Comparison of neural network and hadronic model predictions of the two-photon exchange effect

Krzysztof M. Graczyk*
Institute of Theoretical Physics, University of Wrocław, pl. M. Borna 9, 50-204 Wrocław, Poland
(Received 27 June 2013; revised manuscript received 12 November 2013; published 26 December 2013)

Predictions for the two-photon exchange (TPE) correction to the unpolarized *ep* elastic cross section, obtained within two different approaches, are confronted and discussed in detail. In the first one the TPE correction is extracted from experimental data by applying the Bayesian neural network statistical framework. In the other the TPE is given by box diagrams, with the nucleon and the P_{33} resonance as the hadronic intermediate states. Two different form factor parametrizations for both the proton and the P_{33} resonance are taken into consideration. Proton form factors are obtained from the global fit of the full model (with the TPE correction) to the unpolarized cross-section data. Predictions of the two methods agree well in the intermediate Q^2 range of 1–3 GeV^2. Above $Q^2 = 3$ GeV^2 the agreement is at the 2σ level. Below $Q^2 = 1$ GeV^2 the consistency between the two approaches is broken. The values of the proton radius extracted within the models are given. In both cases predictions for the VEPP-3 experiment have been obtained and confronted with the preliminary experimental results.

DOI: 10.1103/PhysRevC.88.065205 PACS number(s): 13.40.Gp, 25.30.Bf, 14.20.Dh, 84.35.+i

I. INTRODUCTION

Two-photon exchange (TPE) in elastic electron scattering off the proton drew renewed attention of physicists about ten years ago, when the new experimental technique for measurement of the electromagnetic nucleon form factors (FFs) had become available. In this method, later called the polarization transfer (PT) technique, various polarization observables are measured and the form factor ratio,

$$\mathcal{R}_{1\gamma}(Q^2) = \mu_p \frac{G_E(Q^2)}{G_M(Q^2)}, \tag{1}$$

is estimated [1]. Here G_E and G_M are the electric and the magnetic proton form factors, respectively, and $\mu_p = 2.793$ is the proton magnetic moment in units of the nuclear magneton.

The proton electromagnetic FFs are also extracted from unpolarized cross-section (CS) data by applying Rosenbluth [longitudinal-transverse (LT)] separation. As the result, the electric and the magnetic FFs are obtained simultaneously. It turns out that the ratio (1) estimated based on the Rosenbluth FF data is inconsistent with the PT measurements at larger Q^2 values.

It is generally widely accepted that an insufficient estimate of the radiative corrections (RCs) applied in the Rosenbluth data analysis is the main source of inconsistency. In particular, it is argued that a lack of a so-called hard-photon TPE contribution coming from the box diagrams drawn in Fig. 1 is responsible for the disagreement[1] [2–4].

In the old Rosenbluth data analyses the cross-section measurements were corrected by the RCs obtained by Mo and Tsai (MT) [5]. In this approach the TPE corrections were calculated within a soft-photon approximation. Supplementing this contribution with the hard-photon correction changes the results of the Rosenbluth separation and makes them nearly consistent with the PT measurements [2,7].

Recently, several theoretical calculations of the TPE correction have been performed [2,4,7–18]. They have been done within various approaches. (For a review see Refs. [19–21].) The predictions of the TPE effect turn out to be mostly model dependent at larger values of Q^2.

Simultaneously to the theoretical activity phenomenological investigations have been carried out as well. An effort has been made to extract the proton FFs and the TPE term directly from the experimental data [22–35].

The TPE effect can be studied experimentally. The TPE correction for elastic positron-proton scattering has an opposite sign but the same absolute value as the corresponding term in electron-proton scattering. Hence the measurement of the cross-section ratio,

$$R_{+/-} = \frac{\frac{d\sigma}{d\Omega}(e^+ p \to e^+ p)}{\frac{d\sigma}{d\Omega}(e^- p \to e^- p)} \approx 1 - 2(\mathrm{TPE}), \tag{2}$$

gives a direct way to possibly estimating the TPE correction. At present two dedicated $R_{+/-}$ measurement experiments are operating [36,37]. There is also a proposal of a new project, called OLYMPUS, at DESY [38].

