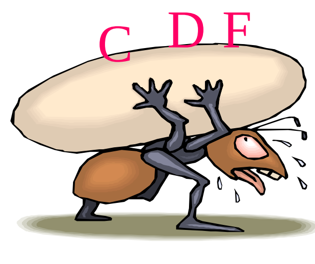
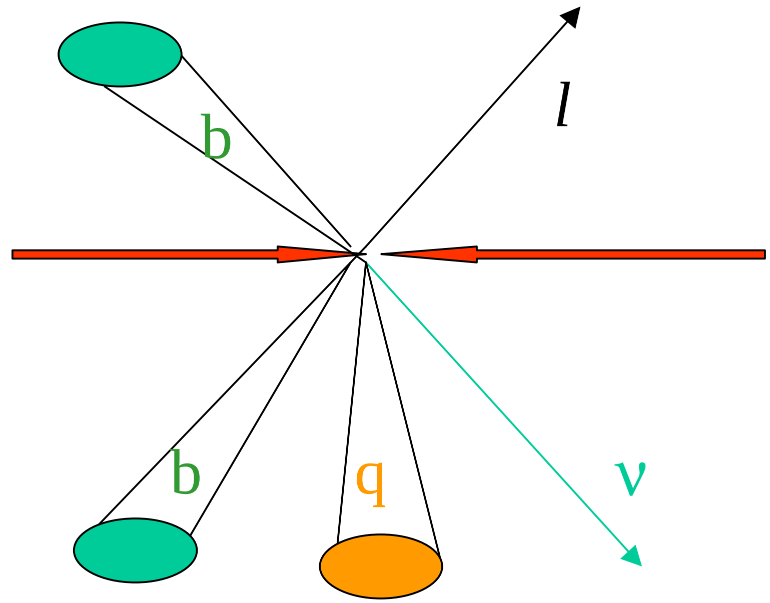
Study of the heavy flavor content of jets produced in association with *W* bosons

Chronicle of a journey to the Pillars of Hercules

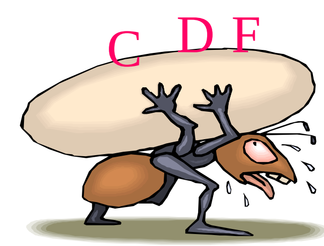


P. Giromini/INFN-Frascati

Data sample : lepton + jet events

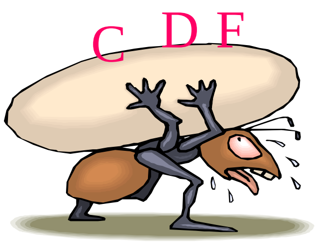


- Kinematic cuts:
  - One high E_T lepton (e , μ)
    - E_T >20 GeV
    - Central (| η | <1)
    - Isolated
  - $\not{E}_T$ > 20 GeV
  - at least 1 jet with E_T >15 GeV and |η| <2.0
- b-identification : – Efficiency
  - SECondary VerTeX (SECVTX) ~ 50%
  - Jet-ProBability (JPB) ~ 50%
  - semileptonic decay (Soft L epton T agging) ~ 20%

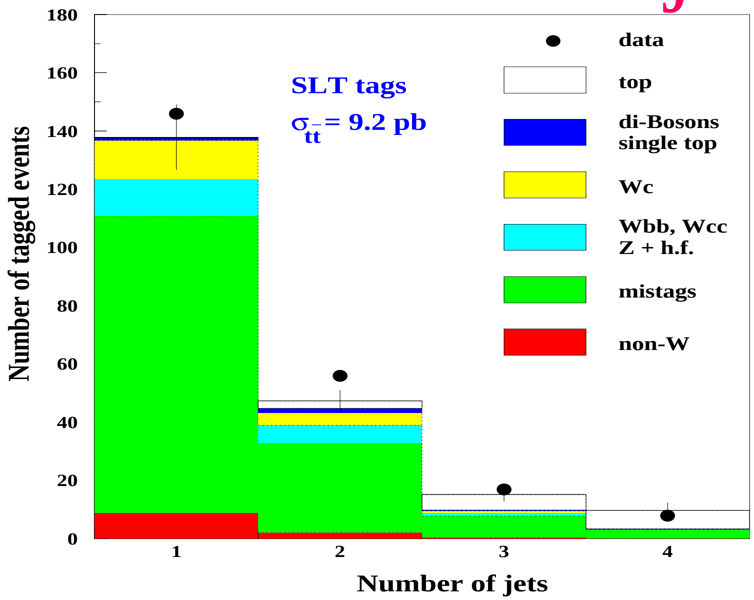


W + jets composition

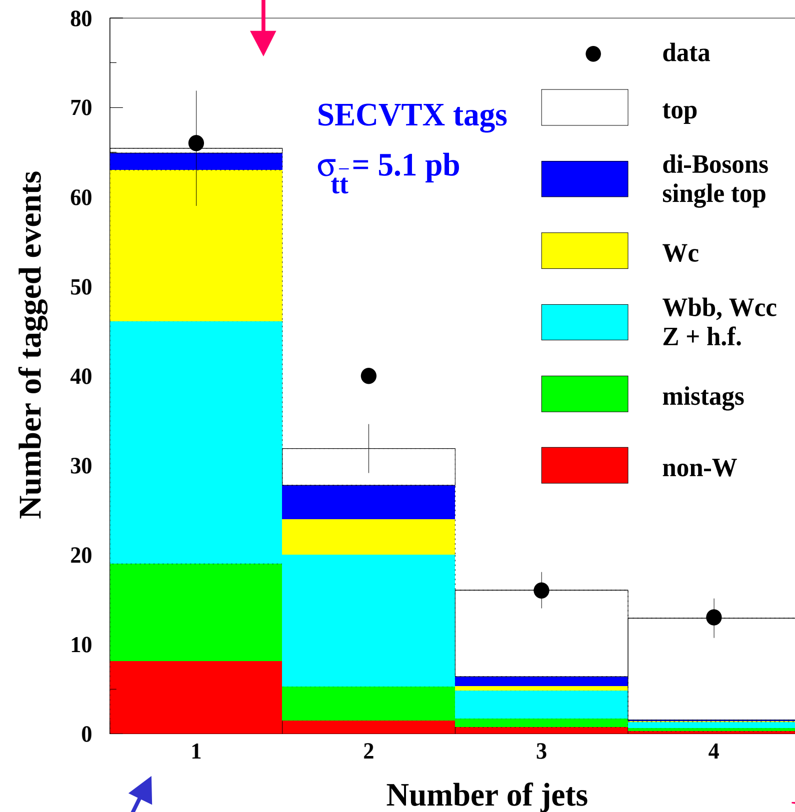
Source	$W + 1\,\text{jet}$	$W + 2\,\text{jet}$	$W + 3\,\text{jet}$	$W + \geq 4\,\text{jet}$
Data	9454	1370	198	54
Non-W	560.1 ± 14.9	71.2 ± 2.7	12.4 ± 2.0	5.1 ± 1.7
WW	31.2 ± 5.4	31.1 ± 5.4	5.2 ± 1.0	0.8 ± 0.2
WZ	4.4 ± 0.9	4.8 ± 1.0	0.9 ± 0.2	0.1 ± 0.0
ZZ	0.3 ± 0.1	0.4 ± 0.1	0.1 ± 0.0	0.0 ± 0.0
Unidentified-Z + jets	234.8 ± 14.5	38.5 ± 5.9	7.9 ± 2.4	0.7 ± 0.7
Single top	14.1 ± 2.1	7.9 ± 1.7	1.7 ± 0.4	0.3 ± 0.1
$t\bar{t}$	1.8 ± 0.5	10.1 ± 2.8	20.3 ± 5.7	21.3 ± 5.9
$W+$ jets without h.f.	7952.0 ± 133.6	1027.7 ± 31.1	121.1 ± 7.7	19.9 ± 6.1
Wc	413.1 ± 123.9	86.8 ± 26.1	11.2 ± 3.4	1.9 ± 0.7
$Wc\bar{c}$	173.1 ± 46.2	61.9 ± 13.6	11.4 ± 2.6	2.3 ± 0.9
$Wb\bar{b}$	69.0 ± 9.5	29.7 ± 5.1	5.7 ± 1.1	1.5 ± 0.5

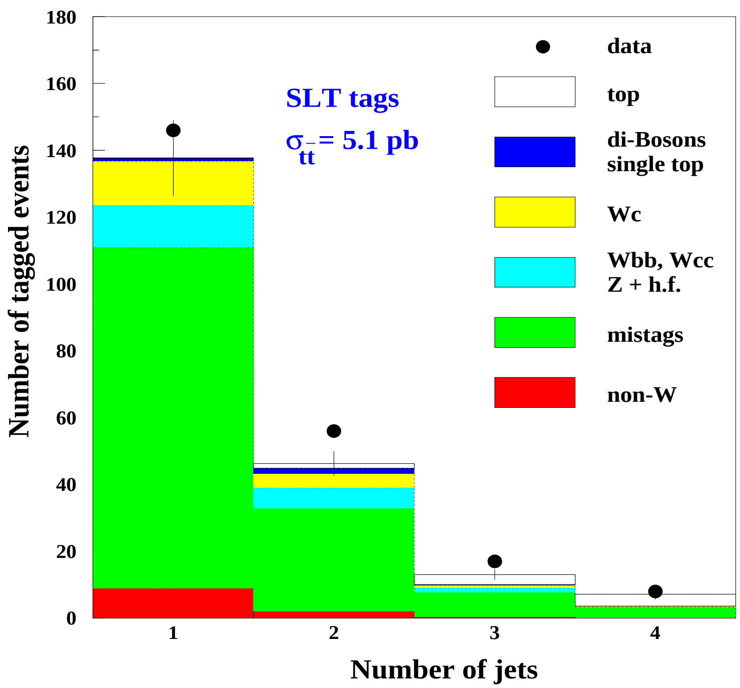


W + jets after tagging

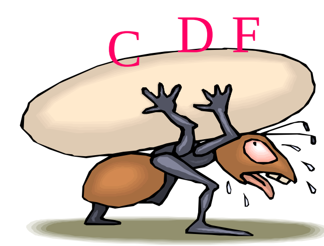


Phys. Rev. D64, 032002 (2001)

