Article

The Asymmetry of Antimatter in the Proton

J. Dove[1], B. Kerns[1], R. E. McClellan[1,18] S. Miyasaka[2], D. H. Morton[3], K. Nagai[2,4],
S. Prasad[1], F. Sanftl[2], M. B. C. Scott[3], A. S. Tadepalli[5,18], C. A. Aidala[3,6],
J. Arrington[7,19], C. Ayuso[3,20], C. L. Barker[8], C. N. Brown[9], W.C. Chang[4], A. Chen[1,3,4],
D. C. Christian[10], B. P. Dannowitz[1], M. Daugherity[8], M. Diefenthaler[1,18], L. El Fassi[5,11],
D. F. Geesaman[7,21], R. Gilman[5], Y. Goto[12], L. Guo[6,22], R. Guo[13], T. J. Hague[8],
R. J. Holt[7,23], D. Isenhower[8], E. R. Kinney[14], N. Kitts[8], A. Klein[6], D. W. Kleinjan[6],
Y. Kudo[15], C. Leung[1], P.-J. Lin[14], K. Liu[6], M. X. Liu[6], W. Lorenzon[3], N. C. R. Makins[1],
M. Mesquita de Medeiros[7], P. L. McGaughey[6], Y. Miyachi[15], I. Mooney[3,24],
K. Nakahara[16,25], K. Nakano[2,12], S. Nara[15], J.-C. Peng[1], A. J. Puckett[6,26],
B. J. Ramson[3,27], P. E. Reimer[7✉], J. G. Rubin[3,7], S. Sawada[17], T. Sawada[3,28],
T.-A. Shibata[2,29], D. Su[4], M. Teo[1,30], B. G Tice[7], R. S. Towell[8], S. Uemura[6,31],
S. Watson[8], S. G. Wang[4,13,32], A. B. Wickes[6], J. Wu[10], Z. Xi[8], Z. Ye[7]

The fundamental building blocks of the proton, quarks and gluons, have been known for decades. However, we still have an incomplete theoretical and experimental understanding of how these particles and their dynamics give rise to the quantum bound state of the proton and its physical properties, such as for example its spin.[1] The two up and the single down quarks that comprise the proton in the simplest picture account only for a few percent of the proton mass, the bulk of which is in the form of quark kinetic and potential energy and gluon energy from the strong force.[2] An essential feature of this force, as described by quantum chromodynamics, is its ability to create matter-antimatter quark pairs inside the proton that exist only for a very short time. Their fleeting existence makes the antimatter quarks within protons difficult to study, but their existence is discernible in reactions where a matter-antimatter quark pair annihilates. In this picture of quark-antiquark creation by the strong force, the probability distributions as a function of momentum for the presence of up and down antimatter quarks should be nearly identical, since their masses are quite similar and small compared to the mass of the proton.[3] In the present manuscript, we show evidence from muon pair production measurements that these distributions are significantly different, with more abundant down antimatter quarks than up antimatter quarks over a wide range of momentum. These results revive interest in several proposed mechanisms as the origin of this antimatter asymmetry in the proton that had been disfavored by the previous results[4] and point to the future measurements that can distinguish between these mechanisms.

The structure of the proton is a prototypical example of a strongly coupled and correlated system with the quarks and gluons interacting according to quantum chromodynamics (QCD). At large energy and momentum scales, the interaction is relatively weak, while at lower energy scales the picture is still clouded by the increasingly strong in-teraction. The original quark model, in which the proton consists of two up quarks (u) and one down (d) quark, has an appealing simplicity, but experiments that measure the distributions of quarks as a function of the fractional momentum of the proton that they carry (x) have re-vealed a rich structure with additional quarks, antimatter quarks (anti-

[1]Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL, USA. [2]Department of Physics, School of Science, Tokyo Institute of Technology, Meguro-ku,Tokyo, Japan. [3]Randal Laboratory of Physics, University of Michigan, Ann Arbor, Michigan, USA. [4]Institute of Physics, Academia Sinica, Taipei, Taiwan. [5]Department of Physics and Astronomy, Rutgers, The State University of New Jersey, Piscataway, New Jersey, USA. [6]Physics Division, Los Alamos National Laboratory, Los Alamos, New Mexico, USA. [7]Physics Division, Argonne National Laboratory, Lemont, Illinois, USA. [8]Department of Engineering and Physics, Abilene Christian University, Abilene, Texas, USA. [9]Accelerator Division, Fermi National Accelerator Laboratory, Batavia, Illinois, USA. [10]Particle Physics Division Fermi National Accelerator Laboratory, Batavia, Illinois, USA. [11]Department of Physics and Astronomy, Mississippi State University, Mississippi State, MS, USA. [12]RIKEN Nishina Center for Accelerator-Based Science, Wako, Saitama, Japan. [13]Department of PHysics,National Kaohsiung Normal University, Kaohsiung City, Taiwan. [14]Department of Physics, University of Colorado, Boulder, Colorado, USA. [15]Department of Physics, Yamagata University, Yamagata City, Yamagata, Japan. [16]Department of Physics, University of Maryland, College Park, Maryland, USA. [17]Institute of Particle and Nuclear Studies, KEK, High Energy Accelerator Research Organization, Tsukuba, Ibaraki, Japan. [18]Present address: Experimental Nuclear Physics Division, Thomas Jefferson National Accelerator Facility, Newports News, IL, USA. [19]Present address: Physics Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA. [20]Present address: Department of Physics and Astronomy, Mississippi State University, Mississippi State, MS, USA. [21]Present address: Argonne Associate of Global Empire LLC, Lemont IL, USA. [22]Present address: Department of Physics, Florida International University, Miami, Florida, USA. [23]Present address: Kellogg Radiation Laboratory, California Institute of Technology, Pasadena, CA, USA. [24]Present address: Department of Physics, Wayne State University, Detroit, MI, USA. [25]Present address: Stanford Linear Accelerator Center, Menlo Park, CA, USA. [26]Present address: Department of Physics, University of Connecticut, Storrs, CT, USA. [27]Present address: Neutrino Division, Fermi National Accelerator Laboratory. [28]Present address: Department of Physics, Osaka City University, Sumiyoshi-ku, Osaka City, Osaka 558-8585, Japan. [29]Present address: Department of Physics, College of Science and Technology, Nihon University, Tokyo, Japan. [30]Present address: Department of Physics, Stanford University, Stanford, CA, USA. [31]Present address: Department of Physics, Tel Aviv University, Tel Aviv, Israel. [32]Present address: Accelerator Systems Division, Argonne National Laboratory, Lemont, IL, USA.
✉email:reimer@anl.gov

Article

quarks), and gluons beyond the minimal three quark Fock state. These additional quarks and antiquarks are referred to as sea quarks. Collectively the quarks and gluons are referred to as partons. It is not possible to identify any individual up or down quark as a sea or valence quark, but antiquarks and strange quarks must belong to the sea and so their study promises to reveal new information about the structure of the proton. Even prior to QCD, hadronic models emphasized the importance of the presence of mesons (e.g., Ref. [5]) and therefore, as was realized later, antiquarks, in the physical state of a proton or neutron. Despite this, the initial naive expectation was that the sea was formed predominantly by gluons splitting into quark-antiquark pairs. Indeed, several authors assumed that at some low momentum scale the sea quarks and gluons vanish and all the sea and the glue at high momentum scales are generated by gluon radiation and then gluon splitting. They were able to describe successfully the existing data in the late 1970's (e.g., Ref. [6,7]).

