

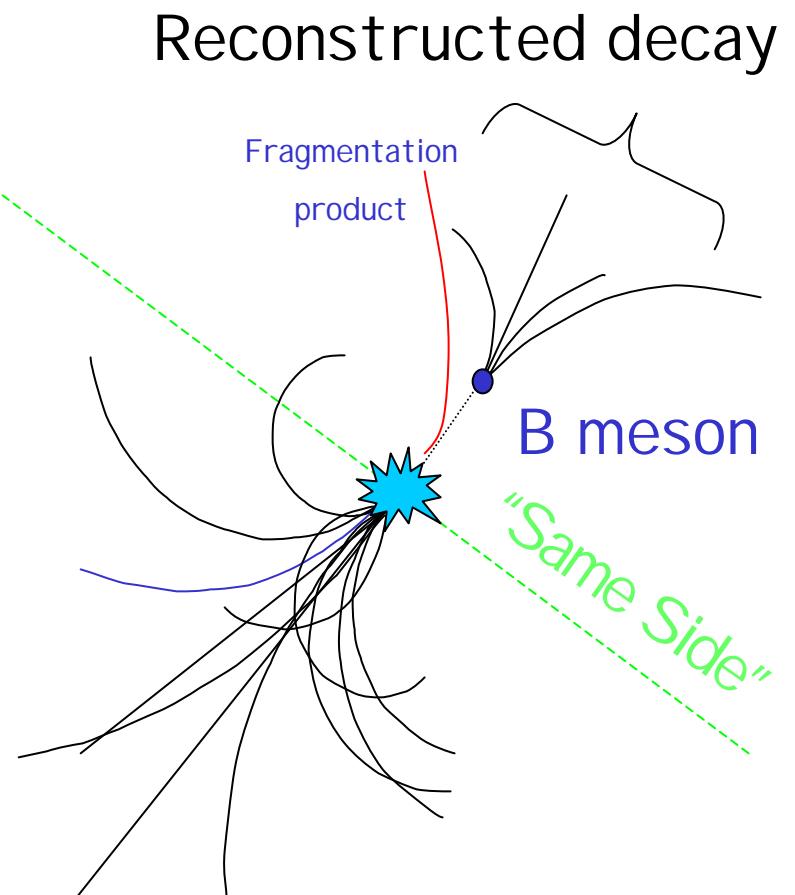
Flavor Tagging

what is going on, **is there room to contribute?**
Where?

Alessandro Cerri, LBNL



$B_s \rightarrow D_s^- \pi^+ : \epsilon D^2$	
Same-Side Kaon	4.2%
μ tag	1.0%
e tag	0.7%
Jet Charge	3.0%
Opp.-Side Kaon	2.4%
Total (correl. small)	11.3%



Tagging

Barry Wicklund, Matthew Jones, Denys Usynin, Joe Kroll, Vivek Tiwari, Gavril Giurgiu, Guillermo Gomez-Ceballos, Sasha Rakitin, Ilya Kravchenko, Ivan Vila, Alberto Ruiz, Jonatan Piedra, Marcin Wolter, Nuno Leonardo, Tania Moulik



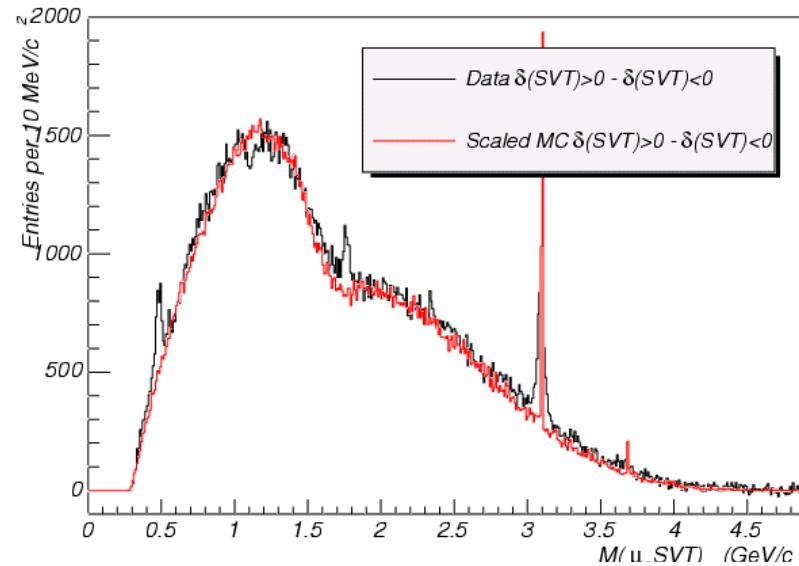
- To some degree, each of these can be developed and checked on the **semileptonic** sample:
 - Soft muon
 - Soft electron
 - Jet Charge
 - OSK
 - Same Side
- We have blessed results for:
 - Soft muon
 - SST
- SeT is still almost there
- JQT: optimized cuts. Still struggling with MC
- First tagger tests ran on $J/\Psi K$ and $D^0\pi$
- OSK still very naïve

muon	raw $D, \%$	$\epsilon, \%$
IMU	43 ± 9	0.18 ± 0.02
CMX	22 ± 5	0.59 ± 0.03
CMU only	13 ± 4	1.20 ± 0.05
CMP only	20 ± 7	0.36 ± 0.03
CMUP	27 ± 5	0.62 ± 0.03
Any	20.2 ± 2.5	2.92 ± 0.07

