Anomalies in Heavy Flavor Jets at CDF

G.Apollinari-Fermilab

For the CDF Collaboration

Les Rencontres de Physique de la Vallée d'Aoste

LaThuile

March 5th,2004

Anomaly in W+2,3 jet Events at CDF PRD 65, 052007(2002)

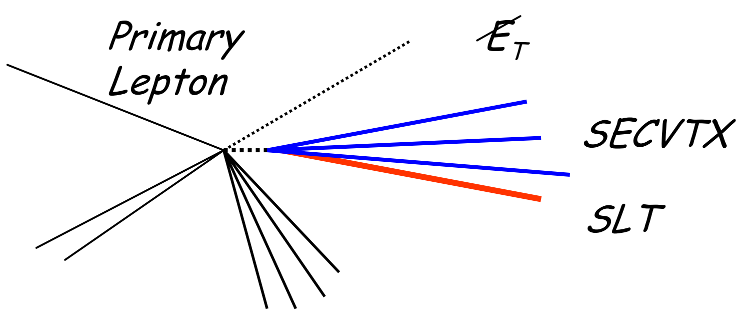
- CDF Data sample used in top quark measurements

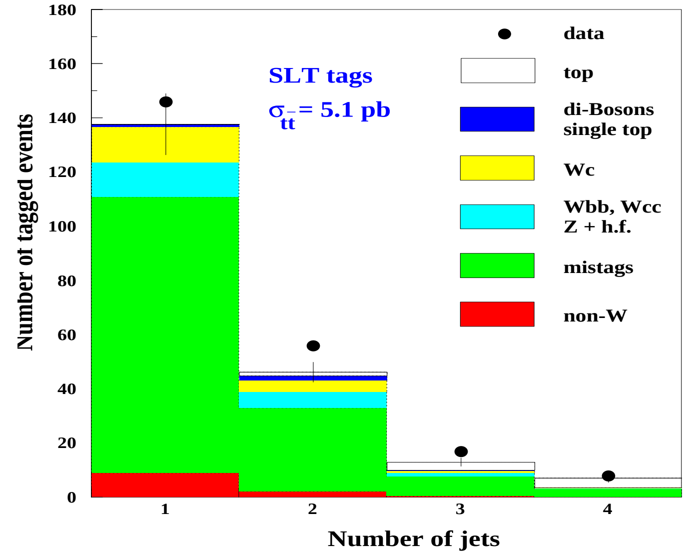
$$p\bar{p} \rightarrow t\bar{t} + X \rightarrow WbW\bar{b} \rightarrow l\nu + 3{,}4 \;\; jets$$

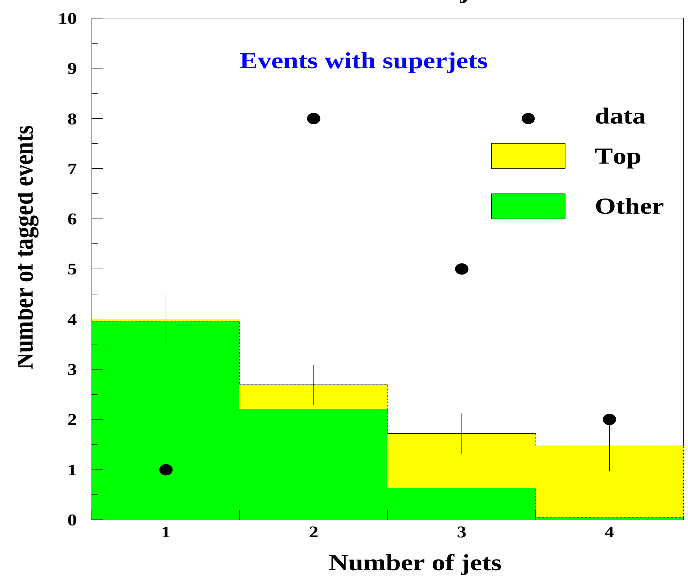
- Heavy Flavor Ident. (tagging) methods

	b	c
SECVTX	43%	9%
JPB	43%	30%
Soft Lepton Tagging	6.4%	4.6%

- Supertag (or superjet): jet containing both a SECVTX and an SLT tag.



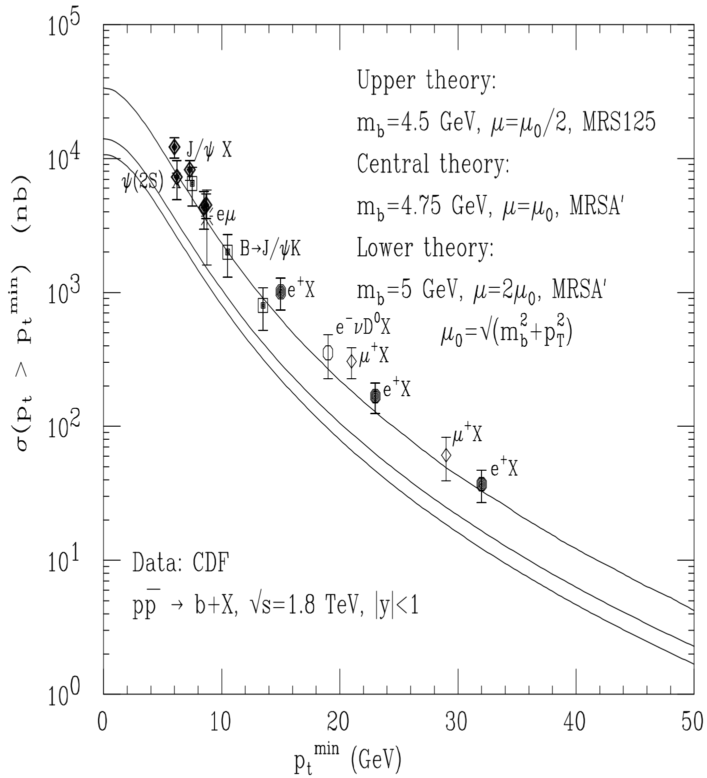


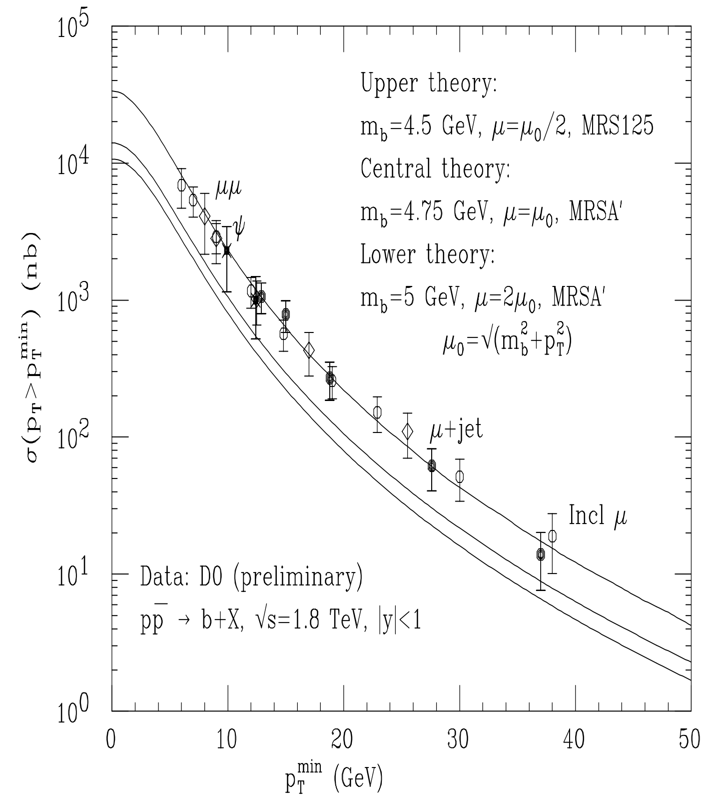


- *The kinematic of the anomalous W+2,3 jets events has a 10^{-6} probability of being consistent with the SM simulation - PRD 64, 032004 (2002)*
- *Superjets modeled by postulating a low mass, strong interacting object which decays with a semileptonic branching ratio of ~ 1 and a lifetime of ~1 ps - hep-ph/0109020*
- *No limit on the existence of a charge –1/3 scalar quark with mass smaller than 7 GeV/c^2 (the supersymmetric partner of the bottom quark, b_s, is a potential candidate) - PRL 86, 4463 (2001).*
- *This analysis is intended to search for evidence either supporting or disfavoring this hypothesis.*

