run 1 has produced 10^9 pairs; why we did not see them ?

Single b cross section

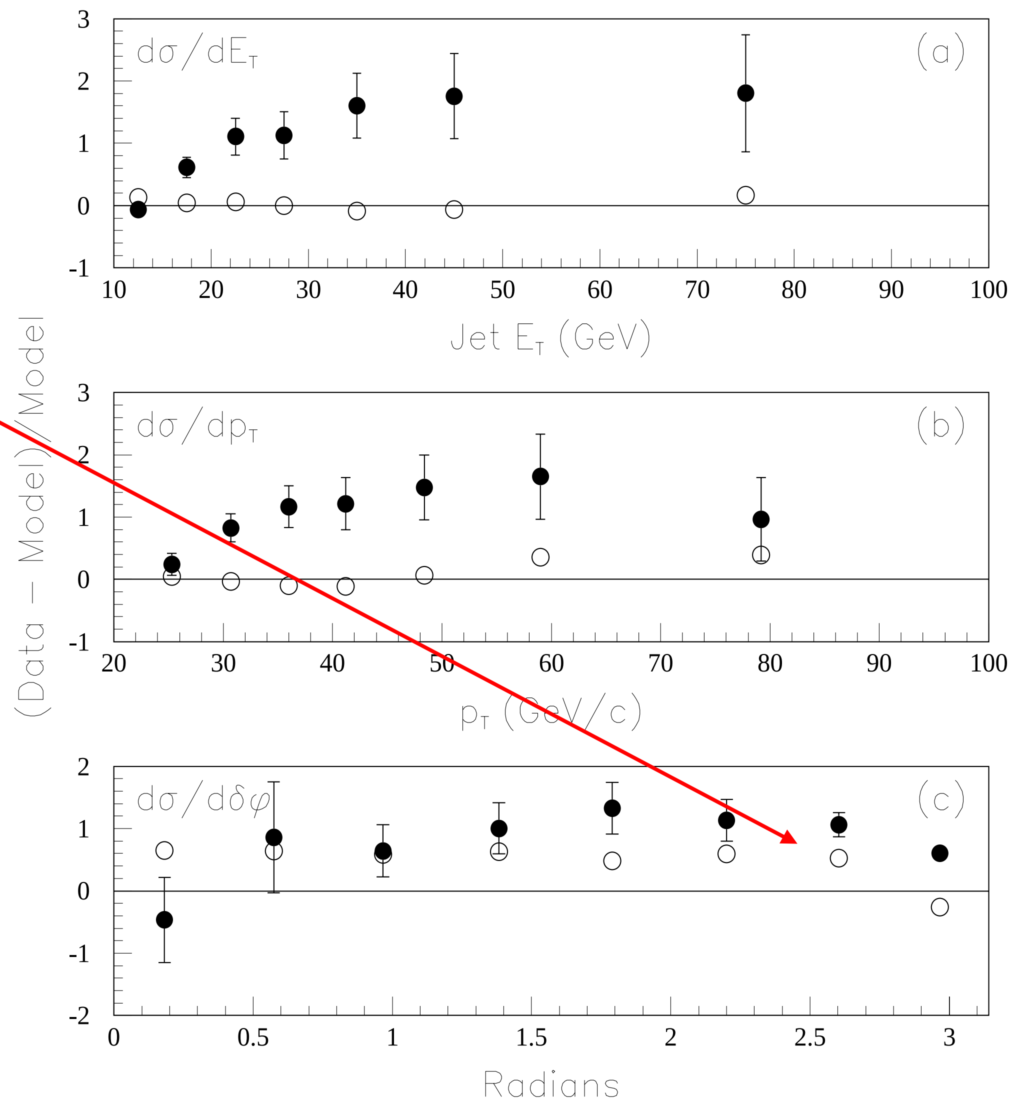




correlated μ+$\bar{b}$-jet cross section

σ_{bb}•BR

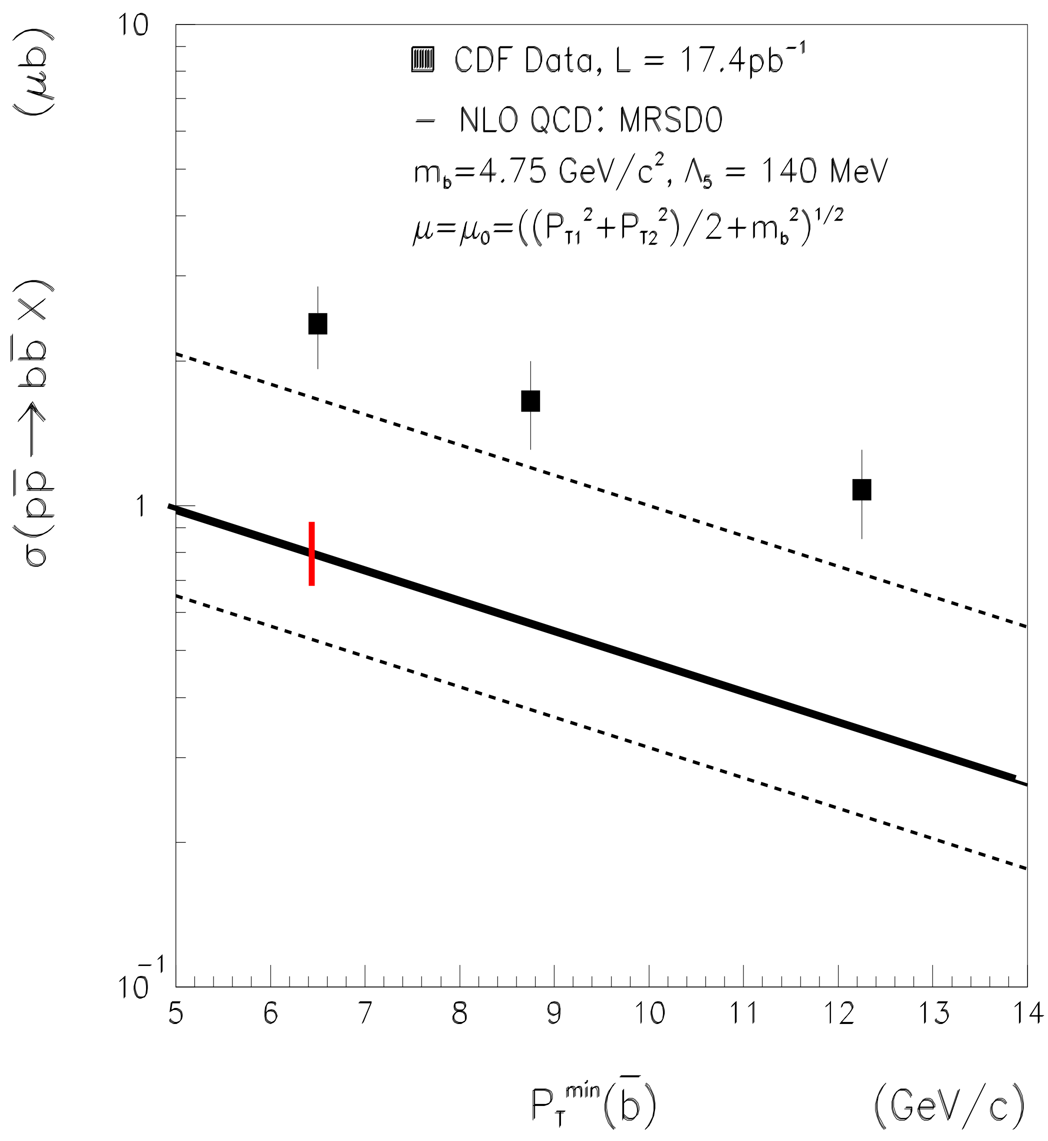
- PRD 53, 1051 (1996)
- Data are 1.5 times larger than the NLO calculation; however
  - ✓ The NLO cross section is not very sensitive to the scales μ
  - ✓ The NLO value is approximately equal to the Born value



$b\bar{b}$ correlations (dimuons)

$\sigma_{bb} \bullet BR^2$

- PRD 55, 2547 (1997)
- Data are 2.2 times larger than the NLO calculation
- DØ has a similar result
- The NLO cross section is not very sensitive to the scales μ (±20%)
- Born and NLO values are within a few percents



What if there is a light sbottom ?

- the NLO calculation of $p\bar{p} \rightarrow \bar{b}_s b_s$ predicts $\sigma = 19.2$ μb for a squark mass of 3.6 GeV/c^2 (Prospino MC generator program) .
- The $b\bar{b}$ production cross section at the Tevatron is $\sigma = 48.1\ \mu b$ (NLO)
- The $c\bar{c}$ production cross section at the Tevatron is $\sigma = 2748.5\ \mu b$ (NLO)

What if….

- We have tuned the heavy flavor parton-level cross sections calculated by Herwig within the theoretical and experimental uncertainties to reproduce the rate of SECVTX and JPB tags observed in generic-jet data.
- In that study we have used jets with with uncorrected E_T>15 GeV and $|\eta|<1.5$; they correspond to partons with transverse energy approximately larger than 18 GeV
- For partons with transverse energy larger than 18 GeV, σ = 84 nb , σ = 298 nb , and σ = 487 nb (10% contamination)
- we could have easily tuned the Herwig generator to explain in terms of SM processes an additional 10% pair production of scalar quarks: σ^f = 382 nb , and σ^f = 487 nb

What if .....

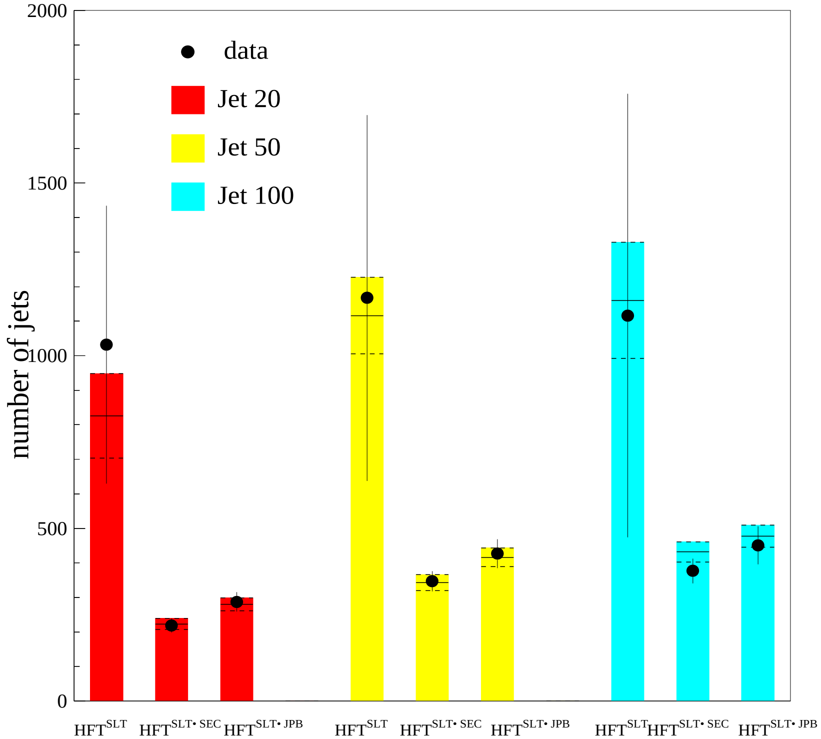
- What if the b_s quark with a 100% semileptonic branching ratio
- In b-quark decays, a lepton is produced in 37% of the cases
- In c-quark decays, a lepton is produced in 21% of the cases
- Compare data and tuned simulation as a function of the number of jets containing a lepton