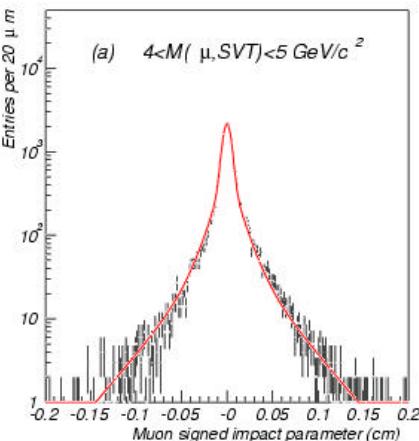
Sample Composition

M. Jones, J. Kroll, A. Wicklund, D. Usynin

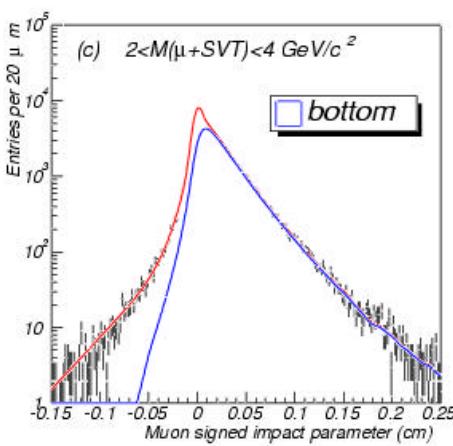
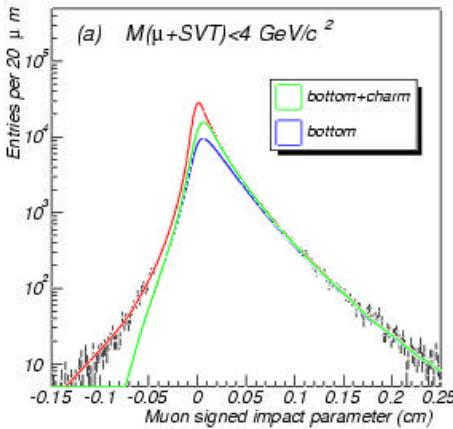
- Starting point for tagging studies: know your sample (how much b? \rightarrow effective dilution)
- Use signed lepton d_0
- Take cc and bb model from MC templates
- Residual background model [from \[4,5\]](#) GeV region + prompt component
- Simultaneous fit:



component	$\mu + SVT$ $0 - 4 \text{ GeV}/c^2$
$b \rightarrow \ell$	$3290 \pm 21 \text{ pb}$
$c \rightarrow \ell$	1234 ± 23
prompt	1018 ± 50
symmetric	382 ± 26



$$D_{\text{sub}} = \begin{cases} 0.641 \pm 0.002 \text{ (stat)}^{+0.014}_{-0.023} \text{ (syst)} & \mu + SVT \\ 0.641 \pm 0.002 \text{ (stat)}^{+0.022}_{-0.037} \text{ (syst)} & e + SVT \end{cases}$$



Prompt charm pollution



- Lifetime riddle tells us that:
 - Background is responsible
 - It has a resonant structure in the D mass
 - Has significantly shorter lifetime source than B
 - Not unlikely that I is fake a significant fraction of the times
 - As is, **not** good for tagging!
 - We can remove this background in a very easy way:

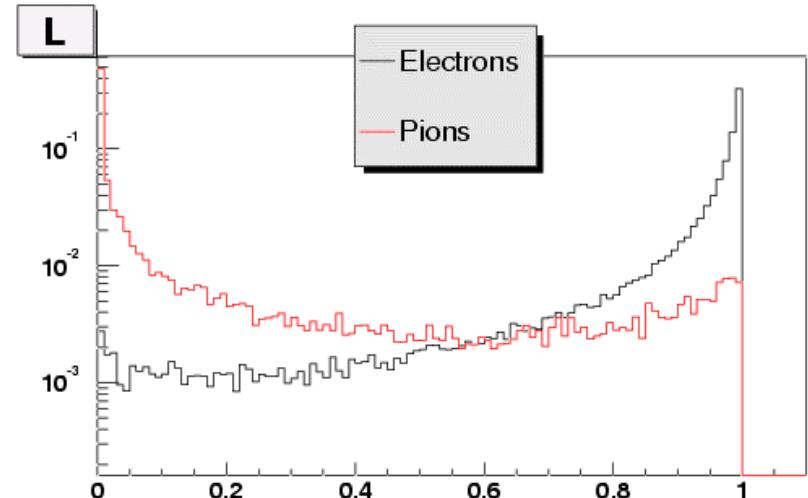
$L_{xy} > 500\mu\text{m}$



SeT

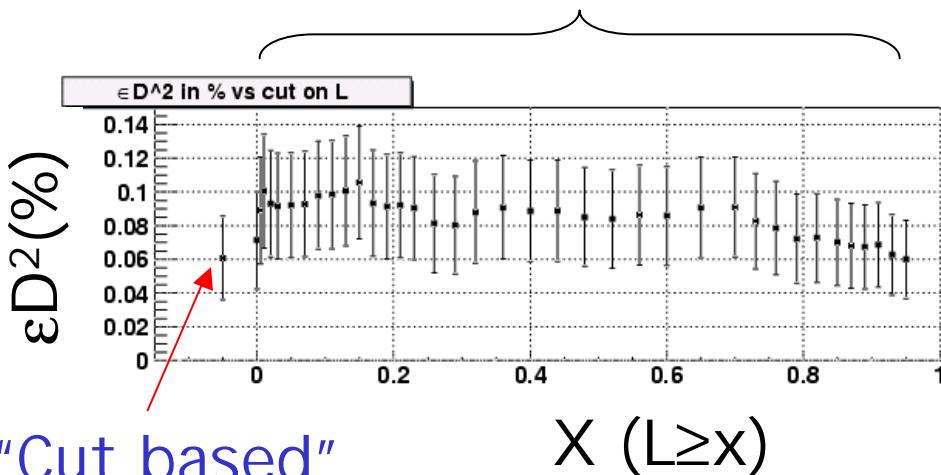
V. Tiwari, G. Giurgiu, M. Paulini, J. Russ, B. Wicklund, T. Moulik

- Two approaches so far:
 - Cut on electron ID
 - Build a likelihood and weight
- Improve efficiency
- Exploit the full rejection power of eid



- Efficiency for Electrons and Fakes

Likelihood based, vs L lower cut



$L \geq$	$\epsilon_\pi(\%)$	$\epsilon_e(\%)$	ϵ_e/ϵ_π
0.01	48	100	2.1
0.05	38	99	2.6
0.10	33	99	3.0
0.15	28	98	3.5
0.20	26	97	3.7
0.50	16	94	5.9
0.70	11	89	8.1
0.90	5	75	15.0

Table 1: Efficiencies of the cuts on L for pure electrons and pions.

L-based SeT performance (cont'd)

V. Tiwari, G. Giurgiu, M. Paulini, J. Russ, B. Wicklund, T. Moulik

- Binning in p_T^{rel}

- Choose the cut on likelihood, $L \geq 0.15$ and bin in p_T^{rel} .

p_T^{rel}	OS/SS	$\epsilon(\%)$	$D(\%)$	$\epsilon D^2(\%)$
0.0	91/59	0.251 ± 0.025	21.33 ± 10.03	0.011 ± 0.011
0.0-0.4	55/22	0.129 ± 0.021	42.86 ± 17.26	0.024 ± 0.019
0.4-0.7	70/66	0.227 ± 0.025	02.94 ± 10.86	0.000 ± 0.001
0.7-1.0	78/60	0.230 ± 0.024	13.04 ± 10.70	0.004 ± 0.006
1.0-1.5	122/56	0.297 ± 0.024	37.08 ± 09.07	0.041 ± 0.020
1.5-2.0	77/28	0.175 ± 0.020	46.67 ± 10.64	0.038 ± 0.018
> 2.0	61/28	0.149 ± 0.020	37.08 ± 12.85	0.020 ± 0.014
Sum	554/319	1.458 ± 0.062		0.138 ± 0.038
Avg	554/319	1.458 ± 0.062	26.92 ± 04.21	0.106 ± 0.033
Cuts	318/183	0.837 ± 0.046	26.95 ± 05.46	0.061 ± 0.025

Table 2: ϵ , D and ϵD^2 in % after binning in p_T^{rel} with the cut $L \geq 0.15$ for the 8 variables case. Also shown are the average numbers from the likelihood and the cut-based approaches without binning in p_T^{rel} .