- *hep-ph/0007318 and hep-ph/0401034 use it to resolve the discrepancy between the measured and predicted values of R for $5 < \sqrt{s} < 10$ GeV and for $20 < \sqrt{s} < 209$ GeV at $e^+ e^-$ colliders*

- *If light b_s existed, Run I has produced 10^9 pairs; why didn't we see them ?*



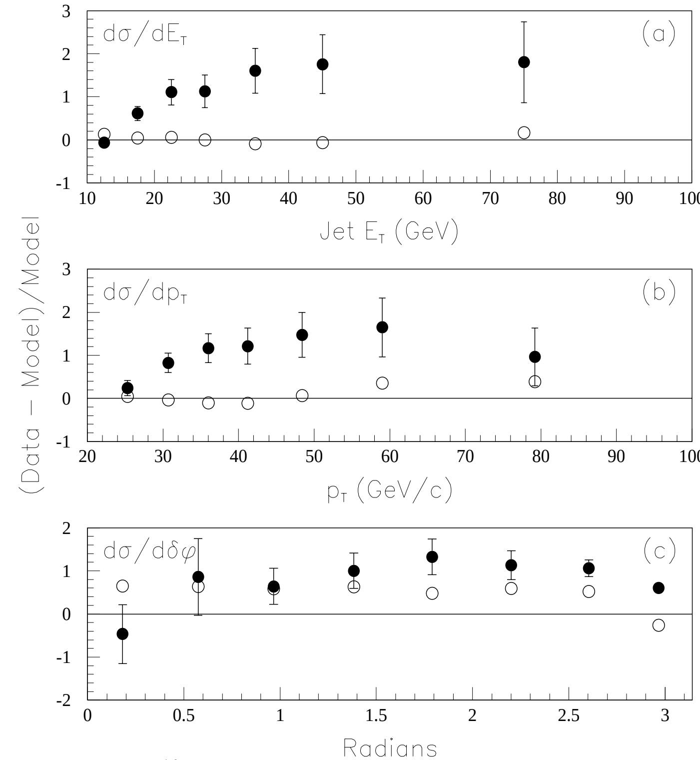


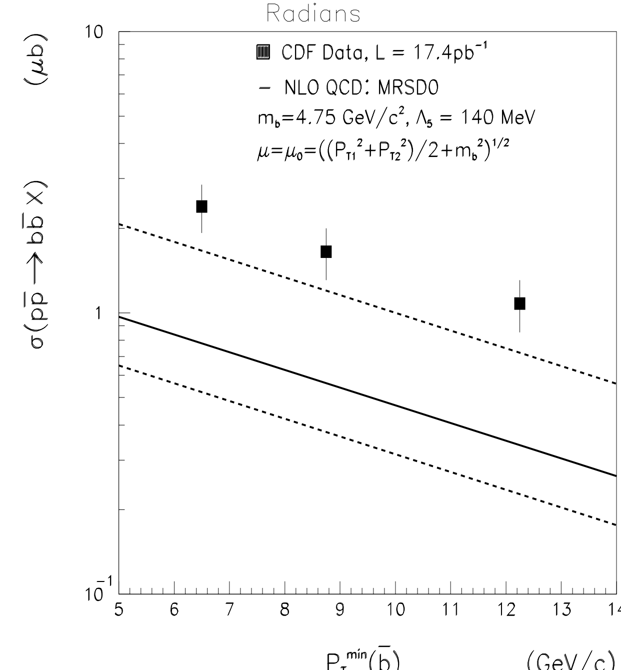
- *PRL 86, 4231 (2001) uses it in conjunction with a light gluino which decays to $b\ b_s$ to explain the difference of a factor of 2 between the measured single-b production cross section and the NLO prediction.*
- *Necessary but not sufficient condition*
    - *NLO not robust*

However….

- *Some interesting CDF & DØ disagreements between Data and Simulation:*
  - *b+μ Production Cross-Section: $\sigma_{b\bar{b}} \cdot BR$*
    - *Data are 1.5 times larger than NLO calculation, LO and NLO calculations are comparable*
    - *PRD 53, 1051 (1996)*
  - *$b\bar{b} \rightarrow \mu^+\mu^-$ Correlations: $\sigma_{b\bar{b}} \cdot BR^2$*
    - *Data are 2.2 times larger than NLO calculation, LO and NLO calculations are within a few percent*
    - *PRD 55, 2547 (1997)*
    - *Phys.Lett. B 487, 264 (2000)*

- *Hint: Data-Simulation discrepancy could increase with the number of leptons in the final state*
  - *Other necessary but not sufficient condition*

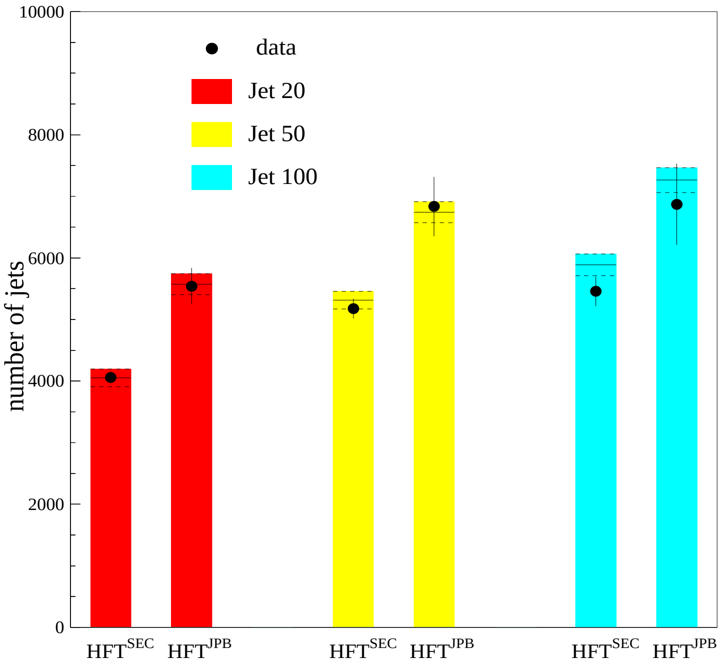




Situation

- The NLO calculation of $p\bar{p} \rightarrow b_s\bar{b}_s$ predicts σ_{bs} = 19.2 μb for a squark mass of 3.6 GeV/c^2 (Prospino MC generator program).
  - σ_{bb} = 48.1 μb (NLO)
  - σ_{cc} = 2748.5 μb (NLO)
- We have used a generic jets data sample with E_T>15 GeV and |η|<1.5 (corresponding to partons with E_T larger than 18 GeV) to calibrate the simulation by using measured rates of SECVTX and JPB.
- Can easily "bend" any Heavy Flavor generator or NLO calculation to explain in terms of SM processes an additional 10% production of scalar quarks