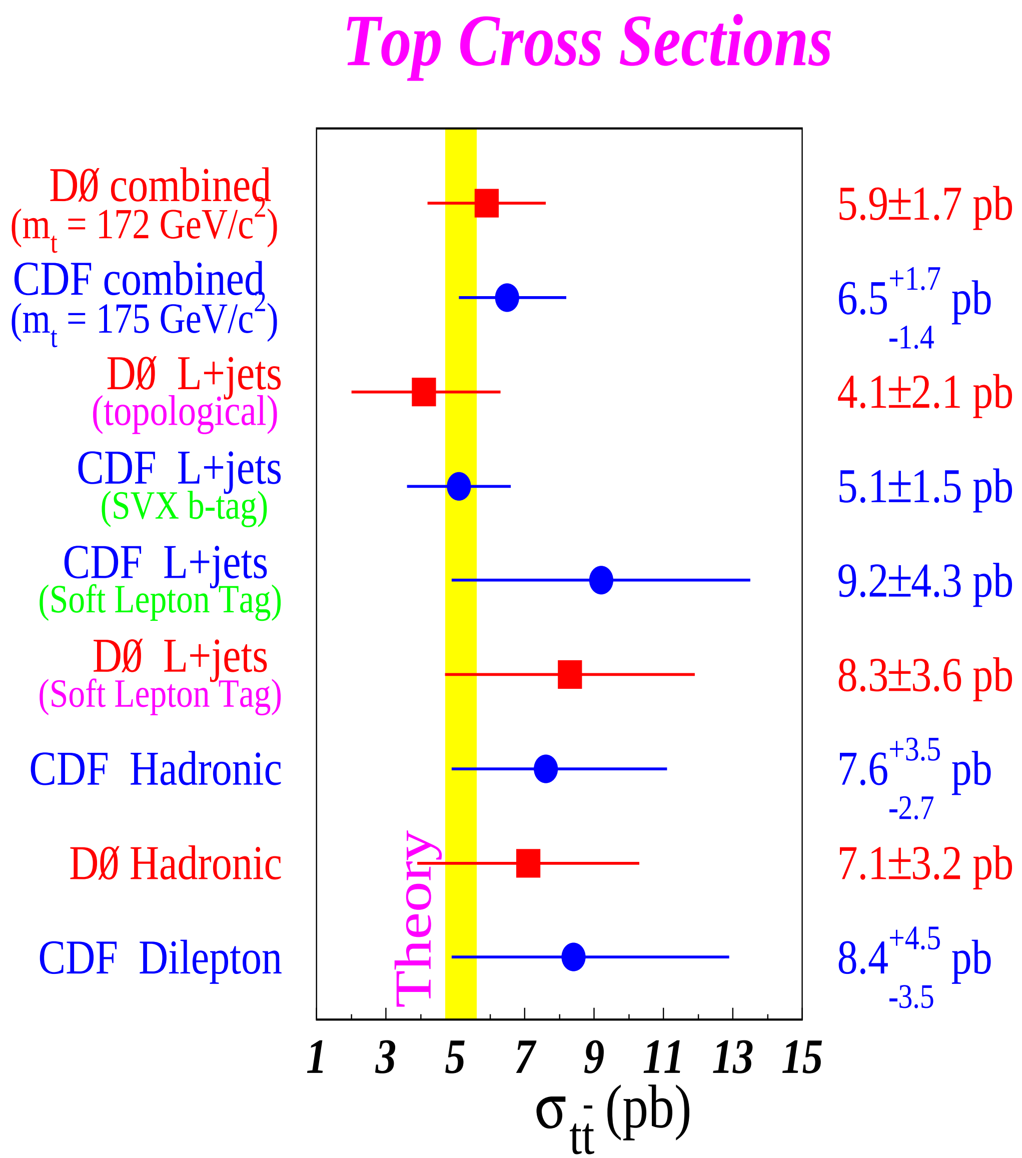


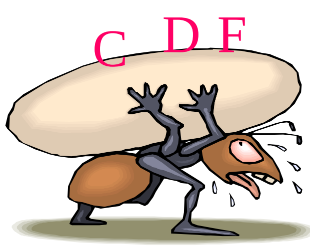


P~50%

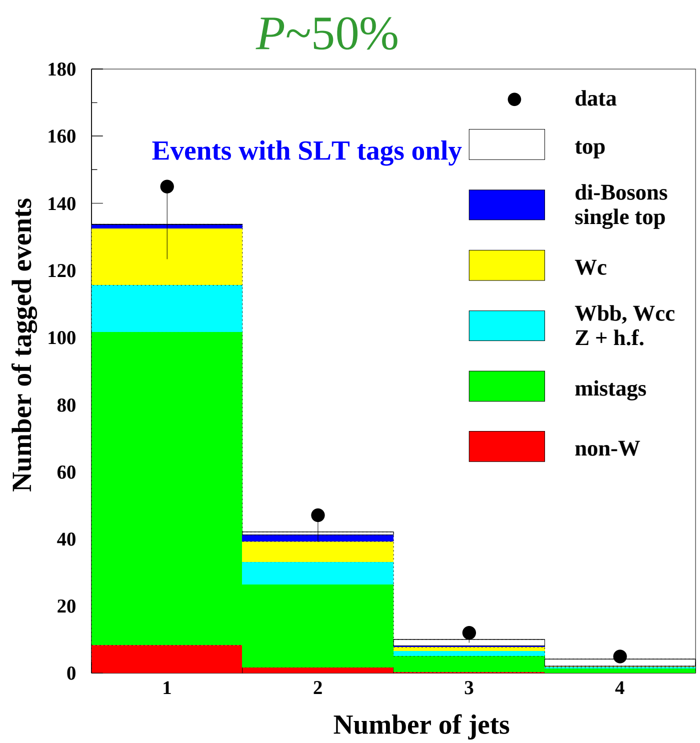


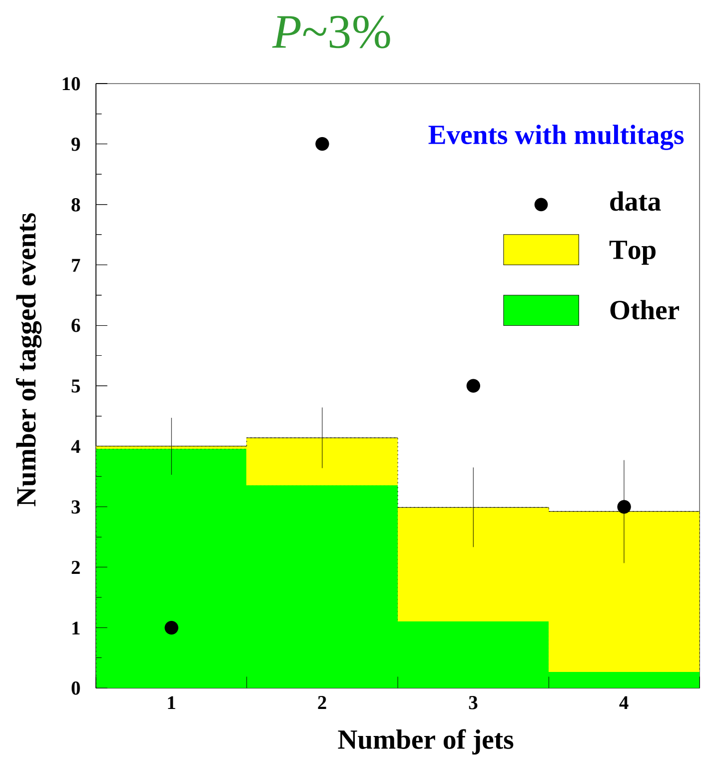
Various topological channels



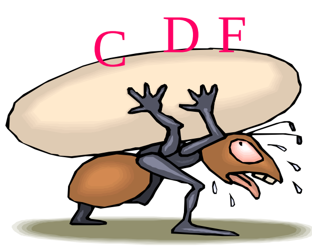


Correlation between SLT and SECVTX tags



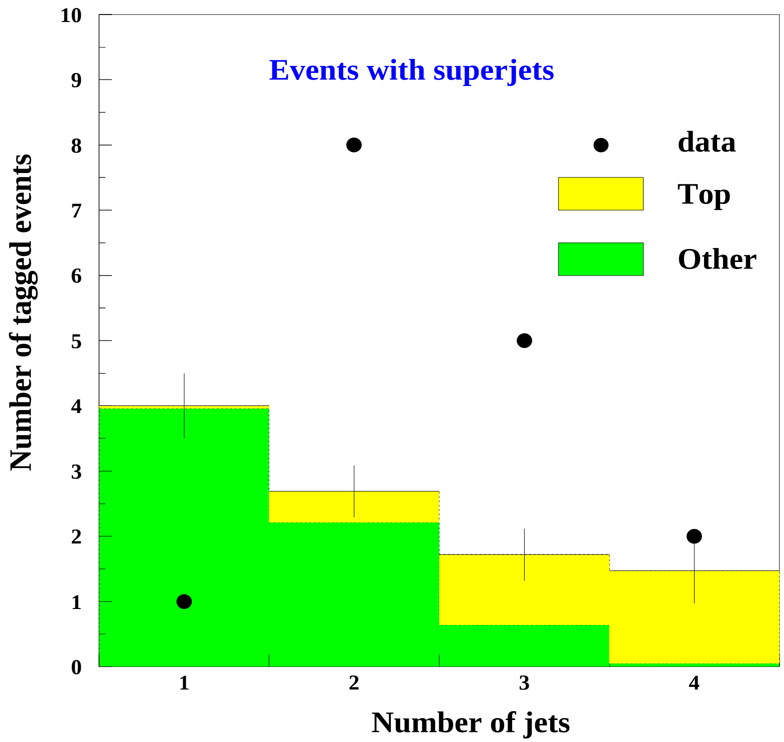


- the numbers of events with SLT and no SECVTX tags (mostly fakes) are consistent with the prediction. The numbers of observed and predicted events with both SLT and SECVTX tags (mostly real heavy flavor) are not very consistent.
- 70% of the events tagged by SECVTX contain b-jets (20% c-jets)
- we check the semileptonic BR in heavy quark jets tagged by SECVTX by studying the fraction of SECVTX tagged jets which contain also a SLT tag (supertag)

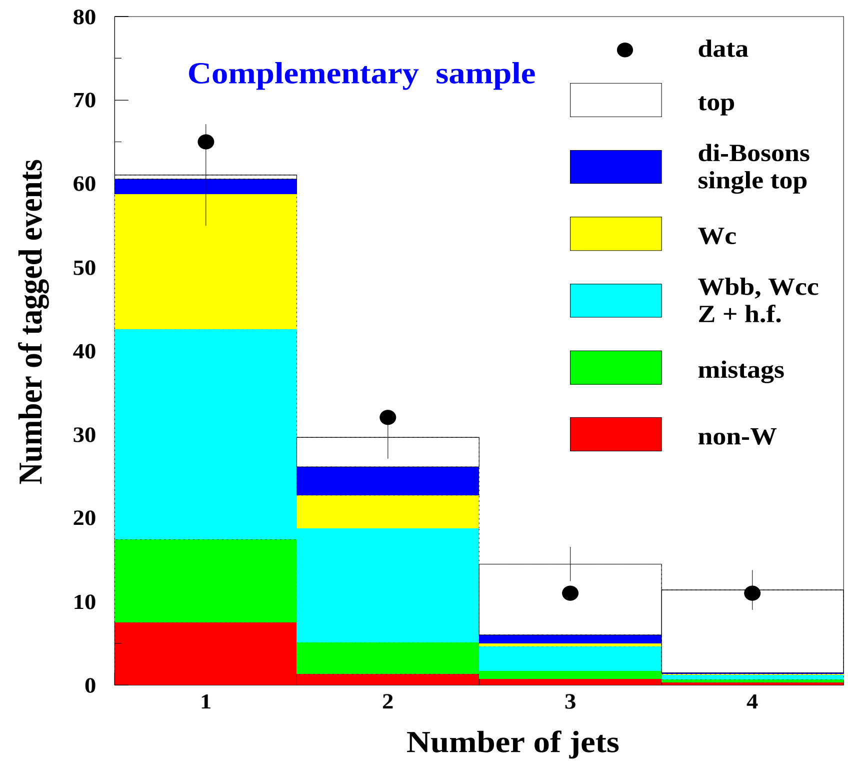


Events tagged by SECVTX with and without (complementary/control sample) supertags

P~0.4%



P~50%

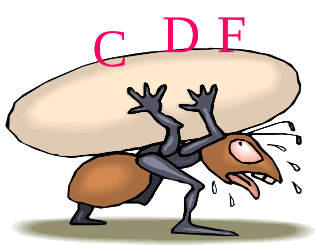


- the complementary/control sample has 43 events and 43.6 ± 3.3 are expected
- the *a posteriori* probability of observing 13 or more events with a superjet when 4.4 ± 0.6 are expected is $P=10^{-3}$
- 4 events with a superjet were included in the data set used to measure the top quark mass

Additional SECVTX tags

Source	$W+1$ jet	$W+2$ jet	$W+3$ jet	$W+\geq 4$ jet
		Events with a superjet		
SM prediction (supertags)	4.00 ± 0.50	2.69 ± 0.41	1.71 ± 0.40	1.47 ± 0.51
Data (supertags)	1	8	5	2
SM prediction (DT)		0.26 ± 0.06	0.36 ± 0.08	0.50 ± 0.13
Data (DT)		2	3	0
		Complementary sample		
SM prediction	61.44 ± 6.09	29.26 ± 2.58	14.39 ± 2.34	11.48 ± 2.37
Data	65	32	11	11
SM prediction (DT)		2.15 ± 0.50	2.87 ± 0.67	3.53 ± 0.90
Data (DT)		3	3	2

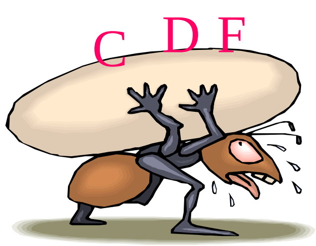
- If superjet events were C.S. alike, we would expect 1.8 DT.
- The likelihood of observing 5 additional SECVTX tags is $4.1\text{x}10^{-2}$.



Cross-check using generic-jet data

Tag type	Data (removed fakes)	Simulation
JET 20 (194009 events)		
SECVTX	4058±92 (616.0)	4052±143
JPB	5542±295 (2801.0)	5573±173
SLT	1032±402 (3962.0)	826±122
SLT+SECVTX	219.8±20 (94.2)	263±29
SLT+JPB	287.3±28 (166.7)	330±29
JET 50 (151270 events)		
SECVTX	5176±158 (1360.0)	5314±142
JPB	6833±482 (4700.0)	6740±171
SLT	1167±530 (5241.0)	1116±111
SLT+SECVTX	347±29 (169.0)	404±22
SLT+JPB	427.5±42 (288.5)	490±32
JET 100 (129434 events)		
SECVTX	5455±239 (2227.0)	5889±176
JPB	6871±659 (6494.0)	7263±202
SLT	1116±642 (6367.0)	1160±168
SLT+SECVTX	377.6±36 (243.4)	508±35
SLT+JPB	451.8±55 (401.2)	563±34

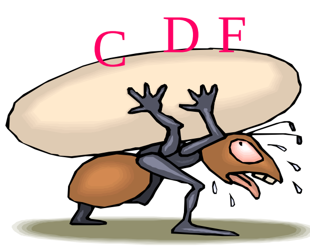
- Herwig simulation
  - tuned by fitting the rates of SECVTX and JPB tags
- (Flavor excitation +direct production) x (1.1±0.16)
- g → bb x (1.4±0.19)
- g → cc x (1.35±0.36)



Cross-check using generic-jet data

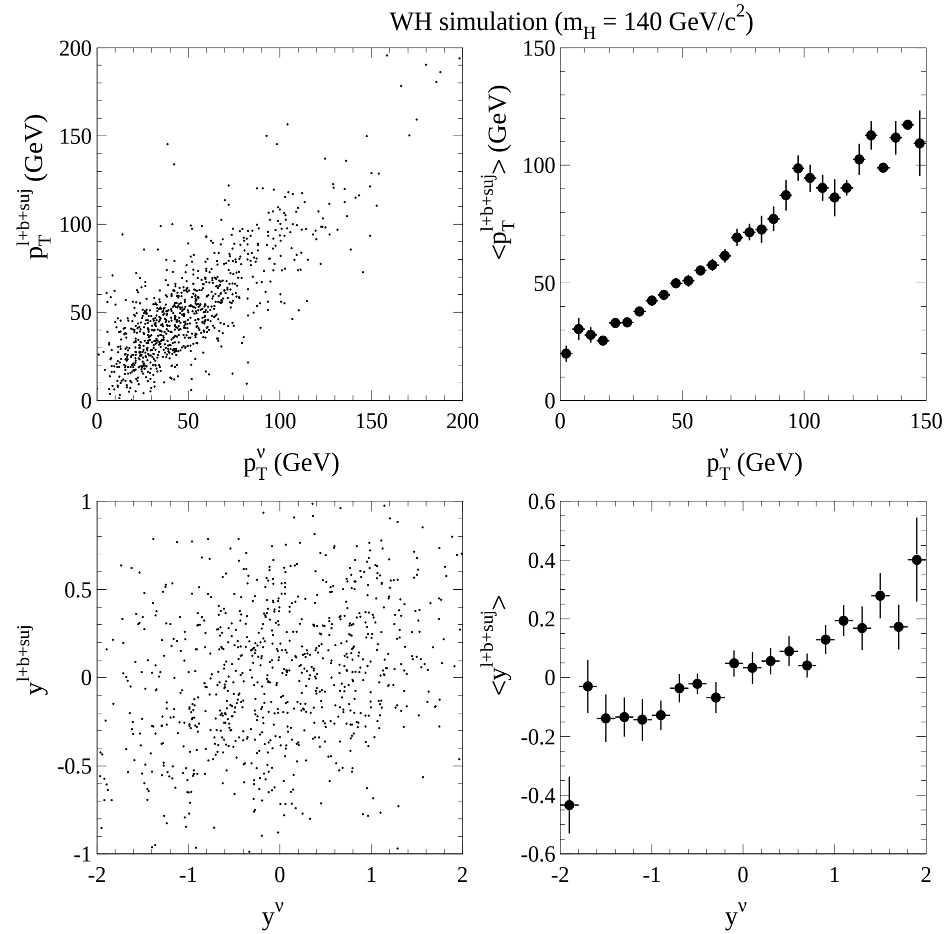
	JET 20		JET 50		JET 100	
	$\frac{\text{SLT+SECVTX}}{\text{SECVTX}}$	$\frac{\text{SLT+JPB}}{\text{JPB}}$	$\frac{\text{SLT+SECVTX}}{\text{SECVTX}}$	$\frac{\text{SLT+JPB}}{\text{JPB}}$	$\frac{\text{SLT+SECVTX}}{\text{SECVTX}}$	$\frac{\text{SLT+JPB}}{\text{JPB}}$
Data	0.054±0.005	0.052±0.006	0.067±0.006	0.063±0.008	0.069±0.007	0.066±0.010
Sim.	0.065±0.007	0.059±0.005	0.076±0.004	0.073±0.005	0.086±0.006	0.077±0.005
Data/Sim.	0.83±0.12	0.88±0.13	0.88±0.09	0.86±0.12	0.80±0.10	0.86±0.14

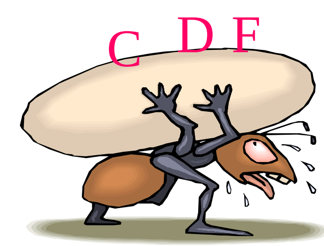
- The efficiency for finding supertags in the data is smaller than in the simulation (85±5)%
- note: after tagging with SECVTX, the heavy flavor composition of generic-jets is similar to W + jet events