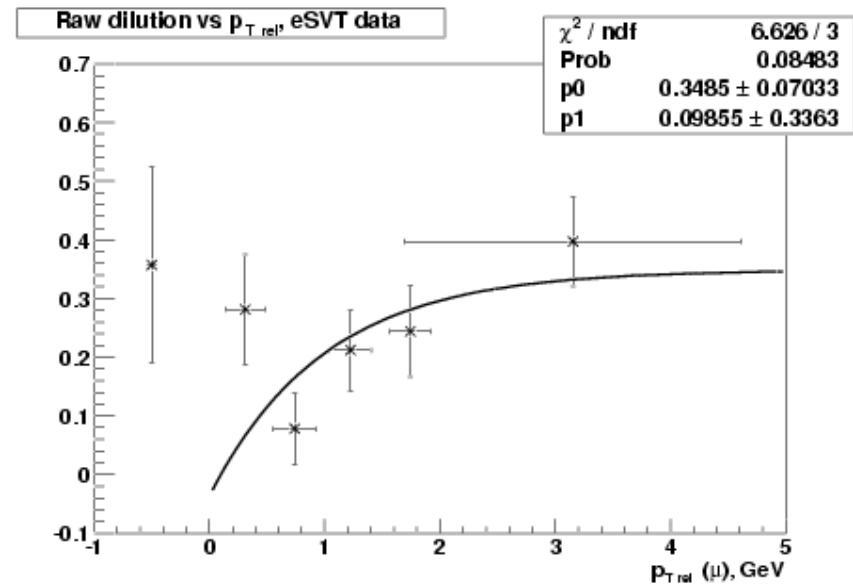
- The total raw ϵD^2 after binning in $p_T^{rel} = 0.180 \pm 0.043$.
- Correcting for trigger side dilution and sequential decays using $D_{sub} = 0.6412 \pm 0.0015$, we get $\epsilon D^2 = 0.438 \pm 0.105$ % for a total efficiency, $\epsilon = 1.927 \pm 0.072$ %.
- Run I number from CDF note 3809, corrected $\epsilon D^2 = 0.34 \pm 0.08$ % (with dE/dx).

Please note:

There are several other very nice works in progress!

- Compared to electrons:

- Higher purity
- Less handles to discriminate fakes
- “Natural” fakes from decays in flight



muon	raw $D, \%$	$\epsilon, \%$
IMU	30 ± 11	0.20 ± 0.02
CMX	26 ± 7	0.57 ± 0.04
CMU only	12 ± 6	0.89 ± 0.05
CMP only	15 ± 9	0.32 ± 0.03
CMUP	37 ± 6	0.54 ± 0.04
Any	22 ± 3	2.52 ± 0.08

muon	raw $D, \%$	$\epsilon, \%$
IMU	39 ± 9	0.23 ± 0.02
CMX	21 ± 6	0.60 ± 0.04
CMU only	9 ± 5	0.90 ± 0.05
CMP only	16 ± 8	0.33 ± 0.03
CMUP	25 ± 5	0.59 ± 0.03
Any	19 ± 3	2.62 ± 0.08

A likelihood-based approach is being developed (?)

Comparisons

Blessed:

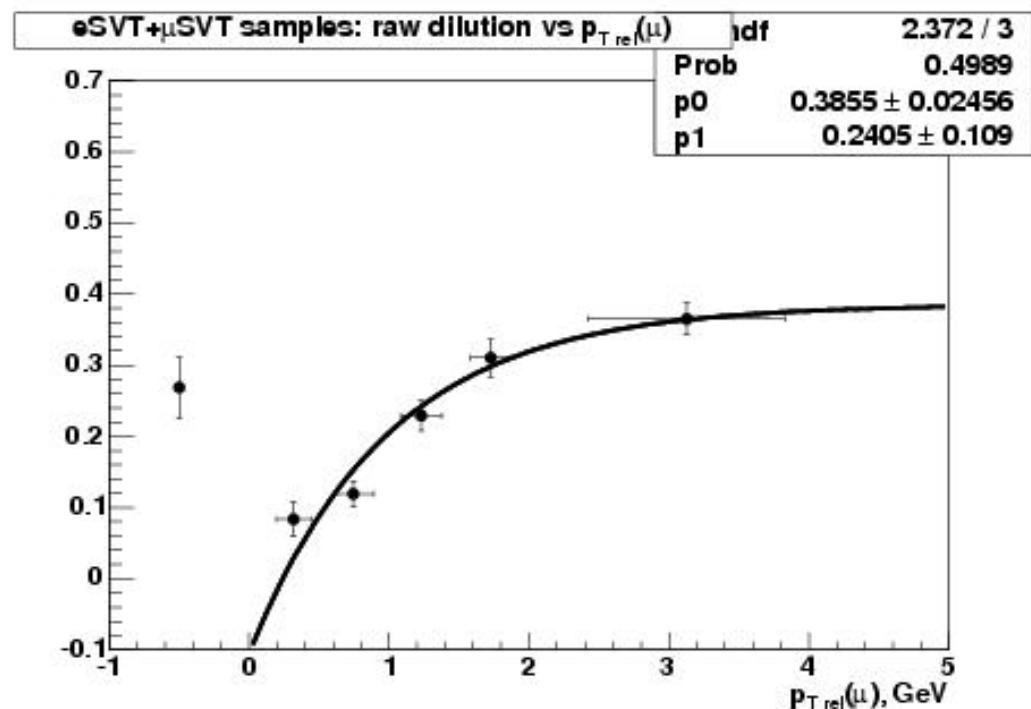
$$\epsilon D^2 = \begin{cases} 0.735 \pm 0.122 \% & \mu + SVT \\ 0.583 \pm 0.134 \% & e + SVT \end{cases}$$

Updated:

$$\epsilon D^2 = \begin{cases} 0.576 \pm 0.056 \% & \mu + SVT \\ 0.515 \pm 0.061 \% & e + SVT \end{cases}$$

Old 64pb run range, with 5.1 data:

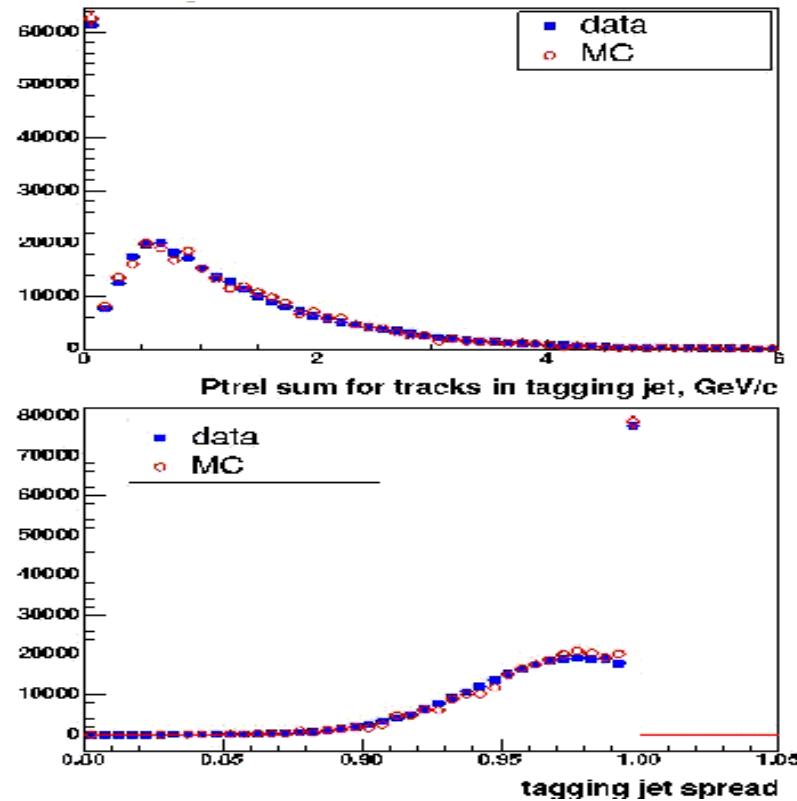
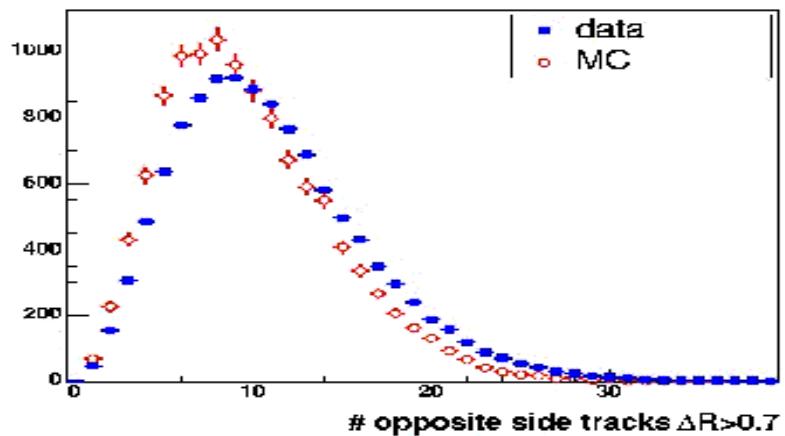
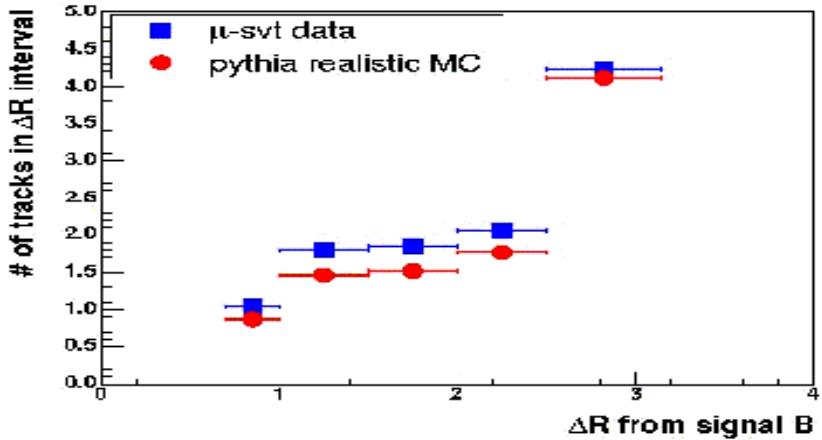
$$\epsilon D^2 = \begin{cases} 0.680 \pm 0.092 \% & \mu + SVT \\ 0.640 \pm 0.090 \% & e + SVT \end{cases}$$



Run I : $0.584 \pm 0.082 \%$ (M. Peters & co.)



- Exercising in a full-fledged data/MC comparison [D^0]
- Still significant discrepancies in the montecarlo: needs tuning!



First Glance Performance:

$$\varepsilon \approx 65\%$$

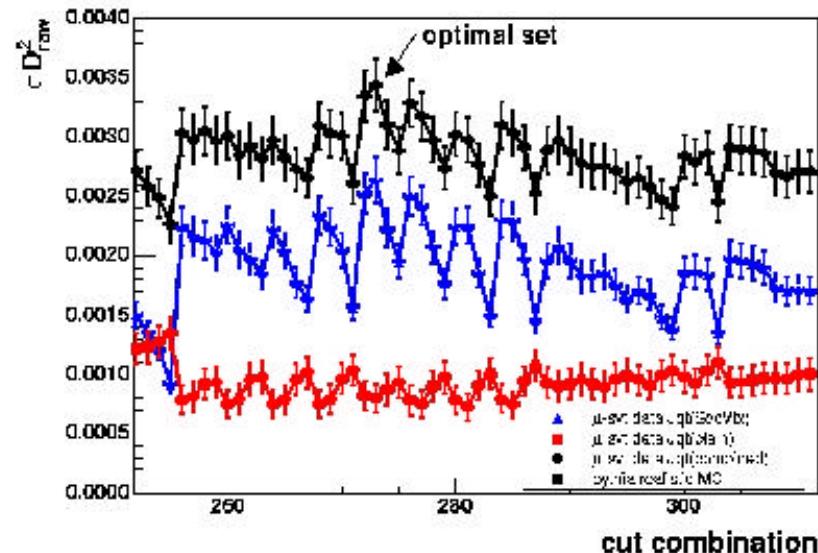
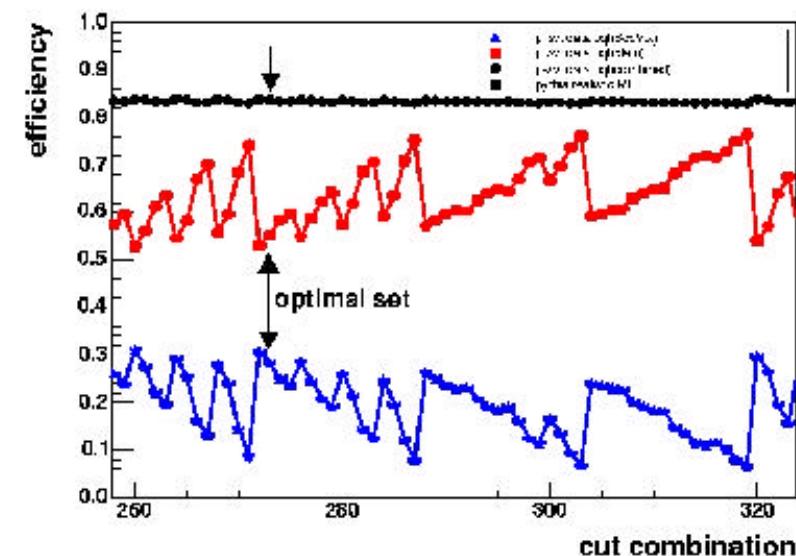
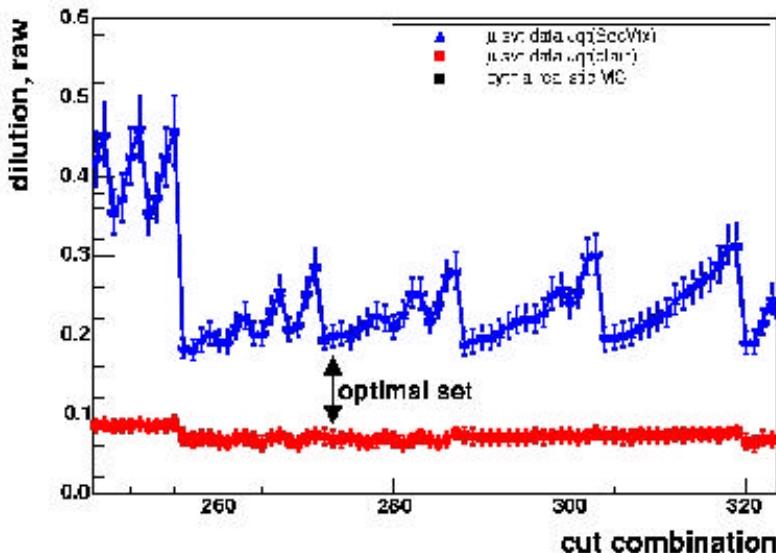
$$D_{\text{raw}} \approx 4\%$$

$$\varepsilon D_{\text{raw}}^2 \approx 0.11\%$$

JQT tuning



- each point: a set of cuts
- full plot runs from 0 to 2559
- part with the best point shown
- optimal set: point 273
 - $dR_{jm} = 1.0$ $dR_{clm} = 0.7$
 - $d0sigp1 = 0.0$ $d0sigp2 = 1.0$
 - $ptp1mx = 0.5$ $ptp2mx = 1.2$
- notice: high SecVtx efficiency



JQT tuning: details

Tuned parameters:

- $dR_{clm} = \{0.7, 0.9, 1.2, 1.5\}$
- $dR_{jm} = \{0.8, 1.0, 1.2, 1.5\}$
- $d0sigp1 = \{0.0, 1.0, 2.0, 3.0\}$
- $d0sigp2 = \{0.0, 1.0, 2.0, 3.0\}$
- $ptp1mx = \{0.5, 1.2, 2.0, 3.0\}$
- $ptp2mx = \{0.5, 1.2, 2.0, 3.0\}$

Scan strategy:

- try all points of 6D grid
- keep $dR_{clm} < dR_{jm}$
- total 2560 sets to explore

Fixed parameters of JQT:

- pttrack 0.4 GeV min
- ptseed 1.0 GeV min
- dR_{iso} 0.7
- ptcutp1 0.4 GeV min
- ptcutp2 0.4 GeV min
- atcutp1 3.0 max
- Lxy 3.0 min
- p1ptupgradep2 false
- 3D vertexing false

Run I comparison

Run I numbers, again, on 8 GeV lepton + vertex:

- average D (cdf4847): $\epsilon D^2 = 0.46 \pm 0.09\%$
- use $D(Q_{jet})$ (cdf4315): $\epsilon D^2 = 0.78 \pm 0.15\%$

More details on Run II numbers with Run I cuts:

- no cut on $P_t(\mu)$

tag type	$\epsilon, \%$, $D, \%$	$\epsilon D^2, \%$
events with SecVtx jet	3.4	20.8 ± 1.4	0.145 ± 0.019
events with no SecVtx jet	27.1	7.2 ± 0.5	0.139 ± 0.019
combined	30.5	—	0.284 ± 0.027

- with cut $P_t(\mu) > 8GeV$

tag type	$\epsilon, \%$, $D, \%$	$\epsilon D^2, \%$
events with SecVtx jet	4.9	25.6 ± 2.3	0.319 ± 0.064
events with no SecVtx jet	34.2	6.3 ± 1.0	0.136 ± 0.042
combined	39.1	—	0.455 ± 0.077

