REPORT IN PROGRESS

HINTS FOR PID IN RICH 1

COMPASS Note - May 2003.

P. Schiavon

1 Introduction

This report is intended to give some hints for the use of the Likelihood values in the RICH Particle IDentification (PID).
The Likelihood used is the 'RING Likelihood' (as presently in RICHONE), in which only the *photons* belonging to the reconstructed *ring* enter the Likelihood computation, as opposed to the 'ALL Likelihood', in which all the *photons* relative to a particle trajectory are used.
The Likelihood values are computed for the π, K or p mass hypothesis and for the background only hypothesis, for reference.
To that purpose the particle **momentum** and the **refractive index** of the RICH radiator are needed.

2 The refractive index

The following analysis has been made on some sample runs (22627, 22642 and 22682), from production of period P2E; the RICH special files (gfiles) have been processed with a value of the refractive index known to the best. Typically production runs are processed with values of the refractive index known with some approximation; a correction procedure has been envisaged (see section 6).

3 Likelihood analysis

The following quantities are used in this analysis:
- the **particle momentum**; the momentum is known from the tracking; the analysis is carried out in the particle momentum range from 8. to 38. GeV/c, split in five momentum bins of 6. GeV/c each. Note that, in the present exercise, (refractive index n = 1.001266), the momentum threshold for K'^s is $\sim$ 9.8 GeV/c and then that a not negligeable number of K'^s under threshold are considered.
- the three **ratios** of the Likelihoods L_π, L_K, L_p to the background Likelihood, L_B; a

particle is said to be identified as, e.g., a K when $R_K = L_K/L_B$ is maximum;
- the three **ratios** of the maximum Likelihood to the second largest Likelihood (L_{MAX2}); e.g. $r_K = L_K/L_{MAX2}$.
A correlation exists between R_j and r_j (for each particle type j), see fig. 1, and a correlation is also expected between R_j and $\Delta\Theta_j = \Theta_{RING} - \Theta_{MASS}$ for mass j, see fig. 2. This means, of course, that a cut in one of the two quantities is not independent on a cut in the second one.

Figure 3, white, shows the mass distribution as calculated from the particle momentum, in the range indicated (8 - 38 GeV/c), and from the reconstructed *ring* angle Θ_{RING}; in color for the particles identified as K'^s, requiring the ratio L_K/L_B to be maximum. Note as the K mass peak is left uncut in the identification.
The width of the K mass peak is large because of the worsening of the mass resolution as a function of the momentum and mostly because of the background under and to the right of the peak.
Θ_{RING} can be badly reconstructed for different reasons: poor resolution around the correct value, e.g., because of a too small number of *photons* per *ring*; wrong ($\simeq$ random) value reconstruction because of competitive background; wrong ($\simeq$ random) value reconstruction because of particles under Cherenkov threshold. The shoulder to the right of the K mass peak is contributed by Θ_{RING} values smaller than Θ_{MASS} (see also fig. 4 (b) to 8 (b)).

Fig. 4 to 8 show, for five momentum bins, some distributions on which cuts are made.
In (a) the plot of L_K/L_B versus the reconstructed mass; L_K/L_B close to 1. means that the Likelihood value is not significantly different from that of the bachground (no signal) and then that the identification is not well defined; the horizontal line is the cut to be applied; with reference to fig. 2, one sees that this cut is strongly correlated with a cut in $\Delta\Theta$. In (b) the corresponding reconstructed mass spectrum is shown (no cut); in (c) the distribution of L_K/L_B for all entries and in (d) the distribution of L_K/L_B for a mass slice ($\pm$ 0.010 GeV/c^2) around the nominal K mass. The vertical lines set the cut to be applied.
Cut1 is then defined, for each momentum bin, as the cut defined in Fig. 4 to 8, (a) or (c) and (d).
Note as, for higher particle momenta, the background to the right of the K mass peak, coming form 'wrong' Θ_{RING} smaller than Θ_{MASS}, is more important compared to the K mass region.

Figure 9 shows the K mass distribution as in fig. 3. In color the mass distribution after having applied Cut1, together with (white) the uncut distribution of fig. 3; note as the background, in particular the shoulders, results reduced.

Fig. 10 to 14 show, again for five momentum bins, after Cut1, in (a) the distributions of L_K/L_{MAX2} versus the reconstructed mass; in (b) the corresponding mass distribution (Cut1 only); in (c) the distribution of the ratio L_K/L_{MAX2}; in (d) the same distribution for a mass slice ($\pm$ 0.010 GeV/c^2) around the nominal K mass. L_K/L_{MAX2} close to 1. in principle means again a loosely defined identification, in the sense that the 'good' mass hypothesis is hardly distinguishable from the 'less good' one.
The lines define a possible second cut, **Cut2**; note that Cut2 includes Cut1.

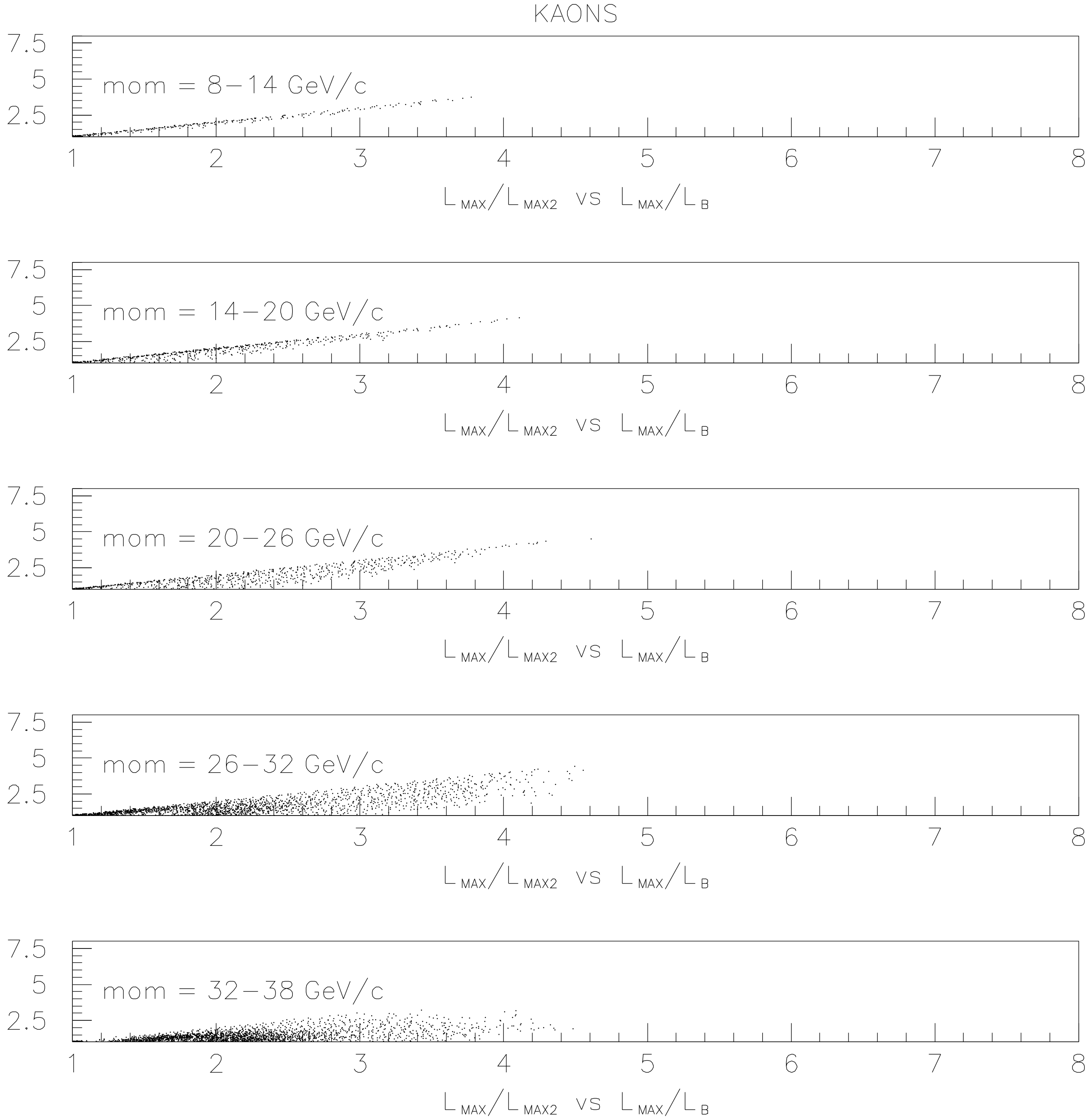

Figure 1: The figure shows, for the momentum bins indicated and for identified kaons, the correlation between the ratio L_K/L_B, when $L_K = L_{MAX}$, and the ratio L_K/L_{MAX2}, where L_{MAX2} is the second maximum Likelihood.

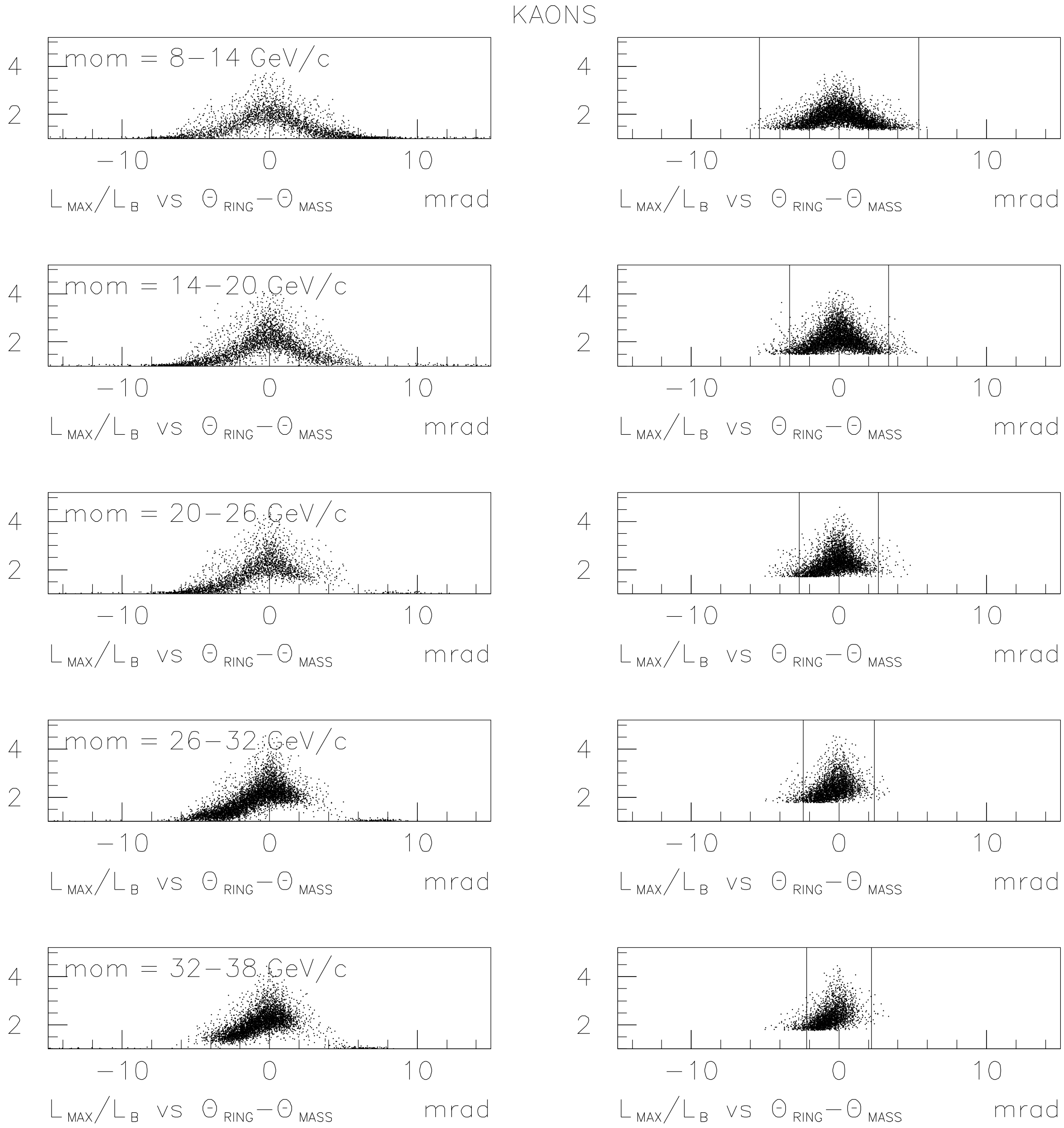

Figure 2: The figure shows, at the left, for the momentum bins indicated and for identified kaons, the plots of the correlation between the ratio L_K/L_B, when L_K is maximum, and the difference $\Theta_{RING} - \Theta_{MASS}$ ($\Delta\Theta$), the Cherenkov angle of the reconstructed *ring* and the Cherenkov angle for the K mass. At the right, the same plots after having applied to the Likelihood the two cuts discussed in the text (Cut1 and Cut2); the correlation with a cut in $\Delta\Theta$ is clear; the lines are the 3 σ limits used in the definition of the monitor of the *ring* reconstruction efficiency (discussed in previous reports).

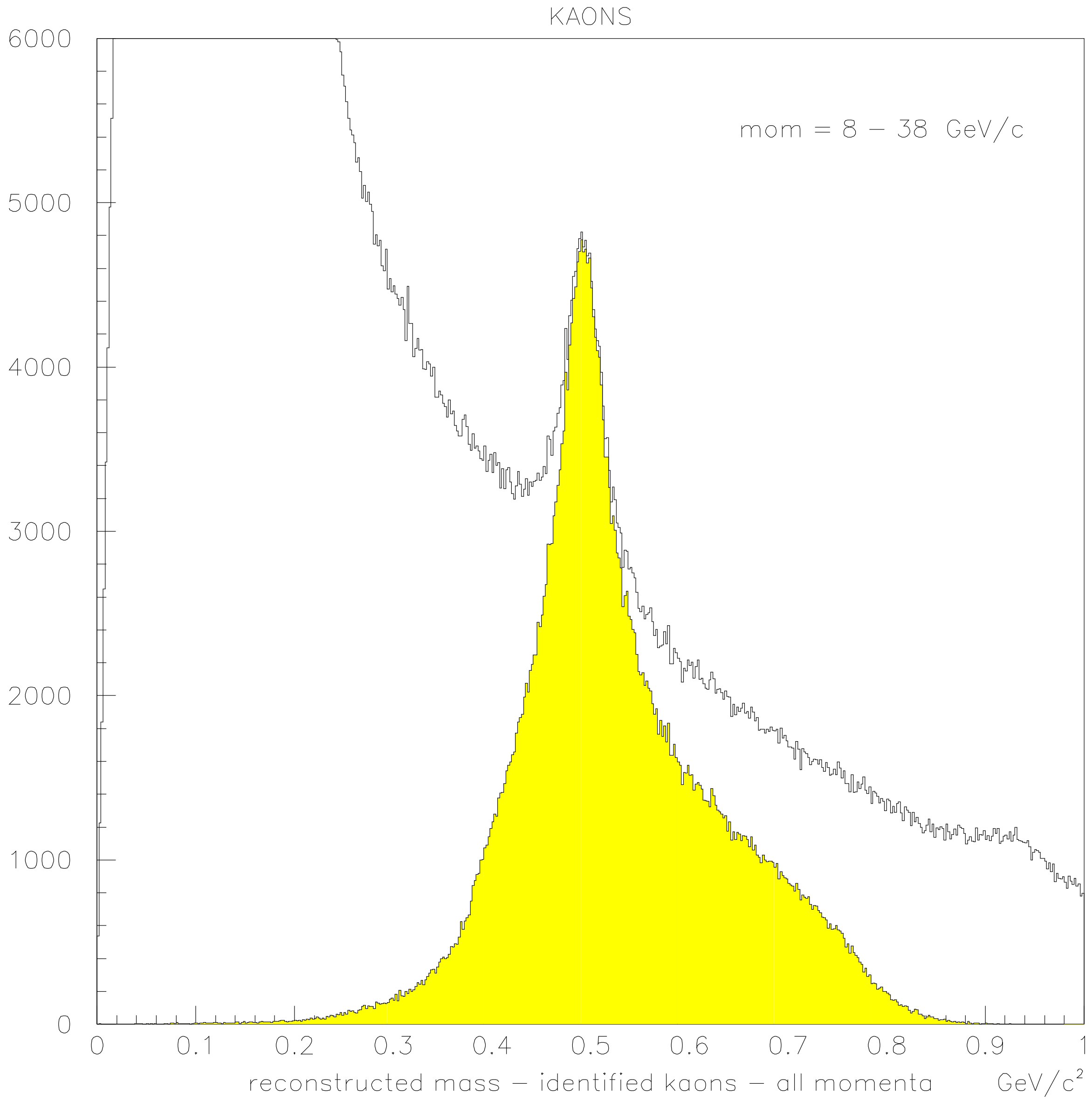

Figure 3: The reconstructed mass spectrum for all particles (white) and for identified kaons (color), within the momentum range indicated.

