

## Leo's questions on $\pi^0 \rightarrow \mu e$

L1. Leo asked several questions that do not relate directly to the analysis, but are rather something that should be included in the paper. These include the question about what limits can be inferred from  $eN \rightarrow \mu X$ , what models suggest looking for this decay (he was referring to  $\pi^0 \rightarrow \mu e$ , but that question applies to all). These have to be answered in the paper, but not here.

L2. The first part of the  $\pi^0 \rightarrow \mu e$  document I want to treat as an archive of what Collin Wolfe did, but is not the way the analysis was done in the end. I am going to change the writeup to make this clear, and let the reader bypass that part completely unless he wants to know about Collin's work. I am also going to make the cuts on all three analyses as similar as possible, which means among other things the cuts on magnet offsets would be the same for electrons and muons.

L3. Leo asked about the opening angle for the charged tracks in  $\pi^0 \rightarrow \mu e$ . I don't have that distribution, nor do my ntuples have the information to make it. Is this distribution important? I do have the invariant mass of the two charged tracks for both signal and normalization (which is  $K \rightarrow 3\pi^0$  with one  $\pi^0 \rightarrow ee\gamma$ ). I think the opening angle is smaller for  $\pi^0 \rightarrow ee\gamma$  because the electrons come from  $\gamma \rightarrow ee$ . I can look at this with a small amount of work.

L4. Leo asked about the difference between the magnet offset cuts for electrons and muons, and more generally about the justification for the selection cuts. "For example how much acceptance was lost by the 1.25cm cut for neutral clusters from the beam hole?"

The original reason for making the cut a little harder for muons than electrons was to try and discriminate a bit more against pion decays to muons. But this is not an important cut and making the magnet offset cut the same for muons and electrons I don't think will matter.

As for justifications of the selection cuts... all these cuts, except for the BA cut, are pretty standard. Do you mean justification for the actual numerical values chosen, or general reason for cutting on each variable? If you want a justification for the exact numerical values chosen, I don't have a strong justification for the specific values ...most cuts could be changed slightly with little consequence.

As for the loss of acceptance due to the 1.25 cm cut around the beam hole—it was a lot, 15%, which is why I removed it. Collin made that cut for the muons, too. Such a cut would have removed the one event I found in the signal box in the 99 data, though! But instead of the beam-hole cut, I use a loose TRD anti-electron cut to be sure the track labeled as a "muon" does not look like an electron to the TRDs. That cut does not impact the acceptance much (98% of all real muons pass), but it does a good job of removing tracks that fake a muon because they are on the edge of the calorimeter.

L5. "The cluster energy cut for the muon was at 1 GeV at the crunch stage?" Yes, the cluster energy cut for muons was set at 1 GeV early on and never changed. The average energy deposited by a muon is more like 300 MeV, but this cut gives good acceptance and still discriminates against pions and electrons.

L6. Why a one-sided cut on  $M(\mu e)$ ?

The one-sided cut is just a first cut which eliminates most of the background. A harder cut, consistent with  $M(\pi^0)$  is made later. The purpose of Figure 1 is just to show what the  $M(\mu e)$  mass distribution looks like, and show

that most events have a mass way above the mass region of  $\pi^0 \rightarrow \mu e$ .

L7. “Why does one not have a cut comparable to the  $M_{ee}$  mass cut in the signal mode?”

I don't think such a cut is needed. In the normalization mode (3pi0D), the  $M_{ee}$  distribution peaks at zero, and the  $M_{ee}$  cut ensures that the tracks are far enough apart to be able to reconstruct them.. In the signal mode, the mu-e system should reconstruct to a pi0 mass, and a cut of  $\pm 3$  Mev around the pi0 mass is made. The signal Monte Carlo shows good resolution in  $M(\mu e)$ , which would not be the case if the tracks were very close together.

L8. “The E/p shift of 0.4% (between data and MC) in the normalization mode is more than I remember from pi0ee. Does it show a variation with electron energy?”

