IPNO DR 02-015

Single muons recombinations as a determination of the combinatorial background in muon pair measurements and estimation of the associated error

4 june 2002

D. Jouan [1], S.Constantinescu [2], S.Dita [2], C.Alexa [2], S.Langlois [1], M.P.Comets [1],Y.Le Bornec [1], M.MacCormick [1], N.Willis [1]

[1] **Institut de Physique Nucléaire d'Orsay, IN2P3/CNRS et Université Paris sud, Orsay, France**
[2] **Institute of Atomic Physics (IFA), Bucharest, Romania**

Abstract

Any pair measurement is accompanied by a combinatorial background of uncorrelated elements. This background increases with the average number of single elements per event.

In ultrarelativistic heavy ion collisions the subtraction of this background is crucial for the study of the muon pair production. After a short description of the principles of an estimate of this background, one of these methods, based on the recombinations of single muons extracted from the measured muon pairs, is extensively studied and the error associated to these recombinations is determined both from analytical calculations and from studies of simulated and experimental samples. The optimization of the number of combinations is also considered.

1 Introduction

In ultrarelativistic heavy ions collisions, extreme conditions of temperature and density are reached and a new phase of matter could be accessed: the plasma of quarks and gluons [1]. It is a challenging experimental task to isolate in the final products of the expansion system significant traces of this transient state.

One of the most promising way is the measurements of muon pairs, which give access to the study of onia production. Thanks to their high masses, their rarity and their short live time, the onia could be the specific probes of the deconfined nature of the medium [2]. An anomalous suppression of the J/ψ has been recently observed by NA50 [3] at the CERN-SPS with 158 GeV/u Pb beams. Beside the resonances, the dimuon mass spectrum [4] displays a continuum due to various sources which can also present interest as a reference or as another potential signal of the plasma formation [5].

Experimentally the true muon pair production is naturally associated to a background of random combinations, which at CERN-SPS energies are mainly due to uncorrelated decays of π and K mesons into muons.

This combinatorial background, increasing with the square of the multiplicity of the produced mesons, increases dramatically with the energy deposited in the collision, and is thus an important experimental problem for dimuon measurements in heavy ion collisions.

Several methods have been proposed in the past years in the NA38 [6] experiment. This work deals mainly with the non trivial estimation of the error associated with the "recombinatorial method", but the framework which have been developed through the years (see [7, 9]) in the NA38 collaboration is also described.

Experiments involved in pair measurements have also to deal with this combinatorial problem [10, 11], but each specific property of the production process can introduce some differences in the background estimation and as a consequence in the associated error. For instance in contrast to pairs of identical particles, the dimuon contains two different particles μ^+ and μ^-, and thus pairs of identical muons $\mu^+\mu^+$ and $\mu^-\mu^-$ which are produced only by combinatorial processes, will provide both a natural normalization and samples of uncorrelated particles. This is also true for the measurement of opposite-sign pion pairs, for which recombinatorial method has also been used in the past [12]. But in contrast with dimuons, for which the region of high mass presents a special interest due to the dimuon production by Drell-Yan hard processes, for opposite-sign pion pairs the high mass part of the spectra can serve as a reference for normalization since no significant correlated production is expected in that region.

2 Estimations of the combinatorial background

The measured muon pairs have several origins and one can try to classify them according to the level of their correlations.

They can be

a) true muon pairs - the muons are produced simultaneously through the decay of a virtual photon

b) true-like muon pairs - they originate from the uncorrelated decay of a correlated heavy quark pair

c) fake muon pairs - from accidental combinations of muons coming from uncorrelated processes.

If the muon multiplicity per event M_μ obeys a Poisson law with average $< M_\mu >$, the yield of measured combinatorial pairs of all signs is simply $< M_\mu >^2 /2$. It is noteworthy that M_μ, can contain also muons from other collisions, when the fortuitous coincidences between collisions become non negligible.

For a given event, the number of different pairs of positive mesons that can be made with N_+ positive mesons is simply

$$N_{++} = N_+(N_+ - 1)/2 \tag{1}$$

and the combinations with the N_- negative mesons cause

$$N_{+-} = N_+ N_- \tag{2}$$

opposite-sign pairs.

From these simple relations applying to a single event, we need to deduce the connections between *averaged* or *differential* numbers of like-sign and opposite-sign pairs.

The above relations remains valid for averaged quantities, but already for the averaged global quantities the way to deduce $< N_{+-} >$, $< N_{++} >$ and $< N_{--} >$ is not straightforward in the general case since

$$< N_{++} >= (< N_+ >^2 + Var(N_+) - < N_+ >)/2$$

and

$$< N_{+-} >=< N_+ >< N_- > + cov(N_+, N_-)$$

This can be simplified in the case of a Poissonian distribution for N_+ and N_- and assuming complete uncorrelation between them, since one have then Var$(N_+)- < N_+ >= 0$ and cov$(N_+, N_-) = 0$ and one can write the relation:

$$< N_{+-} >= 2\sqrt{< N_{++} >< N_{--} >} \tag{3}$$

For small backgrounds, this formula can be approximated by
$< N_{+-} >=< N_{++} > + < N_{--} >$.

We will see in the following that applying the formula (3) to differential distributions assumes a factorization in integrals, and we will then sometimes refer to this method of determining the background as the *"integrated" method.*

Another simple case happens when $< N_- >$ (or $< N_+ >$) is large, in such a way that $< N_- > / < N_- >^2$ can be neglected, whereas $< N_- >$ is proportional to $< N_+ >$. In that case one can again write the relation (3), but the relation (1) does not necessarily hold, depending on the ratio $< N > /Var(N)$. This situation is close to the physical situation of meson production where positively and negatively charged particles are fundamentally linked by total charge conservation.

In [5], a study based on simulations with Venus generator evaluates the impact of the small multiplicities $< N_- >$ and $< N_+ >$ on the relation (3), for p-A (A=Al, Cu, Ag, W), S-U and Pb-Pb collisions in the NA38/NA50 experiments at CERN-SPS. A factor R has been introduced and thus instead of relation (3), the following relation has been used:

$$< N_{+-} >= 2R\sqrt{< N_{++} >< N_{--} >} \qquad (2')$$

for S-U and Pb-Pb interactions the factor R introduced in the relation ($2'$) was found [5] to have values showing a clear departure from R=1 only for peripheral collisions.

The relation (3), eventually corrected by R, allows to deduce the total number of combinatorial opposite-sign pairs from the total numbers of positive and negative like-sign pairs and is the very root of the background subtraction method in muon pairs studies.

It is of the uttermost importance to perform also for differential quantities - as the dimuon mass distribution - a similar combinatorial background determination. But in contrast to what happens for total average numbers, the factorization of the quantities depending on the single muons is generally not possible in the integrals expressing the numbers of pairs, preventing to get a direct relationship between opposite and like-sign pairs numbers.

Indeed, the numbers of muon pairs at a given mass is:

$$dN_{+-}/dM_0 = N^{\mu}_{+-} \int f^+(x^+)f^-(x^-)\delta(M_0 - M(x^+, x^-))dx^+dx^- \qquad (4)$$

$$dN_{++}/dM_0 = N^{\mu}_{++} \int (f^+(x^1)f^+(x^2)\delta(M_0 - M(x^1, x^2))dx^1dx^2$$

$$dN_{--}/dM_0 = N^{\mu}_{--} \int (f^-(x^1)f^-(x^2)\delta(M_0 - M(x^1, x^2))dx^1dx^2$$

where $f^+(x)$ and $f^-(x)$ are the probability distributions for the positive and negative *accepted* muons with respect to the 3-momentum x. Except when f^+ and f^- are proportional, the relation between the above three quantities is not straightforward. Indeed it is generally not possible to factorize the formula (4) with two independent integrals of f^+ and f^-, due to the δ function, and the relation (3) is thus not generally valid for differential quantities.