JQT 2004

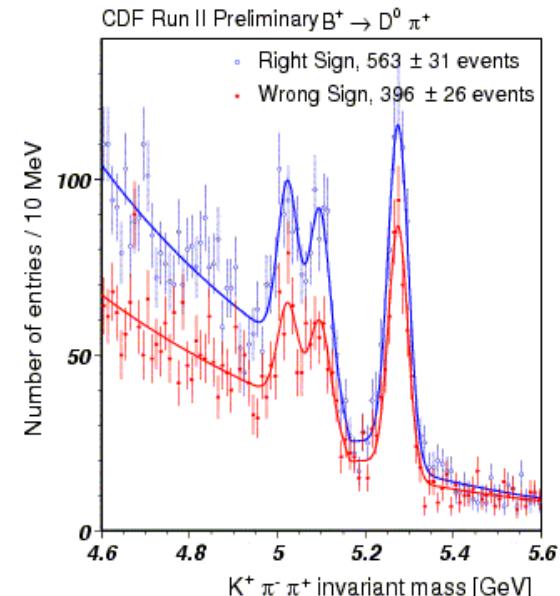
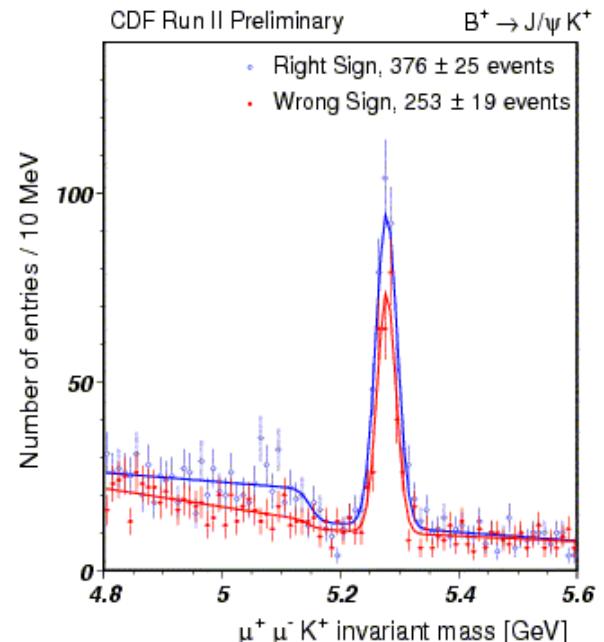
- Got in the phase of “systematic” studies
- Still open issues:
 - X3 discrepancy in D predictions between data & MC
 - Revise SecVtx tuning?
 - How much is L00 helping?
- My favorite topics (I.e. my interests):
 - “Refurbished” vertexing
 - Merging with PID
 - Naïve, single track
 - Full blown likelihood-based (rewarding but tough!)



SST

Gerry Bauer, Guillermo Gomez-Ceballos, Ilya Kravcenko,
Nuno Leonardo, Christoph Paus, Jonatan Piedra,
Sasha Rakitin, Alberto Ruiz, Ivan Vila

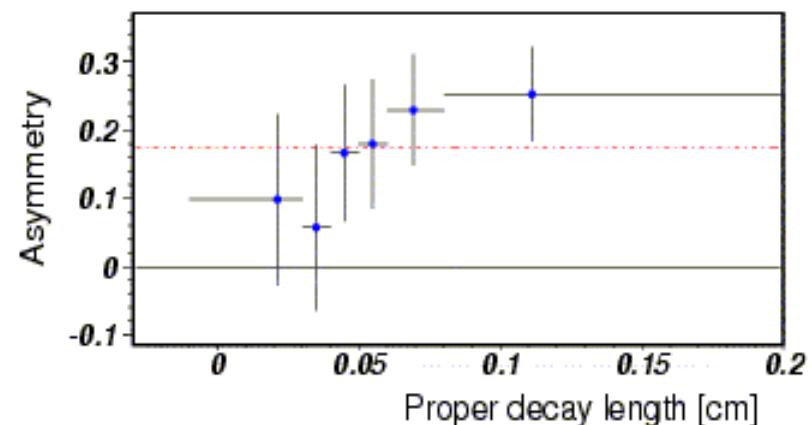
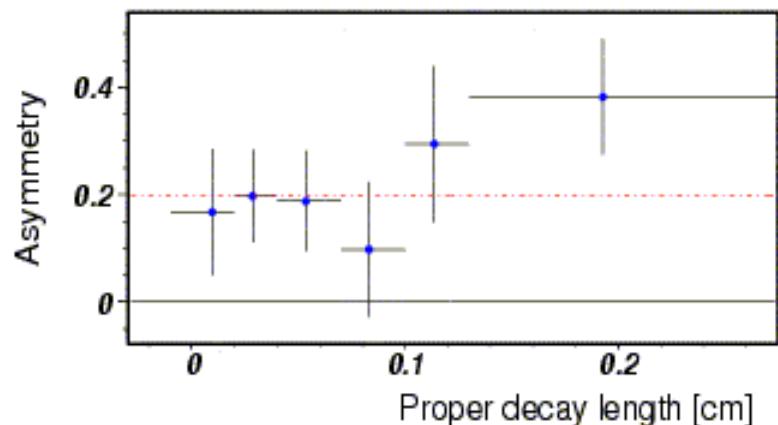
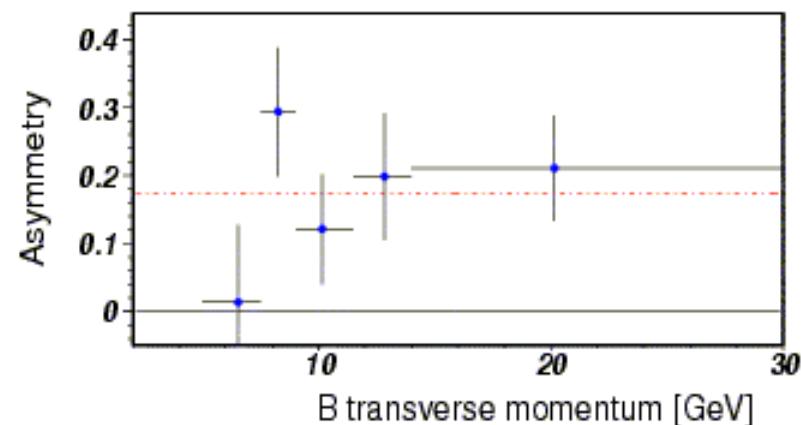
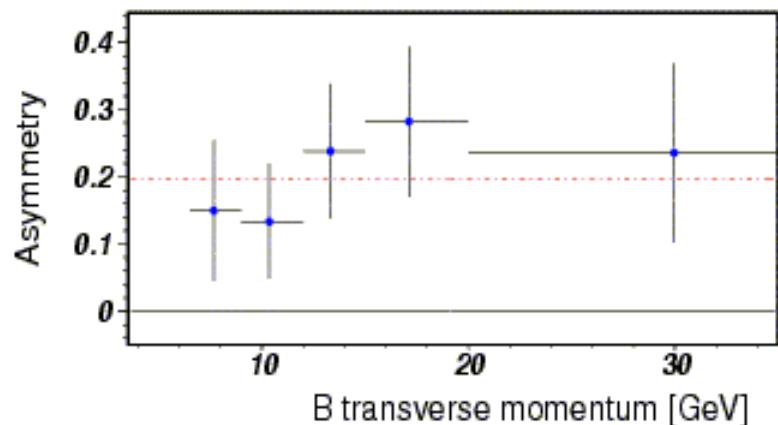
- Run I -like algorithm has been implemented:
 - $\Delta R < 0.7$, $P_T > 0.4$
 - $|d_0/\sigma| < 3.0$
 - Minimum P_T^{rel}
- Results are checked on two samples:
 - $B^+ \rightarrow J/\psi K^+$
 - $B^+ \rightarrow D^0 \pi^+$
- Encouraging results, working on a large statistics study (e.g. ID^0)



SST (cont'd)