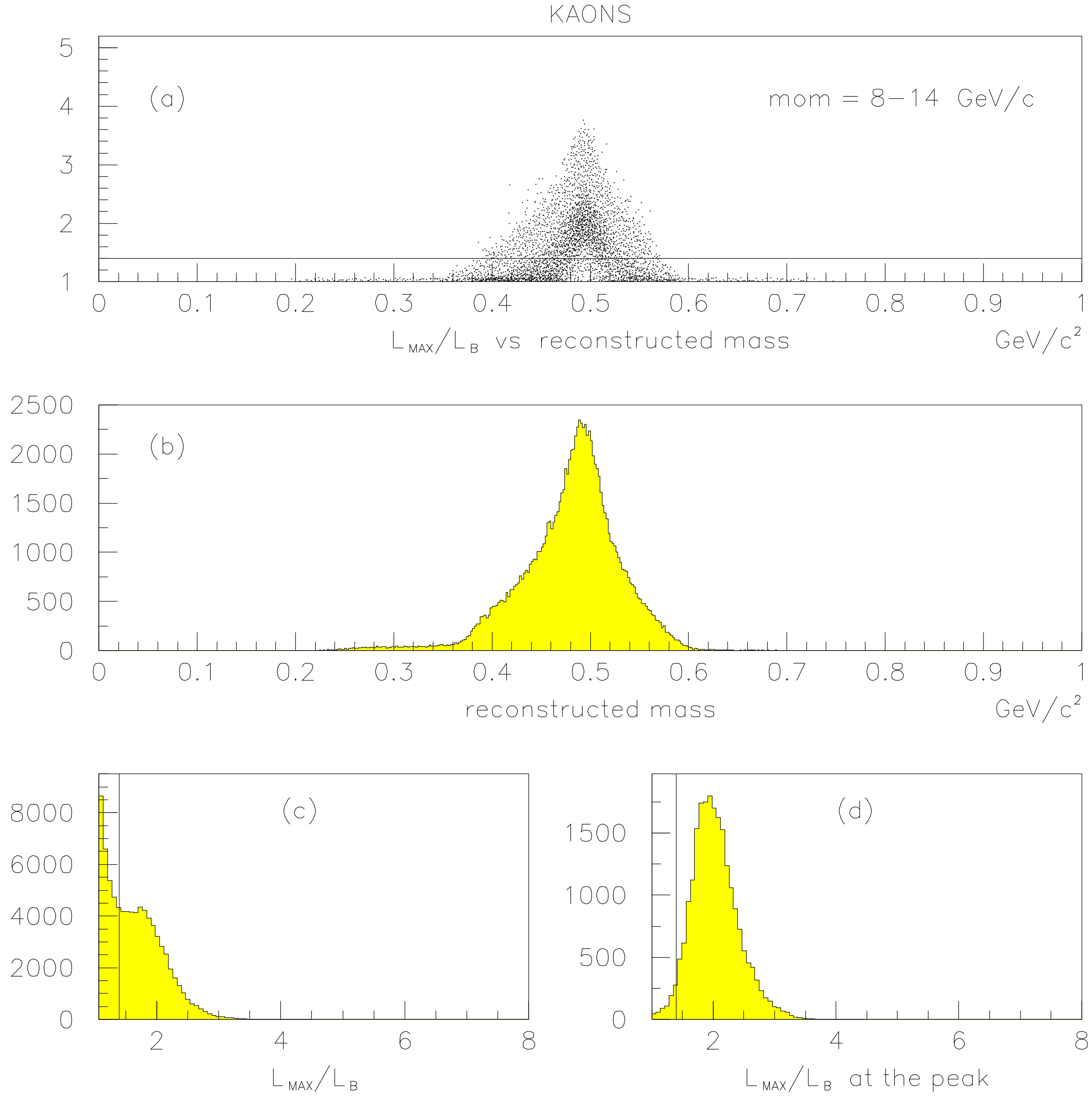

Figure 4: The figure shows, for the momentum range indicated and for identified kaons, in (a), the plot of the ratio L_K/L_B versus the reconstructed mass, in (b) the mass ditribution (no cut), in (c) the distribution of the ratio L_K/L_B and in (d) the distribution of the ratio in a slice around the K mass. The lines in (a), (c) and (d) represent the cut (Cut1) to be applied.

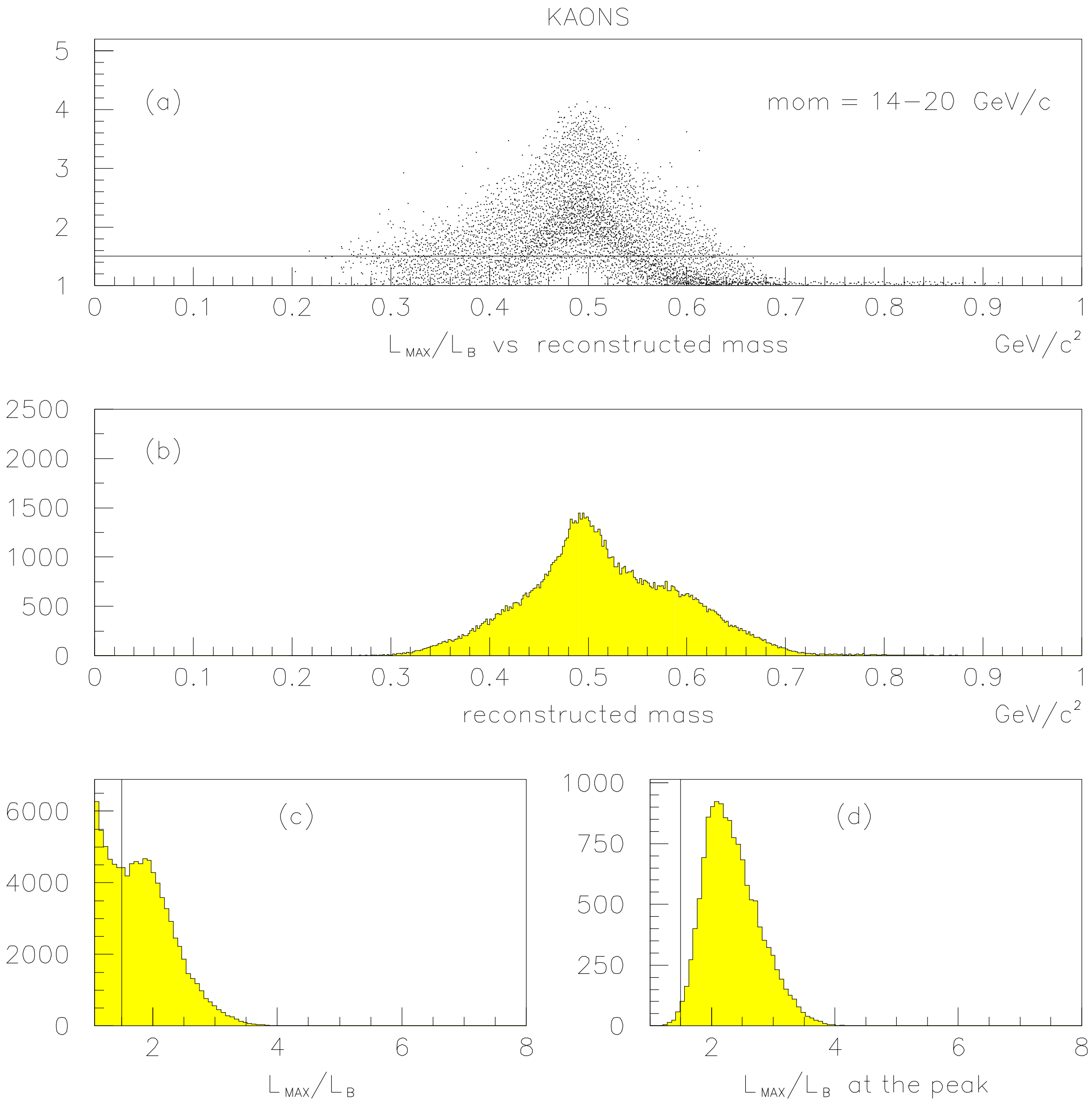

Figure 5: As fig. 4, for the momentum range indicated.

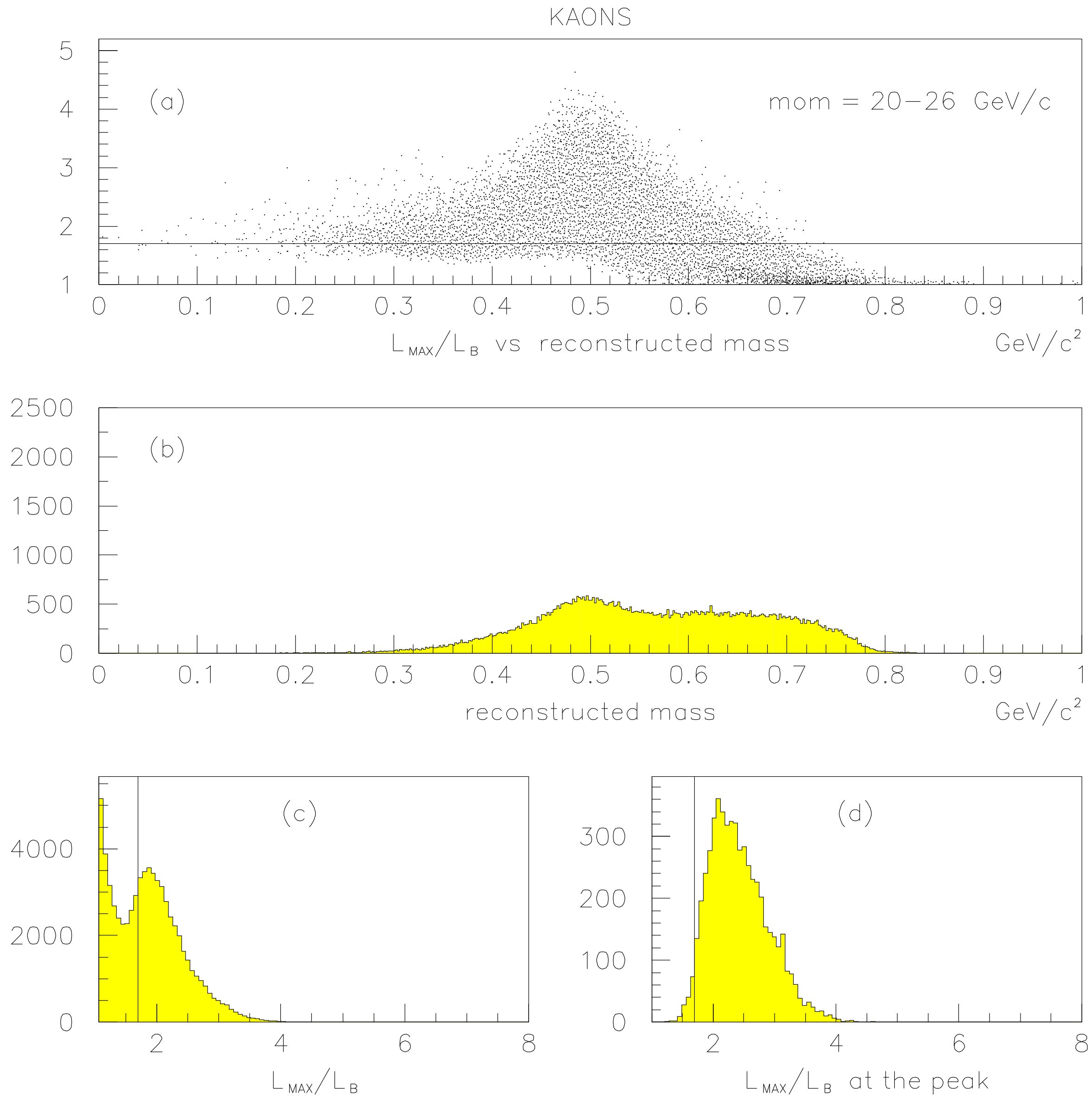

Figure 6: As fig. 4, for the momentum range indicated.

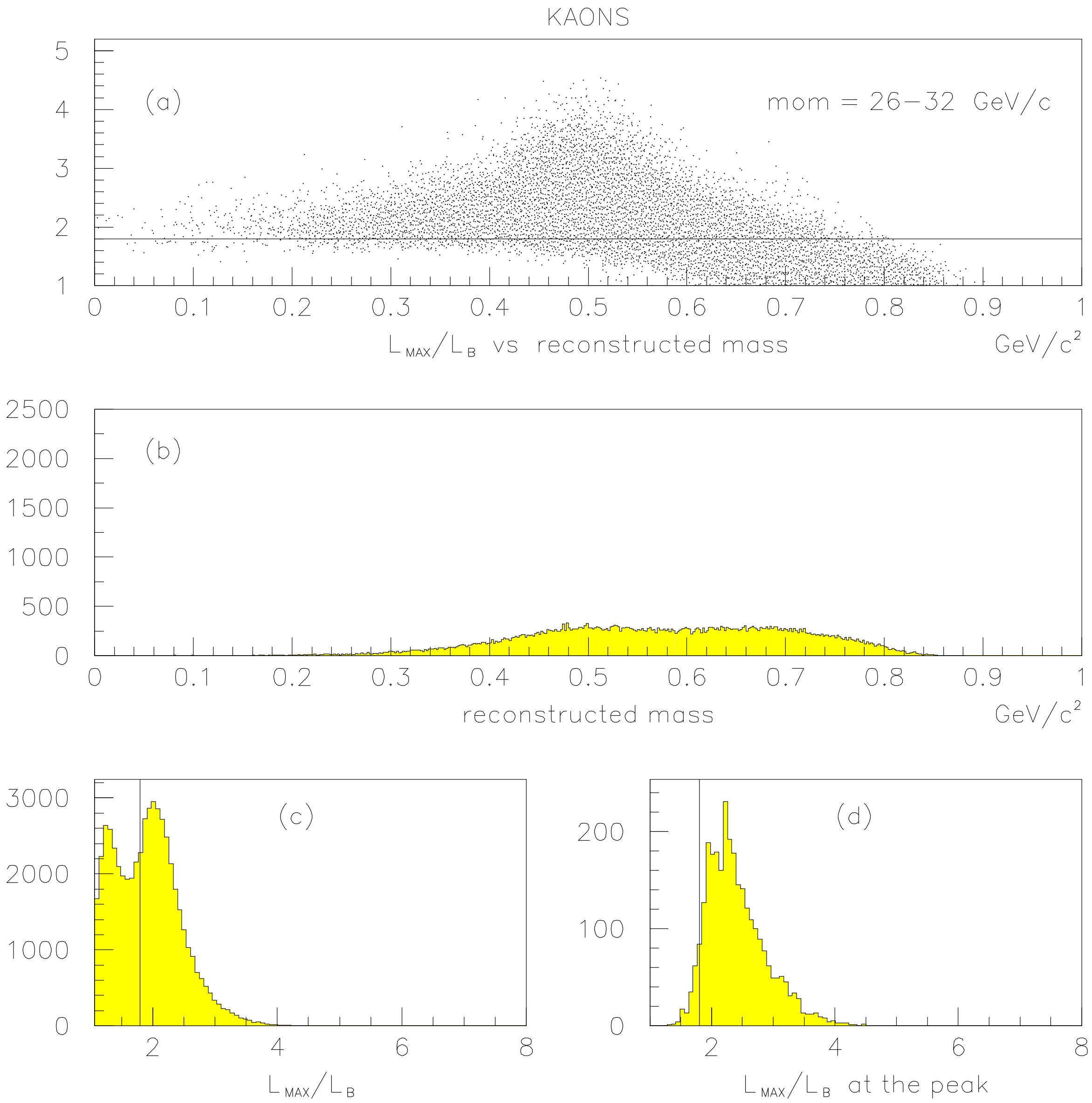

Figure 7: As fig. 4, for the momentum range indicated.