In the experiment NA38 the shapes of positive and negative single muon distributions are very different, which mainly comes from the existing differences in acceptances (see Fig. 1). This effect, due to the magnetic field, has been reduced by imposing a selection on data ("image cut") to equalize the acceptances of the like-sign and opposite-sign pairs of muons. This allows in the NA38 analysis, on the expense of a strong reduction of the acceptance for low mass and low transverse momentum dimuons, to use the following formula between the differential quantities:

$$\frac{dN_{+-}}{dM_0} = 2\sqrt{\frac{dN_{++}}{dM_0} \times \frac{dN_{--}}{dM_0}} \qquad (5)$$

From the previous formulas one can see that another way to estimate the background (4) is to use directly the $f(x)$ distributions. If these distributions are known one can calculate the integral, for instance by Monte Carlo method. A similar procedure can be used with the experimental single muon samples instead of distributions, but then, given their finiteness, it is insignificant to try to build new +- pairs when *all possible couples* have been already done once.

The two different methods - from pairs (formula (5) and using single muons (formula (4)) - derive strictly from the same formula (4), and are then *generally subject to the same limitations concerning their validity.* Only the additional integration needed to establish the formula (5) causes an additional restriction on this method. Determination of the background using random combinations of the single muon distributions has then the advantage of a larger domain of validity, avoiding to restrict the acceptance by cut, and it also leads to a better precision thanks

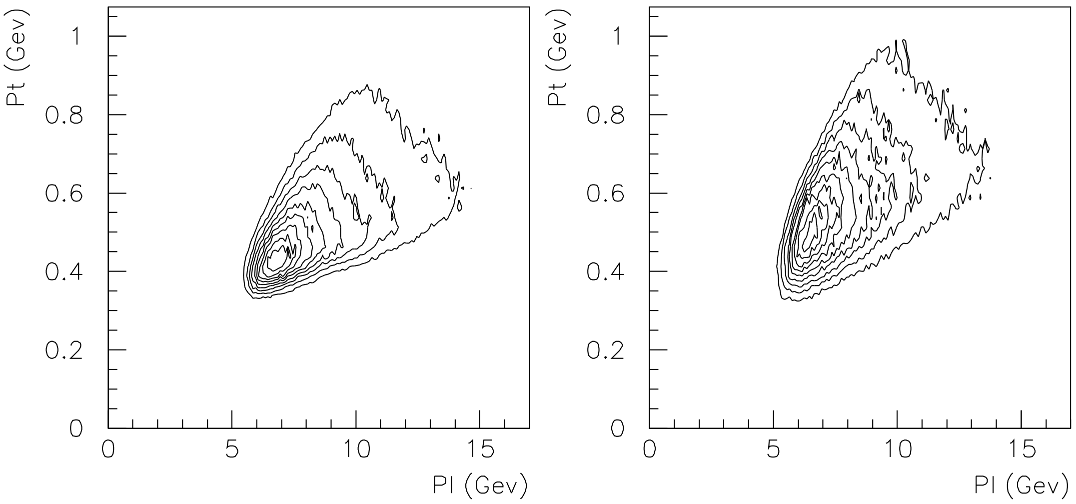

Figure 1: *Experimental distribution of longitudinal and transverse momenta of single muons extracted from like-sign pairs measured, for one direction of the current (positive under NA38 conventions). Left: from positive pairs, right: from negative pairs*

to a much larger number of different pairs, but to the expense of using several times the same muons. The estimation of the associated error is nevertheless not obvious, and this is the main drawback of the method, together with a longer computing time. These are the two questions that we will address in the following.

3 Combinatorial method

As seen in the previous section, similarly to the combinatorial process itself one can randomly combine muons of both signs in order to build the background and this is indeed strictly equivalent to the real combinatorial background, provided the single muon kinematic distributions are available and that the detector and analysis rejections are reproducible. The relation between N_{++}, N_{--} and N_{+-} *total* numbers is not subject to some possible restrictions that could apply to the same differential quantities. Then, in order to normalize the recombined spectrum to the opposite-sign dimuon sample, it is possible to use the relation (3), whatever the positive and negative single muon distributions could be.

The determination of the single muon distributions is then a central point of this method. Since the like-sign pairs originate from random combinations of muons, one can extract from them the single muon experimental distribution. The information contained in these samples has been not completely used in the original direct pair spectrum. Breaking these N pairs, getting all the possible N^2 combinations gives a better estimation of the ideal spectrum.

Practically one stores the single muon momentum for separated sub-classes of the data, each sub-class corresponding to physical or experimental parameters such as deposited energy, magnetic field, position of target, which are the relevant parameters for the analysis or are possibly linked to a change in the shape of the single muon distributions.

In order to get the true single muon distributions, it is also required that there would be *no selection on pair quantities* such as transverse momentum, mass or rapidity of the pair.

Starting from these single muon samples, the basic method combines the positive and negative muons making all the pairs allowed by the (single muon) trigger, and the resulting spectra are normalized thanks to the total quantities N_{++} and N_{--}.

The Fig. 2 shows that the background determined by this combinatorial method and the one determined by the formula (5) are in agreement. The "image cut" has been applied in order to use the formula (5).

In order to verify both the result of the calculations and the assumption of non correlation made for the like-sign pair formation, another comparison has been made between the like-sign experimental mass spectrum and the one obtained by recombinations of the corresponding single muons, normalized by their integrals.

No restrictions on the acceptances have been imposed (no use of "image cut").

Fig. 3 shows that the recombined spectra reproduce the visible differences in shapes due to the different acceptances (see Fig. 1). There is a very good agreement between the original and recombined spectra in the 0-2 GeV/c^2 region where the statistics is meaningful.

4 Study of the error

The main subject and the purpose of this work is the estimate of the error. This is non trivial in the case of a multiple use of the same muons, as have been done in the combinatorial method. In the following, we will aim at estimating the error associated to the number of pairs made in this non independent way, and also look for possible reductions of the number of this recombined pairs in order to reduce the computing time.

This will trigger phenomenological studies using Monte Carlo method as well as analytical calculations. The latter are based on ideal (gaussian) distributions, but also the former uses simple distributions or simplified experimental-like distributions. These simplified distributions are useful to get rid of statistical uncertainties and systematic fluctuations in the experimental sample.

5 Direct estimate of the error

In this section and the following one , we determine the variance on the content of each bin of distributions made by combination of singles element coming from different samples of pairs.

These studies are made for different situations of increasing complexity, from the addition of variables distributed according to 1 or 2 dimensional Gaussian, to dimuon mass spectra obtained from different experimental single muons samples.

5.1 Gaussian distributions

In order to characterize the error for a very simple case, we consider N values x_i and y_i obtained randomly from two Gaussian distributions. The number N is not fixed but has a Gaussian distribution with a root mean square $\sqrt{N}$. Then we consider the distribution obtained by simply adding x_i and y_j. The resulting distribution is also a Gaussian.

