Subject: [FIXED] Tracking + Truth Match Problem
Posted by Klaus Götzen on Thu, 22 Nov 2012 08:24:01 GMT
View Forum Message <> Reply to Message

Dear all,

I'd like to report about a problem I found in simulated event concerning tracking and MC truth match.

The first issue is, that I find from time to time track objects, which appear multiple times in the event. I noticed it on the pid-level, but already after the reco level, these objects are present. This might look like this (analysis level; displayed quantities are: track number in event, 4-vector, charge, PID probabilities, MC truth match index)

trk:0 (-1.06059,0.495762,2.45283;2.72148) -1 PID:0,0,1,0,0 MC:8
trk:1 (0.142945,-0.639217,0.900123;1.12193) 1 PID:0,0,1,0,0 MC:4
trk:2 (0.193793,0.341487,1.30556;1.37045) 1 PID:0,0,0,1,0 MC:6
trk:3 (-0.26782,0.799728,2.97571;3.09607) -1 PID:0,0,1,0,0 MC:7
trk:4 (1.00258,-0.368745,5.14993;5.26141) -1 PID:0,0,0,1,0 MC:3
trk:5 (1.00118,-0.368159,5.14244;5.25377) -1 PID:0,0,0,1,0 MC:3

or this

trk:0 (-0.550406,0.169454,0.525641;0.792111) -1 PID:0,0,1,0,0 MC:8
trk:1 (-0.0351313,0.275942,-0.119804;0.333483) -1 PID:0,0,1,0,0 MC:7
trk:2 (0.0599226,0.085189,0.0741163;0.189264) -1 PID:0,0,1,0,0 MC:7
trk:3 (0.0599226,0.085189,0.0741163;0.189264) -1 PID:0,0,1,0,0 MC:7
trk:4 (0.0599226,0.085189,0.0741163;0.189264) -1 PID:0,0,1,0,0 MC:7
trk:5 (0.0599226,0.085189,0.0741163;0.189264) -1 PID:0,0,1,0,0 MC:7
trk:6 (5.92574,0.770891,34.7574;35.2676) 1 PID:0,0,1,0,0 MC:5
trk:7 (-0.078541,-0.0528536,3.60024;3.60419) -1 PID:0,0,0,1,0 MC:3
trk:8 (-0.250432,-0.622125,7.483;7.51429) 1 PID:0,0,1,0,0 MC:4
trk:9 (0.251804,0.0933929,1.55109;1.58035) 1 PID:0,0,0,1,0 MC:6

What one can see is, that in the frist event tracks 4 and 5 are very similar, having also the same MC index 3. This might be due to tracks broken into tracklets, which are treated as different tracks.

On the other hand, in the second event, there are 4 identical track objects (tracks 2-5), which seem to be true clones.

The second issue is, that there are mutiple reco tracks having the same MC truth index. This might be due to the upper issue, where the clones or very similar tracks have the same index. But I also observe events like the following:

trk:0 (0.607878,0.248934,1.06207;1.25657) -1 PID:0,0,1,0,0 MC:8
trk:1 (0.196574,-0.357057,0.765892;0.87875) 1 PID:0,0,1,0,0 MC:4
trk:2 (0.594232,0.99658,2.77834;3.01412) 1 PID:0,0,0,1,0 MC:6
trk:3 (-0.0748785,-0.0292746,0.268576;0.313172) 1 PID:0,0,0,1,0 MC:3
trk:4 (-0.914007,-0.203599,3.14832;3.28759) -1 PID:0,0,0,1,0 MC:3
trk:5 (-0.338518,-1.15738,3.99636;4.17666) 1 PID:0,0,1,0,0 MC:5

Here tracks 3 and 4 have quite different 4-vectors, nevertheless the MC index is the same (3) in both cases.

I tried to investigate the effect of the multiple tracks issue by removing tracks with 4-vectors deviating by less the 10E-5 in each component by hand.

The results of an analysis of 1000 D+ -> K- pi+ pi+ (+ c.c.) events are shown in the two attached plots. The blue histogram shows all combinations, the red one is the full MC truth matched part.

FIG 1: Combinatorics with original track list. A very spiky histogram is the result.

