γ/Hadron Separation for Milagro and Milagrito

Stefan Westerhoff

March 8, 1999

1 Introduction

A couple of parameters have recently been suggested as tools for γ/hadron separation in Milagro. Some of these parameters are correlated to a certain degree, but their combination may nevertheless lead to an improved quality factor. It therefore is a natural approach not to work out cuts in single parameters, but to combine them and separate γ -and proton showers in the d-dimensional space spanned by the d parameters that have separation power. This is usually done by discriminant analyses, or, in cases where non-linear cuts in parameter spaces are crucial, with neural networks.

I have put together parameters that so far have turned out to be useful for γ/hadron separation in Milagro (see recent memo by Rob and Gaurang) and checked the quality factor obtained by training neural networks with these input parameters.

An important aspect of this analysis is the stability of the separation. Any γ/hadron separation depending on the two-layer structure of Milagro has to tolerate a certain number of dead tubes in both layers. Neural network analyses avoid sharp cuts in single parameters and are therefore relatively stable in case of corrupt data.

2 Input Parameters and Network Topology

The following 15 parameters are used as input parameters for a three layer feed-forward neural network:

- number of tubes hit in the top layer,
- number of tubes hit the bottom layer,

- number of PEs in the top layer,
- number of PEs in the bottom layer,
- PEs of the "hottest" bottom layer tube,
- the number of tubes and PEs in the 5 rings with 5, 10, 15, 20, and 25 m distance to the core as fitted by the center of gravity core fitter (thus basically the distance to the "hottest clump").

The network is trained to give the output "0" for γ- and "1" for proton showers. In order to avoid large numbers as network input parameters, I used the logarithm of the tube and PE numbers, and the number of tubes and PEs in the rings were divided by the total number of tubes or PEs After this "normalization", the network training was stable, and the final performance was largely independent of parameters like the learning rate (chosen to be 0.1), the number of hidden network nodes (10), and the momentum parameter (0.5). For the training, I used 3000 showers (1500 γ-showers and 1500 proton showers) with more than 50 air shower tubes, thrown over an energy-independant area ($300 \times 300\,\mathrm{m}^2$), with energies from 100 GeV to 500 TeV on a 2.4 (2.7) spectrum for γ-(proton) showers. The showers used for training the net were not used for evaluating the network performance. The network performance does not improve after 200 training epochs, and the difference in performance between the training sample and an independent test sample is about 10 %, indicating the stability of the training process.

γ-efficiency ϵ_γ, proton rejection $(1-\epsilon_p)$, and the resulting quality factor Q defined as

$$Q = \frac{\epsilon_\gamma}{\sqrt{1-\epsilon_p}} \tag{1}$$

are now determined using Joe's independent set of Monte Carlo showers. Figure 1 shows the network output for the 53 449 γ- and 17 784 proton showers of this sample, which in contrast to the training sample has an energy-dependent core area. Whereas the shape of the network output for proton showers does not depend on the position of the core, the network output for the γ-showers is considerably different for events with cores on or off the pond. The optimal cut in the network output is a function of the number of tubes in the air shower layer (an observable quantity). Figure 2 indicates that the optimal cut value in the network output moves to 0 with higher number of tubes. It is therefore necessary to either not apply a cut at all and use the network output as a measure of the "γ-likeness" of an event in

