Almost all visible matter is composed of protons and neutrons. Analyses of the scattering of cosmic ray particles off nuclei or of results from fixed target and colliding hadron beam experiments require a quantitative understanding of the partonic structure of nucleons. The theoretical framework for this is known \cite{1,2} since the inception of quantum chromodynamics (QCD); see also Refs. \cite{3,4}.

Of particular importance for the experimental programs at the Large Hadron Collider are unpolarized parton distribution functions (PDFs). These, in the light cone frame, parametrize the likelihood of a parton to carry the Bjorken momentum fraction $x$ at a renormalization scale $\mu$. While these PDFs have been mapped out very well from fits to experimental data, ideally one would wish to evaluate them directly from the underlying fundamental theory, QCD.

The present method of choice is lattice QCD, where in principle all approximations can be removed and systematic uncertainties controlled by taking the limits of infinite volume, of vanishing lattice spacing ($a \to 0$) and of physical quark masses. However, in this Monte Carlo simulation approach to QCD, the statistical errors and the reliability of the extrapolation to the physical point are limited by the power of available computers and the efficiency of numerical algorithms. Moreover, only Mellin moments of the PDFs can be accessed. Thus, present-day lattice simulation cannot compete in terms of precision with determinations of isovector unpolarized PDFs from fits to experimental photon-nucleon scattering data that have been collected over decades of dedicated effort. For a summary of the present status of PDF parametrizations, see Ref. \cite{5}.

The possibility to predict averages over the momentum fraction, however, is complementary to experimental measurements that can only cover a limited range of $x$ values. In particular, the strangeness and gluonic PDFs are determined rather indirectly from experiment, with large uncertainties \cite{6}. Consequently, lattice QCD input is already on the verge of becoming essential to constrain these and other less well-known quantities, e.g., the pion nucleon $\sigma$ term \cite{7–11}, the strangeness fraction of the mass of the proton $f_T$ \cite{8–13}, the strange quark contribution to the spin of the proton $\Delta s + \Delta \bar{s}$ \cite{13,14} or individual (valence and sea) quark contributions to the proton’s momentum $\langle x \rangle_u$, $\langle x \rangle_d$ and $\langle x \rangle_s$ \cite{15}. Naturally, for lattice predictions of such quantities to be trusted, lattice QCD needs to demonstrate its ability to reproduce known observables within similar or smaller errors.

We focus on the second moment of the isovector PDF of the proton $\langle x \rangle_u - d$, i.e., on the difference between the average momentum fractions carried by the up and the down quarks. In the isospin symmetric limit, disconnected quark line contributions cancel in this quantity, which makes it particularly accessible to lattice simulations. Nevertheless, at present $\langle x \rangle_u - d$ seems to be the one benchmark lattice observable that is least reliably determined.

We summarize the present status in Fig. 1 where we plot $\langle x \rangle_u - d$ in the $\overline{MS}$ scheme, at a scale $\mu = 2$ GeV, as a function of the squared pion mass, obtained by different lattice collaborations. Simulations at physically light pion masses are particularly expensive and have only recently become possible for calculating relatively simple observables, e.g., the light hadron spectrum \cite{16}.

The leftmost points (at the physical pion mass $m_\pi = 138$ MeV) are obtained from recent PDF parametrizations \cite{17–19}, evolved to the renormalization scale $\mu = 2$ GeV. The data point next to it (triangle) is the main result of this report. The QCDSF Collaboration points \cite{20,21} stem from simulations employing the same nonperturbatively improved \cite{22} Sheikholeslami-Wohlert $N_f = 2$ quark action at lattice scales ranging from $a^{-1} = 2.60$ GeV to $a^{-1} = 3.26$ GeV. The ETM Collaboration points \cite{23} were obtained using the $N_f = 2$ twisted mass action at maximal twist at $a^{-1} = 2.22$ GeV and at $a^{-1} = 3.52$ GeV while the RBC-UKQCD Collaboration \cite{24} simulated...
\( N_f = 2 + 1 \) flavors using the (approximately) chiral domain-wall action at \( a^{-1} = 1.7 \) GeV. We have omitted data where the renormalization is not exactly known [25] from the figure; see Refs. [26,27] for recent reviews. All the previous data shown in the figure are compatible with a constant \( \sim 0.25 \), far above the expected value \( \sim 0.16 \) and no dependence on the different (quark and gluon) lattice actions, all accurate to \( O(a) \), on the lattice spacing, on the lattice volume, on including strange sea quark or not or on the pion mass can be resolved within statistical errors.