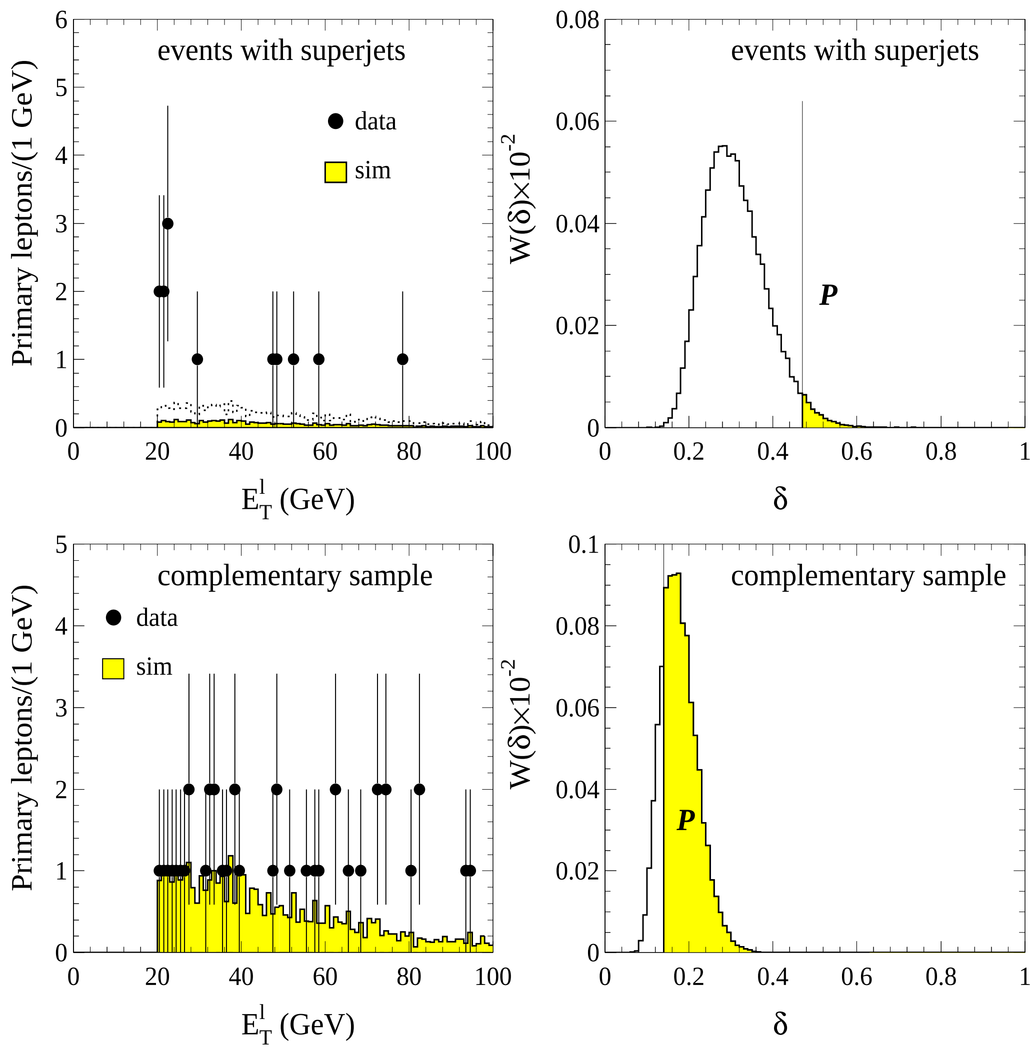
Study of the kinematics

- If the 13 events are a statistical fluctuation, the kinematics of this sample will be consistent with the S.M. simulation and the complementary sample
- We chose two sets of 9 variables to look for differences
- The first set studies $d^2\sigma / (dp_T\, d\eta)$ for every different object in the final state (8 var.)
- Replace $\not{E}$ with the system l+suj+b
- Add the angle between the lepton and *W* (check if events are consistent with the production and decay of *W* bosons)
- This set of 9 variables fully describes the kinematics of the final state with modest correlations
- Compare distributions in the data and the SM simulation using a K-S test. Use the Kuiper's definition.
- the second set shown later

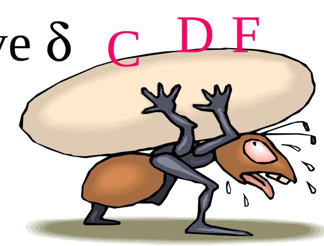




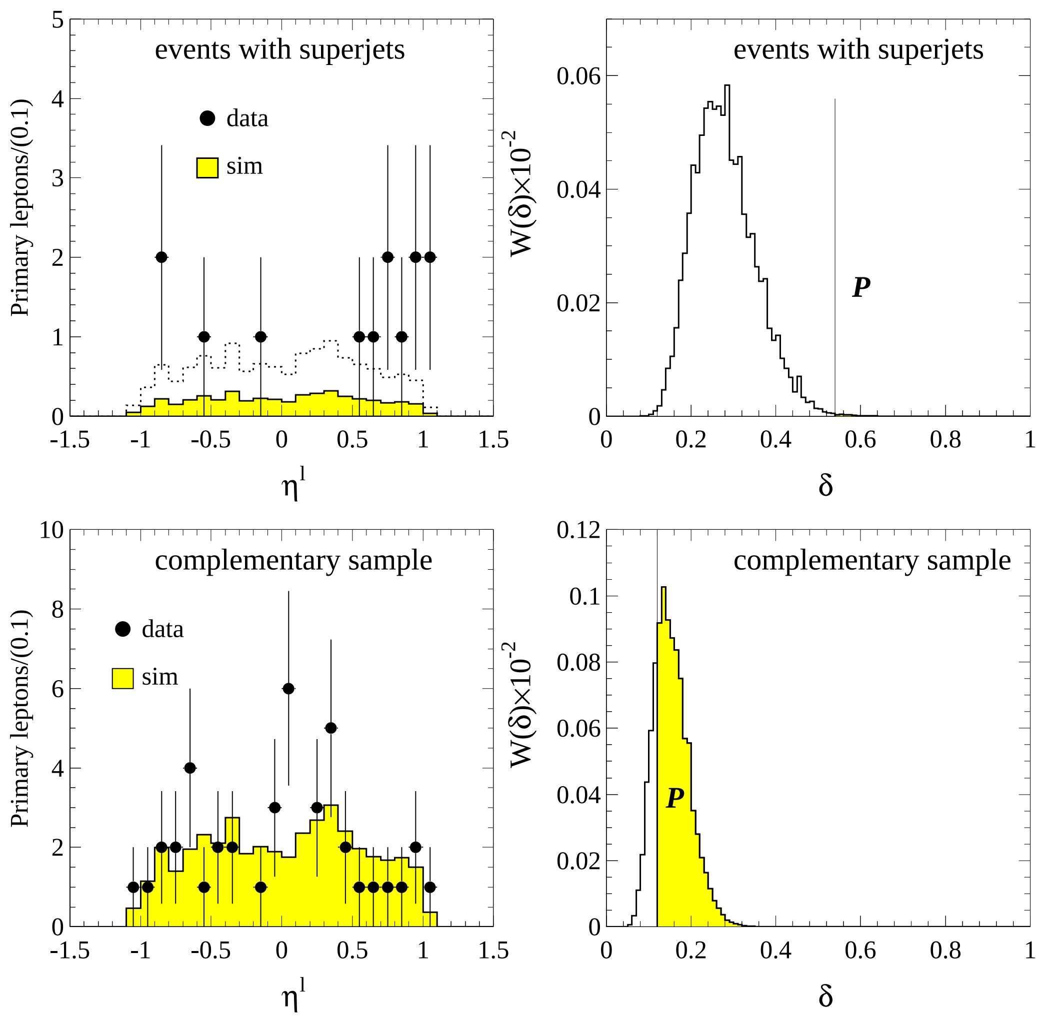
Primary lepton

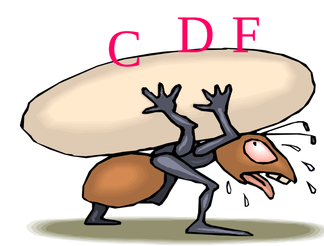


- The probability distribution of the K-S distance, δ, is determined with pseudo-experiments
- In each pseudo-experiment, we construct SM “running” templates which account for Poisson fluctuations and Gaussian uncertainties of each SM process
- From each “running” template we randomly generate a distribution with the same number of entries of the data
- We then compare this distribution to the nominal SM template and derive δ

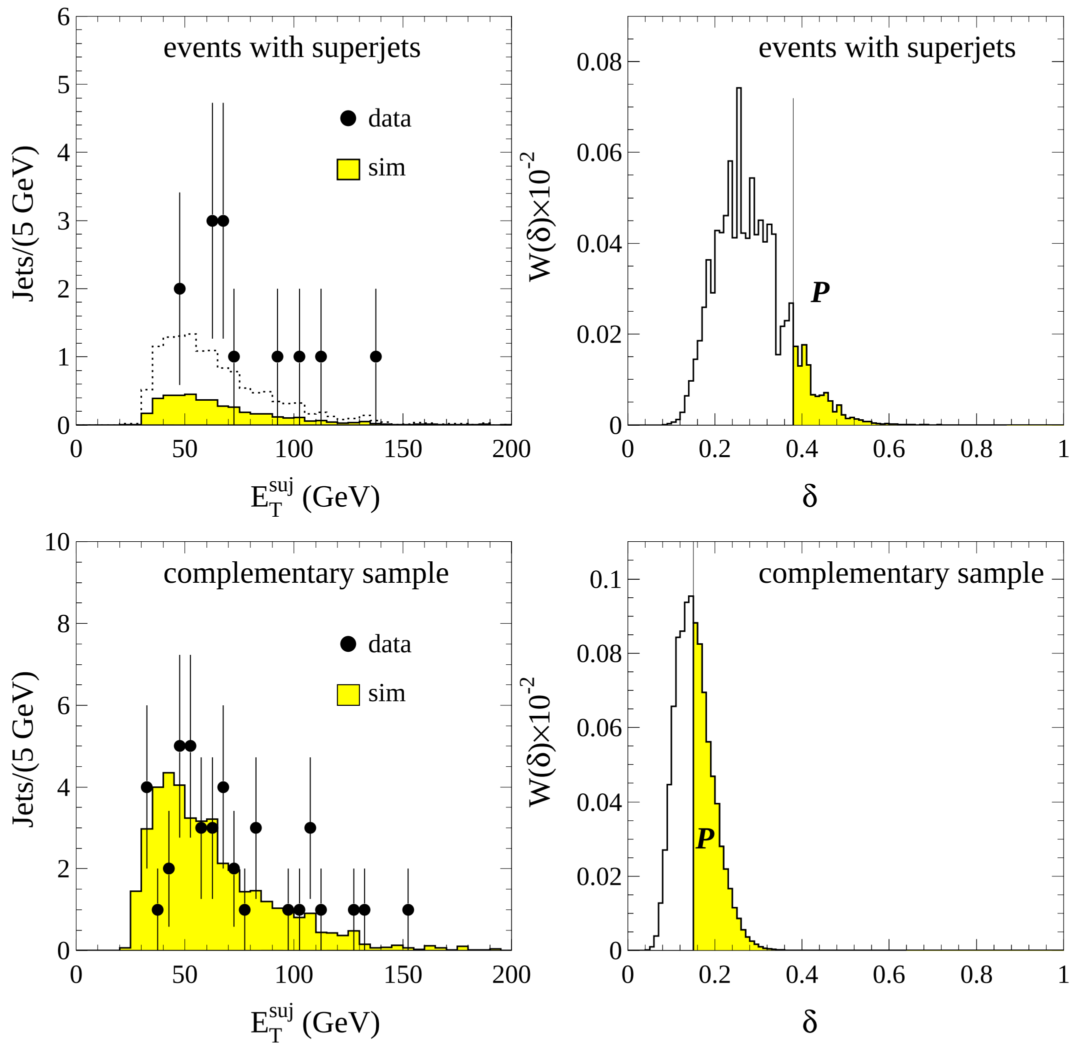


Primary lepton

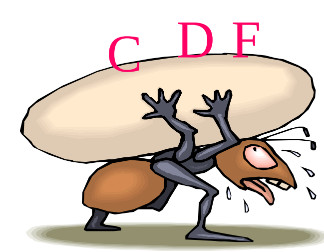




Superjet

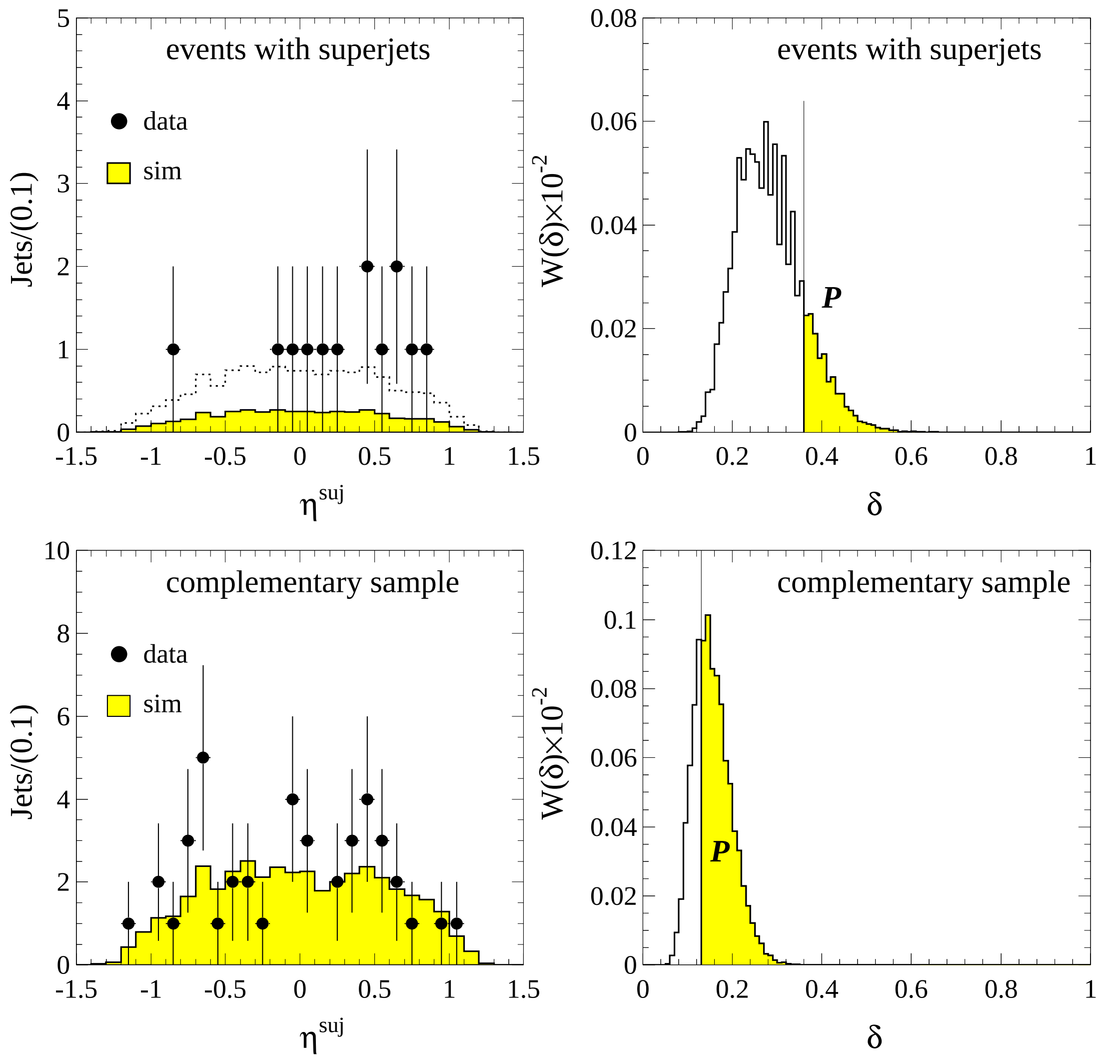


events with superjets
data
sim
Jets/(5 GeV)
E_T^{suj} (GeV)
events with superjets
P
$W(\delta)\times10^{-2}$
δ
complementary sample
data
sim
Jets/(5 GeV)
E_T^{suj} (GeV)
complementary sample
P
$W(\delta)\times10^{-2}$
δ

