A simple algorithm for hadron rejection in Milagro

Gaurang B. Yodh
Robert Atkins

February 24 1999

1 Introduction

We present an algorithm for rejecting cosmic ray showers in Milagro. We know that gamma ray showers are relatively smooth as a function of radial distance from the core while proton showers should be clumpy due to presence of muons, high energy gammas and hadrons. In the top layer their presence is obscured by general illumination by shower particles but may be their presence will show up in the bottom layer. Penetration of one of these particles, with or without interaction in the first 1.5 meters of water should lead to clusters which will have some tubes with large number of pes. This is of course the method for finding presence of muons in Milagro. There may be other circumstances which will also give such clustering. The muon layer also will be illuminated by Cherenkov light from the shower particles hitting the bottom tubes. The number of tubes hit in the bottom layer may be a measure of this light. We have found a simple variable which enhances the difference between cosmic ray and gamma ray showers using just the number of tubes hit in the bottom layer and the maximum pulse height or pe in the tubes hit in the bottom layer. In this memo, Quality factors are defined as the ratio

Q = (fraction of gammas retained after a suitable cut) divided by (square root of the fraction of protons retained after the same cut)

The Q factor defined here is related to statistical signficance from a source by

$$\frac{S}{\sqrt{(BG)}} = Q\frac{F_\gamma}{\sqrt{F_p}}\eta\sqrt{\frac{A_\gamma T}{A_p \Delta\Omega}} \tag{1}$$

where F's are effective fluxes of gammas and protons respectively, η is the fraction of gammas in source bin(0.72), A's are effective areas, $\Delta\Omega$ is the solid angle and T is the exposure time. Note that we are calculating Q for triggered events.

In general, the cut and therefore Q should be function of nhit in the top layer and the core distance from pond center. In what follows, we have used Joe's simulated proton and gamma samples to do this study. The Geant data files used are for events with more than 50 tubes hit in the shower layer.

2 Characteristics of MC events

Figure 1 shows that the nhitbot distributions for protons and gammas are reasonable similar. This means that the overall light that penetrates to the bottom layer is similar for the two species.

