DØ note ????

Combination of Same-Side Taggers for B_s Mesons

Alexander Rakitin
Lancaster University, Lancaster, UK

(Dated: July 28, 2006)

In this paper we apply the technique developed for the combination of opposite-side taggers to the same-side taggers. A few selected same-side taggers are combined into one. Further, same-side and opposite-side taggers are applied together to measure the dilution D and tagging power ϵD^2 of the same-side taggers in 1 fb^{-1} data collected with DØ detector. The dilution of the combined opposite-side tagger was measured previously.

1 Introduction

The oscillation, or mixing, of the b-quark flavor is a well-known effect for B mesons. For example, in B_s meson system the mass eigenstates $|(B_s)_H\rangle$ and $|(B_s)_L\rangle$ with masses M_H and M_L are related to flavor eigenstates $|B_s\rangle = (\bar{b}s), |\bar{B}_s\rangle = (b\bar{s})$ as follows:

$$|(B_s)_H\rangle = \frac{1}{\sqrt{2}}(|B_s\rangle + |\bar{B}_s\rangle),$$

$$|(B_s)_L\rangle = \frac{1}{\sqrt{2}}(|B_s\rangle - |\bar{B}_s\rangle).$$

As a result, a B_s-meson born at time $t = 0$ as $|B_s\rangle$ may decay at time t as $|\bar{B}_s\rangle$ with probability

$$p(B_s \to \bar{B}_s) = \frac{e^{-t/\tau}}{2\tau}(1 + \cos \Delta m_s t).$$

Probability for $|B_s\rangle$ meson to keep its flavor at decay at time t is

$$p(B_s \to B_s) = \frac{e^{-t/\tau}}{2\tau}(1 - \cos \Delta m_s t).$$

Here τ is the B_s-meson lifetime. The parameter $\Delta m_s \equiv M_H - M_L$ is called "mixing frequency". It is important to know this parameter precisely, since it determines the behavior of B_s mesons system. Also, it is important to know the ratio of mixing frequencies $\Delta m_s/\Delta m_d$ for B_s and B_d systems, because it gives us constraint on CKM matrix elements. The Feynman diagrams describing B_s mixing are displayed in Figure 1. The transition is dominated by heavy t-quark, so that the CKM matrix elements V_{tb} and V_{ts} play an important role in the mixing phenomenon. Similar diagram for B_d includes V_{tb} and V_{td}, therefore the ratio $\Delta m_s/\Delta m_d$ allows us to constrain the ratio V_{ts}/V_{td}. More exactly,

$$\frac{\Delta m_s}{\Delta m_d} = \frac{m_{B_s}}{m_{B_d}} \cdot \xi^2 \cdot \frac{|V_{ts}|^2}{|V_{td}|^2},$$

where $\xi = -1.210^{+0.047}_{-0.035}$ (Ref. [1]).

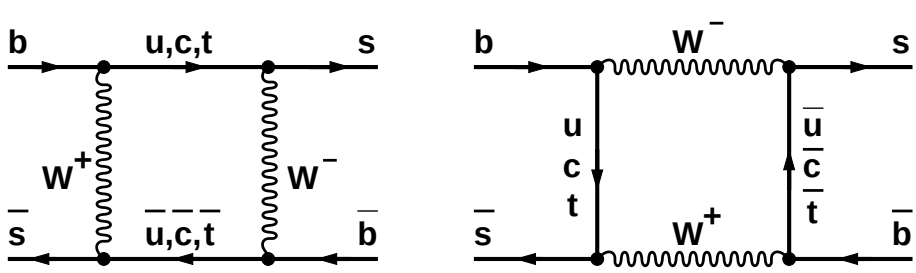

Figure 1: Feynman diagrams for B_s-mixing

In this analysis the tools for Δm_s measurement (tagging methods) are developed.

This note is organized as follows: in the next Section 2 we talk about B-flavor tagging in general. Section 3 describes the data and Monte Carlo samples used for the analysis. In Section 4 we discuss same-side tagging algorithms and their combination in Section 5. Section 6 is devoted to the measurement of the SST dilution by using double-tagged events in the data. Last comes the conclusion.

2 B-Flavor Tagging

To know if B-meson oscillated or not we need to to know (to "tag") its flavor at production and at decay. Let's consider semileptonic decays for simplicity. The flavor at decay can be easily obtained from the sign of trigger lepton (ℓ^+ corresponds to $\bar{b}$-quark, ℓ^- – to b-quark). The flavor at production is usually obtained with a special algorithm called "tagger". There are two main classes of such algorithms: Opposite-Side Taggers (OST) and Same-Side Taggers (SST). OST uses the fact that b-quarks are produced in b-$\bar{b}$ pairs. It determines the flavor of "the other" b-quark in pair which usually hadronizes and then fragments into a jet on "the other" (opposite) side of the event (Figure 2). The average charge of the tracks in this jet gives a charge, and thus a flavor, of "the other" b-quark, from which we infer the initial flavor of the b-quark inside the B-meson. This method is called "jet-charge" tagger. If a soft lepton is found among the tracks in the jet on "the other" side, then the flavor of "the other" b-quark can be obtained from the charge of this lepton only ("soft-lepton" tagger).

SST uses the tracks on the same side of the event as the B-meson. The principle of its work is clear from the Figure 3. The b-quark "picks up" a u, d or s quark from a q-$\bar{q}$ pair and the remaining quark in the pair forms a meson (or even hyperon) with some other quark(s). If this meson is charged we can reconstruct its track and from its charge infer the information about the b-quark flavor at production. B meson born as B_d is likely to have a π^+ nearby (π^- for $\overline{B}_d$). Similarly, B meson born as B_s is likely to have a K^+ nearby (K^- for $\overline{B}_s$).

There are different methods of finding the particle which was born together with B meson (different same-side taggers). All of them can be divided into three groups:

- Taggers using one track, selected according to some kinematic requirements
- Taggers reconstructing resonances, such as K^{*0} or Λ, decaying into two tracks
- Taggers using all the tracks in the vicinity of the B-meson

Particular implementation of these taggers will be considered in Section 4.

For each event taggers conclude if the B-meson did not oscillate (*i.e.* the charge of the tag is the same as lepton charge, "Right-Sign"), or if it did (*i.e.* the charge of the tag is opposite to the lepton charge, "Wrong-Sign"), or if no tag was found ("No-Tag"). The main tagger characteristics are the following:

- tagging efficiency $\varepsilon = \frac{N_{RS}+N_{WS}}{N_{RS}+N_{WS}+N_{NT}}$
- *raw* dilution D_{raw} or asymmetry $A = \frac{N_{RS}-N_{WS}}{N_{RS}+N_{WS}}$
- (*true*) dilution $D = 1 - 2p$, where p is a mistag rate, which can be obtained from MC if true B-flavor at production is known
- tagging power εD^2 or $\varepsilon D^2_{\text{raw}}$

It can be shown that the error on Δm_s is inversely proportional to $\sqrt{\varepsilon D^2}$, so that it is

