H1-06/05-619

VFPS momentum reconstruction using a neural net method

K. Černý[1], P. Van Mechelen[2], E. A. De Wolf[2]

[1]Charles University, Praha, Czech Republic
[2]Universiteit Antwerpen, Belgium

Abstract

A neural net method is proposed for the reconstruction of the proton momentum from the measured impact co-ordinates and slopes in the VFPS detectors. Resolutions of the reconstructed kinematic variables are determined and the effects of detector measurement errors and the uncertainty due to the beam spread and divergence at the interaction point are studied. A quality criterium is developed to distinguish well-reconstructed protons from background and the stability of the method w.r.t. varying beam conditions is studied.

1 Introduction

The Very Forward Proton Spectrometer (VFPS) [1] was installed in the H1 experiment during the 2003 shutdown of HERA. Its aim is to trigger on and measure the momentum of diffractively scattered protons originating from the H1 interaction point (IP). Because of the typically low energy losses ($x_{I\!P} \sim 0.01$) and scattering angles (-1 GeV$^2 \lesssim t$), the VFPS detectors were placed at 220 m distance from the main H1 detector so that the strong spectrometer effect of the horizontal HERA dipole magnets can be used to separate diffractively scattered protons from the nominal proton beam.

The hardware setup of the VFPS is very similar to that of the Forward Proton Spectrometer (FPS) [2] which was previously installed between 60 and 90 m downstream from the H1 detector. The VFPS consists of two Roman Pot stations, approximately 4 m apart and equipped with scintillating fibre detectors which approach the HERA proton beam horizontally from inside the HERA ring. The transverse co-ordinates of proton hits measured in both detector stations are combined into transverse intercepts and slopes with respect to the nominal proton beam at a location halfway between the detector stations.

While the reconstruction code for hit finding, local track (in one station) and global track (combining two stations) reconstruction is common with the FPS, one needs to use a different approach for the determination the proton momentum from the intercepts and slopes. This is partly due to strong non-linear effects present in the HERA beam optics between the H1 interaction point and the VFPS. More importantly, however, a strong inbalance exists between the horizontal and vertical energy dispersion of the proton beam: while a strong dispersion exists in the horizontal co-ordinate, the vertical co-ordinate changes very little within the range of energy deviations relevant for the VFPS. The approach used for the FPS, exploiting both co-ordinate planes for the reconstruction of $x_{I\!P}$ is therefore less adequate in the case of the VFPS.

Regardless of the chosen approach for the reconstruction of the proton momentum, an efficient parameterization of the beam optical functions is required. Although the H1SIM package provides a detailed simulation of the proton trajectories, these routines cannot be used in the reconstruction code because of the large computing time needed for the simulation of one proton trajectory. This computing time would be multiplied by a large factor when testing different assumptions for the proton momentum in a fit to the measured intercepts and slopes. In this note a neural net method is proposed for the parameterization of the dependence of intercepts and slopes at the VFPS location on the initial proton momentum. This method has the added benefit that it is equally possible to parameterize the inverse relation, with the initial proton kinematics depending on the measured intercepts and slopes, thereby solving the reconstruction problem without the need for a complicated fit procedure.

This note is organized as follows. Section 2 discusses the relation between the proton momentum at the interaction point and the intercepts and slopes of the proton trajectory at the location of the VFPS. Section 3 gives a short overview of the neural net method used to parameterize the above relation. Section 4 discusses the different Monte Carlo samples that are used and the resulting parameterizations in both directions are presented in Sec. 5 together with the resolutions of the reconstructed kinematical variables. Finally, Sec. 6 proposes a quality criterium for the reconstruction fit and the effect of varying the proton beam emittance on the obtained resolutions is studied in Sec. 7. A conclusion is given in Sec. 8.

2 VFPS beam optics

The subject of this note is the reconstruction of the proton momentum based on the measurement of impact co-ordinates and slopes in the VFPS. The relation between both is given by the HERA beam optics and can be described in a general way by a matrix equation:

$$\mathbf{X} = \mathbf{T} \cdot (\mathbf{X}_0 + \boldsymbol{\alpha}). \tag{1}$$

The variables in this equation are defined as follows:

- $\mathbf{X}$ is a 4×1 column vector containing the positions (x,y) and slopes (x',y') of the proton at a location halfway between the two VFPS Roman Pot detectors;
- $\mathbf{T}$ is a 4×5 transport matrix describing the optical functions between the IP and the VFPS;

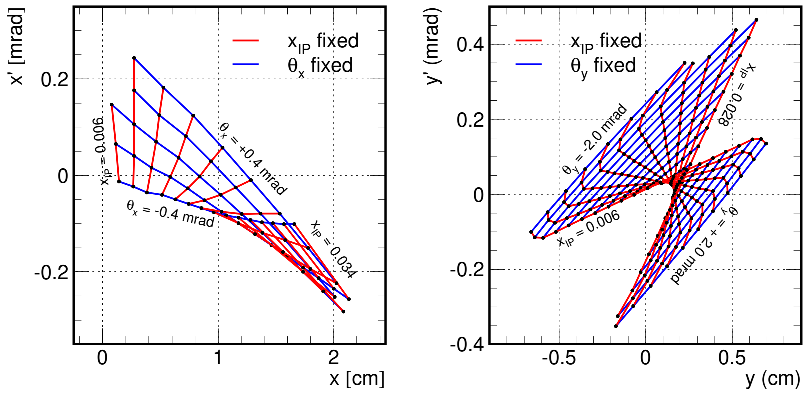

Figure 1: Lines of constant relative energy loss ($x_{I\!P}$) and scattering angles (θ_x, θ_y) are drawn in the horizontal and vertical planes of slope versus position of the proton halfway between the VFPS stations. Lines of constant $x_{I\!P}$ are drawn in steps of 0.002 and lines of constant $\theta_x(\theta_y)$ are drawn in steps of 0.2 mrad. Only points inside the VFPS acceptance are displayed.

- $\mathbf{X}_0$ is a 5×1 column vector containing the positions (x_0,y_0), slopes (x'_0,y'_0) and relative energy deviation (ξ_0) at the IP *before* the diffractive interaction;
- finally, $\boldsymbol{\alpha}$ is a 5×1 column vector describing the diffractive interaction and containing zeroes for the change in position, the horizontal and vertical scattering angles (θ_x and θ_y) for the change in slopes and the relative energy gain ($-x_{I\!P}$) for the change in relative energy deviation.

In order to determine the proton momentum, one needs to solve Eq. (1) for $\boldsymbol{\alpha}$.

No attempt is made to include information on $\mathbf{X}_0$ measured by the main H1 detector. Neglecting the beam spread and divergence at the IP and thus assuming $\mathbf{X}_0 = 0$, results in additional smearing of the measured co-ordinates which has to be included in the uncertainty on $\mathbf{X}$. In this case Eq. (1) reduces to a (over-determined) system with 4 equations and three unknowns:

$$\mathbf{X} = \tilde{\mathbf{T}} \cdot \tilde{\boldsymbol{\alpha}}, \tag{2}$$

where $\tilde{\mathbf{T}}$ is now a 4×3 matrix derived from $\mathbf{T}$ and $\tilde{\boldsymbol{\alpha}}$ is a 3×1 column containing the scattering angles (θ_x and θ_y) and the relative energy gain ($-x_{I\!P}$).

In many applications the transport matrices $\mathbf{T}$ and $\tilde{\mathbf{T}}$ can be taken to be constant such that Eq. (2) is a simple linear system of four equations and three unknowns. Moreover, in the absence of sextupole and higher multipole magnets, the horizontal and vertical co-ordinates decouple, such that the problem is reduced to two independent systems of two equations, that can be solved for (θ_x, $x_{I\!P}$) and (θ_y, $x_{I\!P}$), respectively, yielding two independent determinations of $x_{I\!P}$.

In the case of the VFPS, however, where differences in the measured coordinates of a few tens of microns are relevant, the linear approximation, which is valid for small energy deviations, positions and slopes, is no longer adequate. Strong non-linear effects, originating from the magnets in the cold section of HERA, need to be taken into account, effectively making the matrix elements of $\mathbf{T}$ (and $\tilde{\mathbf{T}}$) dependent on the proton momentum. Moreover the presence of sextupole magnets will introduce a coupling between the vertical and horizontal planes. Figure 1 shows the grid with lines of constant energy loss and

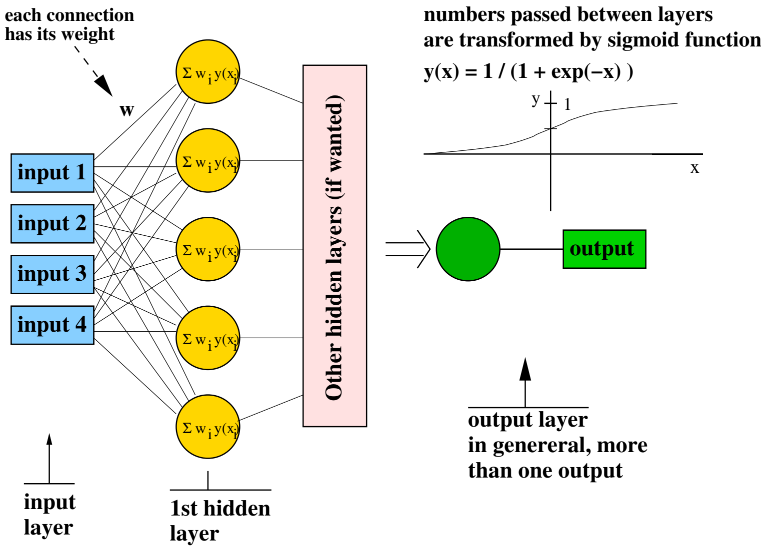

Figure 2: Schematic structure of the Multi Layer Perceptron network. In this case the structure consists of 4 inputs, 5 neurons in the first hidden layer, possible other hidden layers and one output.

scattering angles projected onto the horizontal and vertical planes of slope versus intercept at the VFPS location. As can be seen, the deviations from a rectangular (linear) grid are large. The change in position when keeping the scattering angle fixed is also much smaller in the vertical direction than in the horizontal direction, especially for small scattering angles. This is because the strong energy dispersion only exists in the horizontal plane.

3 MLPfit Neural Net

In this work we concentrate on the *Multi-Layer Perceptron fit* (MLPfit) package [3], but in general any other fitting method could be used. The MLPfit has the advantage of providing a better fit for the same number of parameters than e.g. a simple polynomial fit. Moreover, MPLfit has been implemented in PAW and is therefore a very user-friendly fitting tool.

The basic scheme of the MLP network is sketched in Fig. 2. Input quantities are processed through successive layers of “neurons”. There is always at least an input layer, with a number of neurons equal to the number of variables of the problem, and an output layer where the perceptron response is made available, with a number of neurons equal to the desired number of quantities computed from the inputs. The layers in between are called “hidden” layers. Each neuron of a layer other than the input layer first calculates its input value as a linear combination of the outputs of the neurons of the previous layer plus a bias. The coefficients of the linear combinations and the biases are called “weights”. They are usually determined during training of the network where a set of examples is used to minimize the (Euclidean) norm of the desired output. Neurons in the hidden layer then calculate a non-linear function of their inputs. In MLPfit, the non-linear function is the sigmoid function $y(x) = (1+e^{-x})^{-1}$. Any continuous function of one or more variables can be approximated by a linear combination of sigmoid functions. Various interfaces for the MLPfit package are available; in this work we stick to the PAW interface.

Several points deserve some comments.

- The MLP networks are standard "feed-forward" networks with a maximum number of hidden layers equal to 2. The maximum number of neurons in each hidden layer is 100. The maximum number of inputs and outputs is limited to 29 when data is read from a PAW N-tuple.
- For the training (or learning) phase, the input and desired output values have to be known simultaneously. The learning procedures always attempt to minimize $E = \Sigma_p \, \delta_p$ where p runs over all examples and $\delta_p = \omega_p(\varrho_p - \epsilon_p)^2$, where ω_p is weight of one example (in our case always equal 1), ϱ_p is the response (output value) of the network (this is actually the expression that contains the inter-neuron weights) and ϵ_p is the expected (or desired) output. One minimization iteration is called an "epoch".
- The training data should not coincide with the data that are used to test the performance of the net.
- The inputs are recommended to be rescaled to the interval $[0, 1]$ in order not to lose precision due to the sigmoid transformation. It is convenient to also rescale the output of the network to some reasonable interval like $[0, 1]$ as the output value is a linear combination of sigmoid functions.
- The structure of the network and the number of learning epochs is a matter of proper choice and testing. A network structure that is too small compared to number of examples leads to inaccurate results, while a network that is too large with insufficient statistics may result in over-fitting.

4 Monte Carlo simulation

By means of Monte Carlo simulation one gets access to the full information on the relation between the diffractive kinematics $\tilde{\boldsymbol{\alpha}}$ and the intercepts and slopes $\mathbf{X}$ of the proton track detected by the VFPS. In this study various Monte Carlo samples are produced in order to investigate the effect of different uncertainty sources.

4.1 Uncertainties on the measured coordinates and slopes

As explained in Sec. 2, no attempt is made to measure the initial offset and tilt of the proton entering the diffractive interaction. The beam spread and divergence at the IP will therefore result in a smearing of the positions and slopes of protons at the VFPS, which will limit the attainable resolution on the proton momentum. One can, e.g., never hope to measure the scattering angle with a resolution that is better than the beam spread at the IP. The resulting uncertainties on positions and slopes are highly correlated. Table 1 lists the expected (co-) variances calculated in the linear beam optics approximation (valid for the nominal proton beam) in the horizontal and vertical plane.

A second source of uncertainty is the fibre detector resolution itself. Figure 3 shows the difference between simulated and reconstructed intercepts and slopes as obtained from an H1SIM simulation. The resolutions are 57 μm and 26 μrad for the intercepts and slopes, respectively, in both planes.

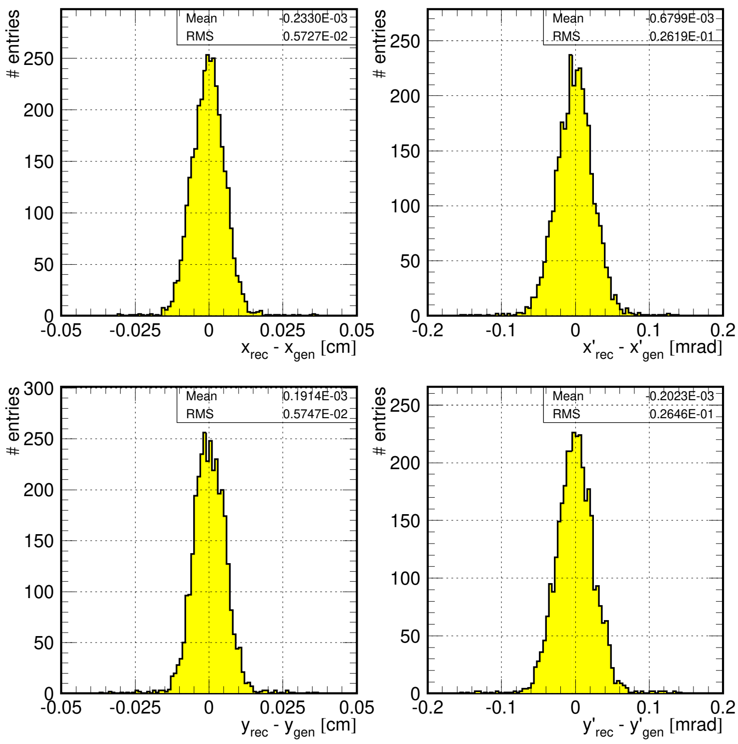

Figure 3: Differences between generated and reconstructed intercepts and slopes are plotted, taking into account the VFPS fibre detector resolution as implemented in H1SIM.

	$\sigma_{x(y)}$ $[\mu m]$	$\sigma_{x'(y')}$ $[\mu rad]$	$\rho_{xx'(yy')}$ $[\%]$
x_0	118	48	0
y_0	320	178	0
x	404	30	-86
y	2027	124	+97

Table 1: Standard deviations and correlation coefficients for the horizontal and vertical positions and slopes of a proton at the IP (x_0, y_0) and at the location of the VFPS (x, y), caused by the beam spread and divergence at the IP.

4.2 Data samples

Monte Carlo samples are obtained by means of the forward beam line simulation as implemented in H1SIM. The range in generated momentum of the scattered proton is defined by $|t| < 4$ GeV^2 and $0 < x_{I\!P} < 0.04$. The following Monte Carlo samples are used and need to be carefully distinguished:

1. A “clean” sample is obtained by switching *off* the beam spread and divergence at the IP, by switching *off* the simulation of all physics processes (except the energy loss process) and by taking the intercepts and slopes at the location of the VFPS *before* any detector simulation. The reason why the energy loss process is kept is technical and has to do with purifying the sample from events in which some background generating phenomena occur (like e.g. protons leaving the beam pipe). This sample is labelled with “CLEAN”.

2. A second sample is obtained by switching *on* the beam spread and divergence at the IP and by smearing the intercepts and slopes at the location of the VFPS according to the resolutions found in Fig. 3. A full simulation of the detector resolutions using H1SIM and H1REC is not possible at this point because physics processes are still switched off and because this would introduce errors in the reconstruction of intercepts and slopes which would deteriorate the quality of the parameterization. This sample is labelled with “VTX+DET”.

3. A sample with all physics processes switched *on* is also obtained. In this case the beam spread and divergence is always present and the detector resolution is taken into account by a full simulation and reconstruction of the VFPS. This sample should resemble the real data and is denoted with “ALL”.

4. For the purpose of testing the stability of the method, a sample with all physics processes switched on was produced using a larger value for the beam emittance, resulting in a larger spread and divergence of the proton trajectories at the interaction point. The beam emittance was changed from $\epsilon_{orig} = 5.69 \times 10^{-9}$ m to $\epsilon_{new} = \epsilon_{orig} \times 1.5$. This sample is denoted with “LRG_EMI”

5 Parameterization of the beam optics

5.1 VFPS coordinates as function of proton kinematics, $\mathbf{X}(\tilde{\boldsymbol{\alpha}})$

The mapping of proton kinematics to VFPS intercepts and slopes,

$$\tilde{\boldsymbol{\alpha}} = \begin{pmatrix} \theta_x \\ \theta_y \\ x_{I\!P} \end{pmatrix} \longrightarrow \mathbf{X} = \begin{pmatrix} x \\ x' \\ y \\ y' \end{pmatrix} \tag{3}$$

is simulated in H1SIM. In order to have faster and more convenient access to this information, the results of H1SIM are parameterized using MLPfit, yielding four functions of the proton kinematics:

$$\mathbf{X}_{nn}(\tilde{\boldsymbol{\alpha}}) = \begin{pmatrix} x_{nn}(\theta_x, \theta_y, x_{I\!P}) \\ x'_{nn}(\theta_x, \theta_y, x_{I\!P}) \\ y_{nn}(\theta_x, \theta_y, x_{I\!P}) \\ y'_{nn}(\theta_x, \theta_y, x_{I\!P}) \end{pmatrix} \tag{4}$$

For this task two independent CLEAN data samples are used for training and for comparison. Figure 4 shows the deviation of the neural net parameterization from the generated intercept and slopes. The resolutions on x, x', y and y' are 25 μm, 1.2 μrad, 19 μm and 0.7 μrad, respectively. Comparing this to the detector resolutions of Fig. 3 one may conclude that the fit is able to describe the proton trajectory with adequate precision. Appendix A lists the training parameters of the neural net used for this parameterization.

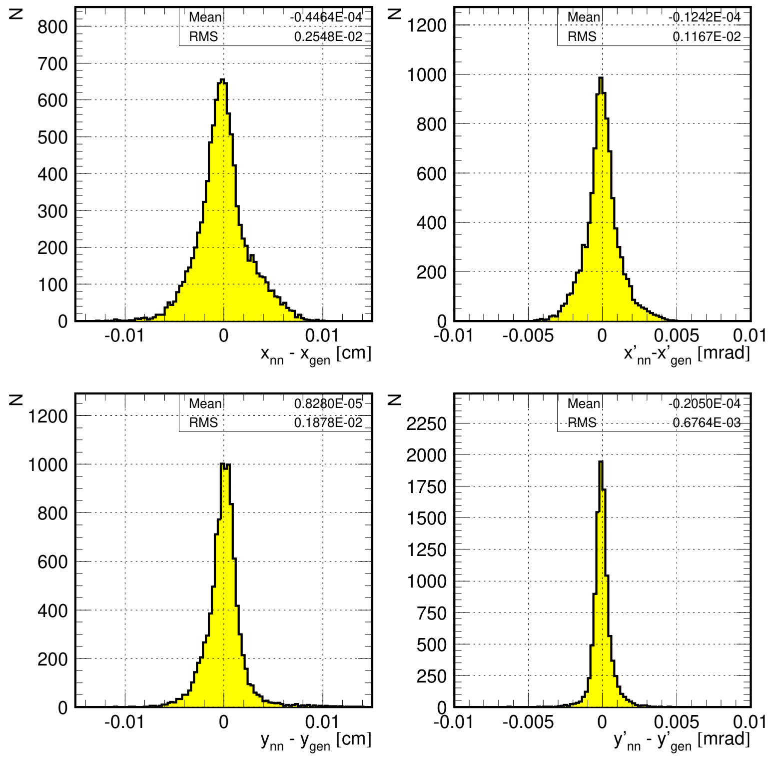

Figure 4: Differences between generated proton track intercepts and slopes and fitted (MLPfit) values are shown. No vertex or detector smearing is used for training and comparison.

5.2 Proton kinematics as function of VFPS coordinates, $\tilde{\boldsymbol{\alpha}}(\mathbf{X})$

The inverse task amounts to the parameterization of the actual reconstruction of diffractive proton kinematics from the VFPS measurement

$$\mathbf{X} = \begin{pmatrix} x \\ x' \\ y \\ y' \end{pmatrix} \longrightarrow \tilde{\boldsymbol{\alpha}} = \begin{pmatrix} \theta_x \\ \theta_y \\ x_{I\!P} \end{pmatrix} . \tag{5}$$

The result of this parameterization are three functions of the intercepts and slopes,

$$\tilde{\boldsymbol{\alpha}}_{nn}(\mathbf{X}) = \begin{pmatrix} \theta_{x,nn}(x,x',y,y') \\ \theta_{y,nn}(x,x',y,y') \\ x_{I\!P,nn}(x,x',y,y') \end{pmatrix} . \tag{6}$$

As the $\mathbf{X}_{nn}(\tilde{\boldsymbol{\alpha}})$ parameterization in the previous section was trained with the CLEAN sample, it is natural to try to use the CLEAN sample for the inverse parameterization as well. Figure 5 displays reconstruction errors on $x_{I\!P}$, $|t|$ and ϕ as functions of $x_{I\!P}$ and $|t|$ for CLEAN-trained networks applied to the CLEAN and VTX+DET sample. Details on the training parameters are again listed in App. A. The natural expectation is confirmed: as the level of smearing of the control sample increases, the reconstruction becomes less precise. One may wonder whether training with already smeared samples could describe the smeared data better. The answer is yes: Fig. 6 shows the results of a VTX+DET-trained network applied to the VTX+DET sample.

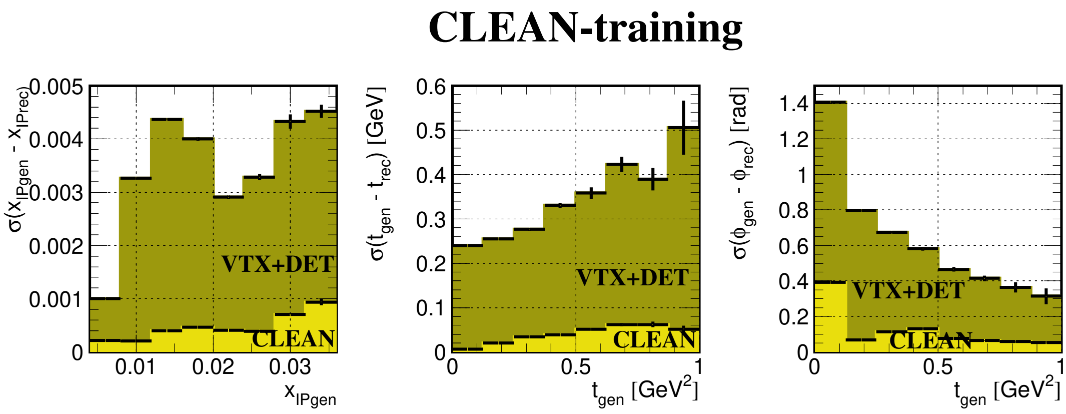

Figure 5: The main kinematic dependencies of the errors on the reconstructed proton kinematics are shown. Here, a CLEAN-trained network is applied to an (independent) CLEAN (light yellow) and a VTX+DET (dark yellow) sample.

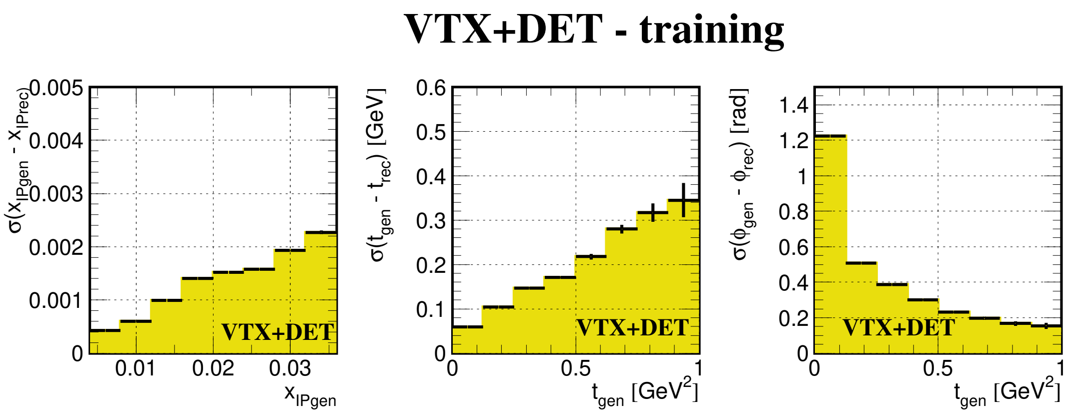

Figure 6: The main kinematic dependencies of the errors on the reconstructed proton kinematics are shown. Here, a VTX+DET-trained network is applied to an (independent) VTX+DET sample.

Since the resolution clearly improves when using a more realistic sample for training of the neural net, one may think that the ALL data sample is preferred for the training of the neural net which is to be used for the reconstruction of real data. However, this sample cannot easily be purified from background events, leading to bad parameterizations. Here, we distinguish background events where the measured intercepts and slopes have no correlation whatsoever with the proton kinematics, from the VTX+DET events where the the relation between co-ordinates and kinematics is merely smeared out. The VTX+DET sample is therefore the most advanced sample for training and this sample will be used in the subsequent sections.

6 Quality assessment

6.1 Definition of a quality estimator

In order to remove background events when applying VTX+DET-trained networks to the ALL data sample, a quality estimator for the reconstructed kinematics needs to be introduced.

As already discussed, the beam spread and divergence at the IP and the finite detector resolution results in the smearing of the reconstructed intercepts and slopes, $\mathbf{X}_{rec}$. A quality estimator can then be defined as

$$Q = {}^{T}(\mathbf{X}_{rec} - \mathbf{X}_{nn}(\tilde{\boldsymbol{\alpha}}_{gen})) \cdot (\mathbf{Cov}_{nn}(\tilde{\boldsymbol{\alpha}}_{gen}))^{-1} \cdot (\mathbf{X}_{rec} - \mathbf{X}_{nn}(\tilde{\boldsymbol{\alpha}}_{gen})), \tag{7}$$

where $\mathbf{X}_{nn}(\tilde{\boldsymbol{\alpha}}_{gen})$ is obtained from the CLEAN sample as in Sec. 5.1 and $\mathbf{Cov}_{nn}(\tilde{\boldsymbol{\alpha}}_{gen})$ represents the 4×4 covariance matrix describing the vertex and detector smearing, which actually also depends on the proton kinematics. The dependence of $\mathbf{Cov}_{nn}$ on $\tilde{\boldsymbol{\alpha}}_{gen}$ also needs to be parameterized using the CLEAN sample.

If one assumes that the smearing of $\mathbf{X}_{rec}$ is Gaussian, then this quality estimator must follow a χ^2 distribution. Indeed, the left plot of Fig. 7 displays the Q distribution as obtained from the VTX+DET sample together with the theoretical χ^2 shape for four degrees of freedom.

The generated kinematics is however unknown in real data processing. Therefore, one has to reformulate Q using $\tilde{\boldsymbol{\alpha}}_{nn}(\mathbf{X}_{rec})$ instead of $\tilde{\boldsymbol{\alpha}}_{gen}$, as sketched in Fig. 8. However, the errors on the "reobtained" $\mathbf{X}$ may now depart from a Gaussian shape. Indeed, Fig. 9 shows the distribution of the difference between $\mathbf{X}_{rec}$ and $\mathbf{X}_{nn}(\tilde{\boldsymbol{\alpha}}_{nn}(\mathbf{X}_{rec}))$ compared to Gaussian distributions of the same mean and variance. On the right plot of Fig. 7 one can see the Q distribution using $\tilde{\boldsymbol{\alpha}}_{nn}(\mathbf{X}_{rec})$. The shape of the Q distribution has changed compared to the original one: although the spread is similar as before and the mean has even decreased, the tail is clearly extending to much larger values of Q. Nevertheless, it is natural to expect larger Q values for events where $\mathbf{X}_{rec}$ and $\mathbf{X}_{nn}(\tilde{\boldsymbol{\alpha}}_{nn}(\mathbf{X}_{rec}))$ differ "too much". This allows to reject $\mathbf{X}_{rec}$ entries which do not correspond to the mapping according to Fig. 1.

6.2 Application of the quality criterion

The selection of high quality events by requesting large Q values inevitably leads to a reduction in the event number. The aim is to apply an efficient cut that is able to reject background events while keeping the loss in statistics reasonably low. Figure 10 shows the relative event loss as a function of the Q cut value for the VTX+DET sample which

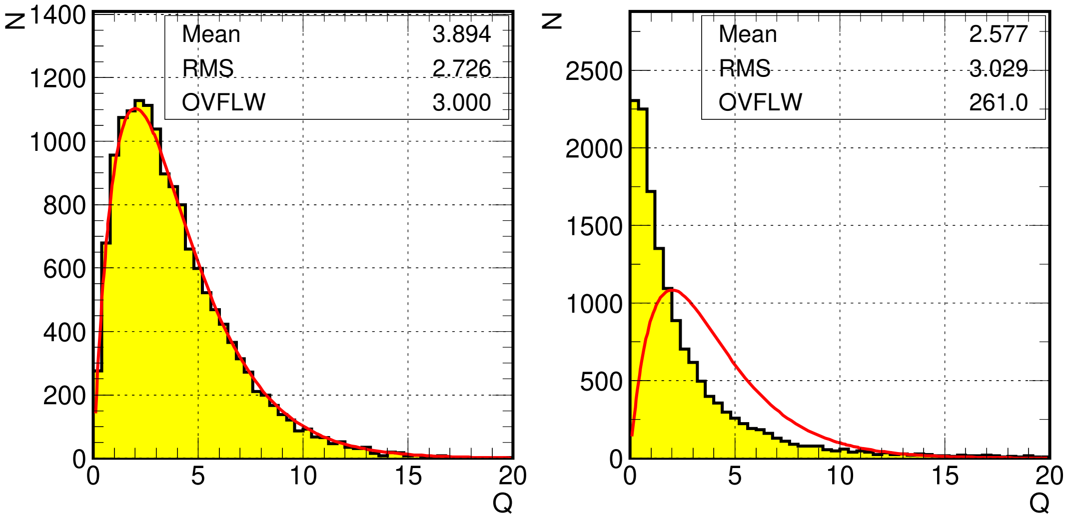

Figure 7: The Q distributions as obtained using $\mathbf{X}_{nn}(\tilde{\boldsymbol{\alpha}}_{gen})$ (left histogram) and using $\mathbf{X}_{nn}(\tilde{\boldsymbol{\alpha}}_{nn}(\mathbf{X}_{rec}))$ (right histogram) for data including the beam spread and divergence at the IP and the detector resolution (VTX+DET). The red curve is the theoretical shape of a χ^2 distribution for four degrees of freedom.

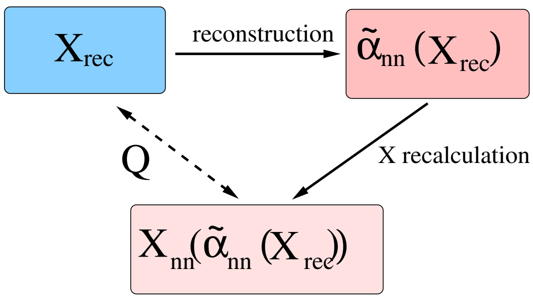

Figure 8: For the calculation of a quality criterium in real data, one first needs to reconstruct the proton kinematics using the reconstructed co-ordinates and a VTX+DET trained network. Then one can use the reconstructed kinematic variables to recalculate the co-ordinates corresponding to this kinematics according to a beam optics parameterization based on a CLEAN training sample. The quality criterium can then be obtained by taking the "χ^2" difference between both sets of co-ordinates.

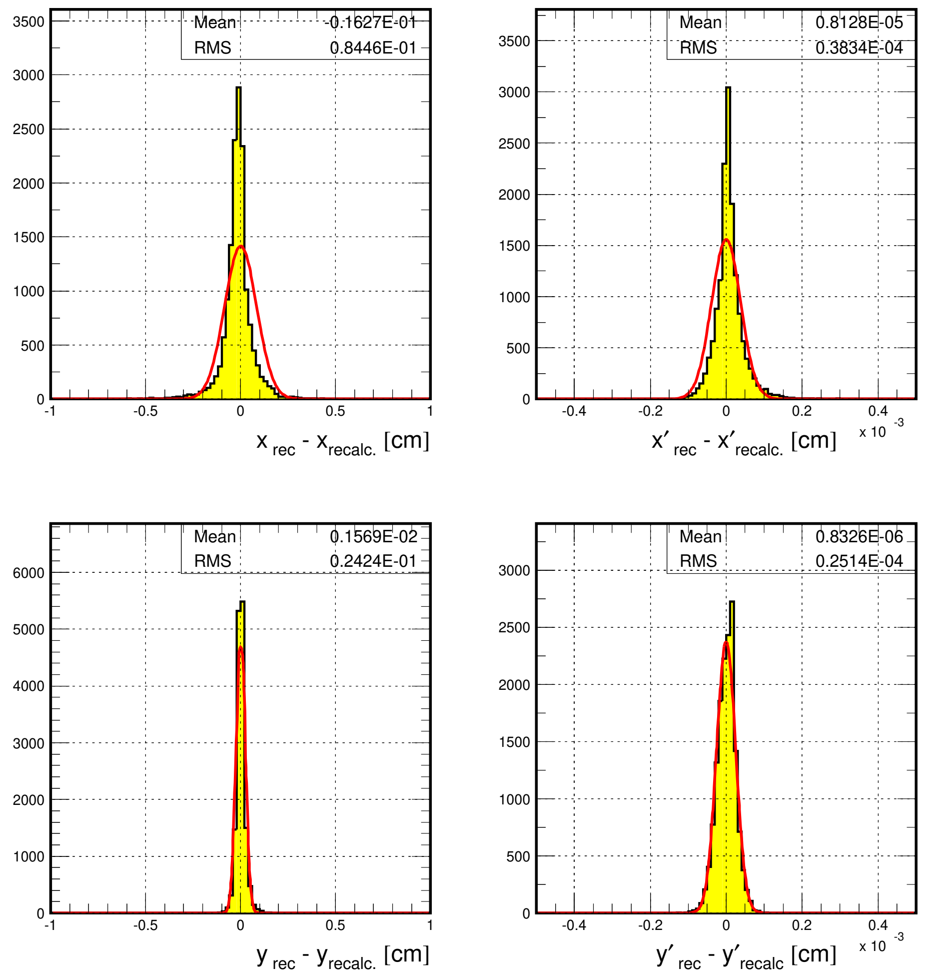

Figure 9: Differences between simulated and reobtained co-ordinates are calculated using the scheme sketched in Fig. 8. The distributions deviate from Gaussian distributions with the same mean and variance.

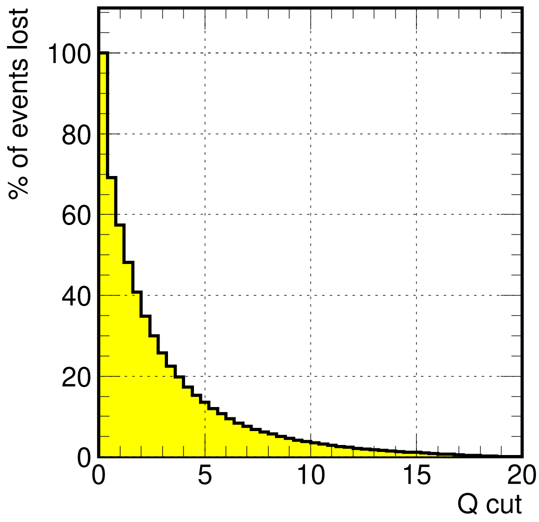

Figure 10: Relative event loss in the VTX+DET data sample as a function of the Q cut.

does not contain any background. A 10% loss of events is achieved with the selection $Q < 7$.

Figure 11 (top) shows the hit distribution in the x'-x and y'-y planes as obtained from the ALL data sample. Due to possible interactions with material inside the beam-pipe, wrong demultiplexing (see [1]), etc. background hits are produced in the VFPS detectors. The power of the Q cut is very well illustrated by Fig. 11 (bottom) where the same plots are obtained after reconstruction and selection of those events with $Q < 7$. The background hits caused by wrong demultiplexing, which show up as additional bands above and below the main hit area have completely disappeared.

It has been checked that events which have reconstructed intercepts and slopes too far away from the generated ones overlap for $\sim 60\%$ with the events that are rejected by the Q selection criterion, proving that this method does indeed remove wrongly reconstructed events.

The final error dependencies for the ALL data sample, after applying the $Q < 7$ cut, are shown in Fig. 12. The error on $x_{I\!P}$ varies more or less linearly from 0.0005 to 0.002 and is always less than 7.5%. The error on t varies from 0.06 to ~ 0.2 GeV^2 and the error on ϕ reaches 0.2 rad at large $|t|$.

7 Stability of the reconstruction method

As discussed in Sec. 5.2, it is important to train the neural net using a data sample which includes the uncertainties on the reconstructed intercepts and slopes. It is therefore important to investigate the performance of this reconstruction method when the actual uncertainties deviate from the one used in the training sample. E.g. the beam quality may vary between lumi runs and a variation in the beam emittance can be expected. This would result in a different beam spread and divergence at the interaction point and therefore yield a different uncertainty on the reconstructed co-ordinates.

In order to study the stability the reconstruction method as function of the beam quality, a data sample with larger beam emittance (LRG_EMI) is generated. This data is then reconstructed using a network trained with the VTX+DET sample with the nominal

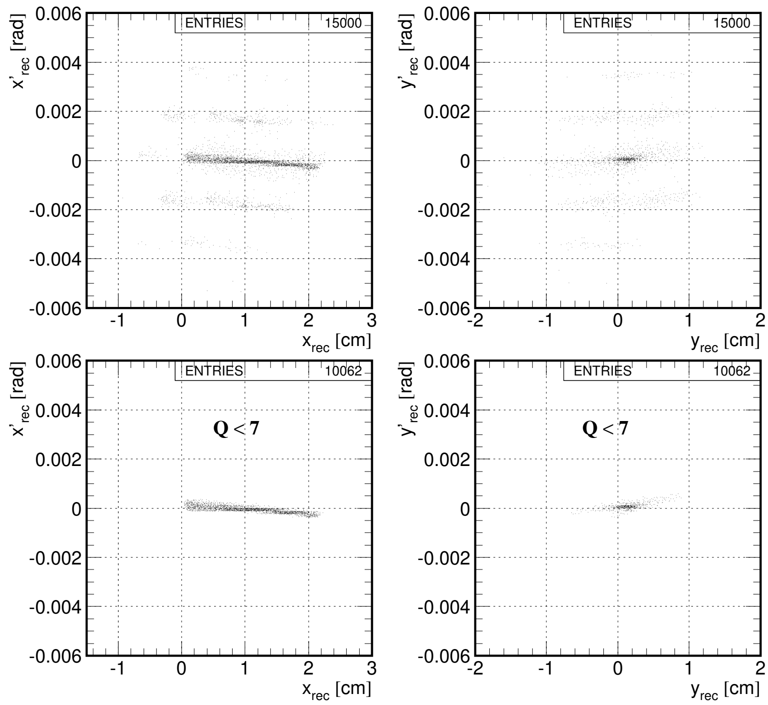

Figure 11: Hit distributions in the x'-x and y'-y planes for the ALL data sample before (top) and after (bottom) selecting events with $Q < 7$.

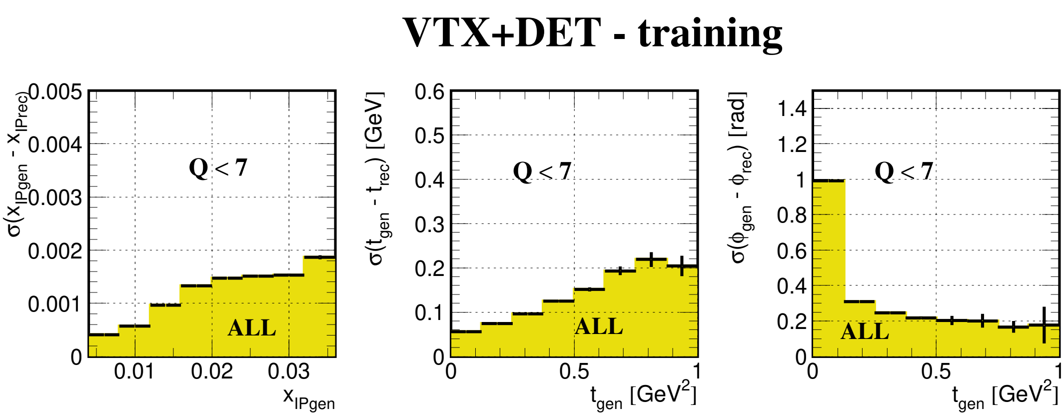

Figure 12: The main kinematic dependencies of the errors on the reconstructed proton kinematics are shown for the ALL data sample obtained with a VTX+DET trained network and after selecting events with $Q < 7$.