Strategy

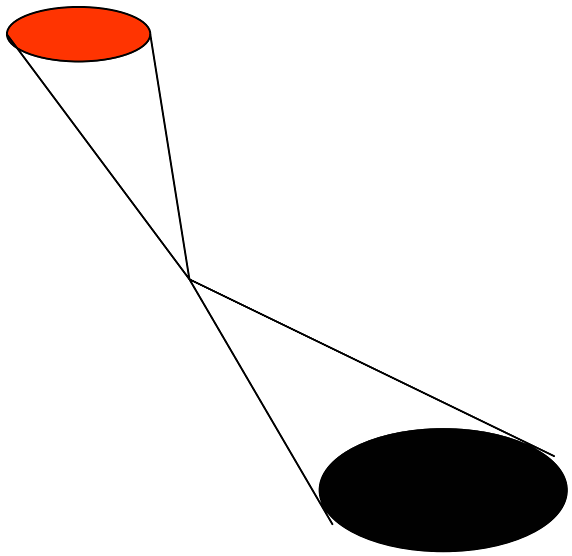
	σ (nb)				b_s(%)	fitted QCD			σ/σ_{QCD}	
	b	c	b_s	total		b	c	total		
generic jets tuned	298	487	84	869	10%	382	487	869	1	
g. j. t. x BR	110	102	84	296	28%	141	102	243	1.2	CS
g. j. t. x BR^2	41	22	84	147	57%	52	21	73	2	
g .j. x BR tuned (or lep-trig. evts)	110	102	84	296	28%	194	102	296	1	
lep-trig. evts. x BR	41	22	84	147	57%	72	21	93	1.5	SS

- Generic-jet comparisons have been already reported in PRD 65 052007 (2002), in which the sample CS has been used to adjust simulated lepton identification efficiencies. After this adjustment one could observe a 30% discrepancy between data and prediction in the sample SS

Control sample



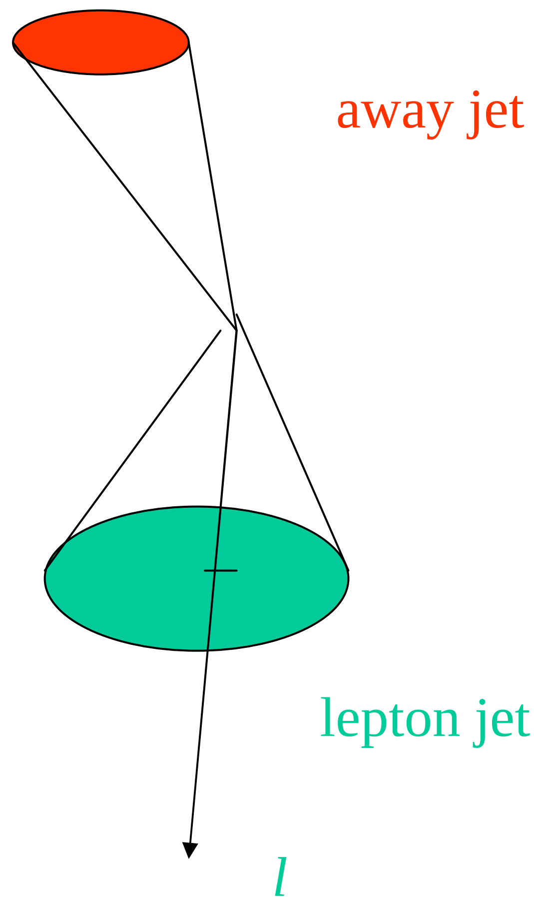
PRD 65 052007



- The simulation of the SLT algorithm uses efficiencies derived from the data (conversions, Z's, and ψ mesons)
- Use generic-jet data to calibrate and cross-check the efficiency for finding SLT tags and supertags
- Efficiency for finding supertags empirically corrected by 15%

Signal sample: lepton-triggered events

- Events with 2 or more jets with E_T > 15 GeV and at least two SVX tracks (taggable,|η|<1.5)
- one electron with E_T> 8 GeV or one muon with p_T > 8 GeV/c contained in one of the jets
- Require *I* > 0.1
- Reject conversions
- Apply all lepton quality cuts used in all previous papers of ours
- 68544 events with an electron jet and 14966 events with a muon jet



Strategy

- Determine the b- and c-quark composition of the data by counting the number of SECVTX, and JPB tags on both the lepton- and away-jets
- Differently from previous analyses, this study checks at the same time the cross section for producing at least 1 b with $|\eta|<1.5$ (imperfect NLO calculation), 1 b +1 $\bar{b}$ with $|\eta|<1.5$ (robust NLO calculation)
- Tune the simulation to reproduce the tagging rates of the data
- This step removes the uncertainty of the theoretical prediction of the heavy flavor cross sections and of the efficiency for finding the trigger lepton
- Then check the semileptonic branching ratio of heavy flavor hadrons by counting the number of a-jets with SLT tags in the data and in the simulation

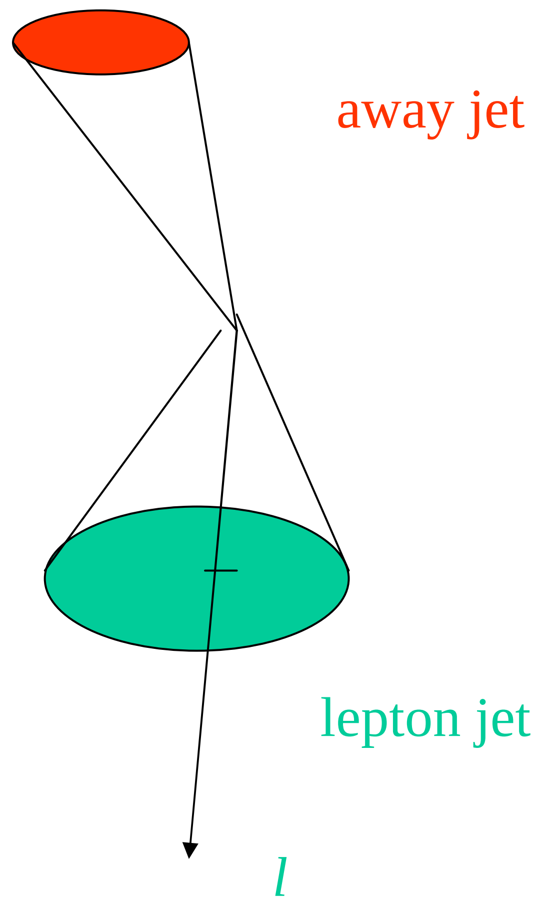
Mistags and tagging efficiencies

- PRD 64, 032002 (2001) and PRD 65, 052007 (2002)
- Mistags (tags in a jet without heavy flavor) are evaluated with parametrized probability functions derived in generic-jet data. We estimate a 10% uncertainty.
- Since we use a parametrized simulation of the detector, we have measured the data-to simulation scale factor for the tagging efficiency of the SECVTX and JPB algorithms.These factors were determined with a 6% accuracy and implemented into the simulation.
- The SLT simulation uses efficiencies for each selection cut measured using data; we estimate a 10% uncertainty, which includes the uncertainty on the semileptonic branching ratio
- The simulated supertag efficiency is corrected for the data-to-simulation scale factor measured in generic-jet data: 6% uncertainty

Evaluation of the heavy flavor content of the data

- Before tagging, approximately 50% of the lepton jets do not contain heavy flavor; they are mostly due to hadrons that mimic the lepton signature.

- The fraction (1-hf) of events in which the l-jet does not contain heavy flavor is not simulated. In these events, away-jets can have tags due to heavy flavor. Their rates are estimated using a parametrized probability of finding a tag due to heavy flavor in generic-jet data. It is a slight overestimate.
- Using a sample of jets containing electrons identified as coming from conversions, which has a 8% heavy-flavor purity, we estimate that the accuracy of the method is better than 10% .

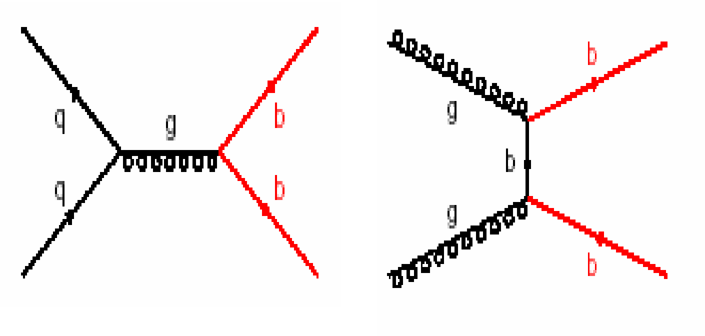


Simulation

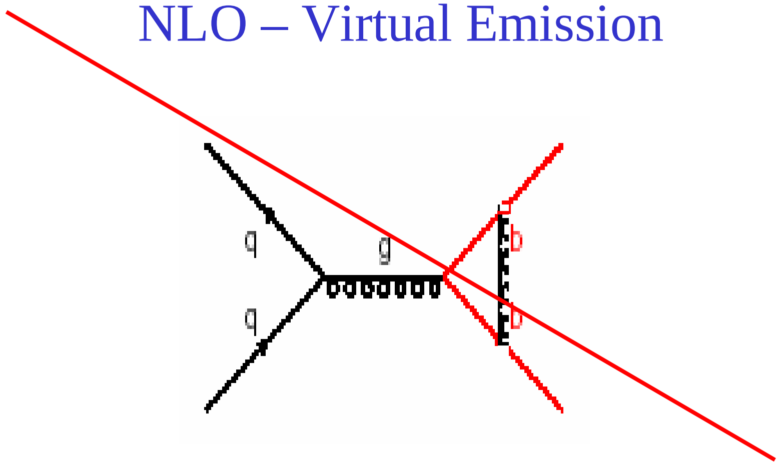
- Use the Herwig generator program (option 1500, generic $2 \rightarrow 2$ hard scattering with $p_T > 13$ GeV/c)
- $b\bar{b}$ and $c\bar{c}$ production are generated through processes of order α^2 such as $q\bar{q} \rightarrow b\bar{b}$
- Processes of order α^3 are implemented through flavor excitation diagrams, such as $g\,b \rightarrow g\,b$, or gluon splitting, in which the process $g\,g \rightarrow g\,g$ is followed by $g \rightarrow b\bar{b}$
- Use MRS (G) PDF’s
- The bottom and charmed hadrons are decayed with QQ (version 9_1)
- We select simulated events which contain hadrons with heavy flavor and at least one lepton with $p_T > 8$ GeV/c
- These events are passed through QFL, a parametrized simulation of the CDF detector and treated as real data
- We have simulated 27156 electron events (98.9 pb^{-1}) and 7267 muon events (55.1 pb^{-1}) with heavy flavor

HERWIG vs exact NLO calculation

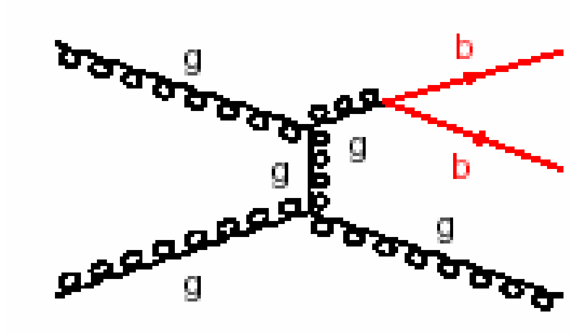
LO – Born term



NLO – Virtual Emission

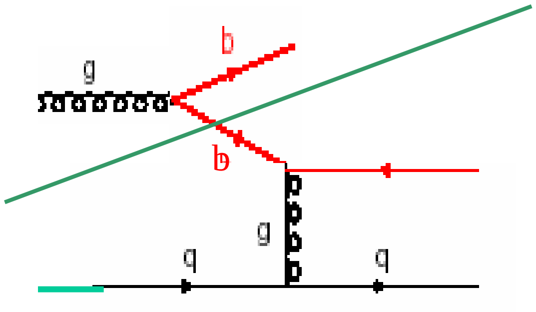


NLO – real emission



Gluon splitting

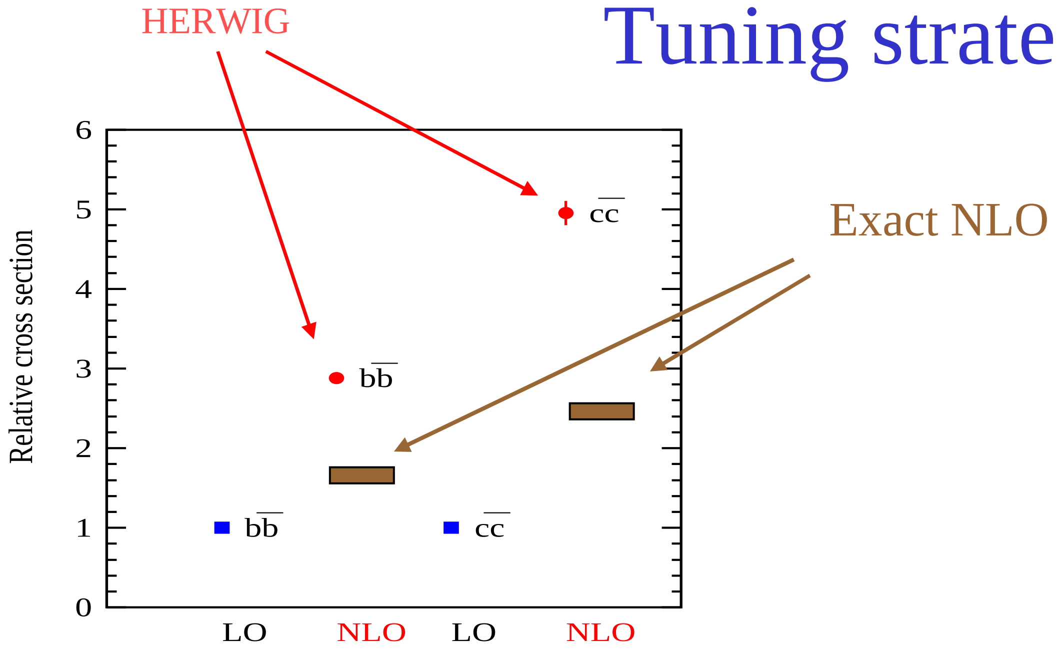
Parton shower

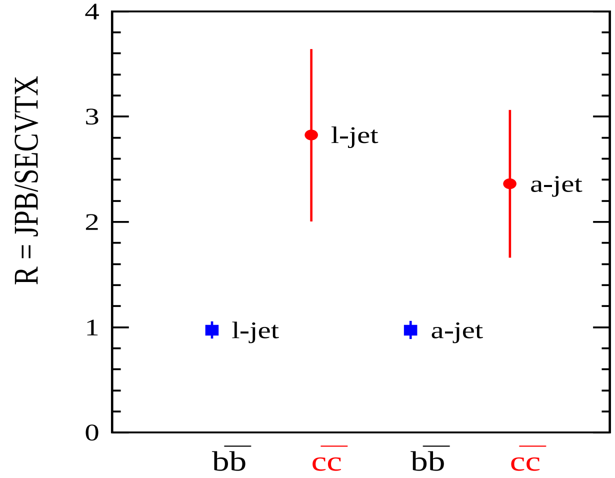


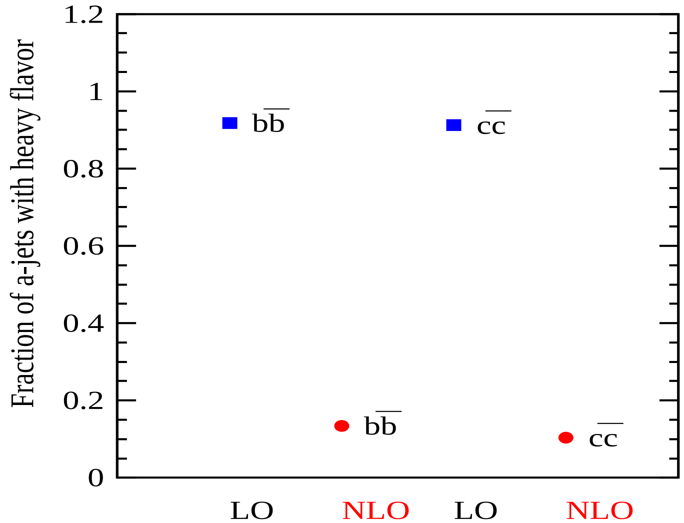
Flavor Excitation

Structure function

Tuning strategy







Fit of the simulation to the data

SECVTX	lepton side	
	away side	
	Both	
JPB	lepton side	
	away side	
	Both	

Fit parameters	Constraints	Error
c dir norm	b dir/c dir ≈ 1	14%
b flav exc norm	b/c ≈0.5	28%
c flav exc norm		
b gluon split norm	1.40	0.19
c gluon split norm	1.35	0.36
Ke norm		
Kμ norm		
SECVTX scale factor, b	1.0	6%
SECVTX scale factor, c	1.0	28%
JPB scale factor	1.0	6%

- Use 6 fit parameters corresponding to the direct, flavor excitation and gluon splitting production cross sections evaluated by Herwig for b- and c-quarks
- K_e and K_μ account for the luminosity and b-direct production
- The parameters bf, bg, c, cf, cg account for the remaining production cross sections, relative to the b-direct production

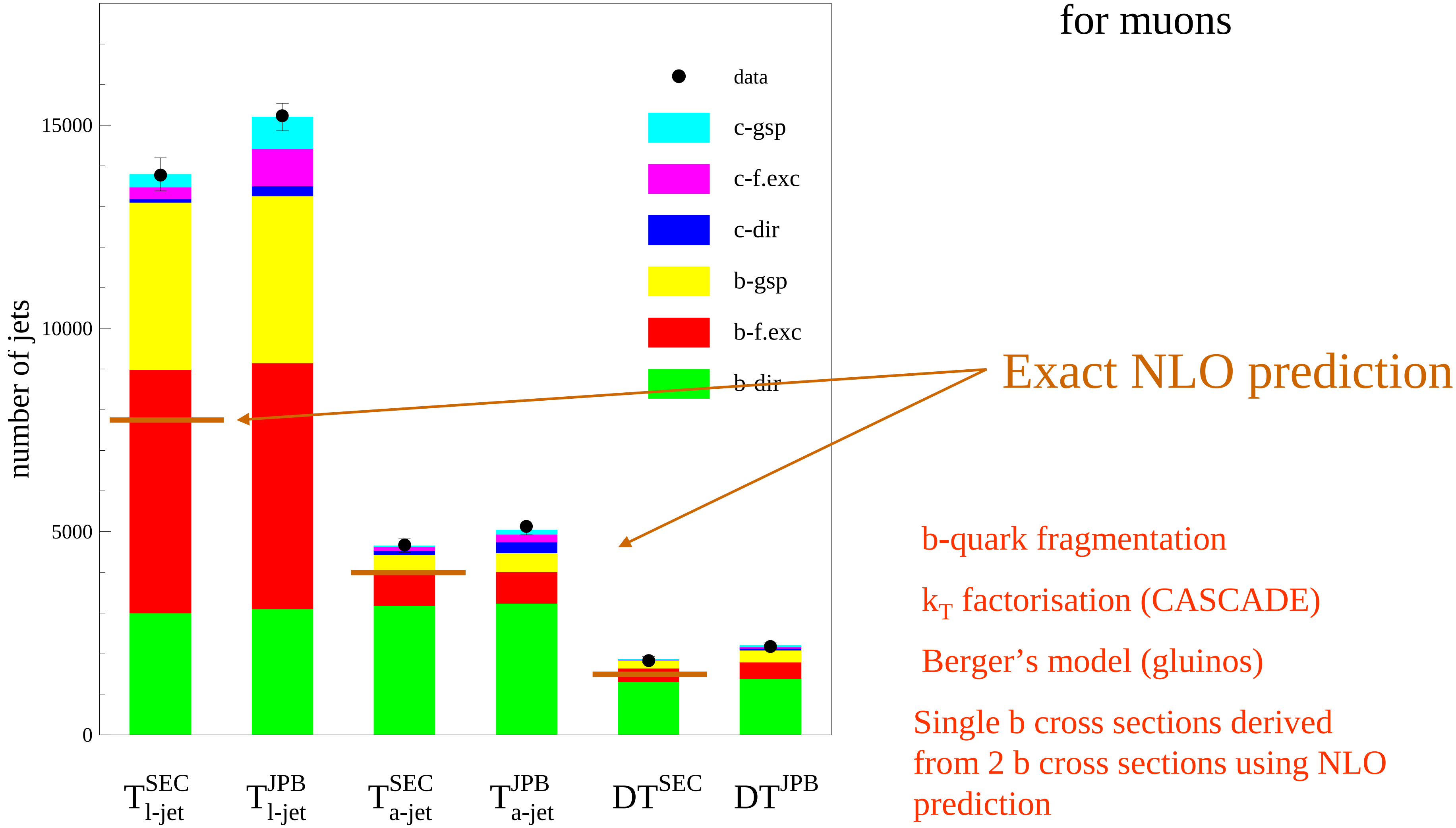
Fit result

SECVTX scale factor	SF_b	0.97 ± 0.03
SECVTX scale factor	SF_c	0.94 ± 0.22
JPB scale factor	SF_{JPB}	1.01 ± 0.02
e norm.	K_e	1.02 ± 0.05
μ norm.	K_μ	1.08 ± 0.06
c dir. prod.	c	1.01 ± 0.10
b flav. exc.	bf	1.02 ± 0.12
c flav. exc.	cf	1.10 ± 0.29
$g \to b\bar{b}$	bg	1.40 ± 0.18
$g \to c\bar{c}$	cg	1.40 ± 0.34

- χ^2/DOF=4.6/9

Tuned HERWIG

- F_{hf} = (45.3±1.9)% for electrons
- F_{hf} = (59.7±3.6)% for muons



b-quark fragmentation

k_T factorisation (CASCADE)

Berger's model (gluinos)

Single b cross sections derived from 2 b cross sections using NLO prediction

LO vs NLO vs mc@NLO vs Herwig

- Single b quark with p_T>18 GeV/c, $|\eta|<5$ (inclusive σ in nb)