In this report we would like to confront the phenomenological estimation of the TPE effect, obtained in our previous paper [33], with the theoretical predictions. In Ref. [33] a global Bayesian analysis of the world elastic *ep* data was performed. The major idea was to build a statistical model based on the experimental measurements with the ability to make predictions of the electromagnetic proton FFs and the TPE term. This was achieved by adapting a Bayesian framework for feed-forward neural networks (BNNs) [39]. This formalism allows one to perform the analysis as model independent as possible. However, because of the incompleteness of the data some additional assumptions had to be made. We applied constraints coming from C parity and the crossing symmetry invariance of the *ep* scattering amplitude [40–42]. But the most important assumption was that for the PT data the TPE effect can be neglected and is only relevant for the cross-section data. This statement is supported by both general arguments [3] and calculations [7,14]. Indeed, the TPE corrections to

*kgraczyk@ift.uni.wroc.pl
[1]We notice an explanation proposed by Bystritskiy *et al.* [6].

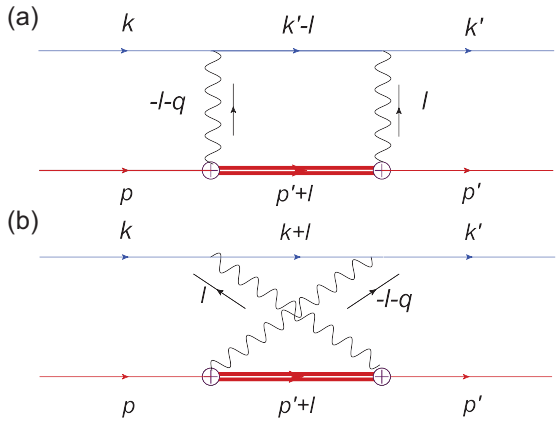

FIG. 1. (Color online) Direct (a) and exchange (b) TPE box diagrams for elastic ep scattering [Eq. (1)]. The intermediate states (thick double lines) are given by either the nucleon or the $P_{33}(1232)$ resonance.

the unpolarized cross-section and the PT ratio $R_{1\gamma}$ data are comparable. But their inclusion in the Rosenbluth analysis affects significantly the results of the FF extraction, while in the case of PT measurements TPE correction is of the order of statistical errors.

The comparison of the BNN and the theoretical predictions allows one to verify the validity of the model assumptions and to confront, in a nondirect way, theoretical model predictions with the data represented by the BNN.

In this paper the TPE corrections are computed in a way similar to that of Refs. [2,7–10], in a quantum-field-theory approach, called later the hadronic model (HM). Hadronic intermediate states in the box diagrams (Fig. 1) are given by the nucleon and the $P_{33}(1232)$ resonance. Heavier resonances are not included because it was shown that their total contribution is negligible for the given kinematics [10] and their inclusion introduces additional model dependence into the discussion.

Our approach should work well at low and intermediate Q^2 range. Its input includes proton and $P_{33}(1232)$ electromagnetic FFs. For the proton we consider two different types of parametrizations. In contrast to the BNN analysis the FF parameters are established from a global fit of the HM to the unpolarized cross-section data only. For the $N \to P_{33}$ electromagnetic transition we consider vertex and FF parametrizations different than those in Ref. [9].

In a wide Q^2 domain the BNN and the hadronic model predictions agree well. The discrepancy appears at low Q^2. However, the value of the proton radius extracted from the BNN fit is consistent with the recent atomic measurement by Pohl *et al.* [43]. It is shown that the low-Q^2 inconsistency between the BNN and the HM is induced by one of the model assumptions mentioned above (with the TPE correction to the PT data neglected).

