E877 Note # PITT-95-004

October 24, 1995

MAXIMUM LIKELIHOOD FIT OF THE CORRELATION FUNCTION WITHOUT BINNING

Sergei VOLOSHIN[a]

University of Pittsburgh, Pittsburgh, PA 15260

ABSTRACT

We discuss an application of the maximum likelihood technique which permits the fit of the two particle correlation function without binning.

4

To be specific we discuss below the fit of the two-particle correlation function constructed in the so called "mixed event" technique. In this approach the two particle correlation function can be defined as the ratio of the distributions of pairs of particles from the same event and pairs from "mixed events":

$$C(\mathbf{p}_1, \mathbf{p}_2) = \frac{dN_{true}(\mathbf{p}_1, \mathbf{p}_2)}{dN_{mixed}(\mathbf{p}_1, \mathbf{p}_2)}. \tag{1}$$

The usual approach for the fit of the correlation function assumes the introduction of binning of the part of or the whole phase space. This approach has many disadvantages. It is sensitive to acceptance and cuts introduced

[a]On leave from Moscow Engineering Physics Institute, Moscow, 115409, Russia

during the analysis, very often some bins have very low number of hits what can introduce an additional bias in the fit. Sometimes the parameterized correlation function has not so many parameters, but just the functional form of the parameterization demands the binning in many dimensions, etc.

It is very attractive in this case to use the maximum likelihood fitting technique, which is free from almost all problems mentioned above. But direct application of this technique is impossible because in this approach one needs to parameterize not only the correlation function itself (what is the goal of the fit), but of the very distributions of true and "mixed" pairs. The latter is extremely difficult especially taking into account numerous nontrivial acceptance and analysis cuts.

Below we propose the method which permits to avoid this problem. For simplicity we consider one dimensional distributions, but the generalization of the method for many dimensions is straightforward. In one dimensional case the correlation function is written as:

$$C(x) = \frac{dN_{true}(x)}{dN_{mixed}(x)}. \tag{2}$$

We denote the probability density for the mixed pairs as $p(x)$:

$$\frac{1}{N_{mixed}} \frac{dN_{mixed}}{dx} = p(x) \tag{3}$$

Then the probability density for true pair distribution can be written as

$$\frac{1}{N_{true}} \frac{dN_{true}}{dx} = A[C]C(x)p(x), \tag{4}$$

where A is the normalization factor

$$\frac{1}{A[C]} = \int dx C(x)p(x). \tag{5}$$

The likelihood function depends on the parameterization of the correlation function in the following way:

$$\ln L[C] = \sum_{k=\{true\ pairs\}} \{\ln A[C] + \ln C(x_k)) + \ln p(x_k)\}. \tag{6}$$

The last term in the sum does not depend on $C(x)$ and can be omitted. But one must evaluated the normalization factor $A[C]$ for different sets of parameters. The problem is that for this calculation one needs to know the probability density $p(x)$. The trick of our method is that the required integral is evaluated using "mixed" pair distribution.

$$\frac{1}{A[C]} = \int dx C(x) p(x) = \int C(x) \frac{dN_{mixed}}{dx} dx \tag{7}$$

In other words the normalization factor is the inverse of the sum over all mixed pairs weighted with the (parameterized) correlation function.

In Fig. 1 we illustrate the method by the results of simple simulation. First we simulate the distribution of "mixed" pairs in the interval $0 < x < 1$ in accordance with the probability density, which for this example we choose proportional to x. The distribution of "true" pairs is simulated for the probability density

$$p_{true}(x) \propto x \cdot \exp(-x^2/\sigma^2) \tag{8}$$

with $\sigma^2 = 10$, what corresponds to the Gaussian form of the correlation function. The "true" and "mixed" pair distributions and their ratio are presented in Fig. 1 (a), (b), and (c), respectively, as histograms (not used in later fit!). The goal of the fit is to evaluate the parameter x_0 assuming the parameterization of the correlation function in Gaussian form

$$C(x) = \frac{dN_{true}(x)}{dN_{mixed}(x)} \propto 1 + exp(-x^2/x_0^2) \tag{9}$$

without any assumption about the distributions dN/dx.