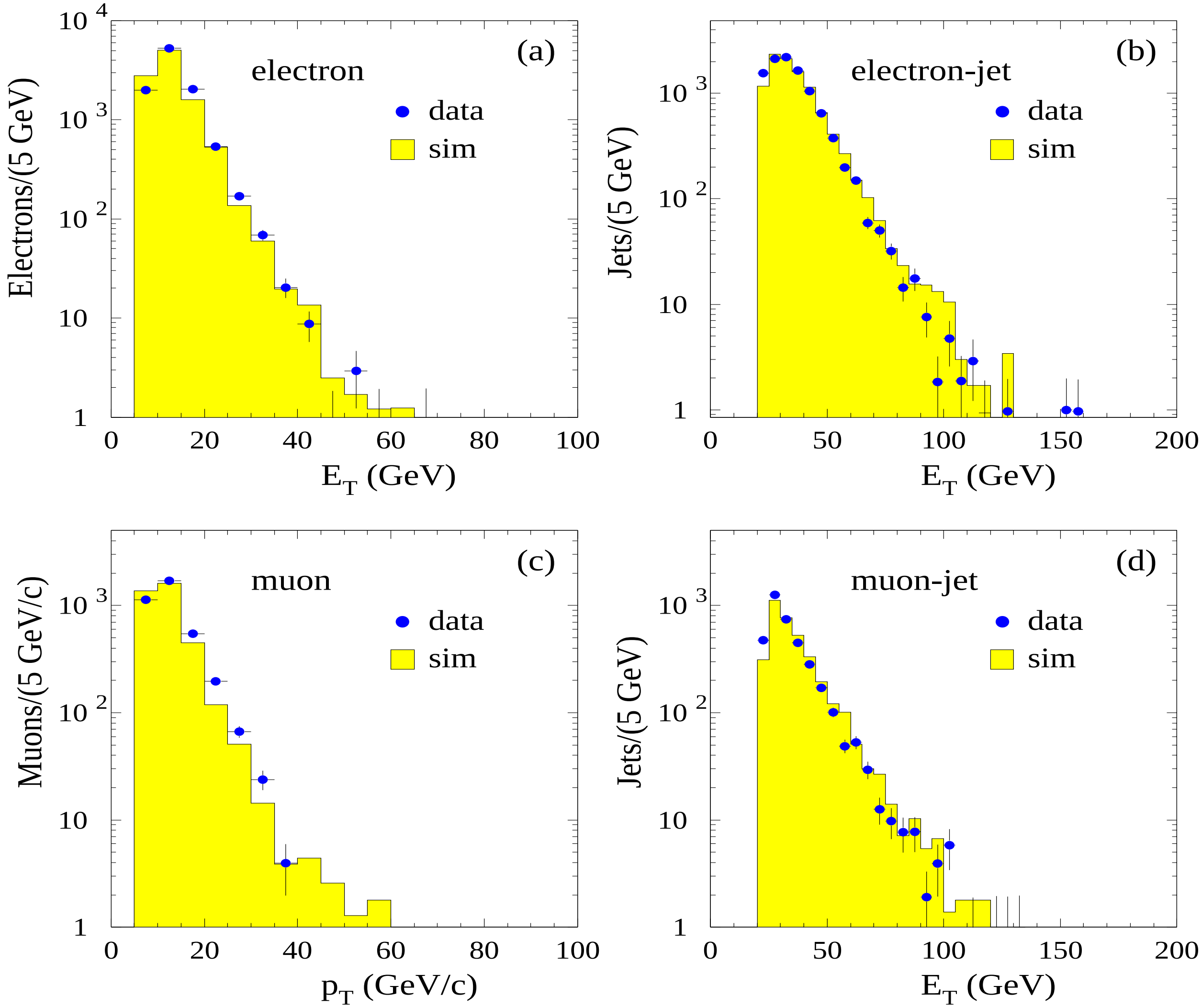
LO	NLO	mc@NLO	Herwig
118	300	430	483

- Single b quark with p_T>18 GeV/c, $|\eta|<1.5$ (inclusive σ in nb)