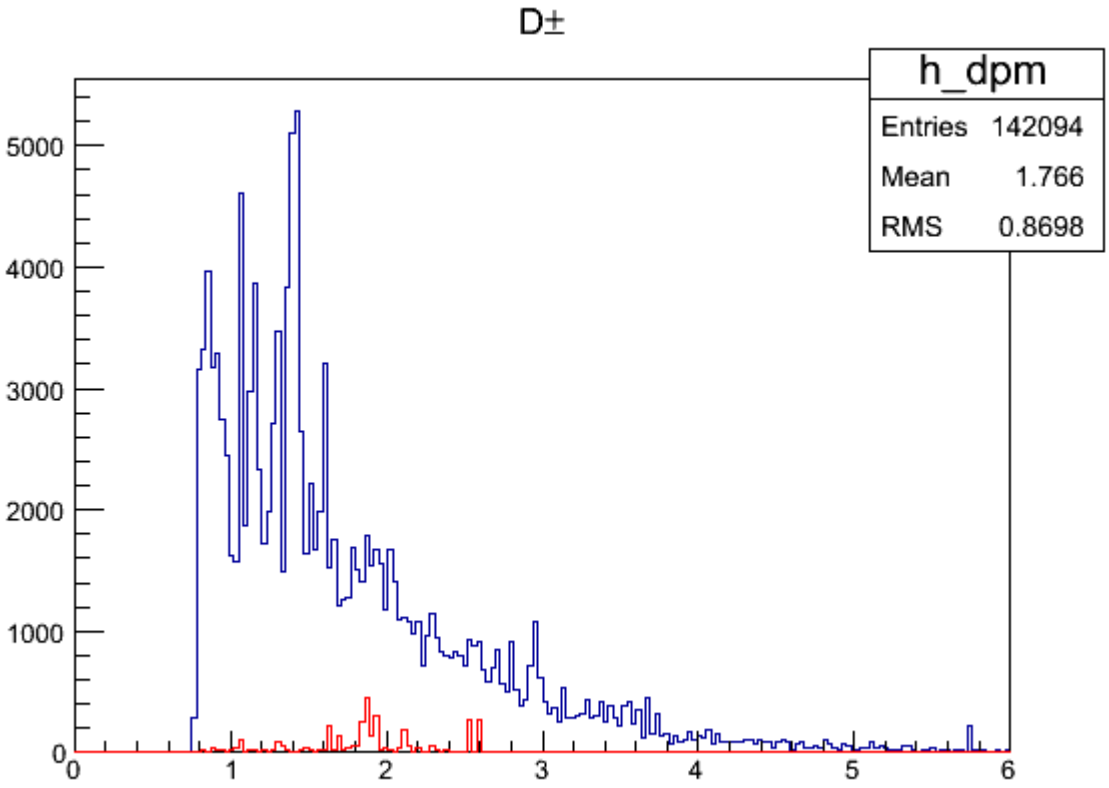
FIG 2: Combinatorics of the same 1000 events with removed double tracks. The shape looks reasonable now. Also note, that the number of combinations is only roughly 25% of those above.

I'd like to ask the (tracking?) experts to take a look to that issue, having a very significant impact on analysis results.

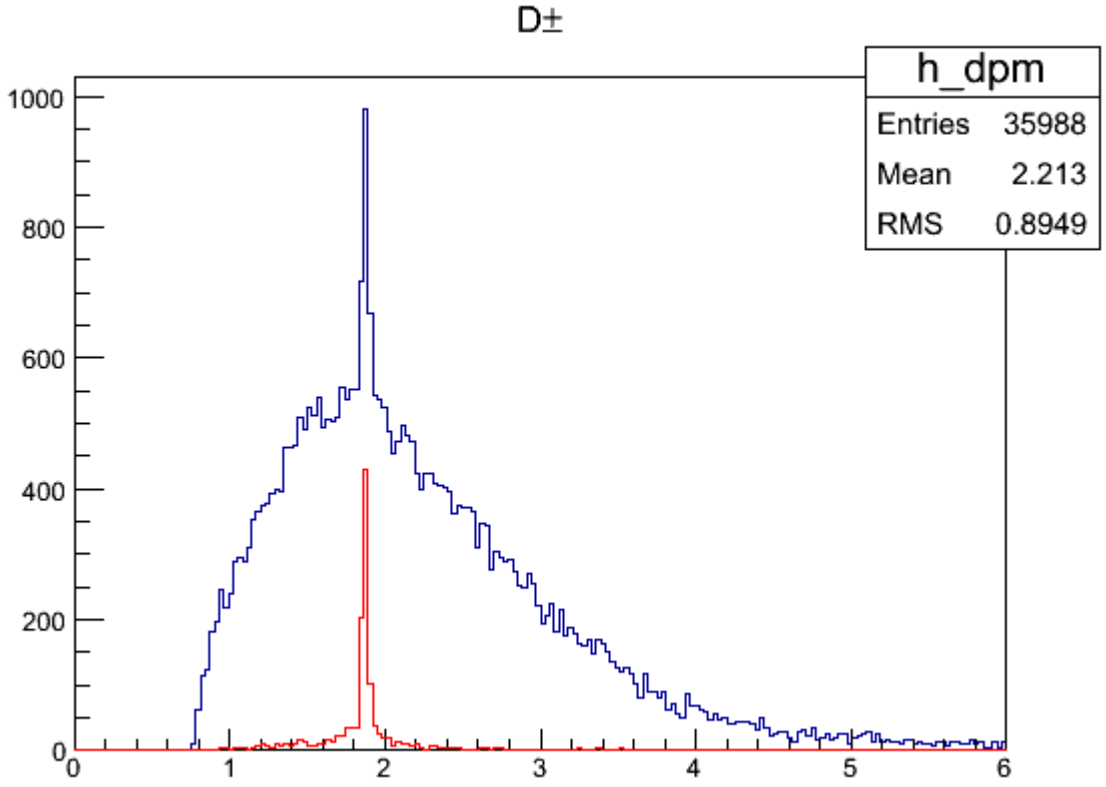
Cheers,
Klaus

## File Attachments

1) dpm55_ftm_std_1k1.gif, downloaded 1531 times




2) dpm55_ftm_rem_1k1.gif, downloaded 1617 times




---

Subject: Re: Tracking + Truth Match Problem

Posted by Tobias Stockmanns on Thu, 22 Nov 2012 08:30:50 GMT
View Forum Message <> Reply to Message

Dear Klaus,

you could use the event display to see which hit points are part of the track. With this method you can see if there are just different combinations of points from one track assigned to different tracklets.
With a track merger one should be able to reduce this effect.