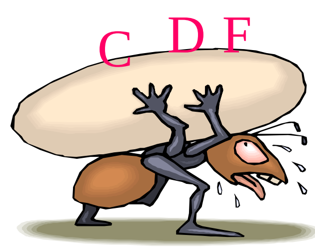


C D F

Superjet

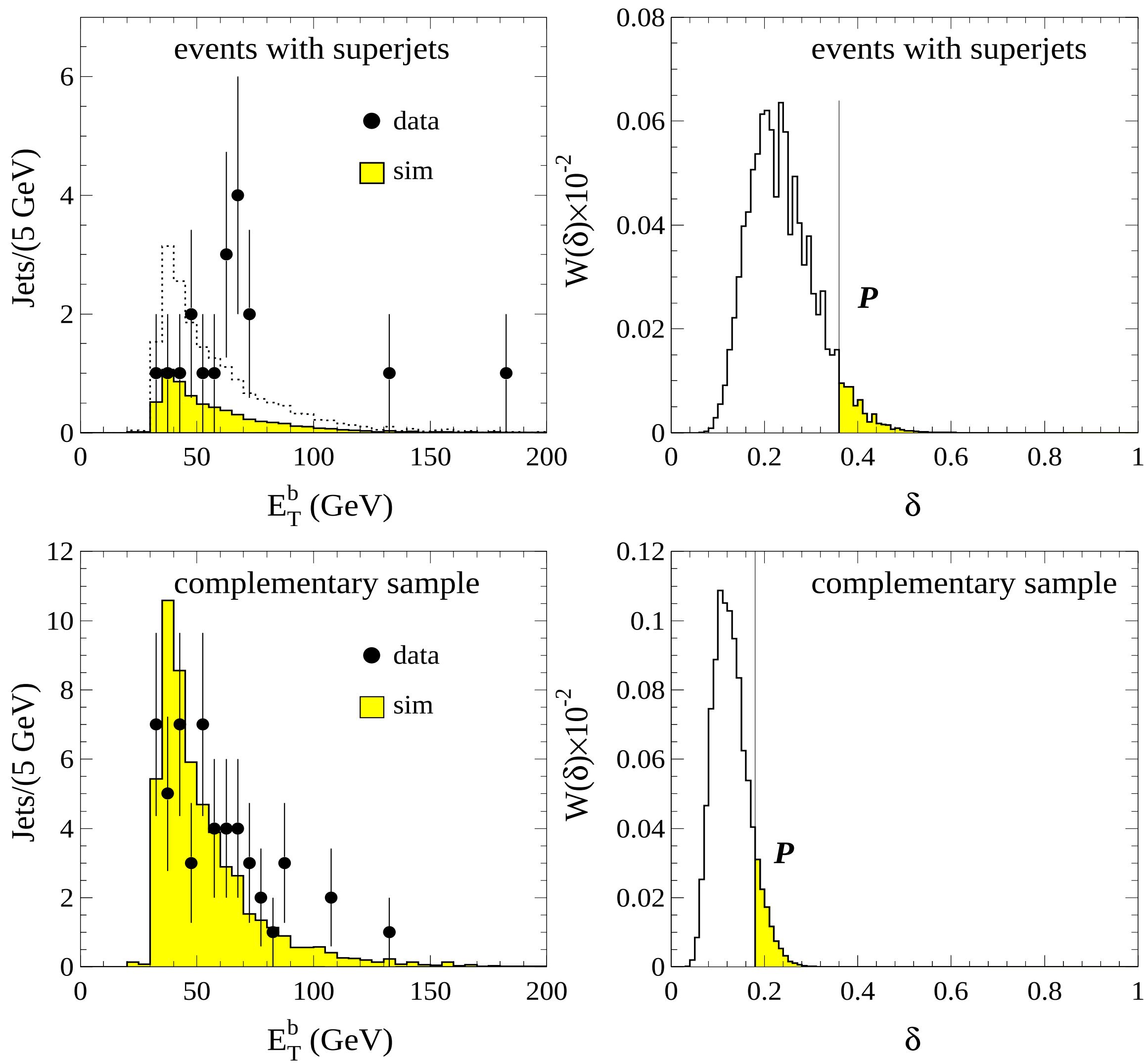


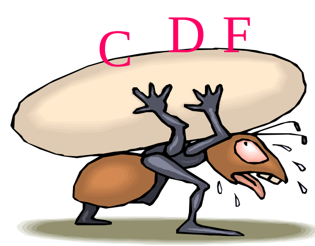
events with superjets
data
sim
Jets/(0.1)
η^{suj}
events with superjets
$W(\delta)\times10^{-2}$
P
δ
complementary sample
data
sim
Jets/(0.1)
η^{suj}
complementary sample
$W(\delta)\times10^{-2}$
P
δ



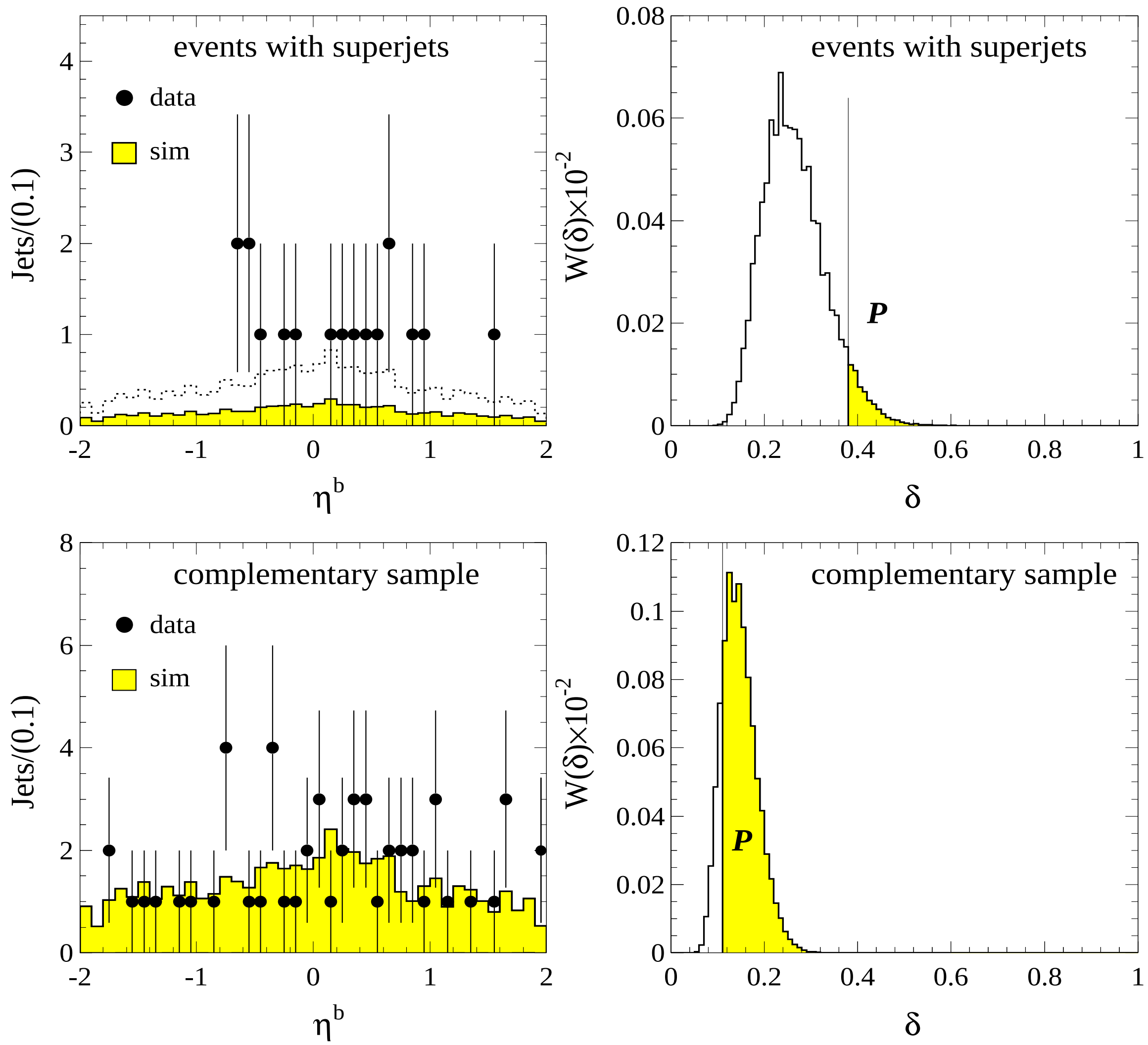
C D F

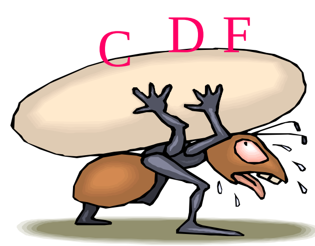
Additional jets (b-jets)



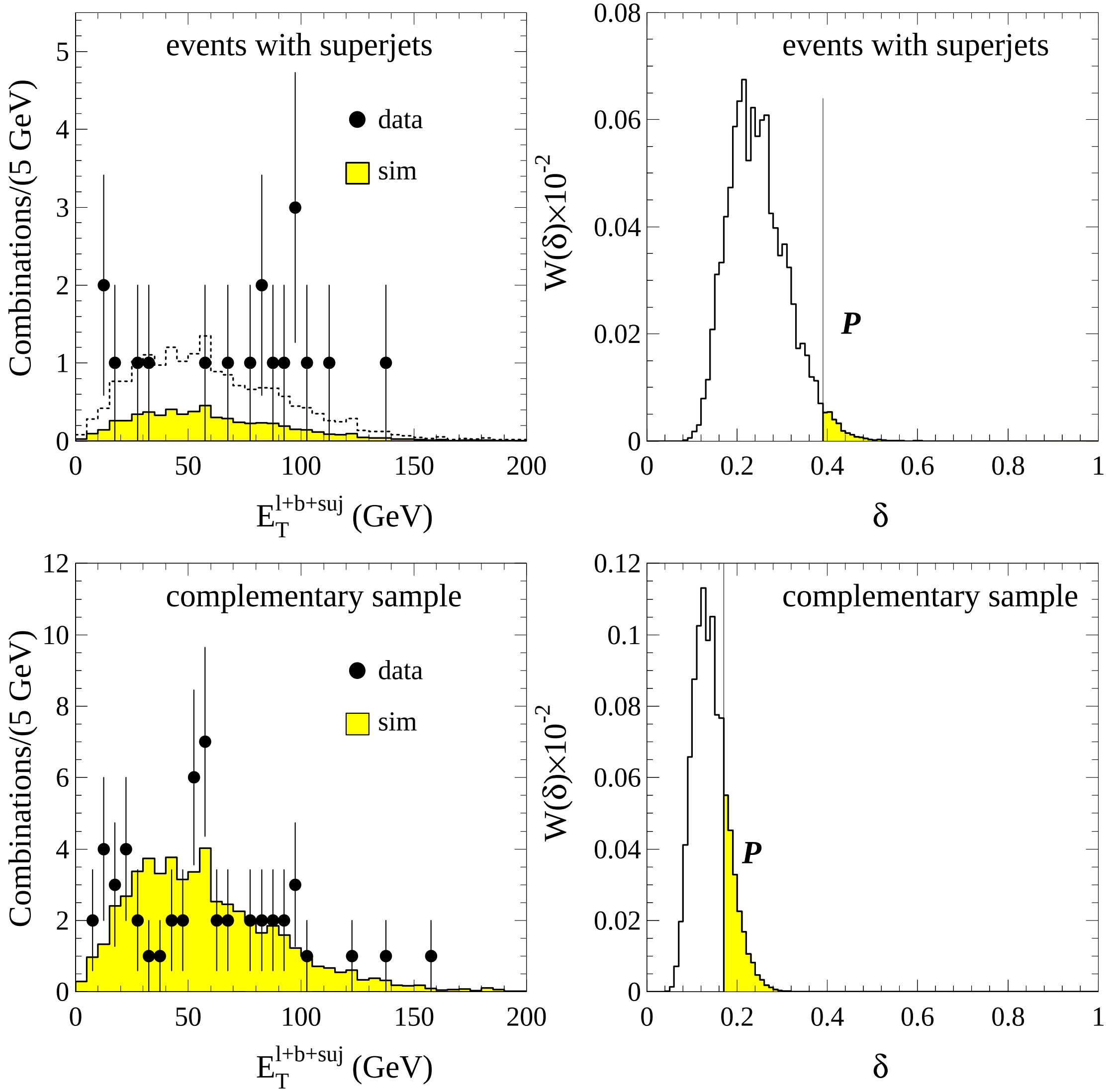


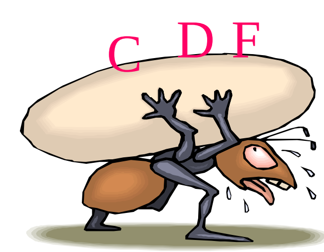
Additional jets (b-jets)