Eventually, we compare the predictions of the $R_{+/-}$ ratio obtained within the BNN and the HM approaches with the preliminary VEPP-3 measurements [37]. The theoretical and the phenomenological predictions are in agreement with the new available data.

The paper is organized as follows. In Sec. II the basic formalism is introduced. In Sec. III the Bayesian neural network approach is briefly reviewed. The hadronic model is described in Sec. IV. A detailed comparison of the BNN and the theoretical predictions is presented in Sec. V. We summarize our results in Sec. VI. Some technical details of the theoretical calculations are enclosed in Appendix A, while Appendix B contains the definition of the χ^2 function used in the data analysis.

II. BASIC FORMALISM

We consider elastic electron-proton scattering,

$$e(k) + p(p) \to p(p') + e(k'). \tag{3}$$

where k and k' and p and p' denote the initial and final electron and proton four-momenta, respectively. A four-momentum transfer is defined as $q = k - k'$ and $q^2 = (k - k')^2 = -Q^2$.

To compute the TPE correction we apply a typical approach to account for the radiative corrections in ep scattering [2,44], using quantum electrodynamics (QED) extended to include the hadronic degrees of freedom, such as the proton and the $P_{33}(1232)$ resonance. The proton and nucleon electromagnetic vertices are expressed in terms of transition FFs.

The matrix element for ep scattering can be written as a perturbative series in $\alpha = e^2/4\pi \approx 1/137$. The first element of the series, $\mathcal{M}_{1\gamma}$, describes an exchange of one photon between the electron and the proton target, and it gives the lowest order contribution of the differential cross section, $d\sigma_{1\gamma} \sim |\mathcal{M}_{1\gamma}|^2$. The $\mathcal{M}_{1\gamma}$ matrix element is a contraction of the one-body leptonic with hadronic currents,

$$i\mathcal{M}_{1\gamma} = i\frac{e^2}{Q^2} j^\mu h_\mu. \tag{4}$$

The leptonic and hadronic currents read

$$j_\mu(q) = \overline{u}(k')\gamma^\mu u(k), \tag{5}$$

$$h^\mu(q) = \overline{u}(p')\Gamma^\mu(q)u(p). \tag{6}$$

Γ^μ is the on-shell proton electromagnetic vertex, given by

$$\Gamma^\mu(q) = \gamma^\mu F_1(Q^2) + \frac{i\sigma^{\mu\nu}q_\nu}{2M_p} F_2(Q^2). \tag{7}$$

F_1 and F_2 are the Dirac and the spin-flip proton form factors, respectively, while $M_p = 938\ \mathrm{MeV}/c^2$ is the proton mass. It is useful to express the above FFs in terms of the electric and the magnetic proton FFs:

$$F_1(Q^2) = \frac{1}{1+\tau}\{G_E(Q^2) + \tau G_M(Q^2)\}, \tag{8}$$

$$F_2(Q^2) = \frac{1}{1+\tau}\{G_M(Q^2) - G_E(Q^2)\}, \tag{9}$$

where $\tau = Q^2/4M_p^2$.

We keep the normalization $G_E(0) = 1, G_M(0) = \mu_p$; hence $F_1(0) = 1$, $F_2(0) = \kappa_p \equiv \mu_p - 1$.

In ep scattering data analysis it is convenient to consider the reduced cross section, σ_R ($d\sigma/d\Omega \sim \sigma_R$), which in the Born approximation is given by the formula

$$\sigma_{R,1\gamma}(Q^2,\varepsilon) = \tau G_M^2(Q^2) + \varepsilon G_E^2(Q^2), \tag{10}$$

where

$$\varepsilon = \left[1 + 2\left(1 + \frac{Q^2}{4M_p^2}\right)\tan^2\left(\frac{\theta}{2}\right)\right]^{-1} \tag{11}$$

is the photon polarization and θ is the angle between the initial and the final electron momenta.

