



# On the mass difference between proton and neutron

J. Gasser<sup>a</sup>, H. Leutwyler<sup>a</sup>, A. Rusetsky<sup>b,c,\*</sup>



<sup>a</sup> Albert Einstein Center for Fundamental Physics, Institute for Theoretical Physics, University of Bern, Sidlerstrasse 5, 3012 Bern, Switzerland  
<sup>b</sup> Helmholtz-Institut für Strahlen- und Kernphysik (Theorie) and Bethe Center for Theoretical Physics, Universität Bonn, Nussallee 14-16, D-53115 Bonn, Germany  
<sup>c</sup> Tbilisi State University, 0186 Tbilisi, Georgia

## ARTICLE INFO

### Article history:

Received 3 December 2020  
 Received in revised form 11 January 2021  
 Accepted 16 January 2021  
 Available online 21 January 2021  
 Editor: G.F. Giudice

### Keywords:

Electromagnetic mass differences  
 Dispersion relations  
 Regge behaviour  
 Structure functions  
 Protons and neutrons  
 Cottingham formula

## ABSTRACT

The Cottingham formula expresses the electromagnetic part of the mass of a particle in terms of the virtual Compton scattering amplitude. At large photon momenta, this amplitude is dominated by short distance singularities associated with operators of spin 0 and spin 2. In the difference between proton and neutron, chiral symmetry suppresses the spin 0 term. Although the angular integration removes the spin 2 singularities altogether, the various pieces occurring in the standard decomposition of the Cottingham formula do pick up such contributions. These approach asymptotics extremely slowly because the relevant Wilson coefficients only fall off logarithmically. We rewrite the formula in such a way that the leading spin 2 contributions are avoided ab initio. Using a sum rule that follows from Reggeon dominance, the numerical evaluation of the e.m. part of the mass difference between proton and neutron yields  $m_{\text{QED}} = 0.58 \pm 0.16$  MeV. The result indicates that the inelastic contributions are small compared to the elastic ones.

© 2021 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>). Funded by SCOAP<sup>3</sup>.

The fact that proton and neutron nearly have the same mass is understood since the 1930s, as consequence of an approximate symmetry: isospin [1]. For a long time, it was taken for granted that the symmetry is broken only by the e.m. interaction – the Cottingham formula [2] did explain the size of the observed mass differences with  $\Delta I = 2$ ,  $m_{\pi^+} - m_{\pi^0}$  for instance. For  $\Delta I = 1$ , however, in particular for  $m_p - m_n$ , the fact that the charged particle is lighter than the neutral one remained mysterious [3–5].

In 1975, Gasser and Leutwyler [6] pointed out that the mystery disappears if the popular conviction, according to which the strong interaction conserves isospin, is dismissed. They showed that a coherent picture of isospin breaking can be reached within the Quark Model, provided the masses of the two lightest quarks are not only very small but also very different. At that time, the experimental results on deep inelastic scattering were consistent with the scaling laws of Bjorken [7]. Evaluating the Cottingham formula with the scarce experimental information available then, they concluded that the elastic contributions dominate over the inelastic ones and arrived at the estimate  $m_{\text{QED}} = 0.7 \pm 0.3$  MeV [6].

Walker-Loud, Carlson and Miller [8] claimed that the analysis in [6] is incorrect and instead arrived at  $m_{\text{QED}} = 1.30(03)(47)$ . This

paper triggered renewed interest and several authors investigated the matter [9–12]. We will briefly discuss the results obtained in these works below. Some of the claims made in [8] are rectified in Appendix E of [13] and in [14,15].

A thorough account of our recent work on the subject with the technical details of the calculation is given in [16]. The aim of the present paper is to describe the basic ideas underlying our analysis and the conclusions drawn from it.

**Cottingham formula.** We denote the e.m. self-energy of the particle by  $m_\gamma$ . As shown by Cottingham [2], it is determined by the spin averaged forward Compton scattering amplitude,

$$T^{\mu\nu}(p, q) = \frac{i}{2} \int d^4x e^{iq \cdot x} \langle p | T j^\mu(x) j^\nu(0) | p \rangle. \quad (1)$$

Current conservation, Lorentz invariance and symmetry under space reflection imply that  $T^{\mu\nu}$  can be expressed in terms of two functions  $T_1(v, q^2), T_2(v, q^2)$  that only depend on  $v = p \cdot q/m$  and  $q^2$  ( $m$  is the mass of the particle). We use the notation of [13]:

$$T^{\mu\nu}(p, q) = (q^\mu q^\nu - g^{\mu\nu} q^2) T_1(v, q^2) \quad (2)$$

$$+ \frac{1}{m^2} \{ (p^\mu q^\nu + p^\nu q^\mu) p \cdot q - g^{\mu\nu} (p \cdot q)^2 \\ - p^\mu p^\nu q^2 \} T_2(v, q^2).$$

\* Corresponding author.

E-mail addresses: [gasser@itp.unibe.ch](mailto:gasser@itp.unibe.ch) (J. Gasser), [leutwyler@itp.unibe.ch](mailto:leutwyler@itp.unibe.ch)  
 (H. Leutwyler), [rusetsky@hiskp.uni-bonn.de](mailto:rusetsky@hiskp.uni-bonn.de) (A. Rusetsky).