Cheers,

Tobias

---

Subject: Re: Tracking + Truth Match Problem
Posted by Gianluigi Boca on Thu, 22 Nov 2012 11:50:36 GMT
View Forum Message <> Reply to Message

Dear Klaus,
I think I can fix this problem, thank you for mentioning.
I am investigating now
cheers und auf Wiedersehen
Gianluigi

Klaus Goetzen wrote on Thu, 22 November 2012 09:24Dear all,

I'd like to report about a problem I found in simulated event concerning tracking and MC truth match.

The first issue is, that I find from time to time track objects, which appear multiple times in the event. I noticed it on the pid-level, but already after the reco level, these objects are present. This might look like this (analysis level; displayed quantities are: track number in event, 4-vector, charge, PID probabilities, MC truth match index)

trk:0 (-1.06059,0.495762,2.45283;2.72148) -1 PID:0,0,1,0,0 MC:8
trk:1 (0.142945,-0.639217,0.900123;1.12193) 1 PID:0,0,1,0,0 MC:4
trk:2 (0.193793,0.341487,1.30556;1.37045) 1 PID:0,0,0,1,0 MC:6
trk:3 (-0.26782,0.799728,2.97571;3.09607) -1 PID:0,0,1,0,0 MC:7
trk:4 (1.00258,-0.368745,5.14993;5.26141) -1 PID:0,0,0,1,0 MC:3
trk:5 (1.00118,-0.368159,5.14244;5.25377) -1 PID:0,0,0,1,0 MC:3

or this

trk:0 (-0.550406,0.169454,0.525641;0.792111) -1 PID:0,0,1,0,0 MC:8
trk:1 (-0.0351313,0.275942,-0.119804;0.333483) -1 PID:0,0,1,0,0 MC:7
trk:2 (0.0599226,0.085189,0.0741163;0.189264) -1 PID:0,0,1,0,0 MC:7
trk:3 (0.0599226,0.085189,0.0741163;0.189264) -1 PID:0,0,1,0,0 MC:7
trk:4 (0.0599226,0.085189,0.0741163;0.189264) -1 PID:0,0,1,0,0 MC:7
trk:5 (0.0599226,0.085189,0.0741163;0.189264) -1 PID:0,0,1,0,0 MC:7
trk:6 (5.92574,0.770891,34.7574;35.2676) 1 PID:0,0,1,0,0 MC:5

trk:7 (-0.078541,-0.0528536,3.60024;3.60419) -1 PID:0,0,0,1,0 MC:3
trk:8 (-0.250432,-0.622125,7.483;7.51429) 1 PID:0,0,1,0,0 MC:4
trk:9 (0.251804,0.0933929,1.55109;1.58035) 1 PID:0,0,0,1,0 MC:6

What one can see is, that in the frist event tracks 4 and 5 are very similar, having also the same MC index 3. This might be due to tracks broken into tracklets, which are treated as different tracks.

On the other hand, in the second event, there are 4 identical track objects (tracks 2-5), which seem to be true clones.

The second issue is, that there are mutiple reco tracks having the same MC truth index. This might be due to the upper issue, where the clones or very similar tracks have the same index. But I also observe events like the following:

trk:0 (0.607878,0.248934,1.06207;1.25657) -1 PID:0,0,1,0,0 MC:8
trk:1 (0.196574,-0.357057,0.765892;0.87875) 1 PID:0,0,1,0,0 MC:4
trk:2 (0.594232,0.99658,2.77834;3.01412) 1 PID:0,0,0,1,0 MC:6
trk:3 (-0.0748785,-0.0292746,0.268576;0.313172) 1 PID:0,0,0,1,0 MC:3
trk:4 (-0.914007,-0.203599,3.14832;3.28759) -1 PID:0,0,0,1,0 MC:3
trk:5 (-0.338518,-1.15738,3.99636;4.17666) 1 PID:0,0,1,0,0 MC:5

Here tracks 3 and 4 have quite different 4-vectors, nevertheless the MC index is the same (3) in both cases.

I tried to investigate the effect of the multiple tracks issue by removing tracks with 4-vectors deviating by less the 10E-5 in each component by hand.

The results of an analysis of 1000 D+ -> K- pi+ pi+ (+ c.c.) events are shown in the two attached plots. The blue histogram shows all combinations, the red one is the full MC truth matched part.

FIG 1: Combinatorics with original track list. A very spiky histogram is the result.