The next order terms of the cross section are given by the interference between $\mathcal{M}_{1\gamma}$ and the second-order amplitude $\mathcal{M}^{(2)}$. In the complete calculation, in order to remove the infrared (IR) divergences, the inelastic Bremsstrahlung contribution must be also taken into account,

$$d\sigma^{(2)} \sim 2\mathrm{Re}\left[(i\mathcal{M}_{1\gamma})^* i\mathcal{M}^{(2)}\right] + d\sigma^{(1)}_{\mathrm{Brem}}. \tag{12}$$

In this paper we focus on the TPE box diagrams (Fig. 1), which describe an exchange of two photons between the electron and the proton target. The intermediate hadronic state is the off-shell nucleon or a resonance. Because the off-shell electromagnetic form factors are not known [45] we make a common ansatz and consider the on-shell vertices instead.

The leading TPE contribution reads

$$\mathcal{I}_{2\gamma} \equiv 2\mathrm{Re}\{(i\mathcal{M}_{1\gamma})^* i\mathcal{M}_{2\gamma}\}. \tag{13}$$

The box diagrams contributing to $\mathcal{M}_{2\gamma}$ are drawn in Fig. 1.

In the old ep data analysis, to account for higher order radiative corrections, the MT approach [5] was usually applied. In this approach the TPE box contribution was computed in the soft-photon approximation. As was pointed out by Blunden *et al.* [2] to properly correct the cross-section data by using the "full" TPE term, one has to subtract first the MT box contribution. Then the redefined TPE correction reads

$$\Delta_{2\gamma} = \delta_{2\gamma}(\mathrm{full}) - \delta_{2\gamma}(\mathrm{MT}), \tag{14}$$

where

$$\delta_{2\gamma} = \frac{\mathcal{I}_{2\gamma}}{|i\mathcal{M}_{1\gamma}|^2}, \tag{15}$$

and it is given by some integral (see Eq. (23) of Ref. [7]).

The inclusion of the TPE correcting term modifies the form of the reduced cross section to

$$\sigma_{R,1\gamma+2\gamma}(Q^2,\varepsilon) \to \sigma_{R,1\gamma}(Q^2,\varepsilon) + \Delta C_{2\gamma}(Q^2,\varepsilon), \tag{16}$$

where $\Delta C_{2\gamma} = \Delta_{\mathrm{TPE}} \cdot \sigma_{R,1\gamma}$.

III. NEURAL NETWORK APPROACH

Artificial neural networks (ANNs) have been used in particle and nuclear physics for many years. ANNs are perfectly dedicated to particle or interaction identification and have been applied in experimental data analyses [46]. Study of the properties of ANNs is an interesting topic by itself. Neural networks have also been investigated within the methods of statistical physics [47].

The feed-forward neural network is a type of ANN and can be applied to interpolating data, parameter estimation, and function approximation problems.

The ANN methodology can be a powerful approach for approximating the physical observables based on measurements if it is difficult to make predictions (based on the theoretical model) of the analyzed quantities or if the theoretical predictions are model dependent but there exist informative experimental data. Then one can construct a model-independent representation (given by a neural network) of the physical observables favored by the measurements. As an example let us mention the parton distributions functions (PDFs), which are parametrized by feed-forward neural networks [48].

As in the case of PDFs, computing the nucleon FFs and the TPE correction from first principles is a difficult task. On the other hand, unpolarized cross-section ratio, $R_{+/-}$, as well as PT ratio data are distributed over a wide kinematical range. Global analyses of these measurements provide reliable information about the FFs and the TPE [27,33].

The BNN approach was adapted by us [33,39] to approximate the nucleon FFs and TPE correction. In the next four sections a short review of the main features of this approach is presented.

A. Multilayer perceptron

The FFs and TPE correction are going to be approximated by using feed-forward neural networks in a multilayer perceptron (MLP) configuration. From the mathematical point of view the MLP, denoted as $\mathcal{N}$, is a nonlinear function, which maps a subset of $\mathbb{R}^{n_i}$ (an input space) into $\mathbb{R}^{n_o}$ (an output space), where $n_i, n_o \in \mathbb{N}$. The given MLP consists of several layers of units, namely, input, hidden, and output layers (see Figs. 2 and 3).