PRD 64, 032002 (2001)



σ_{bs} = 84 nb (Prospino MC)

σ_{bb} = 298 nb (NLO)

σ_{cc} = 487 nb (NLO)

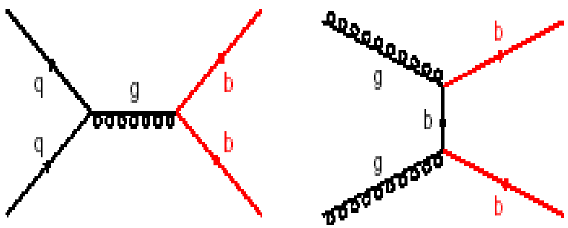
Strategy

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

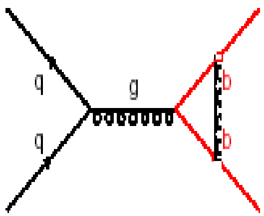
The Control Sample is used to calibrate the SLT efficiency in the simulation and a comparison between the S.S. and the C.S. could have a discrepancy of ~30%.

Models to predict Heavy Flavor Production
HERWIG vs Exact NLO Calculation

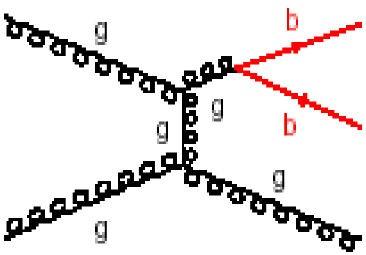
LO – Born term



NLO – Virtual Emission

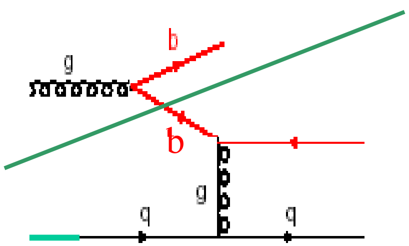


Scattering with 2 b-partons in the final state



Gluon splitting

Parton shower



Flavor Excitation

Structure function

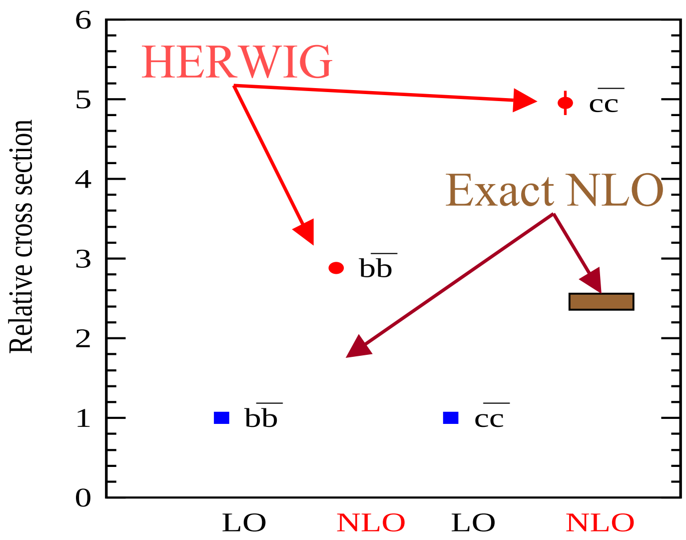
Scattering produces a gluon recoiling against 1 or 2 b-hadrons in the final state

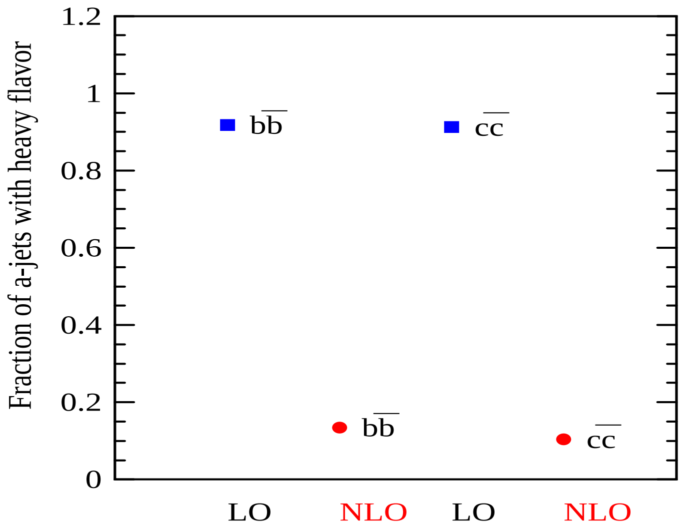
HERWIG vs Exact NLO Calculation

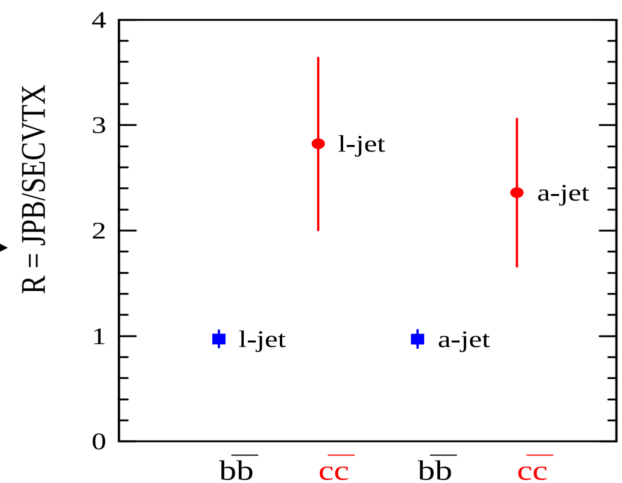
NLO/LO terms can be different for different models (NLO/LO~4 for HERWIG, NLO/LO~2 for NLO Calc.).

Fraction of away h.f. jets in detector acceptance is different for LO vs. NLO terms

