Zemach moments of the proton from Bayesian inference

Krzysztof M. Graczyk* and Cezary Juszczak
Institute for Theoretical Physics, University of Wrocław, plac M. Borna 9, 50-204 Wrocław, Poland
(Received 9 March 2015; published 21 April 2015)

The first and the third Zemach moments are obtained, $\langle r\rangle_{(2)} = 1.1108 \pm 0.0021$ fm and $\langle r^3\rangle_{(2)} = 2.889 \pm 0.008$ fm^3, from the Bayesian analysis of the elastic ep scattering data. The quantitative discussion of the dependence of the results on the parametrization choice is presented and the corresponding systematic uncertainties are estimated—about 0.6% and 1.6% for the first and the third Zemach moments, respectively.

DOI: 10.1103/PhysRevC.91.045205 PACS number(s): 13.40.Gp, 25.30.Bf, 14.20.Dh, 31.30.jr

I. INTRODUCTION

The moments of the electric charge distribution, $\rho_E(\mathbf{r})$, of the proton are the input for the theoretical calculations of hadronic structure corrections to the energy spectrum in simple atomic systems [1,2]. They contain the information about the structure of the proton. For instance, the second moment of charge distribution (mean square radius),

$$r_p^2 \equiv \langle r^2\rangle = \int d^3\mathbf{r}\, r^2\, \rho_E(\mathbf{r}), \tag{1}$$

gives the definition of the proton radius r_p.

The analysis of the measurements of the energy spectrum in simple atomic systems allows one to study the theory of bound states and the relativistic effects of quantum electrodynamics (QED) and also presents the opportunity to investigate the hadronic proton structure, in particular, the electromagnetic proton form factors.

On the other hand the electric, G_E, and magnetic, G_M, proton form factors are extracted from the elastic ep scattering data. In some approximations [3,4] the G_E and G_M proton form factors can be related to the distributions of the electric charge and the magnetic moment inside the proton.

The value of the proton radius obtained from hydrogen atom measurements is consistent with the values of r_p extracted by many groups from elastic ep scattering data [5,6] including recent low-Q^2 ep cross-section measurements [7,8] (for a more complete list of references see Ref. [9]). However, they cannot be reconciled with a recent, currently the most accurate, value of the proton radius obtained from the Lamb-shift measurement in the muonic atom [10]. On the other hand the latter is in agreement with the results of the analysis of the scattering data within the dispersion-based approach [11,12]. The source of this discrepancy is not well understood [13,14] and it is referred to as *the proton radius puzzle*.

The theoretical value of the Lamb shift in the muonic hydrogen is dominated by the QED contribution but the corrections due to the nucleus finite size are non-negligible. They are determined by the moments of the electric charge distribution [15,16]. Indeed, the theoretical value of the Lamb shift reads [1,17,18]

$$L_{\rm th} = 209.9779(49) - 5.2262\frac{\langle r^2\rangle}{\rm fm^2} + 0.00913\frac{\langle r^3\rangle_{(2)}}{\rm fm^3}, \tag{2}$$

where the energy above is in meV, while

$$\langle r^3\rangle_{(2)} = \int d^3\mathbf{r}' d^3\mathbf{r}\rho_E(\mathbf{r}')\rho_E(\mathbf{r})|\mathbf{r}' - \mathbf{r}|^3 \tag{3}$$

is the third Zemach moment [19]. The experimental value is $L_{\rm ex} = 206.2949(32)$ meV [10].

From comparison of $L_{\rm th}$ and $L_{\rm ex}$ it follows that the Lamb-shift measurement can be reconciled with the $r_p \sim 0.87$–0.89 fm supported by many analyses of scattering data and hydrogen spectroscopy, if the third Zemach moment $\langle r^3\rangle_{(2)}$ is large enough—about 36 fm^3 (see discussion in Ref. [20]). However, the value of $\langle r^3\rangle_{(2)}$ calculated from the empirical fits of the form factors ranges from 2.00 to 2.90 fm^3 and is too small to explain the proton radius puzzle [21,22]. Nevertheless, this contribution to $L_{\rm th}$ must be taken into account to obtain an accurate value of the proton radius with a precise estimate of the uncertainty.

The aim of this paper is to infer the values of the Zemach moments from the Bayesian analysis of the ep scattering data. In particular we discuss the dependence of the Zemach moments on the choice of form factor parametrization. We concentrate on $\langle r^3\rangle_{(2)}$ but obtain also the first Zemach moment,

$$\langle r\rangle_{(2)} = \int d^3\mathbf{r}\, d^3\mathbf{r}''\rho_E(\mathbf{r})\rho_M(\mathbf{r} - \mathbf{r}'')|\mathbf{r}''|, \tag{4}$$

which contributes to the finite-size correction to the hydrogen hyperfine splitting interval [2] [$\rho_M(\mathbf{r})$ is the distribution of the magnetic moment inside the proton]. As is discussed in Ref. [23], its accurate estimate can be relevant for the extraction of the magnetic proton radius from spectroscopic measurements.