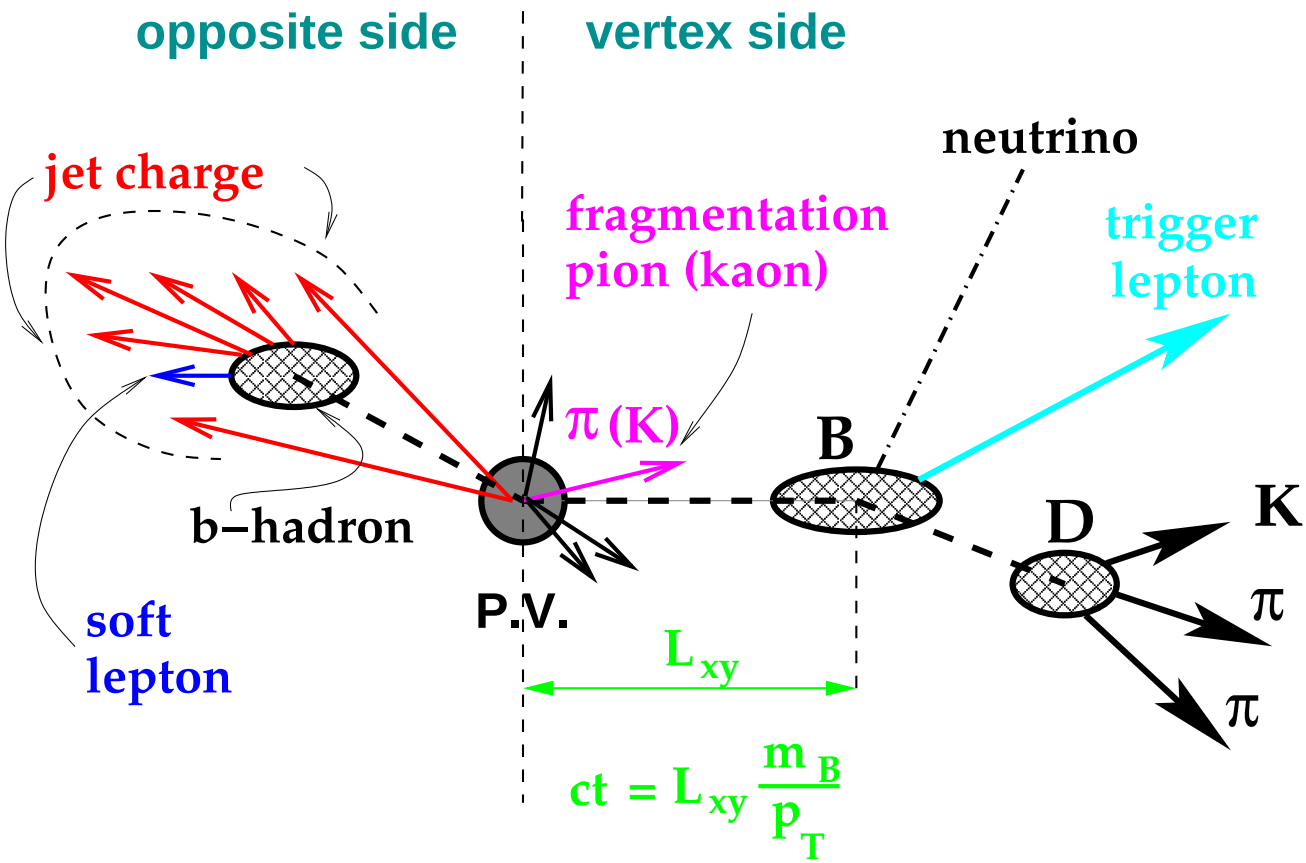

Figure 2: The sketch of the event with a B-meson

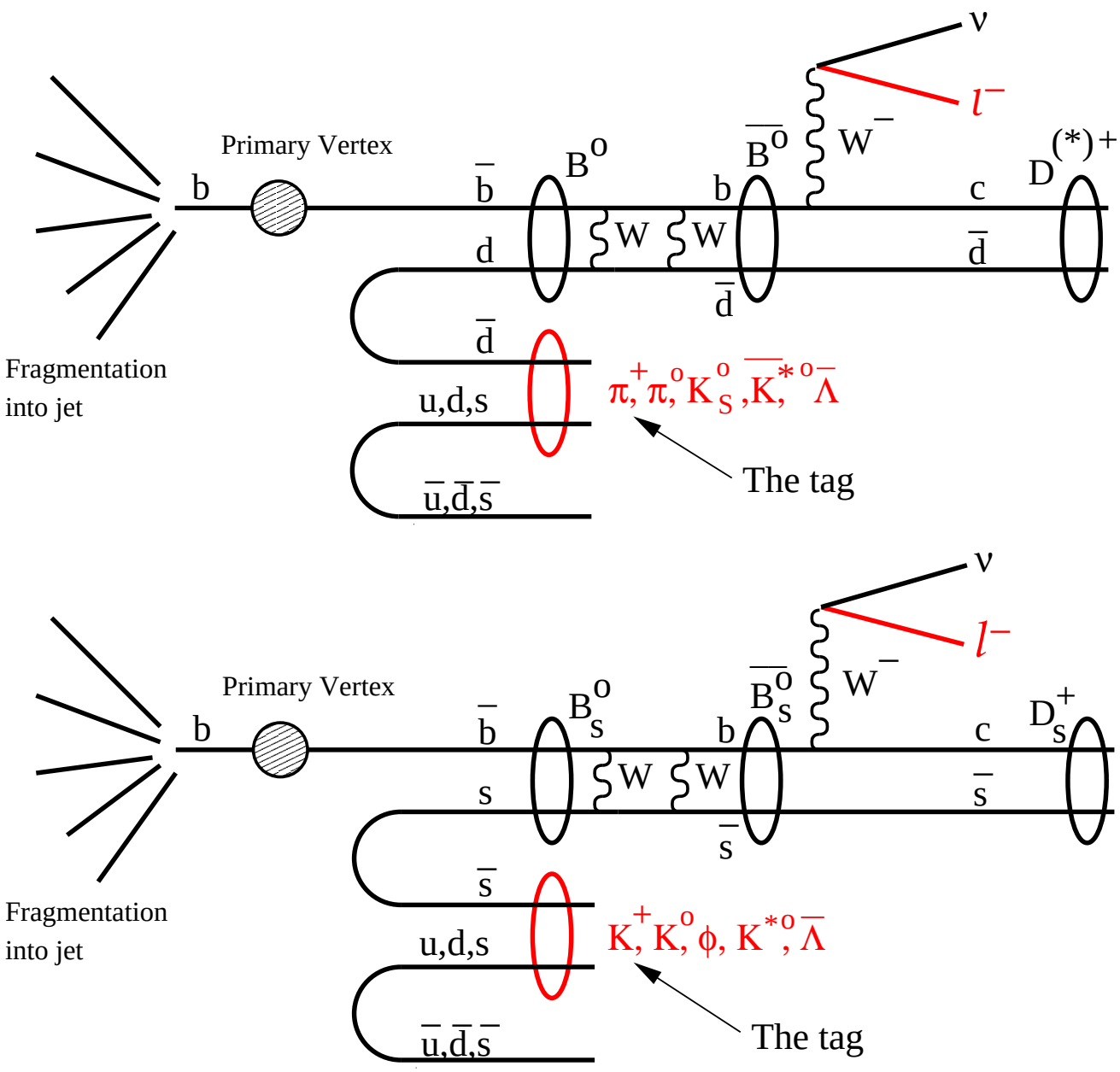

Figure 3: The scheme of SST

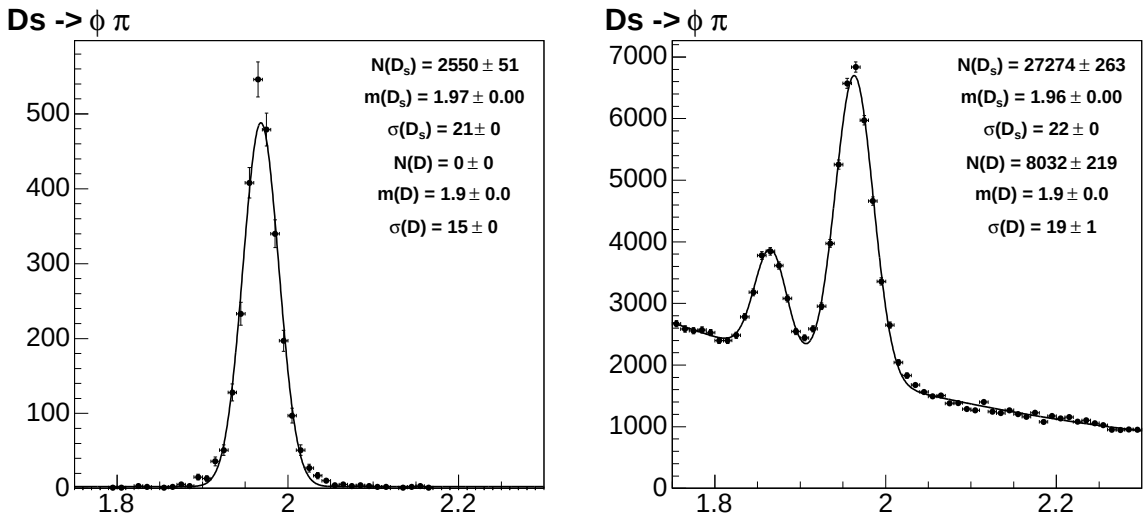

Figure 4: The D_s mass histograms in Monte Carlo and in the data

important to use taggers with high tagging power in the analysis. With the purpose to increase the tagging power we combine the same-side taggers into one "Comb. SST" algorithm. Since the taggers in each group are highly correlated with each other, we select only the best taggers from every group and combine only them (see Section 5).

3 Monte Carlo and Data Samples Used in the Analysis

The following Monte Carlo samples are used to study SST:

- $B_s \to \mu D_s, D_s \to \phi\pi$, p17 requests 29892, 29893 (150K events)
- $B_s \to \mu\mu$, p17 requests 29215, 29216, 29283 (166K events)
- $\overline{B_s} \to \mu\mu$, p17 requests 29213, 29214, 29282 (121K events)
- $B_s \to \mu D_s X, D_s \to K_s K$, p17 request 23838 (180K events)

To make sure that Monte Carlo matches the data we compare various distributions in the first Monte Carlo and in the data selected in `/prj_root/1008/ckm_write/bgv/evt/muphipi-std/` (1 fb^{-1}).

The mass histograms of D_s in the Monte Carlo and the data are shown in Figure 4. It is fitted with two single gaussians plus quadratic background. The sideband-subtracted data distributions are to be compared to the Monte Carlo ones. The signal and sideband regions are chosen to be (1.91,2.01) and (1.75,1.80) $\bigcup$ (2.12,2.17) GeV/c^2.

The Monte Carlo–data match of the important for SST distributions (see details in Section 4.1) is shown in Figure 5 for p17 Monte Carlo $B_s \to \mu D_s, D_s \to \phi\pi$, requests 29892, 29893. Magenta crosses represent the data, green histograms – Monte Carlo. We see very good match except some discrepancies at low p_t^{rel} region.

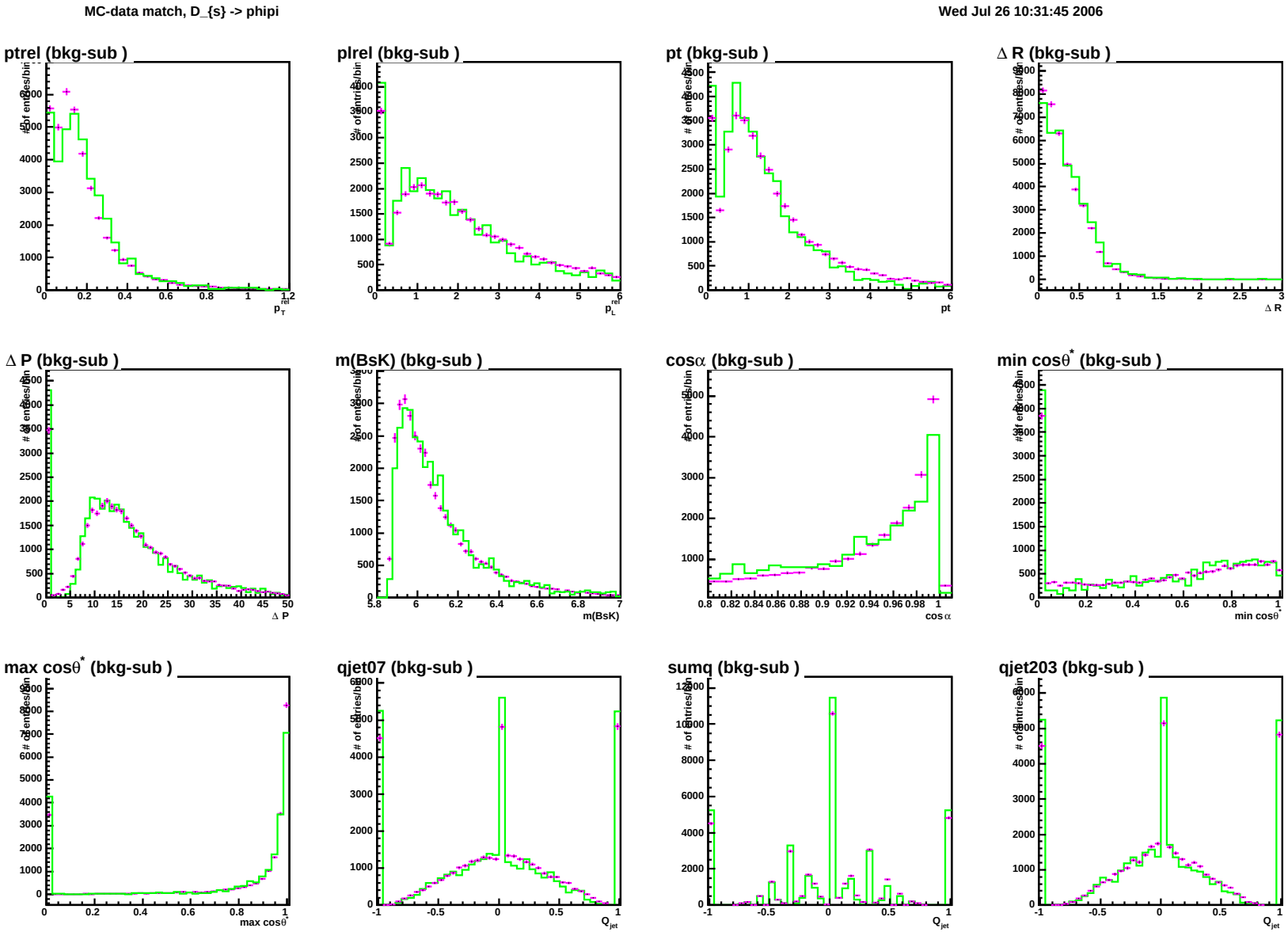

Figure 5: The p17 Monte Carlo–data match for various quantities (see Sections 4.1, 4.3)

4 SST Algorithms and Their Performance in Monte Carlo

Since in Monte Carlo we know the true b-quark flavor at production, we can determine the number of the events with correct charge of the tag ("Right-Tag") and the opposite charge of the tag ("Wrong-Tag"). Thus we obtain the mistag rate $p = \frac{N_{WT}}{N_{RT}+N_{WT}}$, and, therefore, *true* dilution $D = 1 - 2p$ for the taggers above. Also, we obtain the number of the events where tag was not found ("No-Tag") and the tagging efficiency $\epsilon = \frac{N_{RT}+N_{WT}}{N_{RT}+N_{WT}+N_{NT}}$.

To tag the flavor of the B-meson at the production stage we look at the tracks in cone $\cos\alpha < 0.8$ around 3-dimensional momentum of B-meson, $\vec{p}(B_s)$. This requirement comes from consistency with opposite-side tagging (Ref. [3]). The tracks are supposed to have at least 2 SMT axial hits and at least 3 CFT axial hits. The primary vertex of each track must coincide with the production vertex of the B_s meson.

4.1 SST Algorithms Using One Track

The following algorithms were used to select one-track tag (see left part of Figure 6):

- ☞ Min. $p_t^{\rm rel}$
- ☞ Max. $p_L^{\rm rel}$
- ☞ Max. p_t
- ☞ Min. $|\Delta\vec{P}| \equiv |\vec{p}(B_s) - \vec{p}(K)|$
- ☞ Min. ΔR
- ☞ Max. $\cos\alpha$
- ☞ Min. $\cos\theta^*$
- ☞ Max. $\cos\theta^*$
- ☞ Min. $m(B_sK)$
- ☞ Random track
- ☞ ...
- ☞ ...

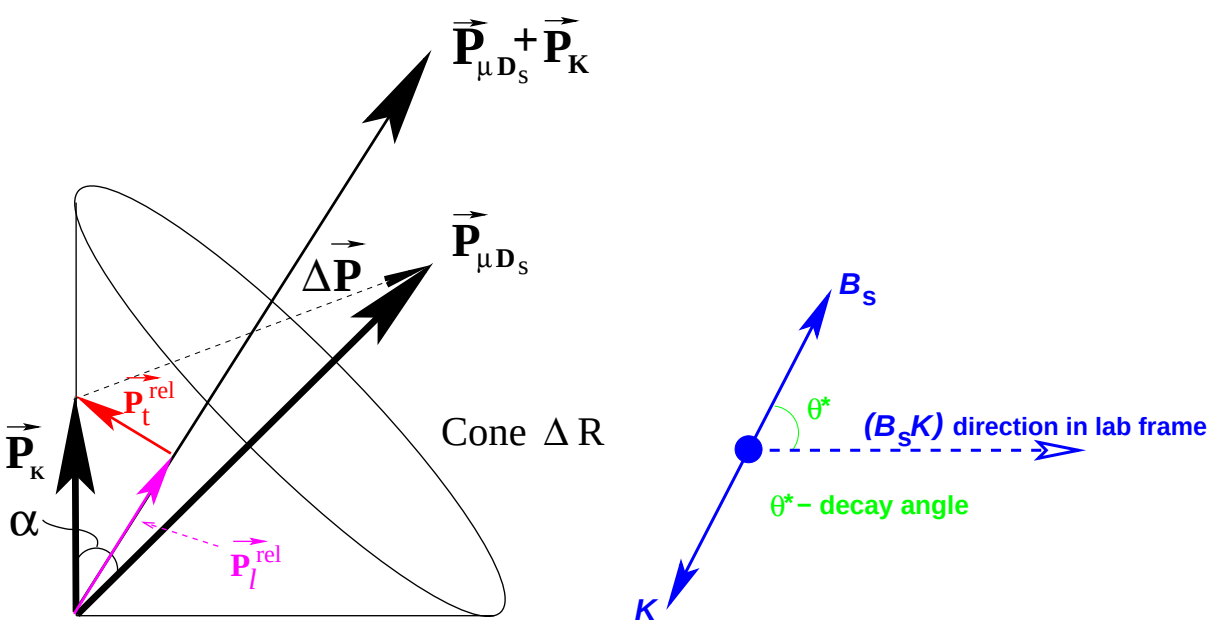

Figure 6: **Left:** One-track SST selection; **Right:** θ^* – decay angle of B_sK-system.

Here $p_t^{\rm rel}$ and $p_L^{\rm rel}$ are $\perp$ and $||$ components of SST candidate's momentum $\vec{p}(K)$ w.r.t $\vec{p}(B_sK)$.