Use tools to disentangle $b\bar{b}$ from $c\bar{c}$ production

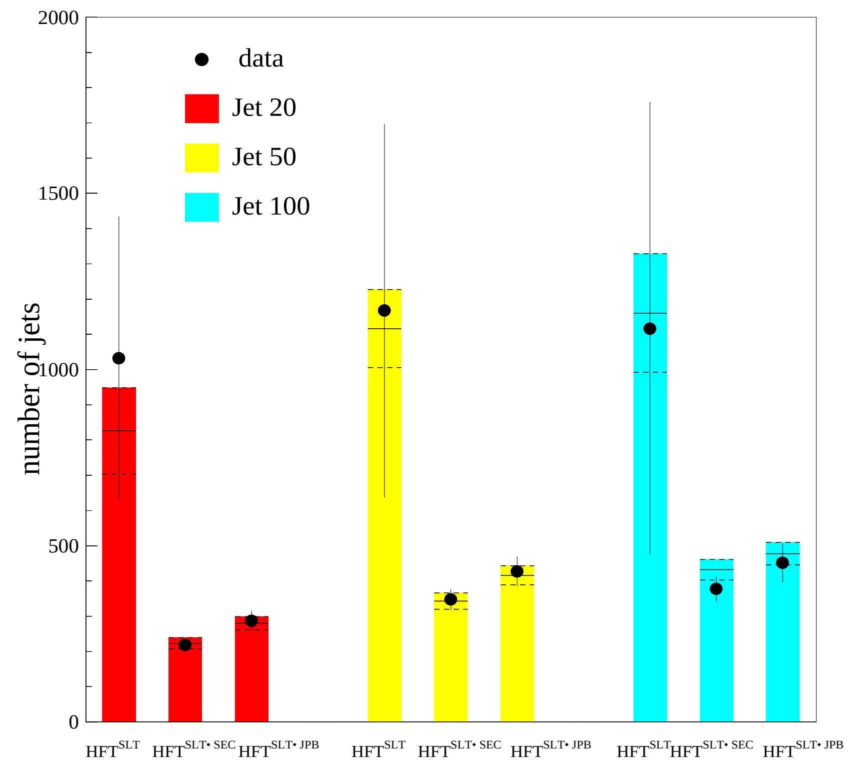






Generic Jet Control Sample

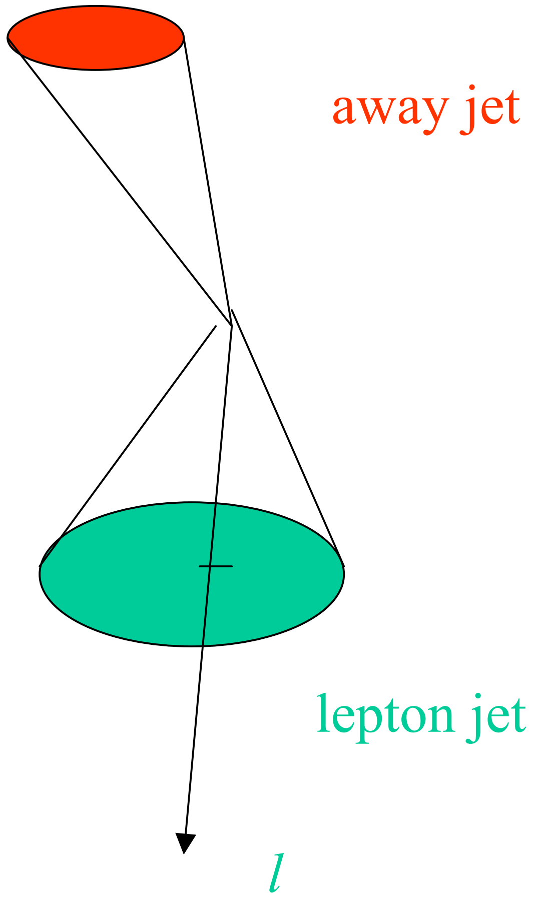
- *The simulation of the SLT algorithm uses efficiencies derived from the data (conversions, Z's and ψ mesons decays).*
- *Use generic-jet data to calibrate and cross-check the efficiency for finding SLT tags and supertags.*
- *Efficiency for finding supertags empirically corrected by 15%*



		σ (nb)				b_s(%)	tuned QCD			σ/ $σ_{QCD}$
		b	c	b_s	total		b	c	total	
	generic jets tuned	298	487	84	869	10%	382	487	869	1
#2	g. j. t. x BR	110	102	84	296	28%	141	102	243	1.2
	g. j. t. x BR^2	41	22	84	147	57%	52	21	73	2

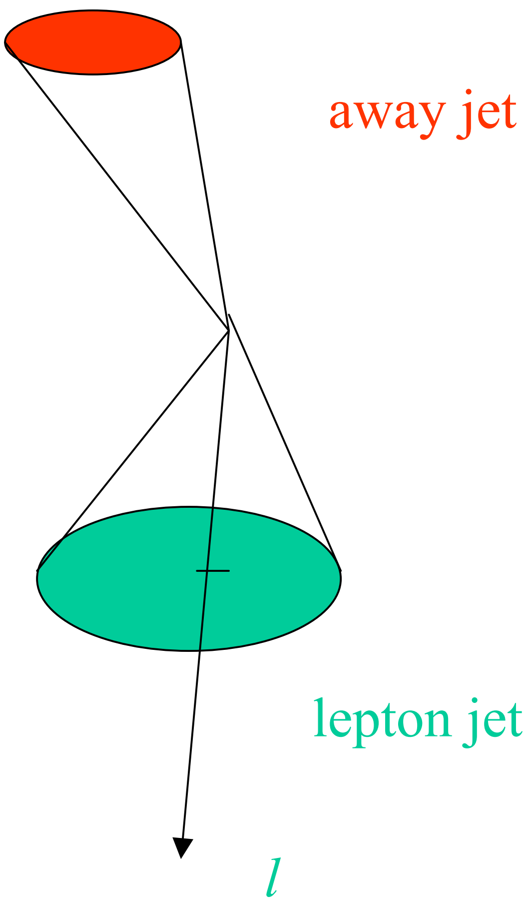
Signal Sample

- *Use sample enriched in Heavy Flavor content*
  - *Events with 2 or more jets with E_T > 15 GeV and at least two SVX tracks (taggable,$|\eta|$<1.5)*
  - *one electron with E_T> 8 GeV or one muon with p_T > 8 GeV/c contained in one of the jets*
- *Counting Experiment:*
  - *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*
  - *Check the semileptonic branching ratio of Heavy Flavor hadrons by counting the number of a-jets with a SLT and in the data and in the simulation*