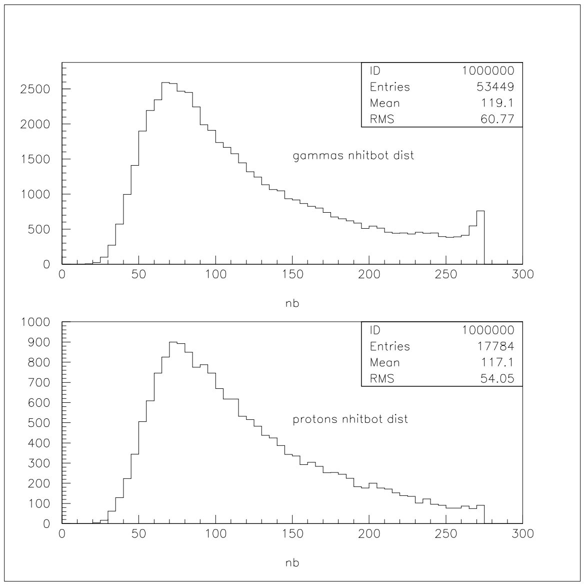


Figure 1: nhitbottom distributions for p and g

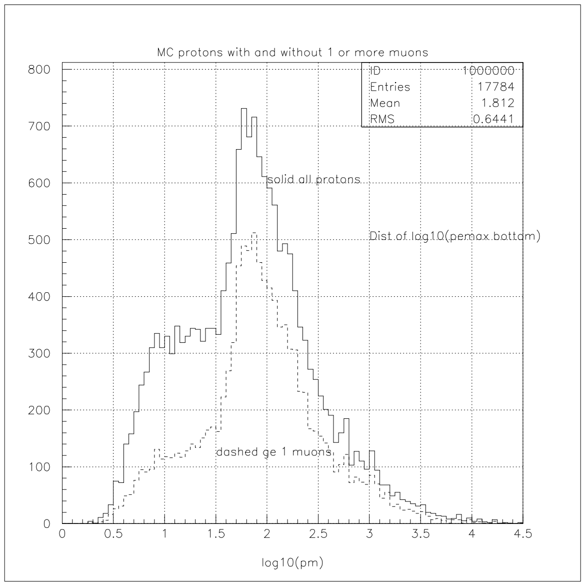


Figure 2:Histogram of log(pmax) in bottom with and without muons for protons

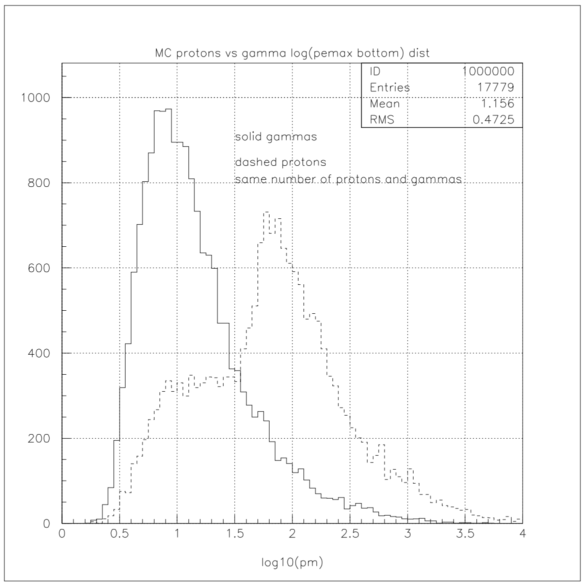


Figure 3: Histogram of log(pemax) in bottom P and G

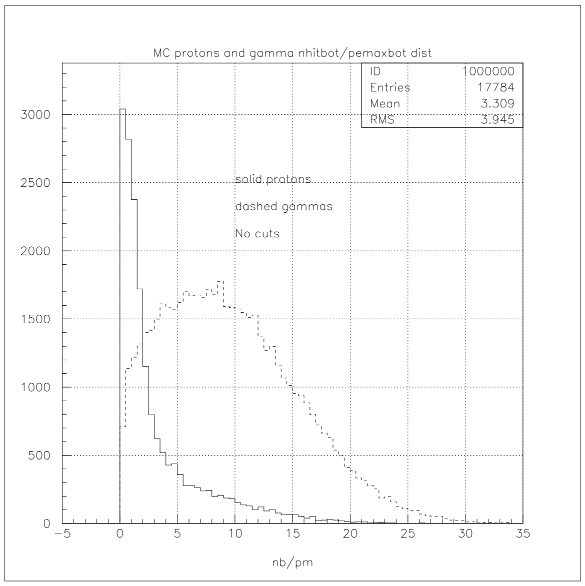


Figure 4:Histogram of nb/pemax for all protons and gammas

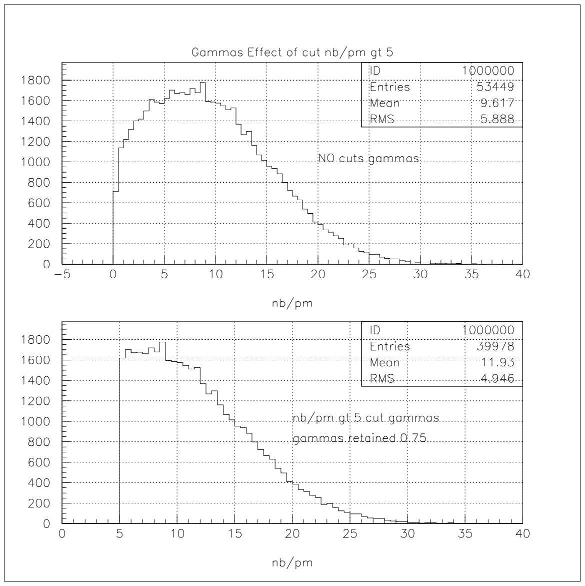
We use the sybol **pm** which corresponds to the number of photo-electrons(pes) of the tube with maximum number of pes in the bottom layer.A plot of log10(pm) shows a peak at 200 pes indicating the presence of penetrating particles in proton showers in Figure 2.

This figure also contains a subset of events in which there are one or more muons in Milagro. The peak for all events is significantly larger than that due to events with muons. In Figure 3 we compare the distributions for proton and gamma showers. A clear difference is seen, suggesting that we could use pm as an indicator for protons.