LO	NLO	mc@NLO	Herwig
100	207	204	383

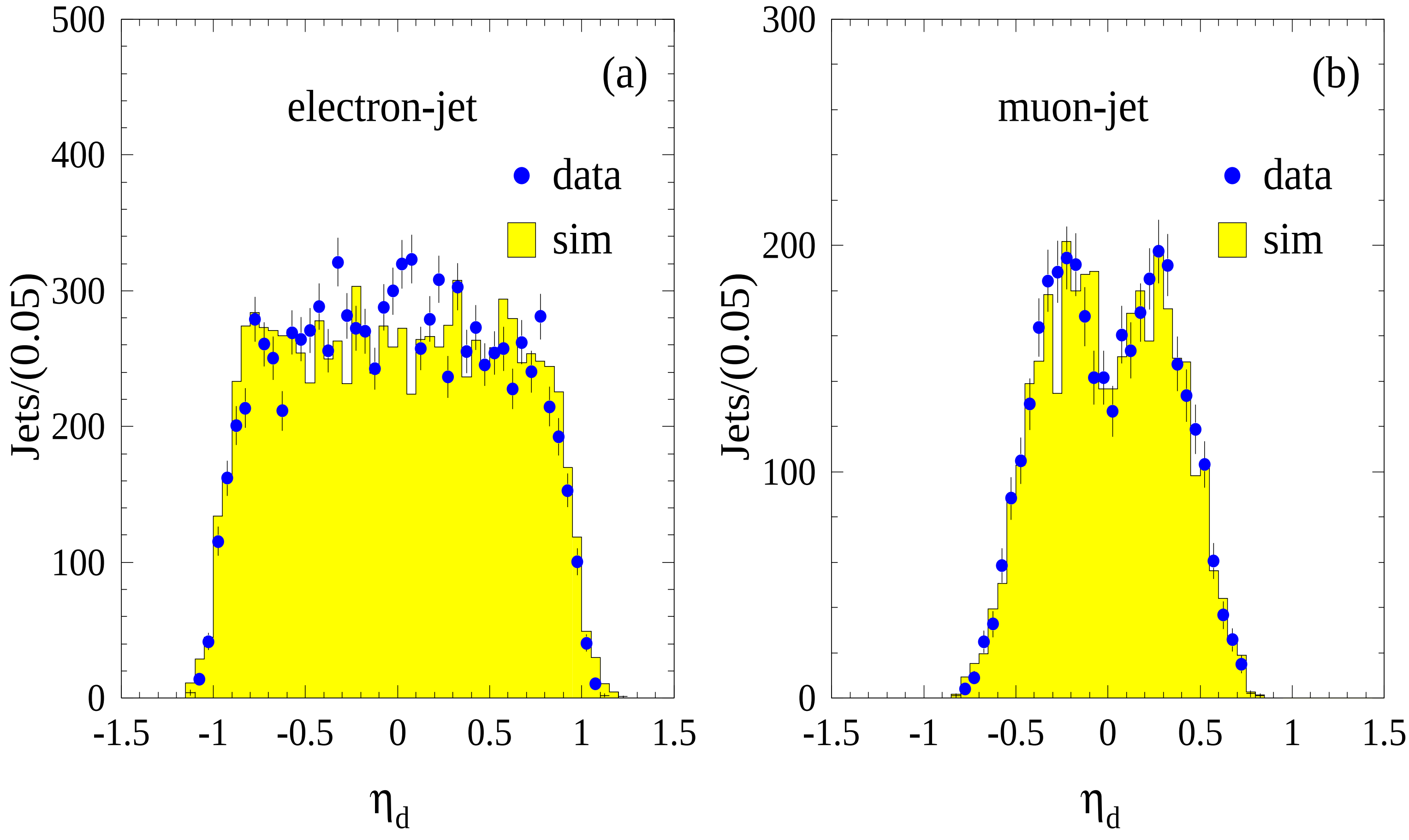
Kinematics

SECVTX tagged



Kinematics

SECVTX tagged

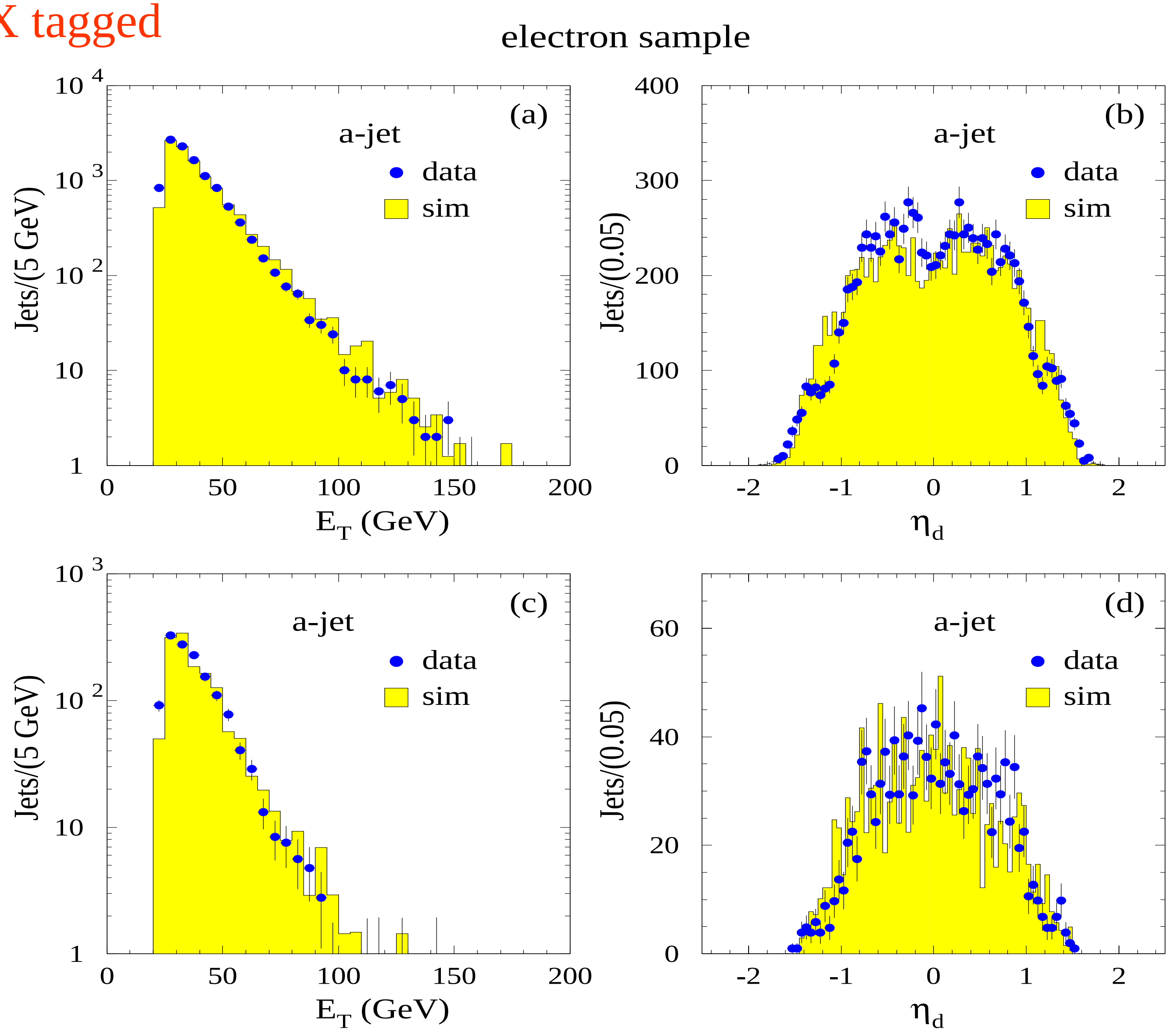


Kinematics

L -jet SECVTX tagged

A-jet

A-jet with SECVTX tags

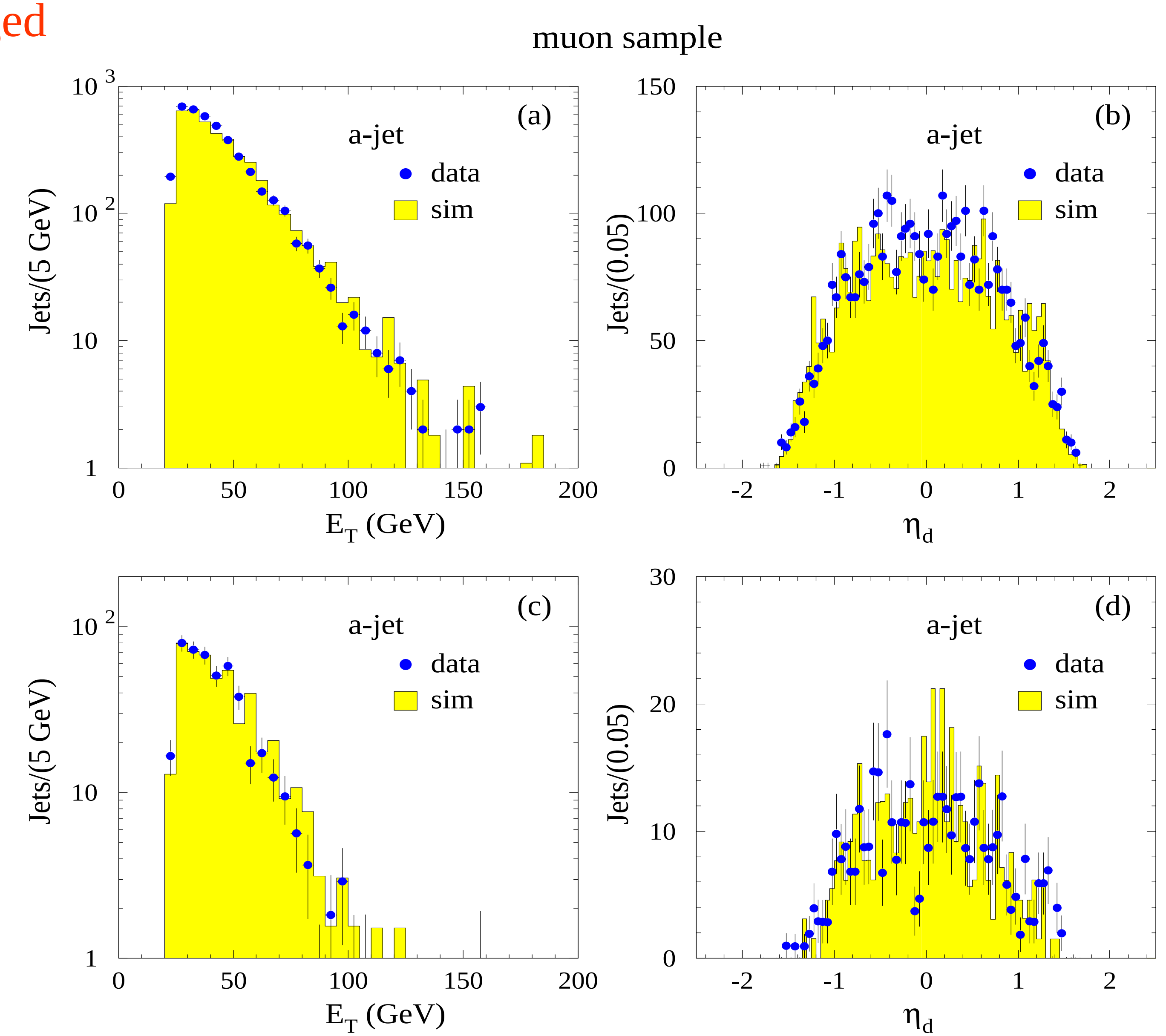


Kinematics

L -jet SECVTX tagged

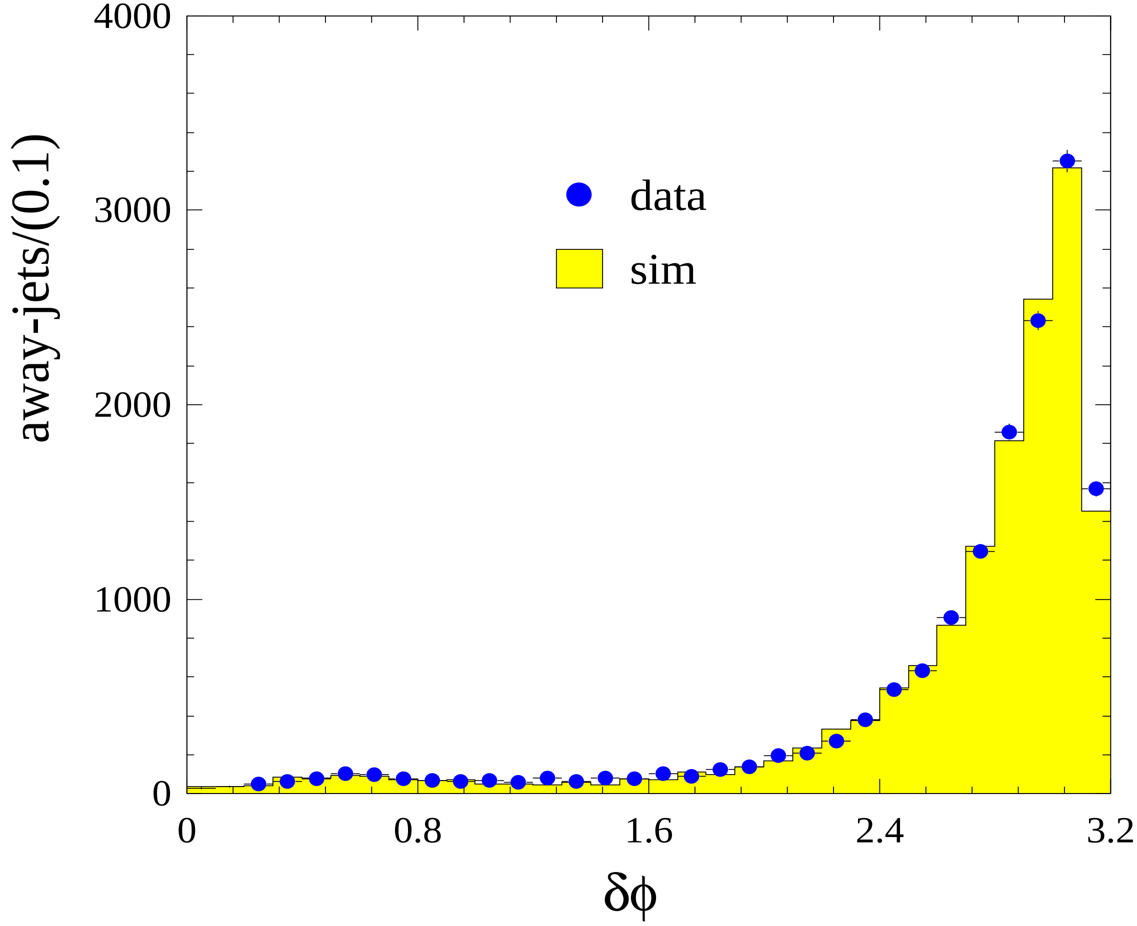
A-jet

A-jet with SECVTX tags



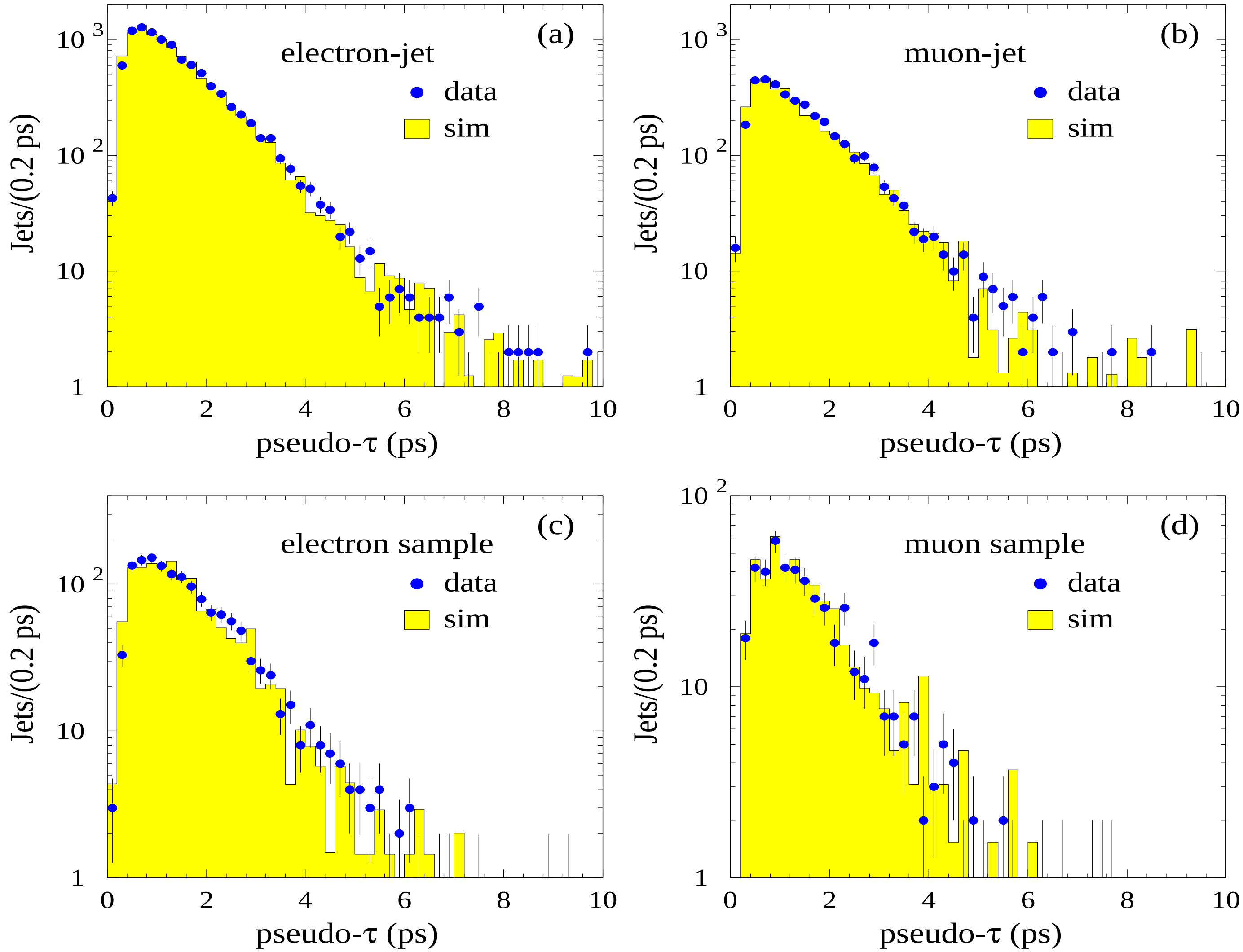
Kinematics

L -jet SECVTX tagged



Kinematics

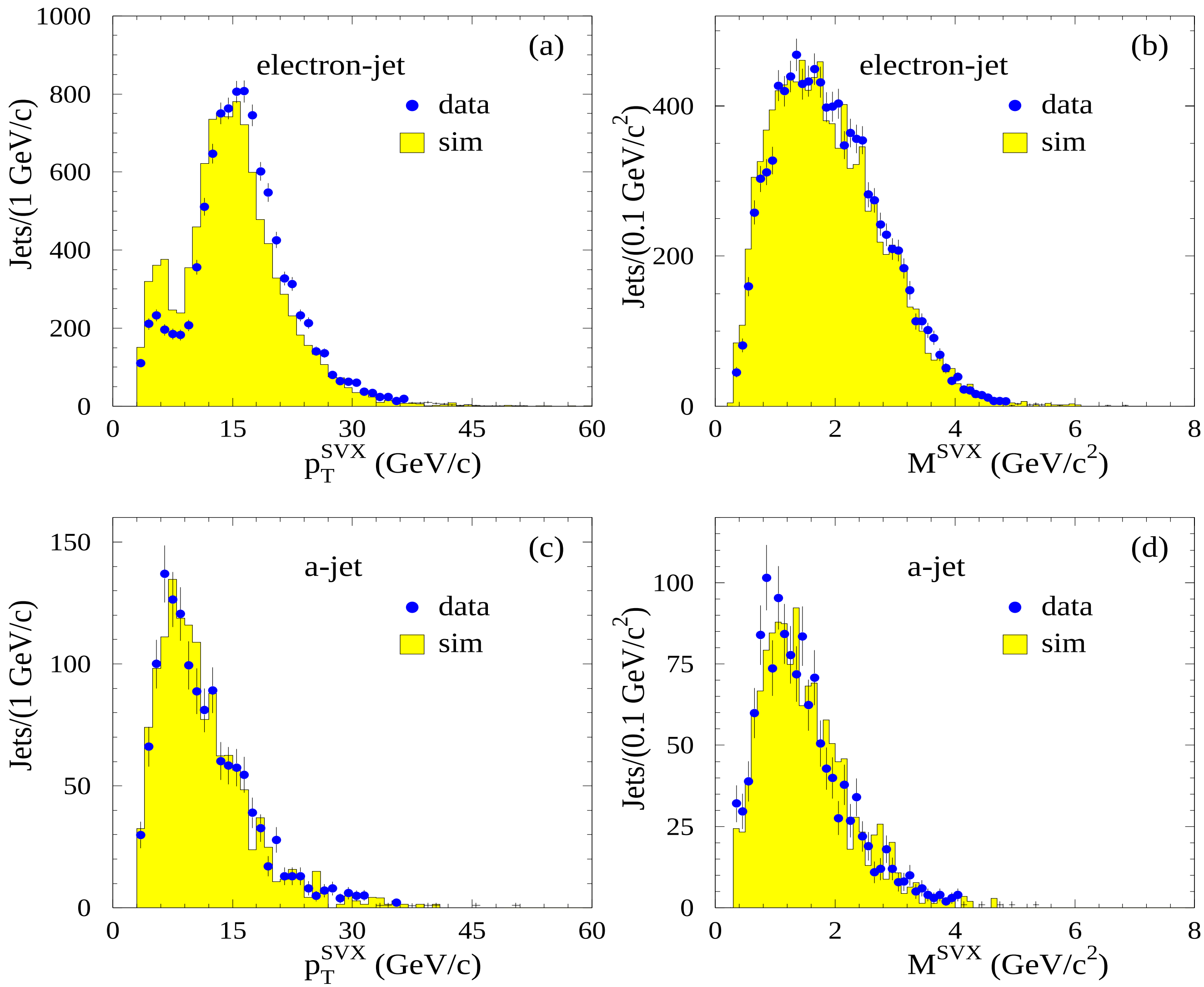
L -jet SECVTX tagged



A-jet with SECVTX tags

Kinematics

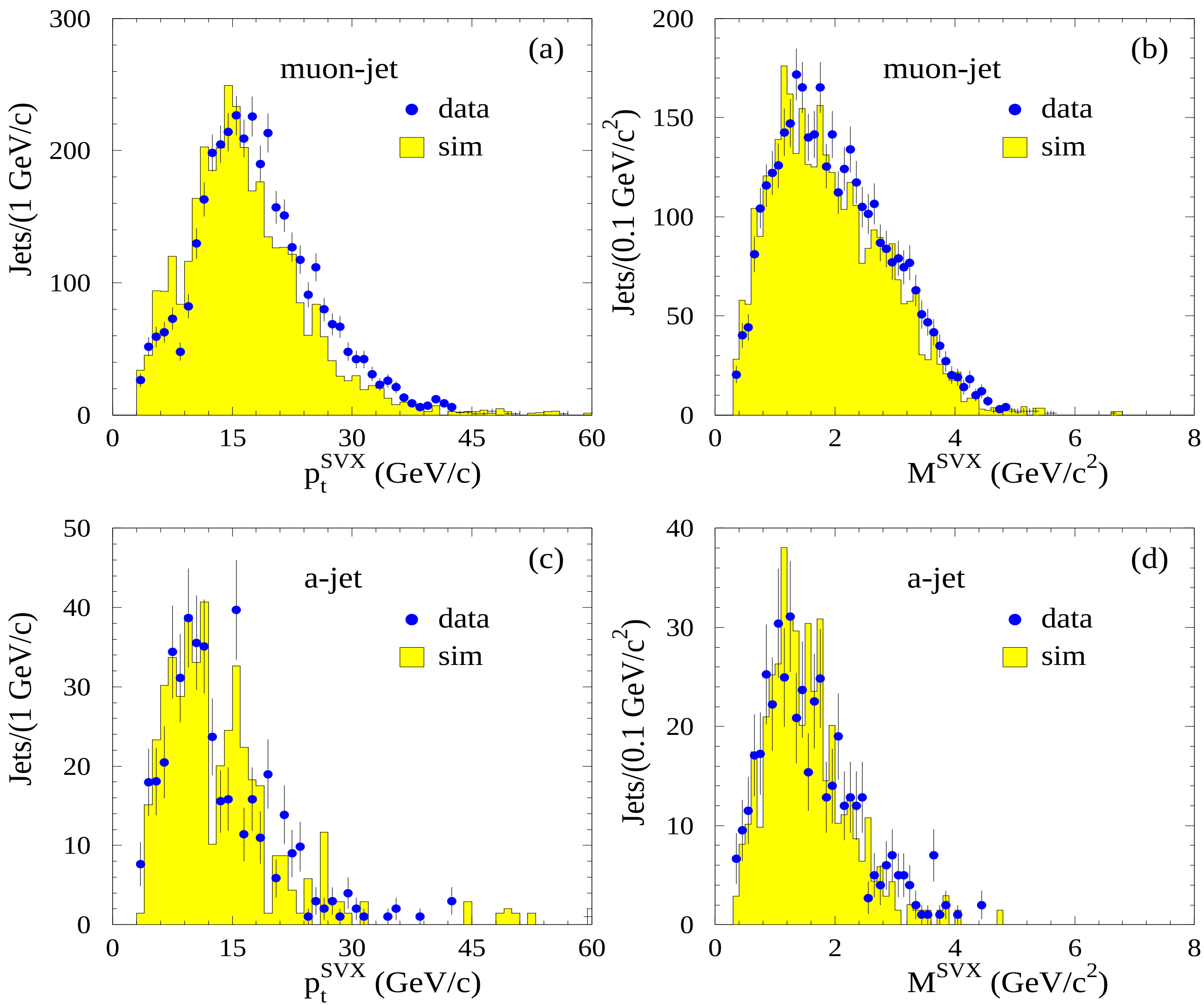
L -jet SECVTX tagged



A-jet with SECVTX tags

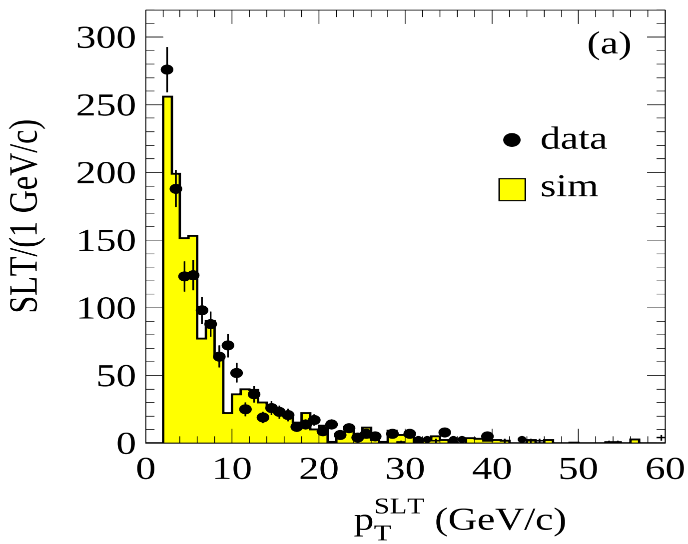
Kinematics

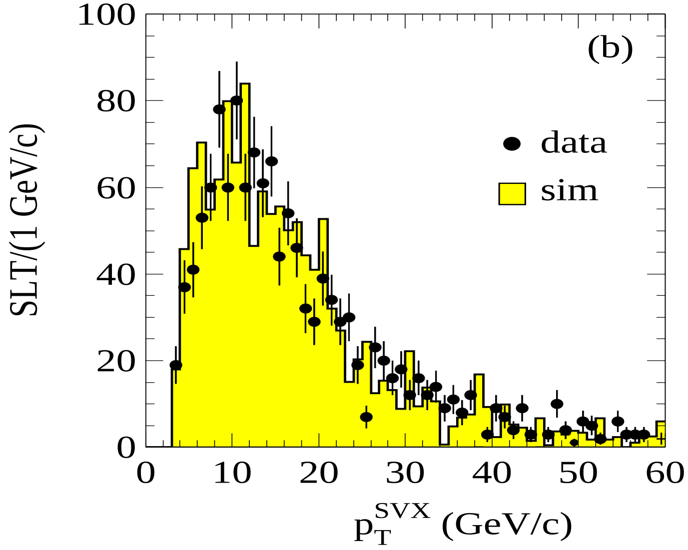
L -jet SECVTX tagged



A-jet with
SECVTX
tags

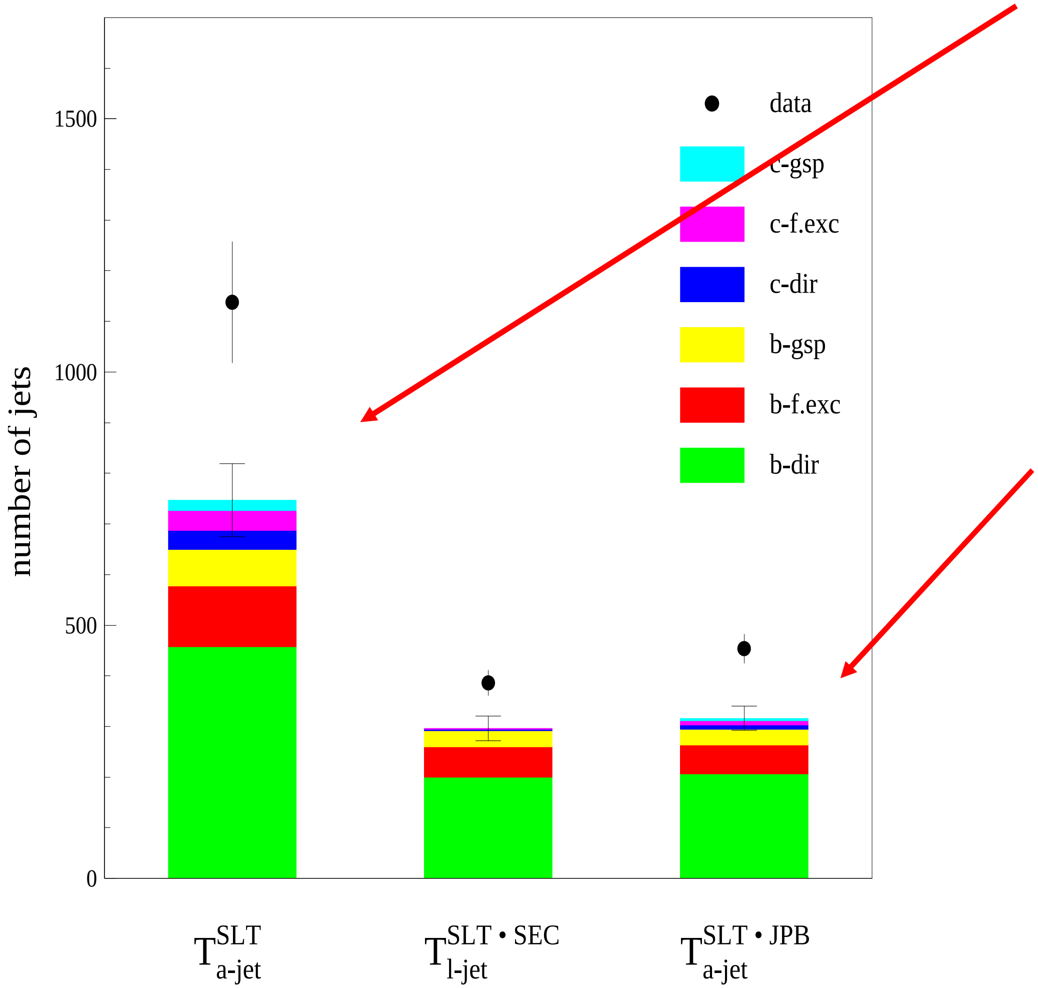
fragmentation in generic-jet data





- 550,000 generic-jet events in the data and in the Herwig simulation (JET20, JET50, and JET100).
  - ✓ 1324 supertags in the data
  - ✓ 1342 simulated supertags

Comparison of a-jets with SLT tags in the data and the tuned simulation



SEEN 1137±140.0
(±51.0 STAT.)

EXPECTED 746.9±75.0 (SYST)

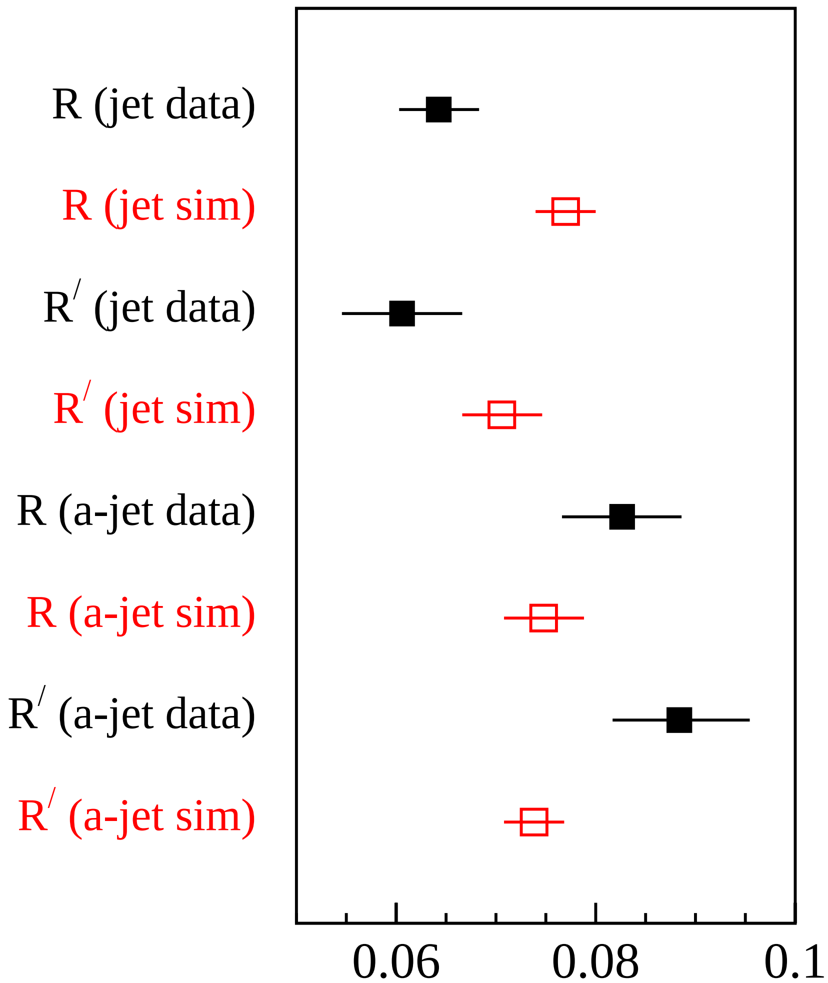
SEEN 453±29.4 (±25 STAT.)