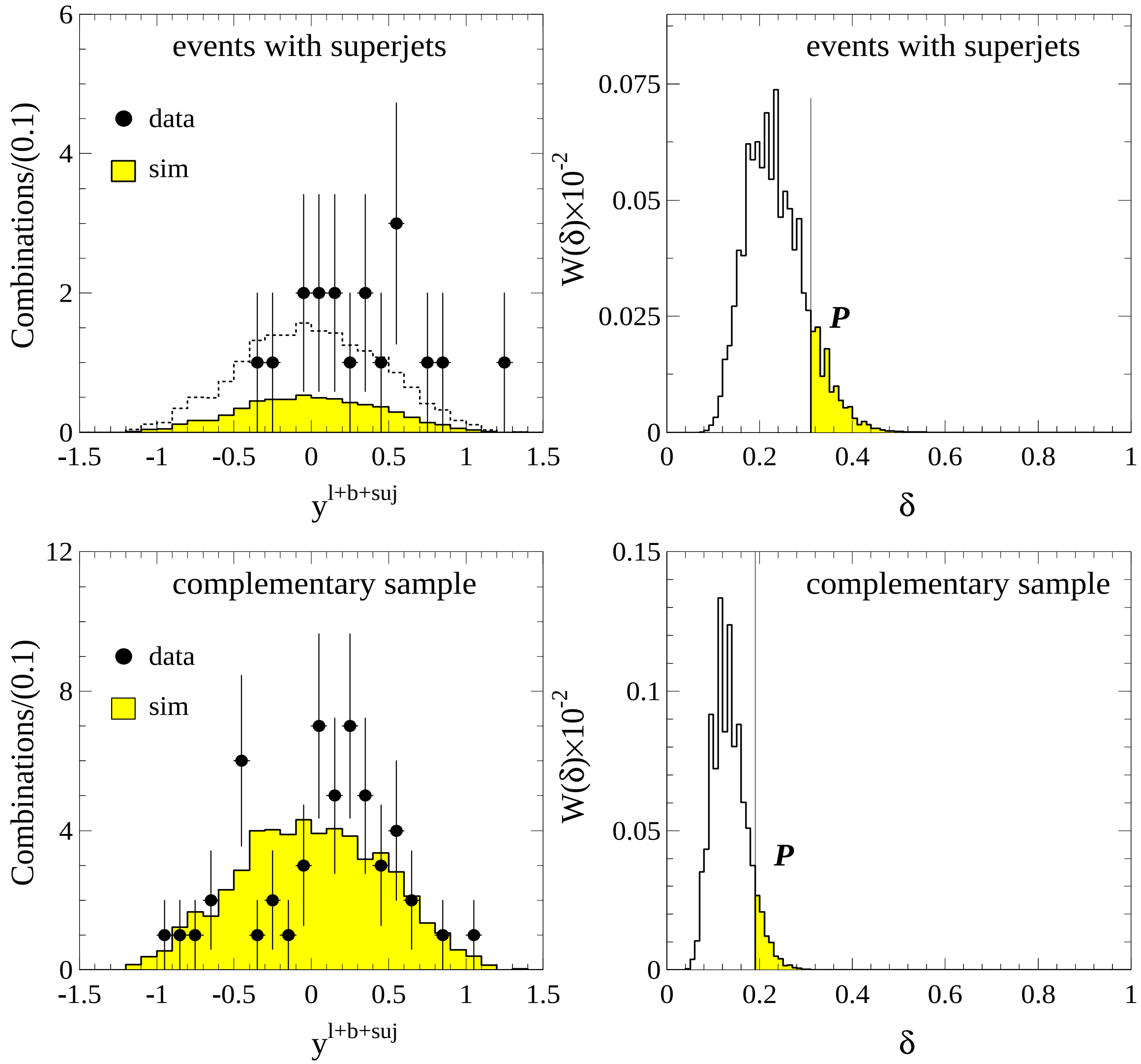


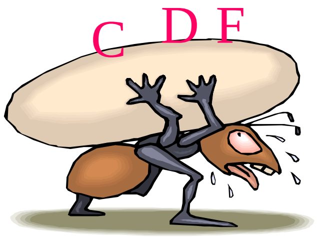
$\not{E}_T$


events with superjets
data
sim
Combinations/(5 GeV)
$E_T^{l+b+suj}$ (GeV)
events with superjets
$W(\delta)\times10^{-2}$
P
δ
complementary sample
data
sim
Combinations/(5 GeV)
$E_T^{l+b+suj}$ (GeV)
complementary sample
$W(\delta)\times10^{-2}$
P
δ

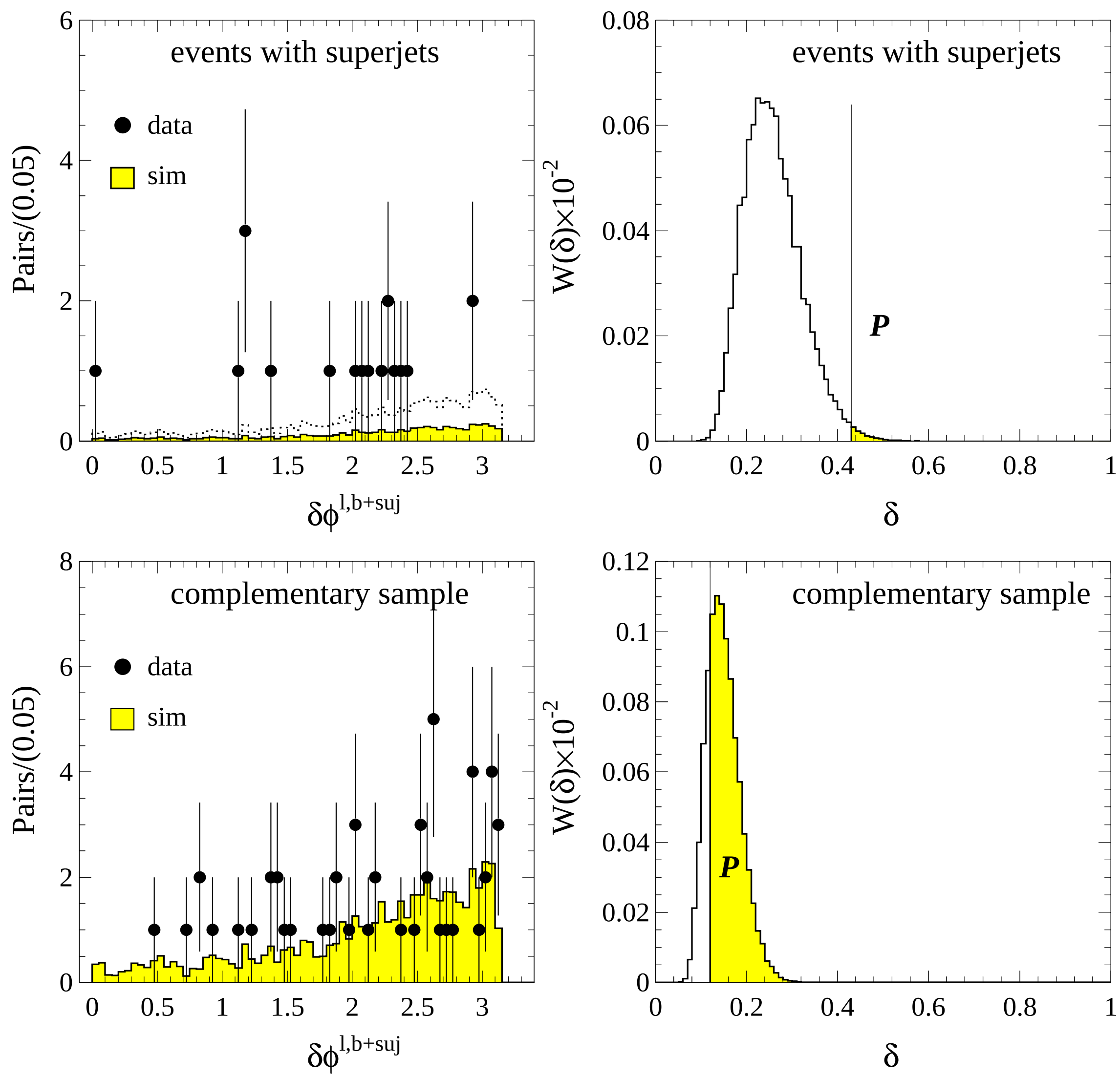

C D F

Longitudinal $\not{E}$

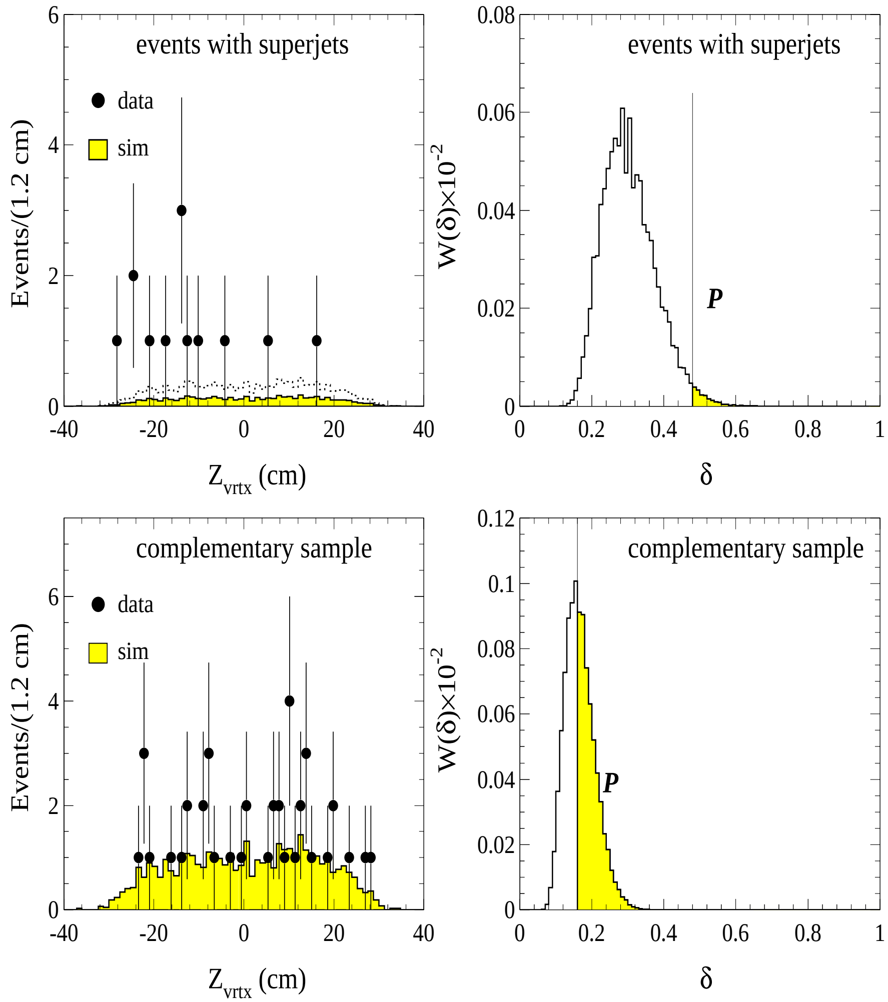




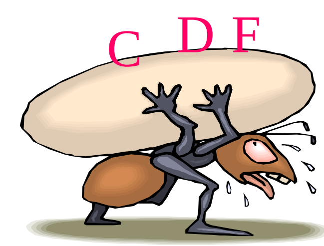
Azimuthal angle (*l*,*W*)



Event vertex

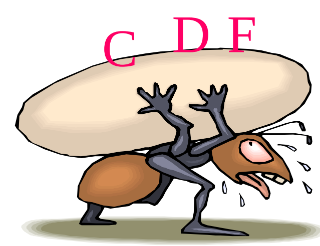


- The binomial probabilities of observing an equal or larger asymmetry in the event vertex and the η distributions are $P=1.1 \bullet 10^{-2}$, $P=13 \bullet 10^{-2}$, $P=4.6 \bullet 10^{-2}$, $P=24 \bullet 10^{-2}$, and $P=1.5 \bullet 10^{-2}$ (3.8 σ effect).
  - we know of no such physics process
  - It may be a low probability statistical fluctuation
  - obscure detector effect; such asymmetries are not visible in any other control sample



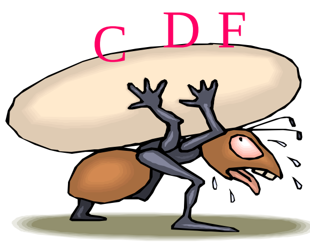
Summary of the probabilities

	Events with a superjet		Complementary sample	
Variable	δ^0	P (%)	δ^0	P (%)
E_T^l	0.47	2.6	0.14	70.9
η^l	0.54	0.10	0.12	72.7
E_T^{suj}	0.38	11.1	0.15	43.0
η^{suj}	0.36	15.2	0.13	73.4
E_T^b	0.36	6.7	0.18	8.6
η^b	0.38	6.8	0.11	80.0
$E_T^{l+b+suj}$	0.39	2.5	0.17	18.8
$y^{l+b+suj}$	0.31	13.8	0.19	7.8
$\delta\phi^{l,b+suj}$	0.43	1.0	0.12	77.9
Z_{vrtx}	0.48	1.7	0.16	50.5

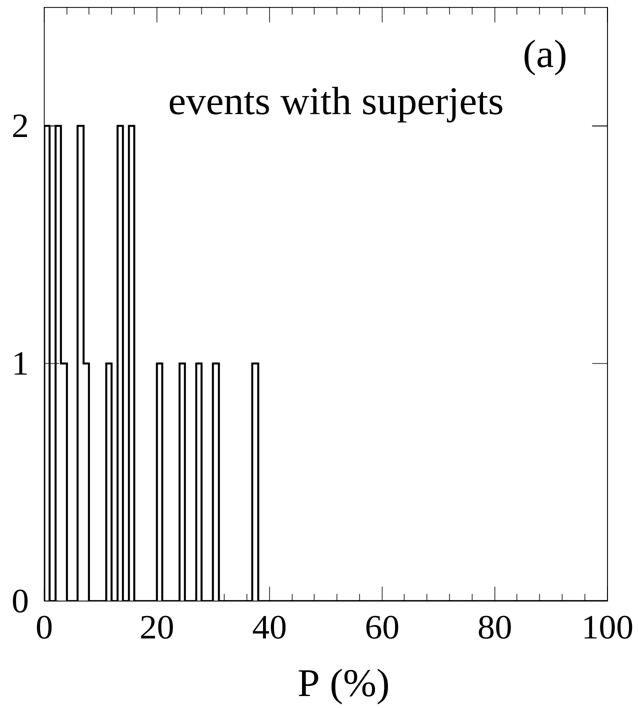


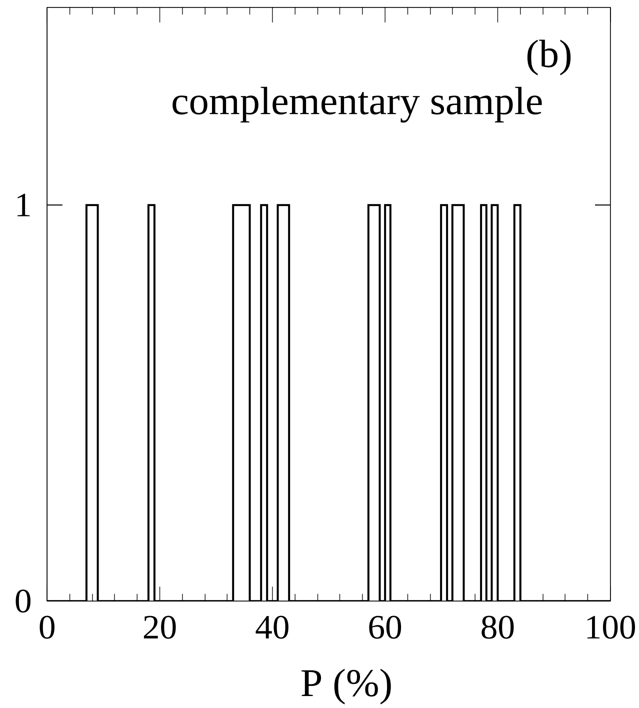
Summary of the probabilities

	Events with a superjet		Complementary sample	
Variable	δ^0	P (%)	δ^0	P (%)
$\not{E}_{\mathrm{T}}$	0.31	27.1	0.14	57.1
M_T^W	0.36	13.1	0.16	38.2
M^{b+suj}	0.36	4.0	0.12	58.9
y^{b+suj}	0.35	7.1	0.14	34.9
E_T^{b+suj}	0.28	24.0	0.10	60.1
$M^{l+b+suj}$	0.31	21.0	0.15	33.6
$\delta\theta^{b,suj}$	0.26	30.1	0.15	41.1
$\delta\phi^{b,suj}$	0.31	15.3	0.10	83.8
$\delta\theta^{l,b+suj}$	0.25	37.3	0.16	35.7