The variable we have chosen is x=nb/pm, where nb is the number of tubes hit in bottom layer. We show that even for all events, summed over all nhittop and all core distances, distributions in variable x can be used as a method to separate proton triggers from gamma triggers.

3 Quality Factors

By examining Figure 5 one sees that an xcut of $x \geq 5$ should be reasonable for separating gammas from protons. Figure 5 shows the effect of such a cut on gammas. With this cut, 75 percent of gammas are retained. Figure 6 shows the same for protons. Only a fraction of 0.2 protons are retained by the same cut. **The quality factor with this simple cut algorithm gives a Q of 1.61**



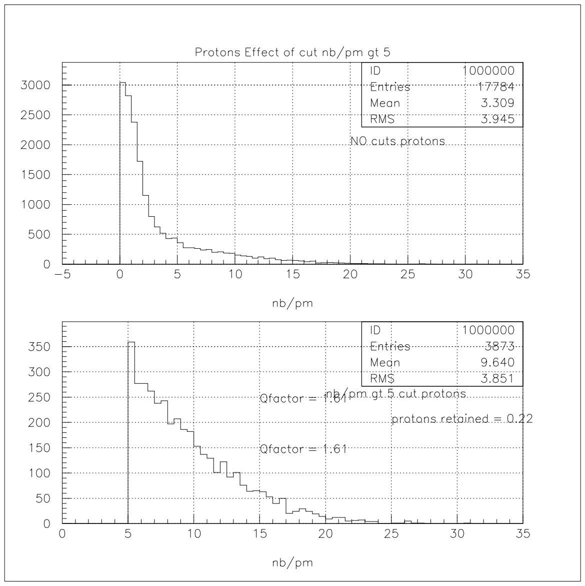


Figure 5: Effect of xcut of 5 or greater gammas

Figure 6:Effect of xcut of 5 or greater on protons

In the next two figures we show that for a given range of nhittop (chosen to be 50-70 hits in Figures 7 and 8) qfactor depends significantly on core distance. Figure 7 shows the x distribution for protons and gammas without and with a cut of $x \geq 5$ for events with core distance between 0 and 20 meters. Figure 8 shows the similar plots for core distance between 60 and 80 meters. We find the following results