The Cottingham formula represents  $m_\gamma$  as an integral over the four components of the photon momentum  $q$ . In the rest frame of the particle, the analytic properties of the time-ordered product allow one to perform a Wick rotation that turns the path of integration in the variable  $q^0$  from the real axis into the imaginary axis,  $q^0 = iQ_4$  [2,13]. The variable  $\nu$  coincides with  $q^0$  and thus becomes purely imaginary. Identifying  $Q_1, Q_2, Q_3$  with the space components of the physical momentum, we have  $q^2 = -Q^2$ , where  $Q$  is the length of the euclidean four-vector  $Q_\mu$ . The Cottingham formula then takes the form of an integral over euclidean space:

$$m_\gamma = \frac{e^2}{2m(2\pi)^4} \int \frac{d^4 Q}{Q^2} \phi, \\ \phi = 3Q^2 T_1 + (2Q_4^2 + Q^2) T_2, \quad (3)$$

where  $T_1, T_2$  are to be evaluated at  $\nu = iQ_4, q^2 = -Q^2$ .

*Operator product expansion.* The asymptotic behaviour of the integrand in formula (3) is controlled by the operator product expansion [17–22]. The leading contributions are determined by the Wilson coefficients of the operators of lowest dimension, which carry either spin 0 or spin 2. The explicit expressions [21–23] show that the contributions from  $T_1$  and  $T_2$  both diverge – the formula (3) must be regularized, e.g. by cutting the integral off with  $Q^2 \leq \Lambda^2$ . We denote the regularized version of  $m_\gamma$  by  $m_\gamma^\Lambda$ .

Since the operators of spin 2 are of anomalous dimension, their contributions to the asymptotic behaviour of  $T_1$  and  $T_2$  involve a negative fractional power of  $\ln Q^2$ , so that asymptotics sets in only very slowly. The key observation in our evaluation of the Cottingham formula is that the spin 2 contributions to  $T_1$  and  $T_2$  are the same – apart from the sign and a factor of 2. In the amplitude

$$\bar{T}(\nu, q^2) = T_1(\nu, q^2) + \frac{1}{2} T_2(\nu, q^2), \quad (4)$$

the leading short distance singularities of spin 2 drop out. Accordingly, this amplitude approaches asymptotics much more rapidly than the individual terms. The advantage of replacing  $T_1$  with  $\bar{T}$  also shows up in the asymptotic behaviour of the integrand in (3): the divergence stems from the first term in the decomposition  $\phi = 3Q^2 \bar{T} - \frac{1}{2}(Q^2 - 4Q_4^2) T_2$ . The angular integration suppresses the second when  $Q^2 \rightarrow \infty$ .

*Dispersion relations.* The amplitudes  $\bar{T}(\nu, -Q^2)$  and  $T_2(\nu, -Q^2)$  obey fixed- $Q^2$  dispersion relations in the variable  $\nu$ . The imaginary parts are determined by the structure functions:

$$\text{Im}\bar{T} = \pi \bar{F}/2xQ^2, \quad \text{Im}T_2 = \pi 2m^2 x F_2/Q^4, \\ \bar{F} = F_L + 2m^2 x^2 F_2/Q^2, \quad (5)$$

with  $x = Q^2/2m\nu$  and  $F_L \equiv F_2 - 2xF_1$ . Regge asymptotics implies that  $T_2$  obeys an unsubtracted dispersion relation, while  $\bar{T}$  requires a subtraction:

$$\bar{T}(\nu, -Q^2) = \bar{T}^{\text{el}}(\nu, -Q^2) + \bar{S}(-Q^2) \\ + (Q^2 + 4\nu^2) \int_0^{x_{\text{th}}} dx \frac{m^2 \bar{F}(x, Q^2)}{(Q^2 + m^2 x^2)(Q^4 - 4m^2 x^2 \nu^2 - i\epsilon)}, \\ T_2(\nu, -Q^2) = T_2^{\text{el}}(\nu, -Q^2) + \int_0^{x_{\text{th}}} dx \frac{4m^2 F_2(x, Q^2)}{Q^4 - 4m^2 x^2 \nu^2 - i\epsilon}. \quad (6)$$

The elastic parts,  $\bar{T}^{\text{el}}$  and  $T_2^{\text{el}}$ , are unambiguously determined by the nucleon form factors [6,13,15]. In the dispersion integrals, we have replaced the variable of integration with  $\nu' = Q^2/(2mx)$ . The

upper limit represents the boundary of the inelastic region,  $x_{\text{th}} = Q^2/(Q^2 + 2mM_\pi + M_\pi^2)$ .

Note that we are not subtracting the dispersion integral for  $\bar{T}$  at  $\nu = 0$ , but at  $\nu = \frac{i}{2}Q$ . This simplifies the analysis further, as it implies that, for  $\nu = iQ_4$ , the subtracted integral picks up the factor  $(Q^2 - 4Q_4^2)$ , so that the angular average suppresses that part as well.

*Renormalization.* In the framework of QCD+QED, the mass of a particle is determined by the bare parameters that occur in the Lagrangian and the cutoff used to regularize the theory. If the electromagnetic interaction is turned off, only the QCD coupling constant, the quark masses and the cutoff are relevant. To order  $e^2$ , the e.m. interaction changes the mass not only by the regularized version of equation (3), but in addition by the contribution  $\Delta m^\Lambda$ , which arises from the change in the bare parameters needed for the mass of the particle to stay finite when the cutoff is removed: the bare quantities depend on the cutoff as well as on  $e$ . The e.m. contribution to the mass is given by

$$m_{\text{QED}} = \lim_{\Lambda \rightarrow \infty} \{m_\gamma^\Lambda + \Delta m^\Lambda\}. \quad (7)$$