Tagger	N_{RT}	N_{WT}	N_{NT}	ε, %	D, %	εD^2, %
Min. $p_t^{\rm rel}$	262644 ± 512	240178 ± 490	105438 ± 325	82.7 ± 0.0	4.5 ± 0.1	0.165 ± 0.010
Max. $p_L^{\rm rel}$	263285 ± 513	239537 ± 489	105438 ± 325	82.7 ± 0.0	4.7 ± 0.1	0.184 ± 0.011
Max. p_t	263870 ± 514	238952 ± 489	105438 ± 325	82.7 ± 0.0	5.0 ± 0.1	0.203 ± 0.011
Min. $\lvert\Delta\vec{P}\rvert$	263809 ± 514	239013 ± 489	105438 ± 325	82.7 ± 0.0	4.9 ± 0.1	0.201 ± 0.011
Min. ΔR	267795 ± 517	235027 ± 485	105438 ± 325	82.7 ± 0.0	6.5 ± 0.1	0.351 ± 0.014
Max. $\cos\alpha$	267412 ± 517	235410 ± 485	105438 ± 325	82.7 ± 0.0	6.4 ± 0.1	0.335 ± 0.014
Min. $\cos\theta^*$	266163 ± 516	236659 ± 486	105438 ± 325	82.7 ± 0.0	5.9 ± 0.1	0.285 ± 0.013
Max. $\cos\theta^*$	258955 ± 509	243867 ± 494	105438 ± 325	82.7 ± 0.0	3.0 ± 0.1	0.074 ± 0.007
Min. $m(B_sK)$	265779 ± 516	237043 ± 487	105438 ± 325	82.7 ± 0.0	5.7 ± 0.1	0.270 ± 0.013
Random track	261105 ± 511	241717 ± 492	105438 ± 325	82.7 ± 0.0	3.9 ± 0.1	0.123 ± 0.009

Table 1: Comparison of different one-track taggers for all four MC samples

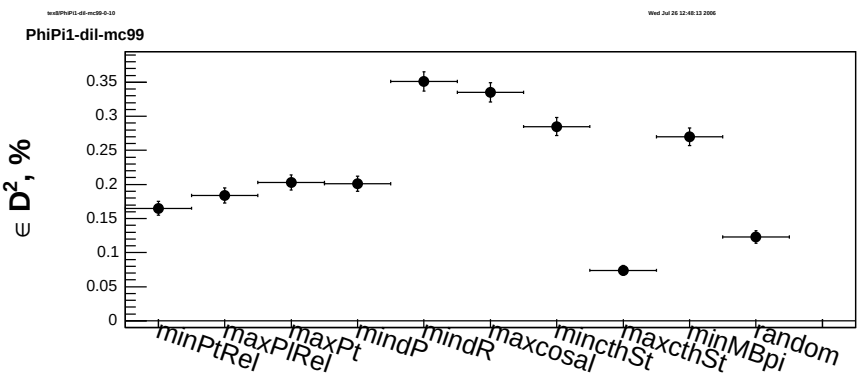

Figure 7: Comparison between εD^2's for one-track taggers for all four MC samples

The $\Delta R \equiv \sqrt{\Delta\phi^2 + \Delta\eta^2}$ and angle α are taken between $\vec{p}(B_s)$ and $\vec{p}(K)$. The θ^* – decay angle of B_sK-system, *i.e.* angle between directions of $\vec{p}(B_sK)$ and $\vec{p}(B_s)$ in reference frame of B_sK system, as shown in the right part of Figure 6.

The *true* dilutions and ϵD^2's for these taggers obtained in all four Monte Carlo samples are given in Table 1 and the ϵD^2's are graphically compared in Figure 7.

Of course, all the one-track taggers are highly correlated to each other. So we will choose one of them which gives the best result (has largest εD^2) and will be using only it in the tagger combination. From Figure 7 one can see that "Min. ΔR" is the best tagger, with "Max. $\cos\alpha$" closely following him. The "Random track" tagger has pretty low tagging power, as one would expect. The "Max. $\cos\theta^*$" has even lower tagging power, because it purposefully selects wrong track as the tag. The tagger selecting better track, "Min. $\cos\theta^*$", has a decent tagging power.

4.2 Using Kaons Coming from K^{*0} and Pions From Λ:

Another group of taggers is based on reconstruction of two-track resonance in the vicinity of B-meson: $K^{*0} \to K\pi$ and $\Lambda \to p\pi$. The sign of kaon (pion for Λ) helps infer the b-quark flavor at production. The K^{*0} is reconstructed out of two oppositely charged tracks assigned with masses of kaon and pion and with invariant mass being $0.862\ \mathrm{GeV}/c^2 < m(K^{*0} \to K\pi) < 0.922\ \mathrm{GeV}/c^2$. The auto-reflection (same track combination with opposite mass assignment) is required to be outside of this mass window so that we could know for sure which track is kaon and which is pion. Both these tracks are required to be within cone $\cos\alpha > 0.8$ around $\vec{p}(B_s)$ and to have at least 2 axial hits in SMT and 3 axial hits in CFT. Also, they must be associated with the same primary vertex as B_s. The B_s daughters are excluded. The mass distributions of so found K^{*0} for MC and data are given in Figure 8. They are sideband-subtracted with respect to D_s mass.

Another implementation of the same tagger ("Optimized K^{*0}") is created with a few additional cuts applied to the reconstructed K^{*0}'s:

- $\Delta R \equiv \sqrt{\Delta\phi^2 + \Delta\eta^2} < 1.5$ between the tracks
- Both tracks' impact parameter $|d_0/\sigma_{d_0}| < 3.0$
- Vertex $\chi^2 < 20.25$
- Isolation $\frac{|\vec{p}(K^{*0})|}{|\vec{p}(K^{*0})| + \sum\limits_{tracks\ with\ \Delta R<0.5} |\vec{p}|} > 0.6$
- Cosine of the angle between the momentum of the positive track and the direction of the K^{*0} in the rest frame of K^{*0}, $|\cos\theta^*| < 0.8$

The mass distributions of K^{*0} for MC and data after these cuts (sideband-subtracted with respect to D_s mass) are demonstrated in Figure 9. The comparison of the taggers is given in Table 2 and, in graphical form, in Figure 10. The dilution of the "Optimized K^{*0}" is higher than that for the unoptimized one, but the efficiency is lower, so that the tagging power $\varepsilon \mathrm{D}^2$ is of the same order. Since the optimization does not give us strong advantages we choose unoptimized "K^{*0}" as the best tagger.

The reconstruction of $\Lambda \to p\pi$ is also performed only with tracks within cone $\cos\alpha > 0.8$ around $\vec{p}(B)$, having 2+ axial SMT hits and 3+ axial CFT hits and associated with the same primary vertex as B_s meson. The B daughters are again excluded. The standard reconstruction algorithm from `AA::vOFinder` from `bana` package is employed. The D_s-sideband-subtracted Λ mass histogram is shown in Figure 11 for data and MC. The dilution for the tagger "Lambda" is very high, but the efficiency is extremely low, so that the tagging power ϵD^2 is low too. Since tagger "Lambda" is uncorrelated with "K^{*0}" by construction, we will be using both of them for the tagger combination.

Tagger	N_{RT}	N_{WT}	N_{NT}	$\varepsilon, \%$	$\mathrm{D}, \%$	$\varepsilon \mathrm{D}^2, \%$
$K^{*0} \to K\pi$	55860 ± 236	52222 ± 229	500178 ± 707	17.8 ± 0.0	3.4 ± 0.3	0.020 ± 0.004
$K^{*0} \to K\pi$(opt)	40798 ± 202	37298 ± 193	530164 ± 728	12.8 ± 0.0	4.5 ± 0.4	0.026 ± 0.004
Λ	1669 ± 41	1246 ± 35	605345 ± 778	0.5 ± 0.0	14.5 ± 1.8	0.010 ± 0.002

Table 2: Comparison of different two-track taggers for all four MC samples

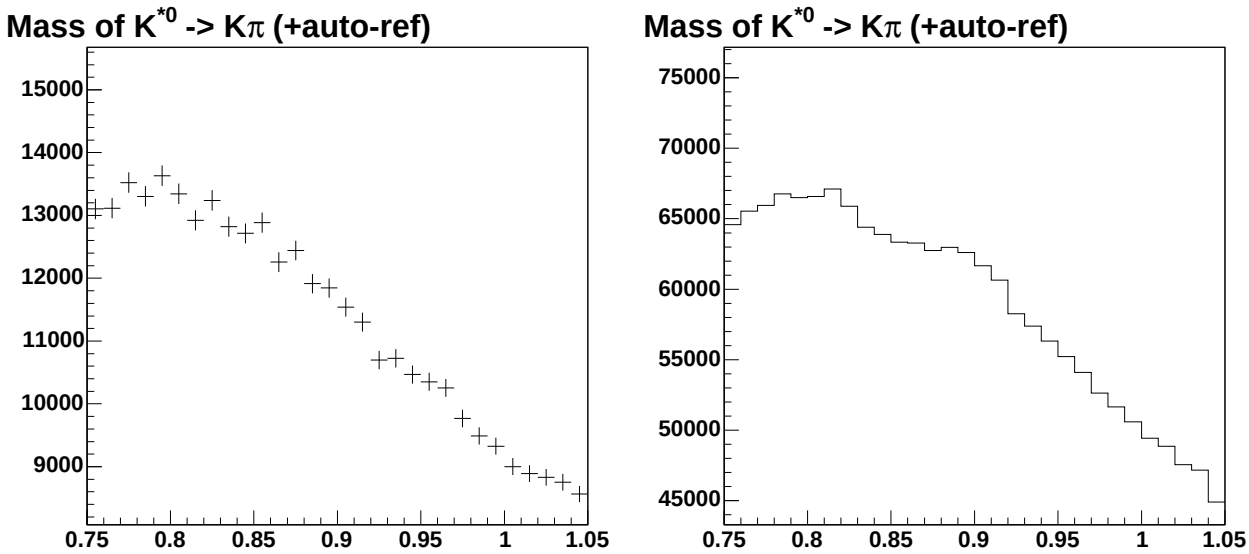

Figure 8: K^{*0} mass: data (left) and MC (right)

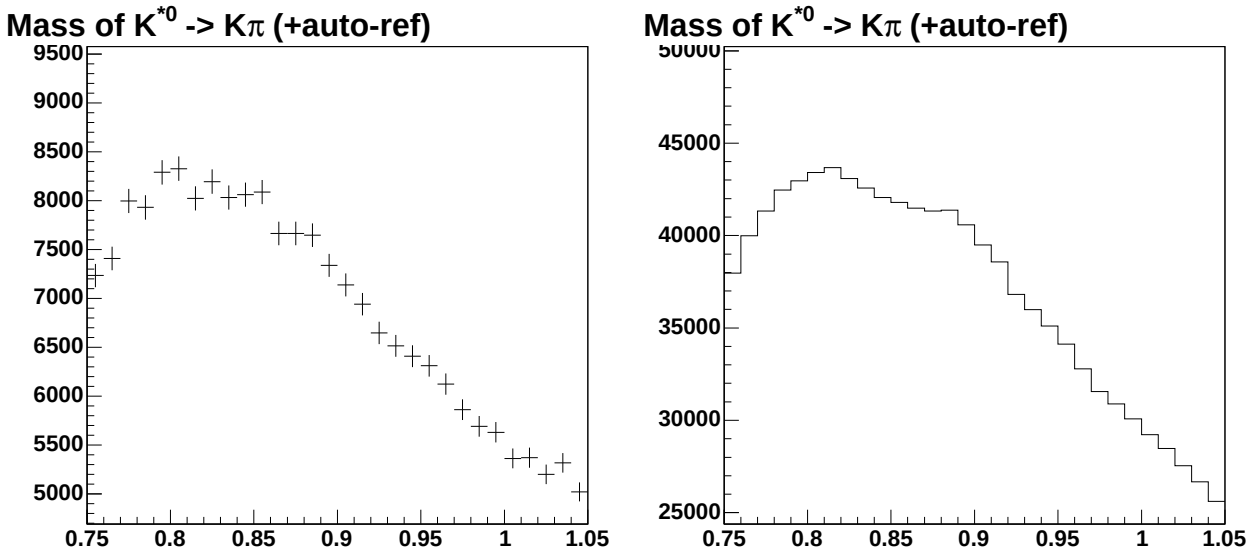

Figure 9: "Optimized" K^{*0} mass: data (left) and MC (right)

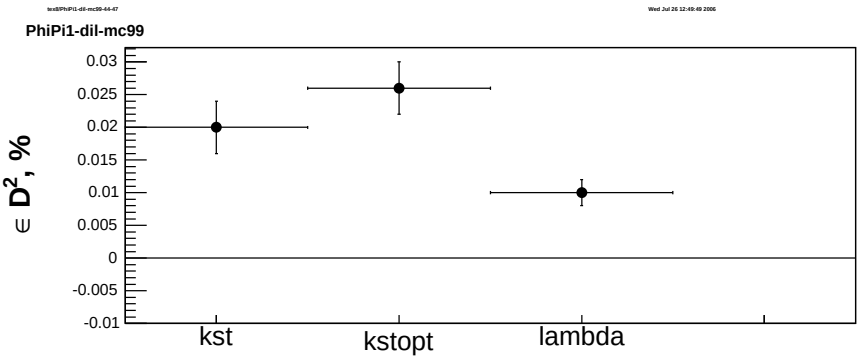

Figure 10: Comparison between εD^2's for two-track taggers for all four MC samples