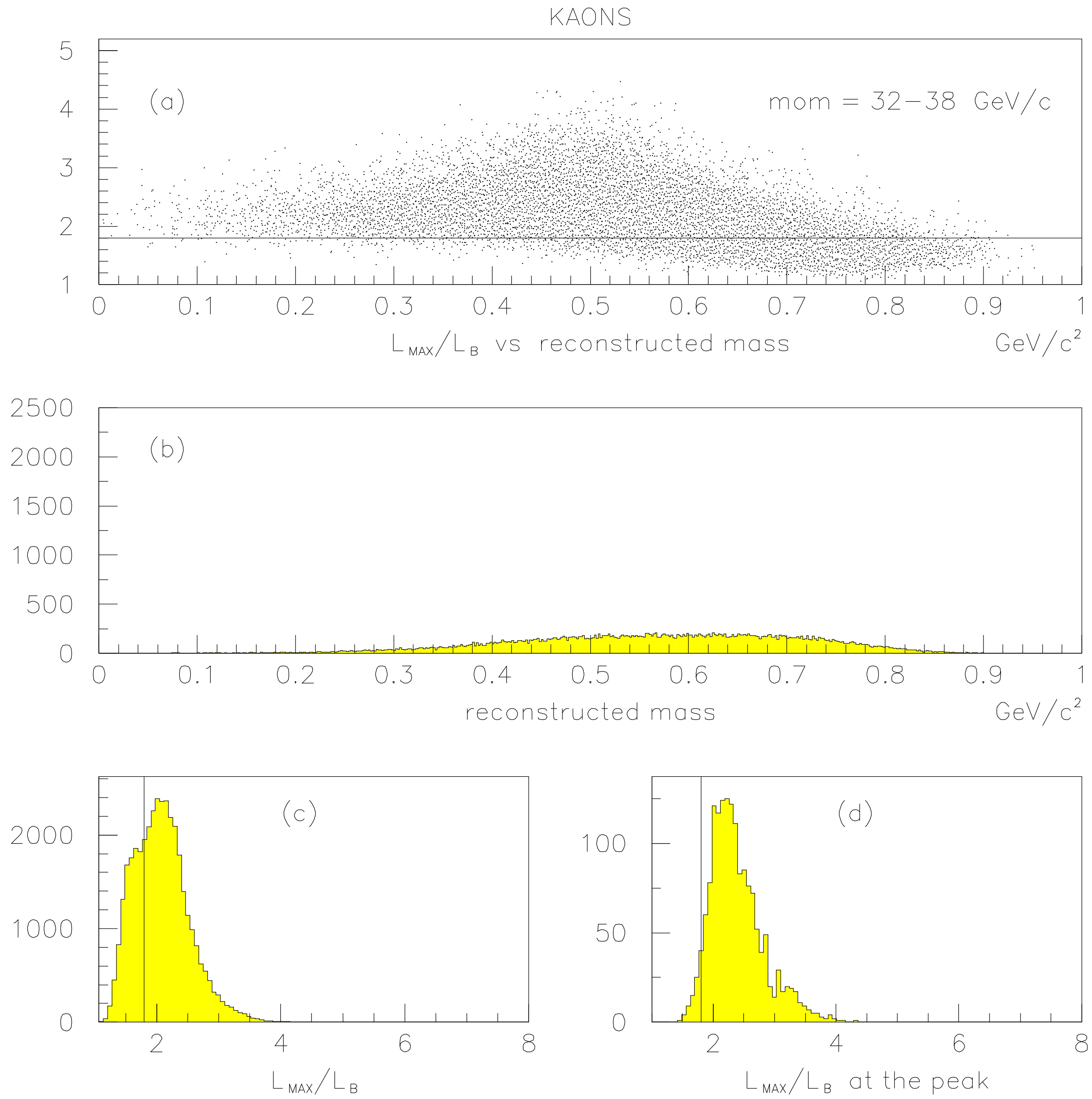

Figure 8: As fig. 4, for the momentum range indicated.

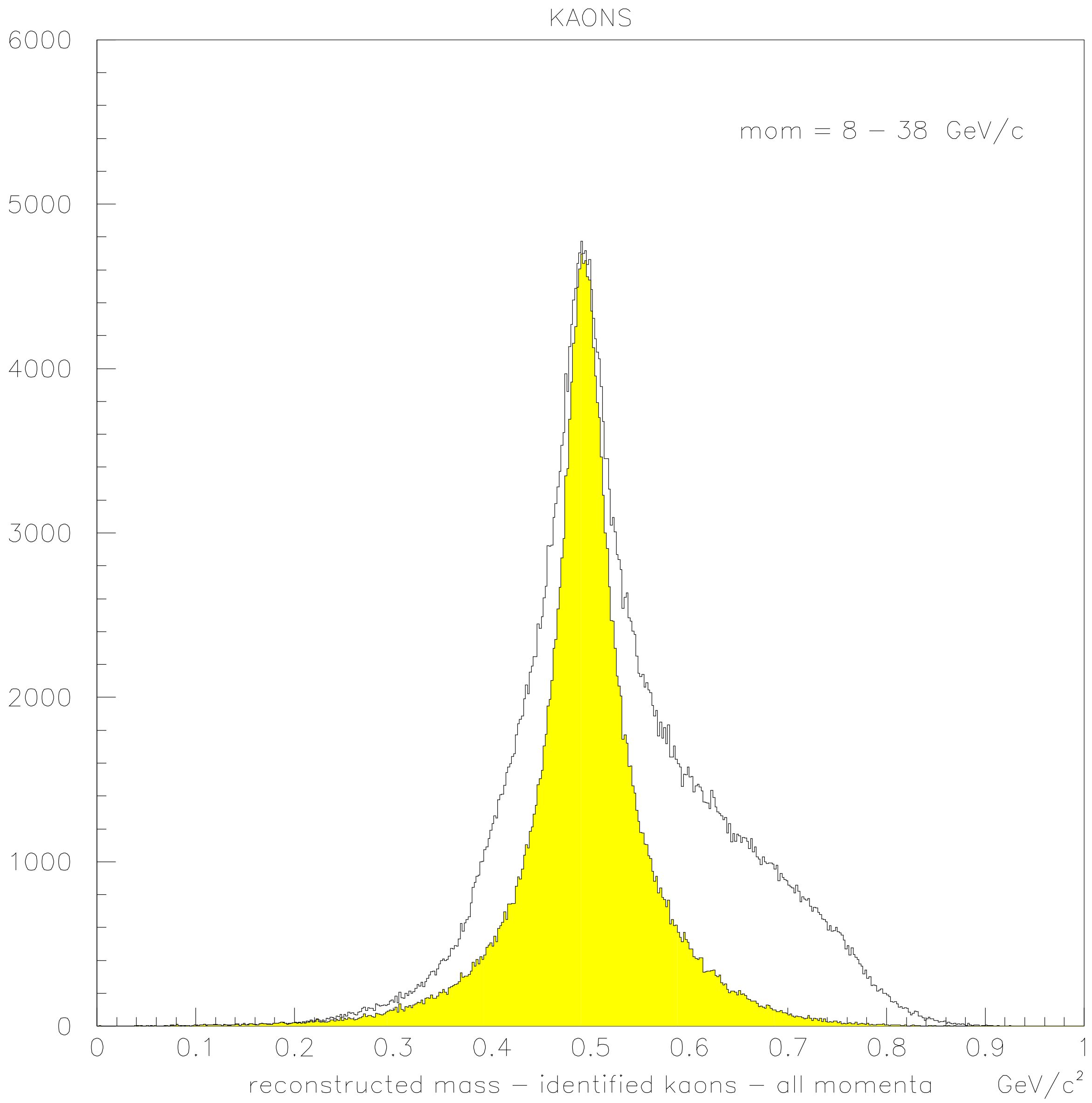

Figure 9: The figure shows, for the momentum range indicated, in color, the reconstructed mass spectrum of identified kaons, after the cuts in L_K/L_B, as described in the text (Cut1), superimposed to the uncut K mass distribution. of fig. 3.

Fig, 15 shows the reconstructed mass distributions after Cut1 and Cut2.

Figure 16 shows, in color, the mass distribution after having applied Cut1 and Cut2, together with (white) the uncut distribution of fig. 3. The background tails result further decreased, but some asymmetry appears. The question arises whether some 'excess' of background is left to right of the mass peak with respect to the left (Cut1 less effective than expected? See fig.s 4 to 8), or Cut2 is not correctly applied, i.e. overdoing; the particles cut by Cut2 correspond to Θ_{RING} larger than Θ_{MASS}, the high beta region, where we expect the Likelihood values for the different mass hypotheses to be very close each other.

A further polishing of the data can be done introducing a cut at 7 in the number of *photons* per reconstructed *ring* (N $\geq$ 7) and a cut at 3.0 in the *ring* reduced χ^2 ($\chi^2/\nu \leq 3.0$); see fig. 17.
Note that the *ring* reduced χ^2 does not peak at ~ 1 as expected, but rather peaks at larger values as the momentum increases. Errors should be better calibrated as a function of the momentum and particle type.
The resulting K mass peak is shown in fig. 18 (color) together with the uncut mass distribution (white) of fig. 3. With respect to fig. 16, the mass peak exhibits a reduced height; the meaning should be that 'signal' events less well measured are discarded.

Looking at the plot R_j vs m_j, the reconstructed mass, of fig. 4 (a) and following, the correlation between R_j and m_j, is such that, for recontructed masses close to m_j, R_j is large; this means that Cut1 (and Cut2 too) cuts doubtful cases, but do not cut the background under the 'signal' peak, as the height of the peak itself indicates. It is then hard to tell how much background remains under the 'signal' peak.

4 Consistency checks

To try to understand how the Likelihood analysis affects the background under the K mass peak, the events have been processed taking a random correspondence between the particle trajectories of one event and the RICH active pads of another one ('scrambled' processing).

In fig. 19, in white, the raw reconstructed mass distribution of fig. 3 (white) together with the 'scrambled' one, in color, normalized assuming equal the number of counts inside the two mass regions defined by the dotted lines.
In fig. 20, in white, the raw reconstructed mass distributions, for the momentum bins indicated, together with the 'scrambled' ones (color), normalized as explained above, to show the dependence on the momentum.
In fig. 19, the shape of the background under the K mass peak and of the 'scrambled' distribution agree well (even if not exactly); an attempt is made of taking the difference of the two distributions, the result being the 'peak' around the K mass, between the two (solid) lines (K mass $\pm$ 150 MeV/c^2).

Fig. 21 (a) shows, in white, the raw reconstructed mass distribution of fig. 3; in color, the identified K mass distribution of fig. 3; inside, the 'scrambled' mass distribu-

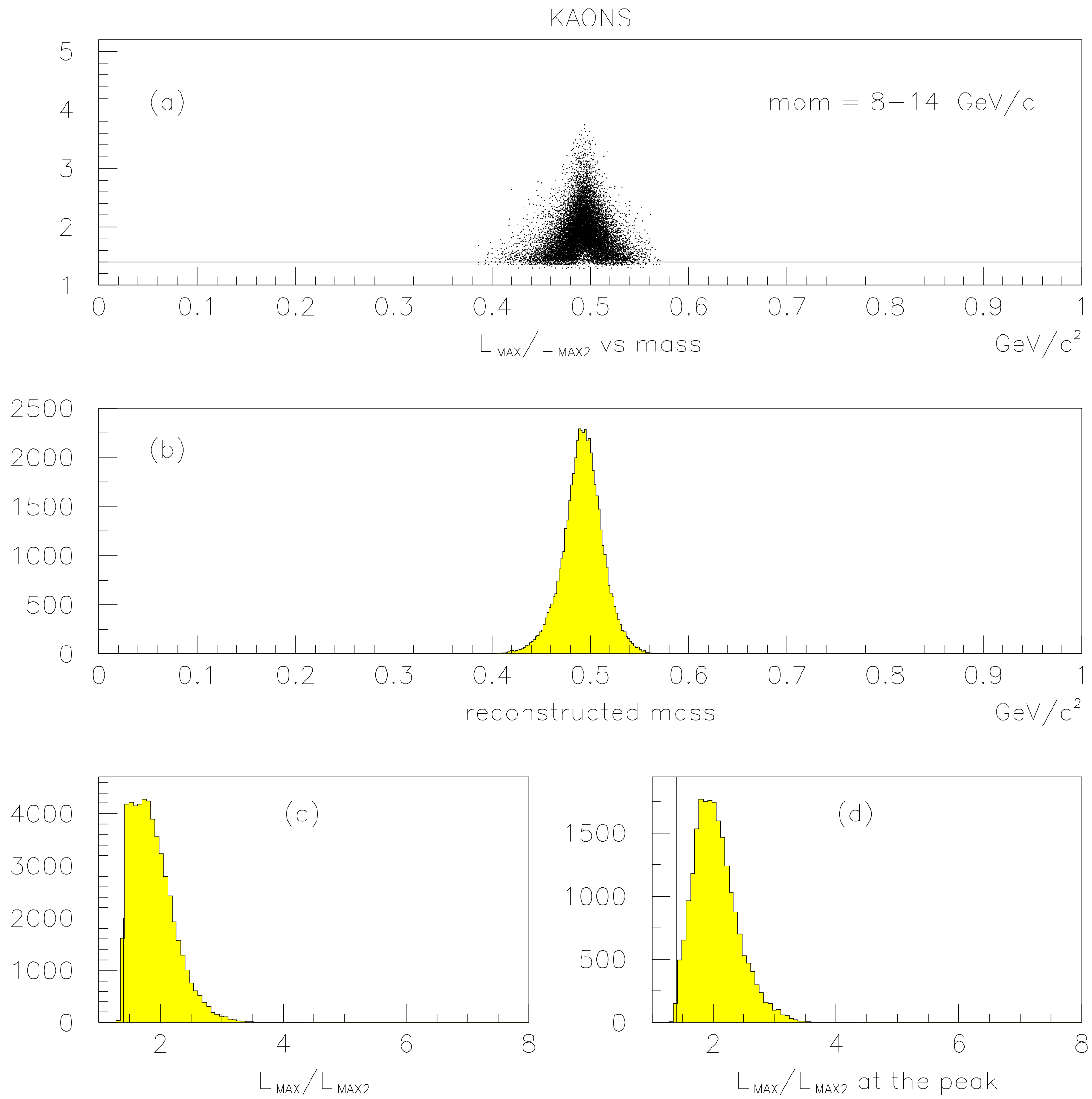

Figure 10: The figure shows, for the momentum range indicated and for identified kaons, in (a), the plot of the ratio L_K/L_{MAX2} versus the reconstructed mass; in (b) the corresponding mass distribution (Cut1 only); in (c) the distribution of the ratio L_K/L_{MAX2} and in (d) the distribution of the ratio in a slice around the K mass. The lines in (a), (c) and (d) represent the cut (Cut2) to be applied.