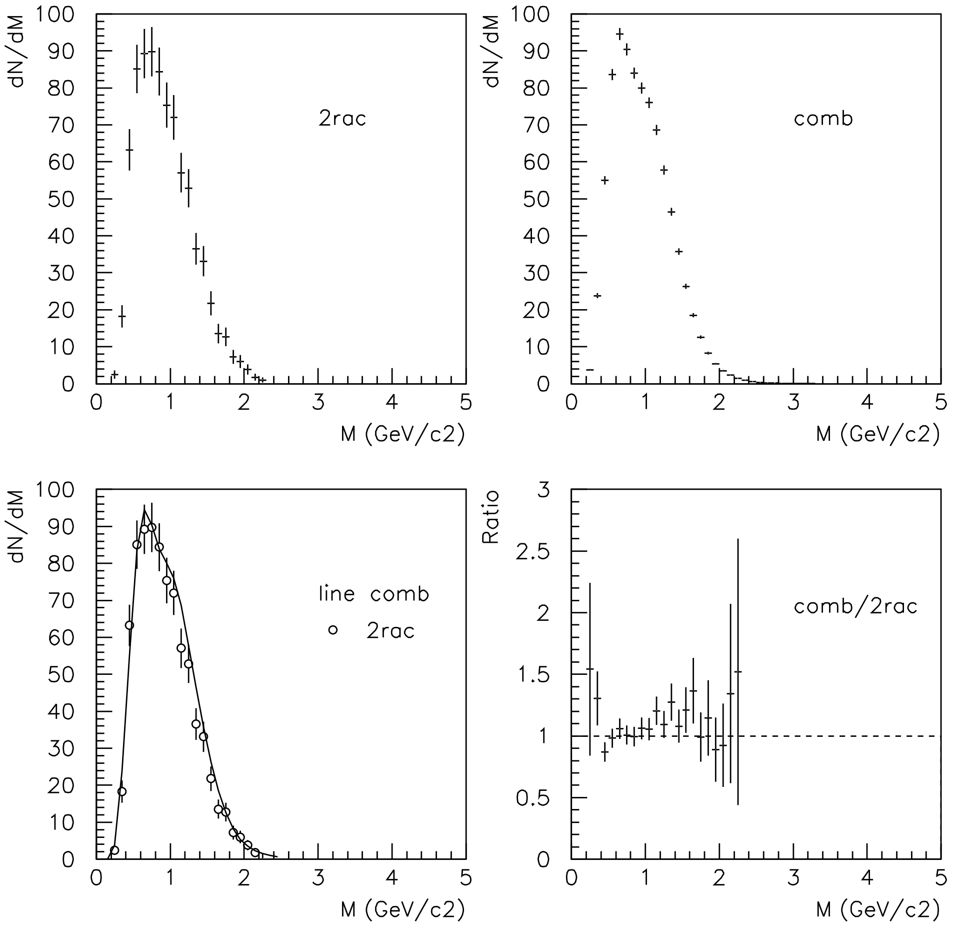

Figure 2: *Background estimated by integrated formula (2rac) and by complete recombination (comb) and their ratio (with image cut)*

.

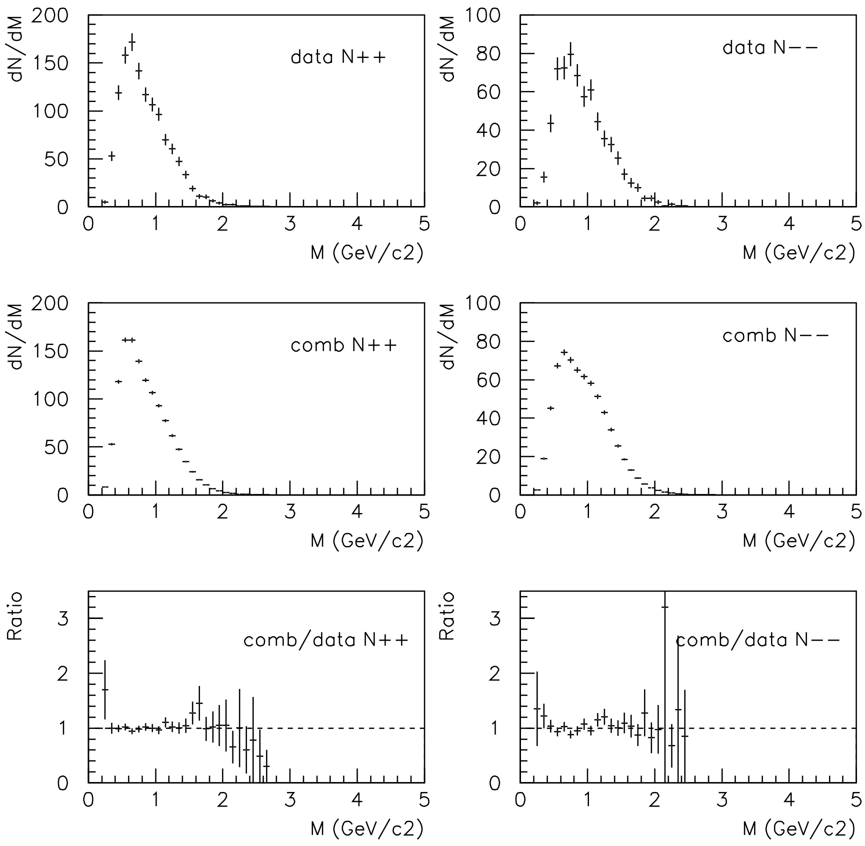

Figure 3: *Experimental mass spectra for like-sign pairs (and recombinations) with transverse momentum in the domain 0.6-1.3 GeV/c, and their ratio, for the highest* E_t *bin and target number 1 in the S-U system*

To study the error, we make several samples in which we consider all the possible pairs (x_i, y_j) (i=1,N ; j=1,N). This leads to the distribution of $x_i + y_j$. Each spectrum is normalized to the initial number of pairs, statistically distributed as previously described.

Fig. 4 displays two examples of the resulting root mean square σ_K of the number of counts in each bin, determined using the known true mean value of N_K. One can see that σ_K behaves as $(N_K)^\alpha$ and the same behavior has also been found for two-dimensional distributions [13]. It is noteworthy that this simple dependence is only obtained when the spread of the normalization has been taken into account, otherwise the distribution tends to saturate at the highest N_K values.

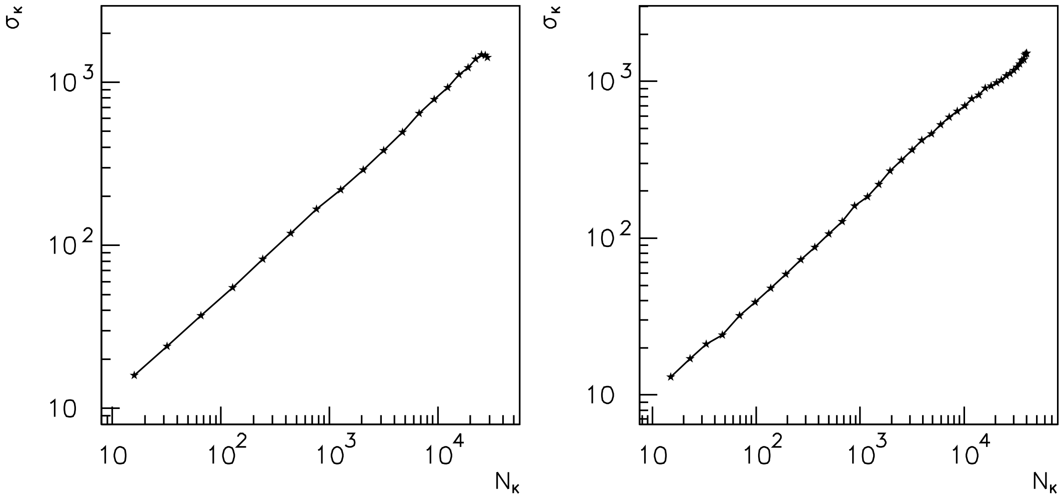

Figure 4: *Increase of error as a function of the number of counts in a bin of the recombined distribution, for different numbers of initial pairs. Left: 400 initial pairs, 50 bins ; right: 1000 initial pairs, 100 bins*

The α parameter if found to be about 0.64, and not to depend on the number of initial pairs (Fig. 4)or on the bin size (Fig. 5). We then need only to determine completely the value of one of the σ_K, the one corresponding for instance to the channel with the highest number of counts N_{max}, to be able to predict the values of all the σ_K. It is also appearing in picture 5(line) that σ_{max} and N_K^{max} keep a simple dependence when the bin size varies.