The present paper is a natural continuation of our previous work [24], where we utilized the Bayesian framework to extract the value of the proton radius from the global analysis of the ep scattering data. The results obtained in Ref. [24] are used here to estimate the Zemach moments.

One of the problems of the extraction of the proton radius from the scattering data is the dependence of the results of the analysis on the model assumptions, in particular, on the choice of the functional form of the form factor parametrization [8,25]. Indeed in Ref. [24] we considered

*krzysztof.graczyk@ift.uni.wroc.pl

about 40 distinct form factor parametrizations and obtained proton radius values ranging from 0.81 to 0.90 fm. Another problem is the choice of the optimal number of parameters in the fit. If this number is large the fits tend to reproduce exactly the data, i.e., overfit data, and lose predictive power. If it is too small the model cannot properly reproduce the data. This is the so-called bias-variance trade-off dilemma.

Such problems are addressed within the framework of Bayesian statistics as described in Ref. [24] where we constructed and calculated a probability measure, which provided a criterion for choosing the model most favorable by the data. In this approach the overfitting problem is also solved because too complicated fits are naturally penalized by the Bayesian algorithm. A different approach to these problems was presented in Ref. [25] where analytic properties of the form factors were utilized to construct the optimal form factor parametrization.

In this paper we use the results of Ref. [24] to estimate the Zemach moments and discuss their dependence on the choice of form factor parametrization. It is done based on the analysis of about 200 statistical models. We obtained $\langle r^3\rangle_{(2)} = 2.889 \pm 0.008 \pm 0.046\,(\text{sys})\,\text{fm}^3$ and for the first Zemach moment we obtained $\langle r\rangle_{(2)} = 1.1108 \pm 0.0021 \pm 0.0071\,(\text{sys})\,\text{fm}$.

This paper is organized as follows. In Sec. II all necessary formulas needed for computing the Zemach moments are introduced. In Sec. III the Bayesian approach is briefly reviewed and the numerical results are presented and discussed.

II. ZEMACH MOMENTS

In elastic ep scattering the energy transfer between the electron and the proton vanishes in the Breit frame; hence in this frame the four-momentum transfer is $q^\mu = (0,\mathbf{q})$. In this frame the distributions of the electric charge $\rho_E(\mathbf{r})$ and the magnetic moment $\rho_M(\mathbf{r})$ can be related to the electric G_E and magnetic G_M form factors of the proton by the Fourier transformations [3]:

$$G_E(-\mathbf{q}) = \int d^3\mathbf{r}\rho_E(\mathbf{r})e^{-i\mathbf{r}\cdot\mathbf{q}}, \tag{5}$$

$$\frac{G_M(-\mathbf{q})}{\mu_p} = \int d^3\mathbf{r}\rho_M(\mathbf{r})e^{-i\mathbf{r}\cdot\mathbf{q}}, \tag{6}$$

where μ_p is the magnetic moment of the proton. The form factors are Lorentz scalars depending only on $Q^2 = -q^\mu q_\mu$, which in the Breit frame is equal to $\mathbf{q}^2$. A comparison of $\rho_E(\mathbf{r})$ distributions obtained from the dipole form factor and the Bayesian fit ($Q^2_{\text{max}} = 1$) [24] of G_E is shown in Fig. 1.

The third Zemach moment can be approximated by the integral [26] (see Eq. (59) in Ref. [17]):

$$\langle r^3\rangle_{(2)} \approx \lim_{c\to\infty} I_3(c,G_E), \tag{7}$$

$$I_3(c,G_E) = \frac{48}{\pi}\int_0^c \frac{dq}{q^4}\left(G_E^2(-q) - 1 + \frac{q^2}{3}\langle r^2\rangle\right), \tag{8}$$

$$\langle r^2\rangle = -6\,\frac{dG_E}{d\mathbf{q}^2}\bigg|_{\mathbf{q}^2=0}, \tag{9}$$