A unit [Fig. 2(a)] is a single-valued real function called an activation function ($f_{\mathrm{act.}}$). For the argument it takes the

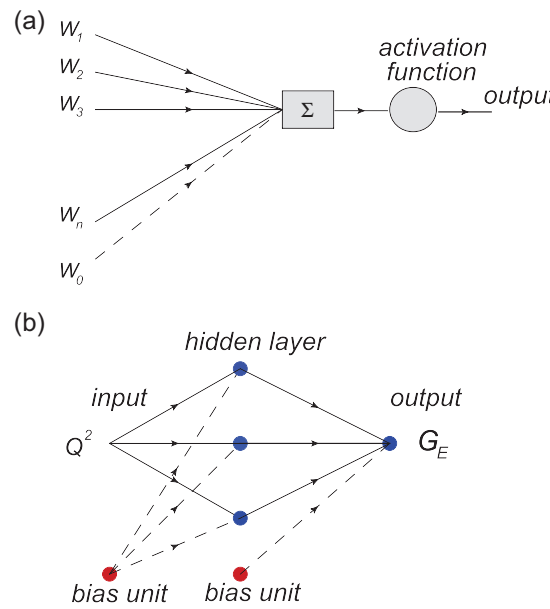

FIG. 2. (Color online) Top: Single unit connected with n units from the previous layer and with one bias unit (ω_0 weight). Bottom: Simple MLP with one hidden layer, used to fit the electric proton form factor data.

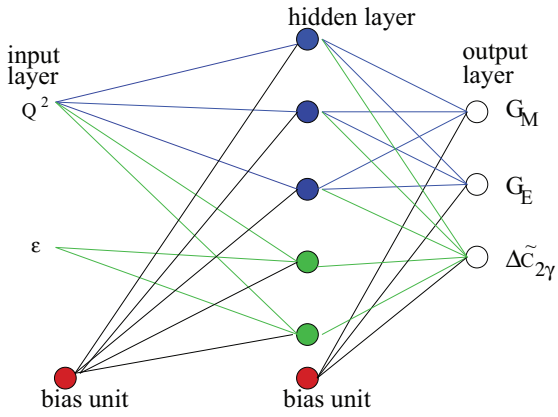

FIG. 3. (Color online) Network $\mathcal{N}_{3,2}$, comprising two input units, one layer of hidden units, and three output units. The FF sector is shown by blue filled units and connections; the TPE sector is shown by green units and connections. Each line corresponds to one weight parameter.

weighted sum of outputs from the the previous layer units,

$$f_{\text{act.}}\left(\sum_{i=0}^{n} w_i\, f^i_{\text{act.}}(\text{previous layer})\right), \tag{17}$$

where w_i $(i = 1,2,\ldots,n)$ is the weight parameter.

An example of a typical unit is drawn in Fig. 2(b). The weight parameters are established during the training, i.e., a process of finding the optimal weight configuration. In reality the optimal weight configuration minimizes some error function.

In our analysis the sigmoid

$$f_{\text{act}}(x) = \frac{1}{1+\exp(-x)} \tag{18}$$

is taken for the activation functions in the hidden layer. But in the case of the output units we consider linear activation functions. The above choice is motivated by the Cybenko theorem [49], which states that it is enough to consider the MLP with one hidden layer and sigmoid-like functions there as well as linear activation functions in the output layer to approximate any continuous function.[2]

Notice that the effective support of (18) is limited. It is a useful feature in the case of numerical analysis (in which the weights are randomly initialized at the beginning of every training).

For a more detailed description of MLP properties, the training process, learning algorithms, etc. see Sec. 2 of Ref. [39].