Fig. 6 displays the evolution of the ratio σ_{max}/N_K^{max} with respect to the number of initial pairs N, and for various numbers of bins. For the limiting case of a unique bin, the ratio σ_{max}/N_K^{max} is equal to $\sqrt{N}/N = 1/\sqrt{N}$. When the number of bins is increased, σ_{max}/N_K^{max} increases too, but only by 40% for the already pessimistic situation of 1000 bins for only 1000 initial pairs. An upper limit for σ_{max}/N_K^{max}, when the spectrum considered has reasonably populated bins, should then be $1.4/\sqrt{N}$

Finally we obtain a phenomenological parameterization of the error value, slightly overestimated, which takes automatically into account the bin size and the eventual normalisation, and uses explicitely the initial number of pairs:

$$\sigma_K = 1.4\left(\frac{N_{max}}{\sqrt{N}}\right)\left(\frac{N_K}{N_{max}}\right)^{0.64} \tag{6}$$

where N is total initial number of "true" pairs, N_K is the number of pairs in the bin K, either

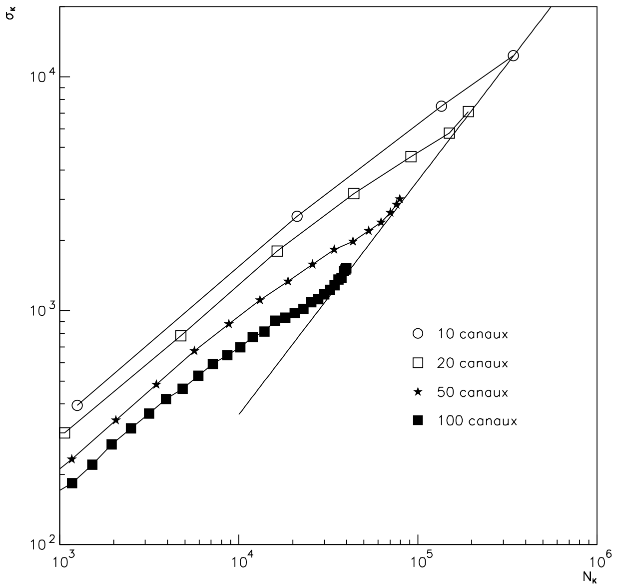

Figure 5: *Error as a function of the number of counts in the bin for different bin sizes (canaux=channels)*

renormalized combined pairs or "true" pairs which are close numbers by construction, and N_{max} is the maximal value of N_K.

As pointed out in [7] the errors for grouped channels are simply the sum of the corresponding errors, not the usual quadratic sum. This property is also verified in this formula, where grouping of channels correspond to a change of binning, that is to a proportial change of N_{max} and indeed of σ_K. This formula has also the correct behaviour, but overestimated by 40%, in the limiting case of a unique bin ($\sigma_K = \sqrt{N}$).

5.2 NA38 distributions

In order to determine the fluctuations for experimental NA38 S-U distributions, subsamples have been extracted in the experimental available data set. The considered data set is also restricted to a transverse energy domain and only one target. Two cases have been considered: 10 subsample corresponding to 800 opposite sign muon pairs, or 17 subsample corresponding to 400 opposite sign pairs. This selection on a fixed number of opposite sign pairs allows statistical fluctuations on the number of positive and negative like sign pairs, without the trouble of the normalisation to the number of collisions and the associated error. But it introduces nevertheless an additionnal fluctuation on the total number of pairs.

For each subsample the opposite sign background is estimated by recombining the muons originating from the like sign pairs, and normalizing the resulting distribution to $2\sqrt{N^{++}N^{--}}$. The root mean square of the number of count obtained in this different sample for the same channel are presented in Figure 7, as a function of the number of counts. A good agreement with the previous formula (6), established for Gaussian distributions, can be seen.

Higher statistic results have also be obtained using simple experimental-like distributions, which will be described in the prescaling section. They are also into a good agreement with the values plotted in the Fig. 7.

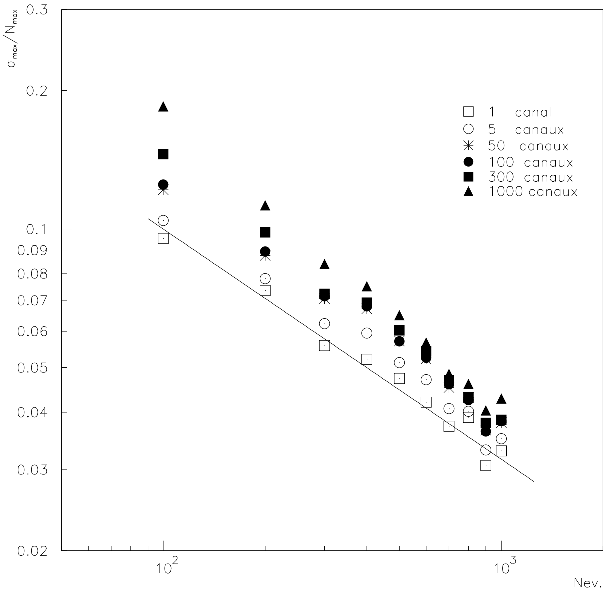

Figure 6: *Relative error at the maximal value of the distribution with respect to the number of initial pairs and for different bin sizes*

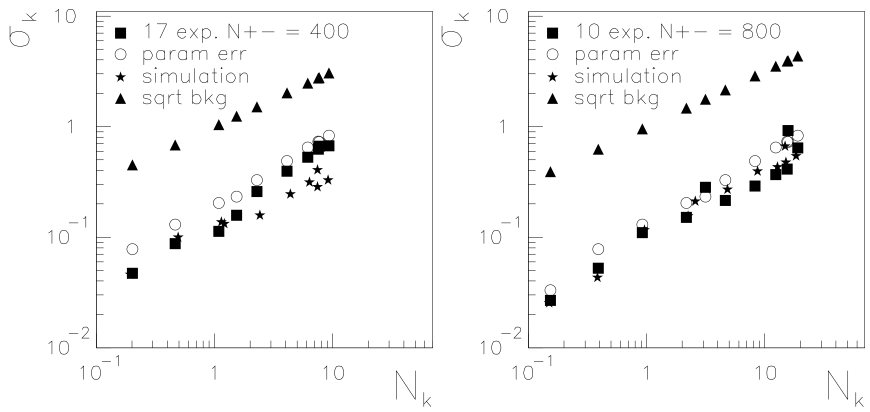

Figure 7: *Observed root mean squares of the number of counts Nk, when using the recombinatorial method (black square) on different subsamples, compared to the parameterization of the error (open points) and to the result of the recombinatorial method applied to a simplified experimental-like distribution (stars). The triangle shows the error associated with the usual "integral" method. Two statistics for the total number of opposite-sign pairs (400 and 800) have been considered.*

5.3 Analytical calculation of the error

Our goal is the estimation of the error on the number of combinatorial pairs p^I in a certain cell I defined by some relevant parameters as: mass, rapidity or transverse momentum for instance.

The distribution p^I is obtained by the recombination of the single muons extracted from like-sign pairs, normalized thanks to the numbers of these initial pairs. There are thus essentially two sources of errors: one related to the fluctuations on the number of combined pairs in one bin, which is not straightforward to estimate due to repeated use of the same muons, and the other one related to the statistical fluctuations on the general normalization. The latter cannot always be neglected. For example, in the limit case of a distribution having only one bin, the error would be unrealistically zero.

In this section we present the analytical calculations of the error made by [7, 8], restricted to an equal number of positive and negative muons for the sake of simplicity. Fluctuations on the normalization have been furthermore taken into account.

The calculation of the error is based on the notion of indicative function $U^I(x_i, y_j)$, which is equal to 1 when the value of the function $C(x_i, y_j)$ (mass, momentum,..) belongs to the interval I. x and y are the kinematic variable of the positive and negative muons.

the estimator $\bar{p}_I$ of p_I, is obtained making *all* the possible combinations in the available sample :
$\bar{p}_I = 1/N \sum_{i=1,j=1}^{N,N} U^I(x_i, y_j)$

The expectation value can be obtained integrating this estimator on the probability distributions $f_A(x)$ and $f_B(y)$ of positive and negative muons and including through the weight p_N the spread of the overall normalization N:

$$E(\bar{p}_I) = \sum_{N=1}^{\infty} p_N \, 1/N \sum_{i=1,j=1}^{N,N} \int U^I(x_i, y_j) f_A(x_1)dx_1 \cdots f_A(x_N)dx_N f_B(y_1)dy_1 \cdots f_B(y_N)dy_N \quad (7)$$

which can also be written

$$E(\bar{p}_I) =< N > \alpha \quad (8)$$

where

$$\alpha = \int U_I(x, y) f_A(x)dx f_B(y)dy \quad (9)$$

and $< N >= \sum_{N=1}^{\infty} N \; p_N$ is the average value of the total number of pairs N.