In the early 1990's the New Muon Collaboration (NMC) reported measurements of the deep inelastic structure functions (F_2) of hydrogen (H) and deuterium (D)[8,9] at $0.004 < x < 0.8$. The cross section for deep inelastic scattering measures the charge–squared-weighted sum of the quark and antiquark distributions, in this case at an average scale of $4\,(\text{GeV}/c)^2$. The integrals of the parton distributions of the proton (p) and neutron (n) were assumed to have charge symmetry, $\int_0^1 u_p(x)dx = \int_0^1 d_n(x)dx$, with similar integrals for the other quark flavors, and nuclear effects in deuterium were assumed to be small $\left(F_2^D = F_2^p + F_2^n\right)$. In that case their measurements and their estimate of the unmeasured region led NMC to conclude

$$\int_0^1 \frac{dx}{x} \left[F_2^p(x) - F_2^n(x)\right] \tag{1}$$
$$= \frac{1}{3} + \frac{2}{3} \int_0^1 dx \left[\bar{u}(x) - \bar{d}(x)\right]$$
$$= 0.235 \pm 0.026,$$

and thus the integral of $\bar{d}(x)$ is greater than the integral of $\bar{u}(x)$:

$$\int_0^1 dx \left[\bar{d}(x) - \bar{u}(x)\right] = 0.147 \pm 0.039, \tag{2}$$

where $\bar{u}(x)$ and $\bar{d}(x)$ are the distributions of up and down antiquarks in the proton respectively.

The Drell-Yan process in hadron-hadron collisions is a reaction where a quark and an antiquark annihilate into a virtual photon and that virtual photon decays into a lepton-antilepton pair.[10] One can isolate the antiquark distributions from the Drell-Yan cross section making use of this property. At lowest order the Drell-Yan cross section is given by

$$\frac{d^2\sigma}{dx_b\,dx_t} = \frac{4\,\pi\,\alpha^2}{9\,s\,x_b\,x_t} \sum_q e_q^2 \left[q(x_b)\bar{q}(x_t) + \bar{q}(x_b)q(x_t)\right], \tag{3}$$

where x_b and x_t are the momentum fractions of the beam and target partons participating in the reaction, e_q is the electrical charge of quark flavor q, α is the fine structure constant, and s is the square of the center of mass energy of the beam and target. In a Drell-Yan measurement at CERN, the NA51 collaboration confirmed[11] that $\bar{d}(x)$ is larger than $\bar{u}(x)$ at an average x value of 0.18.

When a Drell-Yan experiment is performed with a proton beam and kinematic acceptance that selects events with x_b in the valence-quark dominated region and with Feynman x, $x_F \equiv x_b - x_t \gg 0$, the first term in equation 3 dominates. The charge-squared weighting and the fact that $u_v(x)$ is approximately $2d_v(x)$ for the valence quark distributions of the proton beam means that the measurement is, by a factor of approximately eight, more sensitive to $\bar{u}$ quarks in the target than $\bar{d}$. The renormalization and factorization scales for the extraction of parton distributions are usually chosen as the mass squared of the virtual photon times the speed of light squared, $M^2c^2 = x_b x_t s/c^2 - P_T^2$ where P_T^2 is the square of the transverse momentum of the virtual photon and is usually small compared to M^2c^2. Using charge symmetry,[12] to relate proton and neutron parton distributions, $\left(u_p(x) = d_n(x),\, d_p(x) = u_n(x),\, \bar{u}_p(x) = \bar{d}_n(x)\right)$,

$\bar{d}_p(x) = \bar{u}_n(x)$), as is assumed by almost all the global parton distribution fits, and assuming that the nuclear corrections in the deuteron are small, as supported by calculations,[13,14] the ratio of the Drell-Yan cross section on a deuterium target to that on a hydrogen target, $\sigma_D/\sigma_H \approx (\sigma_p + \sigma_n)/\sigma_p \approx 1 + \bar{d}_p(x)/\bar{u}_p(x)$, almost directly measures $\bar{d}(x_t)/\bar{u}(x_t)$.

The Fermilab NuSea/E866 collaboration[4] (whose results are displayed in Fig. 1 and Fig. 2) was able to measure the x_t dependence of the $\sigma_D/(2\sigma_H)$ ratio with an 800 GeV proton beam in the kinematic range $0.015 < x_t < 0.35$ and by extrapolating the results to $x_t = 0$ and $x_t = 1$ obtained a value of the $\int_0^1 dx\left[\bar{d}(x) - \bar{u}(x)\right] = 0.118 \pm 0.012$ at an average scale of $54\,(\text{GeV}/c)^2$. The HERMES collaboration also measured part of this integral and obtained results[15] consistent with NMC and NuSea. One feature of the NuSea results, with admittedly limited statistics, is the suggestion that the ratio of $\bar{d}(x)/\bar{u}(x)$ began to decrease for $x > 0.2$, reaching a value of $\bar{d}(x)/\bar{u}(x) = 0.35 \pm 0.40$ at $x = 0.31$, as seen in Fig. 2.

There are a variety of mechanisms that may account for the antiquark flavor asymmetry of the proton; recent reviews include Ref. [16,17]. Pauli blocking[18] may lead to a flavor asymmetry as the extra u valence quark Pauli blocks some u-$\bar{u}$ pairs from forming, but the x dependence and even the sign of this mechanism are debated in the literature[19,20]. A related approach involves statistical models.[21,22] Another class including chiral soliton models[23] and meson-baryon models emphasize mesonic degrees of freedom in proton structure.[24–26] These latter models (statistical, chiral soliton, and meson-baryon) each attempt to describe the entire non-perturbative composition of the proton. A common feature of these models is a rise in the $\bar{d}/\bar{u}$ flavor asymmetry with x. While at low x this behavior reproduces the NuSea data, none of these models are able to reproduce the fall-off at higher x observed by NuSea. The only *ab initio* technique for calculating the parton distributions of the proton is lattice QCD (recently reviewed by Lin *et al.*[27]). At this time, the lattice results for both quarks and antiquarks are still not in quantitative agreement with global fits of parton distributions to experimental data and the systematic errors are still being evaluated.

The SeaQuest experiment (E906) at Fermi National Accelerator Laboratory (Fermilab) was designed to investigate the flavor asymmetry at higher x_t values than NuSea with the newly constructed experimental apparatus that is described in detail in Ref. [28]. With a proton beam at an energy of $120\,\text{GeV}$, liquid hydrogen and deuterium targets, and a 10 T-m focusing magnet after the target region, the experiment was optimized for the study of the target antiquarks in the intermediate region, with x_t around 0.3, by detecting the $\mu^+\mu^-$ pairs from decays of the virtual photons produced in the Drell-Yan process. The proton beam was extracted from the Fermilab Main Injector using slow-spill extraction for $4\,\text{s}$ every $60\,\text{s}$. The microstructure of the beam consisted of 1 ns long bunches of approximately 0 to 80,000 protons at a $53\,\text{MHz}$ repetition rate. About 6×10^{12} protons were incident on the target in the four second long extraction period. A Cherenkov detector, installed in the beam line, measured the number of protons for each bunch and allowed high intensity bunches (usually greater than 64,000 protons) to be vetoed.