FIG 2: Combinatorics of the same 1000 events with removed double tracks. The shape looks reasonable now. Also note, that the number of combinations is only roughly 25% of those above.

I'd like to ask the (tracking?) experts to take a look to that issue, having a very significant impact on analysis results.

Cheers,
Klaus

dear Klaus,
I think I understood the problem. It is not a bug as I initially thought, but rather it is related to tracks going rather forward and therefore releasing only a few hits in the axial STT OUTER layer.
For a reason rather long to explain, such tracks are found twice.
This undesirable feature will be corrected in the new way of doing the track finding that I am working on (a new way of finding the initial clusters to use in the track finding).
If absolutely necessary I could do an ad hoc patch to eliminate the problem, but I would prefer to solve it in the new scheme directly

cheers Gianluigi

Klaus Goetzen wrote on Thu, 22 November 2012 09:24Dear all,

I'd like to report about a problem I found in simulated event concerning tracking and MC truth match.

The first issue is, that I find from time to time track objects, which appear multiple times in the event. I noticed it on the pid-level, but already after the reco level, these objects are present. This might look like this (analysis level; displayed quantities are: track number in event, 4-vector, charge, PID probabilities, MC truth match index)

trk:0 (-1.06059,0.495762,2.45283;2.72148) -1 PID:0,0,1,0,0 MC:8
trk:1 (0.142945,-0.639217,0.900123;1.12193) 1 PID:0,0,1,0,0 MC:4
trk:2 (0.193793,0.341487,1.30556;1.37045) 1 PID:0,0,0,1,0 MC:6
trk:3 (-0.26782,0.799728,2.97571;3.09607) -1 PID:0,0,1,0,0 MC:7
trk:4 (1.00258,-0.368745,5.14993;5.26141) -1 PID:0,0,0,1,0 MC:3
trk:5 (1.00118,-0.368159,5.14244;5.25377) -1 PID:0,0,0,1,0 MC:3

or this

trk:0 (-0.550406,0.169454,0.525641;0.792111) -1 PID:0,0,1,0,0 MC:8
trk:1 (-0.0351313,0.275942,-0.119804;0.333483) -1 PID:0,0,1,0,0 MC:7
trk:2 (0.0599226,0.085189,0.0741163;0.189264) -1 PID:0,0,1,0,0 MC:7
trk:3 (0.0599226,0.085189,0.0741163;0.189264) -1 PID:0,0,1,0,0 MC:7
trk:4 (0.0599226,0.085189,0.0741163;0.189264) -1 PID:0,0,1,0,0 MC:7
trk:5 (0.0599226,0.085189,0.0741163;0.189264) -1 PID:0,0,1,0,0 MC:7
trk:6 (5.92574,0.770891,34.7574;35.2676) 1 PID:0,0,1,0,0 MC:5
trk:7 (-0.078541,-0.0528536,3.60024;3.60419) -1 PID:0,0,0,1,0 MC:3
trk:8 (-0.250432,-0.622125,7.483;7.51429) 1 PID:0,0,1,0,0 MC:4
trk:9 (0.251804,0.0933929,1.55109;1.58035) 1 PID:0,0,0,1,0 MC:6

What one can see is, that in the frist event tracks 4 and 5 are very similar, having also the same MC index 3. This might be due to tracks broken into tracklets, which are treated as different tracks.

On the other hand, in the second event, there are 4 identical track objects (tracks 2-5), which seem to be true clones.

The second issue is, that there are mutiple reco tracks having the same MC truth index. This might be due to the upper issue, where the clones or very similar tracks have the same index. But I also observe events like the following:

trk:0 (0.607878,0.248934,1.06207;1.25657) -1 PID:0,0,1,0,0 MC:8
trk:1 (0.196574,-0.357057,0.765892;0.87875) 1 PID:0,0,1,0,0 MC:4
trk:2 (0.594232,0.99658,2.77834;3.01412) 1 PID:0,0,0,1,0 MC:6
trk:3 (-0.0748785,-0.0292746,0.268576;0.313172) 1 PID:0,0,0,1,0 MC:3
trk:4 (-0.914007,-0.203599,3.14832;3.28759) -1 PID:0,0,0,1,0 MC:3
trk:5 (-0.338518,-1.15738,3.99636;4.17666) 1 PID:0,0,1,0,0 MC:5

Here tracks 3 and 4 have quite different 4-vectors, nevertheless the MC index is the same (3) in both cases.

I tried to investigate the effect of the multiple tracks issue by removing tracks with 4-vectors deviating by less the 10E-5 in each component by hand.

The results of an analysis of 1000 D+ -> K- pi+ pi+ (+ c.c.) events are shown in the two attached plots. The blue histogram shows all combinations, the red one is the full MC truth matched part.

FIG 1: Combinatorics with original track list. A very spiky histogram is the result.

FIG 2: Combinatorics of the same 1000 events with removed double tracks. The shape looks reasonable now. Also note, that the number of combinations is only roughly 25% of those above.

I'd like to ask the (tracking?) experts to take a look to that issue, having a very significant impact on analysis results.