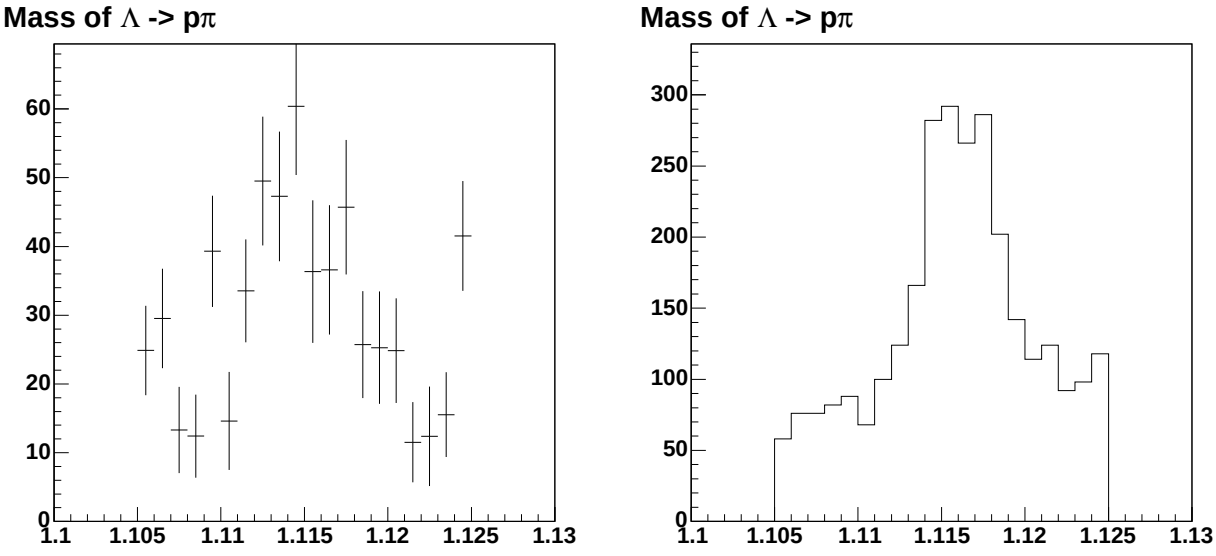

Figure 11: Λ mass: data (left) and MC (right)

Tagger	N_{RT}	N_{WT}	N_{NT}	ε, %	D, %	εD^2, %
Aver. Q	197545 ± 444	172167 ± 415	238548 ± 488	60.8 ± 0.1	6.9 ± 0.2	0.286 ± 0.013
$Q_{jet}(p_t, \kappa = 0.1)$	188330 ± 434	163042 ± 404	256888 ± 507	57.8 ± 0.1	7.2 ± 0.2	0.299 ± 0.013
$Q_{jet}(p_t, \kappa = 0.2)$	188129 ± 434	162808 ± 403	257323 ± 507	57.7 ± 0.1	7.2 ± 0.2	0.300 ± 0.013
$Q_{jet}(p_t, \kappa = 0.3)$	189392 ± 435	163607 ± 404	255261 ± 505	58.0 ± 0.1	7.3 ± 0.2	0.310 ± 0.013
$Q_{jet}(p_t, \kappa = 0.4)$	192047 ± 438	165993 ± 407	250220 ± 500	58.9 ± 0.1	7.3 ± 0.2	0.312 ± 0.013
$Q_{jet}(p_t, \kappa = 0.5)$	195423 ± 442	168852 ± 411	243985 ± 494	59.9 ± 0.1	7.3 ± 0.2	0.319 ± 0.014
$Q_{jet}(p_t, \kappa = 0.6)$	198826 ± 446	172039 ± 415	237395 ± 487	61.0 ± 0.1	7.2 ± 0.2	0.318 ± 0.014
$Q_{jet}(p_t, \kappa = 0.7)$	202067 ± 450	175233 ± 419	230960 ± 481	62.0 ± 0.1	7.1 ± 0.2	0.314 ± 0.014
$Q_{jet}(p_t, \kappa = 0.8)$	205351 ± 453	178374 ± 422	224535 ± 474	63.1 ± 0.1	7.0 ± 0.2	0.312 ± 0.014
$Q_{jet}(p_t, \kappa = 0.9)$	208375 ± 456	181533 ± 426	218352 ± 467	64.1 ± 0.1	6.9 ± 0.2	0.304 ± 0.013
$Q_{jet}(p_t, \kappa = 1.0)$	211435 ± 460	184584 ± 430	212241 ± 461	65.1 ± 0.1	6.8 ± 0.2	0.299 ± 0.013

Table 3: Comparison of different many-track taggers for all four MC samples

4.3 Using Weighted-Average Charge

The last group of same-side taggers is based on the weighted charge of **all** the tracks within cone $\cos\alpha > 0.8$, having 2+ axial SMT and 3+ axial CFT hits with B daughters excluded. We utilize three different methods of averaging:

- $Q_{jet}(p_t, \kappa) = \frac{\sum q \cdot p_t^{\kappa}}{\sum p_t^{\kappa}}$
- $Q_{jet}(p_t^{rel}, \kappa) = \frac{\sum q \cdot (p_t^{rel})^{\kappa}}{\sum (p_t^{rel})^{\kappa}}$
- $Q_{jet}(p_L^{rel}, \kappa) = \frac{\sum q \cdot (p_L^{rel})^{\kappa}}{\sum (p_L^{rel})^{\kappa}}$

Here p_t^{rel} and p_L^{rel} are $\perp$ and $||$ components of SST candidate's momentum $\vec{p}(K)$ w.r.t $\vec{p}(B_s)$. p_t is a transverse component of the SST candidate's momentum w.r.t. the beamline. The parameter κ increases the sensitivity to a particular region of p_t spectrum and must be optimized. If $\kappa = 0$, all three methods give the same answer – average charge of all qualifying tracks in the cone $\cos\alpha > 0.8$ around $\vec{p}(B_s)$.

The events with low absolute value of Q_{jet} have low dilution, so that we want to exclude them. For this reason we impose the cut $|Q_{jet}| > 0.2$. The results for different many-track taggers are given in Tables 3,4,5 and in graphical form in Figure 12. As we can see, all the taggers $Q_{jet}(p_t)$ and $Q_{jet}(p_L^{rel})$ have approximately the same tagging power. We choose the tagger "$Q_{jet}(p_t, \kappa = 0.6)$" as the best one, for consistency with opposite-side tagging.

5 SST Combination Technique

The combination of SSTs is performed according to the algorithm developed for OSTs in Ref. [2]. First, we look for any discriminating variables x_i which have different probability density functions $f_i^b(x_i)$ and $f_i^{\bar{b}}(x_i)$ for b and $\bar{b}$ quarks. Second, we form a ratio $y_i(x_i) = \frac{f_i^b(x_i)}{f_i^{\bar{b}}(x_i)}$. The case $y_i(x_i) > 1$ corresponds to b-quark, the opposite case $y_i(x_i) < 1$ – to $\bar{b}$-quark. Third, we define

Tagger	N_{RT}	N_{WT}	N_{NT}	$\varepsilon, \%$	D, %	$\varepsilon \mathrm{D}^2, \%$
Aver. Q	197545 ± 444	172167 ± 415	238548 ± 488	60.8 ± 0.1	6.9 ± 0.2	0.286 ± 0.013
$Q_{jet}(p_t^{rel}, \kappa = 0.1)$	188202 ± 434	163781 ± 405	256277 ± 506	57.9 ± 0.1	6.9 ± 0.2	0.279 ± 0.013
$Q_{jet}(p_t^{rel}, \kappa = 0.2)$	188421 ± 434	164974 ± 406	254865 ± 505	58.1 ± 0.1	6.6 ± 0.2	0.256 ± 0.012
$Q_{jet}(p_t^{rel}, \kappa = 0.3)$	190866 ± 437	168265 ± 410	249129 ± 499	59.0 ± 0.1	6.3 ± 0.2	0.234 ± 0.012
$Q_{jet}(p_t^{rel}, \kappa = 0.4)$	194136 ± 441	172415 ± 415	241709 ± 492	60.3 ± 0.1	5.9 ± 0.2	0.212 ± 0.011
$Q_{jet}(p_t^{rel}, \kappa = 0.5)$	197372 ± 444	176549 ± 420	234339 ± 484	61.5 ± 0.1	5.6 ± 0.2	0.191 ± 0.011
$Q_{jet}(p_t^{rel}, \kappa = 0.6)$	200681 ± 448	180750 ± 425	226829 ± 476	62.7 ± 0.1	5.2 ± 0.2	0.171 ± 0.010
$Q_{jet}(p_t^{rel}, \kappa = 0.7)$	203989 ± 452	184728 ± 430	219543 ± 469	63.9 ± 0.1	5.0 ± 0.2	0.157 ± 0.010
$Q_{jet}(p_t^{rel}, \kappa = 0.8)$	207201 ± 455	188658 ± 434	212401 ± 461	65.1 ± 0.1	4.7 ± 0.2	0.143 ± 0.009
$Q_{jet}(p_t^{rel}, \kappa = 0.9)$	210377 ± 459	192350 ± 439	205533 ± 453	66.2 ± 0.1	4.5 ± 0.2	0.133 ± 0.009
$Q_{jet}(p_t^{rel}, \kappa = 1.0)$	213159 ± 462	195837 ± 443	199264 ± 446	67.2 ± 0.1	4.2 ± 0.2	0.121 ± 0.009

Table 4: Comparison of different many-track taggers for all four MC samples

Tagger	N_{RT}	N_{WT}	N_{NT}	$\varepsilon, \%$	D, %	$\varepsilon \mathrm{D}^2, \%$
Aver. Q	197545 ± 444	172167 ± 415	238548 ± 488	60.8 ± 0.1	6.9 ± 0.2	0.286 ± 0.013
$Q_{jet}(p_L^{rel}, \kappa = 0.1)$	188378 ± 434	163099 ± 404	256783 ± 507	57.8 ± 0.1	7.2 ± 0.2	0.299 ± 0.013
$Q_{jet}(p_L^{rel}, \kappa = 0.2)$	188263 ± 434	163068 ± 404	256929 ± 507	57.8 ± 0.1	7.2 ± 0.2	0.297 ± 0.013
$Q_{jet}(p_L^{rel}, \kappa = 0.3)$	190070 ± 436	164648 ± 406	253542 ± 504	58.3 ± 0.1	7.2 ± 0.2	0.300 ± 0.013
$Q_{jet}(p_L^{rel}, \kappa = 0.4)$	193133 ± 439	167339 ± 409	247788 ± 498	59.3 ± 0.1	7.2 ± 0.2	0.303 ± 0.013
$Q_{jet}(p_L^{rel}, \kappa = 0.5)$	196695 ± 444	170455 ± 413	241110 ± 491	60.4 ± 0.1	7.1 ± 0.2	0.308 ± 0.013
$Q_{jet}(p_L^{rel}, \kappa = 0.6)$	200279 ± 448	173930 ± 417	234051 ± 484	61.5 ± 0.1	7.0 ± 0.2	0.305 ± 0.013
$Q_{jet}(p_L^{rel}, \kappa = 0.7)$	203674 ± 451	177589 ± 421	226997 ± 476	62.7 ± 0.1	6.8 ± 0.2	0.293 ± 0.013
$Q_{jet}(p_L^{rel}, \kappa = 0.8)$	207043 ± 455	180880 ± 425	220337 ± 469	63.8 ± 0.1	6.7 ± 0.2	0.290 ± 0.013
$Q_{jet}(p_L^{rel}, \kappa = 0.9)$	210187 ± 458	184304 ± 429	213769 ± 462	64.9 ± 0.1	6.6 ± 0.2	0.279 ± 0.013
$Q_{jet}(p_L^{rel}, \kappa = 1.0)$	213312 ± 462	187417 ± 433	207531 ± 456	65.9 ± 0.1	6.5 ± 0.2	0.275 ± 0.013

Table 5: Comparison of different many-track taggers for all four MC samples

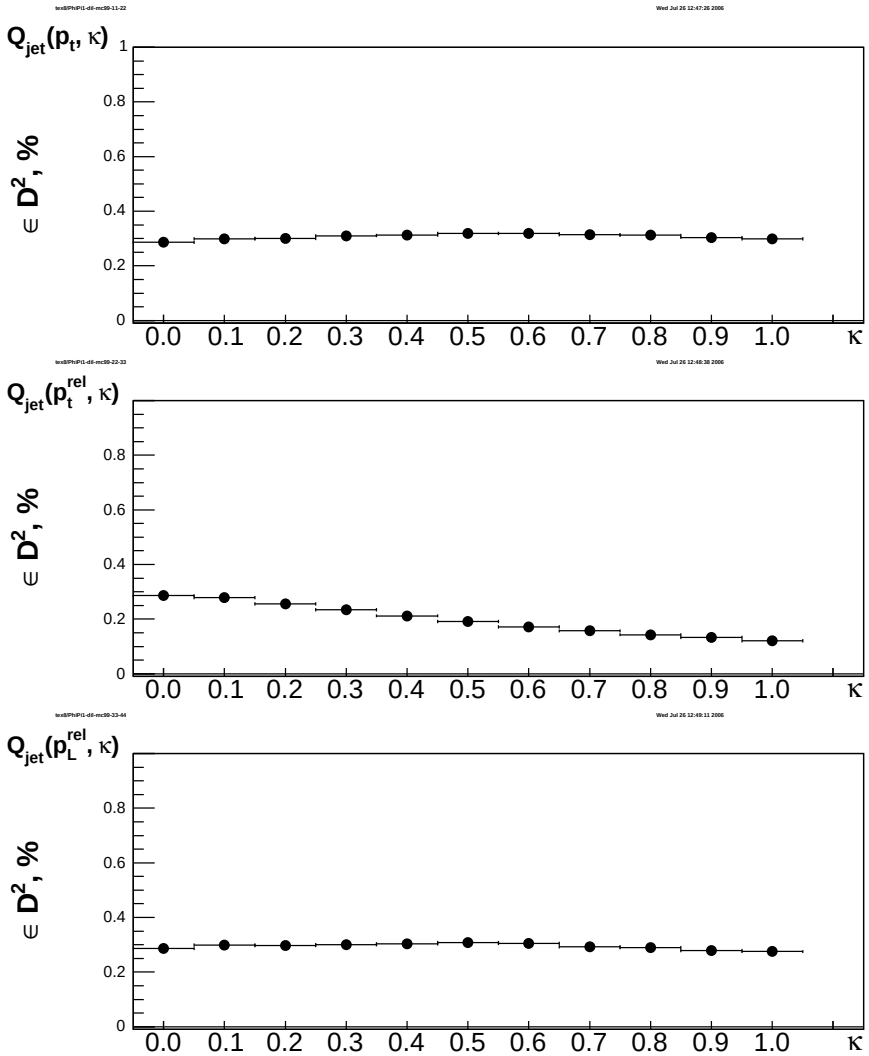

Figure 12: Comparison between εD^2's for many-track taggers for all four MC samples

a variable $y(\vec{x}) = \prod_{i=1}^{n} y_i(x_i)$ which accumulates information from all the discriminating variables. Now, for each event, $y(\vec{x}) > 1$ corresponds to b-quark, the opposite case $y(\vec{x}) < 1$ – to $\bar{b}$-quark. It is more convenient to introduce a *combined dilution* $d = \frac{1-y}{1+y}$ for each event and infer b-quark flavor from its sign. For combination we use four least correlated between themselves taggers from different groups: "Min. ΔR", "K^{*0}", "Lambda" and "$Q_{jet}(p_t, \kappa = 0.6)$". The variables x_i are: $x_1 = q \cdot \Delta R$, $x_2 = q \cdot (m(K^{*0}) - 0.862)/(0.922 - 0.862)$, $x_3 = q \cdot (m(\Lambda) - 1.105)/(1.125 - 1.105)$, and $x_4 = Q_{jet}$, where q is the charge of the tag.