The data analysis procedure is described in the "Methods Summary" section. The primary challenge in the data analysis consisted of rate dependent effects that were exacerbated by large fluctuations in the bunch beam intensity. The average duty factor, defined as the square of the average intensity divided by the average of the intensity squared during the beam spill, $\langle I \rangle^2/\langle I^2 \rangle$, of the beam ranged between 20% and 40%. These intensity variations had two primary consequences: first, a variation in the track reconstruction efficiency, and second, a change in the rate of accidental coincidences. Rather than trying to separate and model all the rate dependent effects, the ratios of the yields on deuterium, Y_D, and hydrogen, Y_H, were analyzed by fitting the ratios as a function of x_t and bunch intensity, I, to a functional form. For the final analysis, the form $\frac{Y_D(x_t,I)}{2Y_H(x_t,I)} = R_{x_t} + aI + bI^2$, was chosen based on the Akaike information criterion.[29] Here, x_t is the bin average of each x_t bin, R_{x_t} is the fitted cross section ratio at zero intensity, $\sigma_D/(2\sigma_H)$, and a and b are constants fitted to the entire x_t range. The results from this form were compared with other functional forms such

Article

Figure 1 | Ratios of σ_D to $2\sigma_H$.
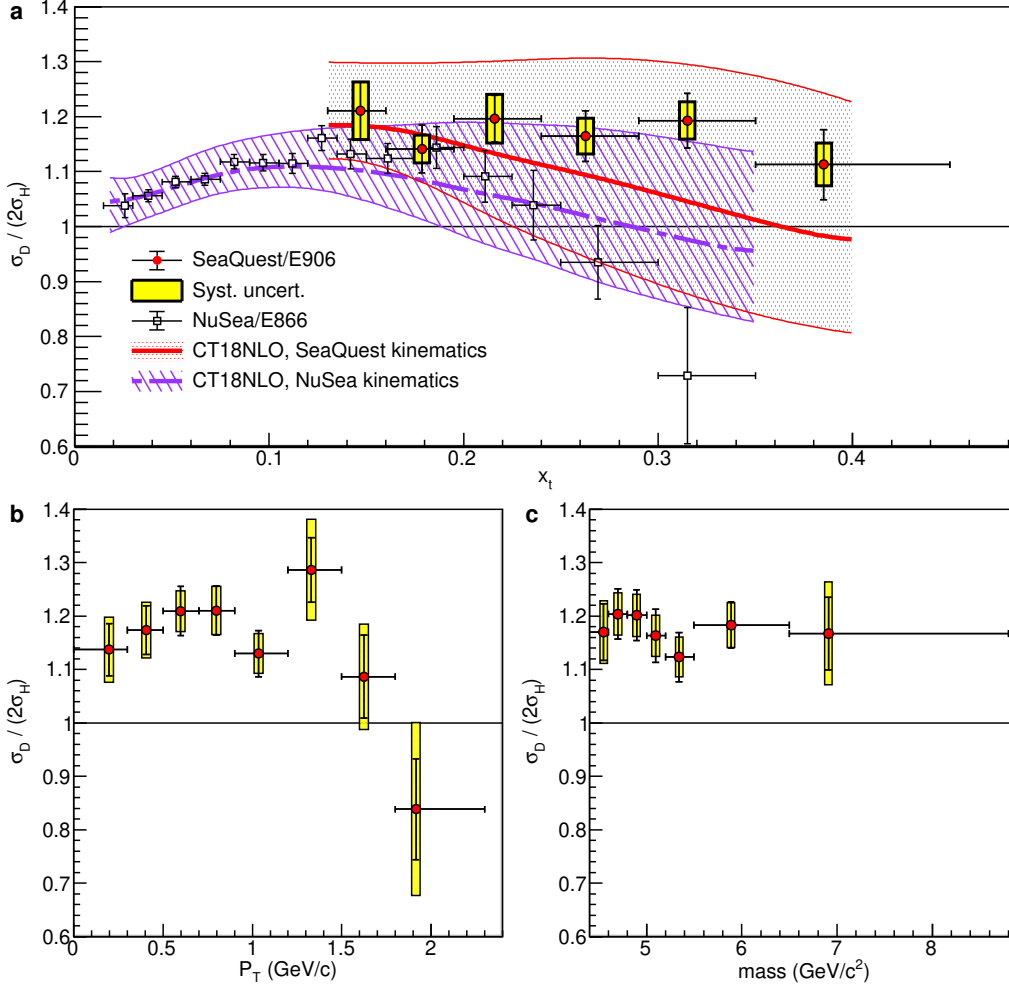
Ratios of $\sigma_D(x_t)$ to $2\sigma_H(x_t)$ (Upper plot, red filled circles) with their statistical (vertical bars) and systematic (yellow boxes) uncertainties as a function of x_t, and (lower plots) as functions of transverse momentum, P_T, and mass, M, of the virtual photon. The cross section ratios are defined as the ratio of luminosity-corrected yields from the hydrogen and deuterium targets. Also shown (open black squares) are the results of the NuSea experiment[4] for the x_t dependence with statistical uncertainties only. NuSea also reports an overall 1% common systematic uncertainty. The mass scale of the NuSea data is up to 50% larger than that of the SeaQuest data and the distributions in the other kinematic variable, x_b, are slightly different due to the differing beam energies and acceptances of the experiments. These differences imply that the cross section ratios do not have to be identical. This is demonstrated by the red solid and violet long-short dashed curves representing NLO calculations of the cross section ratio at SeaQuest and NuSea kinematics using CT18 parton distributions[30]. The horizontal bars on the data points indicate the width of the bins.

as $\frac{Y_D(x_t,I)}{2Y_H(x_t,I)} = R_{x_t} + (a_0 + a_1 x_t)\, I + (b_0 + b_1 x_t)\, I^2$ and the differences between the results with these two forms were taken as part of the x_t dependent systematic error. The only other significant systematic error was an overall 2% uncertainty related to the relative beam normalization. The cross section ratios, $\sigma_D/(2\sigma_H)$, defined as the ratio of luminosity-corrected yields from the hydrogen and deuterium targets, measured by SeaQuest as a function of x_t, M and P_T are shown in Fig. 1 and tabulated in Tab. 1, along with the average x_b, M, and P_T. (The cross section ratios values for the P_T and M plots in Fig. 1 are given in Extended Data Tabs. 3, and 4.) The average values of x_t, x_b, M, and P_T in each bin are the same for deuterium and hydrogen within uncertainties and not corrected for acceptance. The statistical uncertainty on each data point is the uncertainty returned from the fit on the zero intensity (R) parameter, where the uncertainty of the individual ratios only included the counting uncertainties of Drell-Yan yields on the hydrogen and deuterium targets. Since the parameters a and b are fit over all x_t bins, the statistical uncertainty is correlated by 40% - 70% among x_t bins. The covariance matrix is given in the Eq. 9 of the Methods section. The systematic uncertainty is fully correlated among all x_t bins. The observation that both M and P_T (for all but the very highest P_T bins) distributions for deuterium and hy-

drogen have the same shapes helps to confirm that the acceptances are very similar for each target. Also shown in Fig. 1 are the results as a function of x_t from NuSea[4] and the cross section ratio calculated in next-to-leading order (NLO) with CT18[30] parton distribution. The NuSea results are at a different scale, $54\,(\mathrm{GeV}/c)^2$, than the SeaQuest results $22\text{-}40\,(\mathrm{GeV}/c)^2$. The cross section ratios depend on both x_b and x_t, and due to the differing beam energies and acceptances, the x_b distributions are slightly different for SeaQuest and NuSea, the effects of which are shown in Fig. 1 and Extended Data Fig. 3.
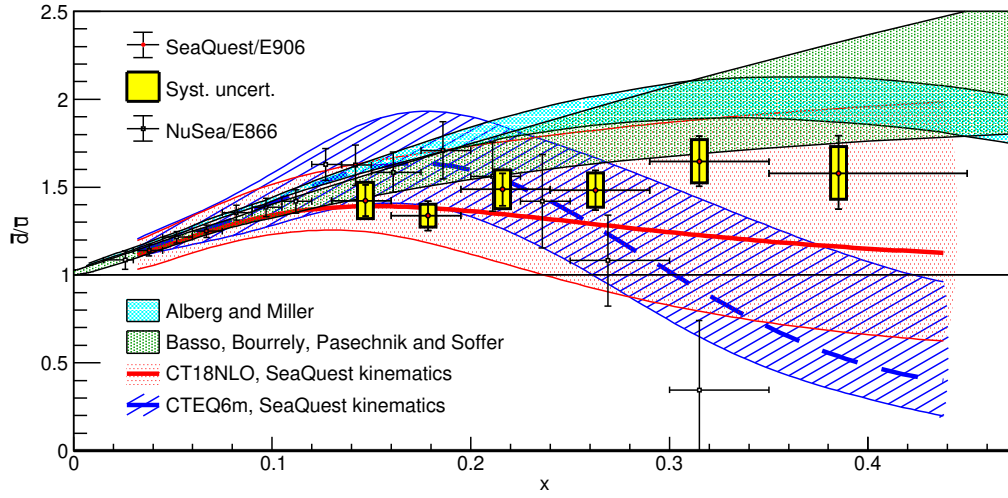
To extract $\bar{d}(x)/\bar{u}(x)$, next to leading order calculations of $\sigma_D/(2\sigma_H)$ were carried out starting from several NLO global fits to the parton distributions (CT10[31], CT14[32], CT18[30], MMHT2014[33]). No nuclear correction for deuterium was applied but a systematic uncertainty of $(0.5+5x_t)\%$ is included based on the range of deviation from unity found in Ref. [13,14]. Holding all other parton distributions fixed including the sum $\bar{d}(x)+\bar{u}(x)$, the ratio $\bar{d}(x)/\bar{u}(x)$ for each x_t bin was varied until the results converged on the measured cross section ratios. The correlations of the statistical uncertainties of $\sigma_D/(2\sigma_H)$ were propagated through the $\bar{d}/\bar{u}$ extraction. The dependence on the value for the