Q(50<nhittop<70, 0<r<20m) = 1.28

Q(50<nhittop<70, 60<r<80m) = 1.78

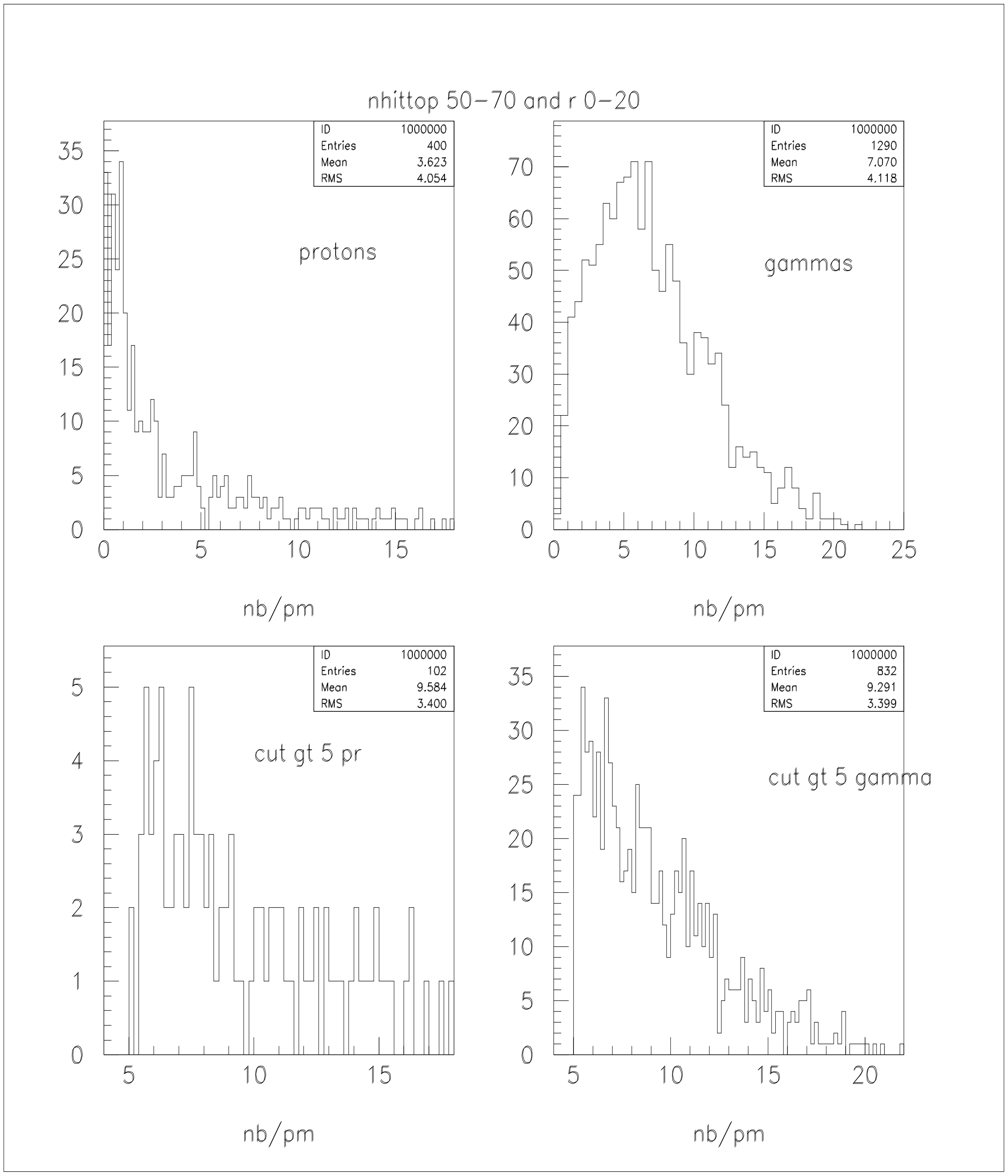


Figure 7: x dist for p and gammas r between 0-20 m

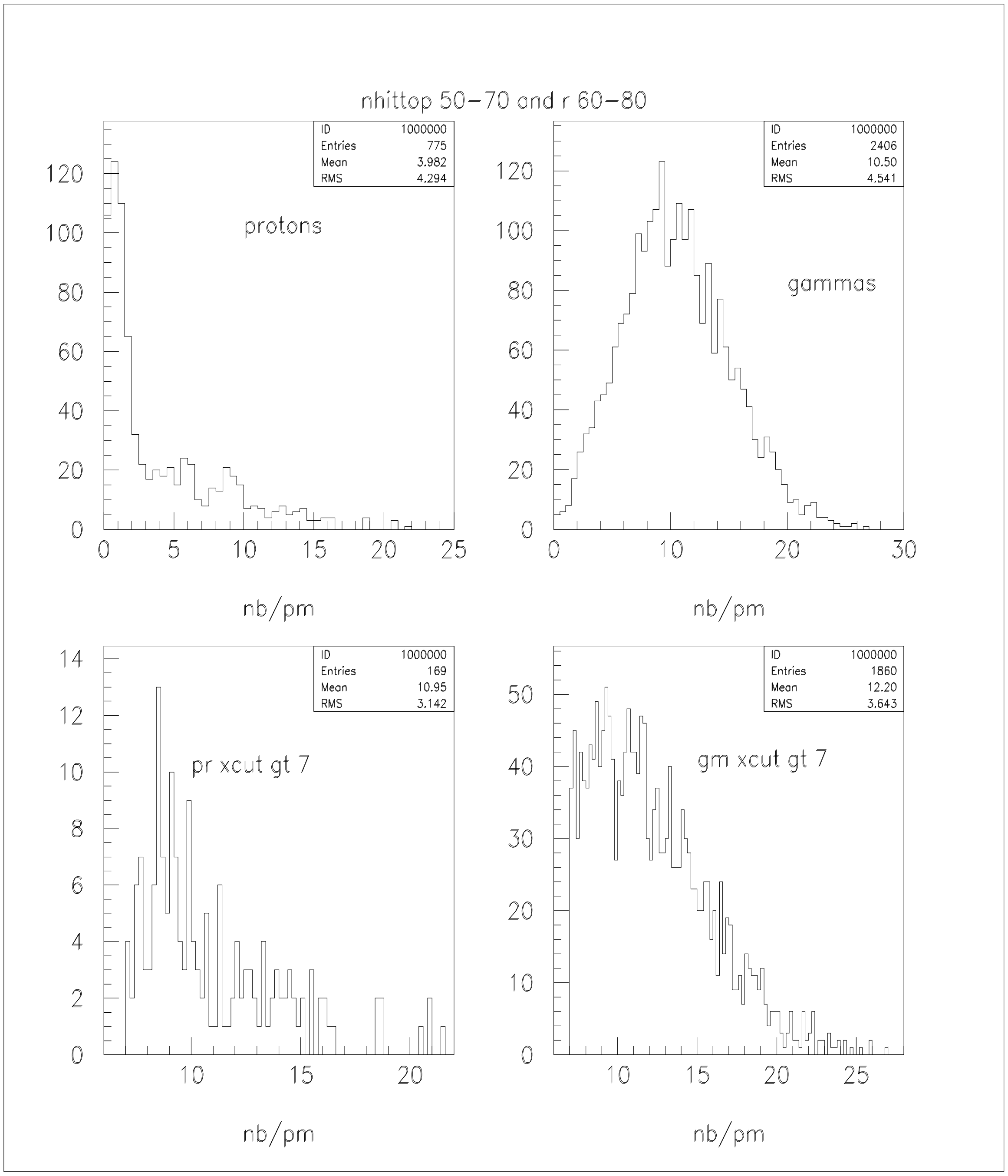


Figure 8: x dist for p and gammas r between 60-80 m

We also see that Q factor depends on xcut. It is important to maximize Q with respect to xcut with a requirement that fraction of gammas retained is greater than 50 percent. Such a calculation was done and the results are shown in Figures 9 and 10. Figure 9 is for a sample with nhittop between 50-70 tubes and Figure 10 is for nhittop between 100-150.