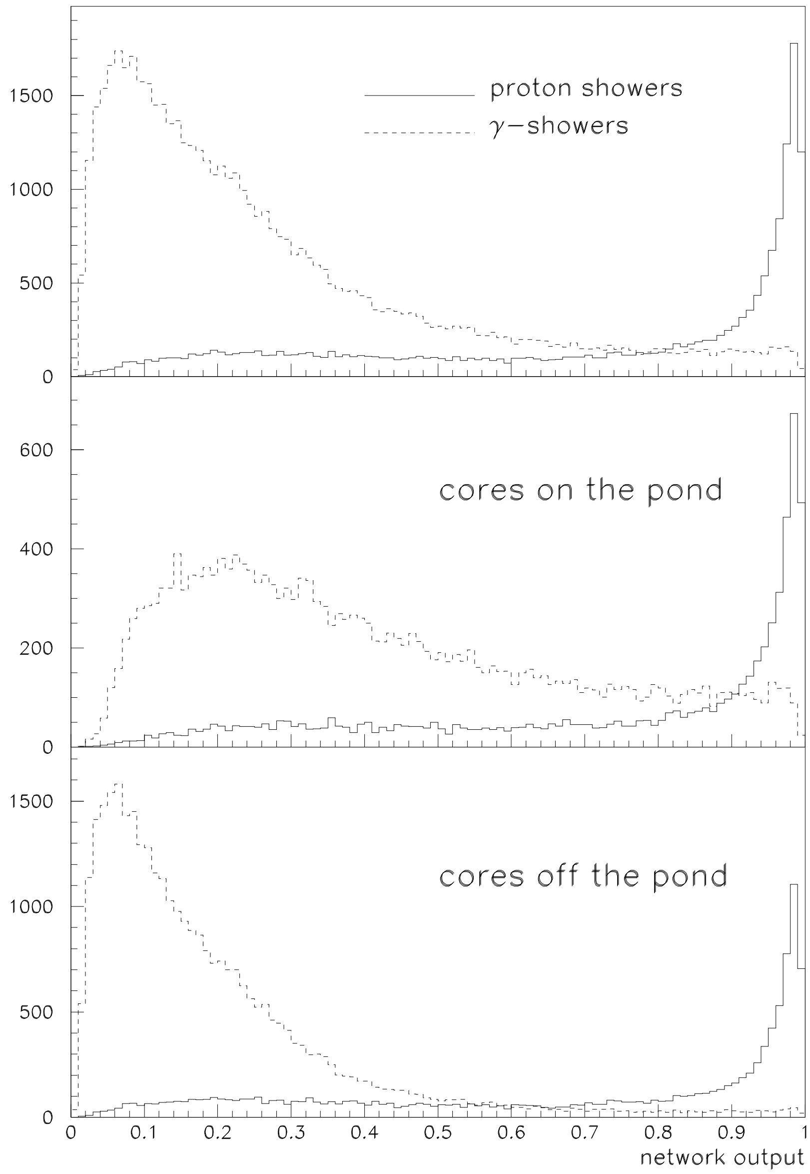

Figure 1: Network output for γ-showers and proton showers.