*Decomposition of the Cottingham formula.* With our decomposition of the Compton amplitude, the renormalized Cottingham formula consists of four parts [16]:

$$m_{\text{QED}} = m_{\text{el}} + m_{\bar{F}} + m_{F_2} + m_{\bar{S}}. \quad (8)$$

While the first term collects the elastic contributions, the second and third ones represent the contributions from the integrals occurring in the dispersion relations for  $\bar{T}$  and  $T_2$ . In the first three parts, the limit  $\Lambda \rightarrow \infty$  can be taken – the explicit expressions involve integrals over the elastic form factors and structure functions of the nucleon [16]. The divergence resides in the fourth term, which contains the contributions from subtraction function and renormalization,

$$m_{\bar{S}} = \lim_{\Lambda \rightarrow \infty} \left\{ N \int_0^{\Lambda^2} dQ^2 Q^2 \bar{S}(-Q^2) + \Delta m^\Lambda \right\}, \quad (9)$$

where the constant  $N$  stands for  $3\alpha_{\text{em}}/8\pi m$ .

*Asymptotic behaviour of the subtraction function.* The operator product expansion implies that  $\bar{S}$  falls off in proportion to  $1/Q^4$  when  $Q^2$  becomes large, while  $\Delta m^\Lambda$  grows logarithmically with  $\Lambda$ . For  $m_{\bar{S}}$  to stay finite, the leading contributions must match:

$$\bar{S}(-Q^2) \rightarrow \frac{C}{Q^4}, \quad \Delta m^\Lambda \rightarrow -NC \ln \frac{\Lambda^2}{\mu^2}. \quad (10)$$

The constant  $C$  is related to the matrix elements of the lowest dimensional operators of spin 0.

In the following, we consider the difference between proton and neutron, without explicitly indicating this in the notation: we write  $\bar{T}$  for  $\bar{T}^{p-n}$  and likewise for  $\bar{S}, C, \bar{F}, T_2, \dots$ . In the proton-neutron mass difference, we work to first order in the isospin breaking parameters  $m_u - m_d$  and  $e^2$  and neglect contributions of  $O[e^2(m_u - m_d)]$ . The constant  $C$  can then be expressed in terms of the proton matrix elements of  $\bar{u}u - \bar{d}d$ :

$$C = \frac{4m_u - m_d}{9} \langle p | \bar{u}u - \bar{d}d | p \rangle. \quad (11)$$

The same matrix element also determines the leading contribution to the QCD part of the mass difference [24]:

$$m_{\text{QCD}} = \frac{m_u - m_d}{2m} \langle p | \bar{u}u - \bar{d}d | p \rangle \{1 + O(m_u - m_d)\}. \quad (12)$$

The crude estimate  $m_{\text{QCD}} \approx -2$  MeV and the known quark mass ratios imply that  $C$  is tiny: with the lattice result  $m_s/m_{ud} = 27.23(10)$  [25] and the value of the ratio  $Q \equiv \sqrt{(m_s^2 - m_{ud}^2)/(m_d^2 - m_u^2)} = 22.1(7)$  extracted from  $\eta$ -decay [26], we obtain  $C \approx 6 \cdot 10^{-4}$  GeV<sup>2</sup>. The approximate chiral symmetry of the Standard Model very strongly suppresses the asymptotic behaviour of  $\bar{S}$ .

The expression (11) for the constant  $C$  receives corrections from higher orders of the perturbation series. These imply that the Cottingham formula contains subleading divergences proportional to  $\ln \ln \Lambda$  – for details, we refer to [16]. Since chiral symmetry suppresses the entire contribution from the region where perturbation theory applies, the corresponding effects in the renormalized mass difference are tiny and can be neglected. Setting  $\bar{\mu} \equiv \exp(-\frac{1}{2})\mu$ , the expression for  $\Delta m^\Lambda$  in (10) differs from  $-NC \int_0^{\Lambda^2} dQ^2 Q^2 / (\bar{\mu}^2 + Q^2)^2$  only by terms of order  $\mu^2/\Lambda^2$ . The contribution to  $m_{\text{QED}}$  that arises from the subtraction function can thus be written in the form [16]

$$m_{\bar{S}} = N \int_0^\infty dQ^2 Q^2 \left\{ \bar{S}(-Q^2) - \frac{C}{(\bar{\mu}^2 + Q^2)^2} \right\}. \quad (13)$$

This distinguishes our evaluation of the Cottingham formula from those in the literature [8–12], where the integral as well as the counter term are evaluated at a finite value of the cutoff around  $\Lambda^2 \approx 2$  GeV<sup>2</sup>. This would be legitimate if asymptotics were reached at such a low scale, but for the subtraction function used in these references, that is not the case.

*Role of spin 2 operators:  $\bar{S}$  versus  $S_1$ .* Traditionally, the subtraction function is identified with a multiple of  $S_1(q^2) \equiv T_1(0, q^2)$ . Asymptotically,  $Q^4 S_1$  approaches the same constant  $C$  as  $Q^4 \bar{S}$ , but this happens much more slowly: in contrast to  $\bar{S}$ , the function  $S_1$  does pick up spin 2 contributions [21] and these are not proportional to the lightest quark masses. This implies that, in the pre-asymptotic region, the two subtraction functions behave quite differently. In  $S_1$ , it takes extremely large values of  $Q^2$  for the spin 0 term to finally win over those of spin 2. When working with a low cutoff, this affects the result for the mass difference quite substantially – see below.