A similar calculation is used for the variance $Var(\bar{p}_I) = E(\bar{p}_I^2) - p_I^2$, with:

$$\begin{aligned} E(\bar{p}_I^2) = \; & \sum_{N=1}^{\infty} p_N \; \int 1/N^2 \sum_{i=1,j=1}^{N,N} U^I(x_i, y_j) \sum_{i'=1,j'=1}^{N,N} U^I(x_{i'}, y_{j'}) f_A(x_1)dx_1 \\ & \cdots f_A(x_N)dx_N f_B(y_1)dy_1 \cdots f_B(y_N)dy_N \end{aligned} \quad (10)$$

using $U^2 = U$, one can separate this sum in terms of different complexity corresponding to $i' = i$ or $j' = j$ or both. For $i = i'$ and $j = j'$ the expectation is α. For $i = i'$ one has a term $< N - 1 > \alpha_A$, with

$$\alpha_A = \int U^I(x, y) U^I(x, z) f_A(x)dx f_B(y)dy f_B(z)dz \quad (11)$$

$$\alpha_A = \int G_B^2(x) f_A(x) dx \tag{12}$$

where $G_B(x) = \int U^I(x,y) f_B(y) dy$.

There is also a similar term $< N-1 > \alpha_B$ for $j = j'$. the terms with $i \neq i'$ and $j \neq j'$ give $< N^2 - 2N + 1 > p_I^2/N^2$ finally one can write, similarly[1] to the result obtained by [7] in the case $\alpha_A = \alpha_B$:

$$E(\bar{p}_I^2) = \alpha + (\alpha_A + \alpha_B)(< N > -1) + \alpha^2 < (N-1)^2 > \tag{13}$$

and

$$Var(\bar{p}_I) = \alpha^2 Var(N) + (\alpha_A + \alpha_B - 2\alpha^2) < N > + \alpha^2 + \alpha - \alpha_A - \alpha_B \tag{14}$$

Assuming large N and $Var(N) =< N >$:

$$Var(\bar{p}_I) \simeq < N > (\alpha_A + \alpha_B - \alpha^2) \tag{15}$$

one can also consider the root mean square:

$$\frac{\sigma(\bar{p}_I)}{\sqrt{< N >}} \simeq \sqrt{\alpha_A + \alpha_B - \alpha^2} \tag{16}$$

It is possible to verify that if the bin I contains everything, $\alpha = \alpha_A = \alpha_B = 1$, and one effectively retrieves the usual statistical fluctuation $\sigma(\bar{p}_I) = \sqrt{< N >}$, or $\sqrt{Var(N)}$ if the distribution of N is not Poissonian.

The estimation of the error values requires the knowledge of the single muons distributions through α, α_A and α_B. In the next section we will apply the previous formulas to a simple analytic case and compare to the simple phenomenological formula 6 of the error obtained in the previous section.

5.4 Test with a Gaussian distribution

We consider the simple case of gaussian distributions f_A and f_B, with the same root mean square σ and means μ_A and μ_B respectively.

x and y are randomly extracted in these distribution, and their sum leads to a new distribution with average $\mu_S = \mu^A + \mu^B$

In the appendix, the detail of the calculation of the error is given, replacing f_A and f_B in the expressions [11] and U by Dirac distributions $U^I(x,y) = \delta(x + y - I)$, leading finally to:

$$\sigma(\bar{p}_I) \simeq \sqrt{\langle N \rangle} \sqrt{\alpha_A + \alpha_B - \alpha^2}$$

$$\sigma(\bar{p}_I) \simeq \sqrt{\langle N \rangle} \sqrt{\frac{1}{\sqrt{3\pi}} \frac{1}{\sigma_S} [f_S]^{\frac{4}{3}} - \frac{1}{4} [f_S]^2}$$

with

$$f_S = \frac{1}{\sqrt{2\pi}\sigma_S} \exp\left\{ -\frac{(\mu_S - I)^2}{2\sigma_S^2} \right\}$$

[1] here the fluctuation on the normalization introduces an additional term due to $< N^2 >=< N >^2 + Var(N)$

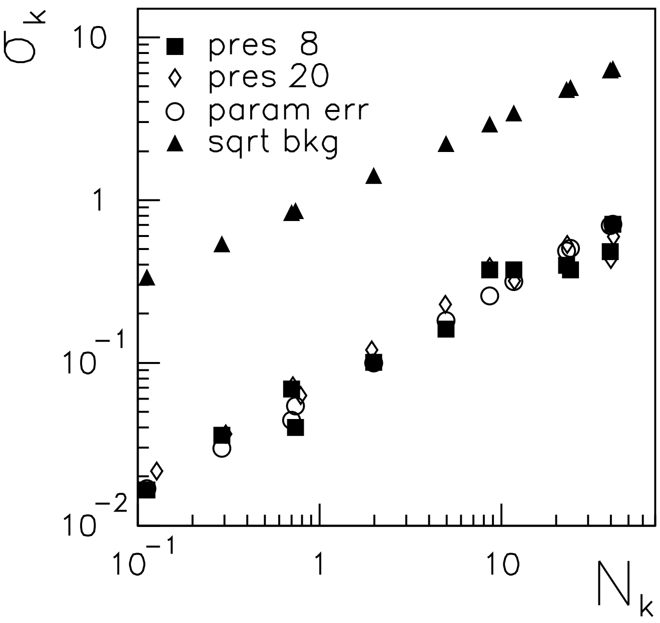

Figure 8: *The estimated error of the background used in our analysis (the average over the N_S evaluations given by sub-samples) obtained by combinatorial method as a function of the number of counts for different values of the prescaling factor N_S. Also shown are the errors corresponding to the integrated method (sqrt). The experimental data set is central S-U collisions, sub-target 1 and positive field in the magnet*

$\sigma(\bar{p}_I)$ is found mainly proportional to $f_S^{2/3}$, which is a similar behavior to $\sigma_K \propto \sqrt{N}(N_K/N_{max})^{0.64}$, the phenomenological formula (6) that has been established in the previous section ($N_K = p_I$).

6 Prescaling

The processing of a too large number of combinations could induce problems related to the requirements of an excessively high computing time and/or memory. Furthermore the generation of all possible combinations is probably not necessary since part of these pairs brings negligible additional information. It is then interesting, both from a practical and from a aesthetic point of view to try to precise the minimal fraction of combinations that will be meaningful.

We introduce a prescaling factor (N_S), which is the number of independent sub-samples extracted from the relevant (transverse energy, target) experimental like-sign sample. These sub-samples, containing each a number of N_{LS}/N_S like-sign muon pairs, are considered separately in the pair generation process.

The construction of these sub-samples is improved by taking one like-sign muon pair and by-passing the N_S-1 following ones. This procedure, not necessary for background calculation, is useful for comparing independent experimental samples which could otherwise be too sensitive to some possible changes occurring in the experimental conditions.

After processing all the possible combinations inside each sub-sample and introducing the normalization to the total data sample, N_S independent background distributions are obtained. They can be used for studying the fluctuations but they can also be summed to calculate the average background that we want to compare to the one obtained after complete recombinations in the maximal data sample.

Finally all the available like-sign muon pairs are involved into the analysis but the total number of generated combinations has been reduced by the factor N_S when the prescaling has been used.

Figure 8 presents the resulting fluctuations, for different prescaling factors and as a function of the number of counts N_K. In this example σ_K remains on the level of the error calculated by formula (6), even for a prescaling factor of 20.

In order to get rid of the statistical limitations imposed by the experimental single muon distributions, a Monte Carlo method using the experimental $(P_t - P_l)$ distribution (Fig. 1) of single positive (negative) muons belonging to N_{++} (N_{--}) pairs has been considered as input, together with a simplified NA10 azimuthal distribution: a 18 degree dead zone every 60 degrees.

One can see in Fig. 9 that these very simple distributions reproduce the details of the mass spectra of the like-sign pairs, although their shape is very different. These simplified NA38 single muons distributions are then valid tools to determine the fluctuations associated to the recombination method in the NA38 experiment.

For different values of the prescaling, left hand side of Figure 10 presents the fluctuations observed when the estimate is based on a sub-sample. The increase of σ_K^S with the prescaling, reaching the level of the precision obtained without recombination for $N_S = 40$, is consistent with the increase with $\sqrt{N_S}$ predicted by formula (6).