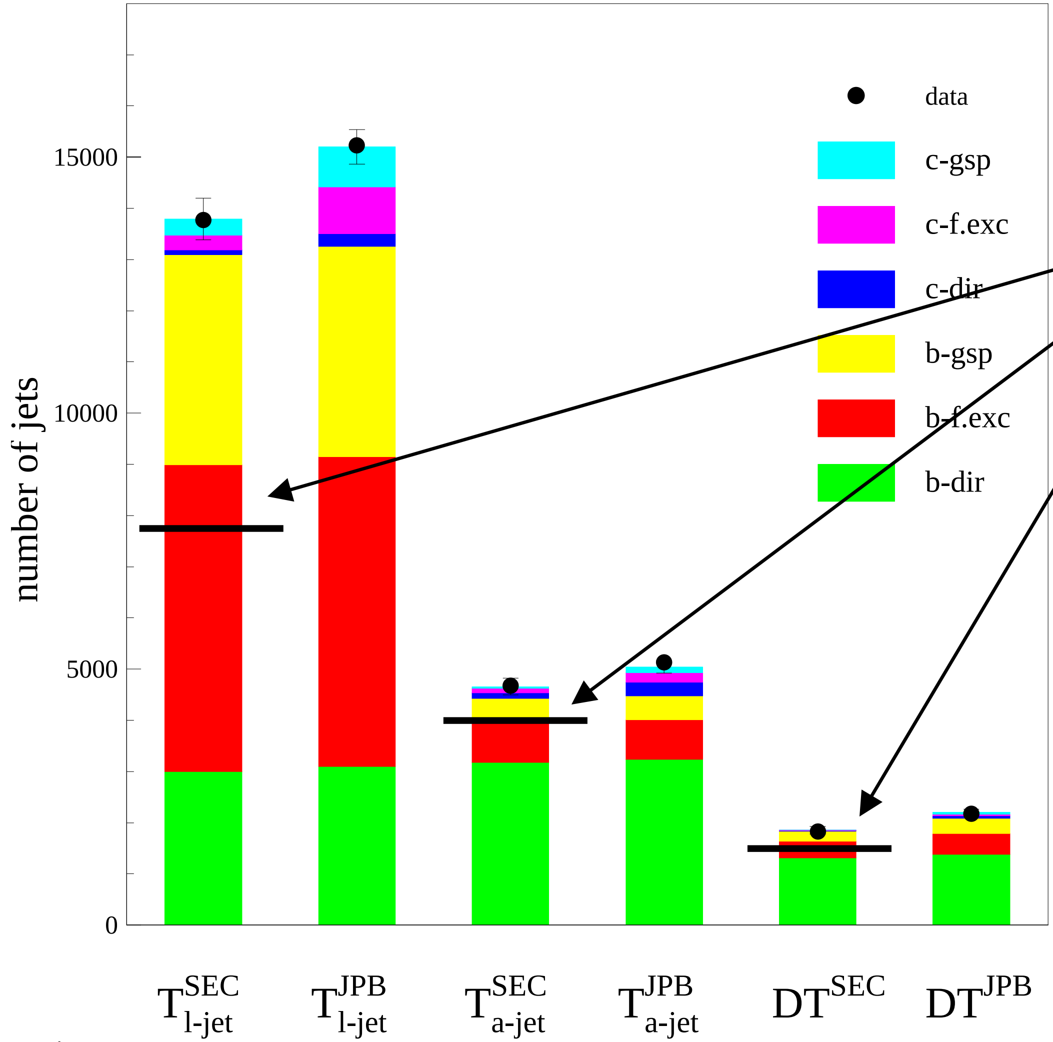


Tuning the Simulation to the data

- "Kitchen Dirty Work" :
  - Mistags evaluated with parameterization (10%)
  - SECVTX-JPB tagging efficiencies measured in data (6%)
  - SLT Efficiency uncertainty (10%)
  - Simulated supertag efficiency (SECVTX+SLT or JPB+SLT) is corrected for the data-to-simulation scale factor measured in the generic-jet sample (85±5%).
  - Take care of tagging rates in the fraction of lepton-trigger events with no h.f. using a parameterized probability of finding a tag due to heavy flavor in generic-jet data.



Tuned HERWIG



- F_{hf} = (45.3±1.9)% for e
- F_{hf} = (59.7±3.6)% for μ

NLO Calculation

Addressed Issues

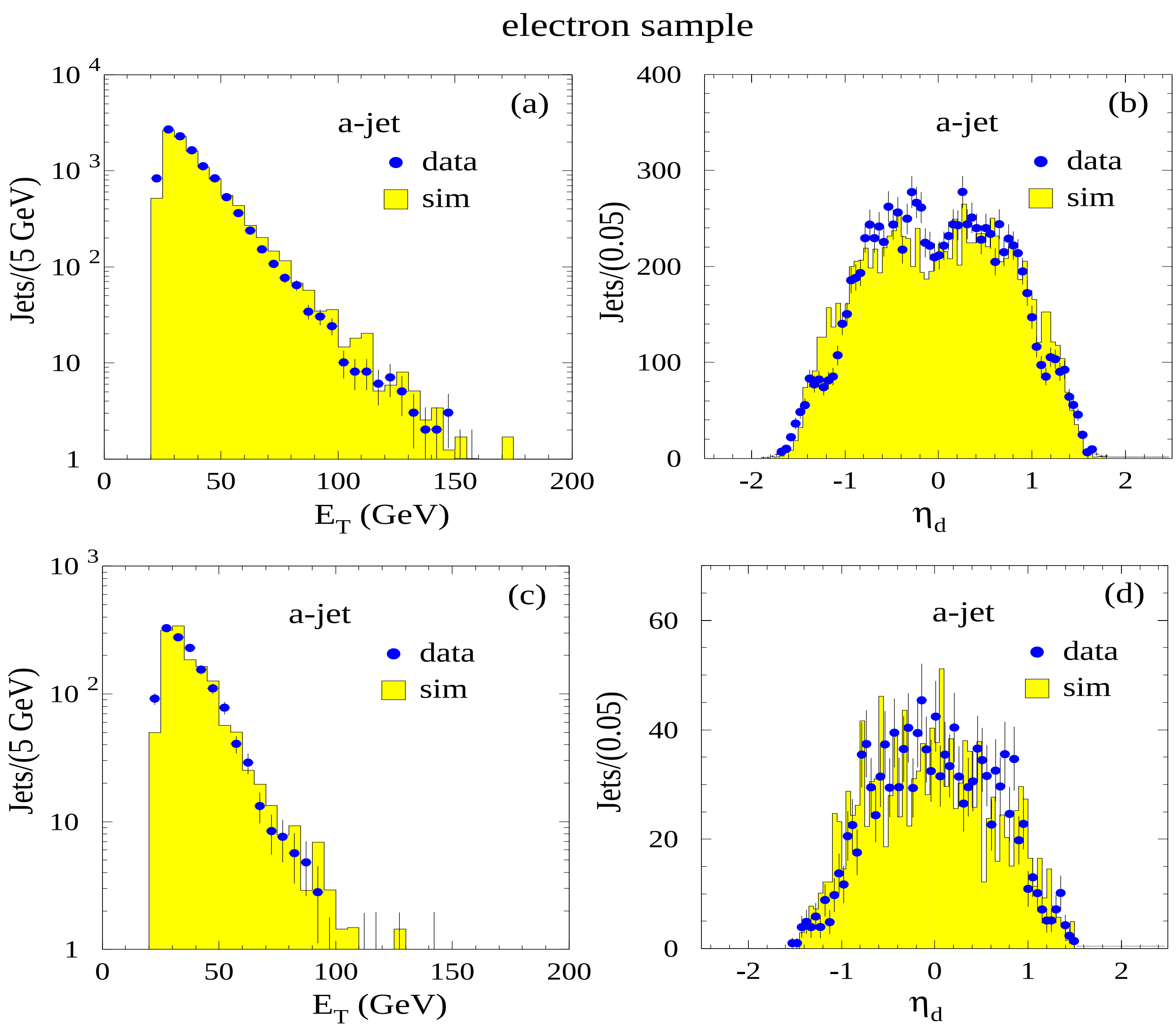
- *b-quark fragmentation*
- *k_T factorization (CASCADE)*
- *Berger's model (gluinos)*
- *Single b cross sections derived from 2 b cross sections using NLO prediction*

							tuned QCD			σ/σ_{QCD}	
#4	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

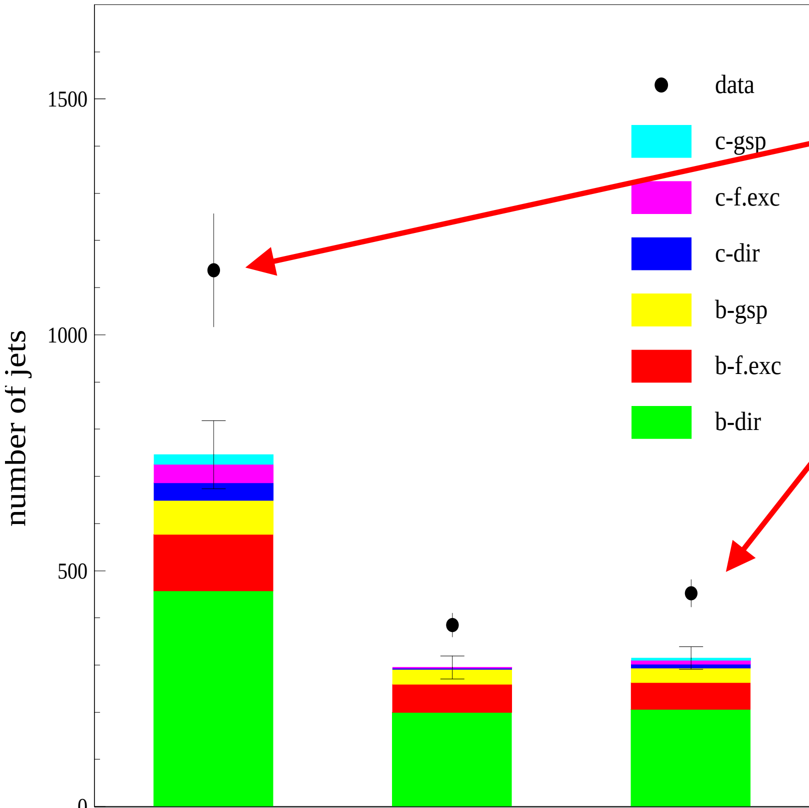
Kinematic Variables Data-Simulation Comparison

a-jet

a-jet with SECVTX tags



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.)