In contrast to the mass itself,  $m_{\text{QED}}$  depends on the renormalization scale  $\mu$ : the splitting into a contribution from QCD and one from QED is a matter of convention. In the decomposition (8), the parameter  $\mu$  resides in  $m_{\bar{S}}$  and enters through the term  $\Delta m^\Lambda$  in (10). The above estimate for the size of the constant  $C$  shows, however, that the sensitivity of  $m_{\bar{S}}$  to  $\mu$  is extremely weak: increasing the scale by a factor of 2 increases the value of  $m_{\text{QED}}$  by about 1 keV. For definiteness, we set  $\mu = 2$  GeV.

*Reggeon dominance.* The behaviour of the Compton amplitude in the limit  $q = \lambda \bar{q}$ ,  $\lambda \rightarrow \infty$  is controlled by the operator product expansion, which implies that both  $\bar{T}$  and  $T_2$  tend to zero in this limit. The gluons as well as the quarks reggeize, however [27–32]: in the limit where  $\nu$  becomes large while  $q^2$  is kept fixed, only  $T_2$  disappears,  $\bar{T}$  diverges.

In the Compton amplitudes of proton and neutron, the leading terms stem from singlet contributions due to Pomeron exchange, but in the difference between the two, these drop out: the dominating contributions to  $\bar{T}$  stem from the exchange of the leading Reggeon with  $I^C = 1^+$ , which we refer to as the  $a_2$ . It generates a Regge pole in the angular momentum plane which moves along the trajectory  $\alpha(t)$ . In the forward direction, only the value at  $t = 0$

matters: at fixed  $q^2$ , the contribution from the  $a_2$  grows with the power  $\nu^\alpha$ :

$$\bar{T}^R(\nu, q^2) = -\frac{\pi \beta(q^2)}{\sin \pi \alpha} \{(-s)^\alpha + (-u)^\alpha\}, \quad (14)$$

where  $\alpha$  stands for  $\alpha(0)$  and the variables  $s = m^2 + 2m\nu + q^2$ ,  $u = m^2 - 2m\nu + q^2$  represent the square of the centre of mass energy in the  $s$ - and  $u$ -channels, respectively. The value of  $\alpha$  is experimentally well determined from the high energy behaviour of hadronic cross sections and is in the vicinity of  $\alpha \approx 0.55$  – the uncertainties in  $\alpha$  are too small to affect our results.

Reggeization implies that the dispersion relation for  $\bar{T}$  requires a subtraction. We assume that the Reggeons fully determine the asymptotic behaviour [4,6,13],

$$\lim_{\nu \rightarrow \infty} (\bar{T} - \bar{T}^R) = 0, \quad (15)$$

and that the remainder disappears sufficiently fast for the difference  $\bar{T} - \bar{T}^R$  to obey an unsubtracted dispersion relation. We refer to this assumption as Reggeon dominance.

A nonzero limiting value in (15) would represent a fixed pole in  $\bar{T}$  at  $\alpha = 0$ . We do not know of a physical phenomenon that could produce such a term – neither causality, nor the short-distance singularities, nor the Reggeons generate terms of this sort. The presence of a fixed pole would mean that the high energy behaviour of the Compton amplitude is not understood.

For  $T_2$ , the contribution from the Reggeons tends to zero in proportion to  $\nu^{\alpha-2}$ . The generally accepted assumption that this amplitude obeys an unsubtracted dispersion relation immediately implies that it also obeys the Reggeon dominance condition (15). Note, however, that the expansion of the dispersion integral for  $T_2$  in inverse powers of  $\nu$  contains a term proportional to  $1/\nu^2$ . As pointed out by Damashek and Gilman [33] and, independently by Dominguez et al. [34], this term corresponds to a fixed pole in  $T_2$ , at  $\alpha = 0$  (there is an analogous term also in  $\bar{T}$ , but it represents a fixed pole with  $\alpha = -2$  and is at most of academic interest).

*Sum rule.* Elitzur and Harari [4] pointed out that if the exchange of Reggeons correctly describes the asymptotic behaviour in the limit  $\nu \rightarrow \infty$  at fixed  $q^2$ , then the subtraction function obeys a sum rule which fully determines it through the cross section of lepton-nucleon scattering. The sum rule relevant for our decomposition of the Compton amplitude exclusively involves the structure function  $\bar{F}$ . At small values of  $x$ , this quantity is dominated by Reggeon exchange:

$$\bar{F}^R = b(Q^2) x^{1-\alpha}, \quad b(Q^2) = 2Q^{2(\alpha+1)} \beta(-Q^2). \quad (16)$$

In the difference  $\bar{F} - \bar{F}^R$ , the leading term cancels out. As demonstrated in [16], the sum rule for  $\bar{S}$  can be brought to the form:

$$Q^2 \bar{S}(-Q^2) = \int_0^{x_{\text{th}}} dx \frac{\bar{F}(x, Q^2) - \bar{F}^R(x, Q^2)}{x^2} - \frac{b(Q^2)}{\alpha x_{\text{th}}^\alpha} - \int_0^{x_{\text{th}}} dx \frac{m^2 \bar{F}(x, Q^2)}{Q^2 + m^2 x^2}. \quad (17)$$

In [6], the violations of Bjorken scaling were ignored: it was assumed that for  $Q^2 \rightarrow \infty$ , the structure function  $\bar{F}$  tends to  $(2xH_1 + F_2)x^2 m^2 / Q^2$ , where  $H_1$  and  $F_2$  only depend on  $x$ . One readily checks that the sum rule (17) then indeed reduces to the relation between the operator matrix element  $C$  and the structure functions given in (5.2), (5.3), (13.14) of [6]. Scaling would imply that the last term on the r.h.s. of (17) tends to zero  $\propto 1/Q^4$ . The scaling violations merely make it disappear less rapidly, in proportion to  $1/Q^2 / (\ln Q^2)^{1+d_2}$  with  $d_2 > 0$  [16].