Article

Table 1 | Ratios of σ_D to $2\sigma_H$.

x_t bin	$\langle x_t \rangle$	$\langle x_b \rangle$	$\langle M \rangle$ (GeV/c^2)	$\langle P_T \rangle$ (GeV/c)	$\sigma_D/(2\sigma_H) \pm$ stat. $\pm$ syst.	δx_t
$0.130 - 0.160$	0.147	0.688	4.71	0.651	$1.211 \pm 0.052 \pm 0.053$	0.013
$0.160 - 0.195$	0.179	0.611	4.88	0.717	$1.141 \pm 0.043 \pm 0.025$	0.016
$0.195 - 0.240$	0.216	0.554	5.11	0.757	$1.196 \pm 0.042 \pm 0.044$	0.019
$0.240 - 0.290$	0.263	0.519	5.46	0.786	$1.165 \pm 0.046 \pm 0.032$	0.022
$0.290 - 0.350$	0.315	0.498	5.87	0.785	$1.193 \pm 0.050 \pm 0.034$	0.026
$0.350 - 0.450$	0.385	0.477	6.36	0.776	$1.113 \pm 0.064 \pm 0.039$	0.030

Cross section ratios $\sigma_D/(2\sigma_H)$ binned in x_t with their statistical and systematic uncertainties and the average values for the kinematic variables of each x_t bin. The cross section ratios are defined as the ratio of luminosity-corrected yields from the hydrogen and deuterium targets. The final column is the experimental resolution in x_t as determined by Monte Carlo simulations.

Figure 2 | Ratios of $\bar{d}(x)$ to $\bar{u}(x)$.

Ratios of $\bar{d}(x)$ to $\bar{u}(x)$ in the proton (red filled circles) with their statistical (vertical bars) and systematic (yellow boxes) uncertainties extracted from the present data based on next-to-leading order calculations of the Drell-Yan cross sections. Also shown in the open black squares are the results obtained by the NuSea experiment with statistical and systematic uncertainties added in quadrature[4]. The cyan band shows the predictions of the meson-baryon model of Alberg and Miller[26] and the green band shows the predictions of the statistical parton distributions of Basso, Bourrely, Pasechnik and Soffer[22]. The red solid (blue dashed) curves show the calculated ratios of $\bar{d}(x)$ to $\bar{u}(x)$ with CT18[30] (CTEQ6[36]) parton distributions at the scales of the SeaQuest results. The horizontal bars on the data points indicate the width of the bins.

ratio $\bar{d}(x)/\bar{u}(x)$ above the measured x region was estimated by varying this value from 1.0 to 0.5 and 2.0. The spread of the results due to the choice of initial parton distributions was always less than half of the statistical error. Each x_t bin was subdivided into multiple x_b sub-bins. The cross sections for hydrogen and deuterium were calculated separately for each sub-bin at the $\langle x_t \rangle$, $\langle x_b \rangle$, and $\langle M \rangle$ of that sub-bin and an acceptance weighted sum was used to determine the final cross section. These distributions are given in Extended Data Tab. 5. Calculations using only one average x_t and x_b for each x_t bin were less reliable. It was also found that a leading order extraction of $\bar{d}(x)/\bar{u}(x)$ using leading order parton distributions and cross section calculations gave very similar results for the ratios compared to the NLO results.

The resulting ratios of $\bar{d}(x)$ to $\bar{u}(x)$ distributions starting with the CT18 distributions are given in Table 2 at the scale of each x_t bin and displayed in Fig. 2, and compared there with the NuSea results. The trends between the two experiments at higher x_t are quite different. No explanation has been found yet for the differing results, even though there is a small overlap in the members of the NuSea and SeaQuest collaborations. The present data are reasonably described by the predictions of the statistical parton distributions of Basso, Bourrely, Pasechnik and Soffer[22] or by the chiral effective perturbation theory of Alberg and Miller,[26] also shown in Fig. 2. These two calculations em-

phasize rather different non-perturbative mechanisms that lead to the differences in $\bar{d}(x)$ and $\bar{u}(x)$. The present data show that $\bar{d}$ is greater than $\bar{u}$ for the entire x range measured by this experiment. This provides important support for these and other non-perturbative mechanisms of the QCD structure of the proton that were disfavored by the NuSea results.

The next major step to help distinguish between the various models is to measure how much the spin and angular momentum of the antiquarks contribute to the total spin of the proton. It has long been realized that these models make rather different predictions for the contribution of the total spin of the proton carried by the antiquarks.[16,17] For example, meson-nucleon models predict little spin is carried by the antiquarks, the statistical model predicts the difference in spin $\Delta\bar{d}(x) - \Delta\bar{u}(x) = -\left[\bar{d}(x) - \bar{u}(x)\right]$ and chiral soliton models[23] predict $\Delta\bar{d}(x) - \Delta\bar{u}(x) = -5/3\left[\bar{d}(x) - \bar{u}(x)\right]$. Experiments are planned or underway at Fermilab, the Thomas Jefferson National Accelerator Facility, Brookhaven National Laboratory, and the European Organization for Nuclear Research to pursue this goal.[16,17]

These results impact the reach of a p-p collider like the Large Hadron Collider for new physics. For example production of high mass Z' and W' particles has been shown to be dominated by light

Article

Table 2 | Ratios of $\bar{d}(x)$ to $\bar{u}(x)$.

$\langle x \rangle$	$\bar{d}(x)/\bar{u}(x)$	$\pm$ stat.	$\pm$ syst.
0.147	1.423	$+0.089$ -0.089	$+0.104$ -0.103
0.179	1.338	$+0.083$ -0.085	$+0.065$ -0.065
0.216	1.487	$+0.092$ -0.092	$+0.111$ -0.110
0.263	1.482	$+0.114$ -0.113	$+0.098$ -0.097
0.315	1.645	$+0.144$ -0.121	$+0.125$ -0.121
0.385	1.578	$+0.214$ -0.203	$+0.153$ -0.148

Ratios of $\bar{d}(x)$ to $\bar{u}(x)$ with their upper and lower statistical and systematic uncertainties. The analysis was based on the present cross section ratio data, and next-to-leading order calculations of the Drell-Yan cross sections using CT18 parton distributions for all except the ratio of $\bar{d}(x)$ to $\bar{u}(x)$. The systematic uncertainty is fully correlated among all x bins. The systematic uncertainty does not include a contribution from the choice of the base (CT18) pdf, which is small if added in quadrature to the other systematic uncertainties.

quark fusion.[34] Calculations with the contrasting statistical distributions (CTEQ6 distributions), which mimic the present (NuSea) data, show that the ratio $\bar{d}/\bar{u}$ is relatively insensitive to scale in each case. At mass scales of 4-5 TeV/c^2 just above current limits on Z' production,[35] cross sections driven by $u_v(x_1)\,\bar{d}(x_2)$ fusion with $x_1 \approx x_2$ in the region of 0.3-0.4 will be enhanced based on the present results and cross sections driven by $u_v(x_1)\,\bar{u}(x_2)$ will be diminished, compared to those calculated with the central values of previous parton distributions.