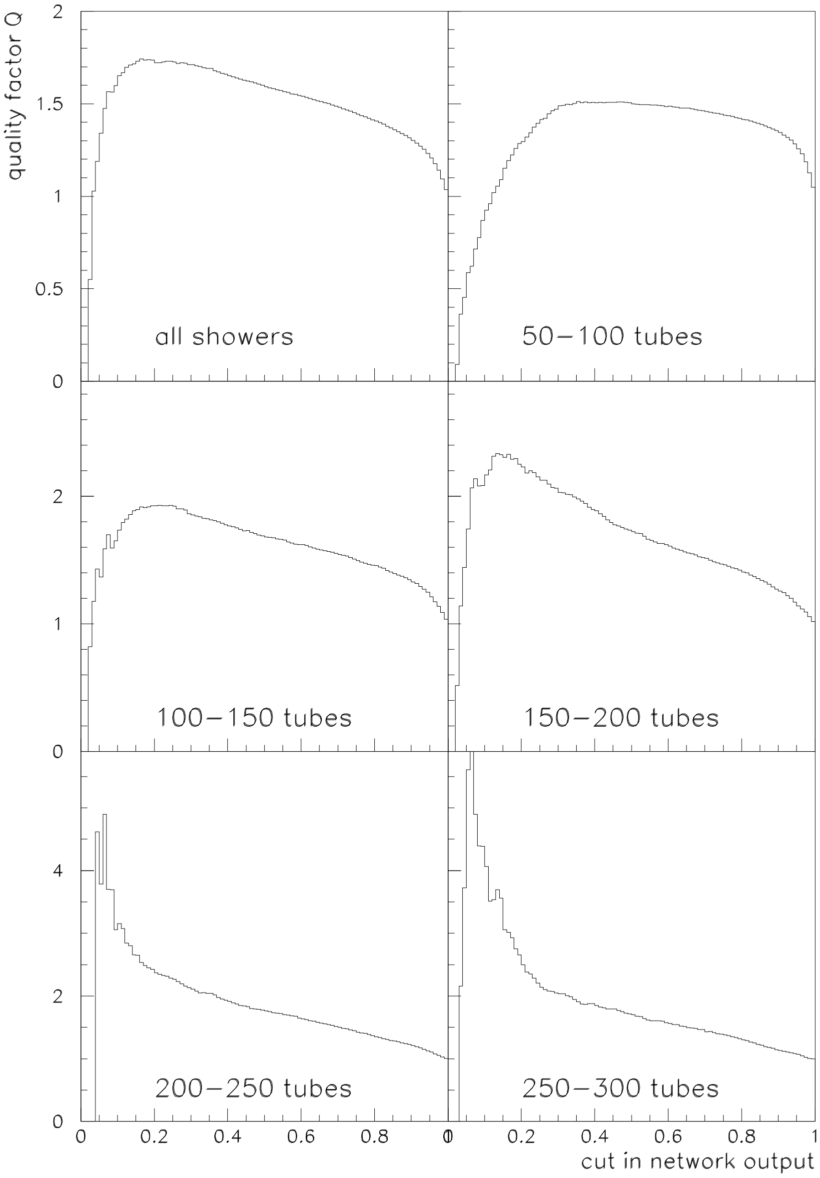

Figure 2: Quality factor Q as a function of the cut in the network output for different event sizes.

a weighted analysis, or have a "running" cut value. For the moment, the latter approach is chosen (a weighted analysis is upcoming). The minimum cut, however, is set to 0.1. This seems necessary as due to statistics, smaller cut values lead to a very small number (up to 0) of remaining protons and thus unnaturally high quality factors. Until the number of Monte Carlo proton showers has substantially increased or the network proton rejection has been evaluated with data, this cut cannot be optimized for large events, and a minimum of 0.1 avoids cutting in the sharp rise of the quality function. Figure 3 shows γ-efficiency, proton rejection, and resulting Q as a function of the event size. The quality factors for events with more than 300 tubes suffer from large statistical uncertainties. Nevertheless, there seems to be a *decrease* in the quality of the separation for very large events. A possible explaination which unfortunately cannot be tested with the current amount of MC data is that the number of large showers in the network training is too small to have the network really "care" about this type of event. In addition, the separation relies heavily on the comparison of the number of tubes in the two layers. For large event sizes, the hadron layer tube number saturates and thus the ratio becomes less meaningful.

As shown in Figure 4, the separation is a strong function of the core position. If we had "on/off" information, the network cut could be optimized for "off" events. However, I should point out once again that Milagro has an intrinsic γ/hadron separation for events on the pond that is not reflected in these plots; only 21 % of all proton showers that trigger Milagro have their core on the pond, whereas 31 % of all γ-showers are on-pond events. Together with the improvement in angular resolution, this leads to an increase in *sensitivity* by a factor of 1.5 for on-pond showers.

3 Important and Less Important Parameters

When training and testing nets on sub-sets of the 15 input parameters, it immediately becomes evident that the most important parameters for the separation are the tube and PE numbers in the two layers. In addition, adding the PE number of the "hottest" hadron layer tube increases the separation substantially. The 10 parameters giving the tubes and PEs in rings around the hottest cluster are almost negligible. However, it is interesting that these parameters give some separation when applied to a *one layer detector*. Figure 5 shows the network output for a 12-10-1 network trained solely on information given by the air shower layer (number of tubes, number