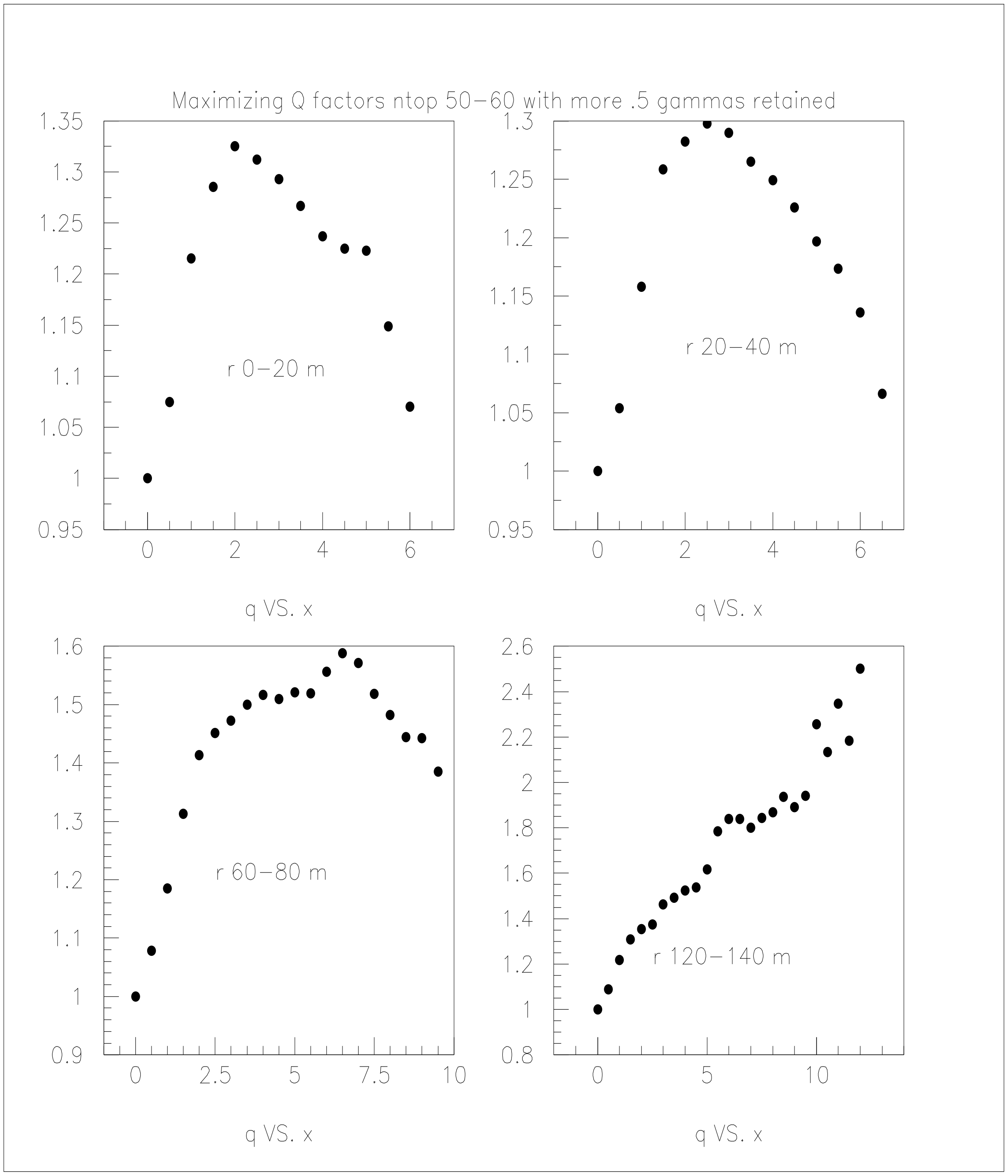


Figure 9: Q as function of x for different core dist bins

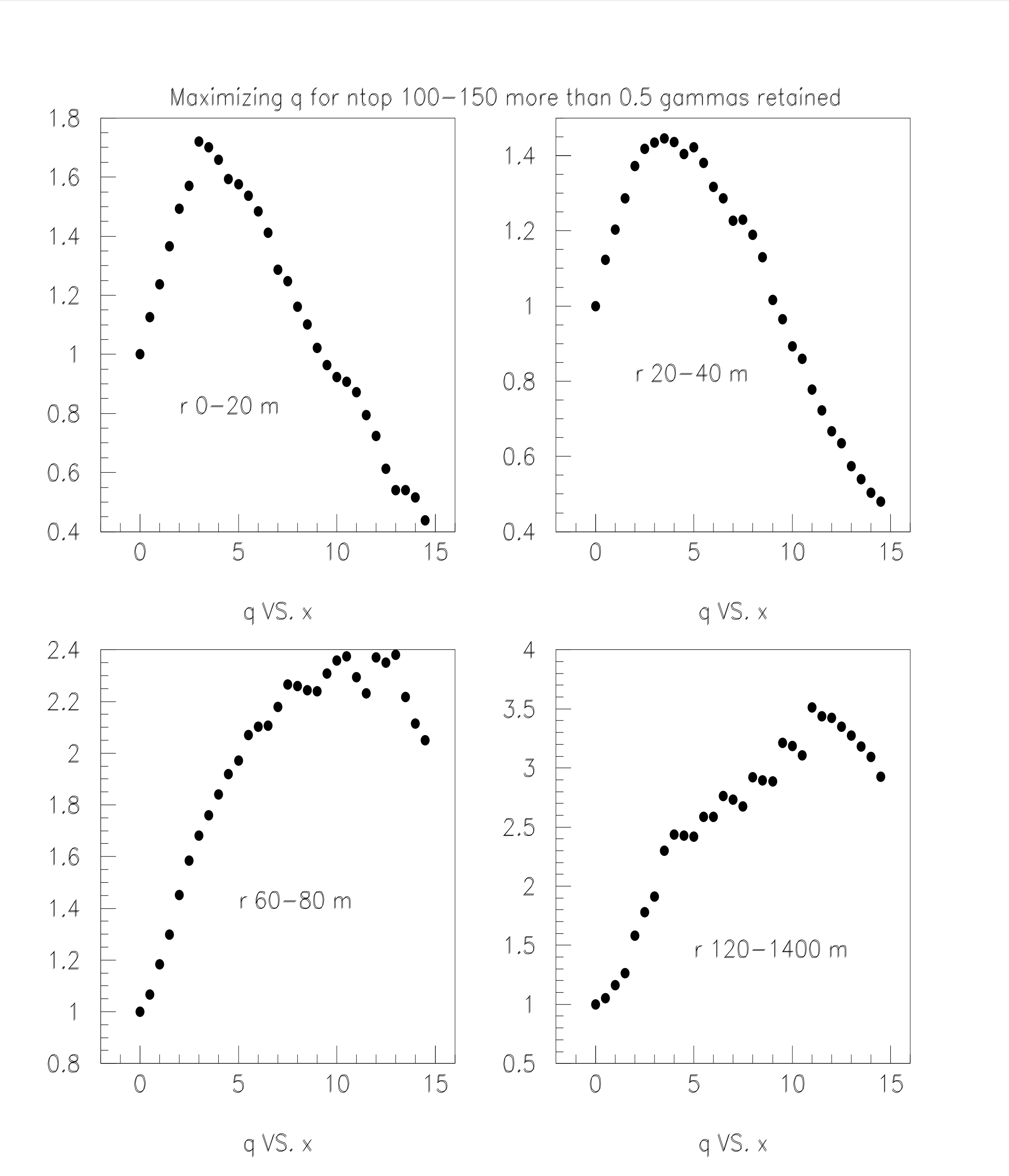


Figure 10: Q as a function of x for different core dist bins

Next we show the maximum q factor, qc, with $f_g \geq 0.5$ as a function of core distance for different bins of nhittop in Figure 11. Figure 12 shows the maximum q factor with $f_g \geq 0.5$ as a function of nhittop for different bins of core distance.