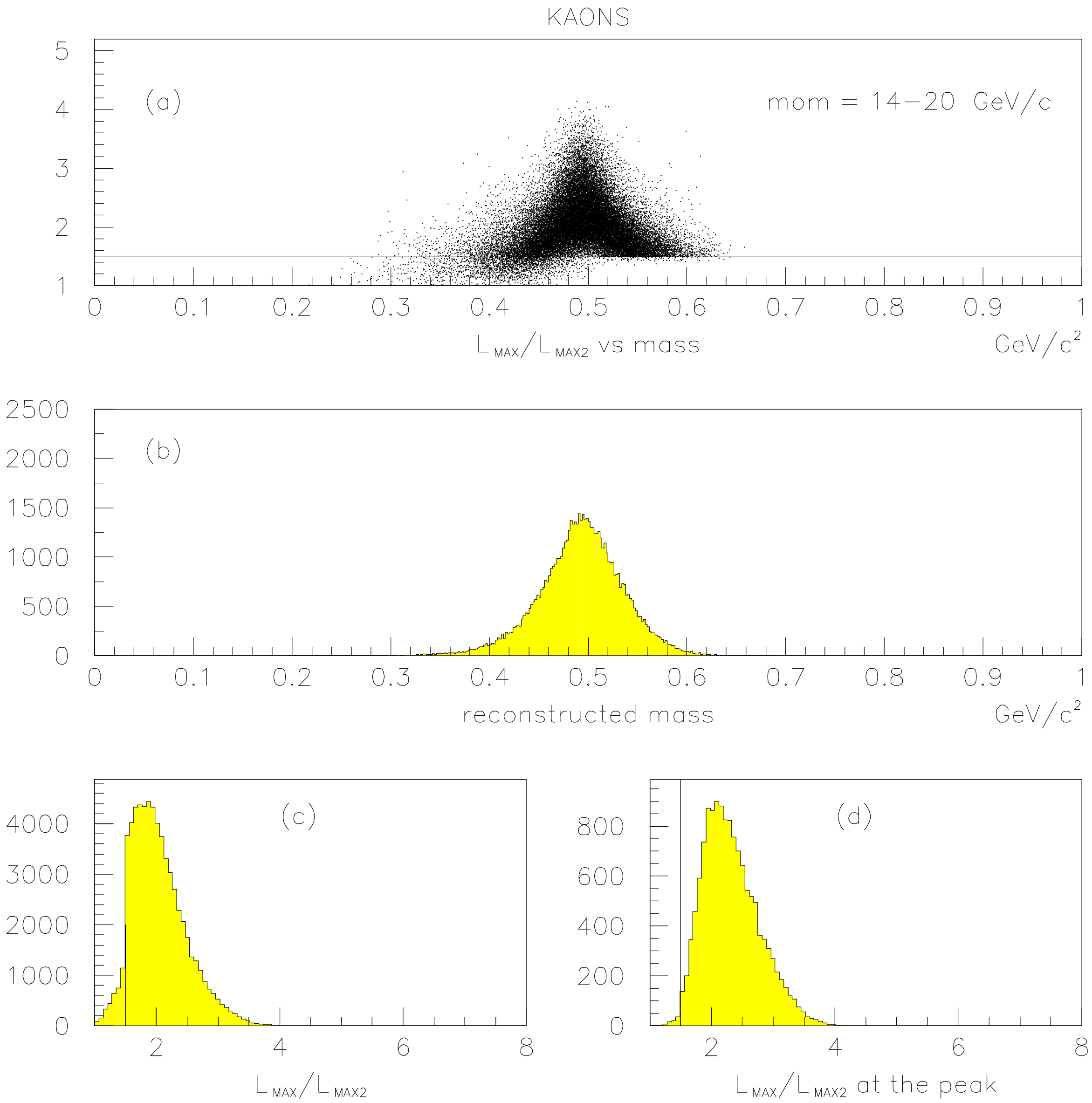

Figure 11: As fig. 10, for the momentum range indicated.

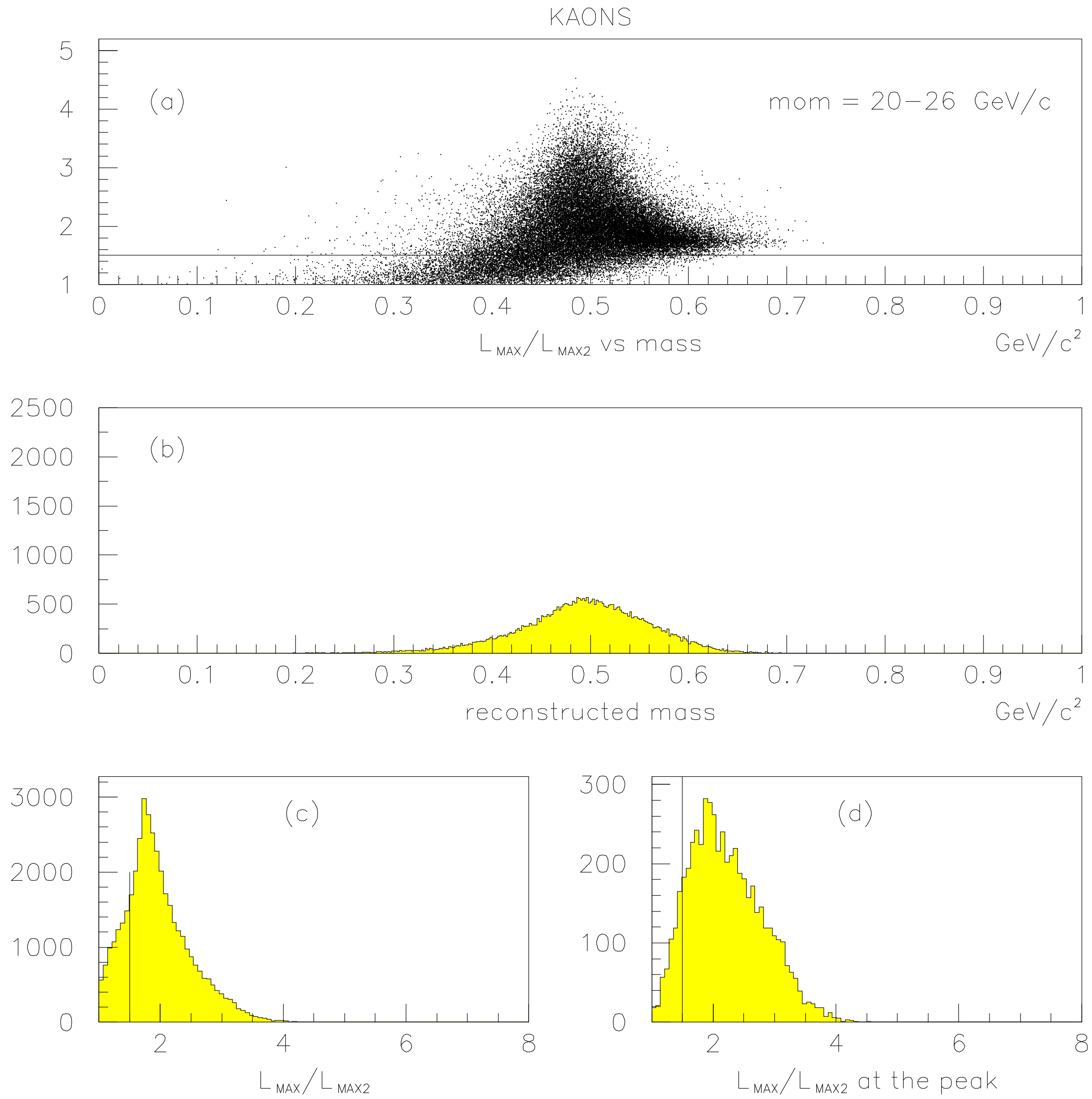

Figure 12: As fig. 10, for the momentum range indicated.

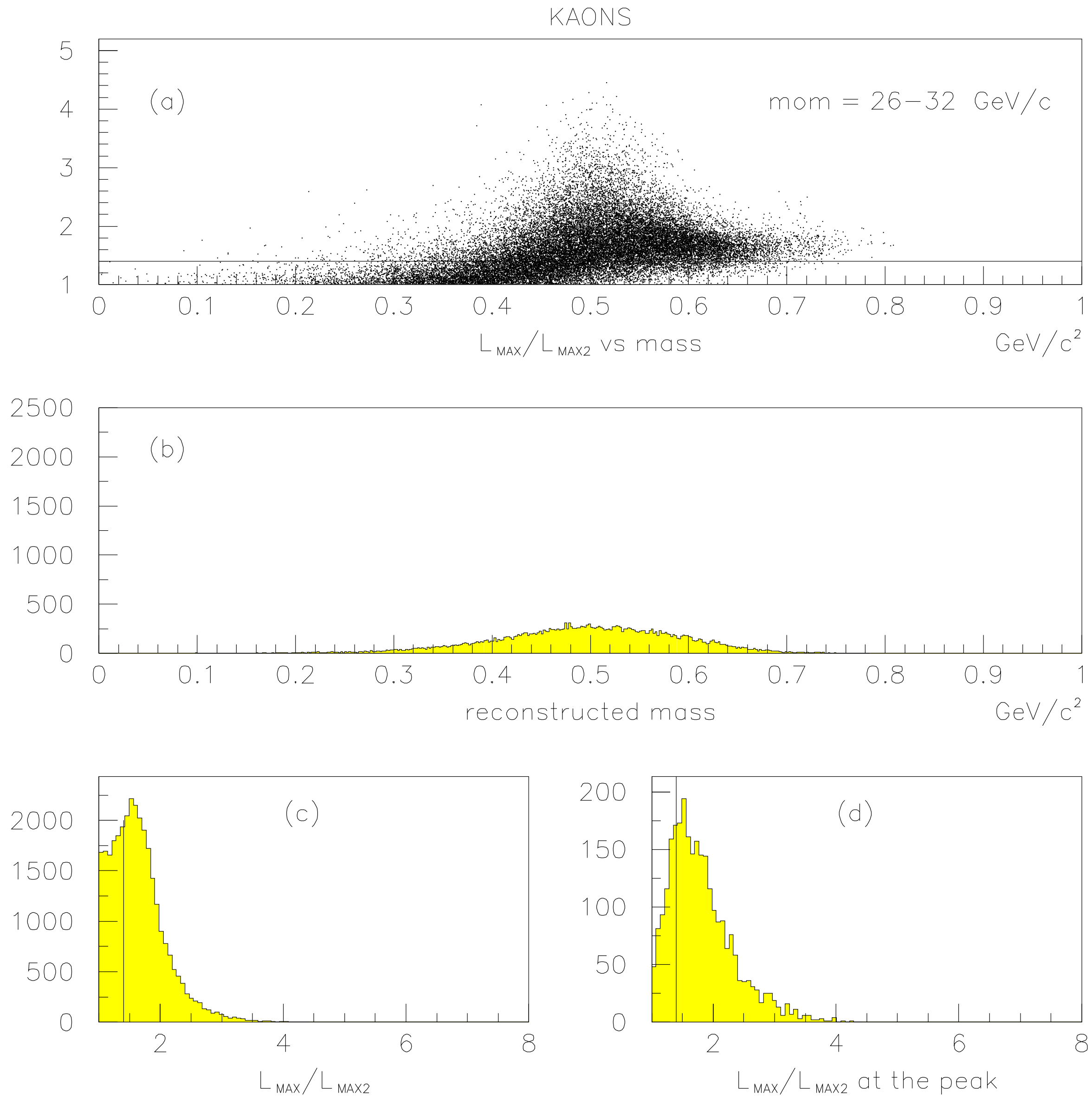

Figure 13: As fig. 10, for the momentum range indicated.

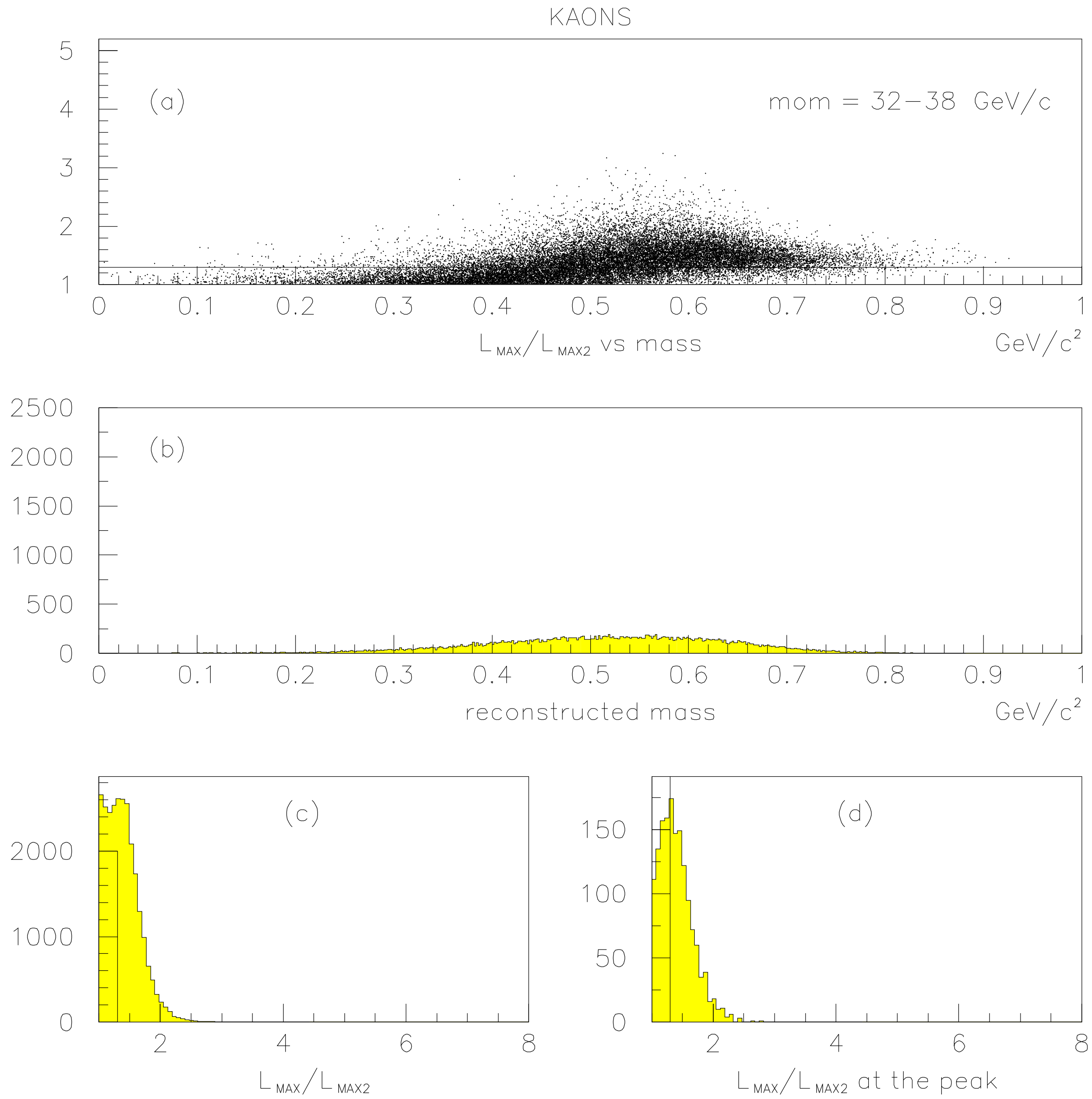

Figure 14: As fig. 10, for the momentum range indicated.