*Elastic contributions.* In recent years, the precision to which the elastic form factors are known has increased significantly [35–38]. The results obtained with the three parametrizations of [35–37] are in the range

$$m_{\text{el}} = 0.75 \pm 0.02 \text{ MeV}. \quad (18)$$

Note that the amplitudes used in the literature often have kinematic zeros – this can make it difficult not only to sort out the asymptotic behaviour, but also to identify the elastic part of the dispersive representation (“Born term”) with the contribution generated by the one-particle intermediate state [13,15]. In [12], for instance, it is assumed that the amplitude  $\hat{T} = q^2 T_1 + \nu^2 T_2$  satisfies the asymptotic condition (15). That assumption, however, requires  $q^2 T_1$  to contain a fixed pole which compensates the one in  $\nu^2 T_2$  and hence violates Reggeon dominance.

*Structure functions, parton distributions.* For the numerical evaluation of the dispersion integrals and of the sum rule for the subtraction function, we need a representation for the difference between the structure functions of proton and neutron, and not only for the relatively well explored quantity  $F_2$ , but also for the longitudinal component  $F_L$ , which is known less well. In the resonance region ( $W < 3$ ), we make use of the parametrizations of the structure functions in [39–43]. For  $W > 3$  and low photon virtualities ( $Q^2 < 1$ ), we invoke the Regge representation of Alwall and Ingelman (AI) [44].

At higher values of  $Q^2$ , the DGLAP equations [45–47] for the parton distribution functions (PDFs) provide a strong constraint for the analysis of the data: at leading order in  $\alpha_s$ , these equations imply that  $F_L$  is given by an integral over  $F_2$ . A vast amount of PDFs is available [48,49] and APFEL Web [50,51] provides a flexible, user-friendly tool for the evaluation of the corresponding structure functions. Since the quark masses  $m_u$  and  $m_d$  are tiny, the  $u$ - and  $d$ -distributions in the neutron must be very close to the  $d$ - and  $u$ -distributions in the proton, respectively. As emphasized e.g. in [52], this ensures that the  $u$ - and  $d$ -distributions can separately be determined by using neutral and charged current data on the proton – scattering on deuterons or heavier nuclei is not needed to sort out the difference between the  $u$ - and  $d$ -distributions of the proton.

As mentioned above, the behaviour in the Regge region ( $x$  small) is dominated by the singlet part of the distributions. This implies that the  $u$ - and  $d$ -distributions must approach one another when  $x \rightarrow 0$ . While the available data strongly constrain the singlet part at small  $x$ , the non-singlet PDFs are much less well determined. The same applies to the non-singlet structure function  $\bar{F}$  which plays a central role in our work.

Since reggeization involves sums to all orders of perturbation theory, it is not a simple matter to analyze the behaviour at small  $x$  in the framework of the DGLAP equations (for a review of the problems encountered in this endeavour, we refer to [53]). In particular, the requirement that the  $u$ - and  $d$ -distributions must approach one another in the Regge limit and that their difference yields a contribution to  $\bar{F}$  that falls off with  $b(Q^2)x^{1-\alpha}$ ,  $\alpha \approx 0.55$  imposes nontrivial theoretical constraints. A coherent parametrization of the PDFs that is consistent not only with the data, but also with these constraints, yet needs to be found.

In our calculation, we rely on the solutions of the DGLAP equations constructed by Alekhin, Blümlein and Moch (ABM) [52,54,55] in the region  $x > 0.01$ . At smaller values of  $x$ , we assume that  $\bar{F}$  is dominated by the Reggeon  $a_2$ , not only at small photon virtualities, but also at higher values of  $Q^2$ . We determine the residue  $b(Q^2)$  by smoothly matching the two parametrizations around  $x = 0.01$ . In the region  $W > 3$ ,  $Q^2 > 1$ , we estimate the uncertainty in our representation for the structure functions  $\bar{F}^{p-n}$  and  $F_2^{p-n}$  at 30%. For details, we refer to [16].

*Inelastic contributions.* The most striking aspect of our numerical result is that the two terms  $m_{\bar{F}}$  and  $m_{F_2}$  turn out to be tiny:  $m_{\bar{F}} + m_{F_2} = -0.004(1)$  MeV. As discussed above, the angular integration suppresses the integrands of these quantities at large values of  $Q^2$ , but the numerical result shows that the suppression is very efficient also in the low energy region. We conclude that – in our decomposition of the Compton amplitude – only the elastic contribution  $m_{\text{el}}$  and the term  $m_{\bar{S}}$  from the subtraction function play a significant role. Note that this statement holds independently of the assumptions used to determine the subtraction function. Since the various attempts at evaluating the Cottingham formula arrive at very similar values for the elastic part, the discrepancies in the results for the mass difference mainly come from the term  $m_{\bar{S}}$ , i.e. from the fact that the parametrizations used for the Compton amplitude yield different representations for the subtraction function  $\bar{S}$ .