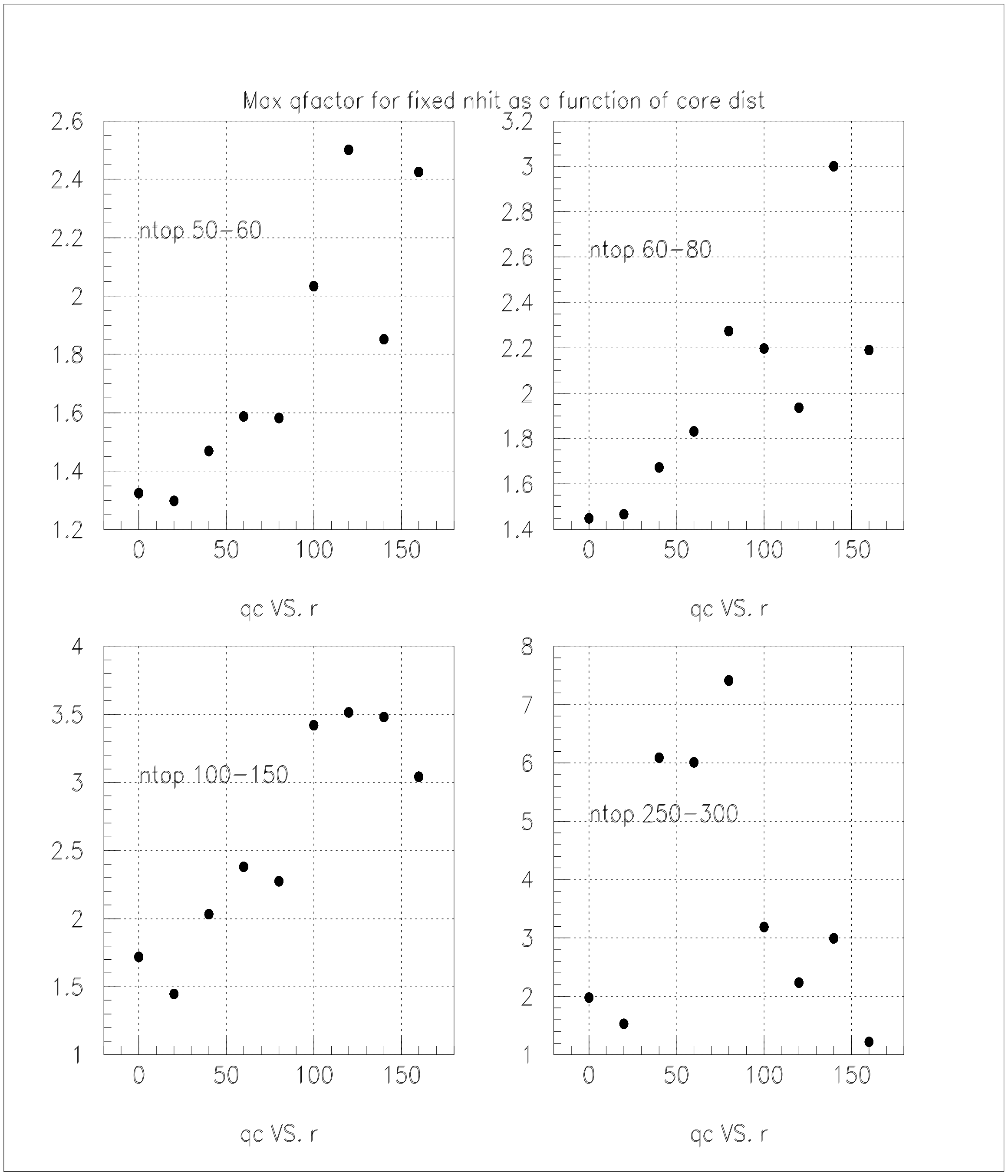


Figure 11: Q as function of r for different nhittop bins

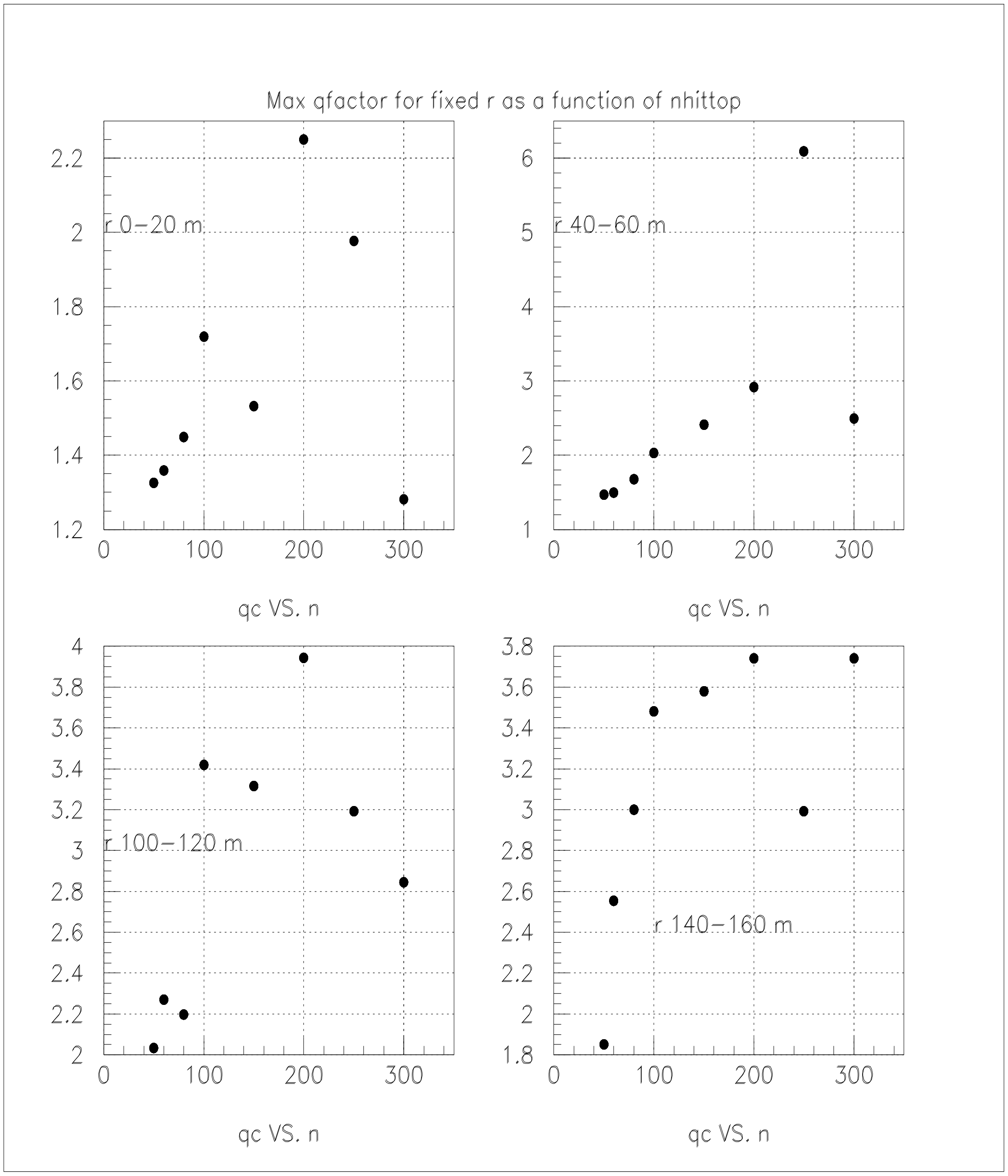


Figure 12: Q as a function of nhittop for different core dist bins

The fraction of protons retained,f_p when max qfactor is obtained is given in Figures 13 and 14 as a function core distance r and nhittop respectively.