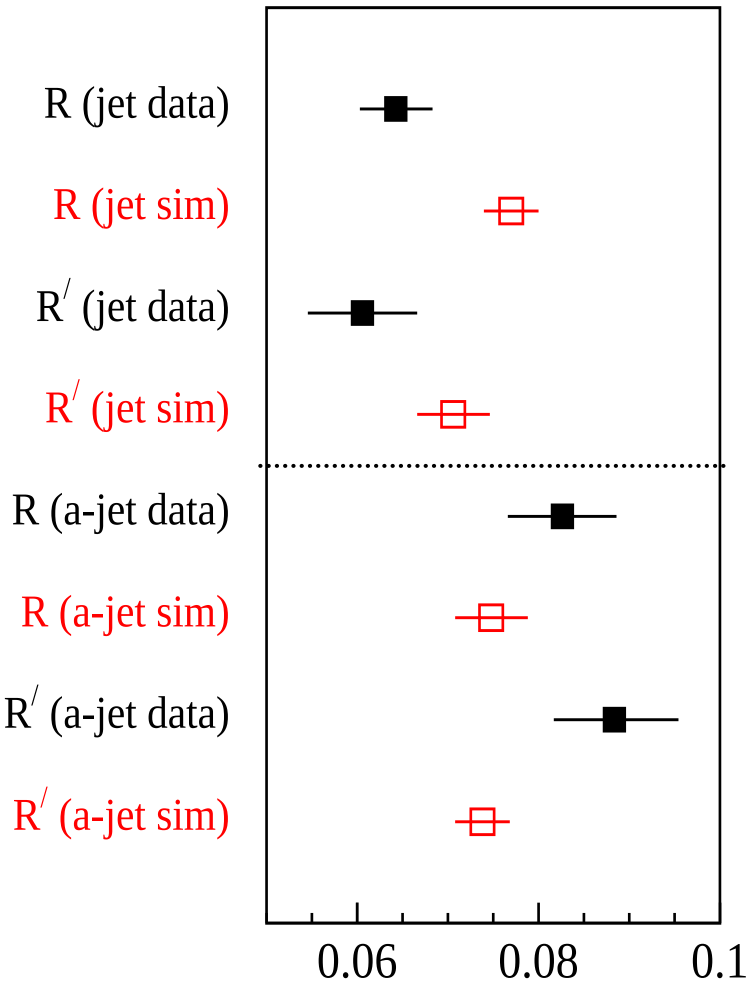
EXPECTED 316.5±25.4 (SYST)

- *~3 σ discrepancy, with errors dominated by systematic effects*

	σ (nb)				b_s(%)	tuned QCD			σ/σ_{QCD}
g .j. x BR tuned (or lep-trig. evts)	110	102	84	296	28%	194	102	296	1
#5 lep-trig. evts. x BR	41	22	84	147	57%	72	21	93	1.5 SS

Supertags

- *Data-Simulation comparison for the yield of R (R'), the ratio of number of jets with a SECVTX (JPB) and SLT tag – supertags - to that with a SECVTX (JPB) tag in the generic jet sample and in the Lepton-trigger sample.*
  - *The tuned QCD Simulation predicts the same yield of supertags in generic jet and lepton-trigger jets*
  - *Data show a ~30% discrepancy between supertags in generic jets and lepton-trigger jets.*
    - *Systematic uncertainties in the SLT simulated efficiency would shift in the same direction the yield R in the generic jets sample and lepton-trigger sample.*



Uncertainty on Mistags and SLT Tagging Efficiency on Heavy Flavors

- SLT mistags and tagging efficiency have been determined historically on data (PRD - 64, 032002) with conservative errors of 10% .
- The availability of a tuned simulation can be used to reduce the previous estimate of the SLT mistags and tagging efficiency systematic errors.
- Fit observed rates of SLT tags in generic jets with

  P_f x fakes + P_{hf} x h.f.
- The fit returns $P_f = 1.017 \pm 0.013$ and $P_{hf} = 0.981 \pm 0.045$, $\rho = -0.77$
- Using this result the SLT expectation in in the SS away-jets is 1362±28 whereas 1757±104 are observed (3.8 σ)
- This discrepancy cannot come from obvious prediction deficiencies

	observed	pred. fakes.	pred. h.f.
SLT's in g. jets	18885	15570±1557	3102 ±403
SLT's in g. jets with SECVTX	1451	999 ±60	508 ±51
SLT's in g. jets with JPB	2023	856 ±86	1175 ±71
SLT 's in a-jets (lep-trig.)	1757	619 ±62	747 ± 75

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*

Support Slides

Tuning 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 results

$\chi^2/DOF=4.6/9$

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 dird.pro	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

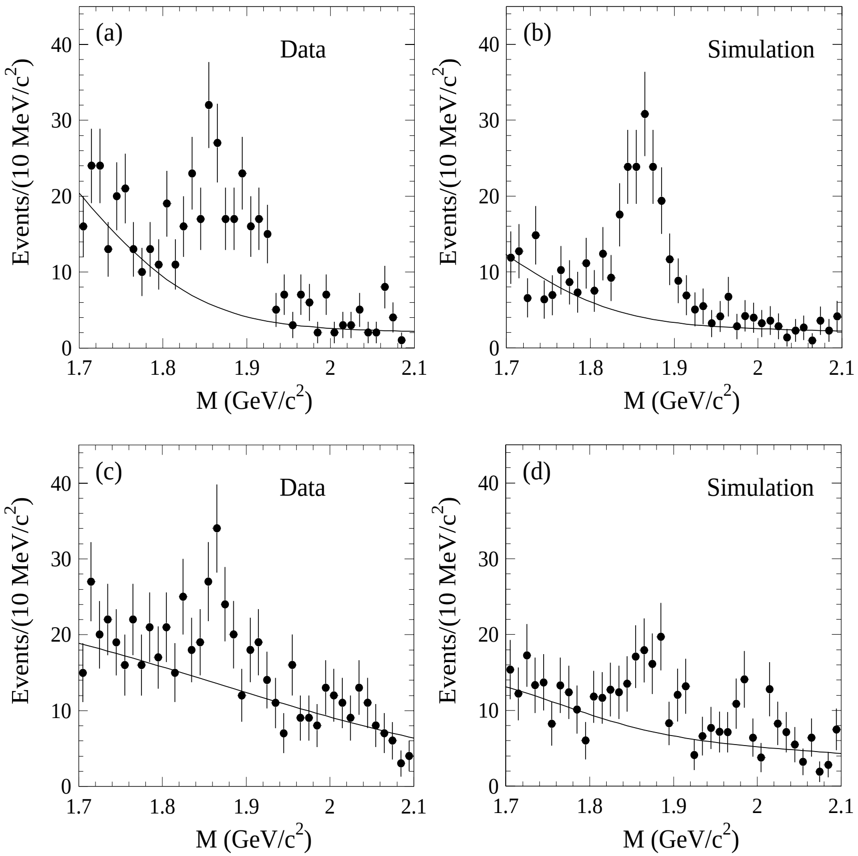
#4

	σ (nb)				b_s(%)	tuned QCD			σ/σ_{QCD}	
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

b-purity (cross-check)