Cheers,
Klaus

---

Subject: Re: Tracking + Truth Match Problem
Posted by Gianluigi Boca on Wed, 09 Jan 2013 22:55:09 GMT
View Forum Message <> Reply to Message

dear Klaus and all,
I put in the repository the new version of the

directory. It contains the version with the new way of doing the clusterization in the offline pattern recognition in the central region.
I think it cures also the problem of the duplicate tracks that was mentioned by Klaus.
Please update and try and let me know the result
thanks
Gianluigi

Klaus Goetzen wrote on Thu, 22 November 2012 09:24Dear all,

I'd like to report about a problem I found in simulated event concerning tracking and MC truth match.

The first issue is, that I find from time to time track objects, which appear multiple times in the event. I noticed it on the pid-level, but already after the reco level, these objects are present. This might look like this (analysis level; displayed quantities are: track number in event, 4-vector, charge, PID probabilities, MC truth match index)

trk:0 (-1.06059,0.495762,2.45283;2.72148) -1 PID:0,0,1,0,0 MC:8
trk:1 (0.142945,-0.639217,0.900123;1.12193) 1 PID:0,0,1,0,0 MC:4
trk:2 (0.193793,0.341487,1.30556;1.37045) 1 PID:0,0,0,1,0 MC:6
trk:3 (-0.26782,0.799728,2.97571;3.09607) -1 PID:0,0,1,0,0 MC:7
trk:4 (1.00258,-0.368745,5.14993;5.26141) -1 PID:0,0,0,1,0 MC:3
trk:5 (1.00118,-0.368159,5.14244;5.25377) -1 PID:0,0,0,1,0 MC:3

or this

trk:0 (-0.550406,0.169454,0.525641;0.792111) -1 PID:0,0,1,0,0 MC:8
trk:1 (-0.0351313,0.275942,-0.119804;0.333483) -1 PID:0,0,1,0,0 MC:7
trk:2 (0.0599226,0.085189,0.0741163;0.189264) -1 PID:0,0,1,0,0 MC:7
trk:3 (0.0599226,0.085189,0.0741163;0.189264) -1 PID:0,0,1,0,0 MC:7
trk:4 (0.0599226,0.085189,0.0741163;0.189264) -1 PID:0,0,1,0,0 MC:7
trk:5 (0.0599226,0.085189,0.0741163;0.189264) -1 PID:0,0,1,0,0 MC:7
trk:6 (5.92574,0.770891,34.7574;35.2676) 1 PID:0,0,1,0,0 MC:5
trk:7 (-0.078541,-0.0528536,3.60024;3.60419) -1 PID:0,0,0,1,0 MC:3
trk:8 (-0.250432,-0.622125,7.483;7.51429) 1 PID:0,0,1,0,0 MC:4
trk:9 (0.251804,0.0933929,1.55109;1.58035) 1 PID:0,0,0,1,0 MC:6

What one can see is, that in the frist event tracks 4 and 5 are very similar, having also the same MC index 3. This might be due to tracks broken into tracklets, which are treated as different tracks.

On the other hand, in the second event, there are 4 identical track objects (tracks 2-5), which seem to be true clones.

The second issue is, that there are mutiple reco tracks having the same MC truth index. This might be due to the upper issue, where the clones or very similar tracks have the same index. But I also observe events like the following:

trk:0 (0.607878,0.248934,1.06207;1.25657) -1 PID:0,0,1,0,0 MC:8

trk:1 (0.196574,-0.357057,0.765892;0.87875) 1 PID:0,0,1,0,0 MC:4
trk:2 (0.594232,0.99658,2.77834;3.01412) 1 PID:0,0,0,1,0 MC:6
trk:3 (-0.0748785,-0.0292746,0.268576;0.313172) 1 PID:0,0,0,1,0 MC:3
trk:4 (-0.914007,-0.203599,3.14832;3.28759) -1 PID:0,0,0,1,0 MC:3
trk:5 (-0.338518,-1.15738,3.99636;4.17666) 1 PID:0,0,1,0,0 MC:5

Here tracks 3 and 4 have quite different 4-vectors, nevertheless the MC index is the same (3) in both cases.

I tried to investigate the effect of the multiple tracks issue by removing tracks with 4-vectors deviating by less the 10E-5 in each component by hand.

The results of an analysis of 1000 D+ -> K- pi+ pi+ (+ c.c.) events are shown in the two attached plots. The blue histogram shows all combinations, the red one is the full MC truth matched part.

FIG 1: Combinatorics with original track list. A very spiky histogram is the result.

FIG 2: Combinatorics of the same 1000 events with removed double tracks. The shape looks reasonable now. Also note, that the number of combinations is only roughly 25% of those above.

I'd like to ask the (tracking?) experts to take a look to that issue, having a very significant impact on analysis results.

Cheers,
Klaus

---

Subject: Re: Tracking + Truth Match Problem
Posted by Klaus Götzen on Wed, 30 Jan 2013 08:57:49 GMT
View Forum Message <> Reply to Message

Hi,

just as an update some plots with the new code.

