QCD form factors and hadron helicity nonconservation

John P. Ralston[1] and Pankaj Jain[2]
[1]*Department of Physics & Astronomy, University of Kansas Lawrence, Kansas 66045, USA*
[2]*Physics Department, I.I.T. Kanpur, India 208016*

(Received 7 February 2003; revised manuscript received 8 January 2004; published 25 March 2004)

Recent data for the ratio $R(Q)=QF_2(Q^2)/F_1(Q^2)$ surprised the community by disobeying expectations held for 50 years. It was widely thought that $R(Q)\sim 1/Q$ was also a rigorous prediction of PQCD. We examine the status of perturbative QCD predictions for helicity-flip form factors. Contrary to common belief, we find there is no rule of hadron helicity conservation for form factors. Numerical calculations support asymptotic analysis and extend it to the regime of laboratory momentum transfer Q^2. Quark orbital angular momentum, an important feature of the helicity flip processes, may play a role in all form factors at large Q^2, depending on the quark wave functions.

DOI: 10.1103/PhysRevD.69.053008 PACS number(s): 13.40.Gp, 12.38.Bx, 13.88.+e

I. THE NATURE OF HIGH ENERGY REACTIONS

There is an asymmetry in high energy reactions due to the Lorentz transformation. The spatial coordinate parallel to a boost direction is Lorentz contracted. The momentum fraction x of partons inside hadrons is thereby distributed over the entire possible range $0<x<1$. This phenomenon is dynamical, because a boost of interacting fields is dynamical, and the x dependence of wave functions cannot reliably be calculated in perturbative QCD. Instead, the x dependence of wave functions is extracted from experimental data.

Meanwhile the spatial coordinate $\vec{b}$ transverse to the boost direction is Lorentz invariant. The transverse coordinate has a certain calculability via perturbative QCD. There is a great deal of interest and controversy associated with the transverse coordinate in exclusive reactions. The transverse spatial coordinate can be probed in reactions sensitive to the angular momentum flow. In some reactions, the sum of the helicities going into a hadronic reaction is automatically conserved. This is the case of the proton's Dirac form factor F_1 in the high energy limit. When the sum of the helicities is not conserved, angular momentum conservation requires either extra constituents, or quark orbital angular momentum. This is the case of F_2, the proton's Pauli form factor.

The Jefferson Laboratory has observed $QF_2(Q^2)/F_1(Q^2)\sim$const up to the highest values of $Q^2\sim 5.8$ GeV^2 yet measured [1,2]. The data was initially very surprising, and the field may have reached a pivotal point in comparison with the quark model. For a long time it was held sacred that $Q^2F_2/F_1\sim$const at large Q^2. This rule appears to predate QCD. It has ancient origins in renormalization questions involving protons as elementary fundamental fields [3].

Meanwhile a perturbative QCD model assuming nonzero quark orbital angular momentum (OAM) predicted $QF_2(Q^2)/F_1(Q^2)\sim$const [4,5]. A relativistic quark model prediction [6] fits the flatness of $QF_2(Q^2)/F_1(Q^2)$ equally well. These papers countered the ancient wisdom, because they shared the common feature of quark orbital angular momentum in the wave functions.

The hypothesis of zero OAM sometimes appeals to the nonrelativistic quark model in the rest frame. Yet the quark model does not require separate conservation of spin and orbital angular momentum, and any naive expectation of small OAM based on atomic physics seems rather weak. Just as importantly, very little of high energy physics and PQCD starts in the rest frame. The perturbative quark wave functions are unrestricted in angular momentum content, except for Lorentz symmetry. Observation of nonzero quark OAM is a leading candidate to resolve the proton "spin crisis," which is the fact that the sum of the perturbative quark spins does not equal the spin of the proton. Consequently the Jefferson lab's data has even broader implications than the mystery of large Q^2 form factors.

Here we address whether $QF_2(Q^2)/F_1(Q^2)\sim$const is a transient feature of comparatively low energy experiments, or a fact destined to persist at higher Q^2. If the flatness of this ratio is due to quark OAM, will the ratio stay flat with increasing Q^2? The question leads us to reexamine the roots of the "hadron helicity conservation" rule [7]. We find that PQCD itself is rather neutral, because the calculations depend on the wave functions. The wave functions needed cannot be obtained by perturbative arguments enforcing any particular model, because the important correlations are nonperturbative. The simplest analysis gives an *inequality* between the powers governing helicity-flip processes and helicity conserving ones. However even the inequality is not general, and it is possible to find helicity flip dominating helicity conservation at large Q. As a result $QF_2/F_1\sim$const may extend to arbitrarily large values of Q^2, without violating anything sacred.

II. DEFINITIONS AND DISCUSSION

Proton form factors. The proton form factors F_1 and F_2 are defined by the standard relation,

$$\langle P',s'|J^\mu|P,s\rangle = \bar{N}(P',s')\left(\gamma^\mu F_1(Q^2)+i\sigma^{\mu\nu}q_\nu\frac{F_2(Q^2)}{2M}\right)N(P,s), \tag{1}$$

where $q=P'-P$, $Q^2=-q^2$, J^μ is the electromagnetic current of the proton and s and s' refer to the spins of the initial and final proton. F_1 and F_2 are the Dirac and Pauli form factors respectively.

Orbital angular momentum. By quark orbital angular momentum (OAM), we refer to an expansion in $SO(2)$ representations (commonly known as L_z states), with quantization axis aligned along the particle 3-momentum $\vec{P}$. The eigenstates of L_z are invariant under boosts along the z axis. We do not use spherical harmonics, and we treat the longitudinal coordinates as scaling variables. Let the transverse spatial coordinate be $\vec{b}=b(\cos\phi,\sin\phi)$. We expand operators or wave functions as

$$\psi=\sum_m e^{im\phi}\psi_m(x,b)$$

where x is the Feynman light-cone fraction of z (or "+") momentum. Perturbative wave functions are theory constructs, and fields at finite separation are technically not gauge invariant, in general. Some basic rules for classifying nonzero OAM wave functions and "dressing" for gauge invariance are developed in Ref. [8]. A 1993 development of generalized parton distributions was partly motivated in order for OAM to be expressed in a manifestly gauge-invariant formalism [9,10]. Observables are also gauge invariant when all the legs of diagrams are contracted properly.

Chirality versus helicity. The proton's Pauli form factor F_2 contains information on the orbital angular momentum of the quarks, but it is indirect. Strictly speaking, the amplitude $i\bar{N}(p',s')F_2(Q^2)\sigma^{\mu\nu}q_\nu N(p,s)$ represents the amplitude for *chirality* of the proton to flip under momentum transfer Q. The chirality (eigenvalue of γ_5) of quarks does not change under a vector gluon interaction to all orders of approximation.[1] Chirality is nearly an exact symmetry of PQCD, with changes in quark propagation order m/P, where m is the quark mass and P its momentum. The chirality also equals the helicity to order m/P. The mass m to be used in PQCD is the so called current quark mass of a few MeV. Adding another constituent to the scattering is also down by $O(1/Q^2)$. Then by this reasoning, at large $Q^2 > \mathrm{GeV}^2$, it is not possible to flip either the chirality or helicity of the proton with a virtual photon, unless there is internal quark orbital angular momentum present to satisfy the selection rules.

What is quark counting? We separate the quark-counting model of Brodsky, Farrar, and independently Matveev et al. [11] from the asymptotic short distance (ASD) model of Brodsky and Lepage [12,13]. The earlier theory is one of counting propagators and the number of scattered constituents. The latter ASD theory is much more detailed, imposing a certain factorization of the hard reaction into components made from the s-wave, $m=0$ Bethe-Salpeter wave functions. The factorization, and selection of priviledged wave functions, does not come by listing all diagrams and evaluating them. Instead the framework is developed by assuming the framework and classifying terms within the assumptions. The ASD approach is characterized by taking the zero-distance limit in the first step., and replacing the rest of the calculation by integrals over Feynman x fractions using "distribution amplitudes," or similar quantities with no transverse information. In comparison the Feynman rules instruct us to leave the longitudinal and transverse integrals coupled. Perform the integrals, and afterwards take the limit of large Q (if wanted). If the two limits are not the same (and they are not in general), then the ASD assumptions can fail.

The transverse coordinates are gone in the ASD approach, being evaluated within $1/Q$ of zero. HHC follows instantly as a test of the framework, independent of the wave functions used. By omitting OAM $m\neq 0$, the prediction for $F_2=0$, and one cannot recover any prediction for finite F_2 in the formalism. As far as we know, previous ASD predictions of F_2 are indirect and deduced by elimination: since $F_1\sim 1/Q^4$, and QF_2 could not be calculated, then $F_2\sim F_1/Q^2$ must lie in the detritus not calculated. Within the framework of a generalized ASD model a recent calculation [14] finds contributions to $F_2\sim 1/Q^6$. It is also possible to modify the model by including quark mass effects [15]. Our goal here is to understand the limit of arbitrarily light quarks and neglecting effects of order m_q/Q.

We separated quark counting from ASD because the two are not the same theory. Should one believe either? There are good indications from the scaling laws that quark counting has some truth. We have no religious commitment here, and the possibility that a fraction of the amplitude is due to "soft physics" must be given credit [16,17]. But amplitudes cannot all be soft, because the form factors are not seen to fall exponentially with Q. Indeed the quark-counting scaling laws generally work well. Meanwhile the inapplicability of HHC to F_2 or any other helicity flip reactions [18] shows we cannot use ASD.

Regarding generalized parton distributions. Generalized parton distributions (GPD) appear to be an ideal way to proceed. Lorentz covariance can be used to set up matrix elements and expressions for the form factor in an apparently model-independent way. As far as we know, form factors were the first instance of GPD, used in the paper of Soper [19] in 1977, which also contains a transformation to a particular transverse spatial coordinate. We will have occasion to revisit the conclusions of that paper in Sec. III F.

Despite our GPD-based predictions [9,4,5] for ratios and the welcome rediscovery of GPD in the field, we chose not to make them the vehicle for this analysis. The reason is that GPD are so general they do not immediately contain the information there are three quarks in the proton. To incorporate the information one can start with wave functions and integrate out all but two quark legs. Since our concern is precisely these integrations, we would have nothing to gain. We caution the reader, in any event, that the method of approximating integrals by ASD methods will have similar limit-interchange problems when clothed in GPD language. Diehl *et al.* [20] discuss general relations, and their formulas codify a relation of F_2 to quark OAM.

[1] The perturbative chiral anomaly can change chirality, but with magnitude too small to be relevant to our discussion.

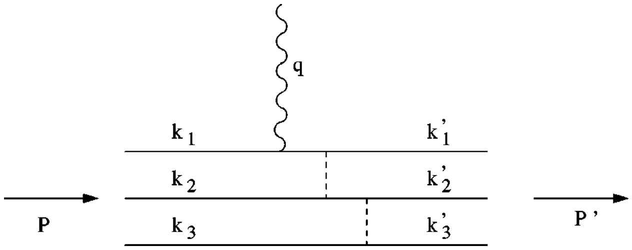

FIG. 1. Definitions of momenta for elastic scattering of a proton from a virtual photon. Here P and P' are the momenta of the initial and final proton, respectively, q is the momentum of the photon and k_i and k_i' the momenta of the initial and final state quarks, respectively.

Some familiar, but inexact assertions. Suppose we have nonzero OAM in the wave functions and we try to use the assertion of the ASD approach,

$$\Delta b \Delta Q_T > 1. \tag{2}$$

Here Δb is the resolved transverse quark separation, in a frame where the momentum transfer $Q \sim \Delta Q_T$ is transverse. We have written the relation like the uncertainty principle, to give it a chance to be seductive. We observe next that, merely from continuity, a wave function ψ_m carrying m units of angular momentum scales like b^m as $b \to 0$. Then under these assumptions each unit of orbital angular momentum of the quarks will lead to amplitudes suppressed by a corresponding power of $1/Q$ at large Q. This familiar assertion has been repeated endlessly in the literature, yet we will show that this type of counting does not represent QCD. It is seductive but it is not right.

Our approach. To address F_2 properly, one must go beyond short-distance to restore the transverse coordinate. This is called "impact-parameter factorization" [21,22]. This well-justified method has dominated recent attention in perturbative QCD [23–25,8,26].

The elastic scattering of a proton from a virtual photon is shown schematically in Fig. 1. In impact-parameter factorization, the three-quark contribution to the proton form factor can be written as

$$\begin{aligned}
&\overline{N}(P',s')\left(\gamma^\mu F_1(Q^2) + i\sigma^{\mu\nu} q_\nu \frac{F_2(Q^2)}{2M}\right) N(P,s) \\
&\quad = \int (d\vec{k}_T)(dx)(d\vec{k}_T')(dx') \\
&\qquad \times \overline{Y}_{\alpha'\beta'\gamma'}(P',\vec{k}_{Ti}',x_i',s') \\
&\qquad \times \Gamma^\mu_{\alpha'\beta'\gamma'\alpha\beta\gamma}(q,\vec{k}_{Ti},\vec{k}_{Ti}',x_i,x_i') Y_{\alpha\beta\gamma}(P,\vec{k}_{Ti},x_i,s).
\end{aligned} \tag{3}$$

We have factored the amplitude into products of a hard scattering kernel Γ^μ and soft initial and final state wave functions Y and Y' respectively. The wave function $Y_{\alpha\beta\gamma}(P,\vec{k}_{Ti},x,s)$ is the Fourier transform of the matrix element $\langle 0|\epsilon^{abc} u^a_\alpha(y_1) u^b_\beta(y_2) d^c_\gamma(0)|P,s\rangle$, where $|P,s\rangle$ is the proton state. The argument $\vec{k}_{Ti}$ of the initial wave function refers to the momenta of the three quarks, $\vec{k}_{T1}$, $\vec{k}_{T2}$ and $\vec{k}_{T3}$. Similarly the argument x_i refers to the longitudinal momentum fractions x_1,x_2,x_3 of the initial proton. The arguments $\vec{k}_{Ti}',x_i'$ of the final state wave function are similarly defined. We use the "brick-wall" frame for our calculations. The integration measures are given by, $(d\vec{k}_T) = d^2k_{T1}d^2k_{T2}d^2k_{T3}\delta^2(\vec{k}_{T1}+\vec{k}_{T2}+\vec{k}_{T3})$, $(dx) = dx_1dx_2dx_3 \times \delta(x_1+x_2+x_3)$. Note that Eq. (3) postpones any assumptions about the dominance of any particular integration region of $\vec{k}_T$ or x. If care is not taken, limit interchange errors leading to the ASD results can result.

To extract the contribution due to quark OAM, it is convenient to work directly with the coordinates $\vec{b}_i$ conjugate to the transverse quark momenta $\vec{k}_{Ti}$. We choose coordinates so that the third quark (down) lies at the origin [24], i.e. $\vec{b}_3 = 0$. The wave function $\tilde{Y}_{\alpha\beta\gamma}$ can be decomposed as a sum of terms [24]:

$$\begin{aligned}
\tilde{Y}_{\alpha\beta\gamma}(P,\vec{b}_i,x_i,s) = \frac{f_N}{8\sqrt{2}N_c}&(\mathcal{C}^1_{\alpha\beta\gamma}A^1(P,\vec{b}_i,x_i) \\
&+\mathcal{C}^2_{\alpha\beta\gamma}A^2(P,\vec{b}_i,x_i)+\mathcal{C}^3_{\alpha\beta\gamma}A^3(P,\vec{b}_i,x_i) \\
&+\mathcal{C}^4_{\alpha\beta\gamma}A^4(P,\vec{b}_i)+\cdots).
\end{aligned} \tag{4}$$

We are concerned with terms which lead by power counting in large momenta. Under a Lorentz transformation along the momentum axes the variables b are invariant. We therefore keep leading wave functions whether or not a power of b occurs. Each power of $b_T \to b_x \pm i b_y$ can be further decomposed into combinations of quark OAM. (At this point, ASD would reject powers of b and OAM$\neq 0$.) The first few operators $\mathcal{C}$ are given by

$$\begin{aligned}
\mathcal{C}^1_{\alpha\beta\gamma} &= (\not{P}C)_{\alpha\beta}(\gamma_5 N)_\gamma \\
\mathcal{C}^2_{\alpha\beta\gamma} &= (\not{P}\gamma_5 C)_{\alpha\beta}N_\gamma \\
\mathcal{C}^3_{\alpha\beta\gamma} &= -(\sigma_{\mu\nu}P^\nu C)_{\alpha\beta}(\gamma^\mu\gamma_5 N)_\gamma \\
\mathcal{C}^4_{\alpha\beta\gamma} &= i(\not{P}\gamma_5 C)_{\alpha\beta}(\not{b}_1 N)_\gamma.
\end{aligned} \tag{5}$$

Here C is the charge-conjugation matrix. Note the operator $\mathcal{C}^4$ which depends on b_1. Here b_i are four vectors with transverse components equal to $\vec{b}_i$ and all other components zero. Similar operators exist for other transverse coordinates.

The Dirac and Pauli form factors can now be obtained from Eq. (3):

$$F_{1,2}(Q)=\int(d\vec{k}_T)(d\vec{k}'_T)(dx)(dx')\times\sum_{s,s'}\overline{Y}_{\alpha'\beta'\gamma'}(P',\vec{k}'_{Ti},x'_i,s')T_\mu^{(1,2)}(P,s,P',s')\times\Gamma^\mu_{\alpha'\beta'\gamma'\alpha\beta\gamma}(q,\vec{k}_{Ti},\vec{k}'_{Ti},x_i,x'_i)\times Y_{\alpha\beta\gamma}(P,\vec{k}_{Ti},x_i,s)$$

$$=\int(d\vec{b})(d\vec{b}')(dx)(dx')\psi^*(b_i,x_i)\tilde{H}_{1,\,2}\times(\vec{b}_i,\vec{b}'_i x_i,x'_i,Q)\psi(b_i,x'_i). \tag{6}$$

Here $T_\mu^{(1,\,2)}$ are tensors that project out $F_{1,\,2}$, and $\tilde{H}_{1,\,2}(\vec{b}_i,\vec{b}'_i,x_i,x'_i,Q)$ are the projected Fourier transforms of the hard scattering kernel, $H(x_i,x'_i,\vec{k}_{Ti},\vec{k}'_{Ti};Q)$. Each projection selects linear combinations of the wave functions A^i defined in Eq. (4), which from the context is obvious, so that subscripts (1,2) in $\tilde{H}$ can be suppressed. The impulse approximation has been used to set the light-cone "time" to zero. There are no *a priori* assumptions in Eq. (6) about short-distance. If Sudakov effects are used consistently, then wave functions in Eq. (6) concentrate important integration regions into one which is perturbatively calculable, *without assuming zero-distance as a starting point.* We do not go to the further extreme of Ref. [23] towards asserting that the ultimate output including Sudakov effects is the ASD model. We find this unjustified, especially since the absence of Sudakov corrections is one of the characteristic shortcomings of the ASD model.[2] To bring the two methods into concordance, special assumptions about the wave functions would need to be made. A detailed calculation of F_1 at leading order in perturbation series is given in Ref. [24].

The form factor F_2 gets contributions from the wave functions, such as $\mathcal{C}^4_{\alpha\beta\gamma}X$, which scale as b in the small b limit. Furthermore the transverse momenta in the quark propagators also contributes to the kernel $\tilde{H}_2$ at leading order. In this paper we are interested in obtaining the scaling behavior of F_2 and not in the detailed leading order calculation. We focus on the b integrations in the limit of zero quark masses. The scaling behavior can be determined by considering the leading order hard scattering kernel whose detailed form is given in Ref. [24]. Since we are working in the impact parameter space, we need to take the Fourier transform of this kernel. In this case the transverse momentum factors such as $k_{T1,i}$, which occur in the Dirac propagators, turn into derivatives $i\partial/\partial b_{1,i}$. The remaining Fourier transform has a form similar to what is obtained for the form factor F_1. The dependence on the impact parameter in this kernel arises through the bessel functions of the form $K_0(\sqrt{x_1x'_1}Q\tilde{b}_{12})$, where $\tilde{b}_{12}=|\vec{b}_1-\vec{b}_2|$. Taking the derivative of this kernel with respect to $b_{1,i}$ gives factors of the form

$$\frac{(b_{1,i}-b_{2,i})}{\tilde{b}_{12}}Q\sqrt{x_1x'_1}K'_0(\sqrt{x_1x'_1Q^2\tilde{b}^2_{12}}).$$

Besides this transverse separation dependence, an additional power of $\vec{b}$ arises from the operator $\mathcal{C}^4$ in the wave function. Hence the scaling of $F_2(Q^2)$, and the form factor ratio, hinges on scaling of the effective transverse separation b at large Q. Counting powers of Q, including one for the prefactor $q^\mu\sigma_{\mu\nu}$, we find that if $b\sim 1/Q$ at large Q, then $F_2/F_1\sim 1/Q^2$ in this limit. This result has recently been confirmed in Ref. [14] which adopts the ASD formalism from the start. Alternatively if $b\sim$const then $QF_2/F_1\sim$const and $F_2\sim 1/Q^5$ in the same limit. Such scaling was predicted in Ref. [4] and is also seen in a relativistic quark model calculation [6]. In the rest of this paper we study in detail how the dominant region of b scales with Q in perturbative QCD.

III. ILLUSTRATIVE CALCULATIONS

With little being known about nonperturbative quark wave functions, there is great latitude for exploring the effects of OAM. We impose general principles and assume that the large b components of wave functions are strongly damped, and the short distance components are smooth, unless otherwise specified. Models which can falsify common prejudices are particularly instructive. Such models can illustrate analytically some general limit-interchange difficulties upsetting Eq. (2). Numerical calculations with more detailed models are presented in Sec. IV.

A. Role of the transverse coordinate

Return to the general expression Eq. (6). To leading order and neglecting transverse momentum in Fermion denominators the hard scattering depends on differences $\vec{k}_T-\vec{k}'_T$ before and after the hard collision. The neglect of Fermion denominators, which is the model of Li and Sterman [23], suffices for our purposes of establishing the general absence of HHC for large Q. Restoring such terms can give subleading dependence on Q which might be important for numerical comparisons with data.

The importance of the variable $\vec{k}_T-\vec{k}'_T$ seems rather general, because the *sum* of the transverse momenta are conjugate to the overall spatial location of the hard scattering, which by translational invariance drops out.[3] In a process with a single hard exchange (a pion or meson form factor) the transverse integrals take the form of convolution:

[2]While Sudakov corrections are mentioned in Ref. [7], they do not come from the operator-product approach of the model itself, but are invoked from outside the model.

[3]From translational invariance, the hard kernel must depend on the difference of space coordinates. The transverse separation b is only one such difference, and the one of interest here. Other differences such as the longitudinal ones also occur.

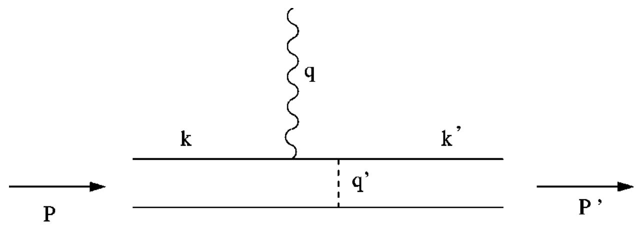

FIG. 2. Definitions of momenta cited for the one gluon exchange kernel, converted to impact parameter coordinates in the text.

$$F(Q)=\int d^2k_T dx dx' \overline{\psi}(\vec{k}_T',x')f(k_T,x,x'),$$

$$f(k_T,x,x')=\int d^2k_T' H(k_T-\vec{k}_T',x,x';Q)\psi(\vec{k}_T,x)$$

$$=\int \frac{d^2b}{(2\pi)^2}e^{i\vec{k}_T\cdot\vec{b}}\tilde{H}(b,x,x';Q)\tilde{\psi}(b,x). \tag{7}$$

Consequently the expressions are diagonal in b, with

$$F(Q)=\int \frac{d^2b}{(2\pi)^2}dxdx'\tilde{\psi}^*(b,x')\tilde{H}(b;x,x',Q)\tilde{\psi}(b,x)$$

$$=\sum_{m_1m_2m_3}\int \frac{d^2b}{(2\pi)^2}dxdx'd\phi$$

$$\times e^{-i(m_1-m_2-m_3)\phi}\tilde{\psi}^*_{m_1}(b,x')$$

$$\times\tilde{H}_{m_2}(b;x,x',Q)\tilde{\psi}_{m_3}(b,x). \tag{8}$$

All contributions to OAM are explicit at this stage. Expansion of the hard scattering H_m was introduced because the kernel can also carry angular momentum and be anisotropic, via the direction of the hard momentum $\vec{Q}$. The selection rules conserving angular momentum will come from the ϕ integrals.

B. Gluon exchange kernel

Consider the simplest one-gluon approximation to the pion kernel (Fig. 2), $H=4\pi C_F\alpha_s/q'^2$, where q' is the gluon momentum. This is written out as

$$H(\vec{k}-\vec{k}_T';x,x',Q)=4\pi C_F\alpha_s\frac{1}{xx'Q^2+(\vec{k}_T-\vec{k}_T')^2},$$

$$\tilde{H}(b;x,x',Q)=8\pi^2C_F\alpha_sK_0(\sqrt{xx'Q^2b^2}). \tag{9}$$

When integrated with nonsingular functions of b and x, the dominant region is not determined by Eq. (2), but instead the Bessel function restricts to

$$\sqrt{xx'}Qb<1. \tag{10}$$

Clearly Eq. (10) is more accurate than Eq. (2), because the partons entering the reaction carry xP, rather than P and can only be scattered through $xx'Q^2$ momentum transfer-squared. This has been recognized for a long time, and common wisdom ascribes an average value of $x\sim 1/3$ for valence quarks. It is commonly accepted that consistent appearance of $xx'Q^2$ postpones any onset of short-distance dominance to higher values of Q^2 compared to the naive implications of Eq. (2).

A physics question. We pause to question how the "uncertainty principle" of Eq. (2), $b\sim 1/Q$, could have misled the field. First, it was not the uncertainty principle, but an assertion of dominant regions. Somehow x, from the longitudinal coordinate, appeared in a relation between transverse things. We seek a physical explanation.

Suppose we stand at impact parameter b beside a fast moving quark with mass m, energy $E_0=\gamma m$. Relativity predicts a pulse of fields with a time scale $\Delta t\sim b/\gamma$.[4] The Fourier modes of the fields coming with the quark must be distributed over time scale Δt. Their individual energies $E_q=xE_0$ are measured in units of $1/\Delta t$, meaning that the fields are functions of

$$E_q\Delta t=xE_0b/\gamma=xbm.$$

If there is no longitudinal momentum transfer, the case $x'=x$, the fields depend on x^2b^2, just as in Eq. (9). While this way of thinking is used in electrodynamics, for "equivalent photons," it is naturally the same result in PQCD, because the gluon propagator coincides with the electromagnetic propagator to leading order.

The longitudinal fraction is coupled to the transverse structure because the time scale depends on x. The time-scale (x) dependence is of course set by the *nonperturbative* part of the problem, the x dependence of wave functions. Since we cannot in principle determine the x dependence in perturbation theory, then no general rules can be made to predict everything about the b dependence, either.

Time scale smearing: Interplay of x and b

Exploring the integrals, the $b^{|m|}\sim 1/Q^{|m|}$ counting will still occur, from dimensional analysis, after doing the b integrations in each x,x' integration bin. It looks like Eq. (2) and its conclusions may win after all. Surprisingly, there is no simple rule *after* the x,x' integrations are done. The scaling indicated by dimensional analysis is "erased" by the x integrals under broad conditions. The physical origin is that the final scale depends on the *dynamical time scale distribution* in the wave functions in x that we are *not* priviledged to predict. The x distribution is set by the proton itself in the quiet of vacuum over infinite time. The "time-scale smearing" destroys naive use of the uncertainty principle, and cor-

[4]This time scale is also a Lorentz-contracted pancake longitudinal distance scale. Our use of "time scale" is consistent with the impulse approximation.

responding counting of powers. However time-scale smearing does not destroy the use of PQCD, which always contains integration over x.

C. Model definitions

It suffices to use the kernel Eq. (9) to show the effects. Since our focus is the orbital angular momentum, we consider the integrals with factors of $b^{|m|}$ explicit. The selection rules from the angular integrals $e^{-i(m_1-m_2-m_3)\phi}$ are obvious.

We explore *Ansätze* for

$$\tilde{\Phi}(b,x,x')=b^A(xx')^{r_1}((1-x)(1-x'))^{r_2}\tilde{\Psi}(b,x,x').$$

The factor of b^A is the phase-space to find an assumed number of quarks close together from naive quark-counting. The remaining factor $\tilde{\Psi}(b,x,x')=\xi_i(b,x)\xi_f(b,x')$ represents the remaining product of the wave functions. Although wave functions symmetric in $x\to 1-x$ are commonly assumed, there is no motivation for this with $m\neq 0$. Exponents r_1,r_2 will be called the "dominant power of x" when they rule singularities.

In *model 1* we use a factored Gaussian function to represent the wave functions cutting off large b,

$$\xi=\exp(-b^2/(2a^2)) \quad model \quad 1.$$

In *model 2* we use a Gaussian in logarithms,

$$\xi=\exp(-\log^2[b^2x(1-x)/a^2]) \quad model \quad 2,$$

to retain some coupling between x and b. In *model 3* we use a Gaussian coupling x and b:

$$\xi=e^{-b^2\sqrt{x(1-x)}/a^2} \quad model \quad 3,$$

where ξ represents the initial or the final soft wave function. These model are chosen for their analytic simplicity to serve as examples.

D. Mellin method

We study the large-Q asymptotics by calculating a Mellin transform $\tilde{F}(N)$ conjugate to Q^N:

$$\tilde{F}(N)=\int_0^\infty \frac{dQ}{Q}Q^N F(Q)$$
$$=\int_0^\infty db\, b^{|m|+A+1}\int_0^1 dx\int dx'$$
$$\times\int_0^\infty \frac{dQ}{Q}Q^N K_0(\sqrt{xx'b^2}Q)\tilde{\Phi}(b,x,x'), \quad (11)$$

where, since the kernel $\tilde{H}$ has been assumed to carry zero orbital angular momentum, $m=|m_1|+|m_3|$. The Q integral is carried out easily,

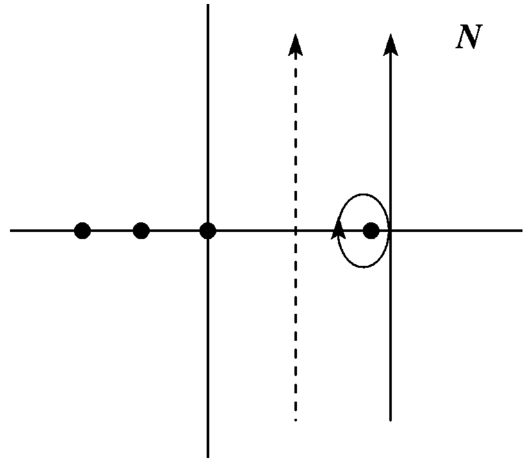

FIG. 3. Inverting the Mellin transform. A contour (dashed) in an analytic strip of the complex N plane is deformed past an isolated singularity (dots) to give the residue plus the contribution of the new contour.

$$\int \frac{dQ}{Q}Q^N K_0(\sqrt{xx'}bQ)=2^{-2+N}(xx'b^2)^{-N/2}\left[\Gamma\left(\frac{N}{2}\right)\right]^2. \quad (12)$$

Note that b^{-N} emerges just as expected from dimensional analysis. If we stopped here, then b^m would be suppressed by Q^{-m}.

The other integrals remain as

$$\tilde{F}(N)=2^{-2+N}\left[\Gamma\left(\frac{N}{2}\right)\right]^2\tilde{\mathcal{G}}(N), \quad (13)$$

and depend on the model.

Some background on Mellin transformations

To keep our presentation complete, we review some background on Mellin transformations [27].

Suppose that a function $f(Q)$ falls like Q^{-A}. The integral of $Q^{N-1}f(Q)$ will converge as $Q\to\infty$ for $N<A$. As N is increased from below, a singularity appears at $N=A$. This singularity is a sign that the function falls like Q^{-A}.

Since Q^N is an analytic function of N, the Mellin transform is then analytic in complex $N=N_x+iN_y$ for a strip of real $N_x<A$. Its analytic continuation is

$$\tilde{f}(N)=\int \frac{dQ}{Q}Q^{N_x+iN_y}f(Q),$$
$$=\int \frac{dQ}{Q}e^{iN_y\log(Q)}Q^{N_x}f(Q). \quad (14)$$

We keep N_x fixed in the strip of convergence, and recognize this as the Fourier transform of $Q^{N_x}f(Q)$ with N_y conjugate to $\log(Q)$. The Fourier transform is inverted by

$$Q^{N_x}f(Q)=\frac{1}{2\pi}\int dN_y e^{-iN_y\log(Q)}\tilde{f}(N),$$

$$f(Q)=\frac{1}{2\pi i}\int dNQ^{-N}\tilde{f}(N). \tag{15}$$

This is the inverse Mellin transform. The contour in the complex N plane runs in the purely imaginary direction in the strip of real-N where $\tilde{f}(N)$ converges (Fig. 3).

The N integral can be done by deforming the contour about singularities one-by-one. At each singularity Q^{-N} is analytic, and is evaluated by the residue theorem. The result generates an asymptotic series in Q. If one pulls the contour to the *left*, then a series in Q^{-N} for increasingly negative N is obtained: a series such as $1,Q,Q^2$ . . . . Such an expansion is appropriate for small Q. For example, Eq. (12) has double poles at $N=0,-1,-2$ . . . . The double pole at $N=0$ needs the N^1 term of $Q^N\sim 1-N\log(Q)-N^2\log^2/2$, and will generate a residue going like $\log(Q)$. Keeping the first 2 singularities, the series is

$$K_0(xbQ)\sim\frac{1}{2}\left[-2\gamma_E-\log\left(\frac{b^2Q^2x^2}{4}\right)\right]$$
$$+b^2Q^2x^2\left(\frac{1}{4}-\frac{\gamma_E}{4}+\frac{\log(2)}{4}+\frac{\log(b^2Q^2x^2)}{8}\right)$$
$$+O((bxQ)^3\log(bxQ)) \tag{16}$$

where $\gamma_E\sim 0.5772$ is the Euler gamma number. The first term is well known, and the method can be extended to get an unlimited number of terms.

E. HHC and non-HHC cases

We use notation $f(Q)\leqslant Q^{-P}$ to indicate dependence falling *at least as fast as* Q^{-P}. Returning to Eq. (13), dependence at *large* Q now comes from singularities of $\tilde{\Phi}(N)$ to the *right* of the convergence strip (Fig. 3).

Model 1. We need

$$\int_0^\infty db\int_0^1 dx(b^2x)^{-N/2}b^{|m|+A+1}x^{r_1}(1-x)^{r_2}e^{-b^2/(2\tilde{a}^2)}$$
$$=\frac{2^{(A+|m|-N)/2}\tilde{a}^{A+|m|-N+2}\Gamma\left(\frac{2+A+|m|-N}{2}\right)\Gamma\left(-\frac{N}{2}+r_1+1\right)\Gamma(1+r_2)}{\Gamma\left(2-\frac{N}{2}+r_1+r_2\right)}. \tag{17}$$

The initial and final state wave functions combine to give the factor $e^{b^2/2\tilde{a}^2}$. An x' integration remains to be done to get $\tilde{\mathcal{G}}$. The integrand is products of powers and analytic. It then suffices to examine the singularities already visible.

The singularities of $\Gamma((2+|m|+A-N)/2)$ are simple poles at

$$N=2+|m|+A+2K,\quad K=0,1,2\ \ldots.$$

These are exactly the singularities creating the naive power counting of HHC. The existence of these singularities implies

$$F(Q)\leqslant Q^{-|m|-A-2}\quad (\text{HHC region only}), \tag{18}$$

from these singularities alone, and barring a zero factor canceling the poles. As expected $b^{|m|}\to Q^{-|m|}$: these are the contributions for which the HHC results are reproduced.

The singularities of $\Gamma(2-N+r_1)$ are simple poles at

$$N=2+r_1+K',\quad K'=0,1,2\ \ldots,$$

yielding a power behavior

$$F(Q)\leqslant Q^{-2-r_1}\quad (\text{dominant power region}) \tag{19}$$

regardless of the value of $|m|$. If such poles dominate, there is no power suppression of OAM $|m|\neq 0$.

If the singularities of $N=2+|m|+A+2K$ and $N=2+r+K'$ coincide, there is a double pole, modifying the result above by a power of $\log(Q)$. If the zero of $1/\Gamma(4-N/2+r_1+r_2)$ meets a pole, then that pole has no residue.

There is considerable structure in these calculations not anticipated in naive power counting.

Model 2. It is inconvenient to combine the exponents of two wave functions, and we can make a point by integrating just one. Although this is not general it indicates the kind of behavior we might expect from a nonfactorizable *Ansatz*. Alternatively we can expand one of the wave functions in powers of b, keeping only the lowest order term. This would be reasonable as long as small transverse separations dominate and hence our results can test the validity of the ASD model in all generality. We need

$$\int_0^\infty db \int_0^1 dx (b^2x)^{-N/2} b^{|m|+A+1} x^{r_1}(1-x)^{r_2} \times e^{-k\log^2(b^2x(1-x)/a^2)}$$
$$= \frac{\sqrt{\pi}}{D} a^{2+A+|m|-N} e^{(2+A+|m|-N)^2/16k} \times \Gamma\left(\frac{-A}{2} - \frac{|m|}{2} + r_1\right)\Gamma\left(\frac{-A-|m|+N+2r_2}{2}\right),$$
$$D = 2\sqrt{k}\Gamma\left(-A-|m|+\frac{N}{2}+r_1+r_2+2\right). \tag{20}$$

There are no singularities for $N>0$, and this model is exponentially damped as $Q\to\infty$. In the argument of the gamma function the effects of $|m|$ and $2r_2$ are equal and opposite: we can have as many powers of $b^{|m|}$ wanted, and produce no effects on the power N provided $|m|-2r_2=$const.

Let us explore this phenomenon qualitatively. Model 2 has $\tilde{\Phi}(b)$ falling faster than any power of b as $b\to\infty$. However as $b\to 0$ the model is also smaller than any power of b. There is no phase space for the quarks to find each other: hence the exponentially damped behavior. While this is not a realistic model feature, it is interesting nonetheless.

The cancellation of $|m|$ and r_2 powers appears to originate in that $b^2x(1-x)x'(1-x')/a^2<1$ sets the scale of b. Large b and unsuppressed OAM would be possible at small x or small $1-x$, except there is no large b allowed anyway. The possibility of large b dominance may be a realistic feature, and deserves further study in model 3.

Model 3. This is the most interesting model. We separate the product of two wave functions by expanding one in powers of b. Integrating the first term, we find

$$\tilde{\mathcal{G}} = \int_0^\infty db \int_0^1 dx (b^2x)^{-N/2} b^{|m|+A+1} x^{r_1}(1-x)^{r_2} \times e^{-b^2\sqrt{x(1-x)}/a^2}$$
$$= \frac{1}{D} a^{2+A-N+|m|}\Gamma\left(1+\frac{A}{2}+\frac{|m|}{2}-\frac{N}{2}\right) \times \Gamma\left(\frac{1}{2}-\frac{A}{4}-\frac{|m|}{4}-\frac{N}{4}+r_1\right)\Gamma\left(\frac{1}{2}-\frac{A}{4}-\frac{|m|}{4}+\frac{N}{4}+r_2\right),$$
$$D = 2\Gamma\left(1-\frac{A}{2}-\frac{|m|}{2}+r_1+r_2\right). \tag{21}$$

This function has its singularities at $N=2+A+2K+|m|$, $K>0$ integer which again are reminiscent of the HHC poles. It also has singularities at

$$N = 2-A+4K-|m|+4r_1, \quad K=0,1,2\ldots.$$

Model 3 is another example where a positive power of $|m|>0$ may be cancelled by a positive power $4r_1>0$. In this model there is *no a priori suppression* of OAM and in some cases OAM could dominate

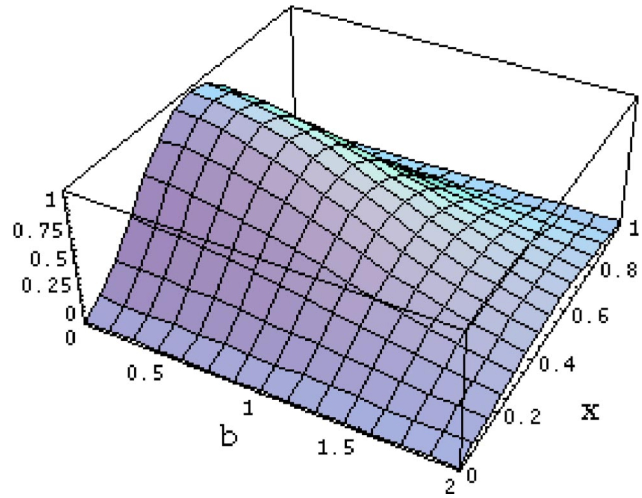

FIG. 4. The wave function of model 3, cited in the text. Despite the width in the b variable going like $1/\sqrt{x(1-x)}$, the wave function is very moderate and conventional in appearance.

Let us understand this qualitatively. The appearance of $\sqrt{x(x-1)}$ in the exponent allows b to take its greatest range at small x or small $1-x$. We have argued this is generic due to time-scale smearing. Instead of the wave function blowing up, however, the range of b is finite, because the prefactors of $x^{r_1}(1-x)^{r_2}$ regulate the phase space. The resulting wave function is quite conventional in appearance, as illustrated in Fig. 4.

Breakdown of series expansions. There appears to be a paradox. If we made a Taylor series for the wave function in powers of b, then term by term powers of $b^{2J}x^J(1-x)^J$ would be suppressed for $J>0$. However when we examine the exponential in model 3, an OAM factor of b^2 is nearly equivalent to a factor of $1/\sqrt{x(1-x)}$, so that powers of $b^{|m|}$ are not necessarily suppressed. This is why x^{r_1} can cancel $b^{|m|}$. The paradox is that the sum acts much different than the separate terms. A simple resolution is that most of the integral is coming from the region where the exponent is of order 1: and a series expansion is not valid. The assumptions of HHC, in contrast, require that a series expansion about $b\sim 0$ be the dominant contribution. Model 3 serves as a counterexample to that expectation.

Rule of thumb. We have extracted an improved HHC-rule of thumb: *the asymptotic Q^2 dependence is determined by the minimum of powers m,r in the integrands.* The rule has been deliberately simplified to powers, suppressing logarithmic factors readily calculated, by further evaluating the Mellin inversion. These logarithms are obtained from the Mellin series and are a separate phenomenon from the asymptotic-freedom logarithms of PQCD. The origin of the variety of predictions is time-scale smearing: due to it, the suppression of OAM expected by naive counting is transferred to smearing inside the longitudinal wave functions, and helicity flips can readily compete with helicity nonflip. There is no power suppression of OAM for $m\neq 0$ in general.

Effects of the large b regions. The singularities to the *left* of the convergence strip determine a series for *small Q* $\to 0$. These singularities are traced to the Gaussian model, Sudakov b cutoff, or the large b cutoff. We are not concerned with these singularities, showing that the large b cutoff drops out of the power laws stated above. However the numerical normalization of integrals depends on the cutoff method. It is a separate question needing separate analysis to find the numerical fit of models to data. Our numerical work is presented momentarily.

On this reasoning there is always a part of the calculation which is "strictly perturbative." Calculations of the large b regions may apply to data only with limited reliability. Regardless of this fact, the small b regions known to be calculable behave as we claim, and are the only regions under conceptual dispute.

F. Discussion

To our knowledge HHC has never previously been observed to fail for large Q^2 form factors. Yet previous work has made related observations that the powers of x can determine the Q^2 dependence. Feynman's mechanism [28] concentrated entirely on the region $x\to 1$ and ignored b. Feynman's work predated PQCD, and there are certainly regions in the b integrals of that mechanism that would not be calculable, so the model is moot. Soper's 1977 paper [19] has a clear statement that the *minimum* power of x or k_T rules the results. The context was the Drell-Yan-West [29] relation, which connects the $x\to 1$ behavior of inelastic structure functions with the Q^2 dependence of the form factor. Soper shows that one cannot prove the Drell-Yan-West relation deductively, but one can get an inequality and force the relation by choosing wave functions.

The literature is clouded here by treatments *assuming* that the two regions of $x\to 1$ and $b\to 0$ must give the same scaling, and forcing a result by circular logic. As we have just shown, the regions are different and no general rule can be made. Underlying this is the fact of perturbative calculability of short-transverse distance not being on the same footing as perturbative *models* of the $x\to 1$ dependence. For instance Brodsky and Lepage [12,13] discuss the end point $x\to 1$ contribution of the proton form factor F_1. The authors argue that for the limit $x_1\to 1$, PQCD implies an $m=0$ wave function of order α_s^2 perturbatively calculated to go like $(1-x_1)^1$. On this basis they find that a contribution to F_1 scaling like $1/Q^4$ *independent of the powers of b* inside the integrals. The Q dependence of this contribution is same as that obtained from the $b\to 1/Q$ region. Although perturbative analysis is not a valid approach to calculating wave functions, the result is consistent with ours.

Our results are due to time-scale smearing and should not be attributed to "end-point singularities" as the term is commonly used. Inside the hard scattering kernels in PQCD are combinations of inverse powers of x or $1-x$. *We are not exploiting these inverse powers* in the evaluation of the Mellin moments. Rather, it is the generic coupling of x and b integrations which makes impossible any general statements of HHC. We reiterate that the $x\to 1$ limits of the $m\neq 0$ wave functions are unknown. End point singularities (if any) can only increase the variety of outcomes.

To summarize the logic so far, for broad classes of x dependence of the wave functions, the nonrule of HHC is transformed to a rule of HHNC, hadron helicity nonconservation. This is a new asymptotic prediction, and proves that $QF_2(Q^2)/F_1(Q^2)\sim$const can be an outcome of the theory up to the highest Q. To be equally fair, different broad classes of wave functions give $Q^2F_2(Q^2)/F_1(Q^2)\sim$const. Our primary accomplishment enlarges the sphere of allowed possibilities. Measurements are still needed to determine what protons are.

IV. NUMERICAL STUDIES, SUBASYMPTOTIC

The scaling rule $QF_2(Q^2)/F_1(Q^2)\sim$const describes experiments at laboratory Q values very far from asymptotic. We explored this region numerically.

A. Studies with the pion

The pion provides a simple test system. We test the dominant integration regions by inserting factors of b^m in the integrands, corresponding to m-units of OAM. The moment $\langle b(Q)\rangle_\pi$ is defined by

$$\langle b(Q)\rangle_\pi=\frac{\int dx dx' d^2b b \mathcal{F}_\pi(Q,x,x',b)}{F_\pi(Q^2)}. \tag{22}$$

The pion form factor $F_\pi(Q^2)$ is given by

$$F_\pi(Q^2)=\int dx dx' d^2b \mathcal{F}_\pi(Q,x,x',b) \tag{23}$$

with the kernel $\mathcal{F}_\pi(Q,x,x',b)$ defined by

$$\mathcal{F}_\pi(Q,x,x',b)=\int dx dx' db b \phi(x')\alpha_s(\mu)e^{-S(x,x',b,Q)} \times K_0(\sqrt{xx'}bQ)\phi(x)\tilde{\Psi}(b). \tag{24}$$

Here $S(x,x',b,Q)$ is the Sudakov form factor. A model soft wave function $\tilde{\Psi}(b)=\exp(-b^2/2a^2)$ is also included, where the parameter $a=1/\Lambda_{QCD}$. Frankfurt, Miller and Strikman in an earlier study [30] found that the b^2 moment falls faster than $1/Q^2$. Their calculation differs in some respects, for example in emphasis on the CZ form and not including a soft wave function. The soft wave function of course appears in the Li-Sterman formalism [23,24], but is replaced by a distribution amplitude when emphasis on Sudakov effects is sought.

In Fig. 5 we plot the moment $\langle b(Q)\rangle_\pi$ using a wave function of the form

$$\phi(x)\sim[x(1-x)]^\delta.$$

This ansatz is deliberately arbitrary, and for our purposes many example wave functions will serve. We study values of $\delta=0.2,1$, with and without the Sudakov form factor. In the latter case the momentum scale $\mu=Q/4$ of the strong coupling is imposed, and the b integrals are cut off at b

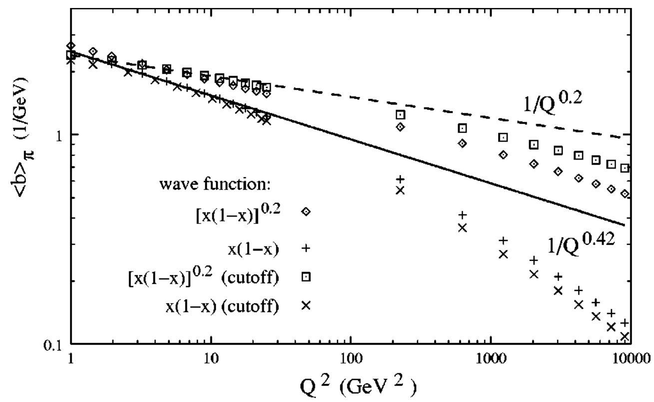

FIG. 5. The moment $\langle b(q)\rangle_\pi$ of the pion form factor kernel, as defined in the text, using the $[x(1-x)]^\delta$ form of wave function for $\delta=0.2,1$. The moment decreases with Q much slower than the $1/Q$ behavior expected in the ASD HHC model. Results are shown with and without including the Sudakov effects. The solid $(1/Q^{0.42})$ and the dashed $(1/Q^{0.2})$ line represent a simple power law fit at small Q^2 for $\delta=1$ and $\delta=0.2$, respectively.

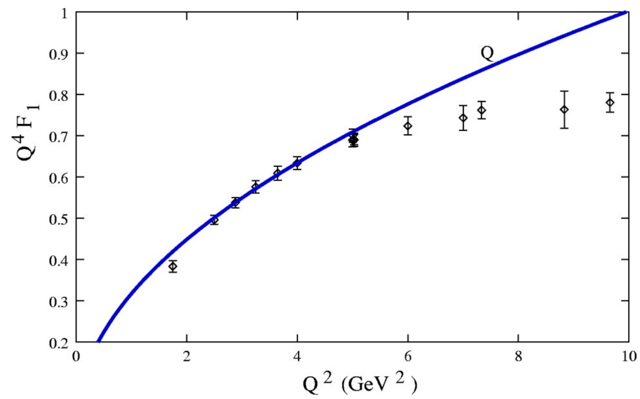

FIG. 6. The proton form factor F_1 for moderate Q^2, extracted using the Jlab data for G_E/G_M and the SLAC data for G_M. The ratio G_E/G_M is obtained from the parametrization [2] $\mu_p(G_E/G_M)=1-0.13(Q^2-0.04)$. The solid line represents $F_1 \sim 1/Q^3$.

$=1/\Lambda_{QCD}$. From Fig. 5 it is clear that for $Q^2<100\ \text{GeV}^2$ the moment falls much slower than the ASD prediction of $1/Q$, regardless of the x dependence of the wave function used [31].

B. Studies with the proton

We turn to the proton form factors. We note that the JLAB data for QF_2/F_1 is flat even below the Q-range where $F_1 \sim 1/Q^4$ begins to fit. In Fig. 6 we show a plot of Q^4F_1. The solid line in the plot corresponds to the behavior $F_1 \sim 1/Q^3$. It is clear from the plot that the scaling $F_1\sim 1/Q^4$ is seen only for $Q^2 \geqslant 6\ \text{GeV}^2$, which is larger than the momentum regime explored at JLAB so far. This is cause for concern.

With 3 parameters and a large logarithm it is possible to fit [32] F_2/F_1 to a form going like $1/Q^2$ asymptotically. The particular fit used in Ref. [32] is

$$\frac{F_2}{F_1}=\frac{\mu_A}{1+\left(\frac{Q^2}{A}\right)\log^C(1+Q^2/B)}. \tag{25}$$

Yet it is certainly possible that the scaling observed *in the ratio* is not overly sensitive to Q^2, and will continue to larger Q^2. We investigate this in greater detail numerically.

To probe the dominant integration regions, we turn to calculating b moments of the proton form factor kernel. These are multidimensional integrals of which our analysis in the previous section determines but a low-dimensional strip. We spare the reader a listing of dozens of Feynman diagrams and substantially complicated kernels listed in the literature [24]. The form factor F_1 in impact parameter coordinates can be written symbolically as [24]

$$\begin{aligned}F_1(Q^2)=&\sum_{j=1}^{2}\frac{4\pi}{27}\int_0^1(dx)(dx')\int_0^\infty b_1db_1b_2db_2\\ &\times\int_0^{2\pi}d\theta[f_N(w)]^2\tilde{H}_j(x_i,x_i',b_i,Q,t_{j1},t_{j2})\\ &\times\Psi_j(x_i,x_i',w)\exp[-S(x_i,x_i',w,Q,t_{j1},t_{j2})],\end{aligned} \tag{26}$$

with $(dx)=dx_1dx_2dx_3\delta(1-x_1-x_2-x_3)$. The variable θ is the angle between $\vec{b}_1$ and $\vec{b}_2$ and x_i and x_i' refer to the initial and final x variables. The expressions for the hard scattering $\tilde{H}_j$, the Sudakov form factor S, and function Ψ_i are given in Ref. [24]. Now defining

$$F_1(Q^2)=\int b_1db_1b_2db_2\mathcal{F}_P, \tag{27}$$

where $\mathcal{F}_P$ can be extracted from Eq. (26), we define moments $\langle b_j(Q)\rangle_p$ as follows:

$$\langle b_j(Q)\rangle_p=\frac{\int b_1db_1b_2db_2\mathcal{F}_Pb_j}{F_1(Q^2)}. \tag{28}$$

The function Ψ_i is where the linear combinations of the products of initial and final x wave functions are found. The most singular part of the kernel in the limit $x_1\to 1$ and $x_1' \to 1$ is obtained from the $\tilde{H}_1\Psi_1$ term, which is of the form

$$\tilde{H}_1\Psi_1 \sim \frac{K_0(\sqrt{(1-x_1)(1-x_1')}Qb_1)K_0(\sqrt{x_2x_2'}Qb_2)\phi(x_i)\phi(x_i')}{(1-x_1)(1-x_1')}. \tag{29}$$

Here K_0 is the modified Bessel function of order zero.

For the test we explore a wave function $\phi(x_i)$ given by

$$\phi(x_i) \sim (x_1x_2x_3)^{\delta}.$$

The numerator in Ψ_1 is then proportional to $(x_1x_2x_3x_1'x_2'x_3')^{\delta}$. The Bessel functions imply that in the limit of large Q, $(1-x_1)(1-x_1')b_1^2 \to 1/Q^2$ and $x_2x_2'b_2^2 \to 1/Q^2$. As long as $\delta \leqslant 0.5$ the time-scale smearing will dominate. For $\delta = 0.5$ we get $F_1 \to 1/Q^4$ even though a dominant region in b_1 and b_2 is independent of Q. For $\delta \leqslant 0.5$ the moments of b_1 and b_2 should also have a region independent of Q.

We check these predictions by performing the calculation using $\delta = 0.5$. We ignore the Sudakov form factor for this test, because it is a side issue. We evaluate the strong coupling at $Q/4$. In Fig. 7 we plot the moment of the transverse separations $b_1(Q)$ and $b_2(Q)$ as a function of Q. We find that for $Q^2 > 100$, the moments $\langle b_2 \rangle_p$ and $\langle b_1 \rangle_p$ fall asymptotically as $1/Q^{0.2}$ and $1/Q^{0.14}$ respectively, in agreement with our analytic expectations. If we include the Sudakov form factor the moment $\langle b_2 \rangle_p$ falls as $1/Q^{0.4}$, asymptotically. This is only a slightly stronger decay compared to the earlier case. In contrast to earlier expectations, the Sudakov form factor does not much suppress the importance of the end-point region.

The results with end point dominated COZ [33,34] x dependence are similar. In this case, as shown in Fig. 8, we find that the $\langle b_2 \rangle_p$ decays very slowly with Q for a wide range of Q. This is quite interesting, because, as shown in Fig. 9, the form factor F_1 itself does not show good $1/Q^4$ scaling in this momentum regime, if the Sudakov form factor is not included. The calculation indicates that the scaling seen in the moment, and hence the ratio QF_2/F_1, is more general than that seen in F_1. This is not totally unexpected in a ratio. At very large Q it starts to fall faster, but only as $1/Q^{0.6}$: well below the supposed $1/Q$ rule of ASD assumptions.

Let us add a remark about the neglect of transverse momentum in Fermion numerators. Formally such neglect is perfectly legitimate and justified at large Q, where the finite k_T integrals cannot add a factor *increasing* with Q. At the finite Q of laboratory data the step defines a model that can be questioned for observables hinging on OAM. For this reason we have presented separate calculations explicitly including k_T effects in order to test the contributions of OAM.

C. The ratio $QF_2(Q^2)/F_1(Q^2)$

Based on these studies we come to a prediction for the ratio $QF_2(Q^2)/F_1(Q^2)$. This prediction is based on power-counting, and the dominant integration regions probed by the moments. It is as complete a prediction as now possible, taking into account that one unit $\Delta m = 1$ of orbital angular momentum is needed for F_2 to proceed for $Q^2 \geqslant \mathrm{GeV}^2$. Up to small and model dependent corrections, of order $Q^{0.05}$, the power counting gives the scaling behavior but not the normalization of the data, via the relation

$$\langle b_2(Q) \rangle_p \sim \frac{QF_2(Q^2)}{F_1(Q^2)} \sim \text{const}. \tag{30}$$

What if this prediction fails? It can fail, according to our asymptotic studies, if the dominant power r is "large." Large Q^2 studies of F_2 not only probe quark OAM, but also they can tell us details of the dominant x power associated with $m \neq 0$.

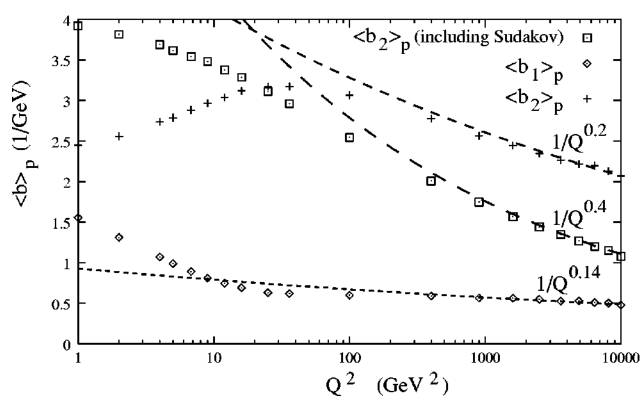

FIG. 7. The moment of the transverse separation b_1 and b_2 for proton using the wave function of the form $(x_1x_2x_3)^{0.5}$. The $\langle b_2 \rangle_p$ moment is shown both with and without including the Sudakov form factor. The small dashed ($1/Q^{0.14}$), medium dashed ($1/Q^{0.2}$) and the large dashed ($1/Q^{0.4}$) lines represent simple power law fits at large Q^2.

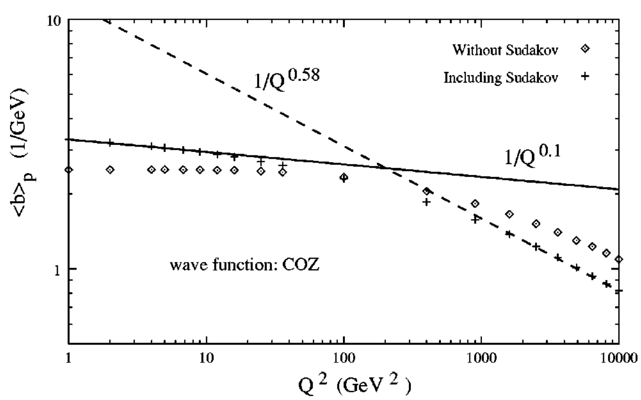

FIG. 8. The moment of the transverse separation $\langle b_2(Q) \rangle_p$ for the proton form factor kernel using the COZ wave function [33]. The solid ($1/Q^{0.1}$) and dashed ($1/Q^{0.58}$) lines represent simple power law fits at small and large Q^2, respectively. The dependence on Q is much weaker than predicted by ASD relations, and supports a flat prediction for QF_2/F_1 at JLAB momentum transfers.

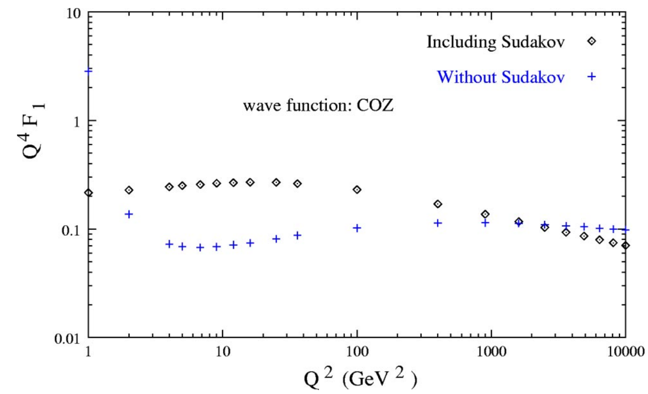

FIG. 9. The Q^2 dependence of the calculated proton form factor F_1 using the COZ wave function [33]. Results are shown both with and without including the Sudakov form factor. Details are in the text.

Positivity of G_E. It is interesting to observe that if the trend of $QF_2(Q^2)/F_1(Q^2) \sim$const continues, the value of G_E will reach zero, and cross to a negative value. One wonders [35] whether there is a physical significance and a barrier to this unusual occurence.

While usually called the ''Sachs'' form factors [36], the Appendix of a 1959 *Reviews of Modern Physics* article [37] list two form factors denoted A and B, now called G_E and G_M, which are linear combinations of Rosenbluth's:

$$G_M = F_1 + \kappa F_2,$$

$$G_E = F_1 - \kappa \tau F_2,$$

$$\tau = \frac{Q^2}{4m_p^2} \tag{31}$$

and κ is the anomalous magnetic moment. Yennie *et al.* gave [37] the alternative linear combinations simply to emphasize that the definitions of form factors was arbitrary. If there is a special significance to $G_E \to 0$ it must be due to a special meaning of G_E.

The meaning of G_E is angular momentum $\Delta J_z = 0$ spin-preservation in the Briet frame. If $G_E = 0$ the proton spin must flip in the scattering. We find nothing special about this. The amplitudes in the spacelike region cannot be limited further than general principles of analyticity, Lorentz and gauge invariance, and so on. It is perfectly consistent to arrange timelike discontinuities so that the sign change of G_E occurs without violating anything holy.

V. CONCLUDING REMARKS

Our calculations show that the (newly revised) asymptotic behavior of powers of b is achieved only at very large Q^2. Physics has many asymptotic predictions, which by construction have suppressed information needed to know when they will apply. To repeat: methods inherent in asymptotic prediction strongly tend *not* to tell you where the prediction will be valid. In QCD it was early thought the ''asymptotic'' regime coincided with the perturbative regime. This was built wrongly into dogma. We have shown here that the scales of asymptopia for form factors are vastly beyond experimental comparisons. At the same time, the perturbative regime of QCD may well extend down to laboratory values of Q^2.

In explicit calculations at finite Q^2, we find that the effects of OAM are suppressed *by even less than the revised asymptotic behavior.* The pyramid of assumptions that the *s*-wave distribution amplitudes are meaningful falls into very grave doubt. To put this more directly, we do not have a reason to use a distribution amplitude any more.

The usual approach to PQCD, in which wave functions are assumed to be unknown, predicts HHNC: *hadron-helicity nonconservation*. The experimental observation of $R(Q) = QF_2/F_1$ appears to have a dual meaning. The ratio is a very robust quantity, which remains flat even in the regime where F_1 is not clearly dominated by quark counting. The same ratio is an important asymptotic quantity, which informs us about the wave functions. The common notion that proton wave functions are ''cubic,'' namely going like $x_1x_2x_3$, is ruled out for wave functions calculating F_2. Applications of the CZ wave functions are ruled out for F_2, if the ratio R continues to be flat at large Q^2. It is important that experimental observations be extended to larger Q^2 values.

Predictions for protons are mirrored in predictions of the same type for neutrons. If the flat $R(Q)$ ratio is followed for neutrons, then the hallowed Galster fits [38] to neutron data will eventually fail [35]. Calculations for the neutron appear in Ref. [6]. We predict a flat R ratio for neutrons on the basis of isospin independence of QCD. However we need to emphasize that it is possible, in principle, for wave functions of substantially different shape to be very close in energy, due to the variational stationarity of an energy eigenstate. If there is a place for Nature to hold surprises it is in the comparison of the neutron to the proton.

The study of polarization transfer [39] in large nuclei, $\vec{e}(A,e')\vec{p}'$ should yield further information. If the existing form factors are dominated by ASD, which we do not believe but is still worth testing, then nuclear filtering will not make the distances shorter, and the ratio R should be flat with $A \gg 1$. The regime of small $A \sim 10$ has no advantages and many complications due to few-body effects, and predictions are more difficult. If the existing form factors are dominated by quark mass effects, then nuclear filtering will not make any difference and the ratio R should again be flat. If quark OAM is responsible for the flatness of R as we believe, then filtering will kill the large transverse extent of large OAM [40,41], F_2 should be depleted relative to F_1, and R will decrease with $A \gg 1$. We intend to dedicate a study to quantifying these predictions.

JLAB has made a pivotal experimental discovery which will be a permanent subject of discussion. The experimentally observed flatness of QF_2/F_1 is a signal of substantial quark orbital angular momentum in the proton. Higher momentum transfer measurements would be helpful in confirming this interpretation. The numerical size of R cannot be converted directly to a wave function, because it is only a single number. But the value of R can rule out models which omit quark OAM. Further studies at higher Q^2 may separate constituent quark models with OAM [6], which tend to have a scale (the quark mass) forcing turnover of R with increasing Q^2.

Note added. After submission of this work, we were informed of important new data [42] from Jefferson Lab on the inclusive neutron spin asymmetry A_1^n in the valence region extending up to Bjorken $x_{Bj}=0.6$ and $Q^2>\mathrm{GeV}^2$. As the paper discusses, both the values and the trend of the data as $x_{Bj}\rightarrow 1$ are inconsistent with HHC models, and appear to confirm the presence of quark OAM.

ACKNOWLEDGMENT

This work was supported by Department of Energy Grant Number DE-FG03-98ER41079.

[1] M.K. Jones *et al.*, Phys. Rev. Lett. **84**, 1398 (2000).
[2] O. Gayou *et al.*, Phys. Rev. Lett. **88**, 092301 (2002).
[3] For example, A. Salam endorsed $Q^2F_2/F_1\sim$const in 1963, when Hofstadter's data was preemininent, without quite explaining why. See *International Conference on Nucleon Structure*, 1963, edited by R. Hofstadter and L. I. Schiff (Stanford University Press, Stanford, 1964).
[4] R. Buniy, J.P. Ralston, and P. Jain, in *VII International Conference on the Intersections of Particle and Nuclear Physics*, Quebec City, 2000, edited by Z. Parseh and W. Marciano (AIP, New York, 2000), hep-ph/0206074.
[5] J.P. Ralston, R.V. Buniy, and P. Jain, Proceedings of DIS 2001, 9th International Workshop on Deep Inelastic Scattering, Bologna, 2001, hep-ph/0206063.
[6] M.R. Frank, B.K. Jennings, and G.A. Miller, Phys. Rev. C **54**, 920 (1996); G.A. Miller and M.R. Frank, *ibid.* **65**, 065205 (2002).
[7] S.J. Brodsky and G.P. Lepage, Phys. Rev. D **24**, 2848 (1981).
[8] T. Gousset, B. Pire, and J.P. Ralston, Phys. Rev. D **53**, 1202 (1996).
[9] P. Jain and J.P. Ralston, in Future Directions in Particle and Nuclear Physics at Multi-GeV Hadron Beam Facilities, Proceedings of the Workshop held at Brookhaven National Laboratory, 1993, edited by D. Geesaman, hep-ph/9305250.
[10] X.D. Ji, J. Tang, and P. Hoodbhoy, Phys. Rev. Lett. **76**, 740 (1996).
[11] S.J. Brodsky and G.R. Farrar, Phys. Rev. D **11**, 1309 (1975); V.A. Matveev, R.M. Muradian, and A.N. Tavkhelidze, Lett. Nuovo Cimento Soc. Ital. Fis. **7**, 719 (1973).
[12] G.P. Lepage and S.J. Brodsky, Phys. Rev. D **22**, 2157 (1980).
[13] S. J. Brodsky and G. P. Lepage, *Perturbative Quantum Chromodynamics*, edited by A.H. Mueller (World Scientific, Singapore, 1989).
[14] A.V. Belitsky, X. Ji, and F. Yuan, Phys. Rev. Lett. **91**, 092003 (2003).
[15] C.R. Ji and A.F. Sill, Phys. Rev. D **34**, 3350 (1986).
[16] N. Isgur and C. Llewelyn-Smith, Phys. Rev. Lett. **52**, 1080 (1984).
[17] A.V. Radyushkin, Phys. Rev. D **58**, 114008 (1998).
[18] A. Wijesoriya *et al.*, Phys. Rev. Lett. **86**, 2975 (2001); R. Gilman, J. Phys. G **28**, R37 (2002).
[19] D.E. Soper, Phys. Rev. D **15**, 1141 (1977).
[20] M. Diehl, Th. Feldmann, R. Jacob, and P. Kroll, Nucl. Phys. **B596**, 33 (2001).
[21] J. Botts and G. Sterman, Nucl. Phys. **B325**, 62 (1989).
[22] J.P. Ralston and B. Pire, Phys. Rev. Lett. **65**, 2343 (1990).
[23] H.-N. Li and G. Sterman, Nucl. Phys. **B381**, 129 (1992).
[24] H.-N. Li, Phys. Rev. D **48**, 4243 (1993).
[25] J. Bolz, R. Jakob, P. Kroll, M. Bergmann, and N.G. Stefanis, Z. Phys. C **66**, 267 (1995).
[26] B. Kundu, H.-N. Li, J. Samuelsson, and P. Jain, Eur. Phys. J. C **8**, 637 (1999).
[27] C. M. Bender and S. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (Springer-Verlag, Berlin, 1999).
[28] R. P. Feynman, *Photon-Hadron Interactions* (Addison-Wesley, Redwood City, CA, 1970).
[29] S.D. Drell and T.M. Yan, Phys. Rev. Lett. **24**, 181 (1970); G.W. West, *ibid.* **24**, 1206 (1970).
[30] L. Frankfurt, G.A. Miller, and M. Strikman, Nucl. Phys. **A555**, 752 (1993).
[31] C. Vogt (private communication) was kind enough to check and verify our numerical work on the pion.
[32] S.J. Brodsky, in *Exclusive Reactions at High Momentum Transfer*, edited by C. Carlson, P. Stoler, and A. Radyushkin (Jefferson Lab, Newport News, VA, 2001), hep-ph/0208158.
[33] V.L. Chernyak, A.A. Ogloblin, I.R. Zhitnitsky, Sov. J. Nucl. Phys. **48**, 536 (1988) [Yad. Fiz. **48**, 841 (1988)].
[34] V.L. Chernyak and A.R. Zitnitsky, Phys. Rep. **112**, 173 (1984); Nucl. Phys. **B246**, 52 (1984).
[35] We thank Charles Perdrisat for discussions on this point.
[36] F.J. Ernst, R.G. Sachs, and K. Wali, Phys. Rev. **119**, 1105 (1960); **126**, 2256 (1962).
[37] D. Yennie, M. Ravenhall, and M. Levy, Rev. Mod. Phys. **29**, 144 (1957).
[38] S. Galster, H. Klein, J. Moritz, K.H. Schmidt, D. Wegener, and J. Bleckwenn, Nucl. Phys. **B32**, 221 (1971).
[39] S. Malov *et al.*, Phys. Rev. C **62**, 057302 (2001).
[40] P. Jain, B. Pire, and J.P. Ralston, Phys. Rep. **271**, 67 (1996).
[41] J.P. Ralston and B. Pire, in *Proceedings of Second Workshop on Hadronic Physics with Electrons Beyond 10 GeV*, Dourdan, France, 1990 [Nucl. Phys. **A532**, 155c (1991)].
[42] X. Zheng *et al.*, Phys. Rev. Lett. **92**, 012004 (2004).