I was also worried about this shift. I asked Pat Toale about it, he also saw this shift If you look in his thesis (pg 88 on the copy from the KTeV web site), you do see a shift. The size of this shift varies with ktevana version, and I think it might be worse for 6.00 than for earlier versions! But unless there's a more up-to-date MC version for 799, I don't know what to do about it. For the  $\pi^0 \rightarrow \mu e$  and  $K \rightarrow \pi^0 \pi^0 \mu e$  I cut fairly loosely in E/p, and I did study the cut variation (I don't think it mattered much). I will check the energy dependence.

L9. “BA1 cut—Collin put a cut in and you removed it, therefore it doesn't belong in the systematic uncertainty list”. No, Collin never made a BA cut, but I did. There was more background for 99 vs 97 and I explored various ways to get rid of it. The BA cut was very effective.

L10. “The section on background seems short”. For the 97 data, we concluded that there was negligible background, so we didn't dwell on it too much. In 99 we discovered there was quite a bit more background, mainly from 3pi0D, which requires an accidental muon. For 99 I believe the ke3, kmu3 and k3pi still don't contribute to the final background. If you look in the writeup about the 99 data there is a discussion of the 3pi0D background, which fits that data quantitatively.

L11. “I don't understand why a cut on charged vs. neutral vertex works”.

There are three vertices in this decay...one charged, two neutral. They should all agree. I calculate an average (weighted) vertex, and use that as the final vertex. The average vertex is dominated by the neutral vertices, which have better resolution.

The final pi0 masses are calculated using this average vertex, and then a cut is made on the pi0 mass. This is also similar but not entirely equivalent to the cut on the vertex difference. In fact a plot of pi0 mass vs. delta-vertex is a diagonal swath. A cut on both pi0 mass and delta-vertex selects the central region of this distribution. A more optimal cut might be something other than a rectangle in these two variables, but I think this cut is OK.

I do make a cut on the charged vertex chisq, but I don't make a chisq cut on the average vertex, which maybe is the way others do it. So the cut on the difference in z position of the (charged-neutral) vertices may be instead of a cut on the chisq of the average vertex.

“Was Fig 22 made before or after a cut on the charged particle mass?” By “charged particle mass” I think you mean  $M_{\mu e}$ , in which case the answer is after. But this plot is for signal MC, and it doesn't matter much if this plot is made before or after a cut on  $M_{\mu e}$ .

“The procedure to determine this cut is dangerous because of the low statistics in the background sample”.

I used signal MC to choose a reasonable place for this cut.

L12. “Re-evaluating the background estimate is intrinsically perilous. You can bias your result by moving the background estimate up almost as quickly as you can by wiggling a cut to push the number of box events down”.

It is certainly true that changing the background estimate can change the limit. But if you miss a background, what do you propose to other than correct your background estimate to include it?

L13. “Bottom of pg 20, using  $M_{\mu e}$  sideband method and MC that is 2x the flux, the background is estimated at  $1.0 \pm .35$  in the study region and  $0.3 \pm .17$  in the signal. box. But the ratio should be much larger, no? The signal box is 50x smaller than the study box”.

The events are not uniformly distributed in the study region, look at fig 26. This ratio in fact is the same for 99 and 97. 99 has more overall expected  $3\pi^0 D$  background but the ratio signal box/study region is the same.

L14. “What about ke4 backgrounds” Haven’t looked at ke4 for  $\pi^0 \rightarrow \mu e$  or  $K \rightarrow \pi^0 \pi^0 \mu e$ . I have a good sample of ke4 MC from trigger 7, so I could . I don’t think it will contribute, though, since it has a small BR and needs an accidental  $\pi^0$  plus a pion decay to fake the signal.

L15. ”What about systematic uncertainties, including cross-trigger normalization issues.

See response to Brad’s question along the same lines.

L16. “The likelihood function—what is the form of the double exponential, is it just the sum of two exponentials? In that case the parameters will have a large correlation.”

Yes it’s the sum of two exponentials. In fact for higher statistics samples the sum of three exponentials works better. Yes the parameters are highly correlated, but as long as the fit is good it doesn’t matter.

L17. “It isn’t obvious to me that the joint PDF comes out normalized to 1”

Of course it does.

L18. “Where is ‘the event’ in fig 29 (ie, in the likelihood variable).”