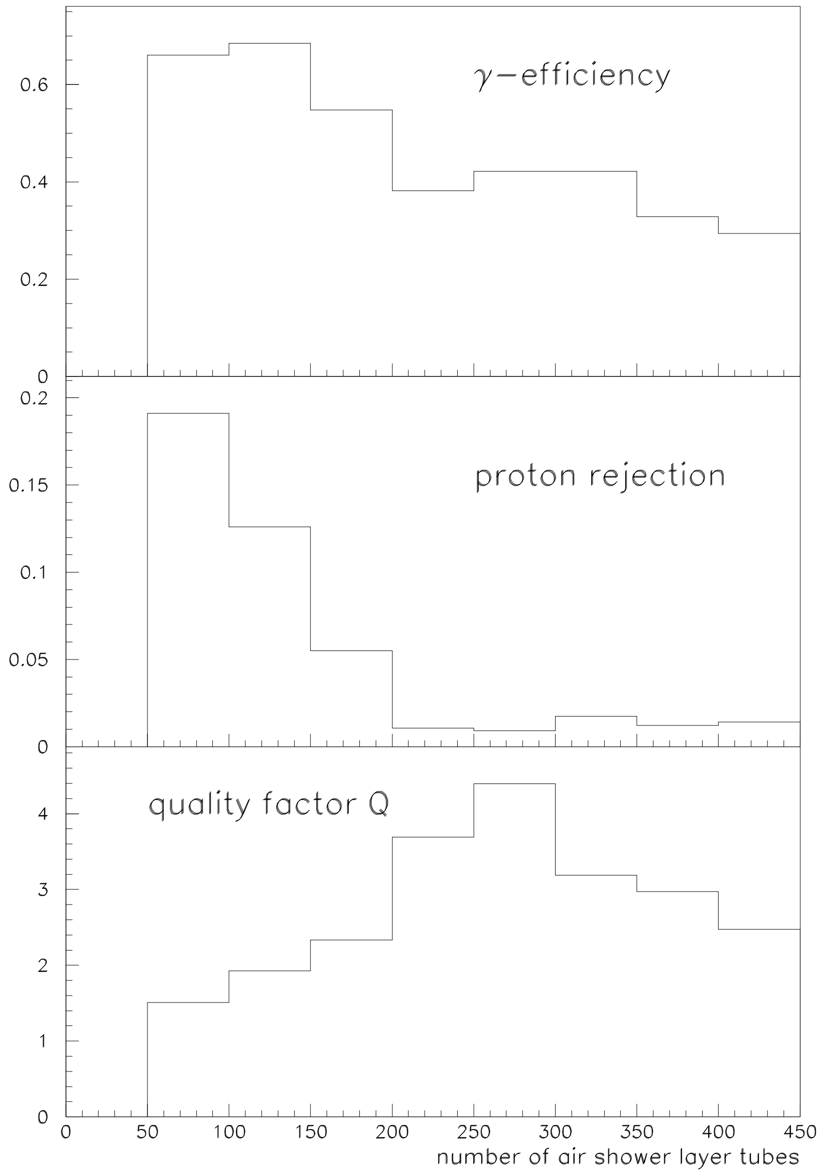

Figure 3: γ-efficiency, proton rejection, and quality factor Q for all events (core on and off).

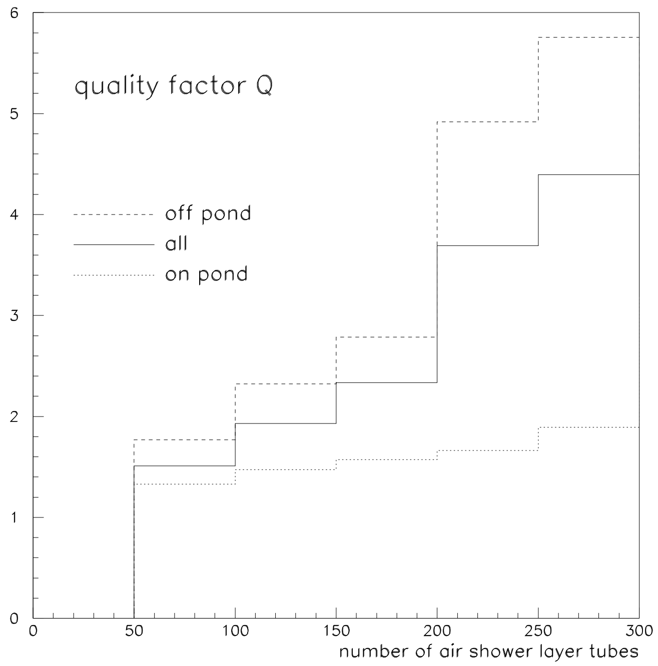

Figure 4: Quality factor Q for events with core on the pond as compared to off-pond events.

of PEs, rings). Again, Joe's independent MC sample is used for testing. The quality factor is small compared to a two-layer detector, but it shows the contribution of the rings. The amount of Milagrito MC data is insufficient to train and independantly test a network for Milagrito, but a superficial test gave about the same quality factor for Milagrito *when plotted vs. the number of tubes.*

4 Dead Tubes

One of the crucial questions is whether any separation based on the hadron layer is of any value when a large amount of hadron layer tubes are dead. Neural network analyses are supposed to be relatively stable when confronted with missing and corrupt data sets. Figure 6 shows the quality factor after excluding 57 hadron layer and 22 air shower layer tubes from the analysis. The network was not re-trained for the corrupt data. While the quality factor decreases for large events, it seems relatively unaffected for events with low tube numbers.

