E877 Note # PITT-94-002

July 18, 1994

PRODUCTION OF UNBIASED DISTRIBUTIONS FROM MULTILEVEL TRIGGERED DATA

Wilfred CLELAND and Sergei VOLOSHIN[a]

University of Pittsburgh, Pittsburgh, PA 15260

ABSTRACT

We discuss a technique which permits the construction of unbiased experimental distributions from multilevel triggered data for the case where the variable of interest is weakly correlated with the variable on which the data are triggered. We illustrate the method by constructing differential cross sections in transverse energy from TCAL for both TCAL- and PCAL-triggered events.

Most of the data collected in E814 and E877 rely on the use of several thresholds on a given trigger quantity, along with downscaling to produce approximately equal statistical weight for the sample over the entire spectrum. It is a technique that is commonly exploited in experiments that measure cross sections which vary by several orders of magnitude over the range of the variable covered in the experiment. In the analysis, natural distributions in the triggered quantity are produced for each trigger level. Events

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

in the threshold region are discarded, and a smooth unbiased distribution is obtained.

Quantities which are only loosely correlated with the trigger variable cannot be handled by the same technique. In the analysis of the 1993 E877 data, we have encountered this problem in calculating the differential cross section in TCAL E_T using data triggered on PCAL E_T, for the region of intermediate centrality, below the TCAL trigger thresholds. This note reports on the resolution of that particular problem and shows how it can be generalized.

The construction of the distribution in a variable which is used for triggering is straightforward. One only needs to multiply the distribution found for the events triggered at different levels by the downscaling factors used for each level. As an example, we show in Fig. 1 the distribution in TCAL E_T from the events triggered on the TCAL FERA information. Due to a nearly perfect correlation between TCAL E_T from the FASTBUS ADCs and the TCAL trigger sums read in the FERA, it is easy to select three regions and find the distributions in each region separately, using the appropriate trigger level for each region. The final distribution so obtained is presented in Fig. 2.

If we use this procedure to calculate the same distribution using PCAL-triggered events, we obtain the result shown in Fig. 3. Although the data triggered on different levels are consistent with each other, due to the lack of a sharp correlation between the triggered and analyzed variables, there are no longer distinct regions in the distribution corresponding to the different trigger levels.

The solution to the problem is based on the use of the expression which gives the distribution of a quantity x as an integral over the distribution in another quantity y, the one on which the trigger decision is based:

$$\frac{dw}{dx} = \int \frac{dw}{dy}(y)\frac{dP}{dx}(x;y)dy, \tag{1}$$

in which dw/dx is the wanted distribution, dw/dy is the distribution in y, and $dP/dx(x;y)$ is the distribution in x for events at a fixed value of y. It is important to recognize that, whereas dw/dy is the distribution of the triggered quantity which is strongly influenced by downscaling, the function $dP/dx(x;y)$ expresses the natural correlation between the two quantities and does not depend on how the data are triggered.

If downscaling is introduced in the trigger, the observed distribution dw/dy will not be the natural distribution. However, the natural distribution can be found from the observed distribution from a knowledge of the downscaling factors. Let us assume that there are 3 trigger levels Y_1, Y_2, and Y_3, and three corresponding downscaling factors[b] d_1, d_2, and d_3. The trigger condition is that if $y > y_i$, the event is written to tape with downscaling factor d_i. The distribution in y of the events which are written to tape is given by:

$$\frac{d\tilde{w}}{dy} = W(y)\frac{dw}{dy}, \tag{2}$$

where $W(y)$ is the probability for the event with certain value of y to be written into the tape and is constant within each trigger level region. The value of $W(y)$ depends only on the value of y (not on whether a particular

[b] The downscaling factor is the number of events which are to be skipped before writing of one event into the tape.

trigger bit is set) and is related to the downscaling factors by

$$W(y) = \begin{cases} W_1 = 1 - \frac{d_1}{d_1+1} & \text{if } Y_1 < y < Y_2 \\ W_2 = 1 - (\frac{d_1}{d_1+1})(\frac{d_2}{d_2+1}) & \text{if } Y_2 < y < Y_3 \\ W_3 = 1 - (\frac{d_1}{d_1+1})(\frac{d_2}{d_2+1})(\frac{d_3}{d_3+1}) & \text{if } y > Y_3 \end{cases} \tag{3}$$