I can’t believe I didn’t show this! It is at 7.7, which is inside my standard blind region ( $pdf > 5$ ) but not in my standard signal region ( $pdf > 10$ ). I will add that to the writeup and also a better plot of the likelihood variable, or an additional one that is more spread out and that shows where the event is.

L19. “How/where do we show that the two variables are uncorrelated?”

I show it by calculating the covariance and showing that it is consistent with zero. It is shown in the update the Angela’s analysis ( $K \rightarrow \pi^0 \mu e$ , fig 2), but I should also show it in this writeup.

L20. “It looks like this analysis was done before the development of run-dependent TRD cuts. In that case it’s hard to say much about a track with probability at 7.4%”

The point of the TRD information for ‘the event’ in  $\pi^0 \rightarrow \mu e$  is that the track labeled as a muon by the calorimeter looks VERY much like an electron in the TRD. That’s the one with the TRD probability (remember that’s the CL that it’s a pion) of .000173, way inside the electron peak. The other track, which the calorimeter thinks is an

electron, has a TRD probability of .07, which in fact is not as favorable, but still consistent with an electron.

For the  $K \rightarrow \pi^0\pi^0\mu e$  analysis I have made a loose anti-electron TRD cut on the muon track, but no cut on the electron track. I think this cut is loose enough that I don't have to worry about the variation over the run, but I will look at it. Of more concern is Angela's analysis, where we did not make time-dependent TRD cuts. Her cut is also fairly loose, so it may be OK, but I will have to check.

### Leo's questions on $K \rightarrow \pi^0\pi^0\mu e$

L21. "What models suggest looking for this decay?"

Again, this is a question for the paper, but Robert Schrock (BNL) says (I exchanged some email with him) that there are mechanisms that allow this but don't allow  $K \rightarrow \pi^0\mu e$ .

L22. "What about ke4 backgrounds?"

Haven't looked, but could. I don't expect it to be large due to the small BR and the need for an accidental pi0.

L23. "Not withstanding the known faults in the MC, the discrepancies in figs 32-35 open the possibility that the background is something we haven't thought of. In addition to ke4 background, maybe something is happening to fake pi0s from satellite clusters."

I'm open to suggestions. The data/MC agreement gets much better with a very loose cut on the difference between the charged and neutral vertices, look at figs 37-40. Maybe there are double decays that are not accounted for by the accidental overlays?

In any case, I don't think we can ever count on MC to give a quantitative estimate of the background due to the poor modeling of the HA and the pion showering, if for no other reasons. For this reason I didn't worry too much about the data/MC mismatch.

L24. "In Fig 38, to my eye it is not obvious that the data distribution is symmetric about the pi0 mass".

I think it is. Data and MC are shifted, though.

"Maybe it is due to visual confusion with the histogram for the MC sample, which looks like it has peaks due to simulation with a small accidental sample."

I used all the accidental events that are on the disks. The MC statistics are not great, but I don't think there is any evidence that a particular accidental overlay is being used over and over.

L25. "Would it clean up the presentation of both analyses to use the same cut on extra in-time hits? Or at least use the same definition of the variable? What about the other cuts on page 32?"

Is it fair to go back and redo the  $\pi^0 \rightarrow \mu e$  analysis after the box is opened? That is the only reason to keep the old definition of in-time DC hits. I am in favor of having a combined analysis for  $K \rightarrow \pi^0\pi^0\mu e$  and  $\pi^0 \rightarrow \mu e$  in which the  $\pi^0 \rightarrow \mu e$  result is just one additional cut (on the  $M_{\mu e}$ ). People have to agree, and then I guess we cannot call the  $\pi^0 \rightarrow \mu e$  result blind.

L26. "pp0kine is the square of the longitudinal component of the missing momentum, in the frame where the

charged tracks have zero net longitudinal momentum, not the K rest frame.”

Really!? Are you sure?

L27 “Is the extrapolation of the pdf for data/background done under the assumption of a constant pdf or a linearly varying pdf?”

I assumed the pdf distribution was flat (zero slope) in the region between -10 and 5, and that it remained flat into the signal region. I do not have any more justification that the bottom plot of fig 51, which does look flat to me. But does it stay flat all the way into the signal region? I don’t know, nor do I know of a good way to check that it does. I am open to suggestions.

**It is exactly this issue for which I wanted some input from the godparent committee. Is this a reasonable way to estimate the background? Do you have any other suggestions, or any other way to estimate the background as a cross-check?**

Leo’s questions on  $K \rightarrow \pi^0 \mu e$

L28. “Angela did/did not vary the TRD cut with run period?” The cuts were quite different for 99 vs 97, but there was only one cut for 99 (also for 97). We did not realize at the time the variation of the TRD for 99. Her cut was loose, so it may be OK, but I will have to check it.

L29. “Are there plots of the 3x3 fusion chisq?”

Yes, look in Angela’s thesis (available from the KTeV web page), figs. 3.27-3.32. She cut at 4 which is hard, but also rescaled the MC to agree better with the data. In  $K \rightarrow \pi^0 \pi^0 \mu e$  I cut at 10. Should I move this cut to 10 for  $K \rightarrow \pi^0 \mu e$  as well? (Which brings up the question of a blind analysis or not again).

L30. “How does 500M k3pi events compare to the size of the data sample?” Only about 1%.

“The CsI rejection factor seems too high” I got this number from ke3s from trigger two and I think it is right. See the response to Taku’s question below, he asked the same thing. “The TRD rejection factor might or might not be too low” Also got it from ke3s. But when I looked at it again recently I get a lower value (5.5 not 10). Angela’s cut was loose, so I don’t expect a large factor. I will have to go back and try to figure out how I got a factor of 10 before. But the exact value of pi/e rejection is not important.

L31. “Is Rick Kessler going to be happy with this treatment of k3pi background?” He better be, it was his idea.

“What number are we assigning to the background from this source?” zero.

L32. “Fig 7 is without a signal box or likelihood cut?” There is a cut on the likelihood variable to keep out of the blind region,  $pdf < 5$ . No (old) signal box cut.

L33. “We should compare pt2 and Mgg for 99 and 97.” Plots should be in Angela’s thesis, but I could easily include them in this short writeup.

L34. “Do we know that relaxing the  $M_{\gamma\gamma}$  cut will not change the pdf?” Hmmm. The pdf is determined from the signal MC distributions in MK and pt2. In the signal MC, there are few events outside of the pion mass region in  $M_{\gamma\gamma}$ . Relaxing this cut has little effect on the pdf distribution since not many events are added.

L35.”A table summarizing the different background estimates of both signal box and likelihood method for the two data periods would help.

I can do that.

### **Taku’s questions on $K \rightarrow \pi^0 \mu e$**

T1. Why is the pi/e rejection factor so high for the CsI? The MC pion shape is not reliable close to 1.

The pi/e rejection factor was determined solely from ke3 events from trigger 2. I looked at this again and still think that value is right given the tight cuts Angela made (although the exact number is not important for the argument I was making). If I take 2-track events with a few standard clean-up cuts (vertex chisquared, track matching in the magnet), then require one track to have an E/p between .95 and 1.05, the other track should be a pion, since these should be mostly ke3 events. Using just these cuts I get a pi/e rejection of about 290. But if I make the additional requirement of a loose TRD cut on the electron, to be sure it is an electron, I get a pi/e rejection of 470. Adding in a fusion  $\chi^2$  cut on the pion track increases the rejection factor to 550. Making harder E/p and fusion  $\chi^2$  cuts on the pion increases the rejection to more than 800.

The figure below shows why the rejection factor for the pion changes when I make a TRD cut on the electron. The black curve is the E/p distribution for the track labeled as a pion when there is an E/p cut of 0.95 to 1.05 on the other track. There is clearly a little peak at E/p=1, indicating that some of these tracks are really electrons. The red curve is the same, but also making a loose TRD cut on the electron (probe<0.05). The green curve has no cut on the TRD information for the electron track, but instead it has a loose anti-electron TRD cut (probpi >0.1).

So I think these numbers are fair, but as I have already noted the exact values of the rejection factors is not important.

T2. “For Ke3 background events, is there a way to estimate the background level using MC?”