The use of all the available sub-samples increases the precision. In the right hand side of Fig. 10 the resulting σ_K remains on the same level even for the highest value of the prescaling considered. From formula (6) one indeed expect to get the same value when considering N_S sub-samples - which multiply the error by $\sqrt{N_S}$ - and making the average of the result - which divide the error by $\sqrt{N_S}$. This implies inversely that the error will not be sensitive to the prescaling as long as the formula (6) remains valid. The validity "limit" of the approximated formula (6) appears already in Fig. 6, for about one hundred pairs in the sub-sample and one hundred bins in the distribution. This gives the order of magnitude of the acceptable prescaling, leading for instance to subsamples of at least 100 pairs before mixing.

Fig. 11 presents a check performed for the S-U system showing that the differences induced by a prescaling factor equal to 150 are still below the expected error.

7 Effects on the experimental dimuon mass spectrum

Finally, in Fig. 12 the signal spectra obtained using the different methods for background estimation are compared. the formula 6 can be used to predict the effect of recombinations compared to the usual "integrated" formula. For the most populated bin this corresponds to $\sqrt{N_{max}/N}$, i.e. to a gain of about 5 if 4% of the N pairs are in this bin. This is not observed when looking at the two top pictures of figure 12, for two reasons:

- the combinatorial method is applied separately to the different sub-targets whose errors combine independently

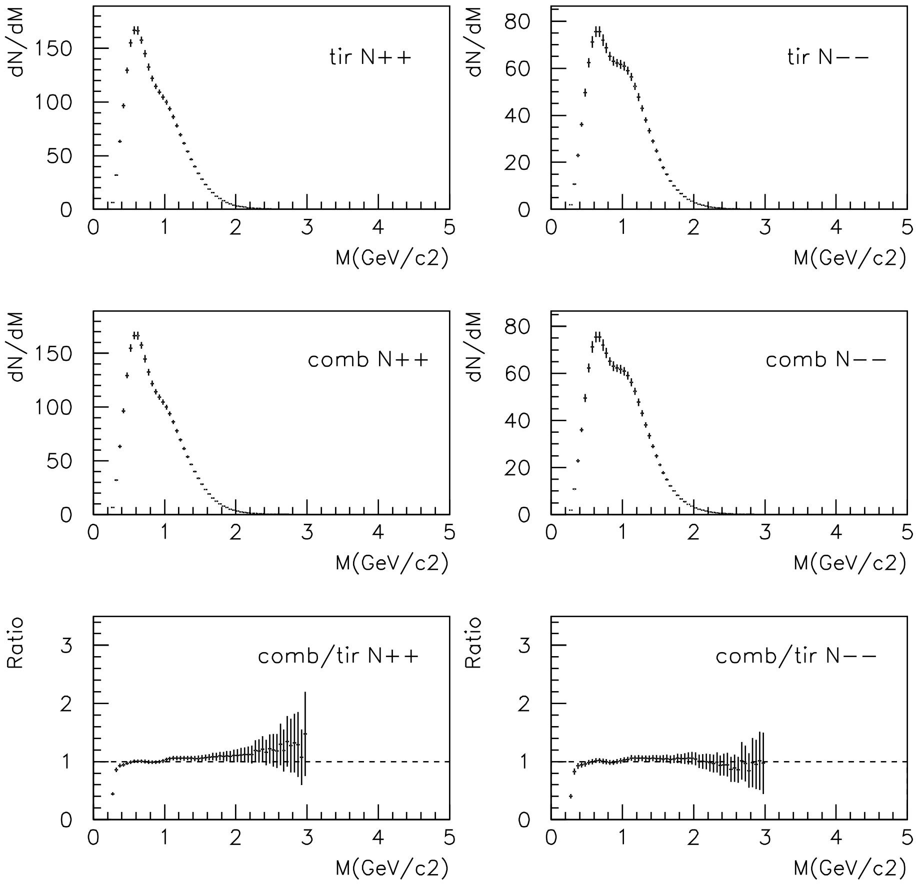

Figure 9: *Comparison of recombinations from simplified (tir) and from experimental (comb) distributions*

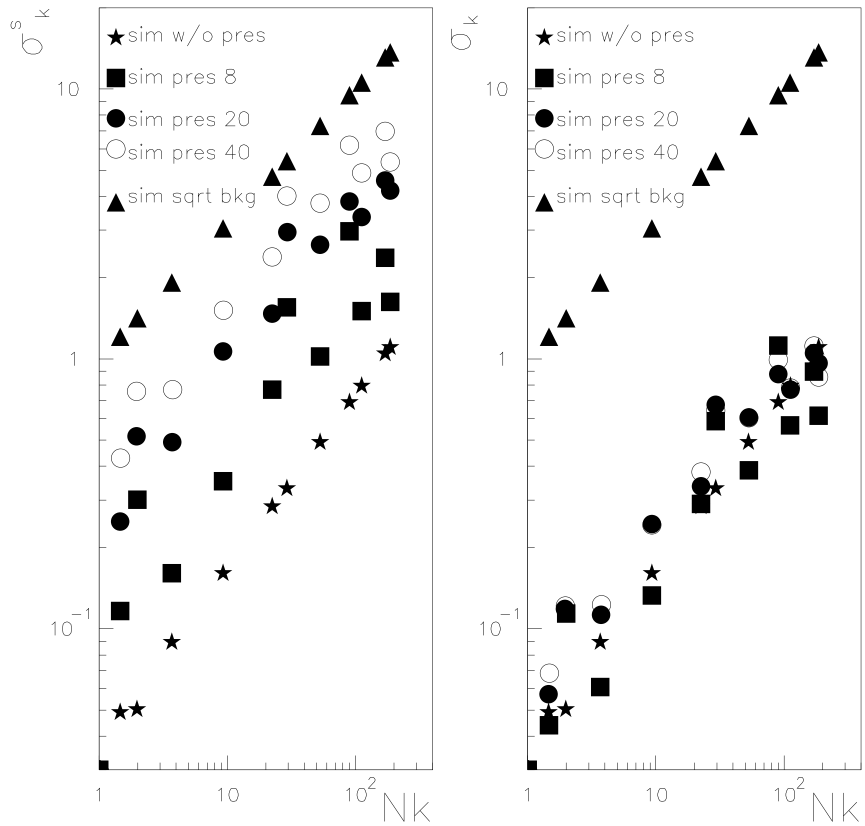

Figure 10: *Estimates of* σ_K^S *(for a sub-sample) and* σ_K *for the background (averaged over all the sub-samples) normalized to the same number of events as a function of* N_K *obtained using recombinations of simulated events and different values for* N_S*,*