$B^+ \rightarrow J/\psi K^+$

$B^+ \rightarrow D^0 \pi^+$



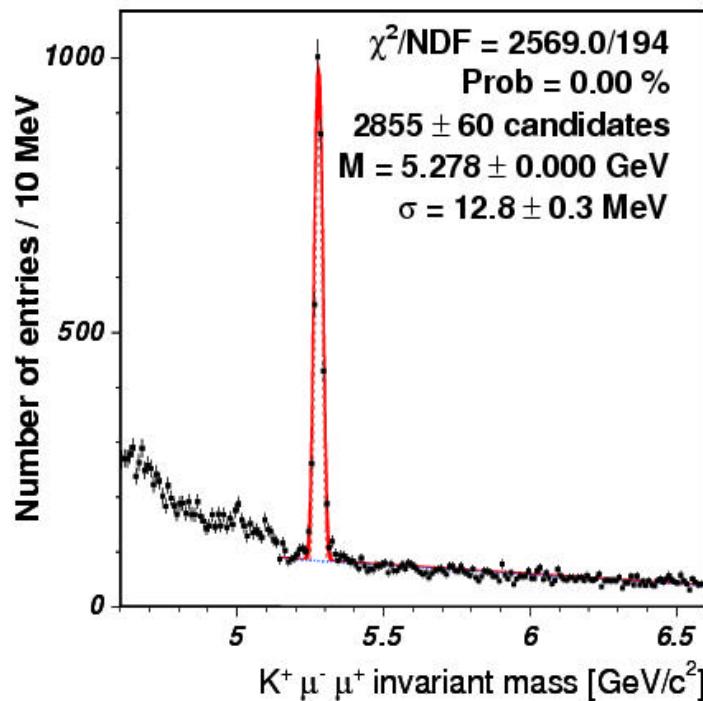
	N_{RS}	N_{WS}	N_{NT}	ϵ	D	ϵD^2
$J/\psi K^+$	376 ± 25	253 ± 19	379 ± 23	62.4 ± 1.8	19.7 ± 4.9	2.4 ± 1.2
$D^0 \pi^+$	563 ± 31	396 ± 26	588 ± 27	62.0 ± 1.5	17.4 ± 4.1	1.9 ± 0.9

SST 2004

$B^+ \rightarrow J/\psi K^+$

- $N(J/\psi K^+) = 2855 \pm 60$

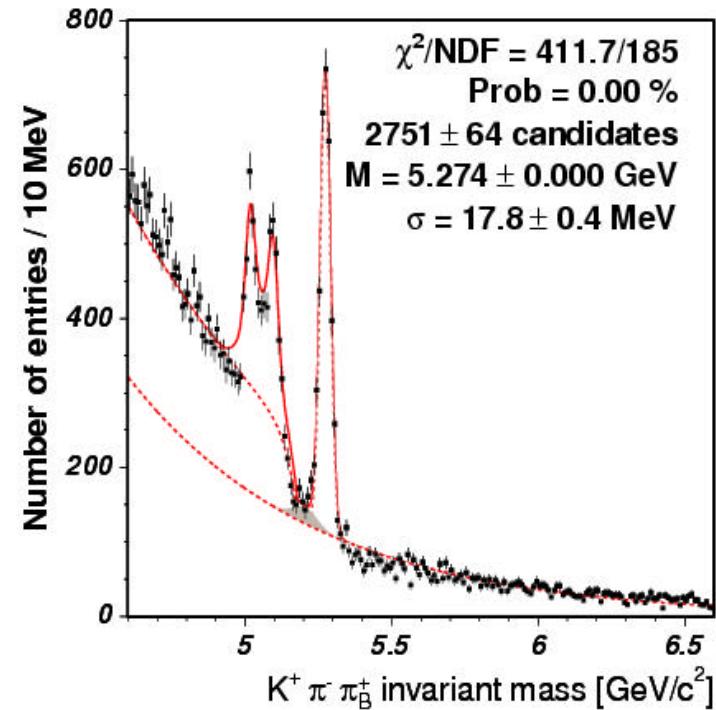
sst all events



$B^+ \rightarrow \bar{D}^0 \pi^+$

- $N(\bar{D}^0 \pi^+) = 2751 \pm 64$

sst all events



Decay	ϵ (%)	D (%)	ϵD^2 (%)
$B^+ \rightarrow J/\psi K^+$	62.5 ± 0.9	19.0 ± 2.3	2.3 ± 0.6
$B^+ \rightarrow \bar{D}^0 \pi^+$	63.0 ± 0.9	20.8 ± 2.4	2.7 ± 0.6

SST 2004 cont'd

- Reasonable agreement with Run I
- Lacking a more systematic approach than just "cut and try":
 - MC tuning
 - Cut optimization
 - Tagger calibration
 - Benchmark on semileptonics
 - Compare with TTT
 - Understand differences (trigger biases?)
 - Combine with PID?

OSK: TOF

J. Piedra, A. Ruiz, I. Vila, M. Wolter and Ch. Paus

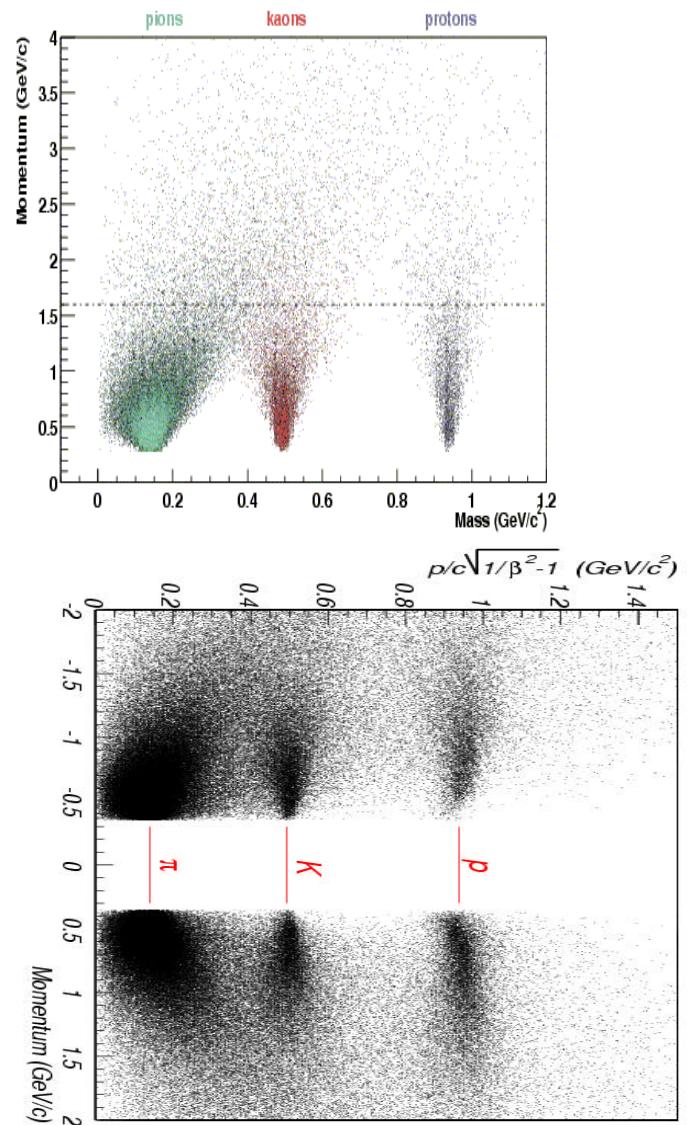
	ϵ (%)	D (%)	ϵD^2 (%)
100% eff., 110ps	12.2 ± 0.3	30.6 ± 2.9	1.14 ± 0.25
GEANT 110ps	10 ± 0.3	28.4 ± 3.2	0.81 ± 0.21
GEANT, $\epsilon=0.8$ t_0 truth	11.2 ± 0.3	26.8 ± 3.8	0.8 ± 0.2
GEANT	11.2 ± 0.3	23.9 ± 3.0	0.64 ± 0.18
GEANT 65% eff.	9.4 ± 0.3	27.0 ± 3.3	0.68 ± 0.19