Angela did a 1x flux ke3 Monte Carlo by doing forced decay and punch through. but that was before the pi-mu decay bug was found. The ke3 MC does have a flat distribution for both  $M_{\gamma\gamma}$  and  $p_t^2$  distribution as was seen in the data. See the two figures below, which are from Angela’s thesis.

T3. It would be nice to confirm that the MC Ke4 background estimate is correct.

If you have a suggestion as to how to do this, I am willing to look into it. I don’t have any ideas.

### **Taku’s questions on $\pi^0 \rightarrow \mu e$**

T4. “Page 5, fig 1, What is the source of the huge bump in the invariant mass at 0.3 GeV?” These are Ke3s.

T5. “P8, Fig 6 What kind of radiative correction treatment has been done in the MC? I used v6\_00 of ktevmc, so whatever is in v6\_00 is what was used.

T6.”figs 11-19, what kind of cuts are being applied to these samples? It seems like not many cuts are being applied, and if that is the case I do not see the validity of data/MC comparisons since the data may have other contributions.”

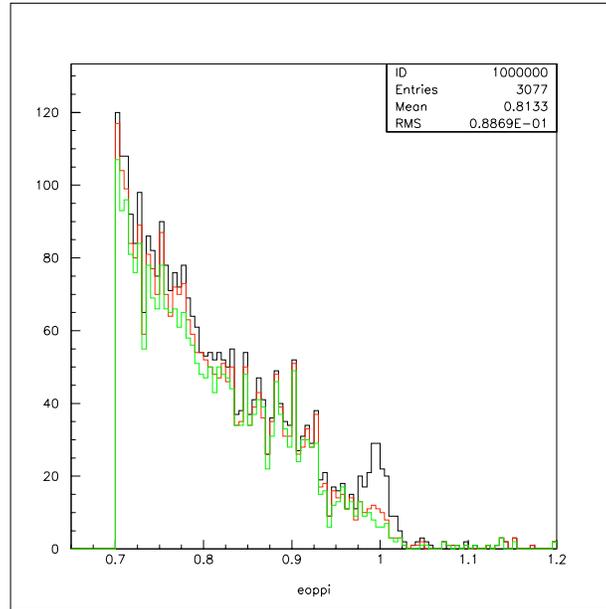


Figure 1: E/p distribution for the “pion” track in two track events. The opposite track has E/p between 0.95 and 1.05 (black). The red curve also has a TRD cut  $probe < 0.05$  on the electron track, ensuring that the track labeled as an electron really is an electron. The green curve has no cut on the electron TRD information, but rather has a cut of  $probpi > 0.1$  on the track labeled as a pion.

The cuts applied in these figure are loose, otherwise there would be no MC events left. Two tracks and two pi0s are required, and there is a loose cut requiring the two pi0s to have vertices not too far apart. But we did look at all the major decay modes and these are the ones that contributed at this level. These plots show Ke3/kmu3 contributions and k3pi0 contributions separately, then combined. The combined plots in fact agree pretty well (figs. 13, 16 and 19).

But in the end, with more cuts, all these events disappear and the only background remaining is 3pi0D. In 99 this background was quite a bit worse due to additional accidental muons in 99.

T7. “What is the unpacking muon bank bug?” A long time ago there was a bug in ktevana so that not all the muon counters were unpacked correctly. Because of this Collin required only 2/3 muon banks to match to the muon track. This bug has long been fixed and is not longer an issue. Data has been reanalyzed.

T8. “Making a cut on charged and neutral verticies differences is similar to a pi0 mass cut and vertex chi-square cut. Why should we introduce a new cut rather than tightening the old ones?”

See the answer to L11, which was the same questions.