- l-D^0 : 126.0 ± 15.5 in the data and 139.9 ± 15.0 in the simulation
- l-$D^{\pm}$: 73.7 ± 17.8 (data) and 68.5 ± 14.1 (simulation).
- J/ψ: 90.8 ± 10.1 (data) and 101.9 ± 11.4 (simulation)
- Ratio of the b-purity in the simulation to that in the data is 1.09 ± 0.11
  - Discrepancy between observed and predicted number of a-jets with SLT tags due to heavy flavor is not due to an underestimate of the bb contribution

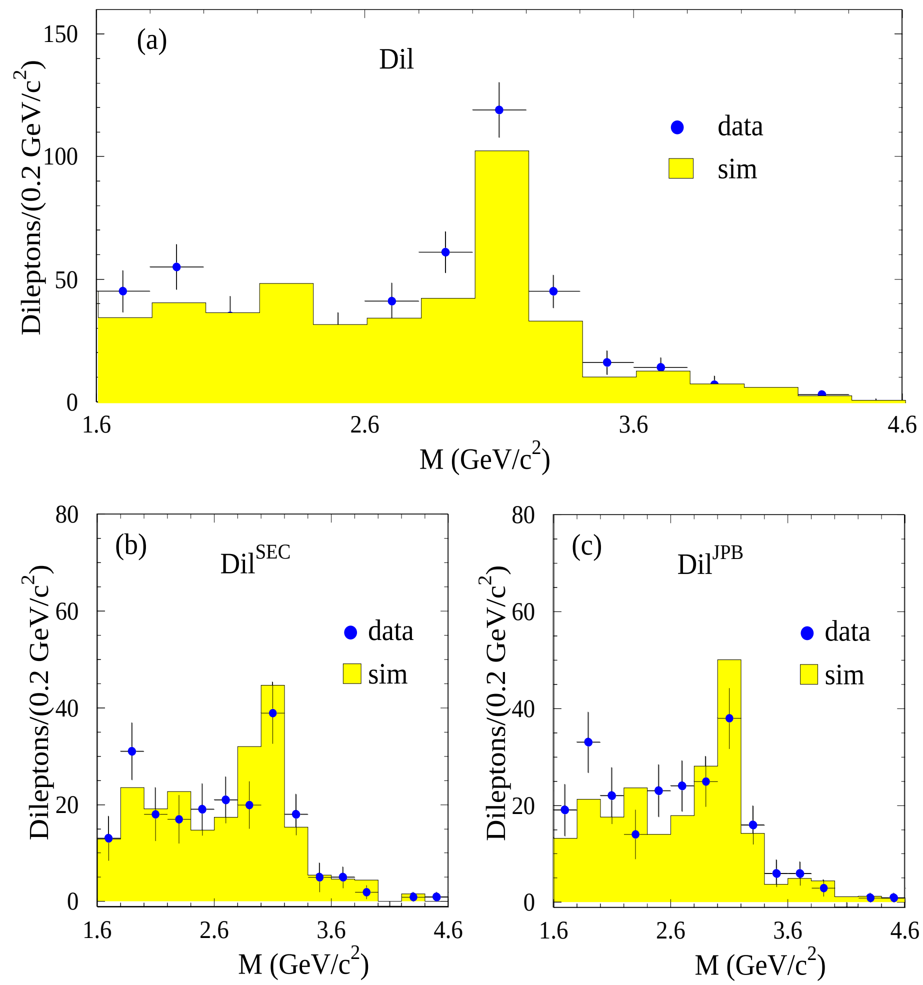
$$D^0 \to K^-\pi^+$$



$$D^- \to K^+\pi^-\pi^-$$

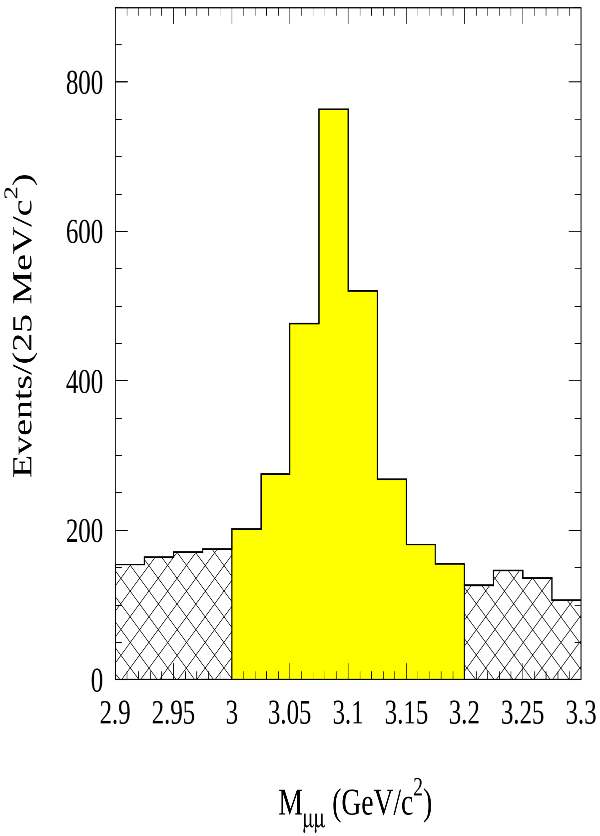
b-purity (cross-check)

- $2.6 < m_{ee} < 3.6\ GeV/c^2$
- $2.9 < m_{\mu\mu} < 3.3\ GeV/c^2$
- SS dileptons (with 10% error) used to estimate and remove bkg. to OS dileptons due to misidentified leptons.
- 259 ± 17.2 and 209.2 ± 23.7 (before tagging)
- 89.7 ± 10.5 and 100.5 ± 12.4 (SECVTX)
- 90.8 ± 10.1 and 101.9 ± 11.4 (JPB)