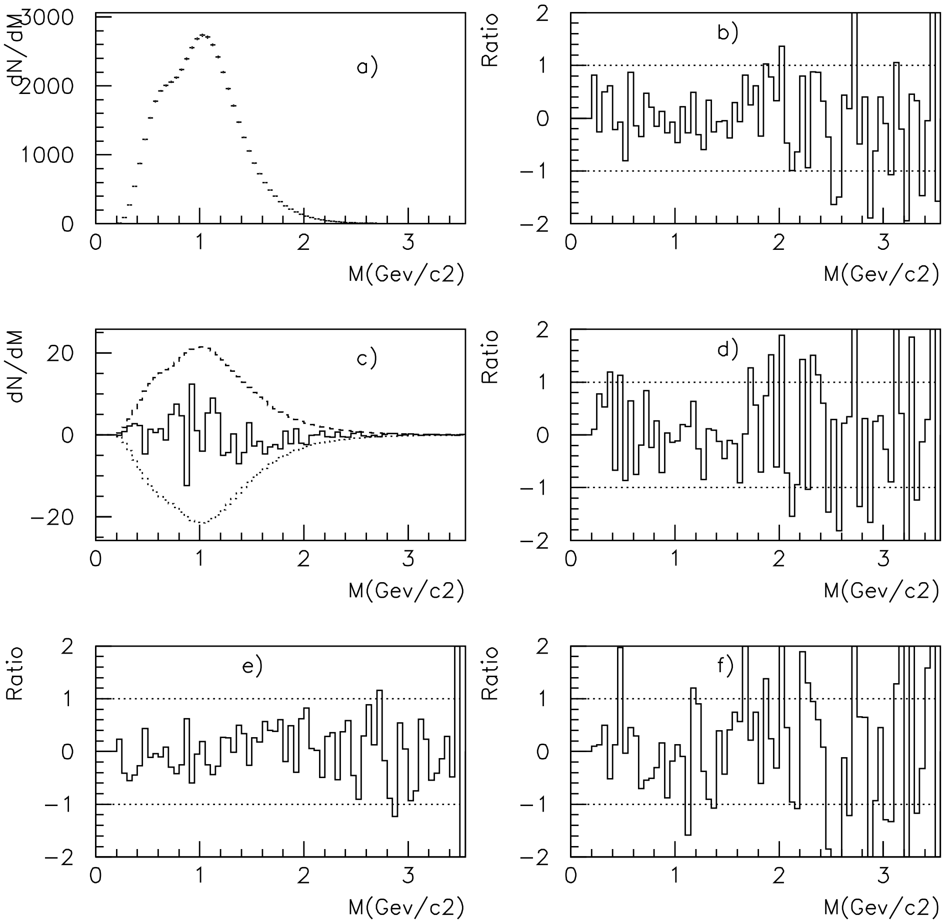

Figure 11: *Comparison of the fluctuations induced by prescaling to the estimated error, in S-U experiment. a) mass distribution of the background obtained for all targets and first transverse energy bin c) difference between the spectra with prescaling 30 and without prescaling, and the calculated error e) ratio of the previous difference and the calculated error and b), d), f): idem with prescaling 50, 100, 150.*

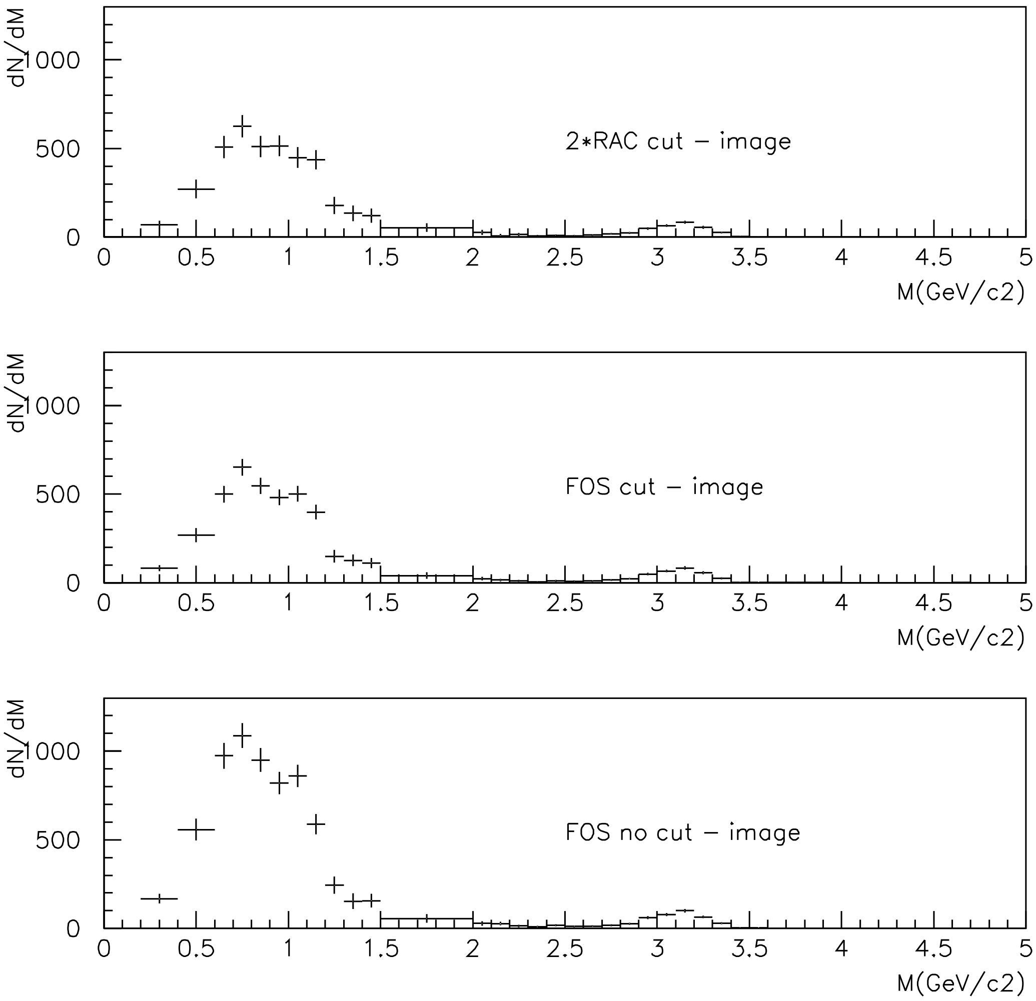

Figure 12: *Comparison of the two estimates of the NA38 experimental S-U signal, for* $0.6 < P_t < 1.2$ *GeV/c with (FOS) or without (2*RAC) the recombinatorial method and with image cut. Also shown is the same experimental result without image cut and using the recombinatorial method*

- about 70% of the signal errors are due to statistical fluctuations on the opposite-sign muon pair spectra, which do not depend on the background determination.

Nevertheless when the "image cut" is not used one observes (Fig. 12) a clear improvement in the relative precision of the low masses signal spectrum, due to the increase of statistic.

The background determination by recombination could then be efficient for improving the measurement thanks to decrease of the error, but also thanks to the possibility of removing completely the "image cut" (which actually remains partially in the data through a online cut performed at trigger level). Such increase of the acceptance should allow to diminish the minimal transverse momentum threshold considered in the low masses analysis [15].

Another interest of the method lies in the low statistic situations, either for the signal as for instance the continuum [5], or for multi-dimensional studies [16].

8 Conclusion

The use of the recombinatorial method to determine the opposite-sign combinatorial background in the dimuon spectra has been presented, and the error has been evaluated by several ways, leading to consistent results.

When different subclasses of events, for instance corresponding to different targets, introduce differences in the single muons distributions, it can be necessary for the validity of the model supporting the various background determination methods, to deal separately with the various subsamples. Compared to the usual, and easier to use, "integrated" method, the recombinatorial method has the advantage of allowing to treat separately different sub-classes of events, not being sensitive to the statistic of the sample. For the same reason this method is also the only practical one for multidimensional studies, where the number of empty cells in pair distributions becomes important.

Furthermore recombinations of single muons do not require restrictions on the kinematical domain, and then permits using of the complete statistics available and maximal acceptance of the detection system.

The non trivial problem of the associated error has been extensively studied and tested in various scenarii and by using various tools, including analytical calculations. An easy-to-use and flexible approximate formula has been derived. The important computing time can be reduced by the introduction of an optimized prescaling factor.

This study deals mainly with the error associated to the recombination process. Some questions can be adressed on possible systematic biases in the estimate of continuum signal [17], in particular when the signal is the origin of a combinatorial background as for the $D\bar{D}$ component at LHC energy. This is outside of the scope of this study.

These works have been partly supported by an IFA-IN2P3/CNRS French Romanian agreement 92-12 and 99-26. NA38 and NA50 collaborations are thanked by the authors of this prepublication for the possibility to use the experimental distributions for the test of the procedure, and for support of the grant.

References

[1] see for instance for a review the quark matter conferences: J.P.Blaizot, Nuclear Physics A698 (2002) 360c-371c

[2] H.SATZ, proceedings of QM95 conference, Nuclear Physics A 590 (1995) 63c

[3] M.C.Abreu et al.,Physics Letters B 477 (2000) 28

[4] M.C.Abreu et al., Physics letters B 410 (1997) 327-336

[5] M.C.Abreu et al., European Physics Journal C13 (2000) 69

[6] C.Baglin et al., Physics letters B 345 (1995) 617-621

[7] S.Papillon thesis, Paris 7 University, IPNO T 91-03

[8] L.Behr, in S.Papillon thesis, private communication

[9] S.Constantinescu, S.Dita, D.Jouan, internal report IPNO-DRE-96-01

[10] W.A. Zajc et al., Phy Rev C 29 (1984) 2173.