*Contribution from the subtraction function.* It is straightforward to evaluate the sum rule for  $\bar{S}$  with the two representations of  $\bar{F}(x, Q^2)$  and  $b(Q^2)$  we are using below and above  $Q^2 = 1 \text{ GeV}^2$ , respectively and to calculate the corresponding contribution to  $m_{\bar{S}}$  with (13). Isospin conservation prevents the most prominent feature in the low energy region, the  $\Delta(1232)$ , to make a significant contribution. Moreover, the regions below and above a centre of mass energy of 3 GeV contribute with opposite sign – within errors, they cancel:  $m_{\bar{S}}(Q^2 < 1 \text{ GeV}^2) = -0.034(68)$  MeV. Note that the error is twice as large as the central value. It is dominated by the uncertainties in the resonance region and is of systematic nature, as it stems from the simplification used in the data analysis of Bosted and Christy [42,43]: the ratio  $R = \sigma_L/\sigma_T$  is assumed to be the same for proton and neutron. In the region where the Pomeron dominates, this holds to good accuracy, but we need the difference between the two, where Pomeron exchange drops out.

At  $Q^2 = 1 \text{ GeV}^2$ , where the representations AI and ABM meet, the results for the contributions to  $\bar{S}$  from  $W > 3 \text{ GeV}$  agree within errors: the two entirely different sources match, both in sign and in size. In order to interpolate between the values of  $Q^2$  where the ABM results provide significant information and the region where asymptotics sets in, we make use of the Generalized Vector Dominance Model of Sakurai and Schildknecht [56], parametrizing the subtraction function in terms of the contributions from  $\rho$ ,  $\omega$  and  $\phi$ . In the difference between proton and neutron, only the off-diagonal terms survive:

$$\bar{S}_{\text{VMD}}(-Q^2) = \frac{1}{m_\rho^2 + Q^2} \left\{ \frac{c_\omega}{m_\omega^2 + Q^2} + \frac{c_\phi}{m_\phi^2 + Q^2} \right\}. \quad (19)$$

The asymptotic condition (10) requires the two terms in the bracket to nearly cancel:  $c_\omega + c_\phi = C$ . This leaves a single parameter free, say  $c_\omega$ . Fitting the parametrization in the range between 2 and  $3.5 \text{ GeV}^2$ , we obtain  $c_\omega = -0.74(49) \text{ GeV}^2$ . We have checked that the outcome for  $m_{\bar{S}}$  is neither sensitive to the specific form of the interpolation nor to the range used in the fit. Numerically, this yields  $m_{\bar{S}}(Q^2 > 1 \text{ GeV}^2) = -0.13(9)$  MeV. Together with the contributions from low virtualities, this yields

$$m_{\bar{S}} = -0.17(16) \text{ MeV}. \quad (20)$$

Because asymptotic freedom fixes the asymptotic behaviour of the subtraction function, the parametrization obtained within Generalized Vector Meson Dominance contains a single free parameter. Instead of fixing it to the results obtained with the ABM solution of the DGLAP equations, we can dismiss the experimental information available for  $W > 3$ ,  $Q^2 > 1$  altogether and determine the free parameter with a fit to the results obtained for  $Q^2 < 1$ . This yields

$m_{\bar{S}} = -0.12(21)$  MeV: the central value stays well within the estimated uncertainty and the error only increases by about 30%. This indicates that our result is not sensitive to the input used in the region where the non-singlet contributions to the structure functions are not yet known well.

*Numerical result.* Collecting the various contributions and using the experimental value of the proton-neutron mass difference, the parts due to the e.m. interaction and to the difference between  $m_u$  and  $m_d$  become

$$m_{\text{QED}} = 0.58 \pm 0.16 \text{ MeV}, \quad m_{\text{QCD}} = -1.87 \pm 0.16 \text{ MeV}. \quad (21)$$

The result for  $m_{\text{QCD}}$  yields a more precise estimate for the leading Wilson coefficient:  $C = 5.7(1.1) \cdot 10^{-4}$  GeV<sup>2</sup>, but the corresponding shift in our results is negligibly small.

The conclusions reached in [6] are thus confirmed:  $m_{\text{QED}}$  is dominated by the elastic contribution. The uncertainty in the old result,  $m_{\text{QED}} = 0.7(3)$  MeV, is reduced by about a factor of two.

It is not difficult to understand why the inelastic contributions are so small: (a) the angular integration suppresses the contributions from the dispersion integrals, (b) if  $Q^2$  is large, the subtraction functions of proton and neutron are nearly the same – in the chiral limit, there is no difference, (c) in the region where Reggeon exchange dominates, the leading term, the Pomeron, is the same, (d) isospin symmetry ensures that the most important resonance, the  $\Delta(1232)$ , contributes equally to proton and neutron and (e) the elastic contributions also dominate the chiral perturbation series of the Compton amplitude: the leading terms exclusively contribute to the elastic part – inelastic processes merely generate higher order corrections (for a recent analysis of the subtraction function in  $\chi$ PT, we refer to [57]).

*Comparison with lattice results.* The determination of  $m_{\text{QED}}$  on a lattice is a very demanding goal. While the numbers in [58] cluster around  $m_{\text{QED}} \approx 0.7$  MeV, in agreement with our result, the values 1.00(7)(14) MeV [59], 1.03(17) [60] and 1.53(25)(50) MeV [61] are higher than ours. Adding statistical and systematic errors in quadrature, the various lattice results differ from the outcome of our calculation by less than two standard deviations. In the framework we are relying on, values like  $m_{\text{QED}} = 1$  MeV or even higher require sizeable positive contributions from  $m_{\bar{S}}$  – this is not compatible with Reggeon dominance.