Covariant baryon chiral perturbation theory suggests \( \langle x \rangle_{u-d} \) to decrease as the physical point is approached, see, e.g., Ref. [28]. However, the onset of this behavior is not at all visible for \( m_\pi > 250 \) MeV, with the possible (but not significant) exception of the RBC-UKQCD Collaboration point with the smallest mass value. Different sources of systematics in lattice extractions of various nucleon matrix elements have been discussed in the past. In particular, contaminations of the ground state signal, due to excited state contributions, have recently gained prominence in the literature [29–32]: reducing these pollutions results in smaller \( \langle x \rangle_{u-d} \) predictions. This effect, though noticeable, however seems to be too small to dominantly contribute to the deviations by factors of about 1.5 of present lattice results from values obtained from PDF parametrizations.

We report on our new result that is obtained simulating \( N_f = 2 \) nonperturbatively improved Sheikholeslami-Wohlert fermions on top of the Wilson gauge action at \( \beta = 5.29 \) and \( \kappa = 0.13640 \), corresponding to \( m_\pi = 157(6) \) MeV, on a volume of \( 48^3 \times 64 \) lattice points. Setting the scale from the (chirally extrapolated) nucleon mass we obtain the inverse lattice spacing [7] \( a^{-1} = 2.76(5)(6) \) GeV, where the errors are statistical and from the chiral extrapolation, respectively. Our spatial linear lattice dimension \( L = 3.43 \) fm is rather small in units of the pion mass: \( m_\pi L = 2.74(3) \). However, in previous simulations at \( m_\pi = 290 \) MeV, differences between results from three volumes of linear sizes \( m_\pi L = 2.5, 3.4 \) and 4.2 could not be resolved [20], within statistical errors smaller than the present one. Hence, we expect finite volume effects to be much smaller than our statistical uncertainty.

The second Mellin moment of the isovector quark distribution in the proton is given by

\[
\langle x \rangle_{u-d} = \int_0^1 dx \{ u(x) + \bar{u}(x) - d(x) - \bar{d}(x) \},
\]

where \( u(x), d(x) \) and \( \bar{u}(x), \bar{d}(x) \) denote the quark and antiquark PDFs, respectively. It is computed by creating a proton at a Euclidean time \( t_0 = 0 \), destroying it at a time \( t_f > 0 \) and inserting the current

\[
\bar{u}(\gamma_4 \vec{D}_4 - \frac{1}{3} \gamma \cdot \vec{D}) u - \bar{d}(\gamma_4 \vec{D}_4 - \frac{1}{3} \gamma \cdot \vec{D}) d,
\]

projected to zero spatial momentum, at an intermediate time \( t < t_f \). For details see, e.g., the review [33]. We use mass-degenerate, electrically neutral \( u \) and \( d \) quarks and therefore all disconnected quark line loops cancel. The resulting three-point function is normalized by dividing out the two-point function of the propagating proton. Keeping \( t_f \) fixed and sufficiently large, for \( t_f \gg t > 0 \) one obtains a plateau in \( t \) from which the lattice matrix element can be extracted. This is then translated into the \( \text{MS} \) scheme at a scale \( \mu = 2 \) GeV, using the renormalization factor determined in Ref. [34], which contains nonperturbative renormalization into the intermediate \( \text{RI}^*\text{MOM} \) scheme and a subsequent perturbative conversion at three-loop accuracy into the \( \text{MS} \) scheme.

FIG. 1 (color online). Pion mass dependence of \( \langle x \rangle_{u-d}^{\text{MS}} \) at \( \mu = 2 \) GeV from \( N_f = 2 \) (new, QCDSF Collaboration [20,21], ETM Collaboration [23]) and \( N_f = 2 + 1 \) (the RBC-UKQCD Collaboration [24]) lattice QCD simulations, together with expectations from PDF parametrizations (NNPDF [17], ABM [18], MSTW [19]).