First naïve attempts on data:

$$\epsilon = 4.34 \pm 0.41\%$$

$$D = 17.43 \pm 9.43\%$$

$$\epsilon D^2 = 0.13 \pm 0.16\%$$



OSK 2004

- TOF reconstruction/simulation are getting better
- We need an effort:
 - Aggressive
 - Systematic
 - Goal-oriented
 - Quick and responsive
- My favorite direction: merging with smart vertexing and JQT (see JQT slide)



Beyond the Baseline

- “Same tags, new samples”: tagging-enhanced samples:

- Hadronic+TOF hit(track)
- Hadronic+away side track
 - How soft?
 - Displaced?
 - TOF-tagged
 - Lepton?
- Hadronic+same side track
 - How soft?
 - TOF-tagged?
- Replace “hadronic” with anything, in principle

Bring the information @ L2?

L1 already knows which XFT tracks have TOF hits!

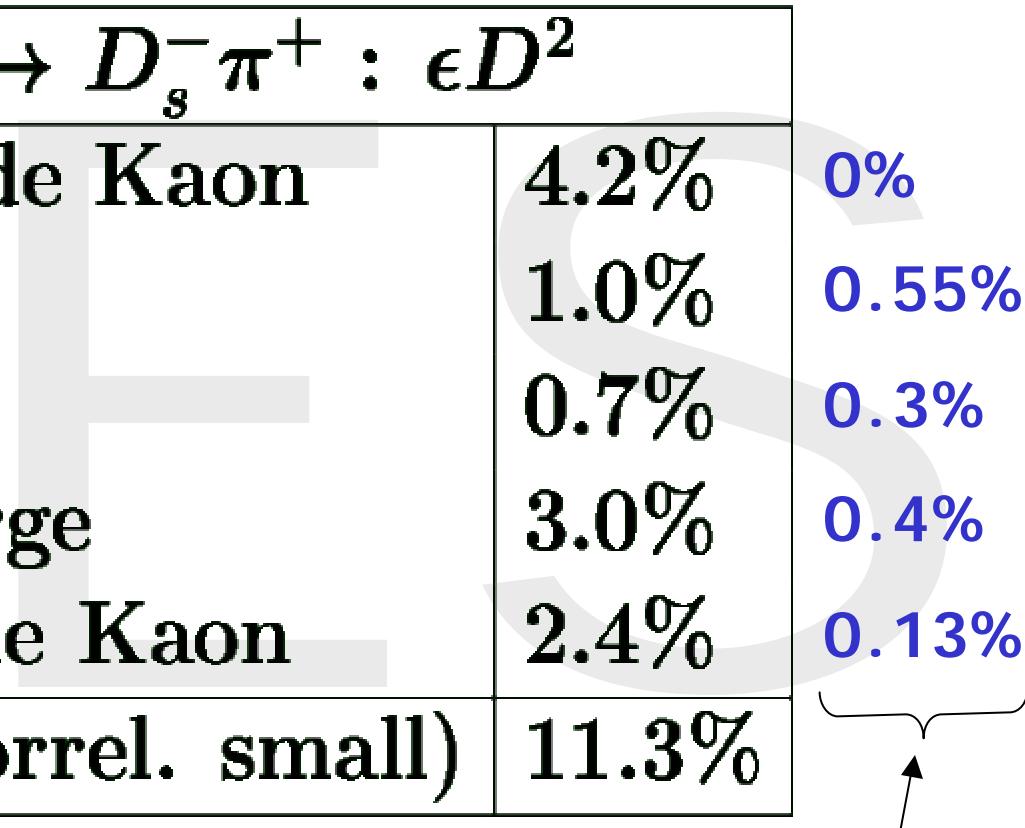
- “New tags”

- Opposite side charm!
 - Ultra-low momentum tracking is **crucial!** ($D^* \rightarrow (K\pi)\pi^*$)
 - Also PID could help ($D^* \rightarrow (K\pi)\pi^*$)?

Tagging category	ε (%)	w (%)	Q (%)
Lepton	10.9 ± 0.3	8.9 ± 1.3	7.4 ± 0.5
Kaon	35.8 ± 0.5	17.6 ± 1.0	15.0 ± 0.9
NT1	7.8 ± 0.3	22.0 ± 2.1	2.5 ± 0.4
NT2	13.8 ± 0.3	35.1 ± 1.9	1.2 ± 0.3
ALL	68.4 ± 0.7		26.1 ± 1.2

Conclusions

$B_s \rightarrow D_s^- \pi^+ : \epsilon D^2$	
Same-Side Kaon	4.2%
μ tag	1.0%
e tag	0.7%
Jet Charge	3.0%
Opp.-Side Kaon	2.4%
Total (correl. small)	11.3%



Wondering if I am
cheating? YES!

Where is the room?

- Overall:
 - So far effort concentrated on reproducing Run I results, can we go beyond?
 - Extended coverage
 - Standalone Si tracks
 - Plug detectors for leptons
 - Montecarlo tuning
 - Detector understanding
 - Improve physics modeling
 - Improved/Alternative taggers
 - OSK
 - SSK
 - “merged” taggers
 - The best tagger is **without cuts**