*Comparison with other dispersive calculations.* The main difference between our analysis and the work reported in [8–11] is that, there, the subtraction function is not calculated, but parametrized with an ansatz in terms of its value at  $Q^2 = 0$  (taken from experiment) and a scale  $m_0$  that specifies the momentum dependence. Moreover, in these models, the parametrization is applied to  $S_1$  rather than to  $\bar{S}$ . As discussed above, the asymptotic behaviour of  $S_1$  picks up contributions from operators with spin 2, which fall off only extremely slowly. Since chiral symmetry suppresses the coefficient  $C$  of the leading asymptotic term, it starts dominating  $S_1$  only if  $Q^2$  becomes very large. This implies that the parametrization used in these models does not behave properly in the pre-asymptotic region, which does make a significant contribution to  $m_{\bar{S}}$ .

The mismatch with the asymptotics disappears if the ansatz in [9] is assumed to be valid for  $\bar{S}$  rather than  $S_1$ . The central value obtained for  $m_{\text{QED}}$  then drops to about 0.7 MeV, in agreement with what we find. The uncertainties in the result for  $m_{\text{QED}}$ , however, are much larger than ours, not only because the experimental values of the magnetic polarizabilities of proton and neutron, which play a key role in those models, are subject to large errors, but also because the result is quite sensitive to the shape of the parametrization used for low values of  $Q^2$ .

*Summary.* We have applied the Reggeon dominance hypothesis to the electromagnetic part of the proton-neutron mass difference. The uncertainty in our final result,  $m_{\text{QED}} = 0.58 \pm 0.16$  MeV, stems from a careful estimate of the errors coming from the different experimental data sets used in the calculations. While this confirms the old result [6], which also relies on Reggeon dominance, recent evaluations of the Cottingham formula [8–12] yield central values around 1 MeV or even higher. The difference stems from the short distance singularities associated with operators of spin 2 that are neglected in those references. The lattice determinations do not yet yield conclusive results, but the method is gradually improved. In the long run, these will achieve comparable accuracy and thereby put Reggeon dominance to a very stringent test.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgements

We thank Johannes Blümlein for providing us with numerical tables for the structure functions based on the ABM solutions of the DGLAP equations, Thomas Becher for information about the anomalous dimensions in QCD+QED, Vadim Lensky and Vladimir Pascalutsa for a Mathematica notebook concerning the representation of the Compton amplitude in  $\chi$ PT and Ryan Bignell, Irinel Caprini, Stefano Carrazza, Gilberto Colangelo, Cesareo Dominguez, Franziska Hagelstein, Bastian Kubis, Ulf-G. Meißner, Sven-Olaf Moch, Gerrit Schierholz and Ignazio Scimemi for comments and useful information. A.R. acknowledges the support from the DFG (CRC 110 “Symmetries and the Emergence of Structure in QCD”), as well as from Volkswagenstiftung under contract no. 93562 and the Chinese Academy of Sciences (CAS) President’s International Fellowship Initiative (PIFI) (grant no. 2021VMB0007).

### References

- [1] W. Heisenberg, Z. Phys. 77 (1932) 1.
- [2] W.N. Cottingham, Ann. Phys. 25 (1963) 424.
- [3] D.J. Gross, H. Pagels, Phys. Rev. 172 (1968) 1381.
- [4] M. Elitzur, H. Harari, Ann. Phys. 56 (1970) 81.
- [5] A. Zee, Phys. Rep. 3 (1972) 127.
- [6] J. Gasser, H. Leutwyler, Nucl. Phys. B 94 (1975) 269.
- [7] J.D. Bjorken, Phys. Rev. 179 (1969) 1547.
- [8] A. Walker-Loud, C.E. Carlson, G.A. Miller, Phys. Rev. Lett. 108 (2012) 232301, arXiv:1203.0254.
- [9] F.B. Erben, P.E. Shanahan, A.W. Thomas, R.D. Young, Phys. Rev. C 90 (2014) 065205, arXiv:1408.6628.
- [10] A.W. Thomas, X.G. Wang, R.D. Young, Phys. Rev. C 91 (2015) 015209, arXiv:1406.4579.
- [11] A. Walker-Loud, PoS CD 2018 (2019) 045, arXiv:1907.05459.
- [12] O. Tomalak, Eur. Phys. J. Plus 135 (2020) 411, arXiv:1810.02502.
- [13] J. Gasser, M. Hoferichter, H. Leutwyler, A. Rusetsky, Eur. Phys. J. C 75 (2015) 375; Erratum: Eur. Phys. J. C 80 (2020) 353, arXiv:1506.06747.
- [14] H. Leutwyler, PoS CD 2015 (2015) 022, arXiv:1510.07511.
- [15] M. Hoferichter, PoS CD 2018 (2019) 028.
- [16] J. Gasser, H. Leutwyler, A. Rusetsky, arXiv:2008.05806.
- [17] K.G. Wilson, Phys. Rev. 179 (1969) 1499.
- [18] D.J. Gross, F. Wilczek, Phys. Rev. Lett. 30 (1973) 1343.
- [19] H.D. Politzer, Phys. Rev. Lett. 30 (1973) 1346.
- [20] S. Weinberg, Phys. Rev. Lett. 31 (1973) 494.
- [21] J.C. Collins, Nucl. Phys. B 149 (1979) 90; Errata: Nucl. Phys. B 153 (1979) 546, Nucl. Phys. B 915 (2017) 392.
- [22] A.J. Buras, Rev. Mod. Phys. 52 (1980) 199.
- [23] R.J. Hill, G. Paz, Phys. Rev. D 95 (2017) 094017, arXiv:1611.09917.
- [24] J. Gasser, H. Leutwyler, Phys. Rep. 87 (1982) 77.
- [25] S. Aoki, et al., Flavour Lattice Averaging Group, Eur. Phys. J. C 80 (2020) 113, arXiv:1902.08191.
- [26] G. Colangelo, S. Lanz, H. Leutwyler, E. Passemar, Eur. Phys. J. C 78 (2018) 947, arXiv:1807.11937.