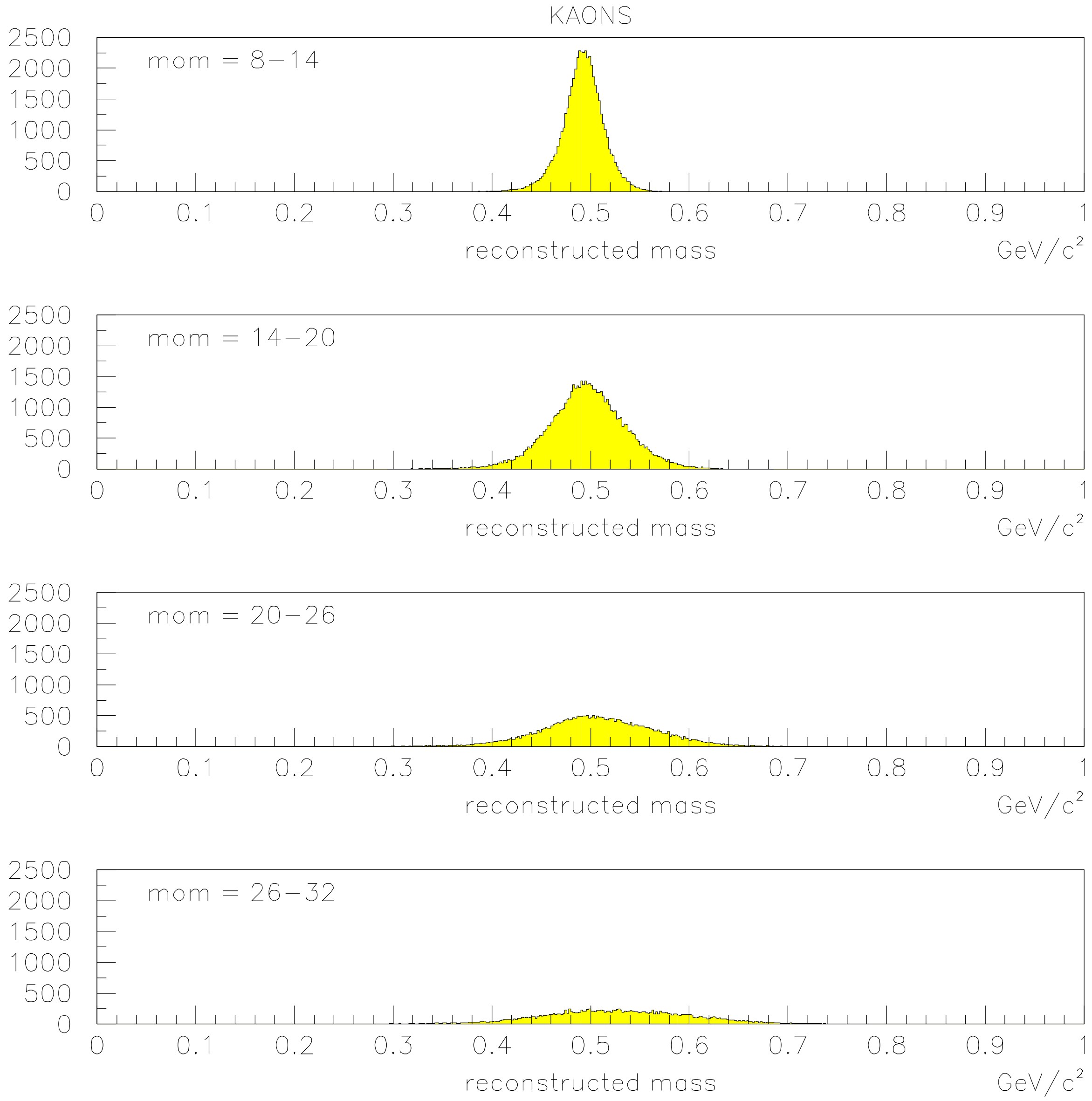

Figure 15: The figure shows, for the momentum bins indicated and for identified kaons, the mass distributions after Cut1 and Cut2.

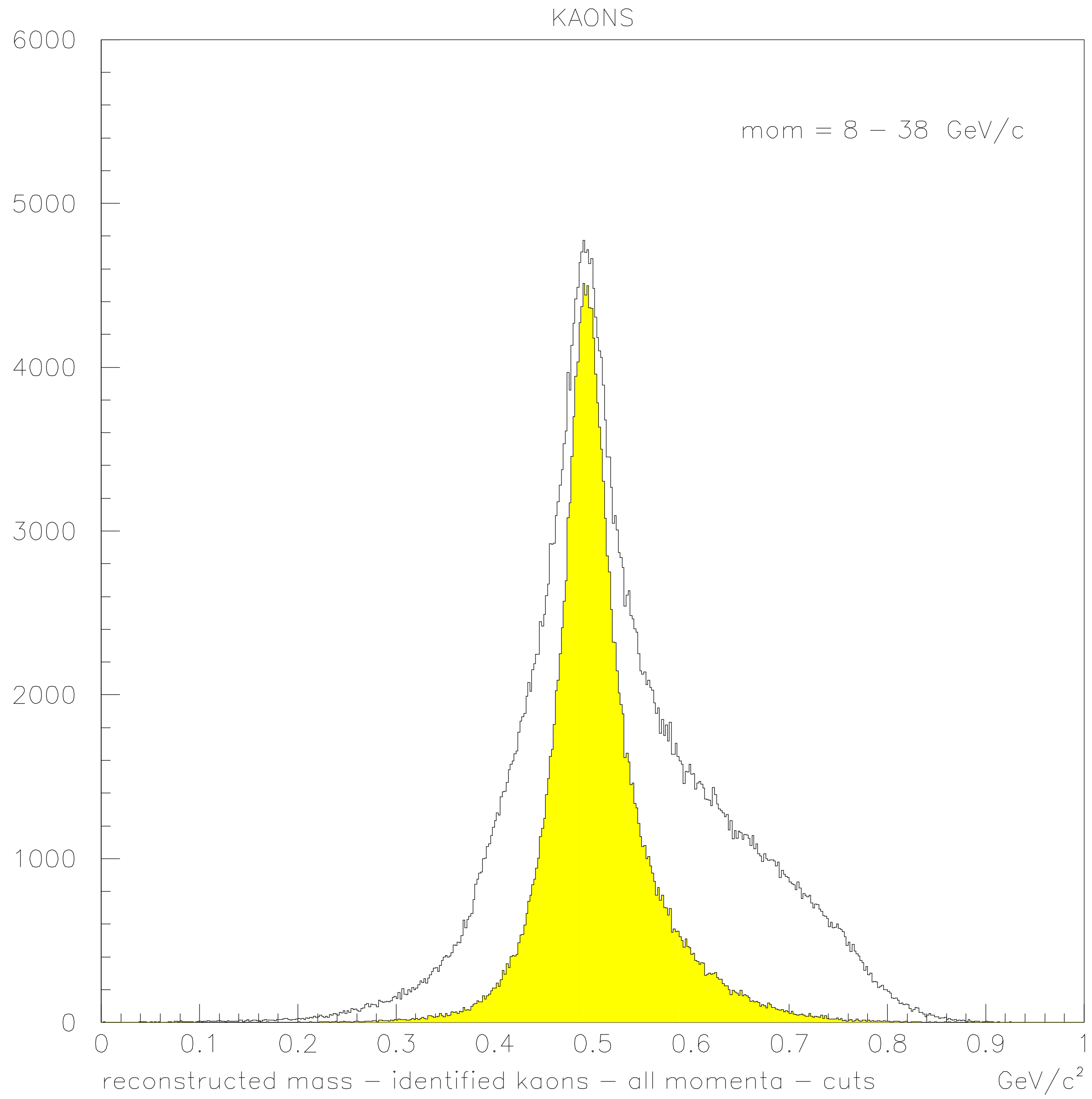

Figure 16: The figure shows, for the momentum range indicated, in color, the reconstructed mass spectrum of identified kaons, after the cuts in L_K/L_{MAX2} (Cut1 and Cut2), as discussed in the text, together with the uncut mass distribution of fig. 3.

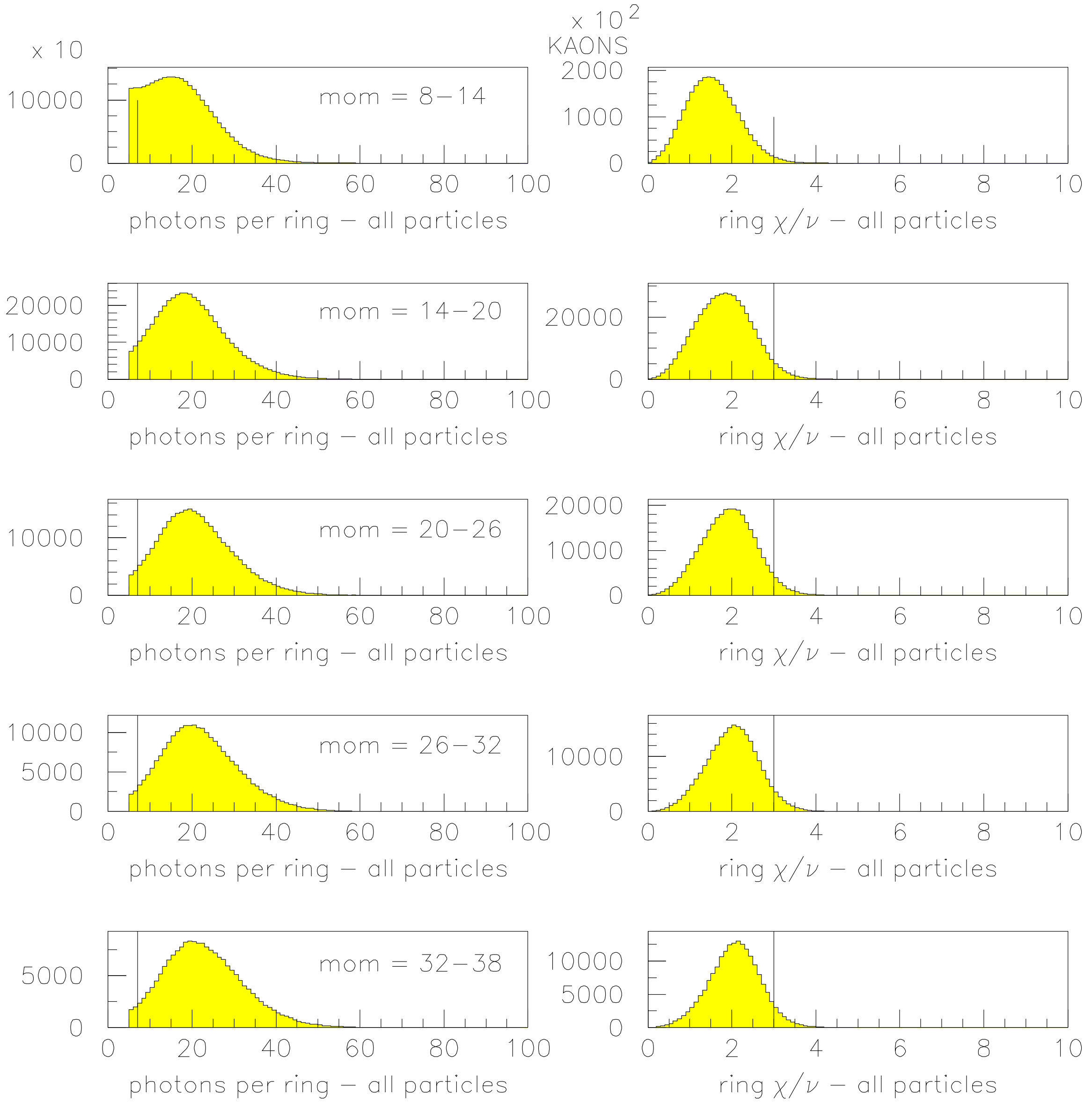

Figure 17: The figure shows, at the left, for the momentum bin indicated and for identified kaons, the number of *photons* per *ring* (already cut at 5), with a suggested cut at 7. At the right, the reduced *ring* χ^2, with a suggested cut at 3: note how the reduced *ring* χ^2 does not peak at ~ 1, as expected.

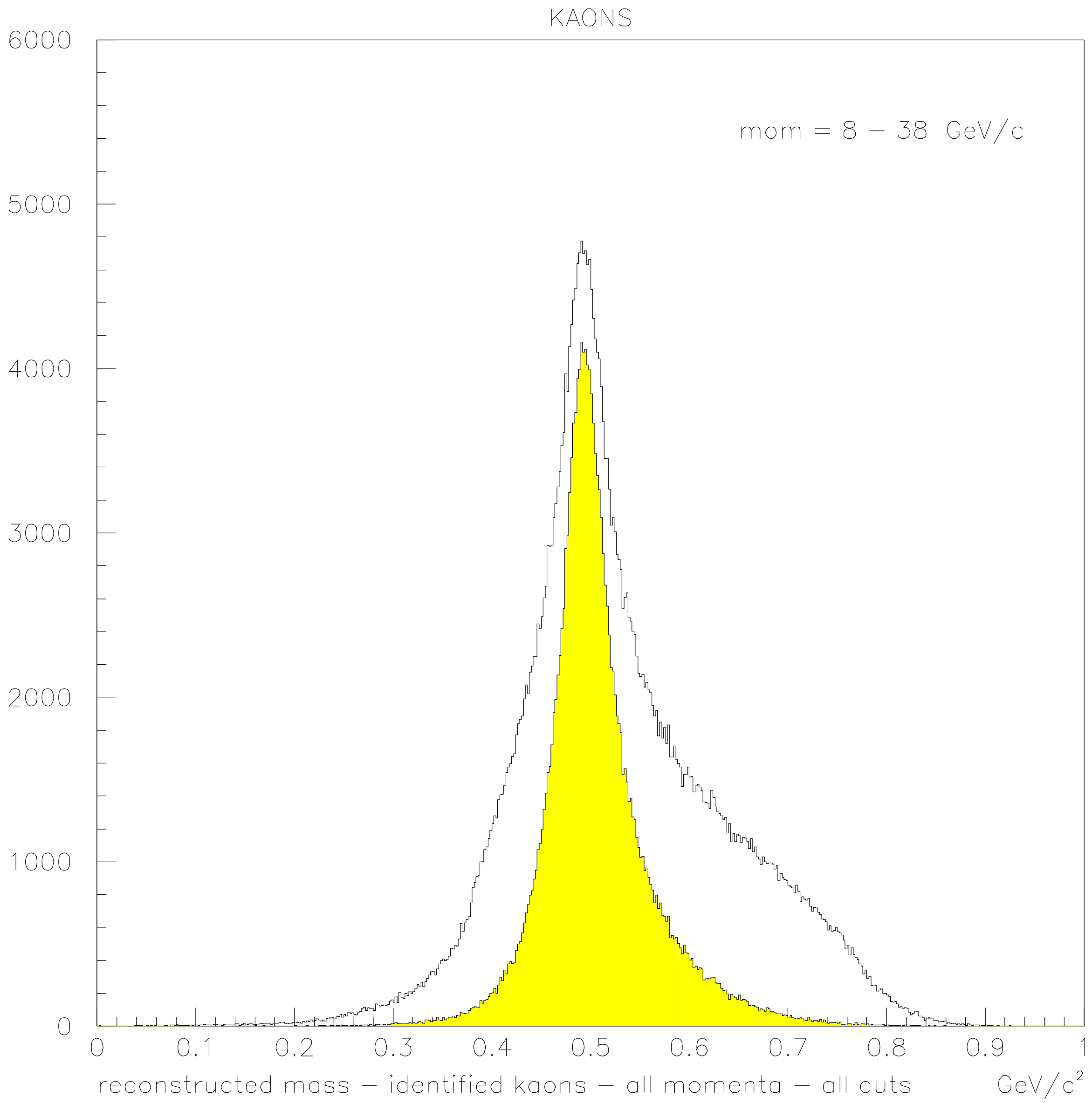

Figure 18: The figure shows, for the momentum range indicated, in color, the reconstructed mass spectrum of identified kaons, after Cut1, Cut2 and cuts in $N_{photons}$ and *ring* χ^2, as discussed in the text, together with the uncut mass distribution of fig. 3.

tion with the normalization as in fig. 19 and, innermost, the 'signal' as the difference of the second distribution minus the third one.
Fig. 21 (b) shows the same distributions as in (a), but after the Likelihood analysis (Cut1 and Cut2).
In both figures the vertical bars at K mass $\pm$ 75 MeV/c^2 represent the cuts to select the useful K mass region used in some analyses.