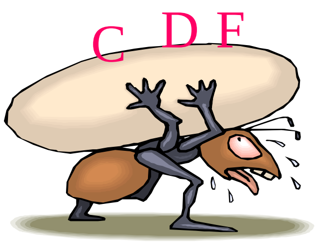
18 Kinematical variables





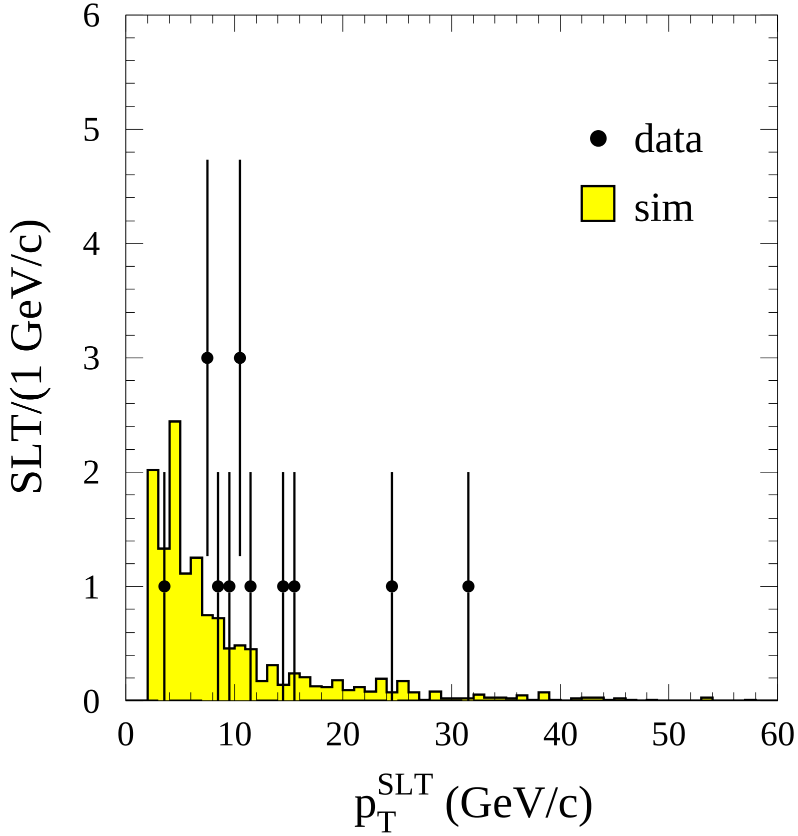
$< P >= 0.13$
RMS=0.11

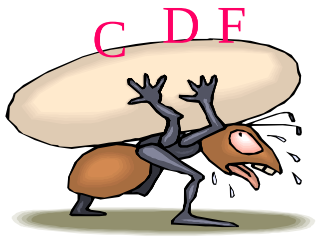
$< P >= 0.50$
RMS=0.24



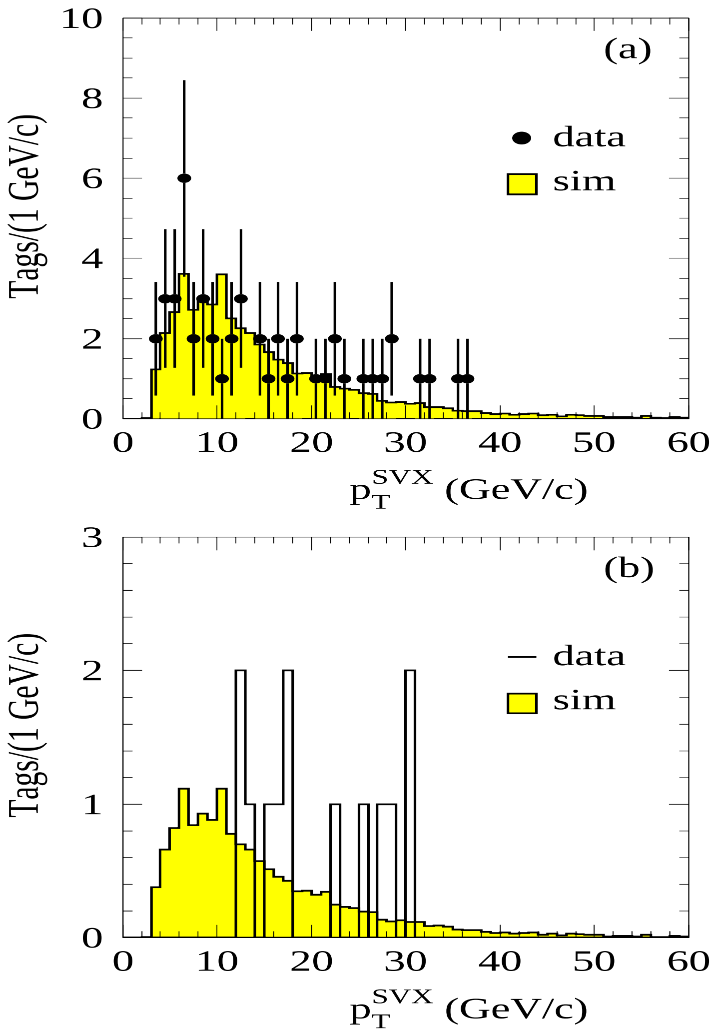
Superjet properties

- Compare to a SM simulation, in which the superjet transverse momentum distribution in each SM process has been sculpted to reproduce the data
- the usual K-S test yields $P = 9 \cdot 10^{-4}$



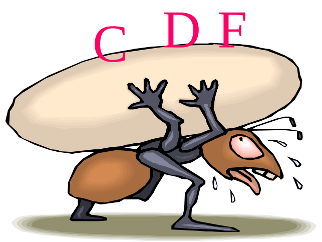


Check of the fragmentation (control sample)

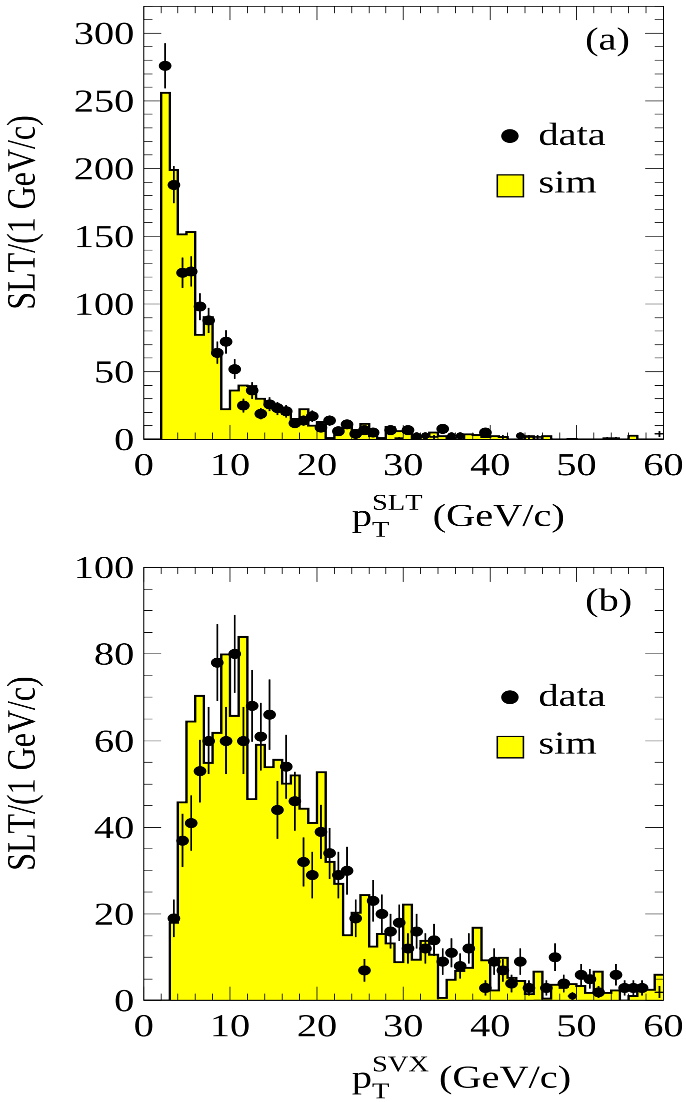


Complementary sample

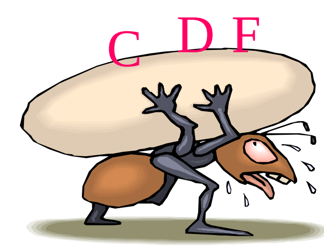
13 events



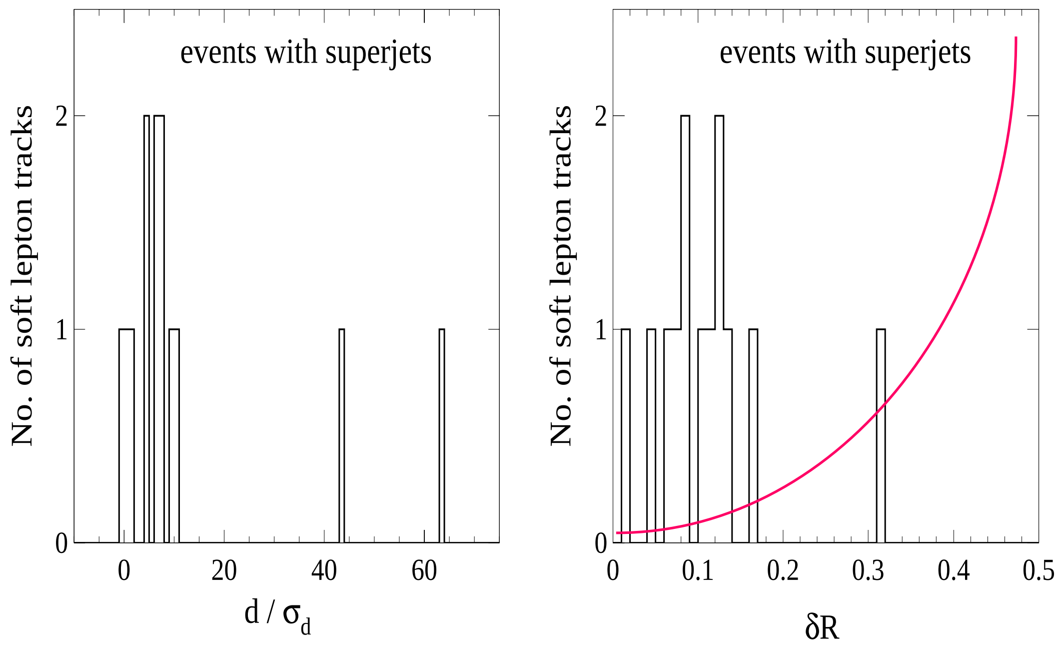
Check of the fragmentation (generic-jet data)



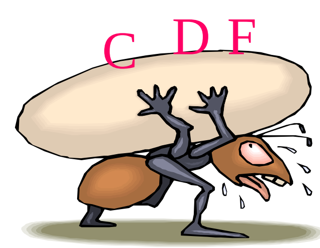
- 550,000 generic-jet events in the data and in the Herwig simulation (JET20, JET50, and JET100).
  - 1324 supertags in the data
  - 1342 simulated supertags



Superjet properties

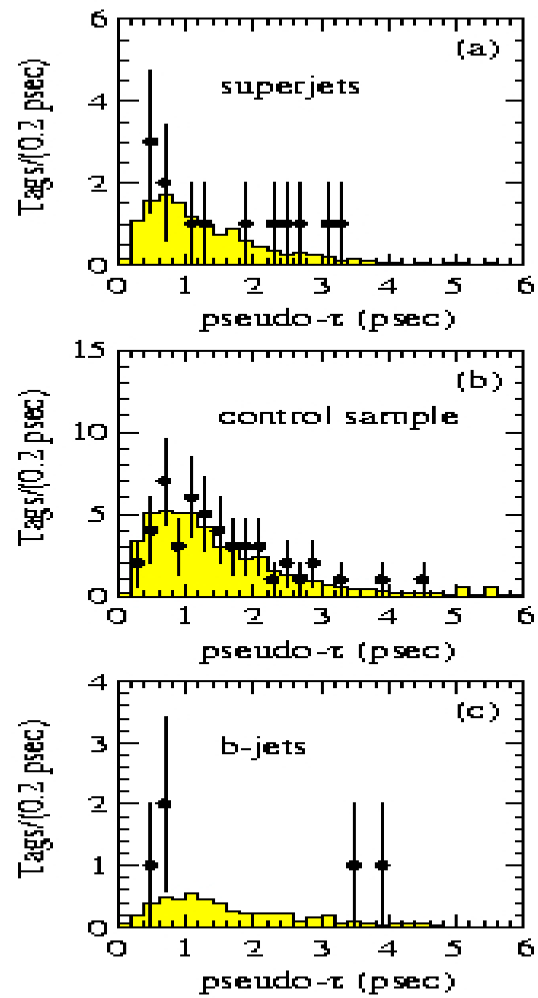


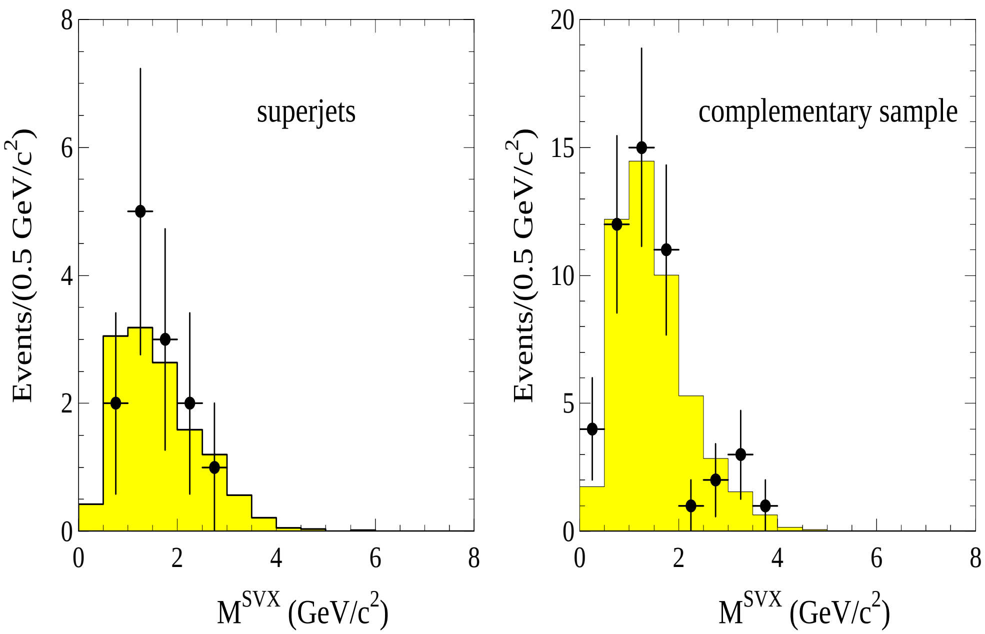
- Soft leptons
- not prompt
  - in 8 cases are part of the SECVTX tag
- emitted along the superjet axis