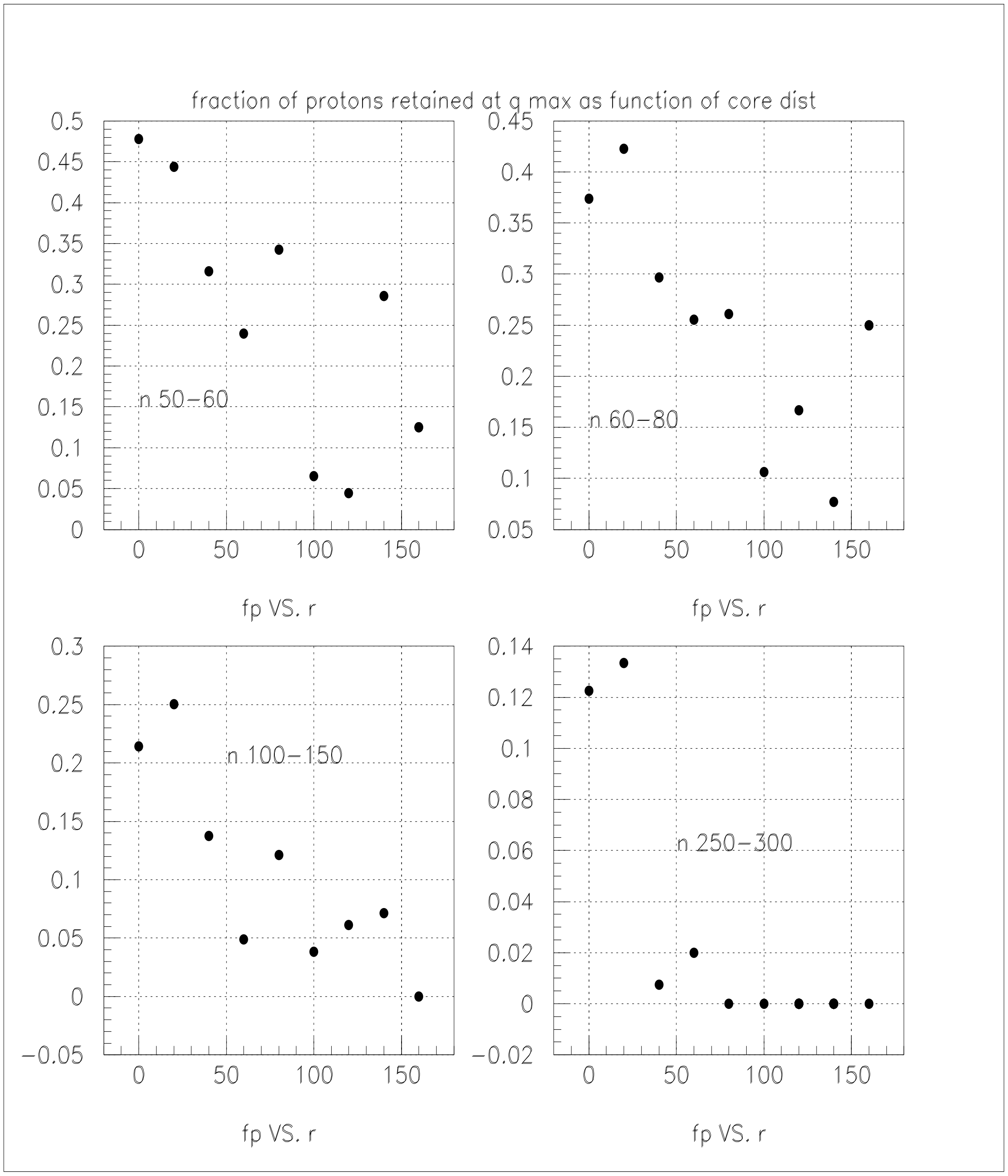


Figure 12: f_p as function of r for different nhittop bins

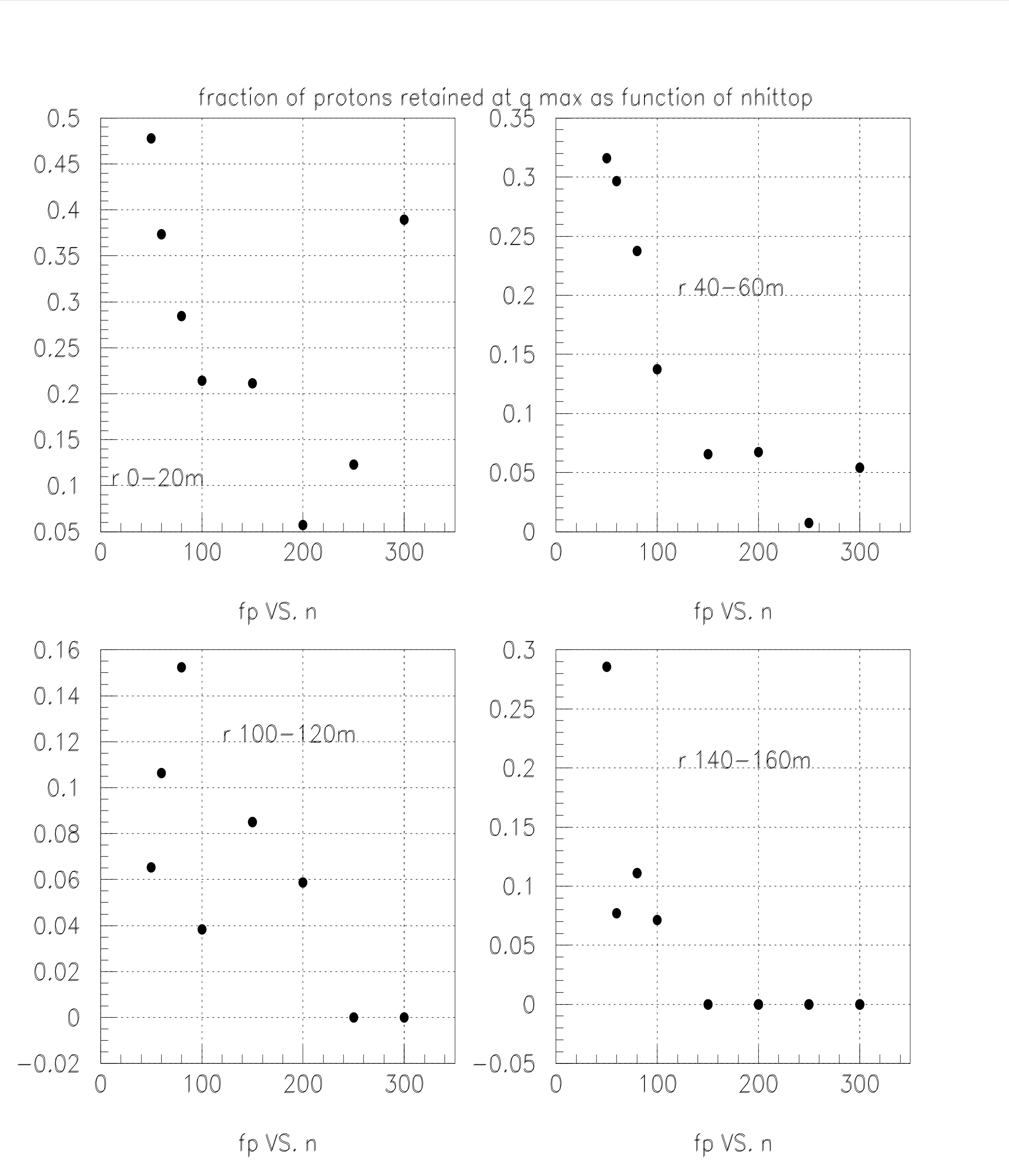


Figure 10: f_p as a function of nhittop for different core dist bins

These figures show the importance of knowing the core distance. For certain core distances and certain nhittops on can reject more than 90 percent of the protons while still retaining more than 50 percent of the gammas. The raggedness of some of the curves is due to paucity of simulated proton triggers in various bins.

When we use Stefan's simulations, in which all energy gammas and protons are thrown over the same area, to find the fraction of events which fall outside a core distance of 20 m, one finds that 90 percent of gamma cores fall outside and 93 percent of proton cores fall outside with a nhittop trigger of 50 or more tubes. Similar numbers for core distance greater than 40 m are 67 % and 78 % for gammas and protons respectively. Except for very small number of tubes in trigger, Figure 12 shows that we can get Q factors **larger than 2** for distances greater than 40 meters.**The need for a good outrigger array to determine core distances is obvious.**

4 Summary

We have shown that using the knowledge of number of hit tubes in the bottom layer and the number of pes in the tube with maximum pulse height inthe bottom layer, we can construct the variable x=nb/pm which is a sensitive variable to distinguish proton showers from gamma showers. We have studied the possibility of only using pm (or the total number of pes in the largest cluster, defined as all tubes within 6 meters of the max tube) for making the distinction but have found that these variables are not as sensitive as using x. If someone is interested we can show these distributions.

The Quality factors depend on both core distance and the number of tubes hit in the air shower layer. Smallest Q factors are obtained for showers whose cores hit the pond, about 1.3. For showers whose cores fall outside the pond most Q factors obtained by maximizing q factors with respect to xcut are larger than 2. As more than 65 percent of our events have triggers outside the pond these cuts should make a significant improvement in our sensitivity. We need to verify that real data, which are mostly protons, have x distributions which correspond to that of protons shown in Figure 4. It is therefore of utmost importance to fix the dead tubes in the bottom layer as soon as possible.

I have also calculated Q factors in which xcut was a function of nhittop and where core was allowed to be anywhere. The result of this calcuation is that **Qfactor varies from 1.78 to 2.8 as the minimum number of tubes hit in the top layer is varied from 50 to 150**. This was done using Stefan's simulations.