FIG. 2 (color online). Plateaus and fitted values (grey bands) of ratios of renormalized three-point over two-point functions at \( \beta = 5.29 \) \( (a^{-1} \approx 2.76 \text{ GeV}) \) for \( \kappa = 0.13632 \) \( (m_\pi \approx 290 \text{ MeV}) \) and \( \kappa = 0.13640 \) \( (m_\pi \approx 157 \text{ MeV}) \).
In Fig. 2 we show the resulting plateau and fit to the renormalized \( \langle x \rangle_{u-d}^{\text{MS}} \) where we use \( t_f = 15a \approx 1.07 \text{ fm} \). We also compare this to our previous result obtained for the same lattice spacing at \( m_\pi = 290 \text{ MeV} \) (\( \kappa = 0.13632 \)), on a \( 40^3 \times 64 \) lattice [20]. The onset and the quality of the plateau depend on the overlap of the interpolating field used to create and to destroy the proton with the physical ground state. Details of our quark and gauge field smearing can be found in Ref. [8]. Here we set the number of Wuppertal smearing iterations to 400. The larger pion mass data [20] were generated with sources and sinks of a smaller (Jacobi) smearing radius and exhibit more curvature as a function of \( t_f/a \), indicating that the ground state overlap is inferior in this case. This is also obvious from a comparison of the respective two-point functions. Improving the smearing, varying \( t_f \) in addition to \( t \) and employing more sophisticated fit functions may somewhat reduce previous values that were obtained at the larger pion masses. Initial tests however indicate that the difference in quality of the interpolators used to create the proton can only explain a fraction of the observed reduction of \( \langle x \rangle_{u-d} \) as the quark mass is decreased.

In Fig. 3 we display results for the generalized form factor \( A_{20}^{u-d} \) as a function of the squared momentum transfer \( Q^2 \) for the present mass point and for the previous two smallest QCDSF Collaboration pion masses [20]. (\( \beta = 5.4 \) corresponds to the somewhat finer lattice scale \( a^{-1} \approx 3.26 \text{ GeV} \).) To leading order in chiral perturbation theory these are expected to extrapolate linearly in \( Q^2 \) to \( \langle x \rangle_{u-d} = A_{20}^{u-d}(0) \), see, e.g., Refs. [28,33]. The results agree with those obtained from the forward matrix elements that are also displayed in Fig. 1 (leftmost points).

This report is based on a (computationally expensive) analysis of the first 1000 configurations of our recent \( N_f = 2 \) simulations at a pion mass \( m_\pi = 157 \text{ MeV} \) and at an inverse lattice spacing \( a^{-1} \approx 2.76 \text{ GeV} \). Altogether, about 3500 configurations have been generated at these parameter values on the QPACE supercomputers [35]. Hence, we expect the present error to reduce by a factor close to 2, once the target statistics are reached. We find \( \langle x \rangle_{u-d} \) to drop substantially, an effect that was not visible at \( m_\pi > 250 \text{ MeV} \). Part of this may be attributed to some excited state pollution of the previous data points. A reanalysis of these is also in progress.

Our value at \( m_\pi = 157 \text{ MeV} \),

\[
\langle x \rangle_{u-d}^{\text{MS}}(2 \text{ GeV}) = 0.207(16),
\]

is still by 2.5 standard deviations larger than the central values of PDF parametrizations [5,17–19]. Figure 1, however, suggests \( \langle x \rangle_{u-d} \) to decrease steeply as the physical pion mass is approached. Moreover, our error, that incorporates the insignificant renormalization uncertainty [34], is purely statistical otherwise. It does not reflect any systematics of the missing continuum limit and infinite volume extrapolations. We remark that the comparison between lattice QCD results and fits to experimental data for \( \langle x \rangle_{u-d} \) is less clean than for the individual moments \( \langle x \rangle_u \) and \( \langle x \rangle_d \). These, however, would require disconnected quark line diagrams to be computed on the lattice. Experimentally, the neutron is usually bound in a deuteron or another nucleus. The EMC effect [36] will affect PDFs at large \( x \) [37] and change their moments, relative to those of a free neutron. On the lattice side, isospin symmetry was assumed. However, a strong slope as a function of the pion mass may also indicate significant corrections to this approximation. We plan to investigate this further.