For each event the B-meson production flavor is obtained from the Monte Carlo truth information and then corresponding histogram is filled to produce p.d.f. $f_i^b(x_i)$ or $f_i^{\bar{b}}(x_i)$. The probability density functions for all four variables x_1, x_2, x_3, x_4 are shown in Figure 13 and their ratios – in Figure 14. We construct the variable y as the product of these ratios and then compute the combined dilution d for each event. The distribution of this quantity d, shown in Figure 15, gives us the largest discrimination between b and $\bar{b}$ quarks. The Right and Wrong-Sign combinations are determined from the correlation of the sign of d variable and the flavor of B_s meson. The events with lower value of d have lower dilution, so that we want to exclude them by imposing a cut $|d| > 0.086$. This number was optimized to obtain the largest tagging power ϵD^2. The resulting dilution is $D = 9.0 \pm 0.2$ % and $\epsilon D^2 = 0.442 \pm 0.016$ %. This is to be compared to the individual ϵD^2's of the taggers:

$$
\begin{aligned}
\epsilon D^2(\text{Min.}\Delta R) &= 0.351 \pm 0.014 \ \% \\
\epsilon D^2(K^{*0}) &= 0.020 \pm 0.004 \ \% \\
\epsilon D^2(\text{Lambda}) &= 0.010 \pm 0.002 \ \% \\
\epsilon D^2(Q_{jet}(p_t, \kappa = 0.6)) &= 0.318 \pm 0.014 \ \%
\end{aligned}
$$

We see the increase in the tagging power due to tagger combination.

6 SST + OST Combination with Double-Tagged Events:

Using both SST and OST on the same data sample allows one to measure the product of their dilutions $D_{OST} \cdot D_{SST}$ (see Appendix A). By using D_{OST} measured earlier ($D_{OST} = (44.3 \pm 2.2)\%$ for combined dilution $|d| > 0.3$, Ref. [3]), we can obtain the D_{SST} purely from the data (Eqn. 3). Then we can compute the tagging power ϵD^2 of the combination of SST and OST (Eqn. 4). We do so for each considered SST algorithm and for final combined SST algorithm for the data in `/prj_root/1008/ckm_write/bgv/evt/muphipi-std/` (1 fb^{-1}). The examples of the fits are displayed in Figures 16 and 17. The distributions are fitted with two single gaussians plus parabolic background. The means and widths of the gaussians and the shape of the background are fixed to the values from the global fit in Figure 4. The resulting numbers coming from the fits, together with calculated dilutions are given in Tables 6 and 7. As we can see, the measured dilutions in the data and *true* dilutions in the MC are reasonably close to each other. The efficiencies in data and Monte Carlo are close too. We display the data-Monte Carlo comparisons for dilutions and efficiencies in Figures 18, 19, 20 and Figures 21, 22, 23, correspondingly.

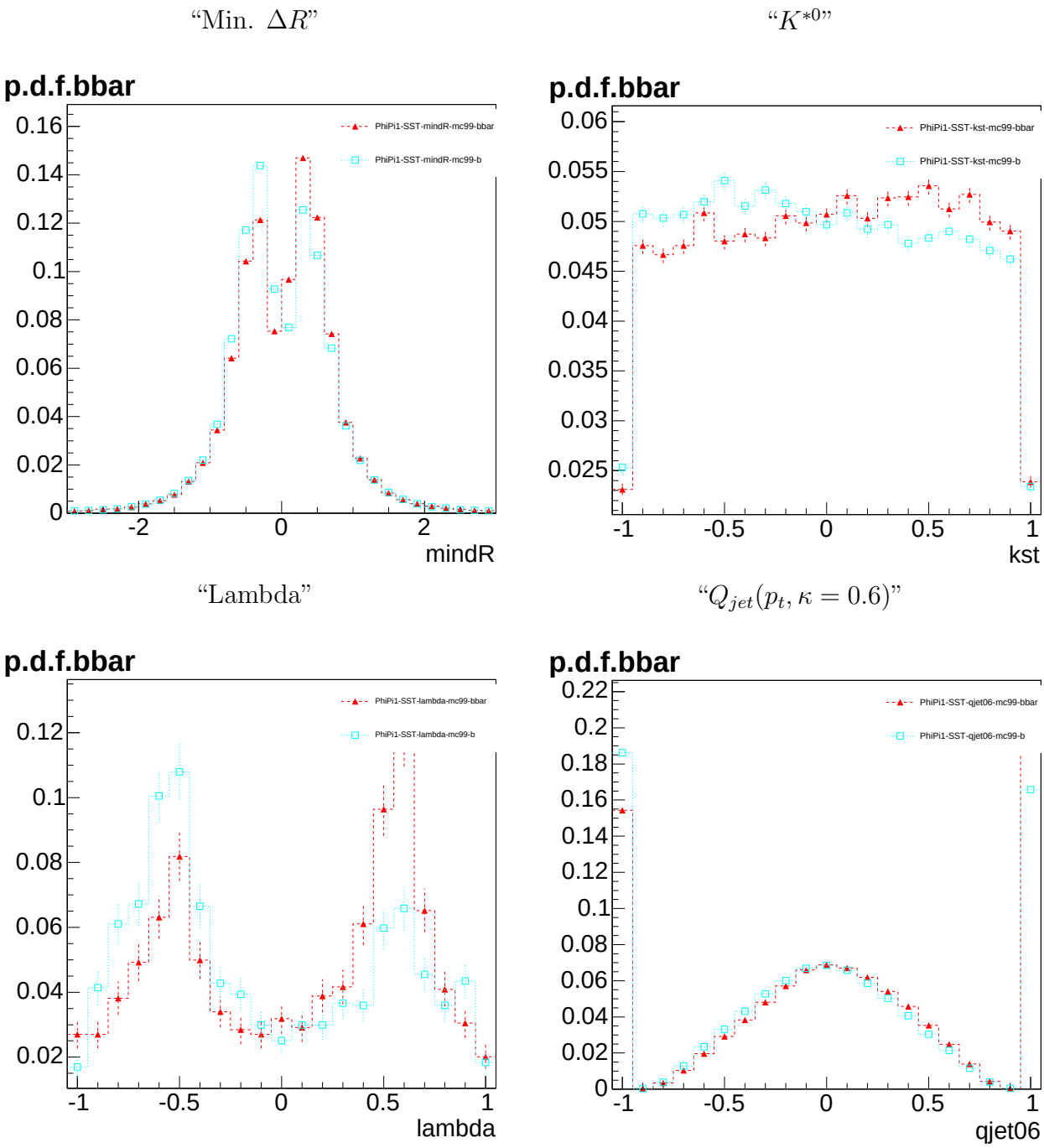

Figure 13: The probability density functions for "Min. ΔR", "K^{*0}", "Lambda" and "$Q_{jet}(p_t, \kappa = 0.6)$". Red full triangles - p.d.f. for $\bar{b}$-quark, cyan open squares - p.d.f. for b-quark. Both histograms are normalized to unity area.

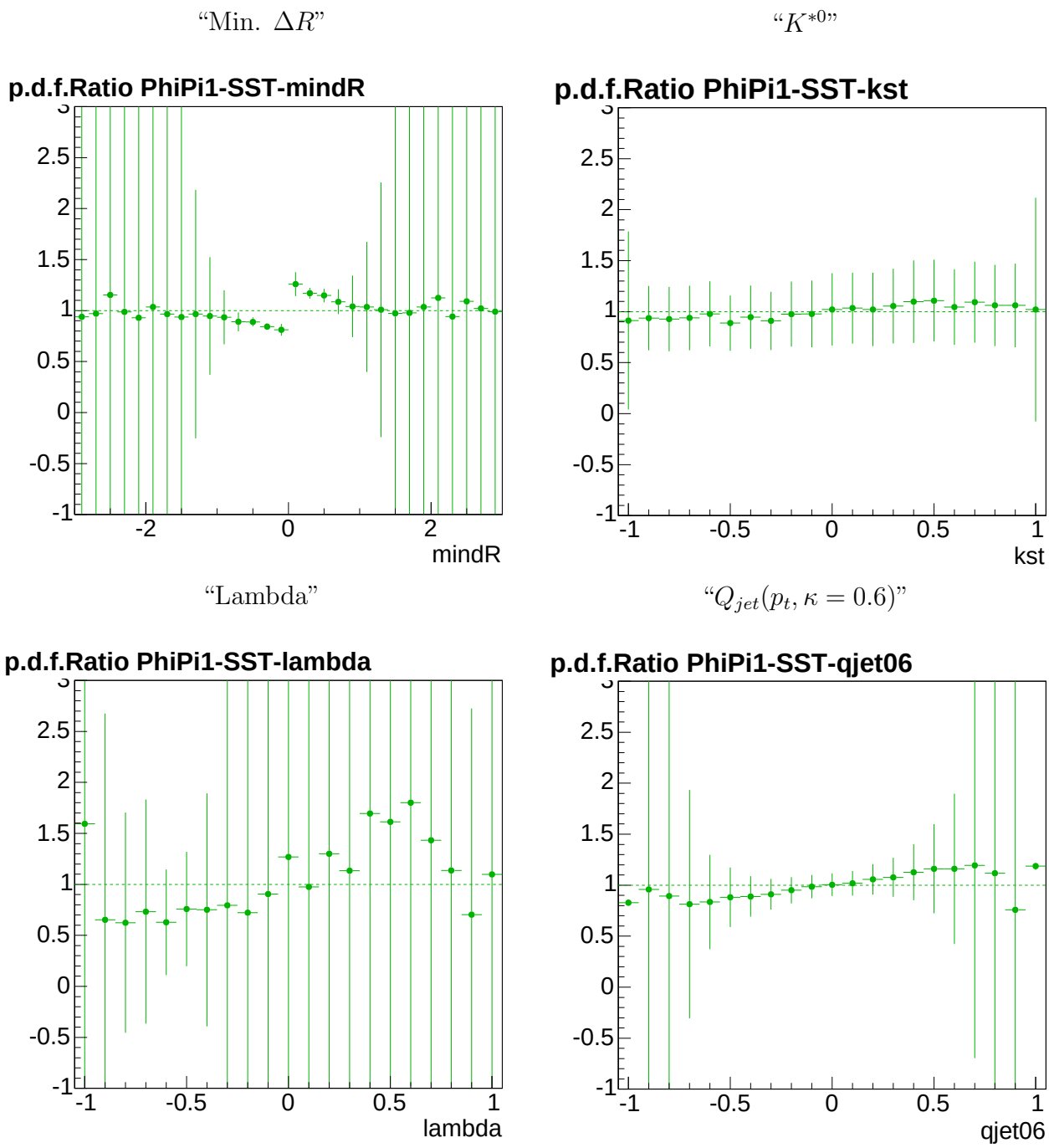

Figure 14: The ratios of the probability density functions for "Min. ΔR", "K^{*0}", "Lambda" and "$Q_{jet}(p_t, \kappa = 0.6)$".