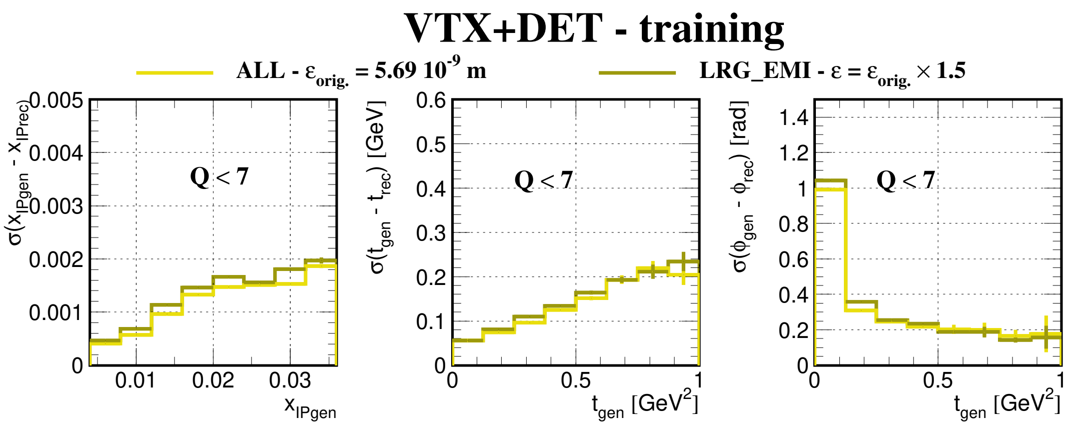

Figure 13: A comparison is shown of the error dependencies of the reconstructed proton kinematics for samples generated with a different beam emittance but reconstructed by a neural net parameterization based on the same (smaller) beam emittance. The selection $Q < 7$ is applied on both data samples.

beam emittance. Figure 13 shows a comparison of the error dependencies between an ALL sample using the original emittance (light yellow) and an ALL sample based on the new emittance (dark yellow). Both error dependencies were obtained with the quality cut $Q < 7$. From the figure it can be concluded that changing the beam emittance by 50% has very little influence on the error dependencies and that the reconstruction method is therefore fairly stable against such effects.

8 Conclusion

A neural net method is proposed to obtain a parameterization of the beam transport functions describing the relation between the proton kinematics at the H1 interaction point and the impact co-ordinates and slopes at the location of the VFPS.

Using this method, the parameterization of the dependence of the impact co-ordinates and slopes as a function of the proton kinematics yields an accuracy of $\delta x = 25\ \mu$m, $\delta x' = 1.2\ \mu$rad, $\delta y = 19\ \mu$m and $\delta y' = 0.7\ \mu$rad for the horizontal position and slope and the vertical position and slope, respectively. This accuracy is at least a factor of two better than the detector resolution as obtained from H1SIM (57 μm and 26 μrad for the positions and slopes, respectively). The uncertainty introduced by the neural net parameterization is therefore negligible w.r.t. the (variable) combined uncertainty originating from the beam spread and divergence at the interaction point and the detector resolution.

For the inverse relation, the parameterization of the proton kinematics as a function of impact co-ordinates and slopes, yields an accuracy of $\delta x_{I\!P}/x_{I\!P} < 4\%$ for the energy loss, $\delta|t| < 0.06$ GeV^2 and $\delta\phi \to 0.05$ rad for large enough $|t|$. Taking the full uncertainty from the beam and detectors into account these numbers become $\delta x_{I\!P}/x_{I\!P} < 7.5\%$ for the energy loss, $\delta|t| = 0.06$–0.2 GeV^2 and $\delta\phi \to 0.2$ rad for large enough $|t|$, which demonstrates that the contribution from the neural net parameterization to the overall uncertainty is much smaller than the combined effect of vertex smearing and detector resolution.

In order be able to estimate how well the momentum is reconstructed, a quality criterion is proposed. When applying this criterion to a fully simulated and reconstructed sample, background hits can be removed efficiently, while keeping the loss of well reconstructed events below 10%.

A Training parameters

A.1 Training of $\mathbf{X}(\tilde{\boldsymbol{\alpha}})$

Table 2 summarizes the CLEAN-training settings for the $\mathbf{X}_{nn}(\tilde{\boldsymbol{\alpha}})$ parameterization.

CLEAN training	x_{nn}	x'_{nn}	y_{nn}	y'_{nn}
in - 1^{st} hidden - 2^{nd} hidden - out	3-10-0-1	3-12-0-1	3-8-0-1	3-13-0-1
# of epochs	1000	1000	1000	1000
# of training samples	15000	15000	15000	15000

Table 2: Structure of the networks for the $\mathbf{X}_{nn}(\tilde{\boldsymbol{\alpha}})$ parameterization.

A.2 Training of $\tilde{\boldsymbol{\alpha}}(\mathbf{X})$

Table 3 lists the training settings for the $\tilde{\boldsymbol{\alpha}}_{nn}(\mathbf{X})$ parameterization for CLEAN and VTX+DET trainings.

CLEAN, VTX+DET training	$x_{I\!P,nn}$	$\theta_{x,nn}$	$\theta_{y,nn}$
in - 1^{st} hidden - 2^{nd} hidden - out	4-17-0-1	4-8-0-1	4-11-0-1
# of epochs	2000	2000	1000
# of training samples	14000	14000	14000

Table 3: Structures of the networks for $\tilde{\boldsymbol{\alpha}}_{nn}(\mathbf{X})$ parameterization for training with CLEAN and VTX+DET data samples.

Acknowledgement

This work was supported by a bilateral grant for scientific cooperation from the Czech and Flemish governments.

References

[1] H1 VFPS group, *Proposal for Installation of a Very Forward Proton Spectrometer in H1 after 2000*, H1 internal note 5/00-582, DESY PRC 01/00.

[2] H1 Collaboration, paper submitted to the 28^{th} International Conference on High Energy Physics (ICHEP96), Warsaw, Poland, July 1996.

[3] J. Schwindling and B. Mansoulié, *MLPfit: a tool for Multi-Layer Perceptrons*, http://schwind.home.cern.ch/schwind/MLPfit.html.