[27] M.T. Grisaru, H. Schnitzer, H.S. Tsao, Phys. Rev. Lett. 30 (1973) 811.

[28] M.T. Grisaru, H. Schnitzer, H.S. Tsao, Phys. Rev. D 8 (1973) 4498.

[29] L.N. Lipatov, Sov. J. Nucl. Phys. 23 (1976) 338, Yad. Fiz. 23 (1976) 642.

[30] E. Kuraev, L. Lipatov, V.S. Fadin, Sov. Phys. JETP 45 (1977) 199, Zh. Eksp. Teor. Fiz. 72 (1977) 377.

[31] I. Balitsky, L. Lipatov, Sov. J. Nucl. Phys. 28 (1978) 822, Yad. Fiz. 28 (1978) 1597.

[32] V. Gribov, The Theory of Complex Angular Momenta: Gribov Lectures on Theoretical Physics, Cambridge Monographs on Mathematical Physics, Cambridge University Press, Cambridge, 2003.

[33] M. Damashek, F.J. Gilman, Phys. Rev. D 1 (1970) 1319.

[34] C.A. Dominguez, C. Ferro Fontan, R. Suaya, Phys. Lett. B 31 (1970) 365.

[35] V. Punjabi, C.F. Perdrisat, M.K. Jones, E.J. Brash, C.E. Carlson, Eur. Phys. J. A 51 (2015) 79, arXiv:1503.01452.

[36] J.J. Kelly, Phys. Rev. C 70 (2004) 068202.

[37] Z. Ye, J. Arrington, R.J. Hill, G. Lee, Phys. Lett. B 777 (2018) 8, arXiv:1707.09063.

[38] K. Borah, R.J. Hill, G. Lee, O. Tomalak, arXiv:2003.13640.

[39] D. Drechsel, S.S. Kamalov, L. Tiator, Eur. Phys. J. A 34 (2007) 69, arXiv:0710.0306.

[40] S.S. Kamalov, S.N. Yang, D. Drechsel, O. Hanstein, L. Tiator, Phys. Rev. C 64 (2001) 032201, arXiv:nucl-th/0006068.

[41] M. Hilt, B.C. Lehnhart, S. Scherer, L. Tiator, Phys. Rev. C 88 (2013) 055207, arXiv: 1309.3385.

[42] P.E. Bosted, M.E. Christy, Phys. Rev. C 77 (2008) 065206, arXiv:0711.0159.

[43] M.E. Christy, P.E. Bosted, Phys. Rev. C 81 (2010) 055213, arXiv:0712.3731.

[44] J. Alwall, G. Ingelman, Phys. Lett. B 596 (2004) 77, arXiv:hep-ph/0402248.

[45] Y.L. Dokshitzer, Sov. Phys. JETP 46 (1977) 641, Zh. Eksp. Teor. Fiz. 73 (1977) 1216.

[46] V. Gribov, L. Lipatov, Sov. J. Nucl. Phys. 15 (1972) 438, Yad. Fiz. 15 (1972) 781.

[47] G. Altarelli, G. Parisi, Nucl. Phys. B 126 (1977) 298.

[48] J.R. Andersen, et al., arXiv:1405.1067.

[49] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht, M. Schönherr, G. Watt, Eur. Phys. J. C 75 (2015) 132, arXiv:1412.7420.

[50] V. Bertone, S. Carrazza, J. Rojo, Comput. Phys. Commun. 185 (2014) 1647, arXiv: 1310.1394.

[51] S. Carrazza, A. Ferrara, D. Palazzo, J. Rojo, J. Phys. G 42 (2015) 057001, arXiv: 1410.5456.

[52] S. Alekhin, J. Blümlein, S. Moch, R. Placakyte, Phys. Rev. D 96 (2017) 014011, arXiv:1701.05838.

[53] A.M. Cooper-Sarkar, R.C.E. Devenish, A. De Roeck, Int. J. Mod. Phys. A 13 (1998) 3385, arXiv:hep-ph/9712301.

[54] S. Alekhin, J. Blümlein, S. Moch, Phys. Rev. D 89 (2014) 054028, arXiv:1310.3059.

[55] S. Alekhin, J. Blümlein, S. Moch, PoS DIS 2019 (2019) 002, arXiv:1909.03533.

[56] J.J. Sakurai, D. Schildknecht, Phys. Lett. B 40 (1972) 121.

[57] J. Lozano, A. Agadjanov, J. Gegelia, U.-G. Meißner, A. Rusetsky, arXiv:2010.10917.

[58] M.G. Endres, A. Shindler, B.C. Tiburzi, A. Walker-Loud, Phys. Rev. Lett. 117 (2016) 072002, arXiv:1507.08916.

[59] S. Borsanyi, et al., Science 347 (2015) 1452, arXiv:1406.4088.

[60] D.A. Brantley, B. Joo, E.V. Mastropas, E. Mereghetti, H. Monge-Camacho, B.C. Tiburzi, A. Walker-Loud, arXiv:1612.07733.

[61] R. Horsley, et al., CSSM QCDSF UKQCD Collaborations, J. Phys. G 46 (2019) 115004, arXiv:1904.02304.