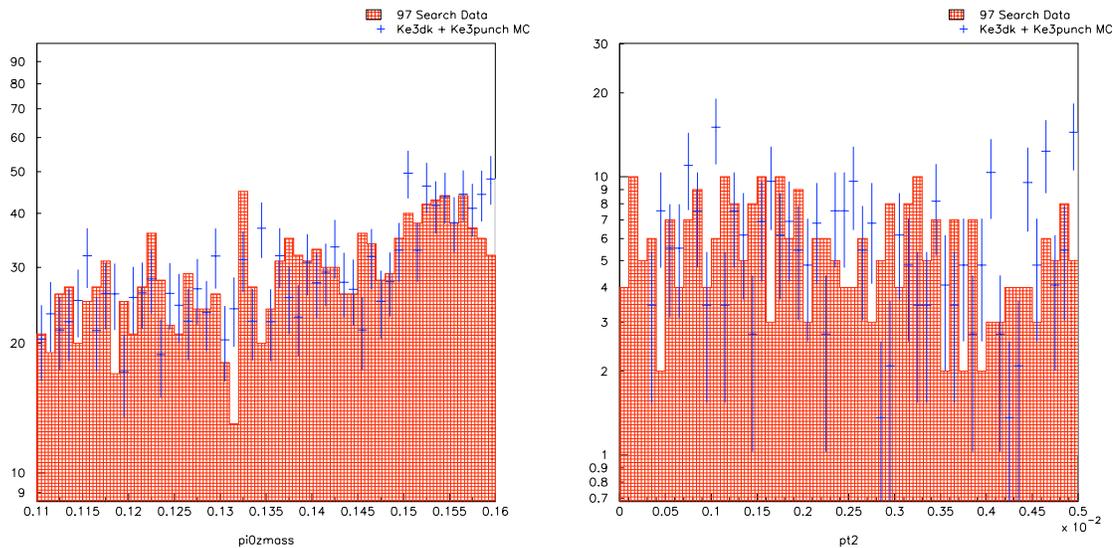


Figure 2:  $M_{\gamma}$  and  $p_t^2$  distributions for data and ke3 Monte Carlo.

T9. "I am surprised that accidental muon would add more backgrounds, since they are mostly suppressed by requiring a good vertex match."

The dominant background is  $3\pi^0D$  which absolutely requires an accidental muon. It is certainly true that most accidental muons don't match to anything, but all we need is one in  $10^8$  or so to generate the background. Look at figure 25 and 26 for example. The  $3\pi^0D$  Monte Carlo reproduces the search data both in shape and absolute normalization.

### Brad's questions

B1. "Have you included any systematics in your limits?" Until I saw how David Smith included systematics I had no idea how to do so. What he did is reasonable...as you said, the bad thing is that it throws away all the benefit from having a normalization mode cancellation of systematics, but I don't have any better ideas. I can do a similar analysis.

B2. "Have you tried Sasha L's combined vertex routine?" No, I don't have it. I did look at a routine that Pat Toale had, though, maybe it is similar. That routine did some momentum rescaling, which I got worried about. If I know I have signal it is OK to rescale momenta to improve resolution. But if all I have is background, I don't want to rescale the momentum to put background into the signal region. But I'll ask Sasha about it.

B3. "There are considerable mismatches/shifts in K,  $\pi^0$  E/p plots of data relative to MC for the normalization mode."

Yes. I asked Pat Toale about this once, he also sees these shifts, which are evident if you look at his thesis. This is a feature of Ktevana v6\_00, which he and I both used. I don't think the Kaon mass is shifted so much, but  $\pi^0$

and E/p are shifted in MC vs data.

For my analysis, I cut loosely on E/p and pi0 mass, so it shouldn't matter too much. These cuts are varied along with others in determining the variation of the apparent flux. If I treat systematics in some way similar to how David Smith did, these effects would be folded into the systematics.

For Angela's analysis, she cut a fixed number of sigma from the mean for the pi0 mass cut, exactly to minimize the effect of this shift (that was Leo's suggestion).

B4. "In general there have been problems determining muon and electron efficiency uncertainties. The electron efficiency due to tails and shifts of E/p, and muons due to the attenuation in the absorber at 8 GeV."

I require muon momenta above 8 GeV, for exactly that reason. I could do a study cutting harder and see if there is a difference. For electrons, the E/p cut is loose, especially for  $K \rightarrow \pi^0 \pi^0 \mu e$  and  $\pi^0 \rightarrow \mu e$ . The E/p for  $K \rightarrow \pi^0 \mu e$  is tighter, I could vary that cut and see how the apparent flux changes.