This plots shows Ds->K+ K+ pi+- without removale of doubles. It looks already quite reasonable (blue=all combinations, red= mc truth matched):

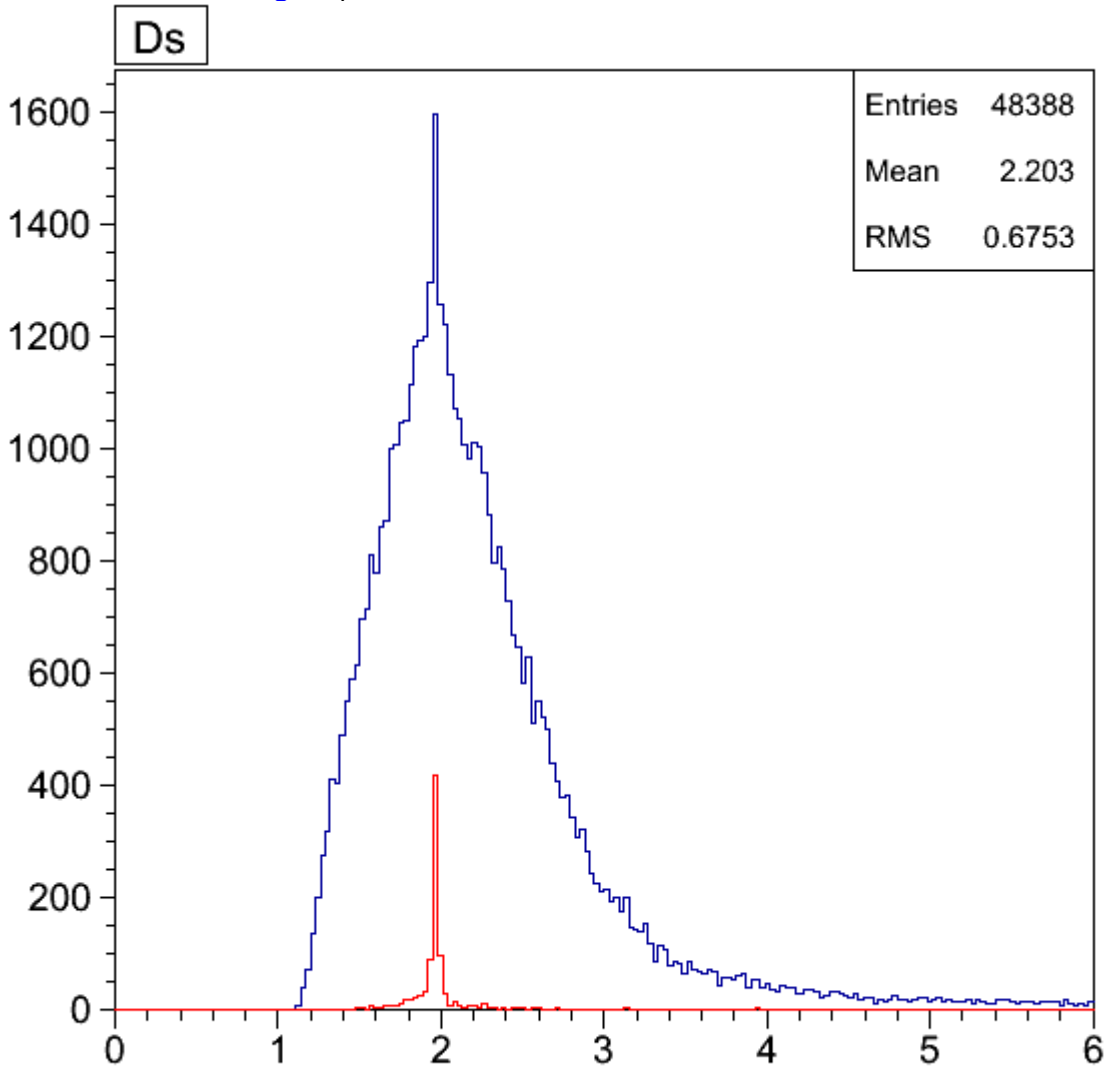
After removing double trackes (deviation of each component <0.001; removed 75/6267 = 1.2%

of the tracks), it's still a bit less twitchy (7% less combinations):