Following the prescription described above we construct the normalization factor

$$A(x_0^2) = \sum_{i=\{mixed\ pairs\}} (1 + exp(-x_i^2/x_0^2)). \tag{10}$$

Then the likelihood function can be written in as

$$\ln L(x_0^2) = \sum_{k=\{true\ pairs\}} \{\ln(1 + exp(-x_k^2/x_0^2)) - \ln A(x_0^2)\}. \tag{11}$$

The dependence of the likelihood function on x_0^2 is shown on Fig. 1(c). One can see that the likelihood function reaches the maximum at the value of $x_0^2 = 10$ which coincides with the input value.

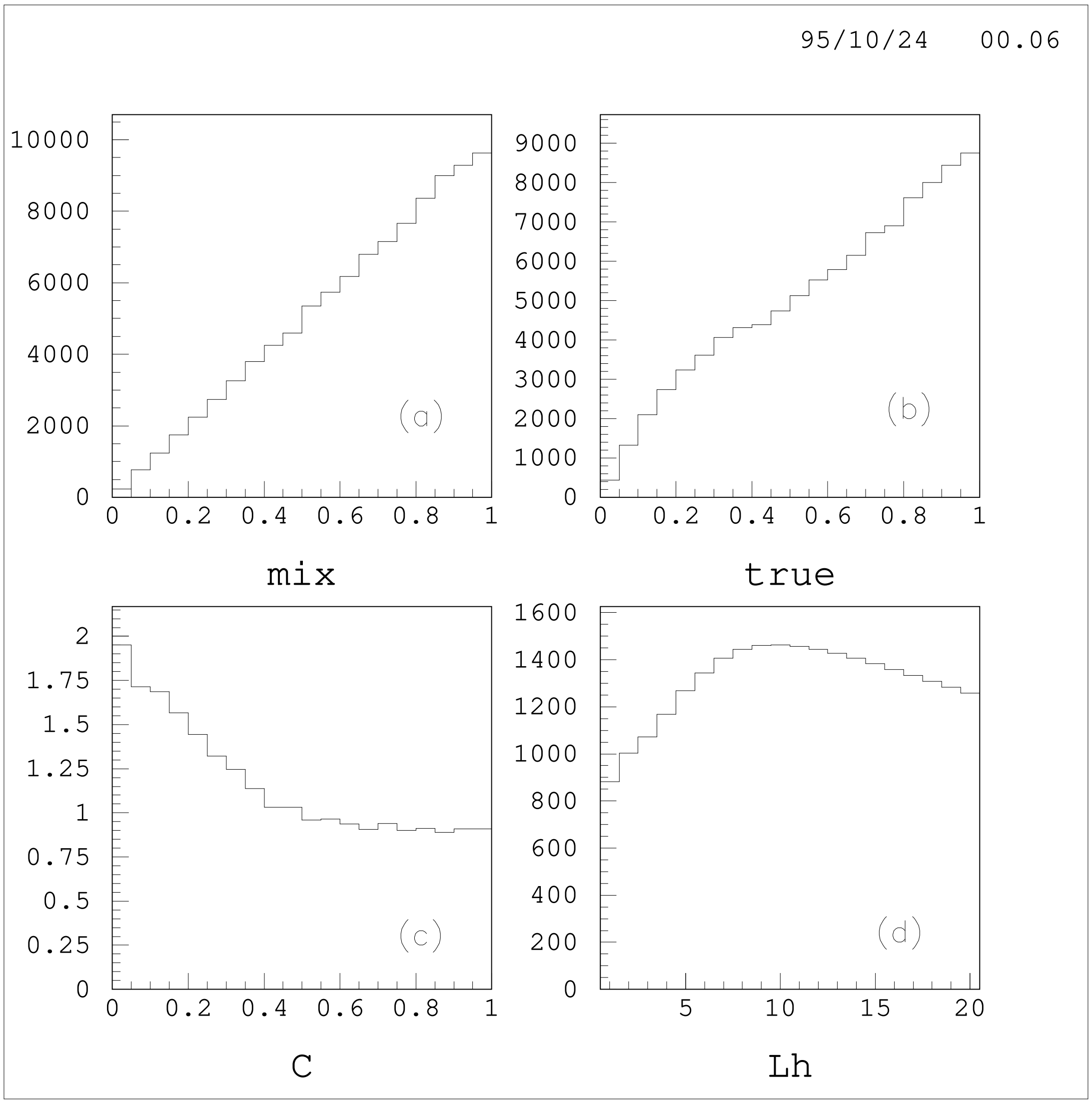

Figure 1:

(a) The distribution of “mixed” pairs.

(b) The distribution of “true” pairs.

(c) The correlation function.

(d) The likelihood function as a function of x_0^2.