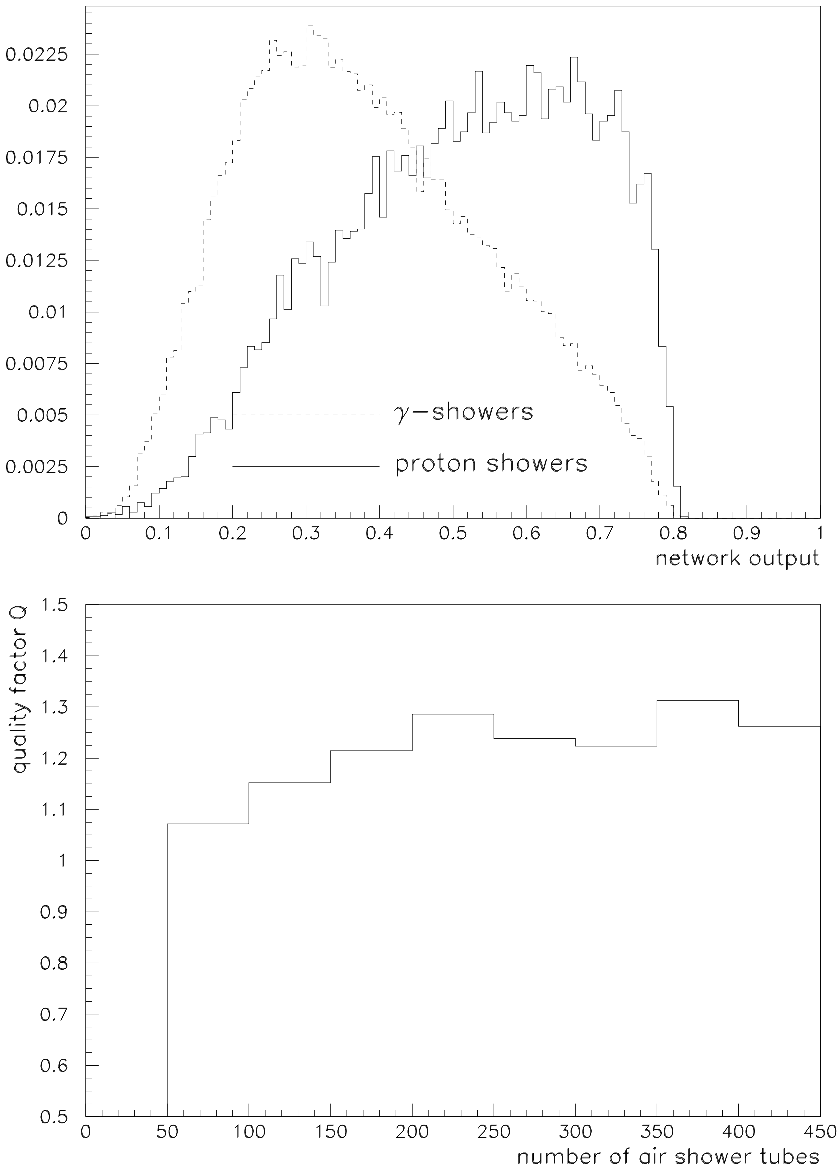

Figure 5: Network output and quality factor for Milagro with only one layer. The number of events is normalized to 1.

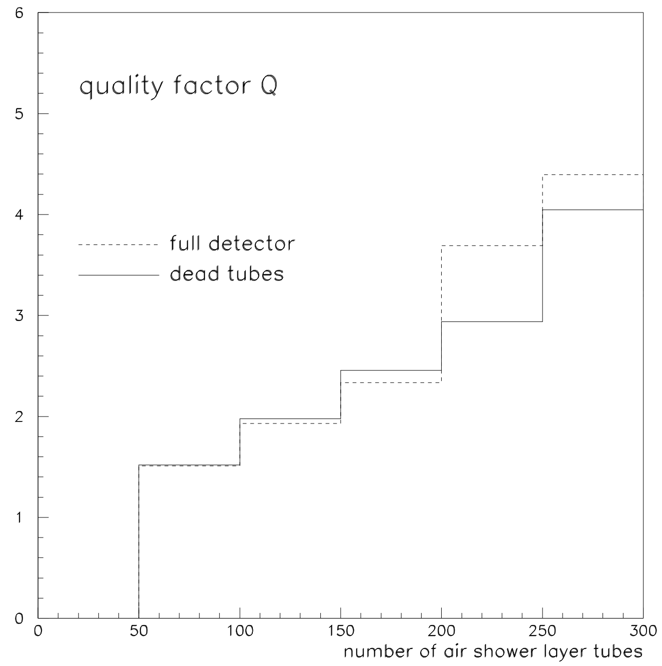

Figure 6: Quality factor for a fully operational detector and a detector with 22 dead air shower and 57 dead hadron layer tubes.