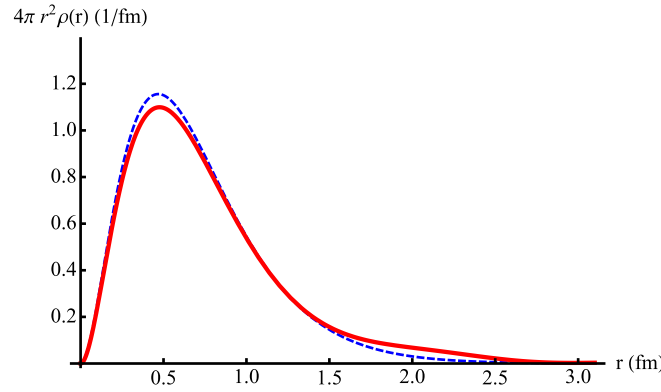

FIG. 1. (Color online) Charge distribution inside the proton obtained from Eq. (5) for the dipole form factor (15) (dashed blue line) and the Bayesian fit ($Q^2 < 1$ GeV^2) from Ref. [24] (solid red line).

while the first Zemach moment (4) is given by the integral:

$$I_1 = \lim_{c\to\infty} I_1(c,G_E,G_M), \tag{10}$$

$$I_1(c,G_E,G_M) = -\frac{4}{\pi}\int_0^c \frac{dq}{q^2}\left(\frac{G_E(-q)G_M(-q)}{\mu_p} - 1\right). \tag{11}$$

III. BAYESIAN ANALYSIS AND NUMERICAL RESULTS

The idea of the Bayesian framework utilized in Ref. [24] to extract the proton radius from scattering data is described in more detail in Refs. [27–29]. In this approach the electric and magnetic form factors are simultaneously parametrized by one feed-forward neural network (with one hidden layer of units) with two outputs:[1]

$$\mathcal{N}_H(Q^2;\{w_i\}) = \left(o_M^H,\quad o_E^H\right), \tag{12}$$

where $\{w_i\}$ is the set of neural network parameters, and o_M^H and o_E^H parametrize the departure from the dipole shape of the form factors:

$$\frac{G_M^H(Q^2)}{\mu_p} = \left[1 - Q^2 o_M^H(Q^2)\right]G_D(Q^2), \tag{13}$$

$$G_E^H(Q^2) = \left[1 - Q^2 o_E^H(Q^2)\right]G_D(Q^2), \tag{14}$$

$$G_D(Q^2) = \left(1 + Q^2/M_V^2\right)^{-2}, \tag{15}$$

where $M_V^2 = 0.71$ GeV^2. H denotes a number of hidden units in the neural network and ranges from 2 to 40.

Within each class of parametrization $\mathcal{N}_H$, with given H, we found the most optimal configuration of neural network parameters $\{w_i\}_{\text{MP}}$ that maximizes the prior probability[2] $P(\{w_i\}|\mathcal{D},\mathcal{N}_H)$, where $\mathcal{D}$ denotes the data, which in this case includes cross-section and polarization-transfer

[1] According to the Cybenko theorem [30–32] any continuous function can be approximated by a function from this class with arbitrary precision.

[2] The construction of the objective prior probability for neural network parametrization is much easier than in the case of the typical form factor parametrization.

TABLE I. The numerical values (in fm^3 units) of the third Zemach moment (last column) and integral $I_3(c,G_E,G_M)$ calculated for several values of c.

c (GeV)	1.0	3.0	10	∞
Fit $Q^2 < 1$ GeV^2	2.091	2.404	2.609	2.865 ± 0.005
Fit $Q^2 < 3$ GeV^2	2.099	2.411	2.617	2.872 ± 0.004
Fit $Q^2 < 10$ GeV^2	2.129	2.440	2.645	2.899 ± 0.012

measurements. From the cross-section data the two-photon exchange corrections, calculated in Ref. [33], have been subtracted. It was an important point of the analysis; indeed the results for two-photon exchange corrected data were characterized by much smaller variance than in the case of not modified data.

For every parametrization class $\mathcal{N}_H$ the evidence $P(\mathcal{D}|\mathcal{N}_H)$ was calculated. This probabilistic measure ranks the models—the most probable model, favorable by the data, is the one that maximizes the evidence.

Certainly the results of the analysis may depend on the selection of the data. We treated each data set on the same footing but to discuss the impact of the data selection on the final results we considered three subsets of the measurements denoted as $\mathcal{D}_{Q^2_{\max}}$ and obtained by removing data points with Q^2 above $Q^2_{\max} = 1$, 3, and 10 GeV^2, respectively. It turns out that the proton radius results weakly depend on this cut.