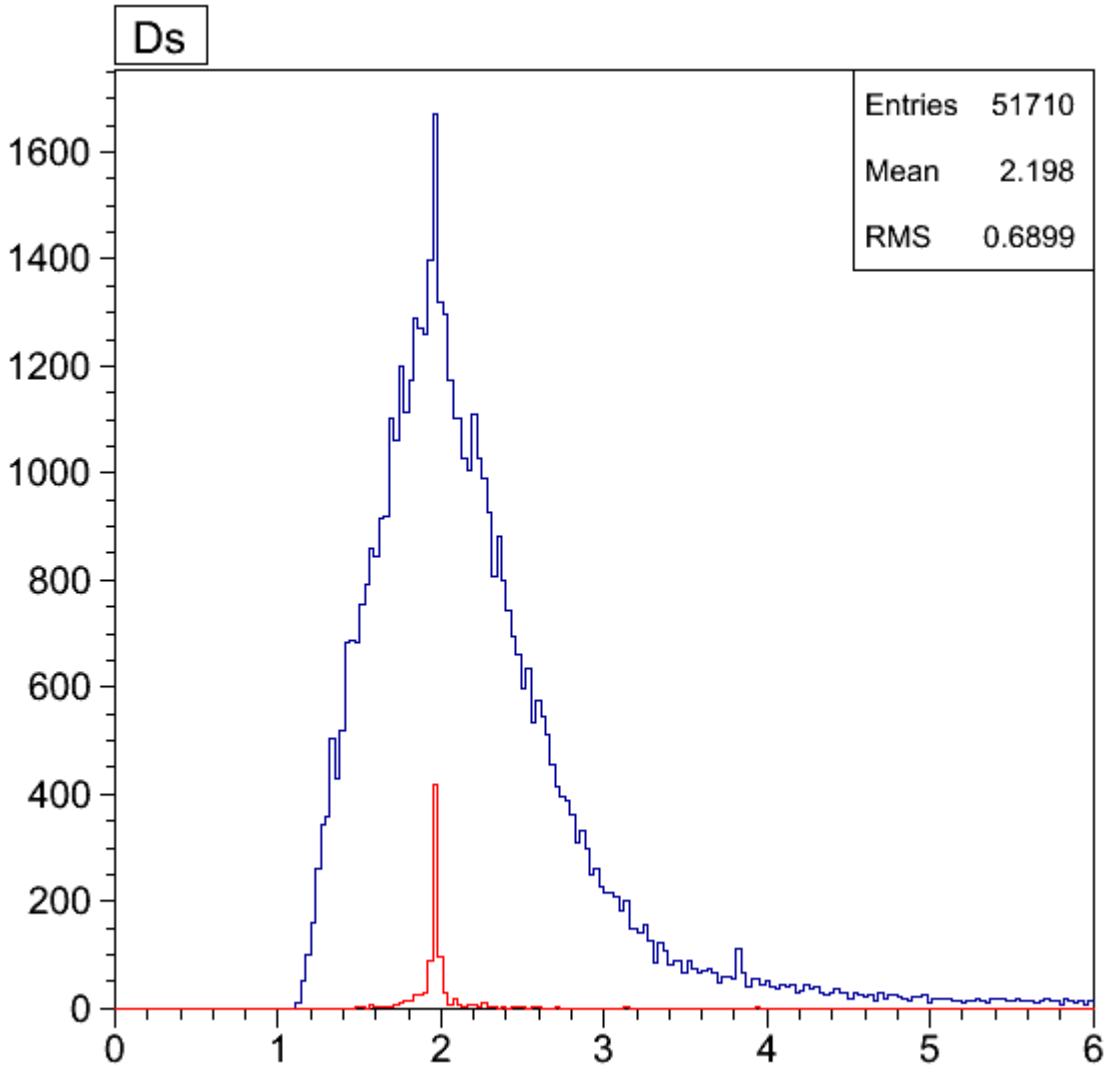
Best regards,
Klaus

## File Attachments

1) ds_default.gif, downloaded 1307 times




2) ds_remove.gif, downloaded 1242 times




Subject: Re: Tracking + Truth Match Problem
Posted by StefanoSpataro on Wed, 30 Jan 2013 09:16:28 GMT
View Forum Message <> Reply to Message

Hi klaus,
thanks for your pictures, I have only one question.

If I compare the new picture to what you showed in the first message, after your doubles cleanup, I can see that in the old plot you have a peak at 1000 and a signal peak at 400, in the new plot you have a peak at 1600 entries and a signal peak at 400.
It seems to me that now we have less doubles, the same signal efficiency, but we have a higher combinatorial after your doubles cleanup than before. In this sense the situation has improved but maybe not so much,
Or have I missed something? What do you think?

Subject: Re: Tracking + Truth Match Problem

Posted by Klaus Götzen on Wed, 30 Jan 2013 14:32:07 GMT
View Forum Message <> Reply to Message

Hi Stefano,

sorry for the confusion, the channel is a different one (Ds -> KKpi instead of D+ -> Kpipi), so you can't compare them directly. I could run on the same dataset and also post the plots for D+ decays.

I just wanted to show the relative improvement of having 3x the combinatorics compared to only 7% more after the fix.