This expression takes into account the fact that some events with the same value of y can be triggered by different trigger levels. We combine Eqs (1)-(3) to obtain

$$\frac{dw}{dx} = \int \frac{1}{W(y)} \frac{d\tilde{w}}{dy}(y) \frac{dP}{dx}(x;y)dy, \tag{4}$$

which can be used directly for the calculation of distributions of quantities from multilevel triggered data. Eq. (4) has a very simple interpretation: in the calculation of distributions of some quantity x each event should be weighed with the inverse probability of the event being triggered. For example, to calculate the natural distribution of events in TCAL E_T, one should evaluate the sum

$$\frac{\Delta N}{\Delta E_T} = \sum_{\text{events in } \Delta E_T} \frac{1}{W(\text{PCal FERA})}, \tag{5}$$

where the sum is taken over all triggered events written to tape with TCAL transverse energy between E_T and $E_T + \Delta E_T$. The variance in this quantity is given by

$$\sigma^2(\frac{\Delta N}{\Delta E_T}) = \sum_{\text{events in } \Delta E_T} \frac{1}{W^2(\text{PCal FERA})}. \tag{6}$$

In Fig. 4 we present the results for the distribution in TCAL E_T from the PCAL-triggered events sample. In Fig. 5 we show that the cross sections found from TCAL-triggered events and PCAL-triggered events are consistent. The loss of events in the low E_T region for the TCAL-triggered data is an obvious trigger bias, which is not present in the PCAL-triggered sample.

Note that, in contrast to the traditional technique of forming distributions of triggered quantities, one is insensitive to errors in the trigger weights. In the normal procedure, if the correlation between the triggered quantity and the analyzed quantity is imperfect, the fraction of events discarded increases. With this technique it is of no consequence; only a knowledge of the downscaling factors and the value of the quantity on which the trigger decision was made is necessary. Note also that the distribution dw/dx can be multidimensional. Consider for example the quantity $d^2w/dE_Td\eta$, which is proportional to the pseudorapidity distribution for events with a given E_T. In this case one should substitute $d^2P/dyd\eta$ instead of dP/dy in Eq. (4), but operationally, the same procedure is followed.

Another useful application of this technique could be in the calculation of distributions from Monte-Carlo generated data, where events are generated in different regions of some variable (e.g. impact parameter). If the desired quantity is only loosely correlated with this variable, the same procedure could be used for obtaining its distribution.

This work was supported by the US DoE under Contract No. DEFG02-87ER40363.

REFERENCE

[1] E877 Logbook. Run 1993.

FIGURE CAPTIONS

1. Au+Au. Triggering on TCal FERA.
2. Au+Au. Final distribution combining three distributions in three different regions.
3. Au+U. Triggering on PCal FERA.
4. Au+U. Final distribution using the proposed technique.
5. Au+Au. Final distributions using the proposed technique based on TCal triggers (solid histograms) and PCal triggering (dashed histogram).

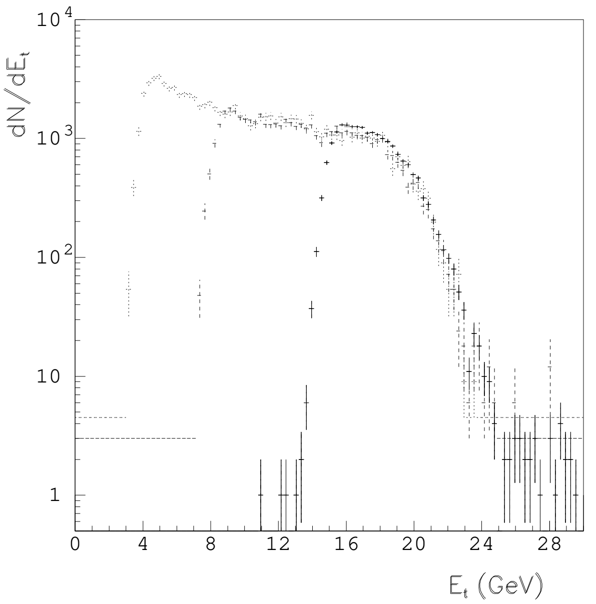

Figure 1:

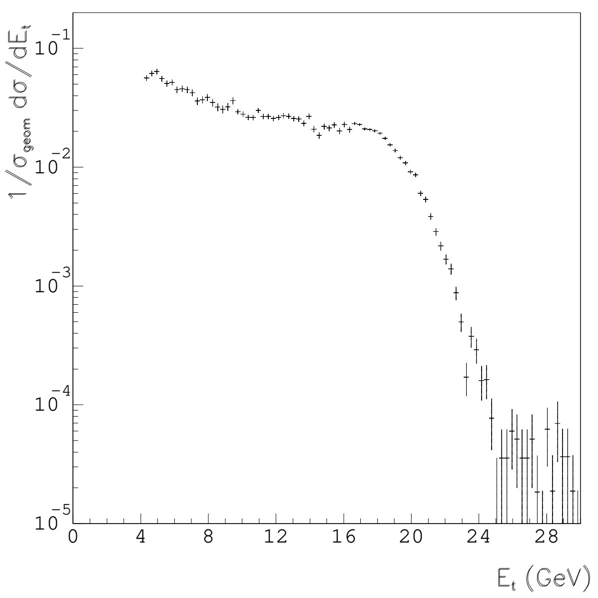

Figure 2:

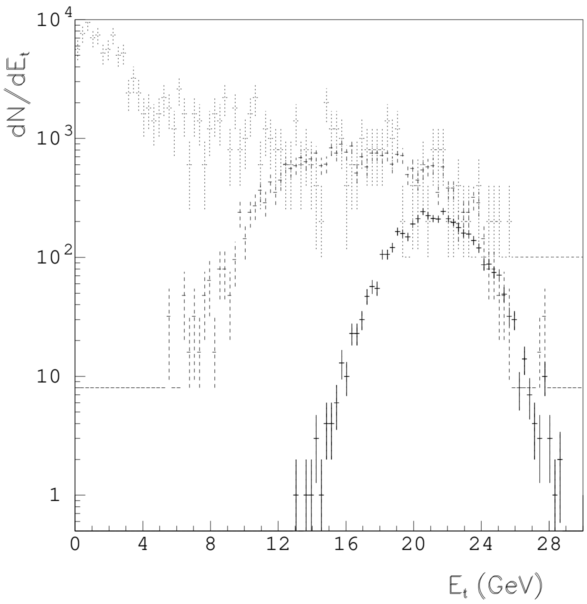

Figure 3:

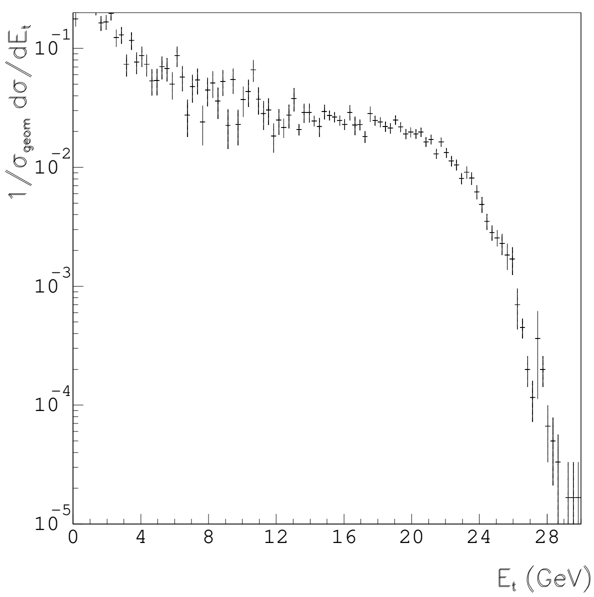

Figure 4:

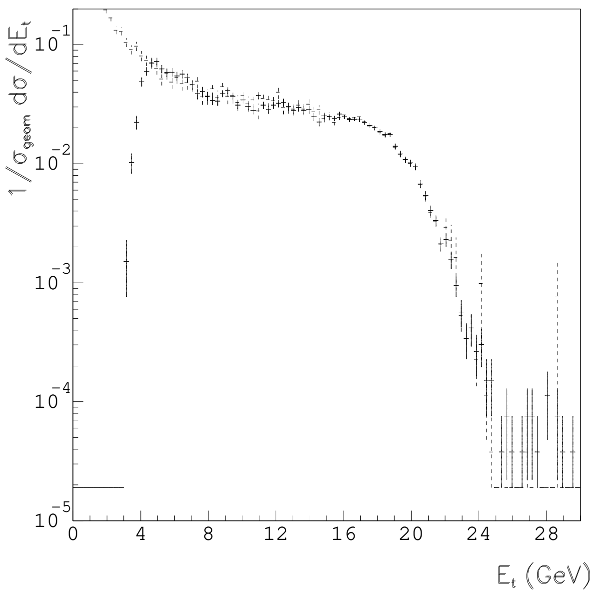

Figure 5: