

**DETERMINATION OF QUARK DISTRIBUTIONS IN ep COLLISIONS**

G. INGELMAN and R. RÜCKL

*Deutsches Elektronen-Synchrotron DESY, Notkestraße 85, D-2000 Hamburg 52, Fed. Rep. Germany*

Received 4 November 1987

A method is presented for extracting various interesting quark distributions from measurements of differential cross sections for neutral and charged current processes at an ep collider such as HERA. The complete structure of ep interactions as described by the parton model in leading-order QCD and lowest-order electroweak theory is taken into account. Using Monte Carlo data we illustrate the experimental feasibility of the proposed unfolding procedure.

In the past, deep-inelastic lepton–nucleon scattering has played a leading role in exploring the interior structure of the nucleon and putting QCD to quantitative tests. At HERA these investigations will be continued and extended to momentum transfers  $Q$  considerably beyond the W and Z masses as well as to values of Bjorken- $x$  below  $10^{-2}$ . One expects to get rid of uncertainties due to non-asymptotic effects from target mass, heavy flavour thresholds, higher twist operators, and quark and gluon fragmentation, and to be able to examine the asymptotic scaling violations predicted by QCD under purer conditions. Furthermore, new determinations of quark distributions and, more indirectly, also of the gluon density should be possible with a resolving power for distances as small as  $10^{-16}$  cm.

Yet, it does not seem to be an easy task at ep colliders to realize these expectations [1]. Apart from various experimental difficulties in performing sufficiently precise and complete measurements of the basic observables, i.e. the differential cross sections in  $x$  and  $Q^2$  for the inclusive charged and neutral current processes  $ep \rightarrow \nu_e X$  (CC) and  $ep \rightarrow eX$  (NC), there are also some problems in principle which have to be solved. For instance, the scale dependences introduced into the cross sections through the W and Z boson propagators have to be unfolded before the QCD scaling violations can be tested. This is relatively simple in the CC case where the W propagator enters as an overall factor. On the other hand, NC scattering proceeds through photon and Z boson exchange including interferences so that the  $\gamma$  and Z propagators cannot simply be factored out from the NC structure functions. Moreover, in ep collisions one does not have the simplifications due to isospin symmetry which exist for lepton scattering on isoscalar nuclear targets and greatly facilitate the determination of singlet and non-singlet (under the flavour group) combinations of quark distributions. As a consequence, the deep-inelastic structure functions directly measurable in ep collisions cannot straightforwardly be connected to results from present-day fixed target experiments. Also, the extraction of particular quark distributions and tests of the QCD renormalization group predictions for scaling violations over the whole accessible range in  $Q^2$  require some new experimental approaches and a more involved data analysis. Finally, electroweak radiative corrections are predicted to be important [2] and must be carefully taken into account.

In order to develop a strategy for structure function studies optimized for high-energy ep collisions it is certainly useful to try different procedures. In this letter we address the problem of extracting quark distributions from differential NC and CC cross sections. Approaches which have been contemplated so far [1,3] include valence quark approximation at large  $x$ , photon exchange approximation at small  $Q^2$ , combination of measurements at different ep CM energies, use of especially tuned polarized  $e^\pm$  beams, and even collisions of an electron beam with an isoscalar deuteron beam. We want to suggest another possibility which admittedly also has its limitations but which is free of the obvious disadvantages of the suggestions just mentioned. In particular, our

method is a priori not restricted to large  $x$  or small  $Q^2$  and, moreover, avoids time consuming runs at different CM energies. After a general explanation of the procedure we illustrate the experimental feasibility of our proposal by a Monte Carlo event simulation.

The method is based on measurements of NC and CC cross sections in suitable bins in  $x$  and  $Q^2$  for  $e^-p$  and  $e^+p$  collisions at a fixed CM energy  $\sqrt{s}$ . It is assumed that the electroweak couplings and W and Z masses are as predicted by the standard model, an assumption which will be tested very precisely at SLC/LEP, and that the radiative effects are subtracted off from the data. Correspondingly, we start from the parton model expressions for the inclusive differential cross sections in the limit of vanishing lepton and quark masses. Effects of heavy quark masses are not essential for the present study. In fact, in a large part of the accessible  $(x, Q^2)$  region at HERA, the contributions from charm and bottom quarks are expected to have a leading log behaviour similar to that of light quarks, while top production is either negligible or directly observable [4]. In the NC case one thus has

$$\frac{d\sigma_{\text{NC}}(e^\mp)}{dx dQ^2} = \frac{4\pi\alpha^2}{xQ^4} [y^2 xF_1(x, Q^2) + (1-y)F_2(x, Q^2) \pm (y - \frac{1}{2}y^2) xF_3(x, Q^2)], \quad (1)$$

where  $\alpha$  is the electromagnetic fine-structure constant and  $y = Q^2/xs$ . The NC structure functions  $F_i(x, Q^2)$  read

$$F_2(x, q^2) = 2xF_1(x, Q^2) = \sum_f A_f(Q^2)[xq_f(x, Q^2) + x\bar{q}_f(x, Q^2)],$$

$$xF_3(x, q^2) = \sum_f B_f(Q^2)[xq_f(x, Q^2) - x\bar{q}_f(x, Q^2)], \quad (2)$$

where  $q_f(\bar{q}_f)$  describes the density of a quark (antiquark) flavour  $f$  in the proton including leading order QCD scaling violations, and the flavour-dependent coefficients  $A_f$  and  $B_f$  are given by

$$A_f(Q^2) = e_f^2 - 2e_f v_e v_f P_Z + (v_e^2 + a_e^2)(v_f^2 + a_f^2) P_Z^2, \quad B_f(Q^2) = -2e_f a_e a_f P_Z + 4v_e v_f a_e a_f P_Z^2. \quad (3)$$

Here,  $e_f$  is the electric charge ( $e_e = -1$ ),  $v_f = [T_{3f} - 2e_f \sin^2 \theta_w / \sin 2\theta_w]$  and  $a_f = T_{3f} / \sin 2\theta_w$  are the NC vector and axial vector couplings expressed in terms of the third component of the weak isospin ( $T_{3e} = -\frac{1}{2}$ ) and the Weinberg angle  $\theta_w$ , and  $P_Z = Q^2 / (Q^2 + m_Z^2)$  denotes the ratio of  $\gamma$  and Z propagators. The individual contributions from  $\gamma$  exchange, Z exchange and  $\gamma$ -Z interference are easily recognizable in eq. (3). Furthermore, the differential CC cross sections can be written in the form

$$\frac{d\sigma_{\text{CC}}(e^-)}{dx dQ^2} = \frac{\pi\alpha^2}{4 \sin^4 \theta_w (Q^2 + m_W^2)^2} \sum_{ij} [ |V_{u_i d_j}|^2 u_i(x, Q^2) + (1-y)^2 |V_{u_j d_i}|^2 \bar{d}_i(x, Q^2) ],$$

$$\frac{d\sigma_{\text{CC}}(e^+)}{dx dQ^2} = \frac{\pi\alpha^2}{4 \sin^4 \theta_w (Q^2 + m_W^2)^2} \sum_{ij} [ |V_{u_i d_j}|^2 \bar{u}_i(x, Q^2) + (1-y)^2 |V_{u_j d_i}|^2 d_i(x, Q^2) ], \quad (4)$$

where  $V_{u_i d_j}$  are elements of the Kobayashi-Maskawa matrix,  $u_i$  and  $d_j$  denote up-type and down-type quark flavours, respectively, and  $i, j$  are family indices. As a good approximation one may consider a world of four massless quark flavours ( $u, d, s, c$ ) and use unitarity relations  $\sum_j |V_{u_i d_j}|^2 = \sum_j |V_{u_j d_i}|^2 = 1$ .

One can now solve eqs. (1), (4) for four quark distribution functions or combinations thereof. It is convenient to choose the valence quark distributions  $u_v(x, Q^2)$  and  $d_v(x, Q^2)$ , and the total up- and down-type quark distributions  $U(x, Q^2) = \sum_i [u_i(x, Q^2) + \bar{u}_i(x, Q^2)]$  and  $D(x, Q^2) = \sum_i [d_i(x, Q^2) + \bar{d}_i(x, Q^2)]$ . Using the abbreviations

$$\sigma_{\pm}^{\text{NC}} = \frac{Q^4}{4\pi\alpha^2} \left( \frac{d\sigma_{\text{NC}}(e^-)}{dx dQ^2} \pm \frac{d\sigma_{\text{NC}}(e^+)}{dx dQ^2} \right), \quad \sigma_{\pm}^{\text{CC}} = \frac{4 \sin^4 \theta_w (Q^2 + m_W^2)^2}{\pi\alpha^2} \left( \frac{d\sigma_{\text{CC}}(e^-)}{dx dQ^2} \pm \frac{d\sigma_{\text{CC}}(e^+)}{dx dQ^2} \right) \quad (5)$$

and the coefficients  $A_f(Q^2)$  and  $B_f(Q^2)$  from eq. (3) one finds

$$u_v(x, Q^2) = f_1^{uv} \sigma_-^{NC} + f_2^{uv} \sigma_-^{CC}, \quad f_1^{uv} = \frac{(1-y)^2}{[1-(1-y)^2][(1-y)^2 B_u + B_d]}, \quad f_2^{uv} = \frac{B_d}{(1-y)^2 B_u + B_d}, \quad (6)$$

$$d_v(x, Q^2) = f_1^{dv} \sigma_-^{NC} + f_2^{dv} \sigma_-^{CC}, \quad f_1^{dv} = \frac{1}{[1-(1-y)^2][(1-y)^2 B_u + B_d]}, \quad f_2^{dv} = \frac{-B_u}{(1-y)^2 B_u + B_d}, \quad (7)$$

$$U(x, Q^2) = f_1^U \sigma_+^{NC} + f_2^U \sigma_+^{CC}, \quad f_1^U = \frac{(1-y)^2}{[1+(1-y)^2][(1-y)^2 A_u - A_d]}, \quad f_2^U = \frac{-A_d}{(1-y)^2 A_u - A_d}, \quad (8)$$

$$D(x, Q^2) = f_1^D \sigma_+^{NC} + f_2^D \sigma_+^{CC}, \quad f_1^D = \frac{-1}{[1+(1-y)^2][(1-y)^2 A_u - A_d]}, \quad f_2^D = \frac{A_u}{(1-y)^2 A_u - A_d}. \quad (9)$$

From these distributions one can construct the non-singlet structure function

$$F_{ns}(x, Q^2) \equiv x[u_v(x, Q^2) + d_v(x, Q^2)] = f_1^{ns} \sigma_-^{NC} + f_2^{ns} \sigma_-^{CC},$$

$$f_1^{ns} = \frac{x[1+(1-y)^2]}{[1-(1-y)^2][(1-y)^2 B_u + B_d]}, \quad f_2^{ns} = \frac{x(B_d - B_u)}{(1-y)^2 B_u + B_d}, \quad (10)$$

and the singlet structure function

$$F_s(x, Q^2) \equiv x[U(x, Q^2) + D(x, Q^2)] = f_1^s \sigma_+^{NC} + f_2^s \sigma_+^{CC},$$

$$f_1^s = \frac{-x[1-(1-y)^2]}{[1+(1-y)^2][(1-y)^2 A_u - A_d]}, \quad f_2^s = \frac{x(A_u - A_d)}{(1-y)^2 A_u - A_d}, \quad (11)$$

which have been measured in present-day experiments (at least in good approximation) and which are particularly suitable quantities for QCD renormalization group states [5]. The above relations make the manipulations explicit which must be applied to the measured cross sections in order to unfold the weak propagator effects and the flavour dependent  $\gamma$  and Z couplings, and to extract particular quark distributions.