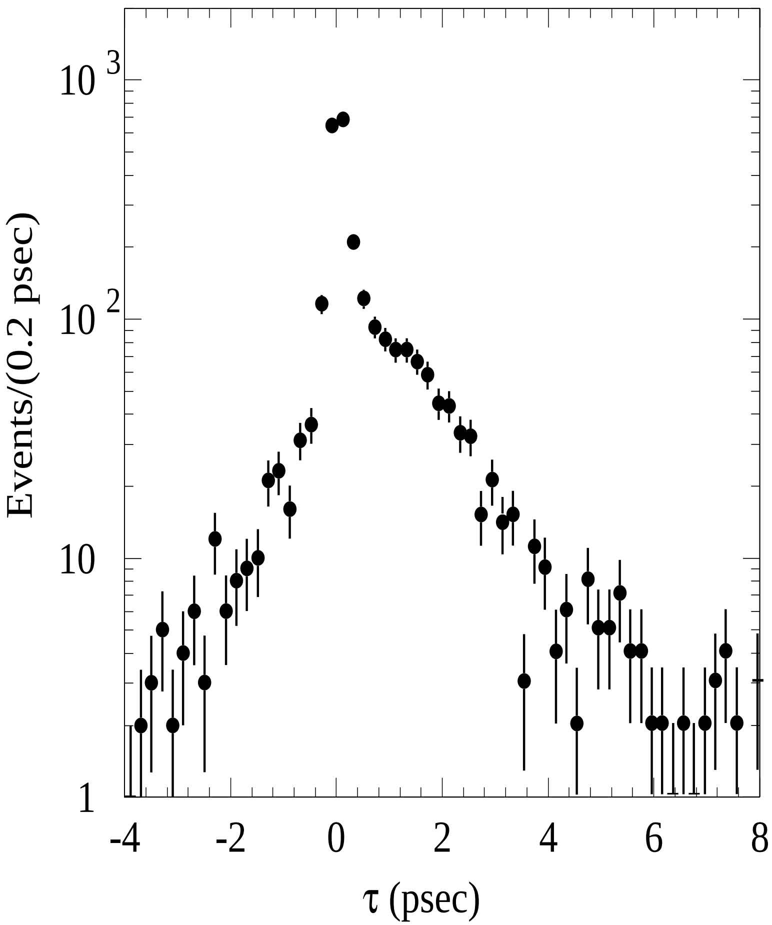
J/ψ mesons from B-decays

- *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 $J/\psi \rightarrow \mu\mu$ 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*
- *The observed number of SLT tags is 1.2 σ lower than the prediction rather than 50% larger as in the inclusive lepton sample.*



Data

	electron data			muon data		
ttag eyp			P_{DGQC}			P_{DGQC}
N_{l-je}	68544			14966		
N_{a-je}	73335			16460		
T^{SEC}_{l-je}	10115.3 ± 101.7	$(10221/105.7)$	0	3657 ± 60.8	$(3689/31.7)$	0
T^{BJP}_{l-je}	11165.4 ± 115.8	$(11591/425.6)$	0	4068.6 ± 66.2	$(4204/135.4)$	0
T^{SEC}_{a-je}	4353.3 ± 68.5	$(4494/140.7)$	1.56%	1054.6 ± 33.3	$(1094/39.4)$	1.67%
T^{BJP}_{a-je}	5018.9 ± 98.9	$(5661/642.1)$	2.45%	1265.2 ± 41.1	$(1427/161.8)$	2.63%
DT^{SEC}	1375.2 ± 37.6	$(1405/29.8)$	0	452.6 ± 21.6	$(465/12.4)$	0
DT^{BJP}	1627.8 ± 43.7	$(1754/126.2)$	0	546.4 ± 25.1	$(600/53.6)$	0

Heavy flavors in the simulation are identified at generator level

electron simulation

ttag eyp	b-dir	c-dir	b-f.exc	c-f.exc	b-gsp	c-gsp
HF_{l-tye}	5671	947	10779	2786	5263	1690
HF_{a-tye}	5848	977	11280	2913	6025	1877
h.f./light	5407/441	899/78	1605/9675	367/2546	707/5318	145/1732
HTF^{SEC}_{l-tye}	1867	52	3624	194	1732	147
HTF^{BJP}_{l-tye}	2392	163	4531	602	2106	356
HTF^{SEC}_{a-tye}	2093	91	480	68	222	15
HTF^{BJP}_{a-tye}	2622	203	584	136	276	58
HDF^{SEC}	678	5	157	4	78	1
HDF^{BJP}	1083	43	303	25	168	18

muon simulation

ttag eyp	b-dir	c-dir	b-f.exc	c-f.exc	b-gsp	c-gsp
HF_{l-tye}	1285	298	2539	942	1455	747
HF_{a-tye}	1358	313	2705	994	1708	816
h.f./prompt	1206/152	278/35	422/2283	124/870	171/1537	48/768
HTF^{SEC}_{l-tye}	569	34	1131	83	652	92
HTF^{BJP}_{l-tye}	707	77	1386	229	830	202
HTF^{SEC}_{a-tye}	498	29	132	13	54	11
HTF^{BJP}_{a-tye}	627	62	173	34	60	21
HDF^{SEC}	218	3	59	2	20	1
HDF^{BJP}	347	12	105	7	50	6

Fit of the simulation to the data

- 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
- The ratio of b to c direct production constrained to the default value (about 1) within 14%
- the ratio of b to c flavor excitation constrained to the default value (about 0.5) with a 28% uncertainty
- bg constrained to (1.4±0.19)
- cg constrained to (1.35±0.36)
- The tagging efficiencies are also fit parameters, and are constrained to their measured values within their uncertainties (6% for b-quarks, 28% for c-quarks)

Fit result-parameter corr. coeff.

	SF_c	SF_{JPB}	K_e	c	bf	cf	bg	cg	K_μ
SF_b	-0.073	0.718	-0.747	0.054	0.346	0.297	-0.062	0.066	-0.715
SF_c		0.358	-0.238	-0.002	0.038	0.147	-0.071	0.086	-0.306
SF_{JPB}			-0.810	0.010	0.363	0.127	-0.009	-0.049	-0.802
K_e				-0.092	-0.641	-0.302	0.071	0.077	0.933
c					0.053	0.020	0.008	0.002	-0.098
bf						0.245	-0.680	-0.199	-0.526
cf							-0.321	-0.164	-0.274
bg								-0.029	-0.019
cg									-0.018

For

NLO and Herwig calculations

- However, in this specific analysis we are interested in comparing rates of a-jet with heavy flavor (signaled by SLT or SECVTX tags) in events in which the l-jet has also heavy flavor
- These jet have $|\eta|<1$ and corresponds to partons with $E_T > 18$ GeV
- In this case Herwig evaluates that the gluon splitting+flavor excitation contribution are 40% of the Born contribution and not a factor of 3 higher
- For this type of kinematics, the ratio of the NLO to Born calculations is also of the order of 1.1-1.3. In addition, for this topology, the NLO calculation depends little on the choice of μ, and it appears to meet general criteria of robustness.

NLO and Herwig calculations

- *Herwig ignores interference terms between the Born approximation and the NLO diagrams, and evaluates a gluon splitting+flavor excitation contribution which is a factor of 3 larger than the Born approximation.*
- *In the NLO calculation the contribution of the Born cross section and of the gluon splitting+flavor excitation are approximately equal using the renormalization scale* μ; *when using the scale the scale* $\mu/2$,*the NLO calculation gets closer to Herwig.*
- *The fact that the ratio between NLO and Born is about two and is not stable as a function of the renormalization scale is taken by the experts as an indication that NNLO corrections are important*
- *The relevance of the Herwig result, which models the data, is the indication that the effect of NNLO correction should be that of canceling the interference terms*

away-jets with SLT tags

ttag eyp	electron data		P_{DGQC}	muon data		P_{DGQC}
T^{SL}_{a-jet}	1063.8 ± 47.0	$(2097/1033.2)$	0.49%	308.6 ± 34.7	$(562/253.4)$	0.54%
$T^{SL \cdot SEC}_{a-jet}$	356.3 ± 22.8	$(444/87.7)$	0.08%	69.3 ± 9.9	$(92/22.7)$	0.09%
$T^{SL \cdot JP}_{a-jet}$	401.3 ± 25.3	$(513/111.7)$	0.13%	112.3 ± 12.3	$(143/30.7)$	0.14%

	Electrons		Muons	
Tage yp	Data	Simulation	Data	Simulation
HF^{SL}_{a-jet}	865.1 ± 114.8	597.6 ± 69.3	272.7 ± 34.9	149.3 ± 21.0
$HF^{SL \cdot SEC}_{a-jet}$	322.6 ± 23.3	242.4 ± 22.5	63.3 ± 9.9	53.8 ± 8.7
$HF^{SL \cdot JPB}_{a-jet}$	350.2 ± 26.3	251.5 ± 21.7	103.2 ± 12.4	65.0 ± 8.9