Superjet lifetime

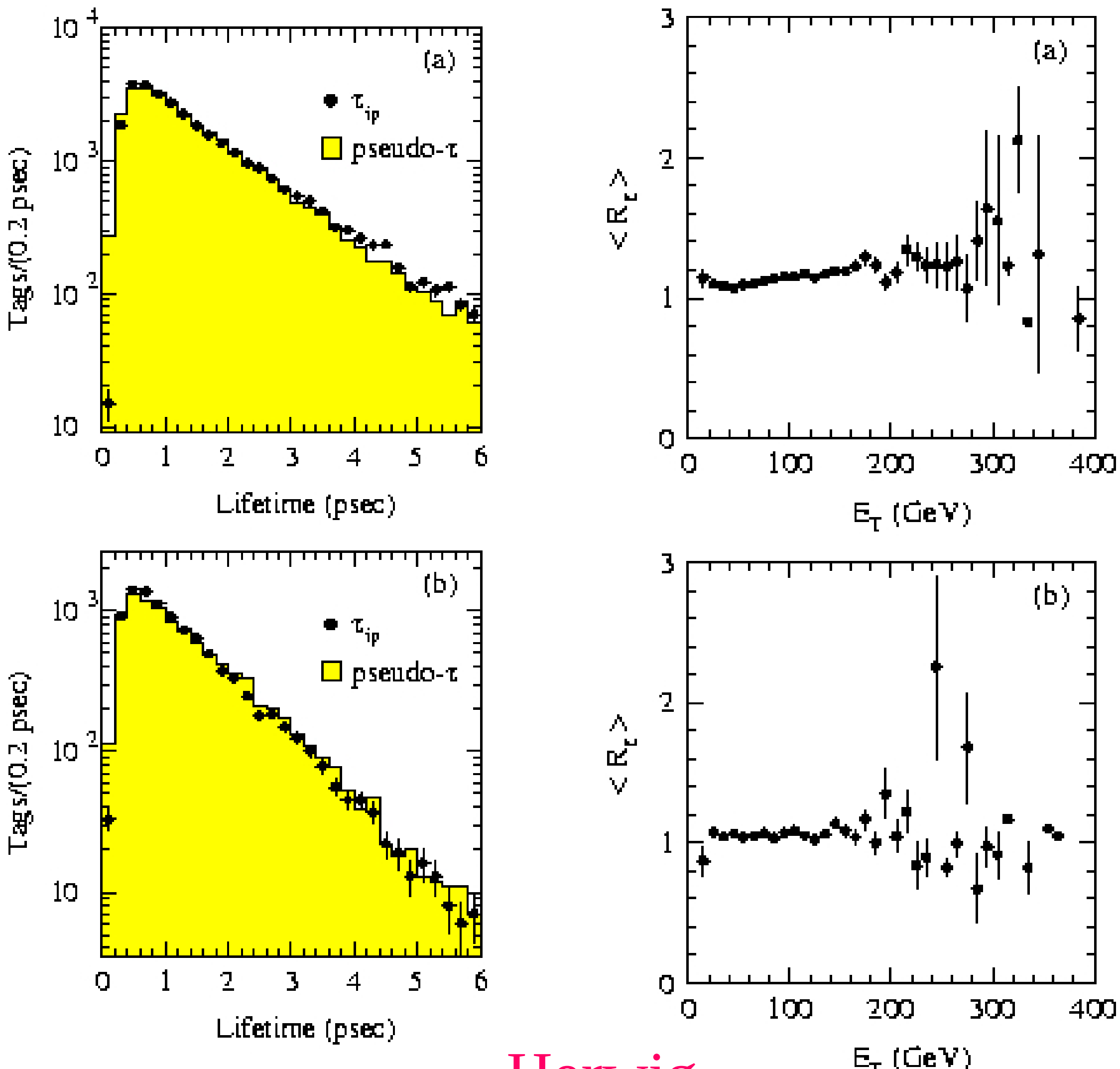
- Lifetime
  - $\text{Pseudo-}\tau = \frac{L_{XY} \bullet M}{c \bullet P_T}$
  - It is related to the lifetime by a kinematical factor K, which accounts for lost neutrals. For b and c-hadrons, K~ 1.1





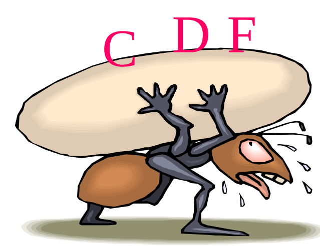
Another method to measure the lifetime

Generic-jet data

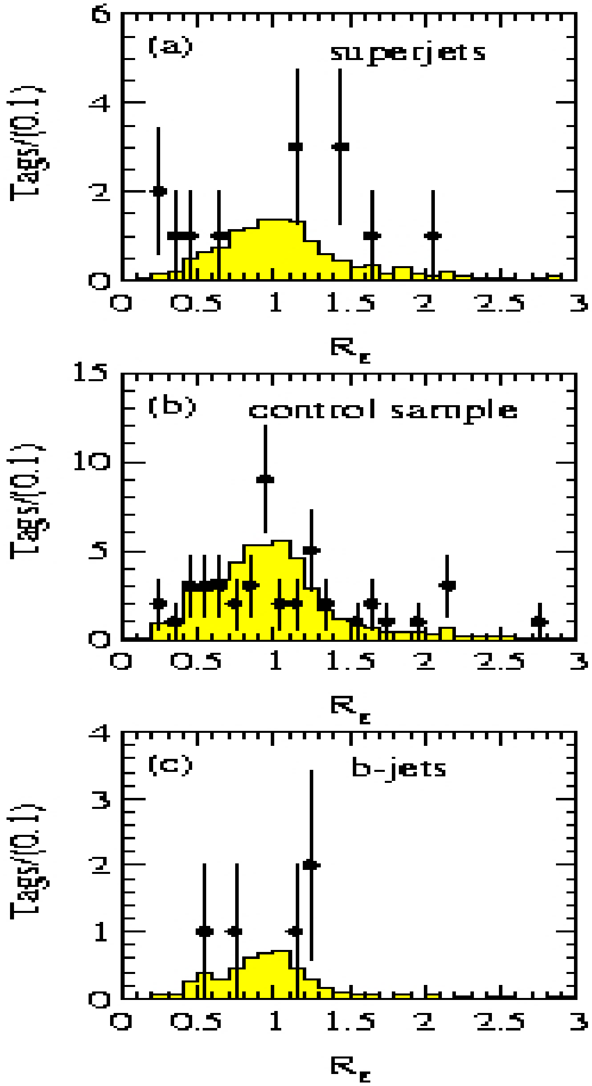


Herwig

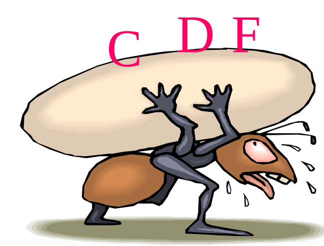
- ❖ $\tau_{ip} = \frac{4\langle d_0\rangle}{c\pi}$
- ❖ $R_\tau = \tau_{ip}/\text{pseudo-}\tau$ tests the kinematical factor K



Superjet lifetime

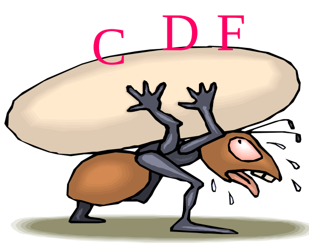


- compare data to the SM simulation with a K-S test
  - control. s. P=0.35
  - b-jets P=0.47
  - superjets: P=0.05



Are superjet events the tail of a large not understood background ?

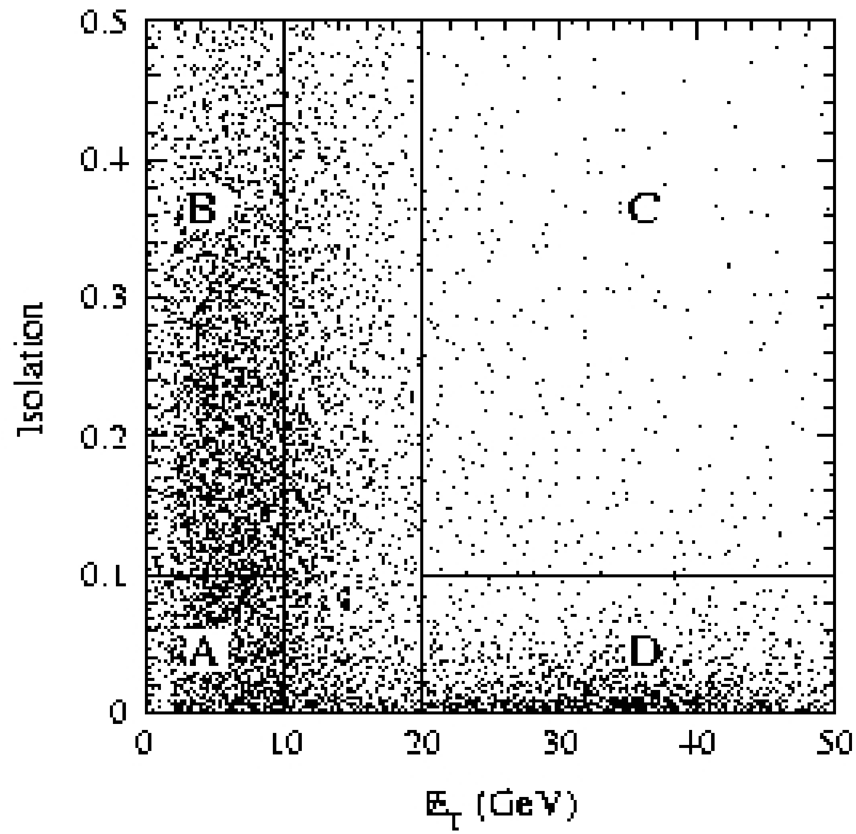
- The selection criteria used in this analysis are the same used to find the top quark
- The original high p_T inclusive lepton data set consists of 82000 events, mostly due to multi-jet production with one jet faking a lepton (it includes a small amount of $b\bar{b}$ and $c\bar{c}$ production in which the primary lepton comes from semileptonic decays of heavy quarks)
- The requirements on $\not\!\!E_T$, and the primary lepton p_T and I produce a sample of 11000 almost pure *W*+jet events

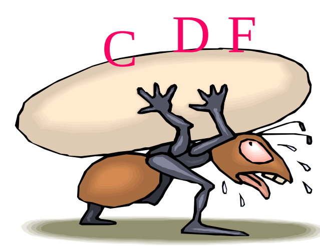


Non-*W* events ?

- The 13 events with a superjet are hardly consistent with W production
- non-W is estimated from the data (A•C/B)
  - 84.6±3.2 before tagging
  - 0.13±0.03 with a supertag due to *bb* and *cc* production
- is the *bb* and *cc* production estimated correctly ?
- What is the ratio of supertags to SECVTX tags in A, B, and C ?

data

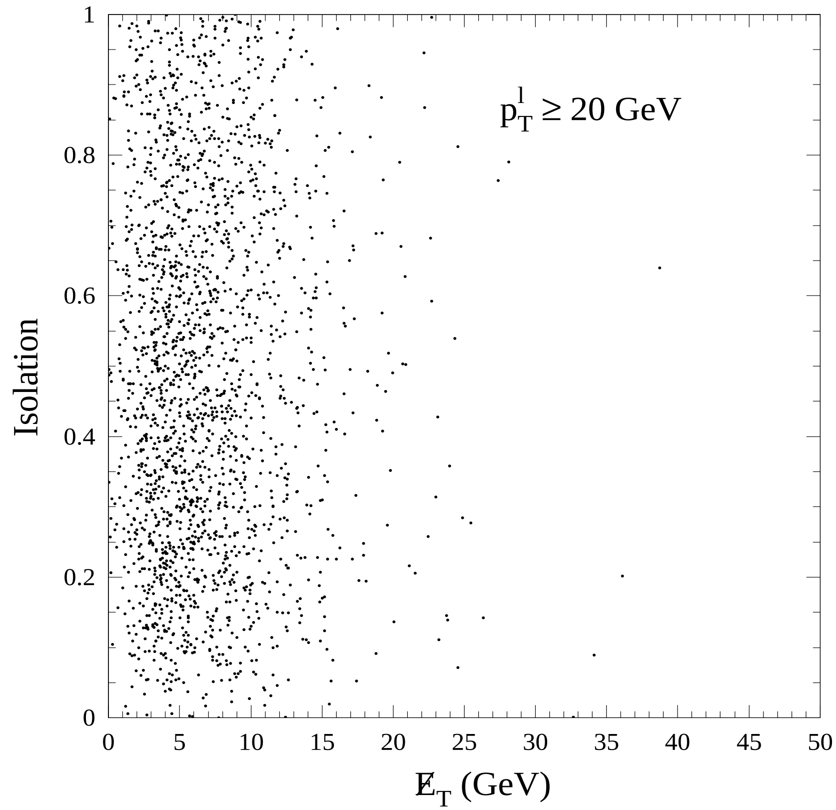


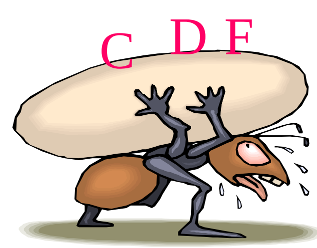


Low-p_T inclusive lepton sample

- Events with one lepton
  - $|\eta|<1$ and $P_T > 8$ GeV/c
  - One or more additional jets
- Herwig simulation (opt. 1500) normalized to the same number of SECVTX and JPB tags of the data
- In the simulation there are 9 events with
  - 1 lepton with $P_T > 20$ GeV/c, $I<0.1$
  - $\not{E}_T > 20$ GeV
  - 2,3 jets
- No supertags