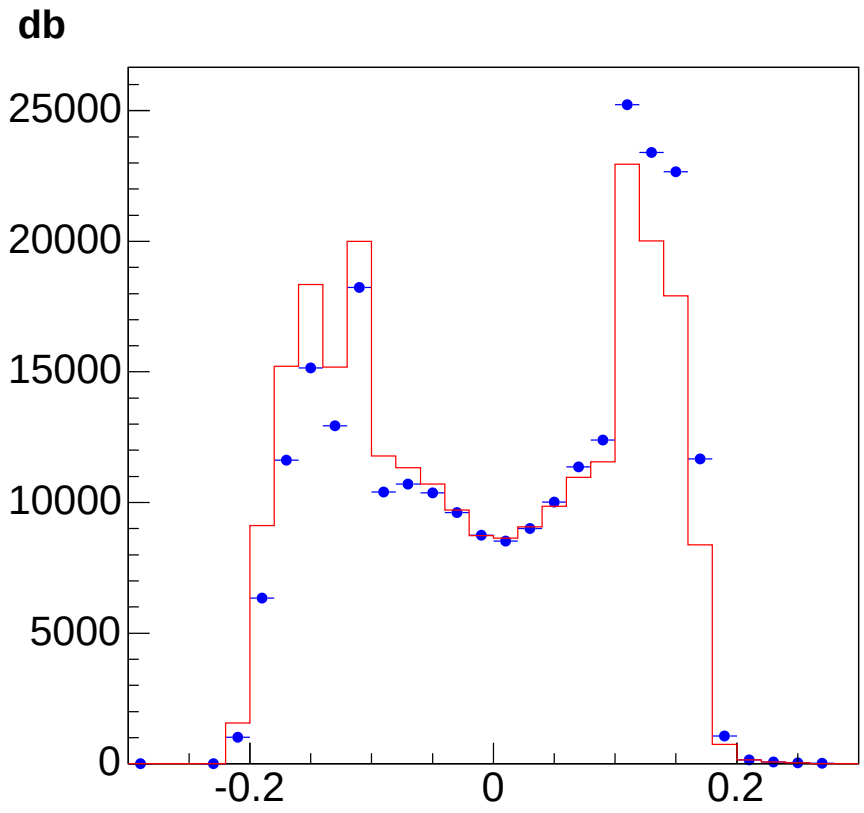

Figure 15: Distribution of combined dilution d

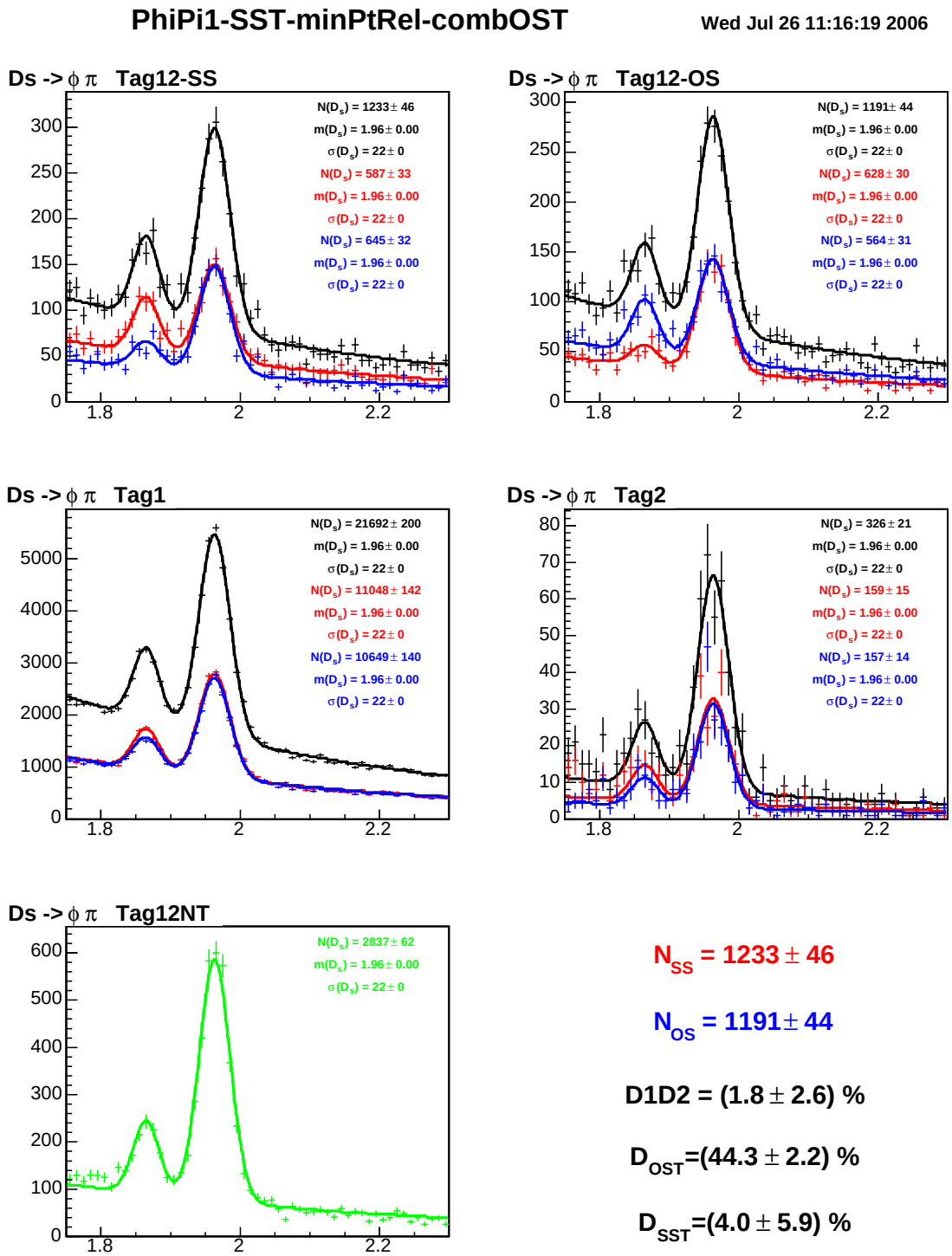

Figure 16: Fit of the events double-tagged with "Min. p_t^{rel}" and OST. Red histograms show events with the sign of the tag identical to the sign of the trigger lepton. Blue histograms display events for the sign of the tag opposite to the sign of the trigger lepton. Green histogram demonstrates the not-tagged events.

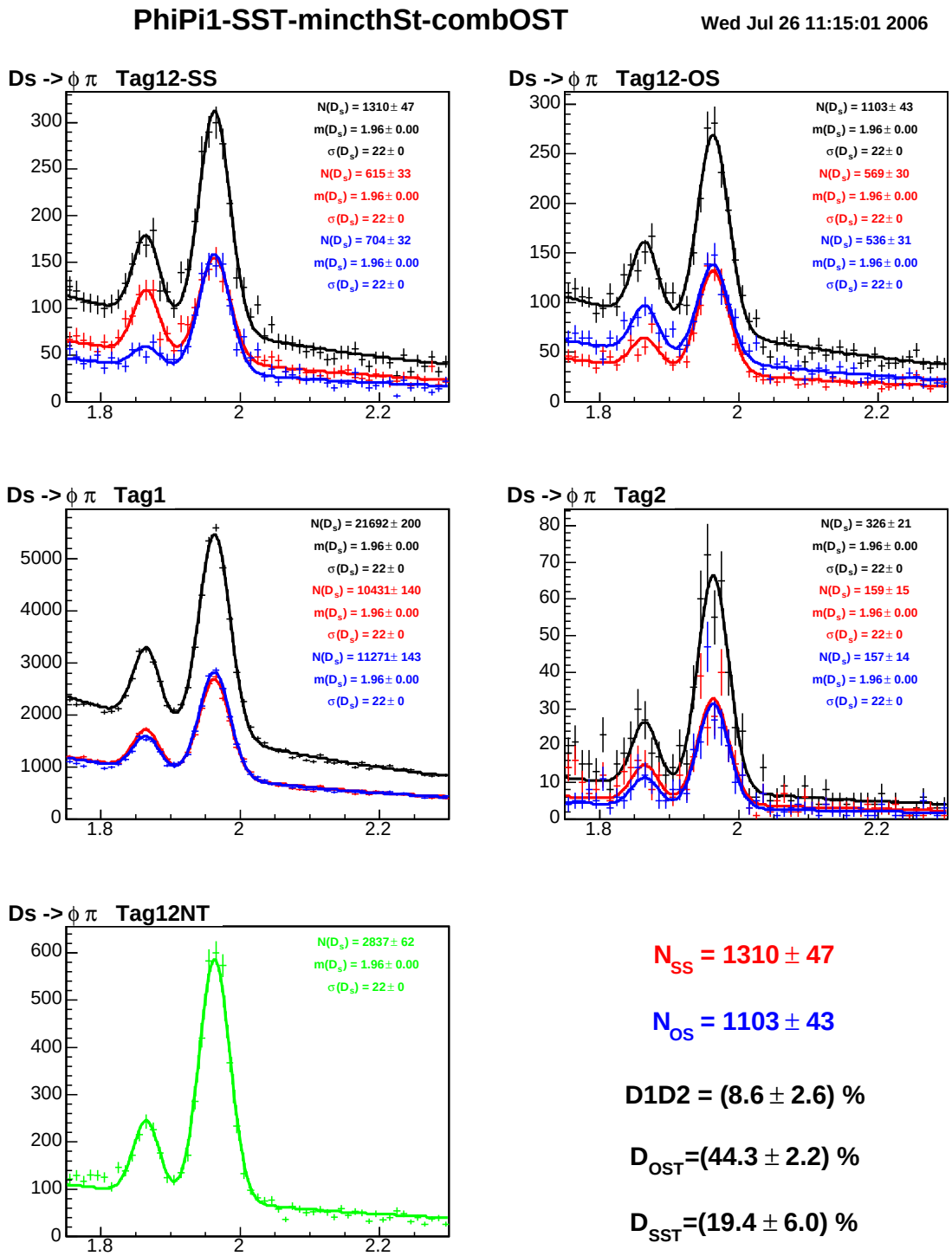

Figure 17: Fit of the events double-tagged with "Min. $\cos\theta^*$" and OST. Red histograms show events with the sign of the tag identical to the sign of the trigger lepton. Blue histograms display events for the sign of the tag opposite to the sign of the trigger lepton. Green histogram demonstrates the not-tagged events.

	N_1	N_2	N_{NT}	N_{12}	$\bar{N}_{12}$	$\frac{N_{12}-\bar{N}_{12}}{N_{12}+\bar{N}_{12}}$	ϵ^{meas}	D^{meas}_{SST}	D^{calc}_{12}	$\bar{D}^{calc}_{12}$	εD^2(calc), %
Min. $p_t^{\rm rel}$	21692±200	326±21	2837±62	1233±46	1191±44	0.018±0.026	89.6±0.1	4.0±5.9	47.4±5.1	41.0±5.4	2.114±0.478
Max. $p_L^{\rm rel}$	21692±200	326±21	2837±62	1300±47	1116±43	0.076±0.026	89.6±0.1	17.2±6.0	57.1±4.5	29.3±6.2	4.494±1.661
Max. p_t	21692±200	326±21	2837±62	1276±47	1138±43	0.057±0.026	89.6±0.1	12.9±6.0	54.1±4.7	33.3±5.9	3.391±1.258
Min. $\lvert\Delta\vec{P}\rvert$	21692±200	326±21	2837±62	1299±47	1118±43	0.075±0.026	89.6±0.1	16.9±6.0	56.9±4.6	29.6±6.1	4.410±1.634
Min. ΔR	21692±200	326±21	2837±62	1280±47	1136±43	0.060±0.026	89.6±0.1	13.4±6.0	54.5±4.7	32.8±5.9	3.514±1.308
Max. $\cos\alpha$	21692±200	326±21	2837±62	1283±47	1137±43	0.060±0.026	89.6±0.1	13.6±6.0	54.6±4.7	32.7±5.9	3.547±1.319
Min. $\cos\theta^*$	21692±200	326±21	2837±62	1310±47	1103±43	0.086±0.026	89.6±0.1	19.4±6.0	58.6±4.5	27.3±6.3	5.176±1.875
Max. $\cos\theta^*$	21692±200	326±21	2837±62	1225±46	1197±44	0.012±0.026	89.6±0.1	2.7±5.9	46.4±5.1	42.1±5.4	2.038±0.389
Min. $m(B_sK)$	21692±200	326±21	2837±62	1271±46	1156±44	0.048±0.026	89.6±0.1	10.8±5.9	52.6±4.8	35.2±5.8	2.968±1.059
Random	21692±200	326±21	2837±62	1219±46	1196±44	0.009±0.026	89.6±0.1	2.1±6.0	46.0±5.2	42.6±5.4	2.012±0.359
$K^{*0} \to K\pi$	5661±111	2160±59	18890±177	291±23	293±24	-0.004±0.057	30.8±0.3	-0.9±12.9	43.6±10.7	45.0±10.5	1.974±0.220
$K^{*0} \to K\pi$ (opt)	4287±96	2316±61	20270±186	205±20	216±20	-0.027±0.068	25.7±0.3	-6.1±15.3	39.3±13.2	49.1±11.9	2.029±0.360
"Comb. SST"	21692±200	326±21	2837±62	1279±47	1139±43	0.058±0.026	89.6±0.1	13.1±6.0	54.3±4.7	33.1±5.9	3.442±1.278

Table 6: D_{SST} measured in data and total ϵD^2 for one-track, two-track and combined taggers