Best,
Klaus

---

Subject: Re: Tracking + Truth Match Problem
Posted by StefanoSpataro on Wed, 30 Jan 2013 15:33:30 GMT
View Forum Message <> Reply to Message

It would be nice to have the same plot as before, just to be sure that fixing one thing we are not destroying something else

---

Subject: Re: Tracking + Truth Match Problem
Posted by Klaus Götzen on Fri, 01 Feb 2013 12:37:10 GMT
View Forum Message <> Reply to Message

Hi,

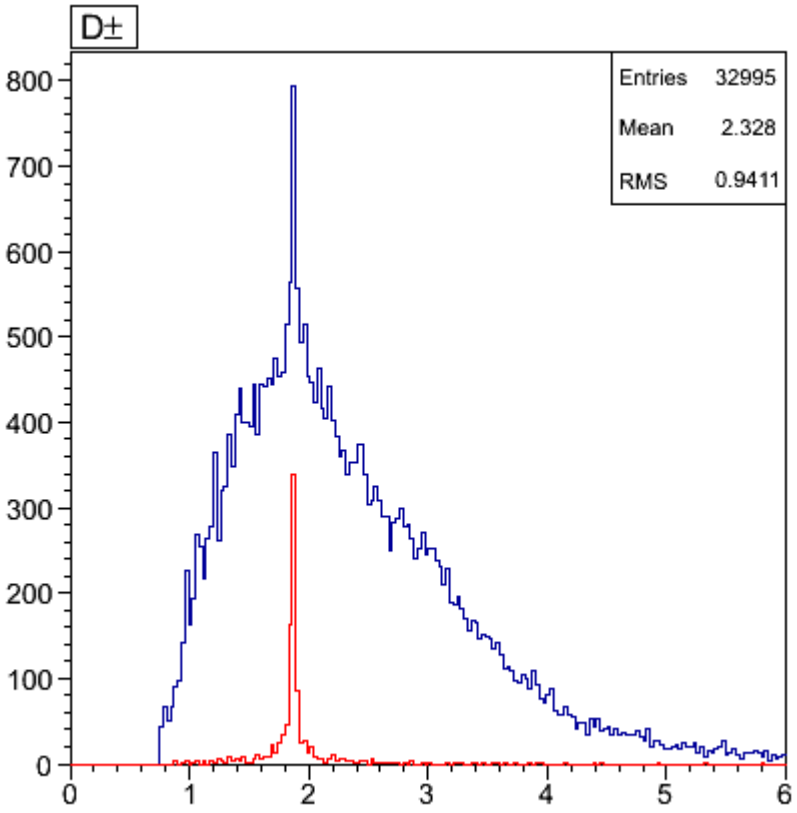
here again the comparison for the D+- -> K+- pi+ pi- spectra (D+ D- @ 15GeV) without (left) and with (right) double track removal after Gianluigis fix.

The number of removed tracks was 92/6553 = 1.4%, there are roughly 9% more combinations (32995/30255 = 1.09).

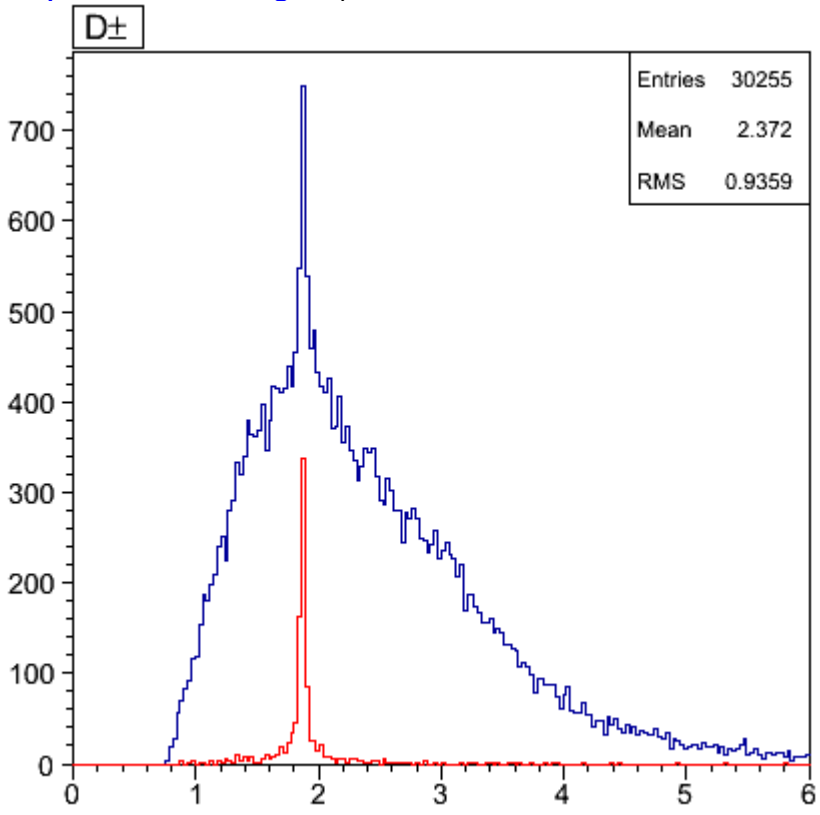
Best,
Klaus

## File Attachments

1) dpdm55_default.gif, downloaded 1307 times




2) dpdm55_rem.gif, downloaded 1223 times



Subject: Re: Tracking + Truth Match Problem
Posted by StefanoSpataro on Fri, 01 Feb 2013 13:11:48 GMT
View Forum Message <> Reply to Message

Thanks Klaus for the additional work,
if I compare the new plots with the ones at the first message, still I can see that the MC signal peak (350) is now lower than before (450). In this sense we have lost efficiency or the resolution has worsen.
Let's see what will come from the tracking efficiency studies.