B. Overfitting problem

A simple example of a one-hidden layer MLP configuration, used to approximate the electric proton FF [39], is shown in Fig. 2(b). The input and output are the one-dimensional vectors (Q^2) and (G_E), respectively. In this case the error function is postulated to be

$$S_{ex}(\mathcal{D},\vec{w}) = \chi^2 = \sum_{i=1}^{N}\left(\frac{G_E(Q_i^2;\vec{w}) - G^{\text{exp.}}_{E,i}}{\Delta G^{\text{exp.}}_{E,i}}\right)^2, \tag{19}$$

where N denotes the number of experimental points, and $(G^{\text{exp.}}_{E,i}, \Delta G^{\text{exp.}}_{E,i})$ is the ith experimental point (its best value and error). $\mathcal{D}$ denotes the experimental data set (or sets).

It is obvious that increasing the number of units (degrees of freedom) improves the ability of the network for representing the data. An MLP with a large enough number of weight parameters may fit the data exactly but in this case the statistical fluctuation of the measurements is reproduced. Such a model has no predictive power and adding new data to the fit spoils its quality. This kind of network overfits the data (or it is said that the network is overlearned). Such a fit is characterized by unrealistic prediction of the uncertainties (see the discussion in Sec. 2.1 and Figs. 3 and 4 of Ref. [39]).

One of the methods for facing the overfitting problem and finding the optimal network configuration is to implement Occam's razor principle. Then in a natural way simpler network configurations are preferred. The simplest idea is to consider a penalty term

$$\alpha E_w, \quad \text{where} \quad E_w = \frac{1}{2}\sum_{i=1}^{W} w_i^2. \tag{20}$$

Including in the error function the above expression may prevent one from getting absolute values of the weight parameters that are too large and as a result overlearned networks.[3]

The parameter α in Eq. (19) is introduced to regularize the penalty term. In general one would consider several α parameters (one each for every distinct class of weights); see, e.g., Sec. 3.2 of Ref. [39] or Chap. 9 of Ref. [50].

The major difficulty is in finding an optimal value of the α parameter. The Bayesian framework offers a mathematically consistent method for getting such α's. Indeed in this approach the penalty term has a natural probabilistic interpretation and α is computed within the objective Bayesian algorithm.

C. Bayesian neural networks

The Bayesian framework for MLPs [50,51] was developed to provide consistent and objective methods, which allow one to

(i) establish the optimal structure of the MLP (number of the hidden units and layers);
(ii) find optimal values of the weights and the α parameters;
(iii) establish optimal values of the learning algorithm parameters;
(iv) compute the neural network output uncertainty and uncertainties for the weight and α parameters; and

[2]According to the Cybenko theorem the discontinuous functions can be approximated well by an MLP with two hidden layers of units.

[3]Usually the MLP that overfits the data contains at least one weight parameter of large absolute value.

(v) classify and compare models quantitatively.

The BNN approach requires minimal input from the user. Indeed, the idea of the approach was to replace the user's common sense by mathematical objective procedures [50]. Obviously, some user input is necessary.

Below we briefly review the BNN approach. For a more detailed description of the BNN see Refs. [39] (Sec. III) and [33] (Sec. III) as well as [50,52].

1. Model comparison

Let us consider a set $\mathcal{S}$, which contains MLPs with different numbers of hidden units. Without lose of generality of the approach we can restrict the set S to MLPs with only one hidden layer (the choice supported by the Cybenko theorem). Each network $\mathcal{N}_\beta \in \mathcal{S}$ $(\beta = 1,2,\dots)$ approximates some physical quantities based on the data $\mathcal{D}$. The models (networks) can be classified by a conditional probability

$$P(\mathcal{N}_\beta|\mathcal{D}). \tag{21}$$

The BNN approach gives a recipe for how to construct and compute the above function.