Having described the method in principle, we now turn to the most crucial question of how well it can be expected to work in practice. Since the achievable accuracy in the determination of the quark distributions from the relations (6)–(11) clearly depends on the properties of the coefficient functions  $f_i$ , it is useful to pay some attention to these quantities. Concentrating on the valence up-quark density  $u_v(x, Q^2)$  and the singlet structure function  $F_s(x, Q^2)$  for brevity, we have plotted the corresponding coefficients  $f_{1,2}^{uv}$  versus  $x$  for some values of  $Q^2$  in fig. 1. As can be seen from figs. 1a and 1b,  $f_1^{uv}$  is much larger than  $f_2^{uv}$  in the considered range of  $Q^2$  except for  $x \rightarrow x_{\min}$  where  $f_1^{uv}$  vanishes proportionally to  $(1-y)^2$ . Hence, NC data are more relevant for the determination of  $u_v(x, Q^2)$  than what one would expect from the difference  $\sigma_-^{NC}$  alone. The importance of the NC input even at moderate values of  $Q^2$  is indeed surprising since  $\sigma_-^{NC}$  gets quite small with decreasing  $Q^2$  as a result of the increasing dominance of the pure photon exchange. Figs. 1c and 1d show that the singlet coefficients  $f_1^s$  and  $f_2^s$  are similar in size and that they are divergent at certain values of  $x$  and  $Q^2$ . These divergencies are caused by the vanishing of the denominators  $(1-y)^2 A_u(Q^2) - A_d(Q^2)$  in eq. (11). For sufficiently low  $Q^2$ ,  $A_f(Q^2) \simeq e_f^2$  so that the problem occurs at  $y \simeq \frac{1}{2}$ , i.e.  $x \simeq 2Q^2/s$ . As  $Q^2$  increases this point is shifted to values of  $x$  somewhat larger than  $2Q^2/s$  due to the Z-boson contributions to  $A_f(Q^2)$  given in eq. (3). The occurrence of divergencies just means that the cross sections (1) and (4) cannot be inverted and solved for  $F_s(x, Q^2)$  along a particular line in the  $(x, Q^2)$ -plane. In practice, the unfolding procedure is bound to fail at values of  $x$  and  $Q^2$  close to this line because of rapid variation of the coefficient functions. One sees from figs. 1c and 1d that the whole branches of  $f_{1,2}^s$  on the left-hand side of the divergencies, corresponding to  $y > 0.5$ , are problematic. This kinematical region is disfavoured also for other reasons, e.g. large radiative corrections. On the other hand, the branches of  $f_{1,2}^s$  on the right-hand side of the divergencies, corresponds to  $y < 0.5$ , are perfectly well-behaved. Concerning the properties of the other coefficient functions which appear in eqs. (7)–(10) we only state that apart from overall signs  $f_{1,2}^{dv}$  and  $f_{1,2}^{ns}$  are similar to  $f_{1,2}^{uv}$ , while  $f_{1,2}^U$  and  $f_{1,2}^D$  resemble  $f_{1,2}^s$ .