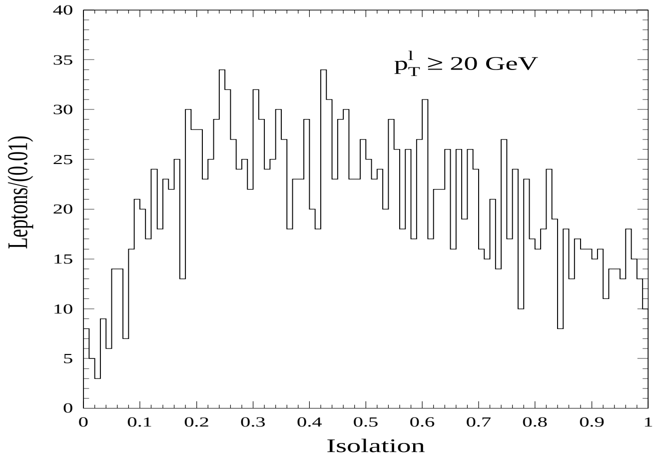
simulation



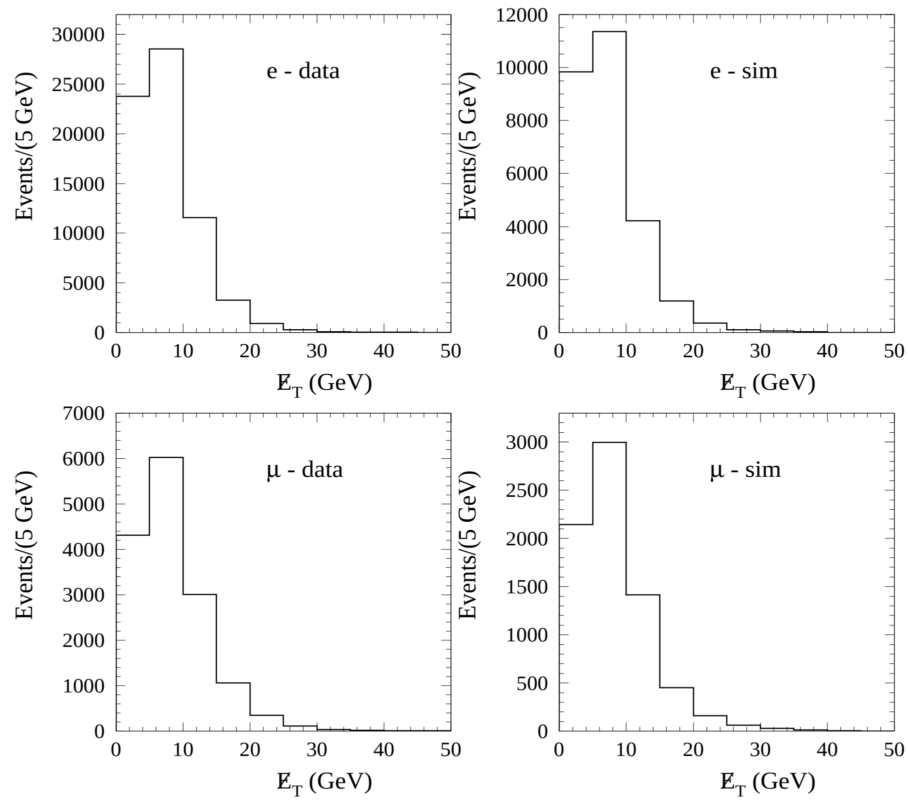


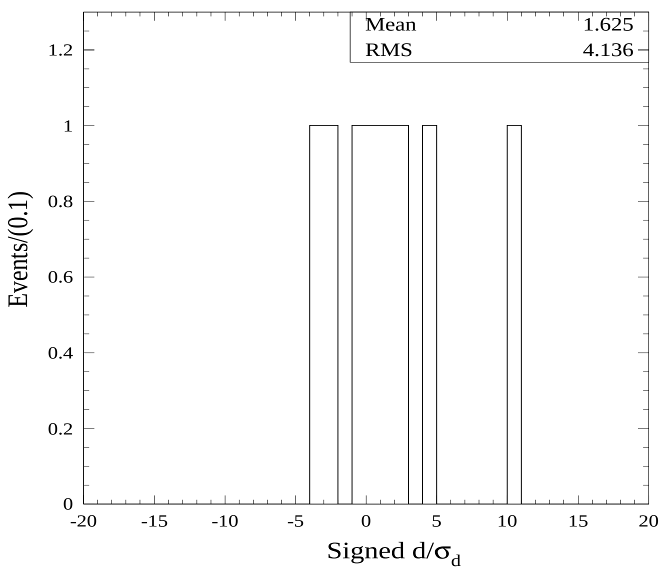
Herwig simulation

isolation

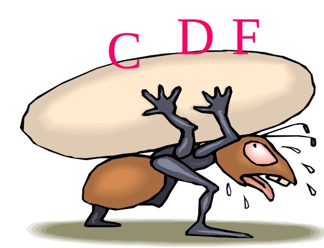


$\not{E}_T$

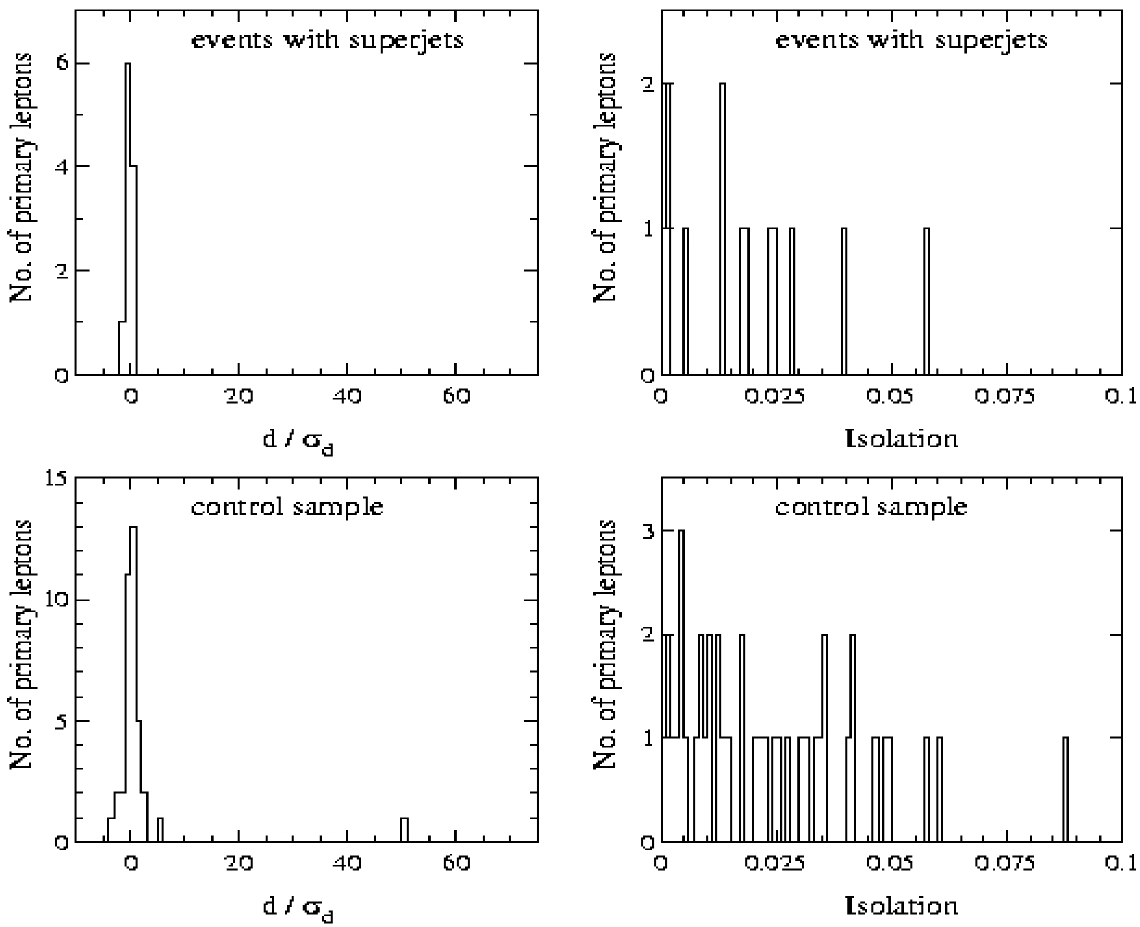




impact parameter significance



Studies of the primary lepton



Studies of the primary lepton

Tag type	1 jet	2 jets	3 jets	$\geq$ 4 jets
SECVTX	168	21	7	6
SECVTX+SLT	12	1	0	0
Prediction	10.2±1.3	1.2±0.2	0.5±0.2	0.5±0.2

$\not{E}_T$ < 20 GeV, and I <0.1

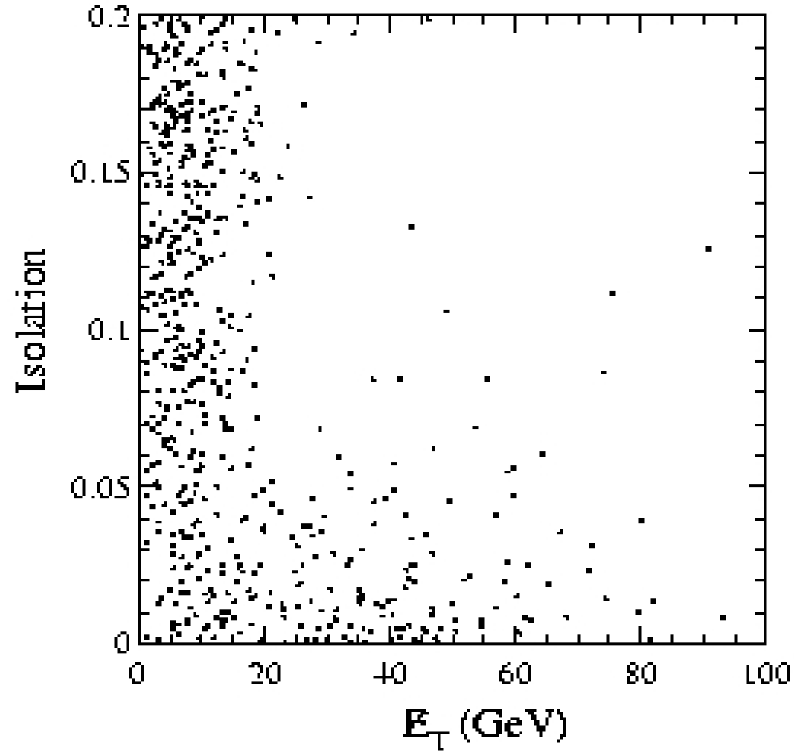
$\not{E}_T \leq 20$ GeV

Tag type	1 jet	2 jet	3 jet	$\geq$ 4 jet
SECVTX	220	33	10	2
SECVTX+SLT	17	4	2	1

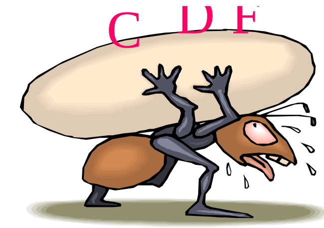
$\not{E}_T \geq 20$ GeV

Tag type	1 jet	2 jet	3 jet	$\geq$ 4 jet
SECVTX	8	3	5	0
SECVTX+SLT	2	0	1	0
Prediction	0.6 ± 0.1	0.4 ± 0.2	1.0± 0.7	0

data

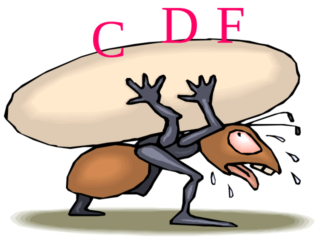


0.1 < I <0.2



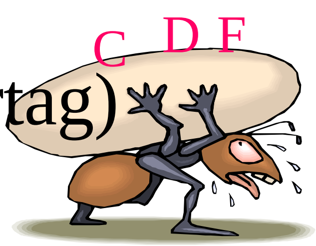
Studies of the primary lepton

- The p_T of the primary leptons has a cluster of events at the 20 GeV/c threshold
- lower the theshold in the high-p_T lepton sample to 18 GeV/c
  - No events with a superjet found
- Search in the low-p_T lepton sample with a 10 GeV/c threshold
  - This sample is prescaled by a factor of about 1.3
  - Smaller integrated luminosity; it corresponds to 8 of the 13 events with a superjet
  - We find 6 of the 8 events, and an additional one containing a primary lepton with p_T=17.7 GeV/c

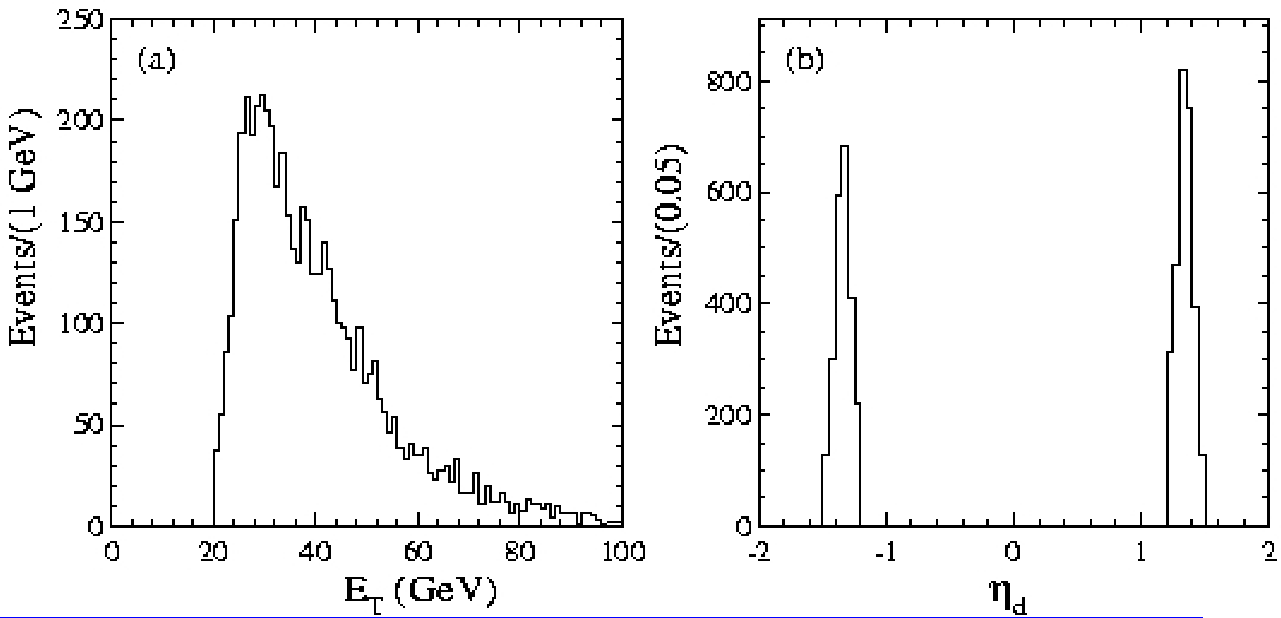


Studies of the primary lepton

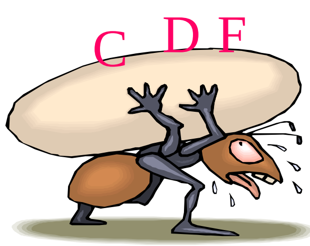
- We require that primary leptons fired the L2 trigger
  - For muons the L2 trigger efficiency is about 70%
- If superjets are for real, we expect to have lost 2 muon events with a superjet because of this request
  - as seen before, 85% of the superjets contain a soft muon above L2 trigger theshold
  - we then expect to recover 1 or 2 *W*+2,3 jet events with a supertag removing the L2 request
- If superjet events are a fluctuation of SM processes, we expect to recover
  - 3 *W*+1 jet, 1.1 *W*+2,3 jet events (0.09 and 0.08 events with a supertag, respectively)
- We find 3 *W*+1 jet events with no supertags, and 1 *W*+2 jet plus 1 W+3 jet event (both with a supertag)



Plug electrons



Source	$W+1$ jet	$W+2$ jet	$W+3$ jet	$W+\geq 4$ jet
Data	1245	243	52	11
SECVTX tags	15	3	1	1
SECVTX+SLT tags	3	2	0	0
S. M. prediction	0.9 ± 0.1	0.24 ± 0.03	0.10 ± 0.02	0.10 ± 0.03



Impact of the new events

- We find 4 additional events when 0.42 ± 0.06 are expected
- In 13 out of 17 events with a superjet, the charges of the primary lepton and the soft lepton are opposite
- No SM process has this kind of charge correlation
  - the probability of an equal or larger fluctuation is $P = 0.024$

C D F

Conclusions

- We have studied the heavy flavor contents of jets produced in association with *W* bosons
- We generally find good agreement between observed and predicted rates of SECVTX (displaced vertex) and SLT (soft lepton) tags
- An exception is the number of events with a superjet (13 observed and 4.4±0.6 expected)
- A detailed examination of the kinematical properties of these events shows that they are statistically difficult to reconcile with a simulation of SM processes. The same simulation models well a complementary sample of W+jet events and larger samples of generic-jet data
- Obscure detector effects can never be ruled out, but extensive studies of these events and investigation of larger statistics data sets have not revealed any effects which might indicate detector problems or simulation deficiencies
- We are not aware of any model for new physics which incorporates the production and decay properties necessary to explain all features of these events