EXPECTED 316.5±25.4 (SYST)

(±15.8 SLT efficiency, ±20 fit)

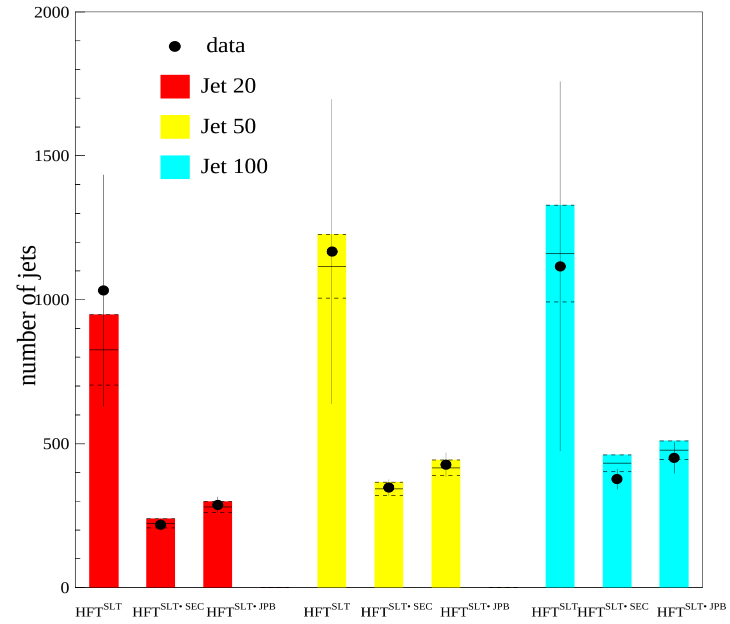
Supertags

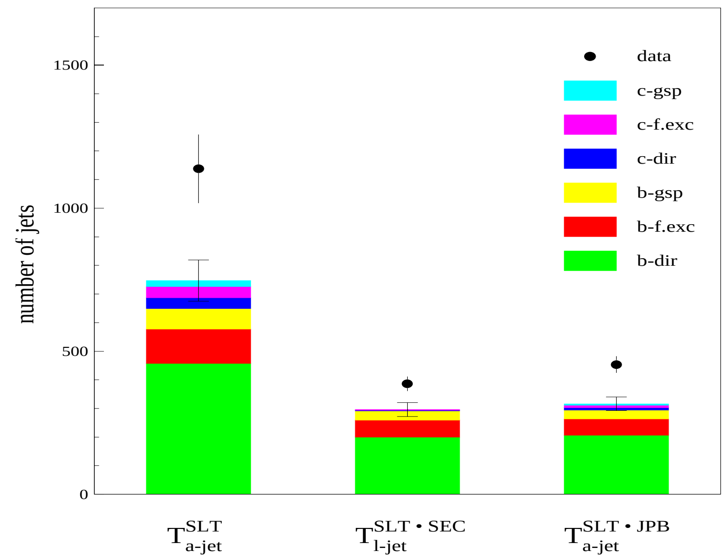
USE JET DATA TO CONSTRAIN THE SYSTEMATIC ERROR



Systematics (SLT tags)

- The discrepancy between expected and predicted rates of away-jet with SLT tags is approximately a 2.5 σ systematic effect due to a 10% uncertainty on the mistag removal and a 10% uncertainty on the tagging efficiency.
- These two systematic uncertainties are quite conservative and had been estimated using other data samples
- One can use generic-jet data to reduce these uncertainties





Uncertainty of (fake + h.f.) SLT tags

- Fit observed rates of SLT tags in generic jets as
  observed tags = P_f x predicted fakes + P_{hf} x predicted h.f.
- The fit returns P_f =1.017±0.013 and P_{hf} =0.981±0.045, ρ = -0.77
- Using this result the SLT expectation in away-jets is 1362±28 whereas 1757±104 are observed
- This discrepancy cannot come from obvious prediction deficiencies

	observed	pred. fakes.	pred. h.f.	Fitted sum
SLT's in g. jets	18885 ±137	15570±1557	3102 ±403	18878 ±130
SLT's in g. jets with SECVTX	1451 ± 38	999 ±60	508 ±51	1514 ±38
SLT's in g. jets with JPB	2023 ±45	856 ±86	1175 ±71	2023 ±45
SLT's in a-jets (lep-trig.)	1757 ±104	619 ±62	747 ± 75	1326 ±28

Conclusions

- We have measured the heavy flavor content of the inclusive lepton sample by comparing rates of SECVTX and JPB tags in the data and the simulation
- We find good agreement between the data and the simulation tuned within the experimental and theoretical uncertainties
- We find a 50% excess of a-jets with SLT tags due to heavy flavor with respect to the simulation; the discrepancy is a 3 σ systematic effect due to the uncertainty of the SLT efficiency and background subtraction. However, comparisons of analogous tagging rates in generic-jet data and their simulation do not support any increase of the efficiency or background subtraction beyond the quoted systematic uncertainties

Conclusions

- A discrepancy of this kind and size is expected, and was the motivation for this study, if pairs of light scalar quarks with a 100% semileptonic branching ratio were produced at the Tevatron
- The data cannot exclude alternate explanations for this discrepancy
- Previously published measurements support the possibility, born out of the present work, that approximately 30% of the presumed semileptonic decays of heavy flavor hadrons produced at the Tevatron are due to unconventional sources