Beyes' theorem connects the probability (21) with the so-called evidence

$$P(\mathcal{N}_\beta|\mathcal{D}) = \frac{P(\mathcal{D}|\mathcal{N}_\beta)P(\mathcal{N}_\beta)}{P(\mathcal{D})}, \tag{22}$$

where $P(\mathcal{D})$ is the normalization factor, which does not depend on the model $\mathcal{N}_\beta$, and $P(\mathcal{N}_\beta)$ is the prior probability. However, at the beginning of any analysis there is no reason to prefer a particular model (network); therefore it is natural to assume that

$$P(\mathcal{N}_1) = P(\mathcal{N}_2) = P(\mathcal{N}_3) = \cdots. \tag{23}$$

Hence the evidence differs from $P(\mathcal{N}_\beta|\mathcal{D})$ by only a constant normalization factor and it can be used to qualitatively classify the statistical hypotheses.

We apply the so-called hierarchical approach [52] to construct and then to compute the evidence. The procedure entails several steps, as described in the next four sections.

2. First step

In the first step of the approximation the posterior probability distribution $P(\vec{w}|\mathcal{D},\alpha,\mathcal{N}_\beta)$ is computed. According to Bayes theorem it reads

$$P(\vec{w}|\mathcal{D},\alpha,\{\mathcal{I}_{\rm Phys}\},\mathcal{N}_\beta) = \frac{\mathcal{P}(\mathcal{D}|\vec{w},\alpha,\{\mathcal{I}_{\rm Phys}\},\mathcal{N}_\beta)\mathcal{P}(\vec{w}|\alpha,\{\mathcal{I}_{\rm Phys}\},\mathcal{N}_\beta)}{\mathcal{P}(\mathcal{D}|\alpha,\{\mathcal{I}_{\rm Phys}\},\mathcal{N}_\beta)}, \tag{24}$$

where $\mathcal{P}(\vec{w}|\alpha,\{I\}_{\rm Phys},\mathcal{N}_\beta)$ is the prior probability, and $\{\mathcal{I}\}_{\rm Phys}$ denotes the set of initial physical assumptions.

The likelihood function $\mathcal{P}\left(\mathcal{D}|\ \vec{w},\alpha,\{\mathcal{I}_{\rm Phys}\},\mathcal{N}_\beta\right)$ does not depend on α but in our analysis it is modified due to the physical constraints $\{\mathcal{I}\}_{\rm Phys}$:

$$\mathcal{P}(\mathcal{D}|\vec{w},\mathcal{N}_\beta,\{\mathcal{I}\}_{\rm Phys}) = \frac{1}{n_\beta}\exp[-S_{ex}(\mathcal{D},\vec{w}) - S_{\rm Phys}(\{\mathcal{I}_{\rm Phys}\},\vec{w})]. \tag{25}$$

The normalization factor n_β is computed in the Hessian approximation (see Eq. (3.8) of Ref. [39]).

The functions $S_{ex}(\mathcal{D},\vec{w})$ and $S_{\rm Phys}(\{\mathcal{I}_{\rm Phys}\},\vec{w})$ are given by some χ^2 distributions. $S_{\rm Phys}(\{\mathcal{I}_{\rm Phys}\},\vec{w})$ is introduced to force the MLP to properly reproduce the form factors at $Q^2 = 0$. (see Sec. III of Ref. [33]). It may also account for other physical constraints.

The prior $\mathcal{P}(\vec{w}|\mathcal{N}_\beta)$ describes only the initial ANN assumptions about the weights. A reasonable approximation is to assume that it is given by the normal distribution, centered at $\vec{w}_0 = \vec{0}$,

$$\mathcal{P}(\vec{w}|\alpha,\mathcal{N}_\beta) = \frac{1}{n_a}\exp[-\alpha E_w], \tag{26}$$

$$n_a = \int d^{W_\beta} w \exp[-\alpha E_w]. \tag{27}$$

The optimal configuration of weights $\vec{w}_{\rm MP}$ maximizes the posterior probability (24). In reality it minimizes the following error function:

$$S_{ex}(\mathcal{D},\vec{w}) + S_{\rm Phys}(\{\mathcal{I}_{\rm Phys}\},\vec{w}) + \alpha E_w. \tag{28}$$

Notice that in this step of the approximation the α parameter is assumed to be known.