1. Ji, X., Yuan, F. & Zhao, Y. Proton spin after 30 years: what we know and what we don't? (2020). To appear in *Nature Reviews*, 2009.01291.
2. Yang, Y.-B. *et al.* Proton Mass Decomposition from the QCD Energy Momentum Tensor. *Phys. Rev. Lett.* **121**, 212001 (2018).
3. Ross, D. & Sachrajda, C. Flavour symmetry breaking in antiquark distributions. *Nuclear Physics B* **149**, 497 – 516 (1979).
4. Towell, R. S. *et al.* Improved measurement of the $\bar{d}/\bar{u}$ asymmetry in the nucleon sea. *Phys. Rev. D* **64**, 052002 (2001).
5. Fermi, E. & Marshall, L. On the interaction between neutrons and electrons. *Phys. Rev.* **72**, 1139 (1947).
6. Glück, M. & Reya, E. Dynamical determination of parton and gluon distributions in quantum chromodynamics. *Nuclear Physics B* **130**, 76 – 92 (1977).
7. Parisi, G. & Petronzio, R. On the breaking of Bjorken scaling. *Phys. Lett. B* **62**, 331 (1976).
8. Amaudruz, P. *et al.* Gottfried sum from the ratio F_2^n/F_2^p. *Phys. Rev. Lett.* **66**, 2712–2715 (1991).
9. Arneodo, M. *et al.* Reevaluation of the Gottfried sum. *Phys. Rev. D* **50**, R1–R3 (1994).
10. Drell, S. D. & Yan, T.-M. Massive lepton pair production in hadron-hadron collisions at high-energies. *Phys. Rev. Lett.* **25**, 316 (1970). [Erratum: Phys. Rev. Lett. **25**, 902 (1970)].
11. Baldit, A. *et al.* Study of the isospin symmetry breaking in the light quark sea of the nucleon from the Drell-Yan process. *Phys. Lett. B* **332**, 244 (1994).
12. Londergan, J. T. & Thomas, A. W. The validity of charge symmetry for parton distributions. *Prog. Part. Nucl. Phys.* **41**, 49 (1998).
13. Kamano, H. & Lee, T.-S. H. Pion-exchange and Fermi-motion effects on the proton-deuteron Drell-Yan process. *Phys. Rev. D* **86**, 094037 (2012).
14. Ehlers, P. J., Accardi, A., Brady, L. & Melnitchouk, W. Nuclear effects in the proton-deuteron Drell-Yan process. *Phys. Rev. D* **90**, 014010 (2014).
15. Ackerstaff, K. *et al.* Flavor asymmetry of the light quark sea from semi-inclusive deep-inelastic scattering. *Phys. Rev. Lett.* **81**, 5519 (1998).
16. Chang, W.-C. & Peng, J.-C. Flavor structure of the nucleon sea. *Prog. Part. Nucl. Phys.* **79**, 95 (2014).
17. Geesaman, D. F. & Reimer, P. E. The sea of quarks and antiquarks in the nucleon. *Rep. Prog. Phys.* **82**, 046301 (2019).
18. Field, R. D. & Feynman, R. P. Quark elastic scattering as a source of high transverse momentum mesons. *Phys. Rev. D* **15**, 2590 (1977).
19. Ross, D. A. & Sachrajda, C. T. Flavor symmetry breaking in anti-quark distributions. *Nucl. Phys. B* **149**, 497 (1979).
20. Steffens, F. M. & Thomas, A. W. Flavor asymmetry of the nucleon sea. *Phys. Rev. C* **55**, 900 (1997).
21. Bourrely, C. & Soffer, J. New developments in the statistical approach of parton distributions: tests and predictions up to LHC energies. *Nucl. Phys. A* **941**, 307 (2015).
22. Basso, E., Bourrely, C., Pasechnik, R. & Soffer, J. The Drell-Yan process as a testing ground for parton distributions up to LHC. *Nucl. Phys. A* **948**, 63 (2016).
23. Pobylitsa, P. V., Polyakov, M. V., Goeke, K., Watabe, T. & Weiss, C. Isovector unpolarized quark distribution in the nucleon in the large-N_c limit. *Phys. Rev. D* **59**, 034024 (1999).
24. Thomas, A. W. A limit on the pionic component of the nucleon through SU(3) flavor breaking in the sea. *Phys. Lett. B* **126**, 97 (1983).
25. Alberg, M. & Miller, G. A. Taming the pion cloud of the nucleon. *Phys. Rev. Lett.* **108**, 172001 (2012).
26. Alberg, M. & Miller, G. A. Chiral light front perturbation theory and the flavor dependence of the light-quark nucleon sea. *Phys. Rev. C* **100**, 035205 (2019).
27. Lin, H.-W. *et al.* Parton distributions and lattice QCD calculations: a community white paper. *Prog. Part. Nucl. Phys.* **100**, 107 (2018).
28. Aidala, C. A. *et al.* The SeaQuest spectrometer at Fermilab. *Nucl. Inst. Meth. A* **930**, 49 (2019).
29. Akaike, H. A new look at the statistical model identification. *IEEE Trans. on Auto. Control* **19**, 716 (1974).
30. Hou, T.-J. *et al.* New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC (2019). arXiv:1912.10053.
31. Lai, H.-L. *et al.* New parton distributions for collider physics. *Phys Rev. D* **82**, 074024 (2010).
32. Dulat, S. *et al.* New parton distribution functions from a global analysis of quantum chromodynamics. *Phys. Rev. D* **93**, 033006 (2016).
33. Harland-Lang, L. A., Martin, A. D., Motylinski, P. & Thorne, R. G. Parton distributions in the LHC era: MMHT 2014 PDFs. *Eur. Phys. J. C* **75**, 204 (2015).
34. Agashe, K. *et al.* CERN LHC signals for warped electroweak neutral gauge bosons. *Phys. Rev. D* **76**, 115015 (2007).
35. Sirunyan, A. *et al.* Search for a W' boson decaying to a τ lepton and a neutrino in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Phys. Lett. B* **792**, 107 – 131 (2019).
36. Pumplin, J. *et al.* New generation of parton distributions with uncertainties from global QCD analysis. *JHEP* **07**, 012 (2002).

Author Contributions P.E.R and D.F.G are the co-spokespersons of

Article

Article

Methods

For the measurement of the pp and pd cross section ratios, 50.8 cm long liquid hydrogen (0.069 interaction lengths) and liquid deuterium targets (0.116 interaction lengths) and an empty target flask were used. The targets were interchanged every few minutes to substantially reduce time-dependent systematic effects.

The SeaQuest spectrometer was constructed for the measurement of muon tracks in the forward region (laboratory angles less than 0.1 radian). It was composed of two magnets, four detector stations, each consisting of fast trigger detectors and drift chambers, distributed over 25 m along the beam direction with a 1 m thick iron muon identification wall before the final detector station. The first magnet provided a 3.07 GeV/c transverse momentum kick between the target and the first detector station to enhance the acceptance for muon pairs resulting from the decay of high-mass virtual photons and to reduce the acceptance for the large background of low mass (less than 4 GeV/c^2) virtual photon events. It was filled with solid iron to absorb the proton beam and all other hadrons and electrons produced in either the target or this beam dump. A second magnet (with a 0.41 GeV/c transverse momentum kick) located between the first and second detector station provided for charge and momentum measurements of the muons. The iron hadron absorber between the third and fourth stations was used to establish the identification of muons.

Opposite sign muon pairs were combined into di-muon candidates. The muons of each candidate were tracked back through the spectrometer to find if they emerged from a common vertex along the beam path and near the target. The resolution of the vertex location was about 30 cm along the beam direction compared with the 170 cm separation between the target center and the average interaction point of protons in the solid iron magnet. Events identified as coming from the target were re-fit using the target-center vertex location, and the di-muon mass, M, longitudinal momentum, P_L, in the laboratory frame, and transverse momentum, P_T, were determined. With this information, the fractional momentum of the beam and target quarks participating in the reaction were calculated as

$$x_b = \frac{p_{\text{target}} \cdot p_{\text{sum}}}{p_{\text{target}} \cdot (p_{\text{beam}} + p_{\text{target}})}, \tag{4}$$

$$x_t = \frac{p_{\text{beam}} \cdot p_{\text{sum}}}{p_{\text{beam}} \cdot (p_{\text{beam}} + p_{\text{target}})}, \tag{5}$$

where p_{target} and p_{beam} are the four momenta of the target and beam, respectively and p_{sum} is the sum of the four momenta of the positive and negative muons. The prominent J/ψ peak (resolution $\sigma = 0.21$ GeV/c^2) and the requirement that events come from the target or beam dump served to calibrate the field integral of the solid iron magnet. The mass spectra are shown in Extended Data Fig. 5. Detailed field maps coupled with Hall probe measurements served to calibrate the second magnet.

Various kinematic constraints were placed on the accepted events, the most important of which were to require the virtual photon mass to be greater than 4.5 GeV/c^2 and the primary vertex to be in the target region. The yields for each target were corrected by subtracting the appropriately weighted yield of the empty target flask. For much of the data sample, the deuterium target had a 8.4% $\pm$ 0.4% per molecule HD contamination and the yields were corrected accordingly. The beam normalization and the uncertainty in the rate dependence corrections were the dominant systematic errors. Other smaller contributions include the uncertainty of the deuterium target purity, uncertainties in the target density and the contribution of the tail of the J/ψ and ψ' to the di-muon mass spectrum greater than 4.5 GeV/c^2.

Instantaneous fluctuations in the beam intensity while the data were being collected presented the main challenge in the data analysis. These fluctuations occurred at the accelerator frequency of 53 MHz and led to a luminosity-normalized rate-dependent variation in the yield of events from the deuterium and hydrogen targets from a number of different sources. Several approaches were considered to account for this variation. Very generally, the simplest approach would have been to reject any event produced when the accelerator was above a certain, arbitrary (fairly low) threshold and absorb the remainder of the effect into a systematic uncertainty. This would have had a significant impact on the statistical significance of the data. A second approach would have been to model each individual effect in Monte Carlo, then parameterize individual effects, and finally apply the combined parameterizations to the data. The systematic uncertainty would need to account for the accuracy of the model and for any still unknown effects. For the present data, a third method was chosen that allowed the full statistical power of the data to be kept without requiring that each and every effect of the intensity variation to be modeled.

This method considers only the final result–the ratio of event yields between the two targets–as a function of intensity. For each bin in x_t, the cross section ratio was plotted as a function of the instantaneous beam intensity when that event occurred, as illustrated in Extended Data Fig. 4. The effect of the intensity dependence on the final result could then be parameterized from the measured data and then extrapolated to zero intensity. The simplicity of this method is that the data alone are being used to measure and correct for the intensity dependence. Since the bin boundaries in x_t are arbitrary relative to the beam intensity, a smooth, common parameterization for the intensity dependence is to be expected. A variety of parametric forms were compared to the data. One such form is

$$\frac{Y_D(x_t, I)}{2Y_H(x_t, I)} = R_{x_t} + aI + bI^2, \tag{6}$$

where is the $Y_{D(H)}$ luminosity-normalized, empty target subtracted yield of events from the deuterium (hydrogen) target. In this form, a and b are parameters of the fit that are common to all x_t bins describing the intensity, I, dependence, and R_{x_t} is the zero intensity intercept for that bin. The intercepts resulting from the simultaneous fit of all x_t bins gives the cross section ratio $\sigma_D/(2\sigma_H)$ for each bin. The common intensity parameters, a and b, correlate $\sigma_D/(2\sigma_H)$ for all bins and are also determined in the simultaneous fit. Other forms were also studied, including, for example

$$\frac{Y_D(x_t, I)}{2Y_H(x_t, I)} = R_{x_t} + (a_0 + a_1 x_t)I + (b_0 + b_1 x_t)I^2, \tag{7}$$

which allows for an x_t correlated intensity dependence. An example of a less conventional extrapolation form that was considered is

$$\frac{Y_D(x_t, I)}{2Y_H(x_t, I)} = R_{x_t} \cos\left(\frac{I}{a_0 + a_1 x_t}\right). \tag{8}$$

In addition, constraining either a or b values to zero and thus eliminating the I or I^2 dependence was explored. Using the Akaike information criterion, to avoid over parameterization, the form given in Eq. 6 was chosen for the extrapolation. The resulting fits from three representative x_t bins are shown in Extended Data Fig. 4. A comparison with a fit to the parameterization in Eq. 7 was used to estimate the systematic uncertainties. The covariance matrix for the intercepts, R_{x_t} resulting from the fit to Eq. 6 is

$$\begin{pmatrix} 2.70 & 1.19 & 1.15 & 1.20 & 1.09 & 1.16 \\ 1.19 & 1.87 & 1.25 & 1.31 & 1.19 & 1.26 \\ 1.15 & 1.25 & 1.79 & 1.25 & 1.15 & 1.21 \\ 1.20 & 1.31 & 1.25 & 2.14 & 1.20 & 1.27 \\ 1.09 & 1.19 & 1.15 & 1.20 & 2.49 & 1.16 \\ 1.16 & 1.26 & 1.21 & 1.27 & 1.16 & 4.06 \end{pmatrix} \times 10^{-3}. \tag{9}$$

The same technique was independently applied to the data binned in transverse momenta, P_T, and mass, M.

The cross section ratios shown in Fig. 1 and listed in Tab. 1 are not corrected for acceptance. To compare any calculation with the present data, it is necessary to consider the SeaQuest spectrometer's acceptance in x_b. The appropriate theoretical cross section ratio may be calculated for a x_t bin i as

$$\left(\frac{\sigma_D}{2\sigma_H}\right)_i = \frac{\sum A_{ij}\sigma_D^{\text{calc}}(x_t, x_b, M)}{2\sum A_{ij}\sigma_H^{\text{calc}}(x_t, x_b, M)}, \tag{10}$$

where the subscript j denotes the jth x_b sub-bin of the ith x_t bin, and A_{ij} is the acceptance for that bin, tabulated in Extended Data Tab. 5. Finally, $\sigma_{D(H)}^{\text{calc}}(x_t, x_b, M)$ is the calculated cross section, where the dependence on x_t, x_b, and M has been made explicit. SeaQuest used code for the NLO calculation of $\sigma_{D(H)}^{\text{calc}}(x, M)$ provided by W. K. Tung of CTEQ.

Article

Extended Data

Figure 3 | Comparison of NuSea and SeaQuest to NLO calculations.
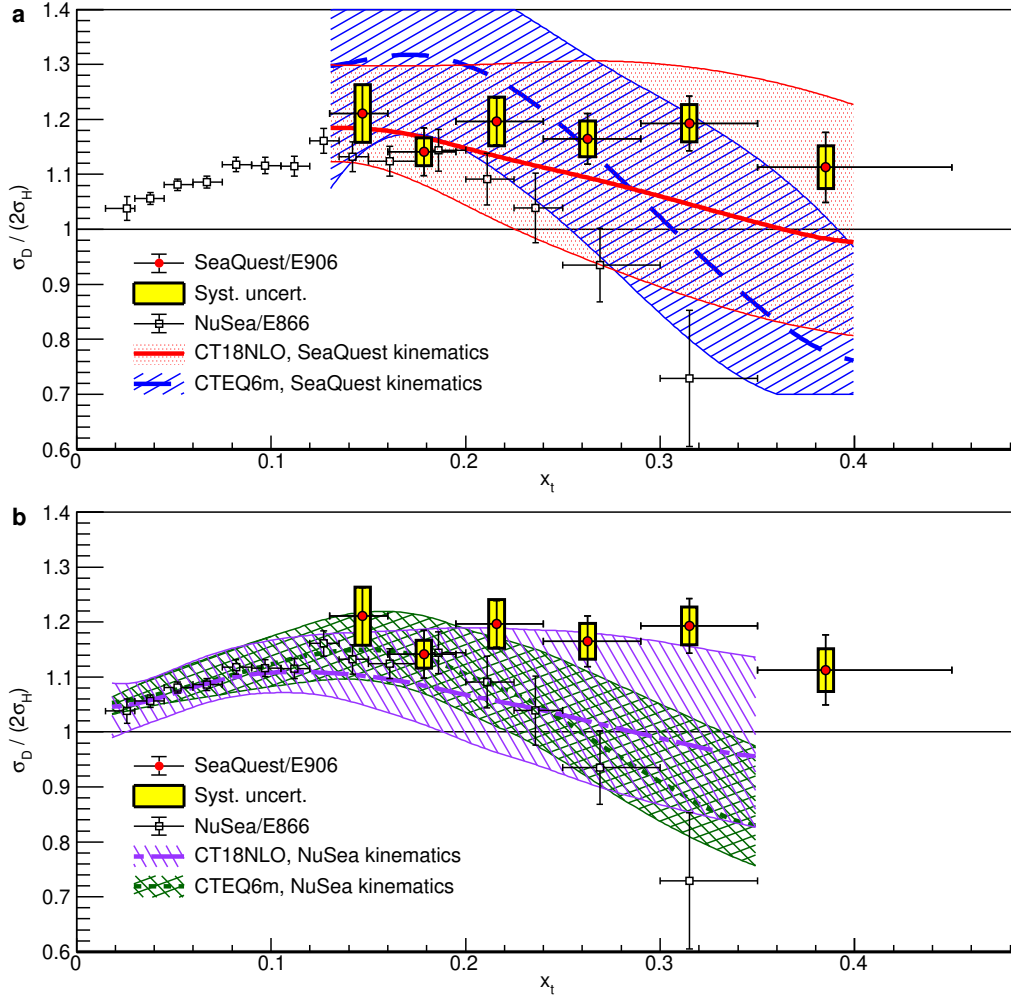
A comparison of the data from the present result and the NuSea measurements with NLO calculations made at the integrated kinematics of SeaQuest (a) and average kinematics NuSea (b) based on the CT18 and CTEQ6m parton distributions. Events in SeaQuest were produced by a 120 GeV proton beam while NuSea's were from an 800 GeV beam. In addition, the spectrometers, while similar in concept, had different acceptances. As a consequence, the cross section ratios, which convolve x_t with x_b, are expected to differ because of their distinct distributions in accepted x_b. These kinematic effects can clearly be seen by the difference in the curves. An acceptance table analogous to Extended Data Tab. 5 was not available for NuSea, so those calculations used $\langle x_t \rangle$, $\langle x_b \rangle$, and $\langle M \rangle$ of the NuSea data. Both CTEQ6m and CT18 have included the NuSea data in their global analysis, so calculations based on those PDFs are expected to agree better with the NuSea data. The red (violet) curve in the upper (lower) plot is the same as in the main paper's Fig. 1 upper plot and repeated here for comparison.

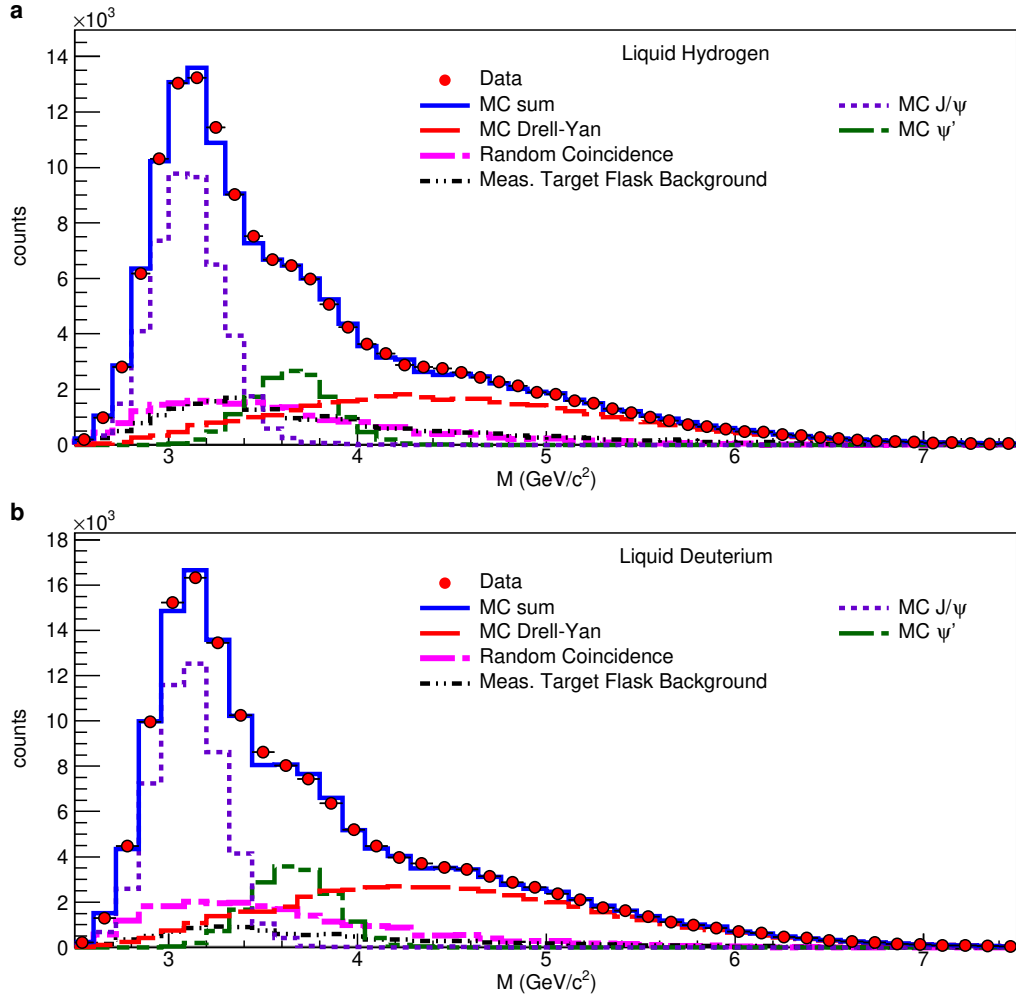### Figure 4 | Extrapolation to zero intensity.
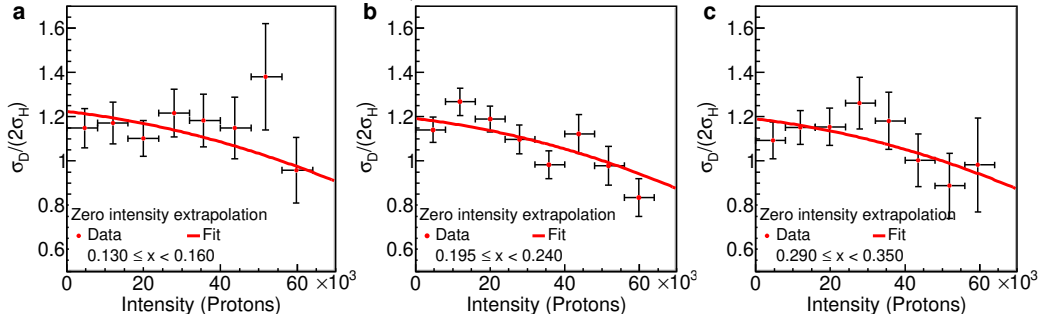
Extrapolation to zero intensity fits for representative x_t bins ($0.13 \leq x_t < 0.16$, $0.195 \leq x_t < 0.240$, and $0.290 \leq x_t < 0.350$). The I (intensity) and I^2 coefficients are common to all bins. The $\chi^2/\text{dof} = 38.7/40$ for the simultaneous fit of all x_t bins.

Figure 5 | Reconstructed invariant mass spectra.

a

b

The reconstructed muon pair invariant mass spectra for the liquid hydrogen (a) and liquid deuterium (b) targets. In the lower mass region, the predominant signal is produced by $J/\psi \rightarrow \mu^+\mu^-$ decay, followed by the $\mu^+\mu^-$ decay of the ψ'. The prominence of the J/ψ provides a calibration point for the absolute field of the solid iron magnet. At invariant masses above 4.5 GeV/c^2 the Drell-Yan process becomes the dominant feature. The data are shown as red points. Additionally, Monte Carlo (MC) simulated distributions of Drell-Yan, J/ψ, and ψ' along with measured random coincidence and empty target backgrounds are shown. The sum of these is shown in the blue solid curve labeled MC sum. The normalizations of the Monte Carlo and the random background were from a fit to the data.

Table 3 | Ratios of σ_D to $2\sigma_H$ as a function of P_T.

cP_T (GeV/c)	$c\langle P_T \rangle$ (GeV/c)	$\frac{\sigma_D}{2\sigma_H} \pm$ stat. $\pm$ syst.	cP_T (GeV/c)
$0.0 - 0.3$	0.198	$1.137 \pm 0.049 \pm 0.061$	0.161
$0.3 - 0.5$	0.405	$1.174 \pm 0.045 \pm 0.052$	0.177
$0.5 - 0.7$	0.599	$1.209 \pm 0.046 \pm 0.038$	0.188
$0.7 - 0.9$	0.797	$1.210 \pm 0.046 \pm 0.045$	0.194
$0.9 - 1.2$	1.035	$1.130 \pm 0.043 \pm 0.037$	0.198
$1.2 - 1.5$	1.330	$1.287 \pm 0.061 \pm 0.094$	0.201
$1.5 - 1.8$	1.625	$1.087 \pm 0.078 \pm 0.099$	0.206
$1.8 - 2.3$	1.915	$0.838 \pm 0.095 \pm 0.162$	0.204

Ratios of σ_D to $2\sigma_H$ with their statistical and systematic uncertainties as a function of transverse momentum, P_T. The cross section ratios are defined as the ratio of luminosity-corrected yields from the hydrogen and deuterium targets. The final column, δP_T is the experimental resolution in P_T as determined by Monte Carlo simulation.

Table 4 | Ratios of σ_D to $2\sigma_H$ as a function of M.

$\mathrm{c}M$ $\left(\mathrm{GeV/c^2}\right)$	$\mathrm{c}\langle M\rangle$ $\left(\mathrm{GeV/c^2}\right)$	$\dfrac{\sigma_D}{2\sigma_H} \pm$ stat. $\pm$ syst.	$\mathrm{c}\delta M$ $\left(\mathrm{GeV/c^2}\right)$
$4.4 - 4.6$	4.55	$1.170 \pm 0.053 \pm 0.059$	0.24
$4.6 - 4.8$	4.70	$1.204 \pm 0.047 \pm 0.039$	0.24
$4.8 - 5.0$	4.90	$1.202 \pm 0.048 \pm 0.039$	0.25
$5.0 - 5.2$	5.10	$1.163 \pm 0.050 \pm 0.039$	0.26
$5.2 - 5.5$	5.34	$1.123 \pm 0.046 \pm 0.037$	0.26
$5.5 - 6.5$	5.89	$1.183 \pm 0.043 \pm 0.042$	0.28
$6.5 - 8.8$	6.91	$1.167 \pm 0.068 \pm 0.096$	0.30

Ratios of σ_D to $2\sigma_H$ with their statistical and systematic uncertainties as a function of mass, M. The cross section ratios are defined as the ratio of luminosity-corrected yields from the hydrogen and deuterium targets. The final column, δM is the experimental resolution in M as determined by Monte Carlo simulation.

Table 5 | Spectrometer acceptance.

x_t \ x_b	0.30–0.35	0.35–0.40	0.40–0.45	0.45–0.50	0.50–0.55	0.55–0.60	0.60–0.65	0.65–0.70	0.70–0.75	0.75–0.80
0.130–0.160						0.0007 0.589 0.158 4.54	0.0064 0.628 0.153 4.60	0.0175 0.675 0.148 4.68	0.0304 0.723 0.144 4.77	0.0370 0.772 0.144 4.92
0.160–0.195				0.0007 0.489 0.191 4.56	0.0071 0.528 0.184 4.63	0.0188 0.576 0.178 4.74	0.0299 0.624 0.176 4.91	0.0366 0.673 0.176 5.09	0.0432 0.722 0.176 5.27	0.0471 0.772 0.175 5.45
0.195–0.240		0.0001 0.394 0.235 4.55	0.0023 0.433 0.225 4.65	0.0105 0.477 0.217 4.78	0.0205 0.524 0.216 4.99	0.0298 0.574 0.215 5.21	0.0384 0.623 0.215 5.43	0.0456 0.672 0.214 5.63	0.0510 0.722 0.215 5.84	0.0557 0.772 0.214 6.04
0.240–0.290		0.0015 0.383 0.267 4.76	0.0078 0.427 0.264 4.99	0.0176 0.475 0.263 5.24	0.0270 0.524 0.262 5.50	0.0364 0.574 0.262 5.75	0.0436 0.623 0.262 6.00	0.0499 0.672 0.262 6.24	0.0550 0.722 0.261 6.46	0.0591 0.771 0.262 6.69
0.290–0.350	0.0002 0.341 0.324 4.95	0.0035 0.379 0.319 5.18	0.0120 0.426 0.316 5.46	0.0207 0.475 0.316 5.76	0.0298 0.524 0.316 6.05	0.0379 0.574 0.315 6.33	0.0455 0.623 0.315 6.60	0.0518 0.673 0.315 6.85	0.0544 0.722 0.314 7.10	0.0568 0.771 0.314 7.34
0.350–0.450	0.0006 0.339 0.384 5.38	0.0052 0.377 0.390 5.72	0.0125 0.425 0.386 6.04	0.0203 0.474 0.386 6.38	0.0268 0.524 0.385 6.69	0.0336 0.573 0.384 7.00	0.0374 0.623 0.384 7.29	0.0405 0.672 0.384 7.58	0.0415 0.722 0.383 7.85	0.0413 0.771 0.382 8.11

The acceptance relative to a 4π detector and average kinematic values for bins in x_t and x_b. Each cell gives the acceptance (top), $\langle x_b\rangle$, $\langle x_t\rangle$, and $\langle \mathrm{mass}\rangle$ (bottom) for each sub-bin.