The main result of the analysis [24] is the family of the statistical models. Each model contains the optimal configuration of parameters $\{w_i\}_{\mathrm{MP}}$ and the corresponding posterior probability as well as the evidence:

$$\mathcal{G}_{Q^2_{\max}} = \left\{\left(G_E^H, G_M^H\right), \quad P\left(\{w_i\}_{\mathrm{MP}}|\mathcal{D}_{Q^2_{\max}},\mathcal{N}_H\right), P\left(\mathcal{D}_{Q^2_{\max}}|\mathcal{N}_H\right)\right\}, \tag{16}$$

where $H = 2,\ldots,40$. This distribution of models is used to calculate the Zemach moments and to discuss their dependence on the choice of the functional form.

In principle, the given fit is valid for Q^2 lower than cutoff $Q^2_{\max}$. Above this limit some functional form factor parametrization should be assumed to calculate the integrals (8) and (10). However, for $Q^2 > Q^2_{\max}$ the dominant contributions to the moments in Eqs. (8) and (10) are given by the last terms of the integrand functions, namely, r_p^2/q^2 and $1/q^2$, respectively. They do not depend on form factor parametrization. We verified that assuming a typical high-Q^2 form factor tail (see the Appendix) leads to systematic uncertainty smaller than $10^{-4}\%$.

In Tables I and II the numerical values of the third and first Zemach moments obtained for the best fit within each selection of the data are presented. The 1σ uncertainty is calculated from the covariance matrices of the fits. To investigate the dependence of the Zemach moment on the choice of the form factor functional form we calculate the average values of I_3 and I_1 over the $\mathcal{G}_{Q^2_{\max}}$:

$$\langle I_i\rangle = \frac{1}{N_{\mathcal{N}}}\sum_{H=2}^{40} I_i(\mathcal{N}_H)\mathcal{P}(\mathcal{N}_H|\mathcal{D}), \quad i = 3,1, \tag{17}$$

$$N_{\mathcal{N}} = \sum_{H=2}^{40}\mathcal{P}(\mathcal{N}_H|\mathcal{D}), \tag{18}$$

where $\mathcal{P}(\mathcal{N}_H|\mathcal{D}) \approx \mathcal{P}(\mathcal{D}|\mathcal{N}_H)$. The averaged values are shown in Table III. The systematic uncertainty due to parametrization choice is given by the square root of the variance due to the evidence probability distribution.

TABLE II. The numerical values (in fm units) of the first Zemach moment (last column) and integral $I_1(c,G_E,G_M)$ calculated for several values of c.

c (GeV)	1.0	3.0	10	∞
Fit $Q^2 < 1$ GeV^2	0.852	0.957	1.022	1.1018 ± 0.0004
Fit $Q^2 < 3$ GeV^2	0.853	0.958	1.024	1.1030 ± 0.0004
Fit $Q^2 < 10$ GeV^2	0.877	0.983	1.048	1.1276 ± 0.0025

Similarly as in the case of the most probable values for Zemach integrals (Tables I and II) the average values weakly depend on the cutoff $Q^2_{\max}$.

Since we do not know the relative normalization between the evidence probability distributions calculated for different subsets of data $\mathcal{D}_{Q^2_{\max}}$ the best fits from these analyses cannot be quantitatively compared but we see that the Zemach moments do not significantly depend on $Q^2_{\max}$. As the final result we take the average of the three cutoff results from Tables I and II, respectively:

$$I_3 = 2.889 \pm 0.008 \ \pm 0.046\,(\mathrm{sys})\ \mathrm{fm}^3 \tag{19}$$

$$I_1 = 1.1108 \pm 0.0021 \pm 0.0071\,(\mathrm{sys})\ \mathrm{fm}. \tag{20}$$

The systematic uncertainty is also the appropriate average.

Our estimate of the third Zemach moment is consistent with the results of other groups, 2.71(13) fm^3 in Ref. [26], 2.5–2.7 fm^3 in Ref. [21], and 2.85 fm^3 in Ref. [22], but is much smaller than the value of 36.6(6.9) fm^3 obtained in Ref. [20]. Our estimate of the first Zemach moment is larger then the previous results: 1.047(16) fm [34], 1.037(16) fm [35], and 1.082(37) fm [36].

Summarizing, we presented the results of the estimate of the Zemach moments obtained from the Bayesian analysis of elastic ep scattering data. Within this approach we conducted

TABLE III. The expected values of the third and first Zemach moments [see Eq. (17)] together with systematic uncertainty due to the choice of parametrization.

	$\langle r^3\rangle_{(2)}$ (fm^3)	$\langle r\rangle_{(2)}$ (fm)
Fit $Q^2 < 1$ GeV^2	2.865 ± 0.003	1.1018 ± 0.0002
Fit $Q^2 < 3$ GeV^2	2.872 ± 0.014	1.1031 ± 0.0018
Fit $Q^2 < 10$ GeV^2	2.865 ± 0.122	1.1153 ± 0.0194

a quantitative discussion of the dependence of the results on the choice of form factor parametrization. The corresponding systematic uncertainties are 1.6% and 0.6% for the third and first Zemach moments, respectively.