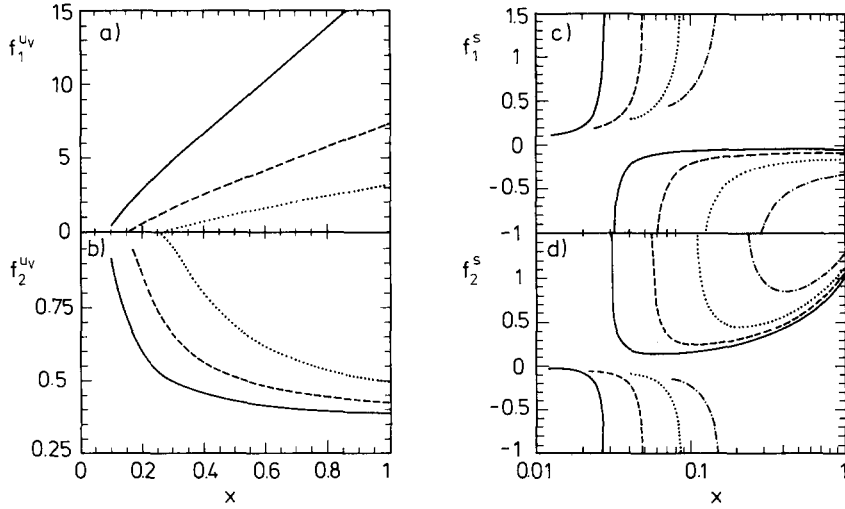


Fig. 1. Coefficients of the unfolding relations for (a), (b) the valence up-quark distribution at  $Q^2(10^3 \text{ GeV}^2) = 7.8$  (full), 14 (dashed), 25 (dotted), and (c), (d) the singlet structure at  $Q^2(10^3 \text{ GeV}^2) = 1.4$  (full), 2.5 (dashed), 4.4 (dotted), 7.8 (dash-dotted).

In order to examine our extraction for realistic data samples we use a Monte Carlo event generator [6] to simulate NC and CC scattering of 30 GeV electrons and positrons on 820 GeV protons. The Monte Carlo generator is based on the differential cross sections eqs. (1) and (4) taking the quark distribution functions given by parametrization I of ref. [7] and the following values for the electroweak parameters:  $\alpha = 1/137$ ,  $\sin^2 \theta_w = 0.226$  and  $m_w = m_z \cos \theta_w = 38.68 \text{ GeV}/\sin \theta_w$ . An integrated luminosity of  $200 \text{ pb}^{-1}$  per  $e^\mp$  beam is assumed. The number of events as well as the corresponding cross sections are recorded in bins of  $x$  and  $Q^2$  for  $x \geq 0.01$  and  $Q^2 \geq 10^3 \text{ GeV}^2$ . In order to ensure samples with reasonable statistics, the bin-size is increased as  $x$  and  $Q^2$  increase. For  $Q^2$  we have chosen four bins per decade which are equally large on a logarithmic scale, i.e.  $\Delta \log Q^2 = 0.25$ , and in  $x$  we have taken the bins  $\Delta x = 0.05$  for  $x \leq 0.5$ ,  $\Delta x = 0.1$  for  $0.5 \leq x \leq 0.8$ , and the bin  $0.8 \leq x \leq 1$ . Of course, larger bins can be used to decrease the size of the statistical errors, but we prefer a not too coarse binning in order to indicate the possible resolution.

These Monte Carlo data sets simulate the real data obtainable in a few years of experimentation at HERA and serve as the input in eqs. (6)–(11). There is some freedom in the choice of the point within a  $(x, Q^2)$ -bin that should be used for the numerical evaluation of the coefficient functions in eqs. (6)–(11) and the factors multiplying the brackets in eq. (5). As a reasonable choice we have taken the weighted bin center as obtained from the event samples. Figs. 2 and 3 show the resulting momentum distribution for the valence up-quark,  $xu_v(x, Q^2)$ , and the singlet structure function  $F_s(x, Q^2)$ , in comparison to the corresponding input distributions used in the generation of the event samples. We find that NC and CC measurements are both necessary to reconstruct  $xu_v(x, Q^2)$  and  $F_s(x, Q^2)$  in the intermediate  $Q^2$  range (figs. 2a and 3d). However,  $F_s(x, Q^2)$  is essentially determined by CC data at relatively low  $Q^2$  (fig. 3a), while  $xu_v(x, Q^2)$  is dominated by the NC input at very high  $Q^2$  (fig. 2c).

The errors displayed in figs. 2 and 3 are obtained from the statistical errors on the number of NC and CC events in a given bin in  $x$  and  $Q^2$ ,  $N_{\text{NC}}(e^\mp)$  and  $N_{\text{CC}}(e^\mp)$ , propagated through the unfolding relations. Formally, one has

$$\delta = \sqrt{(f_1)^2 \left( \frac{\tilde{\sigma}_{\text{NC}}^2(e^-)}{N_{\text{NC}}(e^-)} + \frac{\tilde{\sigma}_{\text{CC}}^2(e^+)}{N_{\text{CC}}(e^+)} \right) + (f_2)^2 \left( \frac{\tilde{\sigma}_{\text{CC}}^2(e^-)}{N_{\text{CC}}(e^-)} + \frac{\tilde{\sigma}_{\text{CC}}^2(e^+)}{N_{\text{CC}}(e^+)} \right)}, \quad (12)$$

where  $f_{1,2}$  denote the appropriate coefficient functions, that is in the case at hand either  $x f_{1,2}^{uv}$  or  $f_{1,2}^s$ , and  $\tilde{\sigma}$