Fig. 22 is intended to show how the Likelihood analysis does not cut the K mass 'signal'; the mass distribution in color is the 'signal' as the difference defined in fig. 19 ($\sim$ 55k); the distribution as solid line is the 'signal' as the difference defined in fig. 21 (a) ($\sim$ 68k); the dotted line is the 'signal' as the difference defined in fig. 21 (b) ($\sim$ 62k).

Fig. 23 shows the raw reconstructed mass distributions, in the momentum range indicated, cut at K mass $\pm$ 75 MeV/c^2 (white), together with the K mass distribution after the Likelihood analysis (Cut1 only) (color).
The vertical bars are the 1σ or 2σ limits, around the K mass, where σ is the standard deviation (gaussian fit) of the K mass 'peak', after Cut1 only, extrapolated to 38 GeV/c; see fig. 24. As a further check, the function of fig. 24, of the form $\sigma = k\ p^2$, with k $\simeq$ 0.00011, is in good agreement with the expression $\sigma_{mK} = p^2/m_K\ n^2\ \cos\theta\ \sin\theta\ \sigma_\theta$, with $\theta \simeq 0.05$ rad and $\sigma_\theta \simeq 0.001$ rad, as from data analysis.
Assuming, as from fig. 22, a 'signal' of $\sim$ 55k, the 'signal' to background ratio for the K mass, after the Likelihood analysis (Cut1 only), is $\sim$ 44% and the reduction of the background from the white distribution of fig. 23 to the colored one, within the 2σ limits, is of the order of 40%.

5 Comparison with the MONTECARLO

Montecarlo events (produced with PYTHIA generator), with the 2002 apparatus geometry, have been used for comparison. The same likelihood analysis has been applied.

Fig. 26 is the analogous of fig. 3; it is worth to notice that the K mass peak is much narrower in the simulation (FWHM of $\sim$ 25 MeV/c with respect to $\sim$ 50 MeV/c for data); the better mass resolution is due to the better Cherenkov angle resolution because of an average number of photons per *ring* a factor of two larger and an apparatus fully aligned. The factor of two in the K mass width is in agreement with the difference in Cherenkov angle resolution (250 μrad for MC, 400 μrad for data), already discussed elsewhere.

Fig. 25 corresponds to fig. 4; note the larger range of the likelihood ratio in (a).
Fig. 27 corresponds to fig. 18, where all the cuts of the previous likelihood analysis have been applied.

6 Refractive index correction

The output stream from RICHONE, together with the Likelihood values, contains the derivatives of the Likelihood with respect to the refractive index (dL/dn), for the mass of π, K and p; they are intended to be used in correcting the Likelihood values for 'small'

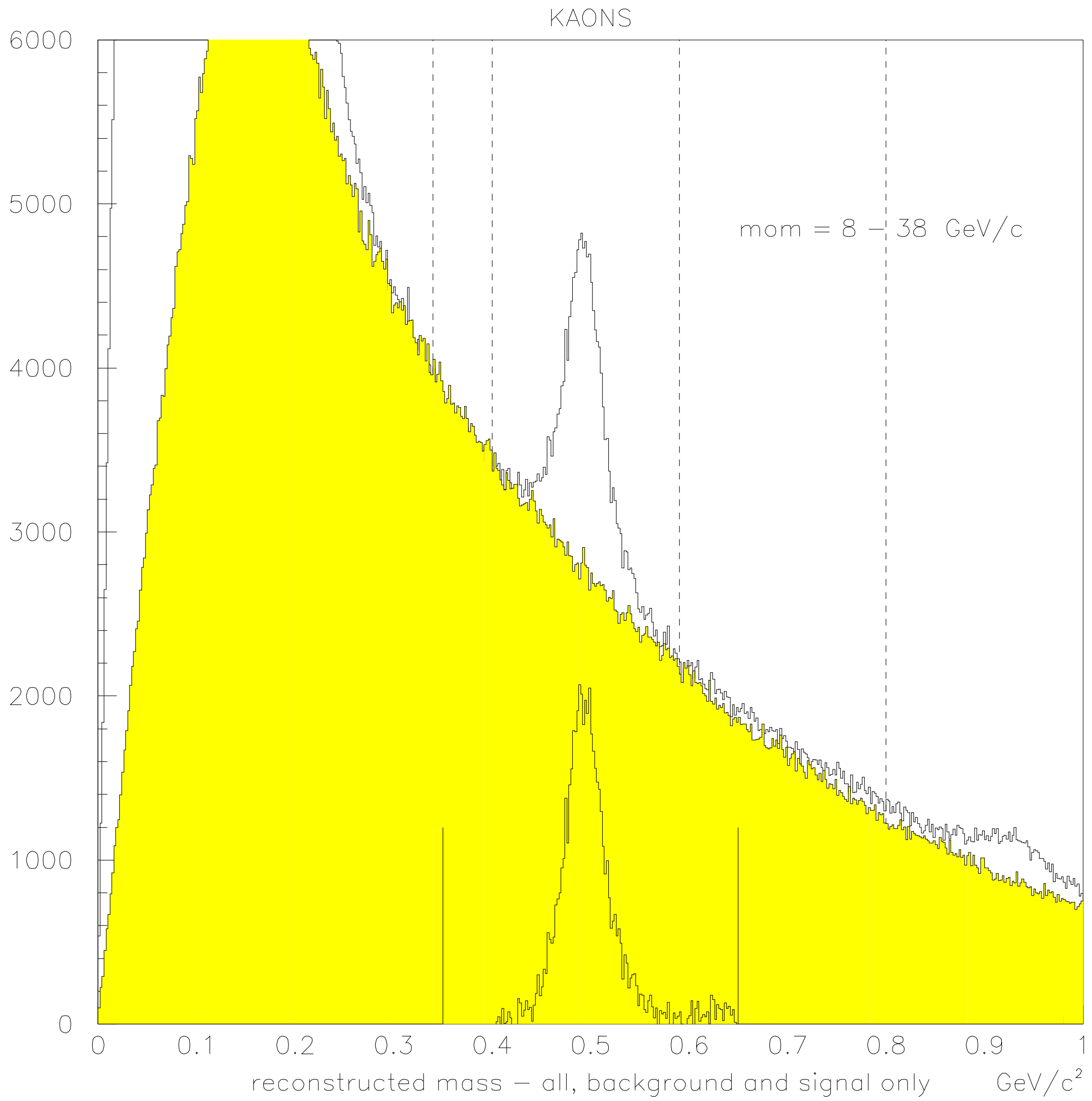

Figure 19: In white, the raw reconstructed mass distribution of fig. 3 (white) together with the 'scrambled' one, in color, normalized assuming equal the number of counts inside the two mass regions defined by the dotted lines. Also shown the difference of the two distributions, the result being the 'peak' around the K mass, between the two (solid) lines (K mass $\pm$ 150 MeV/c^2).

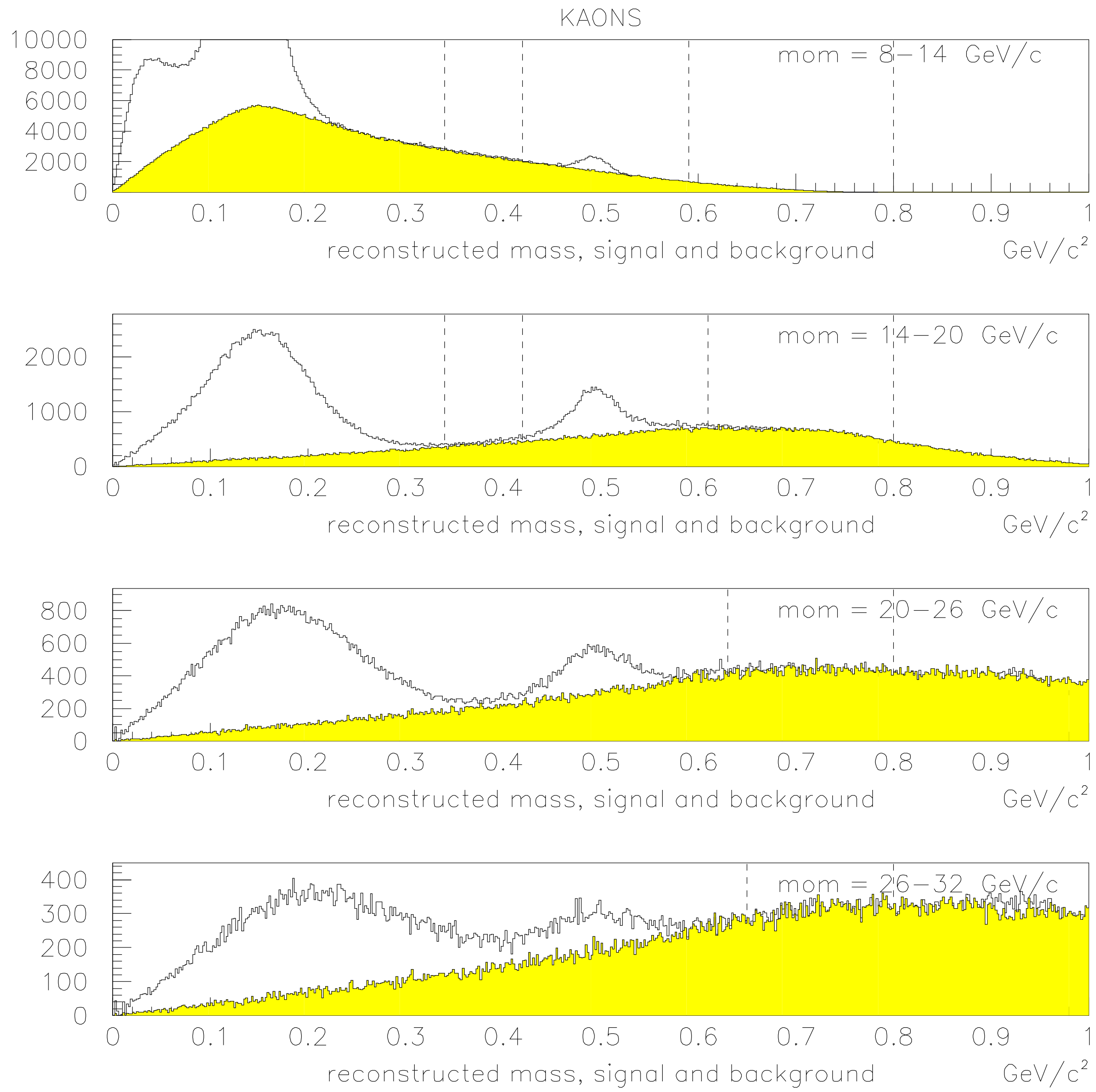

Figure 20: in white, the raw reconstructed mass distributions, for the momentum bins indicated, together with the 'scrambled' ones (color), normalized as explained in fig. 19, to show the dependence on the momentum.

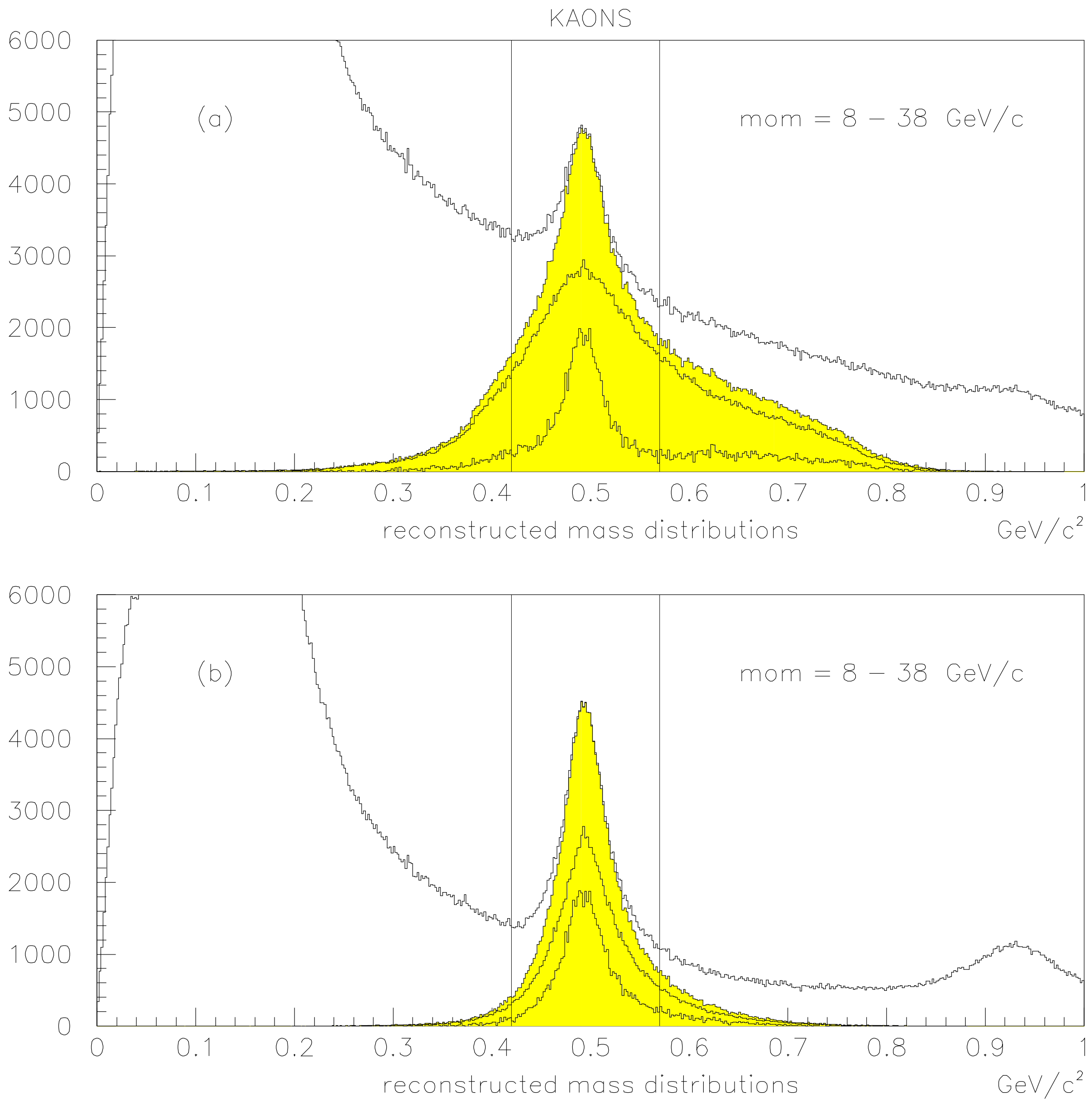

Figure 21: Fig. (a) shows, in white, the raw reconstructed mass distribution of fig. 3; in color, the identified K mass distribution of fig. 3; inside, the 'scrambled' mass distribution with the normalization as in fig. 19 and, innermost, the 'signal' as the difference of the second distribution minus the third one. (b) shows the same distributions as in (a), but after the Likelihood analysis (Cut1 and Cut2). In both figures the vertical bars at K mass $\pm$ 75 MeV/c^2 represent the cuts to select the useful K mass region used in some analyses.