We conclude that lattice simulations are likely to reproduce known moments of unpolarized PDFs in the near future. This prediction is based on our observation of a steep decrease of the value of \( \langle x \rangle_{u-d} \) towards small pion masses. Clearly, high statistics simulations at pion masses smaller than 200 MeV and with good control over excited state contributions are required.

We thank Y. Nakamura and J. Zanotti for their support. This work was funded by the DFG Sonderforschungsbereich/Transregio 55 in part and is supported by the EU Initial Training Network STRONGnet No. 238353. S. Collins acknowledges support from the Claussen-Simon-Foundation (Stifterverband für die Deutsche Wissenschaft) and A. Sternbeck from the EU (IRG 256594). Computations were performed on the SFB/TR55 QPACE computers [35], the BlueGene/P (JuGene) of the Jülich Supercomputing Centre, the SuperMIG/MUC machine of the Leibniz-Rechenzentrum Munich and Regensburg’s iDataCool cluster. The CHROMA software suite [38] was used and gauge configurations were generated with the BQCD code [39].
[1] H. Georgi and H. D. Politzer, Phys. Rev. D 9, 416 (1974).
[2] D. J. Gross and F. Wilczek, Phys. Rev. D 9, 980 (1974).
[3] N. H. Christ, B. Hasslacher, and A. H. Mueller, Phys. Rev. D 6, 3543 (1972).
[4] C. G. Callan, Jr. and D. J. Gross, Phys. Rev. D 8, 4383 (1973).
[5] S. Alekhin et al. (PDF4LHC Working Group), arXiv:1101.0536.
[6] G. Aad et al. (ATLAS Collaboration), Phys. Rev. Lett. 109, 012001 (2012).
[7] G. S. Bali et al. (QCDSF Collaboration), Nucl. Phys. B 85, 054502 (2012).
[8] S. Dürr et al., Phys. Rev. D 85, 014509 (2012).
[9] D. Pleiter et al. (QCDSF and UKQCD Collaborations), Proc. Sci., LATTICE 2010 (2010) 153 [arXiv:1204.0685].
[10] R. Babich, R. C. Brower, M. A. Clark, G. T. Fleming, J. C. Osborn, C. Rebbi, and D. Schaich (Disco Collaboration), Phys. Rev. D 85, 054510 (2012).
[11] W. Freeman et al. (MILC Collaboration), arXiv:1204.3866.
[12] Y. Aoki, T. Blum, H.-W. Lin, S. Ohta, S. Sasaki, R. Tweedie, J. Zanotti, and T. Yamazaki (RBC and UKQCD Collaborations), Phys. Rev. D 82, 014501 (2010).
[13] J. D. Bratt et al. (LHP Collaboration), Phys. Rev. D 82, 094502 (2010).
[14] M. Gong, A. Li, A. Alexandru, T. Draper, and K.-F. Liu (C31 QCD Collaboration), Proc. Sci., LATTICE 2011 (2011) 156 [arXiv:1203.6388].
[15] K.-F. Liu et al. (C31 QCD Collaboration), Proc. Sci., LATTICE 2011 (2011) 164 [arXiv:1203.6579].
[16] Z. Fodor and C. Hoelbling, Rev. Mod. Phys. 84, 449 (2012).
[17] R. D. Ball et al. (NNPDF Collaboration), Nucl. Phys. B855, 153 (2012).
[18] S. Alekhin, J. Blümlein, and S. Moch, arXiv:1202.2281.
[19] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, Eur. Phys. J. C 64, 653 (2009).
[20] A. Sternebeck et al. (QCDSF Collaboration), Proc. Sci., LATTICE 2011 (2011) 177 [arXiv:1203.6579].
[21] K. Jansen et al. (ALPHA Collaboration), Nucl. Phys. B530, 185 (1998); B643, 517(E) (2002).