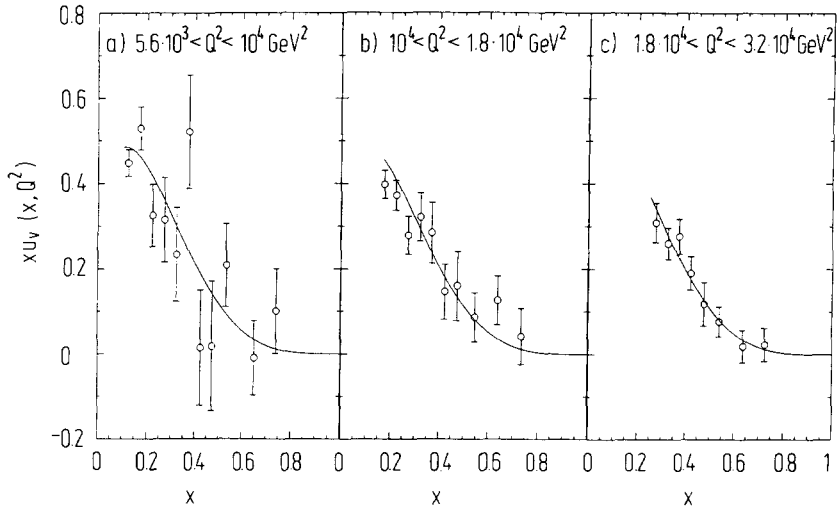


Fig. 2. Valence up-quark distribution extracted from Monte Carlo data samples for  $e^-p$  and  $e^+p$  scattering with statistical errors corresponding to  $200 \text{ pb}^{-1}$  per lepton beam, in comparison to the input distribution (full curves).

represents the differential cross sections given in eqs. (1) and (4) scaled with the same propagator factors as used for  $\sigma_{\pm}$  in eq. (5). For bins with zero or only a few events in a data sample, we use Poisson statistics instead of eq. (12) to obtain the proper error estimates. Fig. 2 demonstrates that for  $u_v(x, Q^2)$  the statistical precision of the unfolding procedure improves as  $Q^2$  is increased. It becomes quite satisfactory in the very high  $Q^2$  range until  $Q^2$  exceeds about  $4 \times 10^4 \text{ GeV}^2$  where one has only very few events. On the other hand, in the intermediate

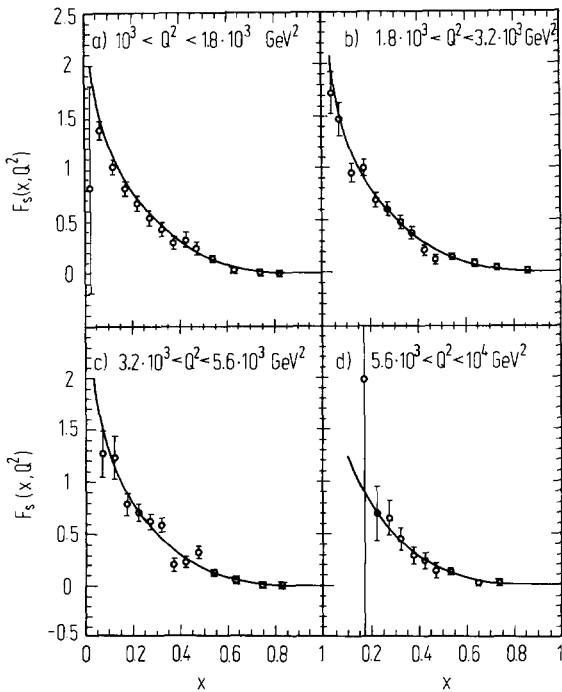


Fig. 3. Singlet structure function  $F_s(x, Q^2) = \sum_i [xq_i(x, Q^2) + x\bar{q}_i(x, Q^2)]$  extracted from Monte Carlo data (specified in fig. 2) and compared to the input distribution (full curves).