	N_1	N_2	N_{NT}	N_{12}	$\bar{N}_{12}$	$\frac{N_{12}-\bar{N}_{12}}{N_{12}+\bar{N}_{12}}$	ϵ^{meas}	D^{meas}_{SST}	D^{calc}_{12}	$\bar{D}^{calc}_{12}$	εD^2(calc), %
Aver. $Q(\kappa=0.0)$	15771±166	983±40	8769±125	941±39	819±36	0.069±0.030	67.9±0.2	15.6±6.8	56.1±5.1	30.8±6.8	3.488±1.255
$Q_{jet}(p_t,\kappa=0.1)$	14847±161	1077±42	9689±134	895±38	774±34	0.073±0.030	64.5±0.2	16.4±6.9	56.6±5.2	30.1±6.9	3.548±1.257
$Q_{jet}(p_t,\kappa=0.2)$	14806±161	1086±42	9732±134	890±38	765±35	0.076±0.031	64.3±0.2	17.1±7.0	57.1±5.2	29.5±7.0	3.666±1.318
$Q_{jet}(p_t,\kappa=0.3)$	14916±161	1103±42	9624±133	889±38	759±35	0.079±0.031	64.7±0.2	17.9±7.0	57.6±5.2	28.7±7.1	3.845±1.392
$Q_{jet}(p_t,\kappa=0.4)$	15114±162	1068±42	9429±132	911±38	769±35	0.084±0.031	65.5±0.2	19.0±7.0	58.4±5.1	27.6±7.1	4.124±1.487
$Q_{jet}(p_t,\kappa=0.5)$	15387±164	1028±41	9152±130	940±39	781±35	0.092±0.030	66.5±0.2	20.9±6.9	59.6±5.0	25.8±7.2	4.609±1.638
$Q_{jet}(p_t,\kappa=0.6)$	15664±166	995±40	8878±127	957±39	794±35	0.093±0.030	67.5±0.2	21.1±6.9	59.8±4.9	25.6±7.2	4.714±1.679
$Q_{jet}(p_t,\kappa=0.7)$	16019±168	959±39	8520±125	991±40	797±36	0.108±0.030	68.8±0.2	24.4±6.8	62.0±4.8	22.3±7.4	5.725±1.976
$Q_{jet}(p_t,\kappa=0.8)$	16362±170	922±38	8178±122	1010±40	817±36	0.106±0.030	70.0±0.2	23.9±6.8	61.7±4.8	22.8±7.3	5.644±1.956
$Q_{jet}(p_t,\kappa=0.9)$	16684±172	891±38	7857±119	1016±41	841±37	0.094±0.029	71.2±0.2	21.3±6.7	59.9±4.8	25.4±7.1	4.953±1.769
$Q_{jet}(p_t,\kappa=1.0)$	16966±174	859±37	7573±117	1034±41	851±37	0.097±0.029	72.2±0.2	21.9±6.7	60.3±4.8	24.8±7.1	5.164±1.837
$Q_{jet}(p_t^{rel},\kappa=0.1)$	14854±161	1074±42	9683±133	900±38	773±35	0.076±0.031	64.5±0.2	17.1±7.0	57.0±5.2	29.5±7.0	3.677±1.314
$Q_{jet}(p_t^{rel},\kappa=0.2)$	14915±162	1063±42	9619±133	922±38	764±35	0.094±0.030	64.7±0.2	21.1±7.0	59.8±5.0	25.6±7.3	4.596±1.625
$Q_{jet}(p_t^{rel},\kappa=0.3)$	15091±163	1030±41	9445±131	928±38	786±35	0.083±0.030	65.4±0.2	18.7±6.9	58.1±5.1	28.0±7.0	4.041±1.441
$Q_{jet}(p_t^{rel},\kappa=0.4)$	15386±165	1003±40	9149±129	932±39	812±36	0.069±0.030	66.5±0.2	15.5±6.8	56.0±5.2	30.9±6.8	3.432±1.221
$Q_{jet}(p_t^{rel},\kappa=0.5)$	15787±167	934±39	8749±125	975±39	839±36	0.075±0.029	67.9±0.2	17.0±6.7	57.0±5.0	29.6±6.8	3.767±1.339
$Q_{jet}(p_t^{rel},\kappa=0.6)$	16187±169	909±38	8348±123	986±40	852±37	0.073±0.029	69.4±0.2	16.4±6.7	56.6±5.0	30.1±6.7	3.696±1.326
$Q_{jet}(p_t^{rel},\kappa=0.7)$	16445±171	858±37	8090±121	1015±40	868±37	0.078±0.029	70.3±0.2	17.6±6.6	57.4±4.9	29.0±6.7	3.973±1.423
$Q_{jet}(p_t^{rel},\kappa=0.8)$	16739±173	825±36	7798±118	1029±41	884±38	0.076±0.029	71.4±0.2	17.2±6.6	57.1±4.9	29.4±6.7	3.910±1.408
$Q_{jet}(p_t^{rel},\kappa=0.9)$	17083±174	801±36	7458±115	1048±41	887±38	0.083±0.029	72.7±0.2	18.8±6.6	58.2±4.8	27.8±6.8	4.341±1.564
$Q_{jet}(p_t^{rel},\kappa=1.0)$	17366±176	757±35	7175±113	1067±41	914±38	0.077±0.028	73.7±0.2	17.5±6.5	57.3±4.8	29.1±6.6	4.053±1.460
$Q_{jet}(p_L^{rel},\kappa=0.1)$	14850±161	1080±42	9687±134	899±38	767±35	0.079±0.031	64.5±0.2	17.9±7.0	57.6±5.1	28.7±7.1	3.844±1.378
$Q_{jet}(p_L^{rel},\kappa=0.2)$	14812±161	1085±42	9725±134	903±38	759±35	0.087±0.031	64.4±0.2	19.6±7.0	58.8±5.1	27.1±7.2	4.212±1.508
$Q_{jet}(p_L^{rel},\kappa=0.3)$	15033±162	1081±42	9512±133	914±38	753±35	0.096±0.031	65.1±0.2	21.8±7.0	60.3±5.0	24.9±7.4	4.777±1.699
$Q_{jet}(p_L^{rel},\kappa=0.4)$	15234±163	1046±41	9306±131	925±38	779±35	0.086±0.030	65.9±0.2	19.4±6.9	58.6±5.0	27.3±7.1	4.228±1.515
$Q_{jet}(p_L^{rel},\kappa=0.5)$	15516±165	1016±40	9023±129	945±39	789±35	0.090±0.030	66.9±0.2	20.3±6.9	59.3±5.0	26.3±7.2	4.501±1.612
$Q_{jet}(p_L^{rel},\kappa=0.6)$	15817±167	969±39	8716±125	970±39	805±36	0.093±0.030	68.0±0.2	21.0±6.8	59.7±4.9	25.7±7.1	4.712±1.675
$Q_{jet}(p_L^{rel},\kappa=0.7)$	16104±169	944±39	8435±124	992±40	809±36	0.101±0.030	69.1±0.2	22.8±6.8	61.0±4.8	23.9±7.2	5.274±1.850
$Q_{jet}(p_L^{rel},\kappa=0.8)$	16435±171	889±38	8101±121	1013±40	841±37	0.093±0.029	70.3±0.2	21.0±6.7	59.7±4.8	25.7±7.0	4.820±1.712
$Q_{jet}(p_L^{rel},\kappa=0.9)$	16785±173	869±37	7751±118	1026±41	851±37	0.093±0.029	71.6±0.2	21.0±6.7	59.8±4.8	25.7±7.0	4.890±1.749
$Q_{jet}(p_L^{rel},\kappa=1.0)$	17069±174	855±36	7469±116	1039±41	852±37	0.099±0.029	72.6±0.2	22.3±6.7	60.6±4.8	24.4±7.1	5.309±1.887

Table 7: D_{SST} measured in data and total ϵD^2 for many-track taggers

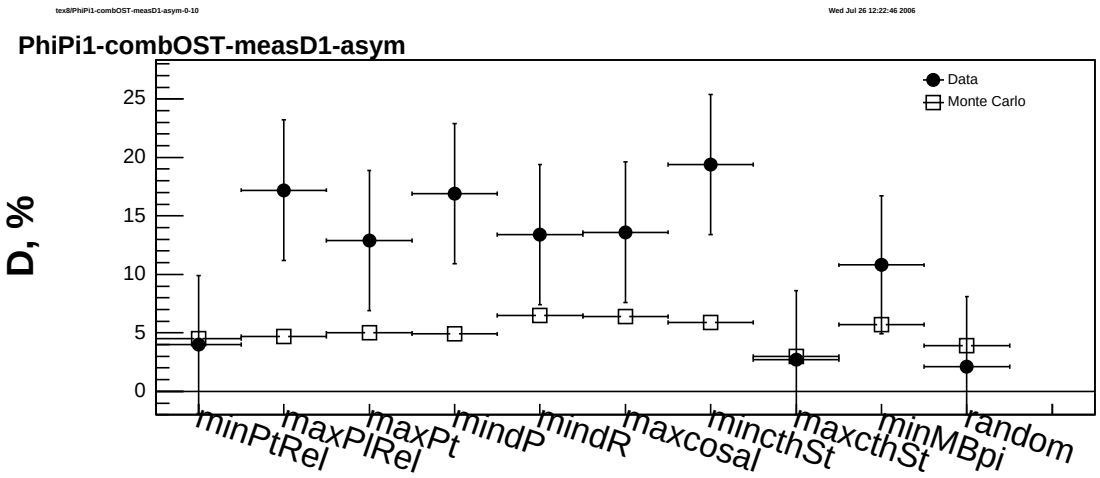

Figure 18: Data – Monte Carlo comparison for dilutions for one-track taggers.

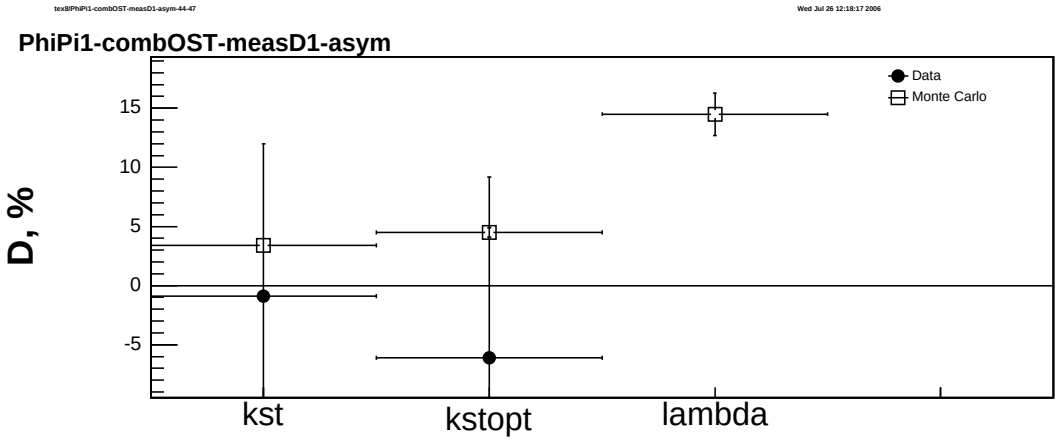

Figure 19: Data – Monte Carlo comparison for dilutions for two-track taggers.

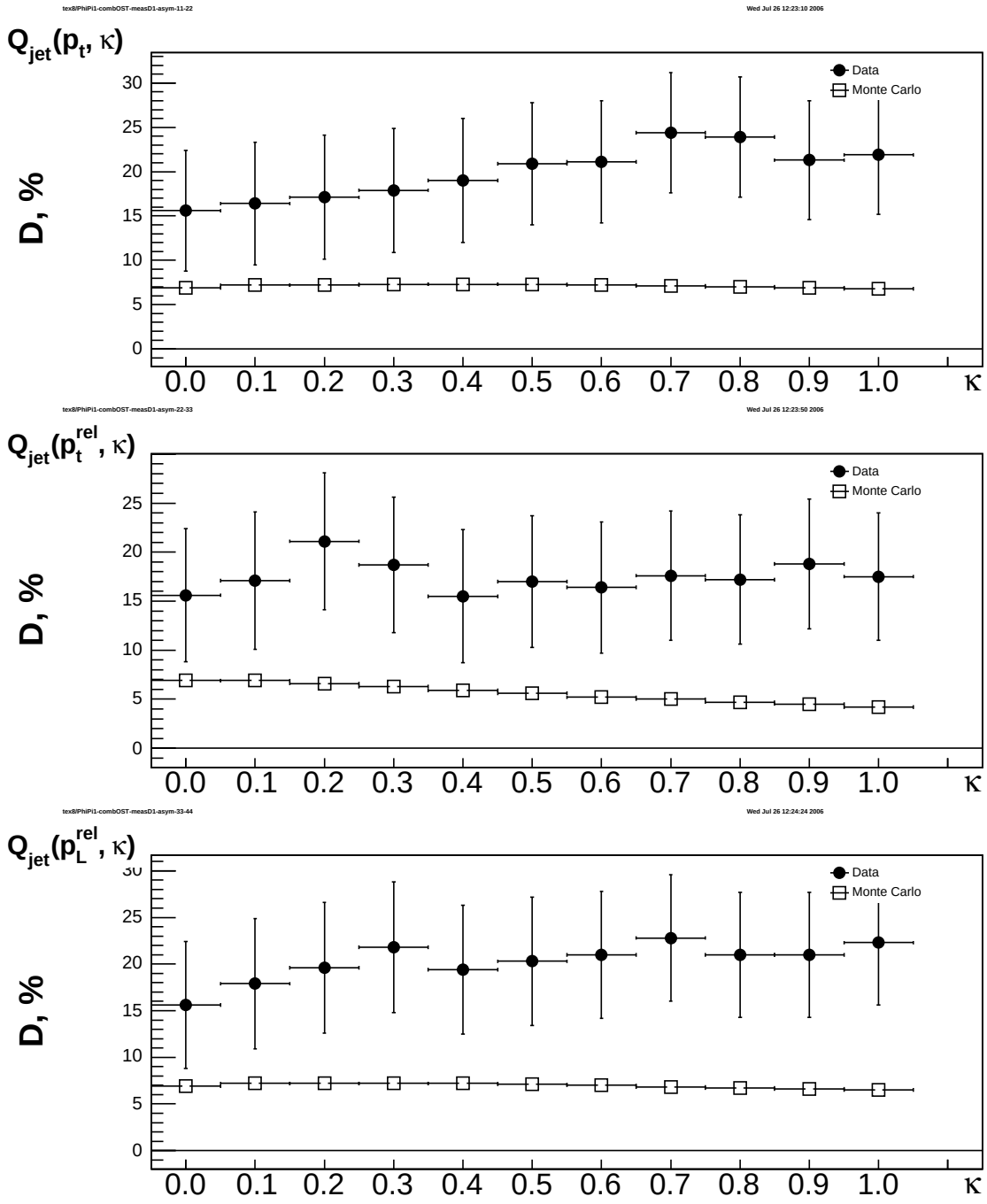

Figure 20: Data – Monte Carlo comparison for dilutions for many-track taggers.