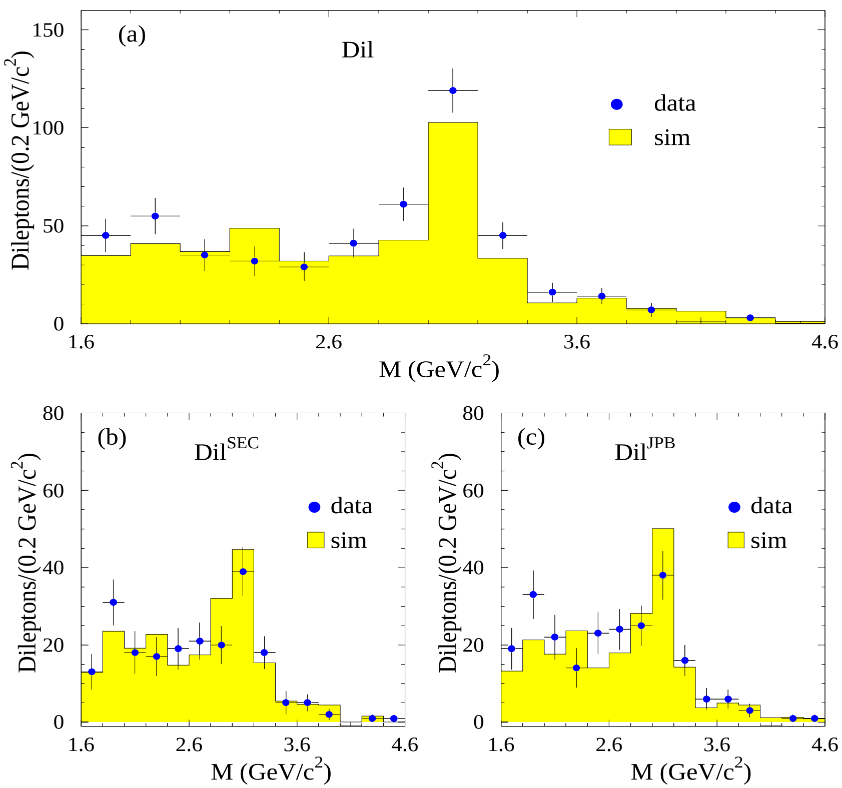
The end

backup slides follow

b-purity (cross-check)

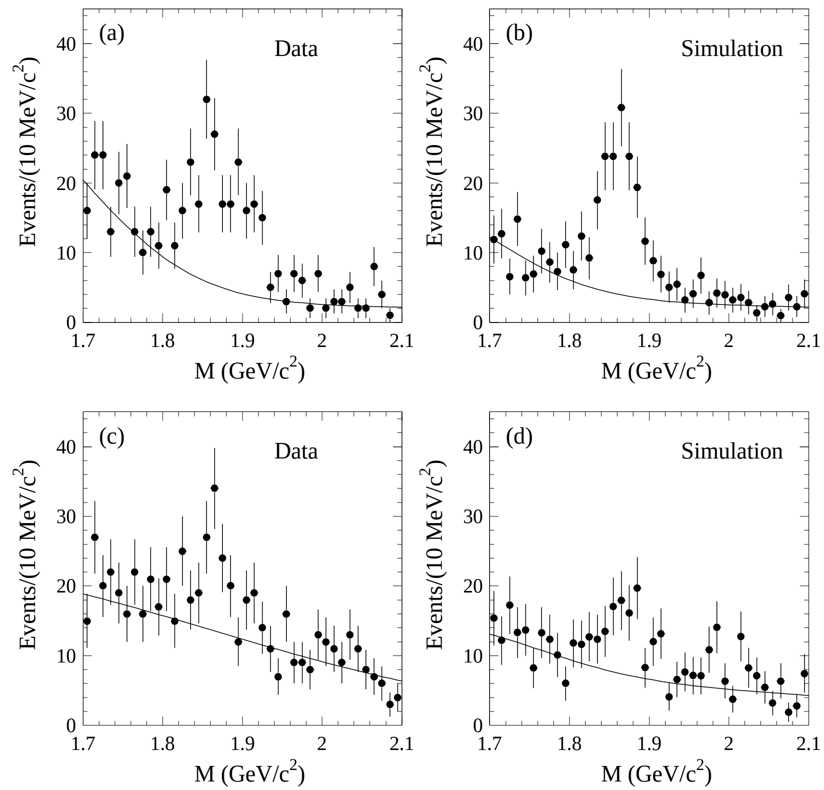
- D^0 : 126.0 ± 15.5 in the data and 139.9 ± 15.0 in the simulation
- $D^{\pm}$: 73.7 ± 17.8 and 68.5 ± 14.1
- J/ψ: 90.8 ± 10.1 and 101.9 ± 11.4
- Ratio of the b-purity in the simulation to that in the data is 1.09 ± 0.11

J/ψ



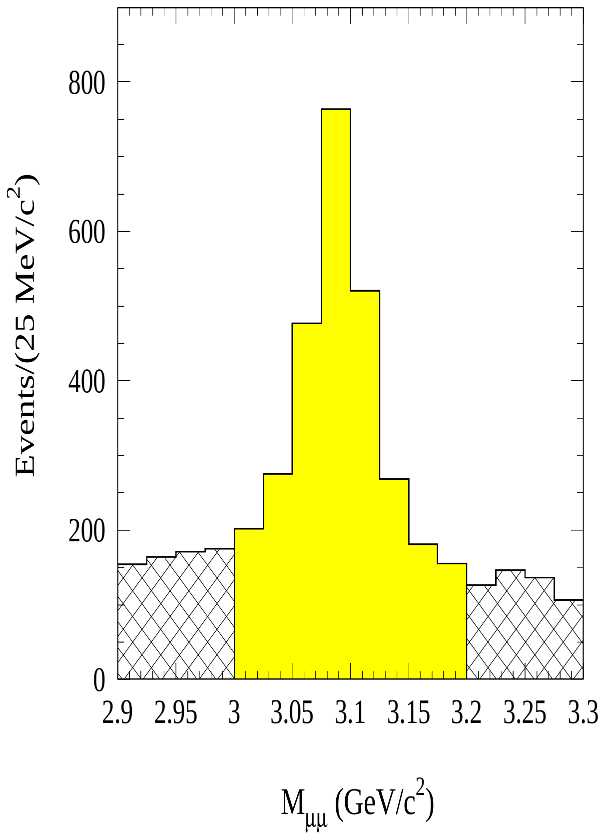
D⁰

D±



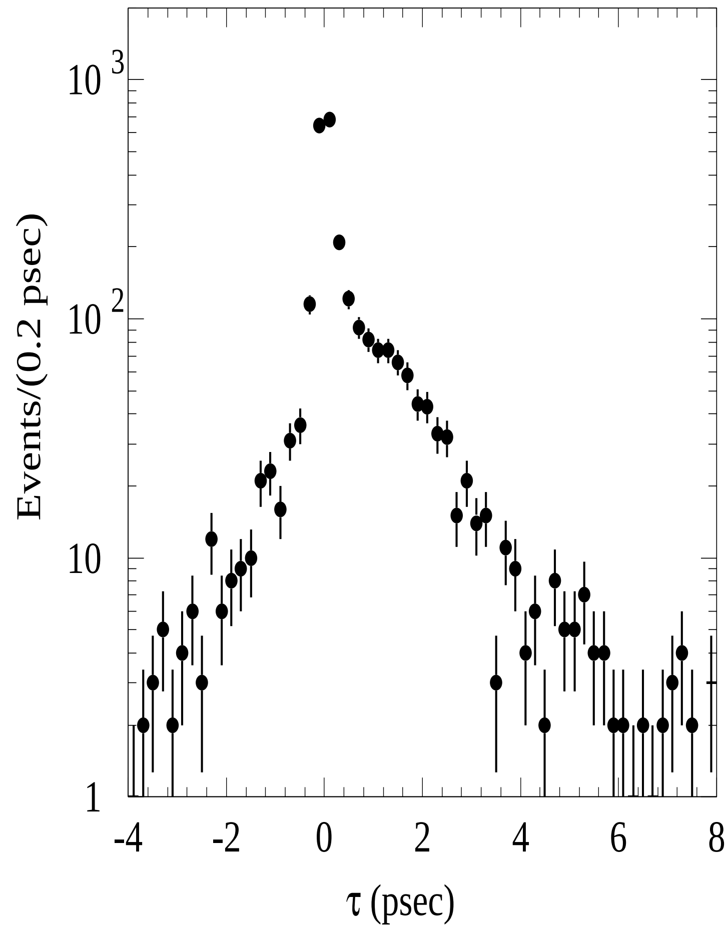
Cross check with J/ψ mesons from B-decays

- In generic-jet data we do not have any excess of jets with SLT tags or supertags
- We do observe an excess after enriching the b-purity of the QCD data by requiring a lepton-jet
- We study a sample of jets recoiling J/ψ mesons from B-decays. We use the same $J/\psi \to \mu\mu$ data set and selection used for the measurement of the J/ψ lifetime and fraction from B-decays
- 1163 J/ψ over a background of 1179 events estimated from the side-bands (SB)



J/ψ lifetime

- The number of J/ψ mesons from B-decays is $N_\psi = (\psi^+ - \psi^-) - (SB^+ - SB^-)$ =561, which is 48% of the initial sample
- In the 572 away-jets we find 48.0 ± 15.1SECVTX, 61.7 ± 17.3 JPB tags, and –9.4 ± 14.4 SLT tags
- In the simulation we expect 8.1 ± 1.1 SLT tags



No dependence of the result on the heavy flavor tuning

- SLT tags: observed 1137±140 and 747 ± 75 predicted by the tuned simulation
- N_b and N_c be the numbers of a-jets with bottom and charmed flavor
- ε_b^{JPB}=0.43, ε_c^{JPB}=0.30, ε_b^{SLT}=0.064, ε_c^{SLT}=0.047
- $\varepsilon_c^{JPB} / \varepsilon_b^{JPB} = \varepsilon_c^{SLT} / \varepsilon_b^{SLT}$ (finesse)
- HFT^{SLT}(a-jet)= $\varepsilon_b^{SLT} (N_b + \varepsilon_c^{SLT} / \varepsilon_b^{SLT} N_c) \varepsilon_b^{JPB} / \varepsilon_b^{JPB}$ = $\varepsilon_b^{SLT} / \varepsilon_b^{JPB}$ HFT^{JPB}(data) = $\varepsilon_b^{SLT} / \varepsilon_b^{JPB}$ (5127±147) =763±80

Systematics (away-jets with SLT tags)

- In events due to heavy flavor, there is an excess of 391 a-jets with a SLT tag with respect to the simulation (1137.8 observed and 746.9 expected), having removed 619.3 fake tags [the events in which the l-jet does not have heavy flavor contain 901.9±91 a-jet with SLT tags (74% fake+ 26% heavy flavor): slight overestimate].
- If one could increase the fake rate in events with heavy flavor by 60%, the excess would disappear. However, in generic-jet data, the fake rate is 74% of the SLT tagging rate.
- The 10% uncertainty of the fake removal is due to the method used to estimate the 26% content of heavy flavor: track impact parameter distribution (very hard to simulate correctly).

Systematics (fake SLT tags)

- The heavy flavor content of generic-jet data has been evaluated using SECVTX and JPB tags
- In generic-jet data the number of SLT tags due to heavy flavor is therefore known with a 13% error, mostly due to the 10% uncertainty of the SLT tagging efficiency
- Therefore the real uncertainty on the fake rate is no larger than 2.6%

Fakes SLT=Data – simulated H.F.=15783±423
Parametrized fake SLT =15570

Systematics (SLT efficiency)

- Away-jets in the inclusive lepton have a higher heavy flavor content (26%) than generic-jet data (13%) .
- Could the fake rate in jets with heavy flavor be anomalously large ? Could the SLT efficiency or the semileptonic branching ratio in the simulation be grossly wrong ?
- Jets with SECVTX or JPB tags in generic-jet data have a heavy flavor content ranging from 86% (JET 20) to 71% (JET 100) . The rate of SLT tags in these jets is not higher than in the simulation