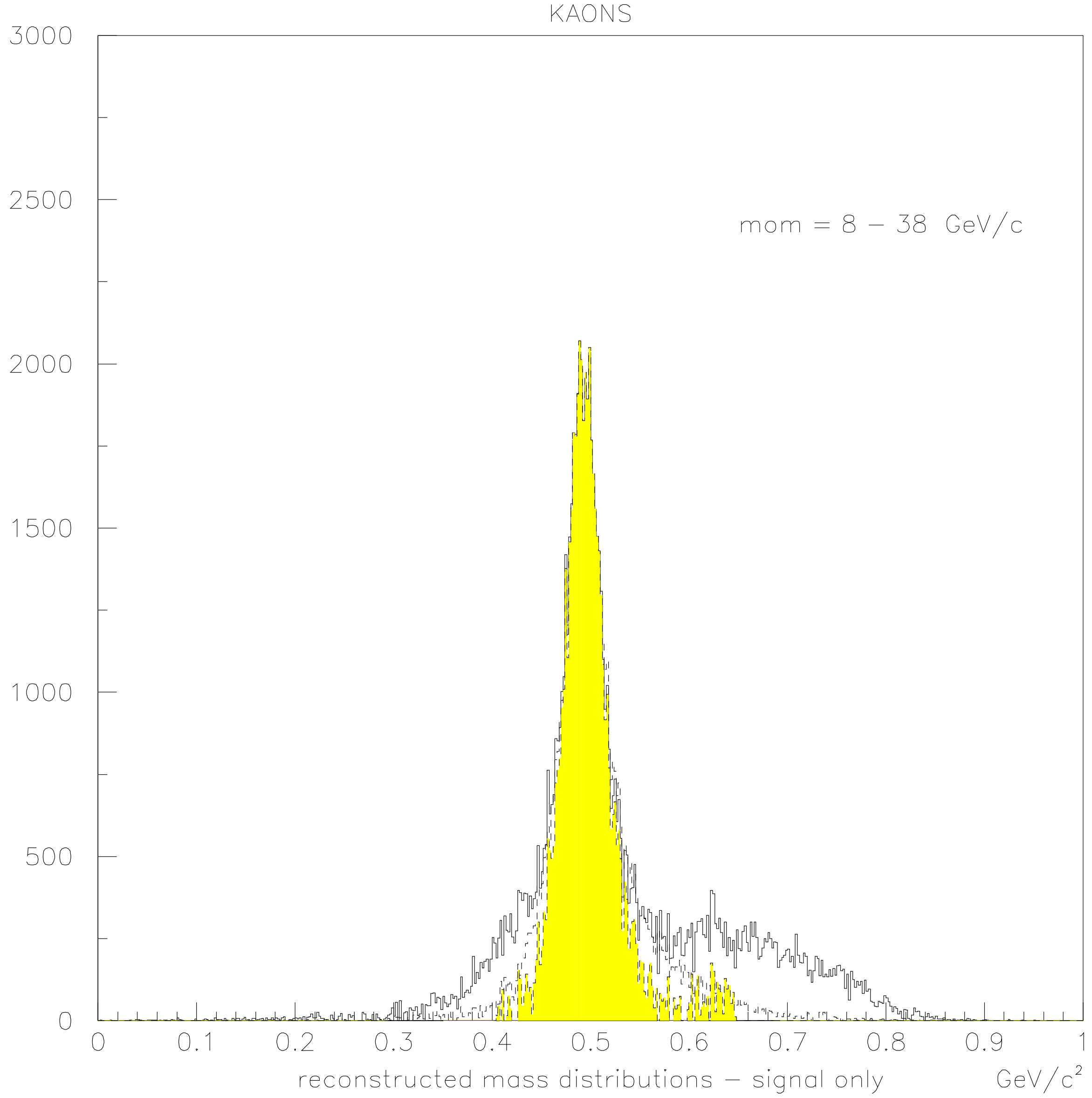

Figure 22: The figure is to show that the Likelihood analysis does not cut the K mass 'signal'; the mass distribution in color is the 'signal' as the difference defined in fig. 19; the distribution as solid line is the 'signal' as the difference defined in fig. 21 (a); the dotted line is the 'signal' as the difference defined in fig. 21 (b).

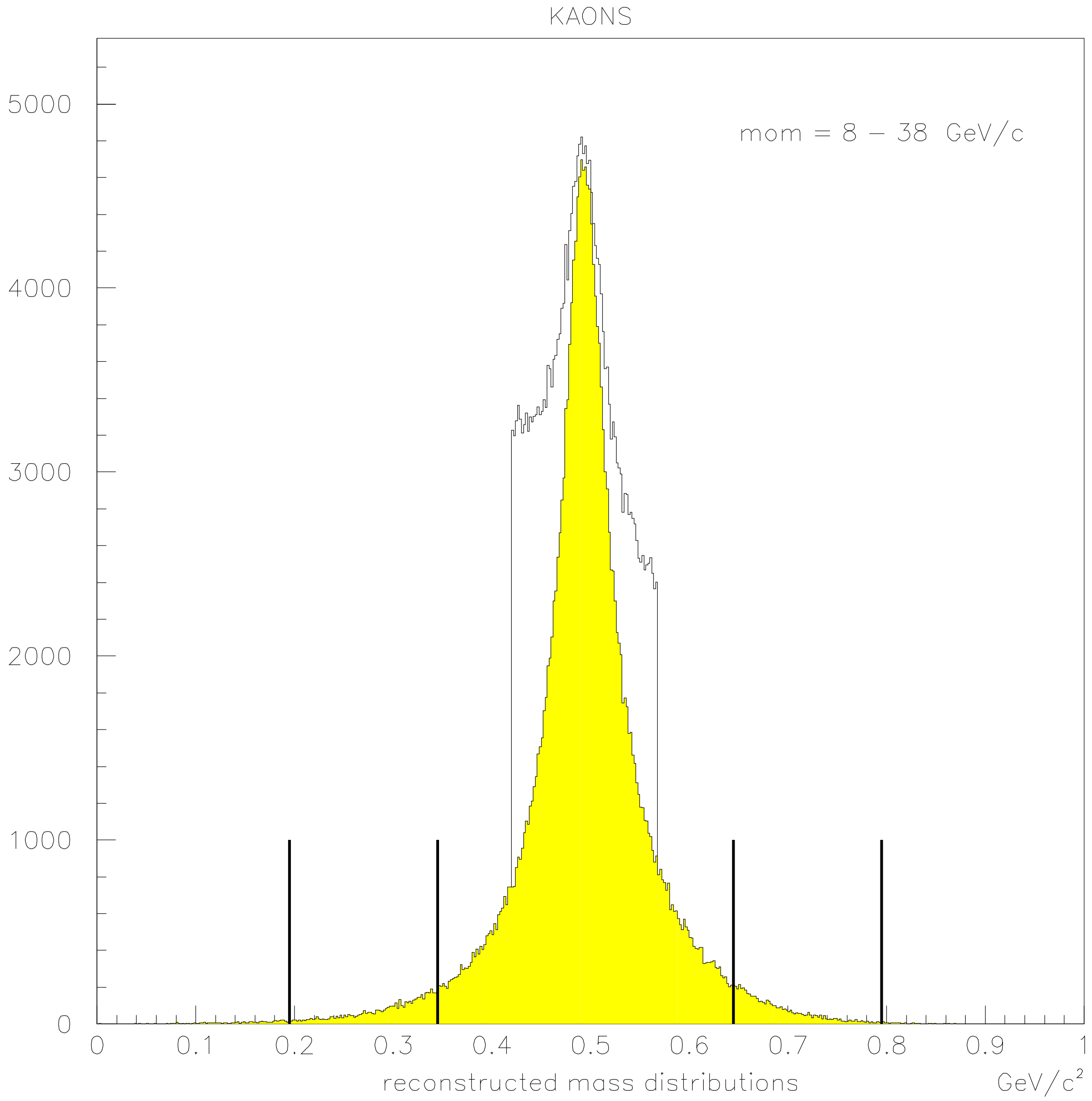

Figure 23: White: the raw reconstructed mass distributions, cut at K mass $\pm$ 75 MeV/c^2, together with the K mass distribution after the Likelihood analysis (Cut1 only) (color). Assuming a 'signal' as from fig. 22, the reduction of the background, going from the white distribution to the colored one, within the 2σ vertical bars ($\pm$ 200 MeV/c^2), is of the order of 40%.

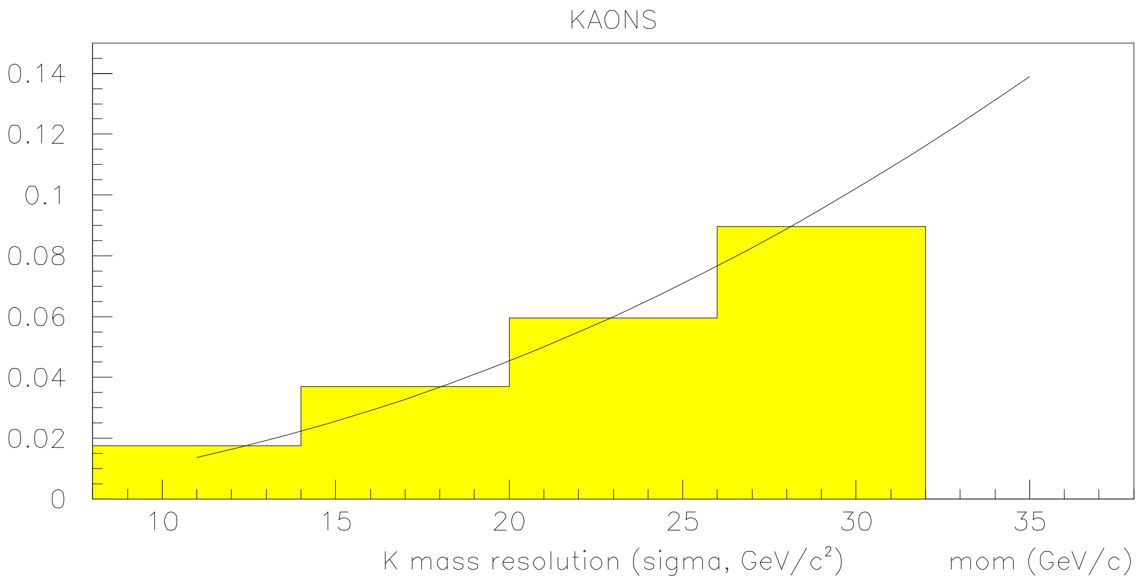

Figure 24: The value of σ from a fit to a gaussian shape of the K mass 'peak', after Cut1, as a function of the momentum and its extrapolation.

differences of the value of the refractive index used in the 'production' with respect to the correct value, as mentioned in section 2.

To give an example, fig. 28 (a) shows, solid line, for the momentum range indicated, the reconstructed mass distribution of identified kaons, when a 'wrong' refractive index, $\Delta(n-1) \simeq 13\%$, is used in processing; the dotted line, the same distribution after correction of the Likelihood, using its derivatives with respect to the refractive index. Figure (b) shows, for comparison, the mass distribution of identified kaons (solid), when the correct value of the refractive index is used in processing; dotted, the same as in (a). Note that no Likelihood analysis has been performed.
The figure 29, (a) and (b), show the same as in fig. 28; in this case a refractive index, with $\Delta(n-1) \simeq 6\%$), has been used in processing. Note as the correction is much more effective.
The result of the corrections seems rather satisfactory. A $\simeq 10\%$ error in $n-1$, seems to be an upper limit.

7 Conclusions

The use of Cut1 in the Likelihood analysis is worth to refine the K mass peak and to reduce the background under it, without cuts in the mass range, dangerous at higher momenta. On the contrary, Cut2 is doutfull, certainly overdoing and then not to be used without a better understanding. The cuts in number of photons per *ring* and *ring* χ^2 have to be well tuned because they cut part of the 'signal'.

The comparison with the MONTECARLO shows that the factor of two in the Cherenkov angle resolution, presently missing, is indeed responsible for the low PID resolution. Possibly the halo background simulation is not yet fully adequate.

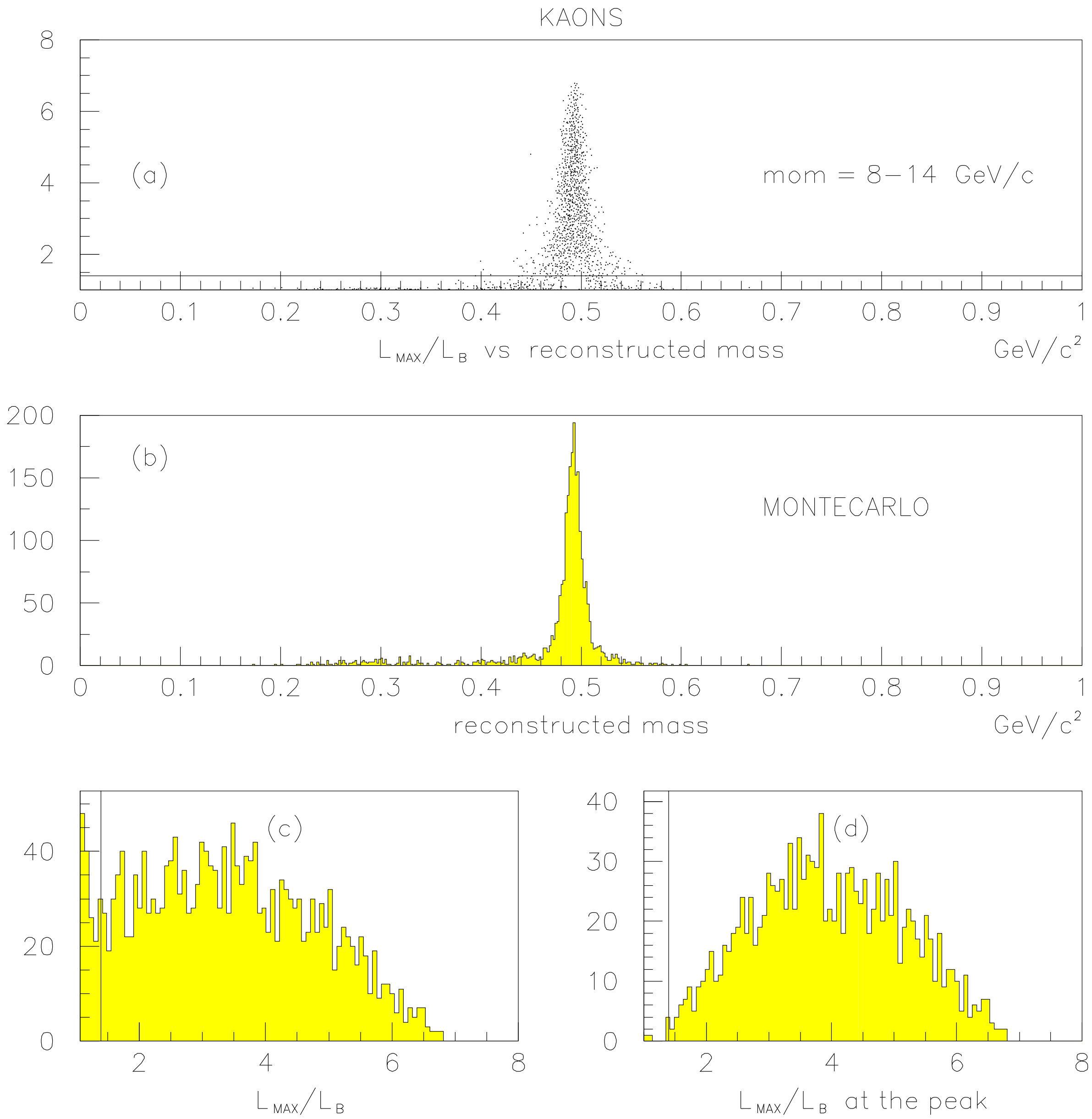

Figure 25: The figure corresponds for MONTECARLO to fig. 4; note the larger range of the likelihood ratio in (a).