[11] D.Drijard et al., NIM A 225(2) (1984) 367.

[12] G.Jancso et al., Nuclear Physics B 124 (1977) 1-11

[13] S.Langlois, rapport de stage du DEA instrumentation et commande, 24 septembre 1993, unpublished

[14] L.Anderson et al., NIM 223 (1984) 26-39

[15] M.C.Abreu et al., Physics Letters B368 (1996) 238
M.Abreu et al., Nuclear physics A638 (1998) 483c-486c
M.Abreu et al., Nuclear physics A638 (1998) 487c-490c

[16] M.C.Abreu et al., Nucl.Instrum.Meth.A405:139-152,1998

[17] M.Gazdzicki, M.I.Gorenstein, hep-ph/0003319, J.Phys.G27(2001)41

9 Appendix.

Starting from the general expressions 16, one must evaluate:

$$\alpha_A = \int G_B^2(x) f_A(x) dx = \int G_A^2(y) f_B(y) dy$$

$$\alpha = \int G_A(y) f_B(y) dy = \int G_B(x) f_A(x) dx$$

where $G_B(x) = \int U^I(x,y) f_B(y) dy$

We consider the case of gaussian distributions for f_A and f_b with similar width σ and means μ_A and μ_B:

$$f_A(x) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma} \exp\left\{ -\frac{(\mu^A - x)^2}{2(\sigma)^2} \right\}$$

and a simple indicative function $U^I(x,y) = \delta(x+y-I)$

Inserting these formulas for α_A and α_B and performing the integrations using the δ Dirac distribution one gets:

$$\begin{aligned}
\alpha &\sim \int dx \exp\left\{ -\tfrac{1}{2\sigma^2} \left[\left[(I-\mu_B) - x \right]^2 + (\mu_A - x)^2 \right] \right\} \\
\alpha_A &\sim \int dx \exp\left\{ -\tfrac{1}{2\sigma^2} \left[2\left[(I-\mu_B) - x \right]^2 + (\mu_A - x)^2 \right] \right\} \\
\alpha_B &\sim \int dy \exp\left\{ -\tfrac{1}{2\sigma^2} \left[2\left[(I-\mu_A) - y \right]^2 + (\mu_B - y)^2 \right] \right\}
\end{aligned}$$

By using the following formulas:

$$\begin{aligned}
(a-x)^2 + (b-x)^2 &= 2\left[x - \tfrac{a+b}{2} \right]^2 + \tfrac{(a-b)^2}{2} \\
2(a-x)^2 + (b-x)^2 &= 3\left[x - \tfrac{2a+b}{3} \right]^2 + \tfrac{4}{3}\tfrac{(a-b)^2}{2}
\end{aligned}$$

one gets, introducing $\sigma_S = \sqrt{2}\sigma$:

$$\begin{aligned}
\alpha &\sim \int dx \exp\left\{ -\tfrac{1}{\sigma^2} \left[x - \tfrac{a+b}{2} \right]^2 - \tfrac{(a-b)^2}{4\sigma^2} \right\} \\
&= \exp\left\{ -\tfrac{(a-b)^2}{4\sigma^2} \right\} \int dx \exp\left\{ -\tfrac{1}{\sigma^2} \left[x - \tfrac{a+b}{2} \right]^2 \right\} \\
&= \exp\left\{ -\tfrac{(a-b)^2}{2\sigma_S^2} \right\} \int dx \exp\left\{ -\tfrac{1}{\sigma^2} \left[x - \tfrac{a+b}{2} \right]^2 \right\}
\end{aligned}$$

and

$$\begin{aligned}
\alpha_A &\sim \int dx \exp\left\{ -\tfrac{3}{2\sigma^2} \left[x - \tfrac{2a+b}{3} \right]^2 - \tfrac{2}{3}\tfrac{(a-b)^2}{2\sigma^2} \right\} = \\
&= \exp\left\{ -\tfrac{2}{3}\tfrac{(a-b)^2}{2\sigma^2} \right\} \int dx \exp\left\{ -\tfrac{3}{2\sigma^2} \left[x - \tfrac{2a+b}{3} \right]^2 \right\} = \\
&= \exp\left\{ -\tfrac{4}{3}\tfrac{(a-b)^2}{2\sigma_S^2} \right\} \int dx \exp\left\{ -\tfrac{3}{2\sigma^2} \left[x - \tfrac{2a+b}{3} \right]^2 \right\}
\end{aligned}$$

using $\int_0^\infty dx \exp\left\{ -a\,(x-b)^2 \right\} = \frac{\sqrt{\pi}}{\sqrt{a}}$, and

$$a = \mu_A; b = I - \mu_B; a - b = \mu_A + \mu_B - I = \mu_S - I$$

it comes:

$$\alpha \sim \sigma\sqrt{\pi}\exp\left\{-\frac{(a-b)^2}{2\sigma_S^2}\right\} = \frac{\sigma_S\sqrt{2\pi}}{2}\exp\left\{-\frac{(\mu_S-I)^2}{2\sigma_S^2}\right\}$$

and

$$\alpha_A = \alpha_B \sim \frac{\sigma\sqrt{2\pi}}{\sqrt{3}}\exp\left\{-\frac{(a-b)^2}{2\sigma_S^2}\right\}^{\frac{4}{3}} = \frac{\sigma_S\sqrt{\pi}}{\sqrt{3}}\exp\left\{-\frac{(\mu_S-I)^2}{2\sigma_S^2}\right\}^{\frac{4}{3}}$$

If we identify

$$f_S = \frac{1}{\sqrt{2\pi}\sigma_S}\exp\left\{-\frac{(a-b)^2}{2\sigma_S^2}\right\} = \frac{1}{\sqrt{2\pi}\sigma_S}\exp\left\{-\frac{(\mu_S-I)^2}{2\sigma_S^2}\right\}$$

and take into account the proportional factors we obtain:

$$\alpha_A + \alpha_B = 2\frac{1}{\left(\sqrt{2\pi}\right)^3\sigma_S^3}\frac{\sigma_S\sqrt{\pi}}{\sqrt{3}}\exp\left\{-\frac{(\mu_S-I)^2}{2\sigma_S^2}\right\}^{\frac{4}{3}} =$$
$$= \frac{1}{\pi\sigma_S^2\sqrt{6}}\frac{\sqrt{2\pi}\sigma_S}{\sqrt{2\pi}\sigma_S}\exp\left\{-\frac{(\mu_S-I)^2}{2\sigma_S^2}\right\}^{\frac{4}{3}}$$

$$\alpha_A + \alpha_B = \frac{1}{\sigma_S\sqrt{3\pi}}\left[N_S\right]^{\frac{4}{3}}$$

and

$$\alpha = \frac{1}{\left(\sqrt{2\pi}\right)^2\sigma_S^2}\frac{\sigma_S\sqrt{2\pi}}{2}\exp\left\{-\frac{(\mu_S-I)^2}{2\sigma_S^2}\right\} =$$
$$= \frac{1}{2\sqrt{2\pi}\sigma_S}\exp\left\{-\frac{(\mu_S-I)^2}{2\sigma_S^2}\right\} =$$

$$\alpha = \frac{1}{2}f_S$$

finally:

$$\frac{\sigma(\bar{p}_I)}{\sqrt{\langle N\rangle}} \simeq \sqrt{\alpha_A+\alpha_B-\alpha^2}$$

$$\frac{\sigma(\bar{p}_I)}{\sqrt{\langle N\rangle}} \simeq \sqrt{\frac{1}{\sqrt{3\pi}}\frac{1}{\sigma_S}\left[f_S\right]^{\frac{4}{3}} - \frac{1}{4}\left[f_S\right]^2}$$

The highest power of f_S inside $\frac{\sigma(\bar{p}_I)}{\sqrt{\langle N\rangle}}$ is 1 coming from α^2 and the next one is 0.66 coming from $\alpha_A+\alpha_B$, which is consistent with the result of the phenomenological evaluation of the fluctuations.

Contents