ACKNOWLEDGMENTS

The calculations were carried out at the Wrocław Centre for Networking and Supercomputing (http://www.wcss.wroc.pl), and the work was supported by Grant No. 268.

APPENDIX: FORM FACTOR TAIL

The form factors parametrizations discussed for $Q^2 > Q^2_{\max}$ are as follows:

$$G^H_{E,M}(Q^2) = C \begin{cases} \frac{G_D(Q^2)}{G_D(Q^2_{\max})} \\ \left(\frac{Q^2_{\max}}{Q^2}\right)^n \end{cases}, \tag{A1}$$

where $C = G^H_{E,M}(Q^2_{\max})$, $n = 3$, 4, and 5.

[1] M. I. Eides, H. Grotch, and V. A. Shelyuto, Phys. Rep. **342**, 63 (2001).
[2] S. G. Karshenboim, Phys. Rep. **422**, 1 (2005).
[3] F. Ernst, R. Sachs, and K. Wali, Phys. Rev. **119**, 1105 (1960).
[4] A. V. Belitsky, X. Ji, and F. Yuan, Phys. Rev. D **69**, 074014 (2004).
[5] I. Sick, Prog. Part. Nucl. Phys. **67**, 473 (2012).
[6] D. Borisyuk, Nucl. Phys. A **843**, 59 (2010).
[7] J. Bernauer *et al.* (A1 Collaboration), Phys. Rev. Lett. **105**, 242001 (2010).
[8] J. Bernauer *et al.* (A1 Collaboration), Phys. Rev. C **90**, 015206 (2014).
[9] Z. Epstein, G. Paz, and J. Roy, Phys. Rev. D **90**, 074027 (2014).
[10] R. Pohl, A. Antognini, F. Nez, F. D. Amaro, F. Biraben *et al.*, Nature (London) **466**, 213 (2010).
[11] I. Lorenz, Ulf-G. Meissner, H.-W. Hammer, and Y.-B. Dong, Phys. Rev. D **91**, 014023 (2015).
[12] I. Lorenz and Ulf-G. Meissner, Phys. Lett. B **737**, 57 (2014).
[13] J. C. Bernauer, arXiv:1411.3743.
[14] C. E. Carlson, arXiv:1502.05314.
[15] J. L. Friar, Ann. Phys. **122**, 151 (1979).
[16] L. Borisoglebsky and E. Trofimenko, Phys. Lett. B **81**, 175 (1979).
[17] K. Pachucki, Phys. Rev. A **53**, 2092 (1996).
[18] A. P. Martynenko, Phys. Rev. A **71**, 022506 (2005).
[19] A. Zemach, Phys. Rev. **104**, 1771 (1956).
[20] A. De Rujula, Phys. Lett. B **693**, 555 (2010).
[21] I. C. Cloet and G. A. Miller, Phys. Rev. C **83**, 012201 (2011).
[22] M. O. Distler, J. C. Bernauer, and T. Walcher, Phys. Lett. B **696**, 343 (2011).
[23] S. G. Karshenboim, Phys. Rev. D **90**, 053013 (2014).
[24] K. M. Graczyk and C. Juszczak, Phys. Rev. C **90**, 054334 (2014).
[25] R. J. Hill and G. Paz, Phys. Rev. D **82**, 113005 (2010).
[26] J. L. Friar and I. Sick, Phys. Rev. A **72**, 040502 (2005).
[27] K. M. Graczyk, P. Plonski, and R. Sulej, J. High Energy Phys. 09 (2010) 053.
[28] K. M. Graczyk, Phys. Rev. C **84**, 034314 (2011).
[29] K. M. Graczyk and C. Juszczak, J. Phys. G **42**, 034019 (2015).
[30] G. Cybenko, Math. Control Signals Syst. **2**, 303 (1989).
[31] K.-I. Funahashi, Neural Networks **2**, 183 (1989).
[32] K. Hornik, M. Sinchcombe, and W. Halbert, Neural Networks **2**, 359 (1989).
[33] K. M. Graczyk, Phys. Rev. C **88**, 065205 (2013).
[34] A. Volotka, V. Shabaev, G. Plunien, and G. Soff, Eur. Phys. J. D **33**, 23 (2005).
[35] A. Dupays, A. Beswick, B. Lepetit, C. Rizzo, and D. Bakalov, Phys. Rev. A **68**, 052503 (2003).
[36] A. Antognini *et al.*, Science **339**, 417 (2013).