$Q^2$  range below  $10^4 \text{ GeV}^2$  the statistical errors are large. This somewhat surprising result can be understood from the properties of NC coefficient  $f_1^{u\nu}$  shown in fig. 1a and from the error formula eq. (12). More precisely, since  $f_1^{u\nu} \gg f_2^{u\nu}$  the charged current statistics gives a negligible contribution to the statistical errors on  $u_\nu(x, Q^2)$  in the whole  $Q^2$  range considered and, moreover, the statistical uncertainties in the NC data are strongly amplified by the rapid increase of the size of  $f_1^{u\nu}$  as  $Q^2$  decreases. Quite in contrast, the singlet structure function  $F_s(x, Q^2)$  can be well extracted in the intermediate  $Q^2$  region as illustrated in fig. 3. Here, the statistical errors are mainly determined by the CC data samples and appear acceptable at least up to  $Q^2 \approx 10^4 \text{ GeV}^2$ . The large errors on the leftmost points in figs. 3a and 3d reflect the drastic changes of the coefficient functions  $f_{1,2}^s$  under small variations of  $x$  near the divergencies shown in figs. 1c and 1d. We want to emphasize, however, that the break-down of the extraction method only affects the low- $x$  end of the spectra and is typically localized to one or two  $x$ -bins.

In summary, we have derived relations unfolding particular quark structure functions from the differential neutral and charged current cross sections for  $e^-p$  and  $e^+p$  scattering. It should be noted that simplifications such as valence quark and photon exchange approximations have not been used. In other words, modulo electroweak radiative effects and higher order QCD corrections, our procedure is as general as possible. Furthermore, based on a Monte Carlo event simulation we have studied the experimental feasibility of the proposed method at HERA energies. As examples, we have shown the statistical precision with which the valence up-quark distribution and the singlet structure function can be reconstructed. The results are quite encouraging. Of course, one has to keep in mind that systematic errors and uncertainties due to radiative corrections have not yet been considered. We have also studied other quark distributions, without including them in this letter, and find that the method works well for some but not all distributions one might wish to determine. In general, the regions in  $x$  and  $Q^2$  where a sufficiently precise extraction is possible are different for different structure functions as exemplified by  $u_\nu(x, Q^2)$  and  $F_s(x, Q^2)$ . Improvements of our method and combinations with more conventional approaches are possible and will be investigated.

## References

- [1] E. Longo, Proc. Workshop on Experimentation at HERA (Amsterdam, 1983), DESY HERA 83/20, p.285;  
F. Brasse, Proc. Workshop on Experimentation at HERA (Amsterdam, 1983), DESY HERA 83/20, p. 499.
- [2] D.Yu. Bardin, O.M. Fedorenko and N.M. Shumeiko, J. Phys. G. 7 (1981) 1331;  
M. Böhm and H. Spiesberger, Würzburg Universität preprint (December 1986);  
J. Kripfganz and H.-J. Moering, KMU HEP internal report (May 1987);  
D. Yu. Bardin, C. Burdick, P. Ch. Christova and T. Riemann, Dubna preprint E2-87-595.
- [3] M. Klein and T. Riemann, Z. Phys. C 24 (1984) 151;  
M. Klein, in: Proc. DESY Workshop on Physics at HERA (October 1987), to be published.
- [4] M. Glück, R.M. Godbole and E. Reya, Dortmund Univ. preprint DO-TH 87/12.
- [5] G. Altarelli, Phys. Rep. 81 (1982) 1.
- [6] G. Ingelman, LEPTO version 5.2, DESY preprint, in preparation.
- [7] E. Eichten, I. Hinchliffe, K. Lane and C. Quigg, Rev. Mod. Pys. 56 (1984) 579; 58 (1986) 1047.