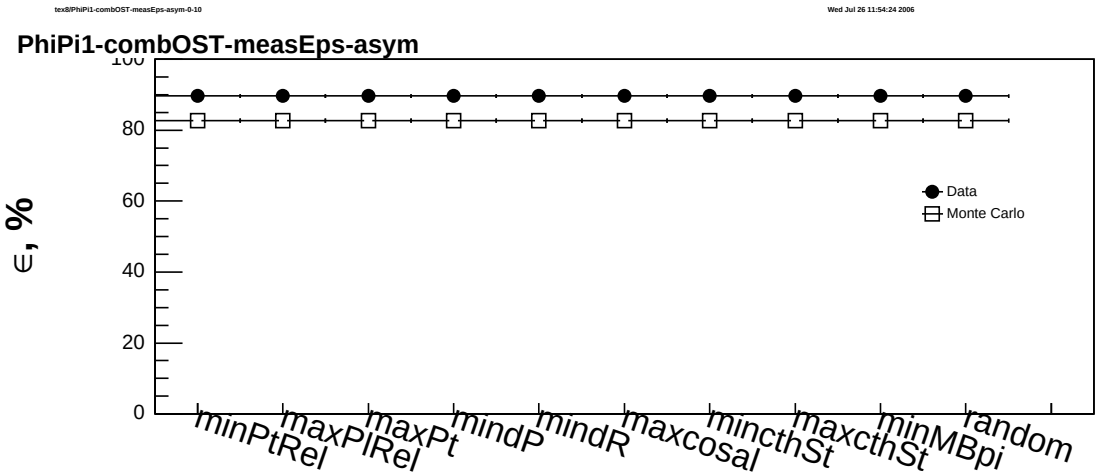

Figure 21: Data – Monte Carlo comparison for efficiencies for one-track taggers.

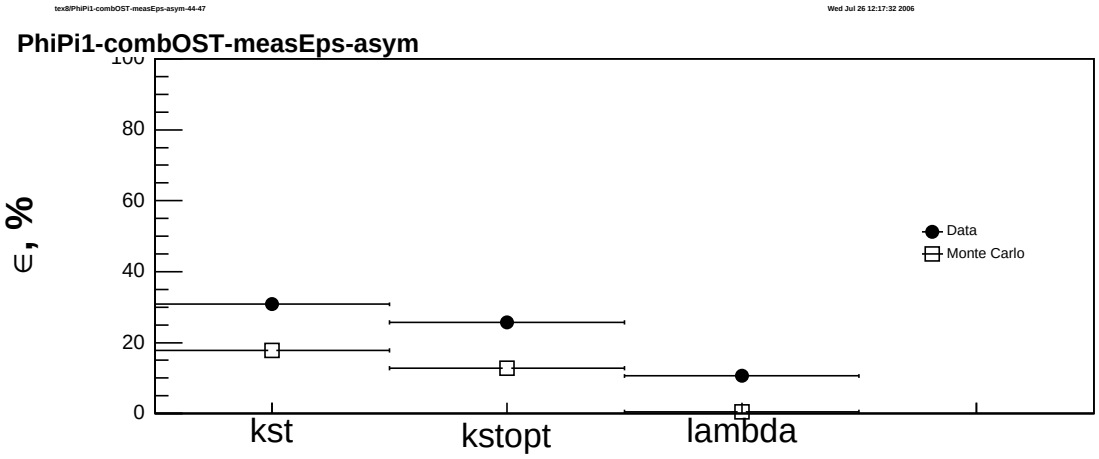

Figure 22: Data – Monte Carlo comparison for efficiencies for two-track taggers.

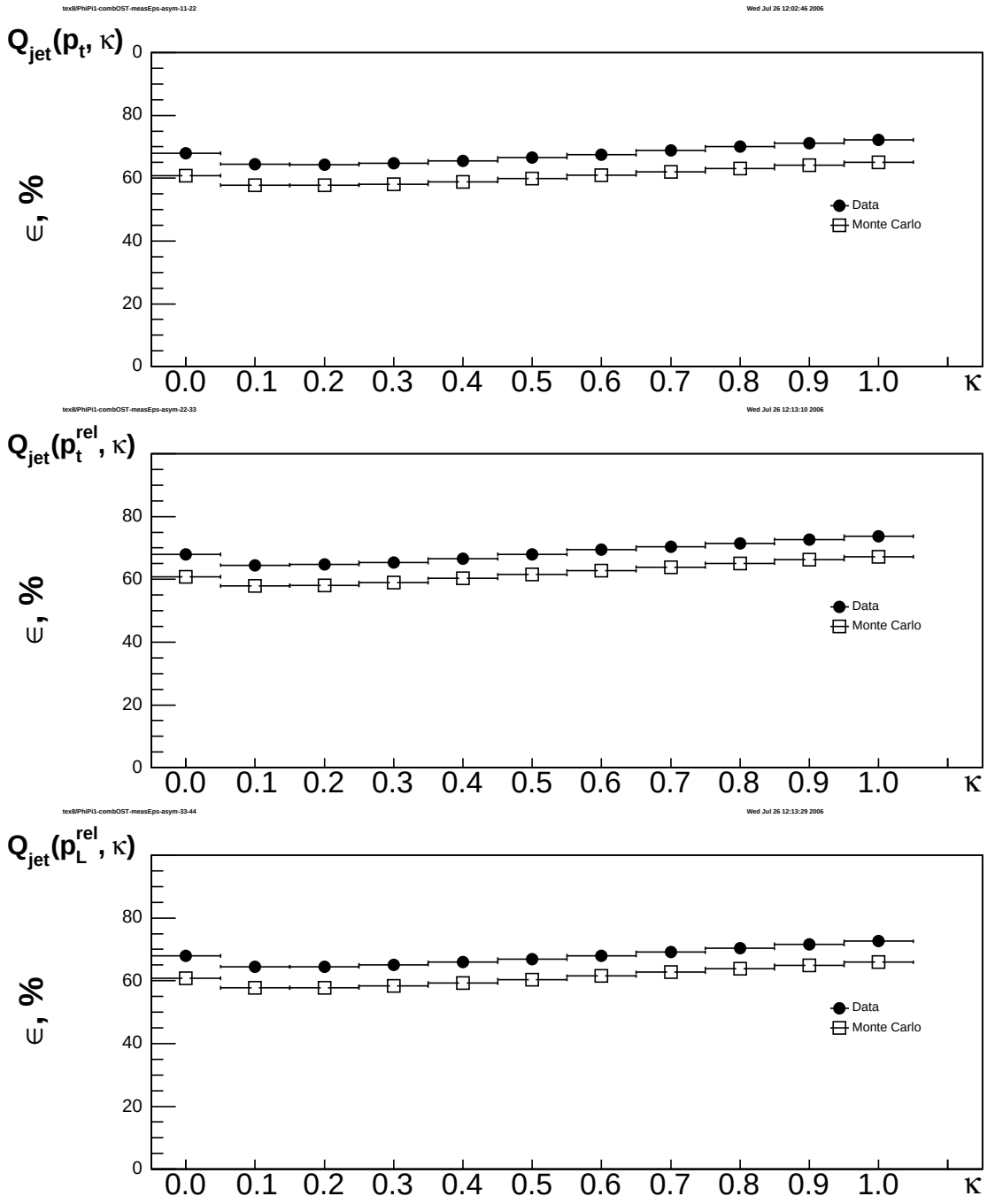

Figure 23: Data – Monte Carlo comparison for efficiencies for many-track taggers.

7 Conclusion

In this note a technique developed for combination of opposite-side taggers (Ref. [2]) is implemented for same-side taggers. A few same-side taggers are chosen ("Min. ΔR", "K^{*0}", "Lambda" and "$Q_{jet}(p_t, \kappa = 0.6)$") to be combined together. The resulting dilution of the combined same-side tagger in MC is $D = (9.0\pm0.2\ )\%$ and the combined tagging power is $\epsilon D^2 = (0.442\pm0.016\ )\%$

Also, double-tagged events are used to apply both "Comb. SST" and previously developed "Comb. OST" to the data. This allows to measure SST dilution from data only, by using earlier measured OST dilution (Ref. [3])). The total ϵD^2's for combination SST + OST are given in Tables 6 and 7. The dilution for "Comb. SST" is $D = (13.1\pm6.0\ )\%$ and the total tagging power for "Comb. SST" and "Comb. OST" is $\epsilon D^2 = (3.442\pm1.278\ )\%$ This number is higher than ϵD^2 for OST only ($\epsilon D^2 = 2.19\pm0.22$, Ref. [3]), as one would expect.

References

[1] ArXiv.org: hep/lat-0510113

[2] G. Borissov *et al.*, *Combined Opposite-side Flavor Tagging*, **D0 Note 4875**, (July 8, 2005).

[3] G. Borissov *et al.*, *B_d mixing measurement using Opposite-side Flavor Tagging*, **D0 Note 4991**, (Feb 17, 2006).

[4] P. Sphicas, *Combining Flavor-Taggers*, **CDF Note 3425**, (Dec 9, 1995).

A Appendix: Double Tagging with Uncorrelated Taggers

Suppose that we have two uncorrelated taggers with mistag rates p_1 and p_2 and that we apply them to a sample of N events. This sample can be separated into five subsamples:

- N_1 events tagged only by first tagger with dilution $D_1 = 1 - 2p_1$
- N_2 events tagged only by second tagger with dilution $D_2 = 1 - 2p_2$
- N_{12} events tagged by both taggers identically with *true* dilution D_{12}
- $\bar{N}_{12}$ events tagged by both taggers differently with *true* dilution $\bar{\mathrm{D}}_{12}$
- N_{NT} events not tagged by both taggers

Let us consider the third group of events. The probability for the event to fall into this group is equal to $p_1p_2+(1-p_1)(1-p_2)$. The mistag rate for this group, then, equals $p_{12} = \frac{p_1p_2}{p_1p_2+(1-p_1)(1-p_2)}$ and the dilution is $D_{12} = 1 - 2p_{12} = \frac{(1-p_1)(1-p_2)-p_1p_2}{(1-p_1)(1-p_2)+p_1p_2}$. This expression can be rewritten as

$$D_{12} = \frac{1 - 2p_1 + 1 - 2p_2}{1 + (1 - 2p_1)(1 - 2p_2)} = \frac{D_1 + D_2}{1 + D_1D_2}. \tag{1}$$

Similarly, for the fourth group,

$$\bar{D}_{12} = \frac{|D_1 - D_2|}{1 - D_1D_2}. \tag{2}$$

This derivation for the case of more than two taggers can be found in Ref. [4].

If the event is tagged with both taggers, then the probability to obtain the identical results ("Same Sign") from them is

$$p_{SS} = N_{12}/(N_{12} + \bar{N_{12}}).$$

Similarly, for opposite tagging results ("Opposite Sign"):

$$p_{OS} = \bar{N_{12}}/(N_{12} + \bar{N}_{12}).$$

These probabilities can be easily expressed in terms of the probabilities p_1 and p_2:

$$p_{SS} = p_1 p_2 + (1 - p_1)(1 - p_2), p_{OS} = p_1(1 - p_2) + (1 - p_1)p_2.$$

From these formulae it follows that

$$p_{SS} - p_{OS} = \frac{N_{12} - \bar{N}_{12}}{N_{12} + \bar{N}_{12}} = (1 - 2p_1)(1 - 2p_2) = D_1 \cdot D_2$$

i.e. the dilution of one of the taggers, D_1, can be measured in the data if the second dilution, D_2, is known:

$$D_1 = \frac{1}{D_2} \cdot \frac{N_{12} - \bar{N}_{12}}{N_{12} + \bar{N}_{12}}. \tag{3}$$

Then the other two dilutions, D_{12} and $\bar{D}_{12}$, can be calculated according to Eqns. 1 and 2. Of course, these formulae only apply in case of uncorrelated taggers. The "ϵD^2" can be obtained as follows:

$$\epsilon D^2 = \frac{N_1}{N} \cdot D_1^2 + \frac{N_2}{N} \cdot D_2^2 + \frac{N_{12}}{N} \cdot D_{12}^2 + \frac{\bar{N}_{12}}{N} \cdot \bar{D}_{12}^2 \tag{4}$$

where $N = N_1 + N_2 + N_{12} + \bar{N}_{12} + N_{NT}$ is the total number of events.