B5. "There is a philosophical issue for keeping or rejecting the one event in the signal region for  $\pi^0 \rightarrow \mu e$ "

Brad argues that when we use the TRD (and this event is clearly ruled out), we are bringing a new detector into play and therefore we should remove that event.

I agree, it is what I would like to do, if everyone else is in agreement.

"Is the issue that a TRD cut would cost more acceptance than it is worth?"

No, the TRD cut I would make is very loose and costs almost no acceptance.

B6. "Why is the 97 data cleaner than the 99 data? Is it totally due to the higher rate of muon accidentals due to the higher rates?"

I think it is just due to the higher rates. The higher background in 99 is seen in several modes, it is not unique to this analysis. If you have an idea of how to check this I could have a look.

B7. "Are there any systematics that might arise due to having the signal mode in trigger 7 and the normalization in trigger 1? Is there any problem with using 3pi0D events as a normalization mode as compared to  $K \rightarrow \pi^0 \mu e$ , which would have a much larger opening angle than the Dalitz pairs".

For  $K \rightarrow \pi^0 \mu e$ , the normalization is K3pi from trigger 2, not 3pi0D. However, I still rely on the MC to give the correct efficiencies for the muon banks and the HCC requirements in the trigger. This is worth checking, which I guess I could do with Ke3s and Kmu3s from trigger 2.

For  $K \rightarrow \pi^0 \pi^0 \mu e$  and  $\pi^0 \rightarrow \mu e$ , the normalization mode is 3pi0D from trigger 1. The muon bank efficiency is an issue for these modes as well.

We should talk about the best way to answer this question.

B8. "Your estimates of background for  $\pi^0 \rightarrow \mu e$  in the 99 data are much larger than seen in the study region of the data. Do you have an explanation? The estimates of the backgrounds for the 97 data seem much closer to the data."

I don't have an explanation other than to say it is a fluctuation, which I really think is the case. If you compare my background estimates over a larger  $M(\mu e)$  region, the agreement is good. I think there was a downward fluctuation in the study region that fooled me into thinking the background was lower than it really was.

B9. "You imply that you cannot use the pdf technique to the  $\pi^0 \rightarrow \mu e$  mode since you have opened the box. I'm not sure why you would take this position. I think the pdf technique is somewhat different than changing cuts. This should be discussed."

What I really would like to do is to make a combined analysis for  $K \rightarrow \pi^0 \pi^0 \mu e$  and  $\pi^0 \rightarrow \mu e$ , in which I use the same cuts and use the pdf method for both. Then  $\pi^0 \rightarrow \mu e$  is just one additional cut on  $M(\mu e)$ . If everyone agrees, that is what I would like to do.

B10. "I presume the pdf region for  $\pi^0 \rightarrow \mu e$  and  $K \rightarrow \pi^0 \pi^0 \mu e$  have not yet been opened. Are you asking for an opinion from the godparents about whether we think you are ready to open them?"

Yes.

B11. "I have been dubious about using the BA since it is not well simulated in the MC. Since the ratio of data to MC for the two track modes seemed to be the same for  $ke3$ ,  $k\mu3$ , and  $3\pi^0 D$ , the hope is that any uncertainties will cancel out in the BR calculation. It would be nice to see a plot of the ratios for the three modes for each run for the 97 and 99 data to see how stable this ratio is vs time. There must be a change from 97 to 99 due to the increased rate. Also a plot of the ratio for  $ke3$  and  $k\mu3$  divided by the ratio of  $K3\pi^0 D$  vs run number would be interesting".

I am a bit confused by the last sentence, but I think you are asking for  $Ke3/k\mu3$  vs run number, and  $ke3/K3\pi^0 D$  vs run number?

This is a fair question. I looked at several runs sprinkled throughout the data-taking period. But I have not looked systematically vs time. I can do the  $Ke3/3\pi^0 D$  reasonably easily, since I have the data. For  $k\mu3$  I would have to analyze the trigger 2 data again, but it is not too much work.

Between 97 and 99 the configuration of the BA changed, so this ratio will be different not just because the rate is different, but also because the detector was different.