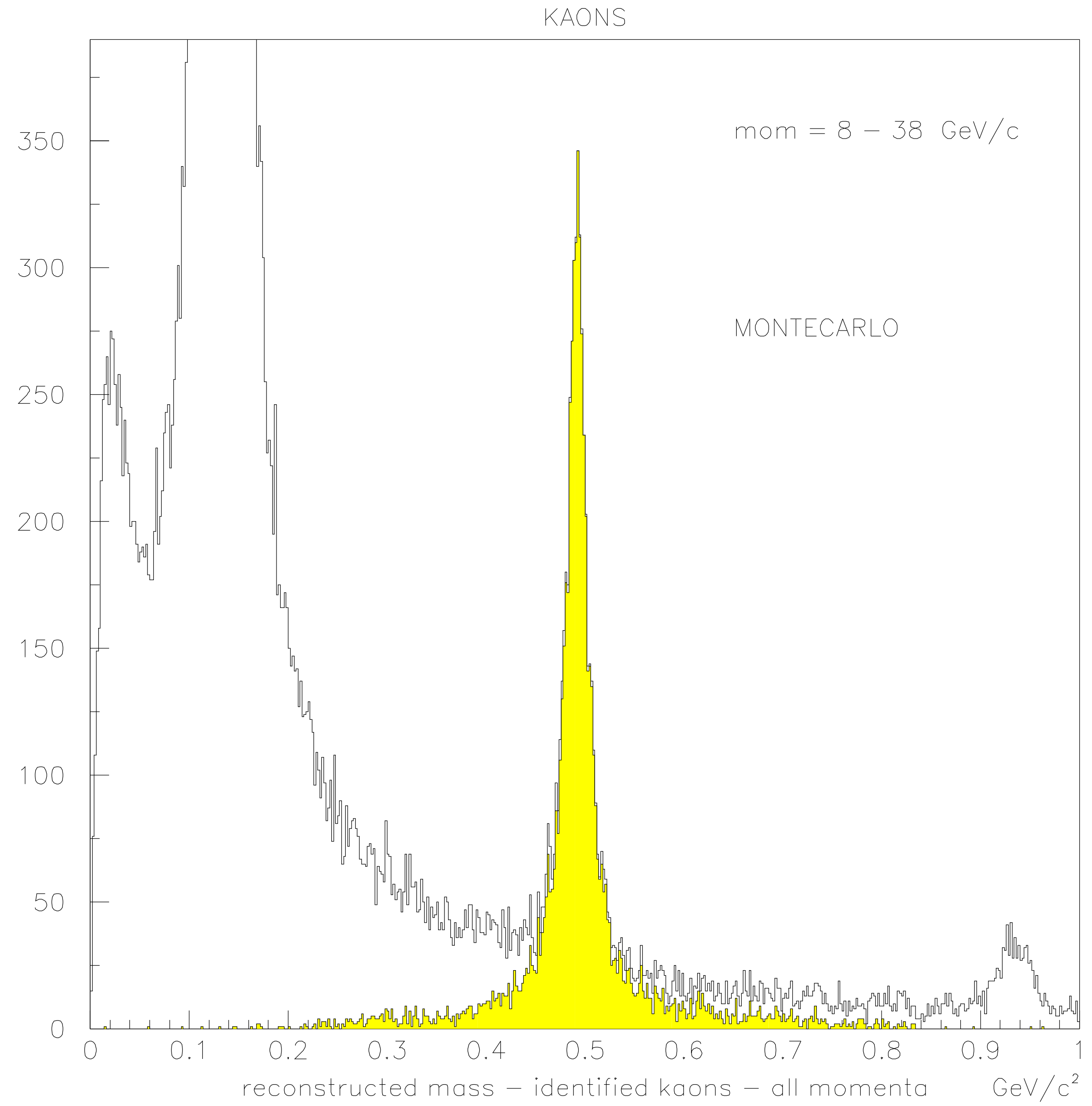

Figure 26: The figure is the analogous for MONTECARLO of fig. 3; it is worth to notice that the K mass peak is much narrower in the simulation (FWHM of $\sim$ 25 MeV/c with respect to $\sim$ 50 MeV/c for data).

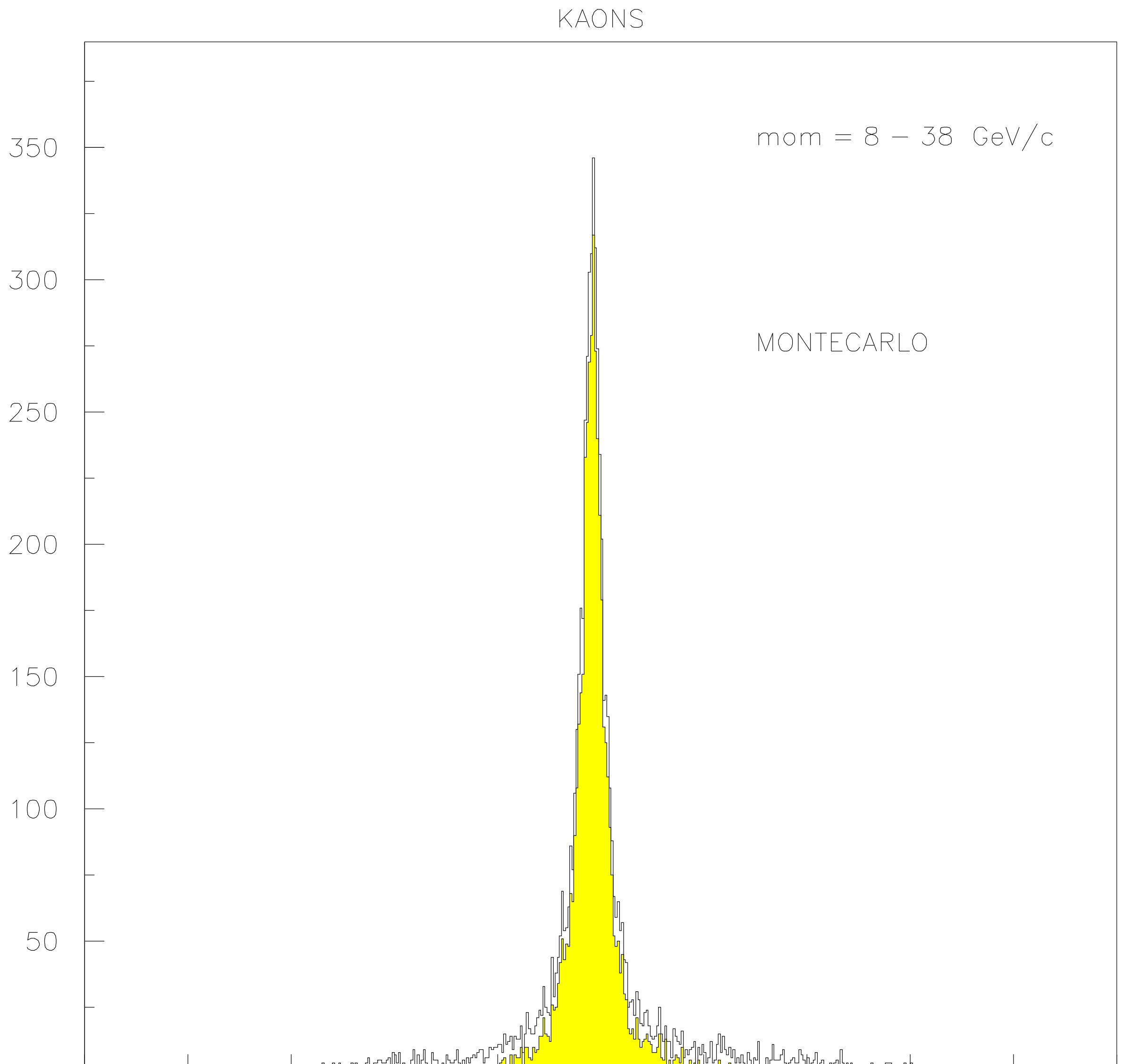

Figure 27: The figure corresponds for MONTECARLO to fig. 18, with all the cuts of the previous likelihood analysis applied.

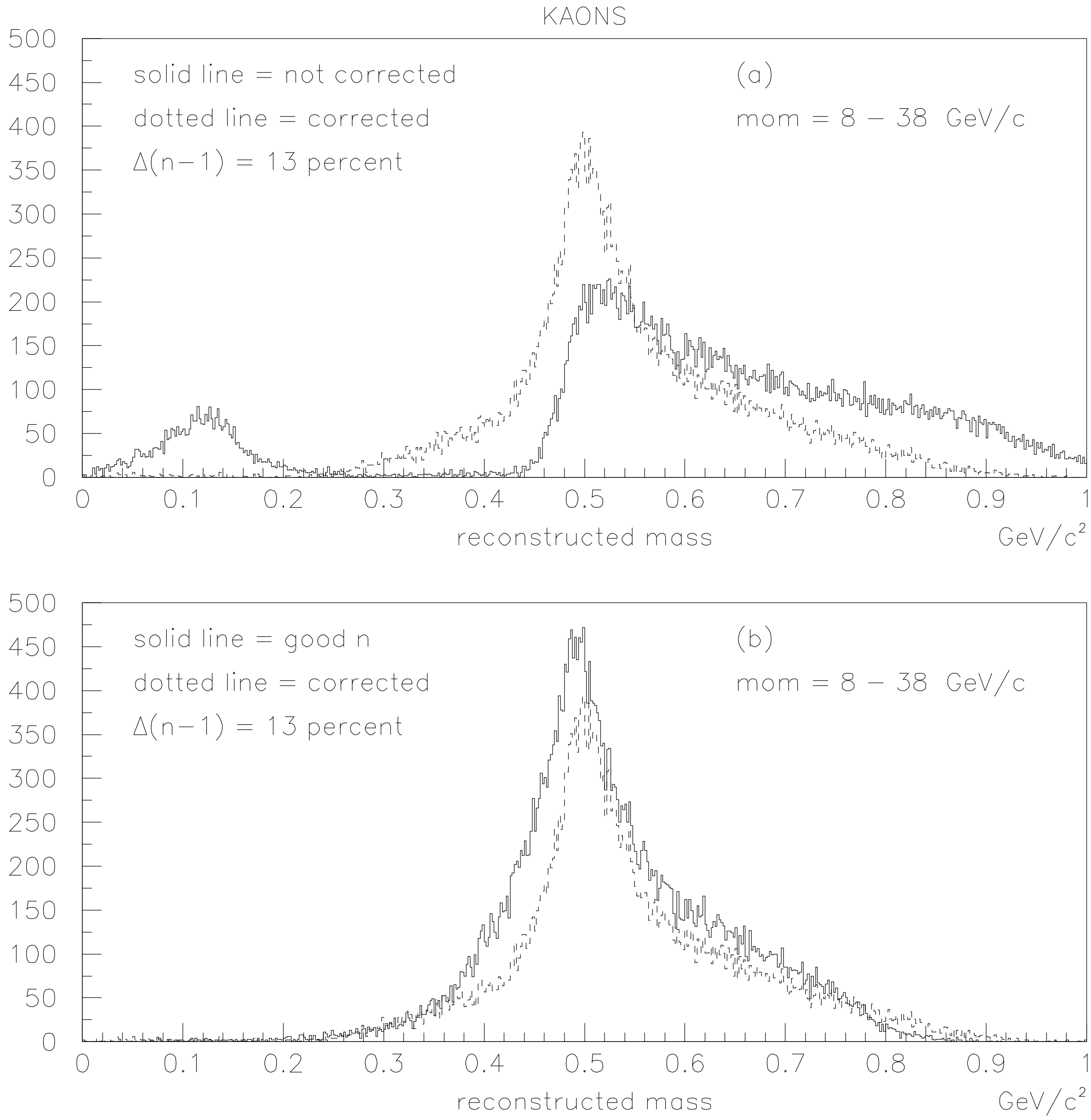

Figure 28: Figure (a) shows, for the momentum range indicated, the reconstructed mass distribution of identified kaons (solid line), when a wrong refractive index, $\Delta(n-1) \simeq 13\%$, is used in processing; as dotted line, the same distribution after correction of the Likelihood, using its derivatives with respect to the refractive index. Figure (b) shows, for comparison, the mass distribution of identified kaons (solid), when the correct value of the refractive index is used in processing; as dotted, the same as in (a). Note that no Likelihood analysis has been performed.

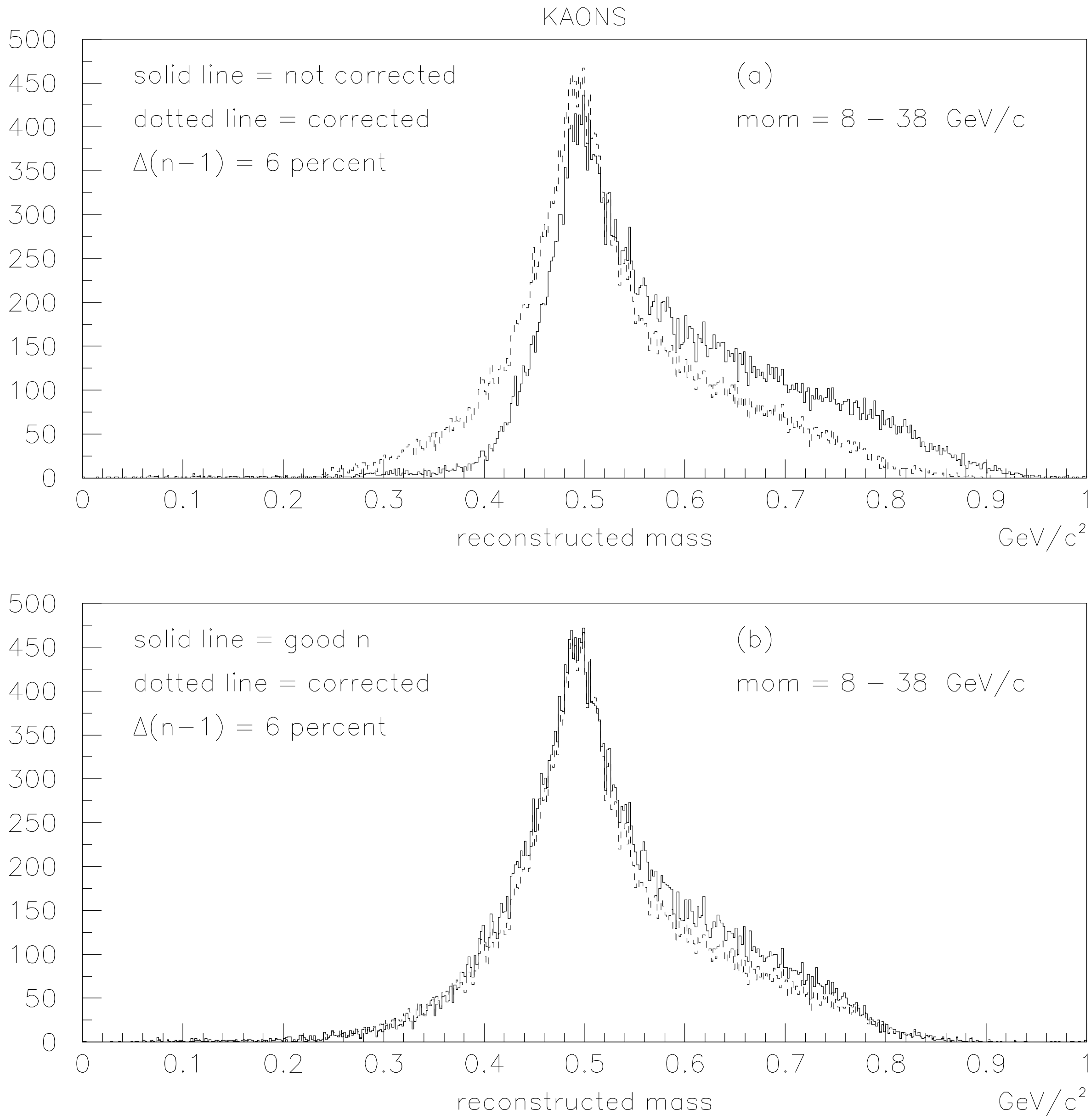

Figure 29: Figures (a) and (b) show the same as in fig. 28; in this case a less wrong refractive index, $\Delta(n-1) \simeq 6\%$), has been used in processing. Note as the correction is much more effective.

The correction of the value of the RICH refractive index, n, as used in the DST production, is possible by means of the Likelihood derivatives, provided n-1 is off by no more than $\sim$ 10-15%.