The 1σ error of any physical observable $\mathcal{O}$, which depends on the network response, is the square root of the variance,

$$(\langle\Delta\mathcal{O}\rangle)^2 = \langle\mathcal{O}^2\rangle - \langle\mathcal{O}\rangle^2,$$

$$\langle\mathcal{O}\rangle = \int d^{W_\beta} w\, \mathcal{O}(\mathcal{N}_\beta)P(\vec{w}|\mathcal{D},\alpha,\{\mathcal{I}_{\rm Phys}\},\mathcal{N}_\beta). \tag{29}$$

The above integral is computed in the Hessian approximation.

3. Second step

The optimal value of the α parameter ($\alpha_{\rm MP}$) maximizes the posterior probability,

$$\mathcal{P}(\alpha|\mathcal{D},\{\mathcal{I}_{\rm Phys}\},\mathcal{N}_\beta) = \frac{\mathcal{P}(\mathcal{D}|\alpha,\{\mathcal{I}_{\rm Phys}\},\mathcal{N}_\beta)\mathcal{P}(\alpha|\{\mathcal{I}_{\rm Phys}\},\mathcal{N}_\beta)}{\mathcal{P}(\mathcal{D}|\{\mathcal{I}_{\rm Phys}\},\mathcal{N}_\beta)} \tag{30}$$

(where the denominator of the above expression is obtained in the previous step of the approximation).

The necessary condition which must by satisfied by $\alpha_{\rm MP}$ reads

$$\frac{\partial}{\partial\alpha}\mathcal{P}(\alpha|\mathcal{D},\mathcal{N}_\beta)\Big|_{\alpha=\alpha_{\rm MP}} = 0. \tag{31}$$

It can be shown that in the Hessian approximation the above equation can be written as

$$2\alpha_{\rm MP}E_w(\vec{w}_{\rm MP}) = \sum_{i=1}^{W_\beta}\frac{\lambda_i}{\lambda_i + \alpha_{\rm MP}} \equiv \gamma(\alpha_{\rm MP}), \tag{32}$$

where λ_i are the eigenvalues of $H_{kj} = \nabla_k\nabla_j(S_{ex} + S_{\rm Phys})$, with $k,j = 1,\dots,W_\beta$ and $\nabla_k \equiv \partial/\partial w_k$.

In practice, the λ_i's depend on α. Hence to get the optimal $\alpha_{\rm MP}$, the value of α is iteratively changed during the training, $\alpha_{k+1} = \gamma(\alpha_k)/2E_w$. This means that the optimal weights and α parameter are established during the same training process.

The initial value of the α parameter is taken to be large, which corresponds to the prior assumption that at the beginning of the analysis almost all relevant values of weights are probable.

4. Third step

In this step the evidence is computed. Notice that it is the denominator of the right-hand side of Eq. (30). Careful calculations lead to the following expression for the natural logarithm of the evidence (see Sec. 3.1 of Ref. [39]):

$$\ln \mathcal{P}(\mathcal{D}|\{\mathcal{I}_{\rm Phys}\},\mathcal{N}_\beta) = -S_{ex}(\mathcal{D},\vec{w}_{\rm MP}) - S_{\rm Phys}(\{\mathcal{I}_{\rm Phys}\},\vec{w}_{\rm MP}) \tag{33}$$

$$-\alpha_{\rm MP}E_w(\vec{w}_{\rm MP}) - \frac{1}{2}\ln\det A + \frac{W}{2}\ln\alpha_{\rm MP} - \frac{1}{2}\ln\frac{\gamma}{2} \tag{34}$$

$$+(g+t)\ln(2) + \ln(g!) + \ln(t!), \tag{35}$$

$A = H(\vec{w}_{\rm MP}) + \alpha_{\rm MP} I$.

Expression (33) is the misfit of the approximated data. It is usually of low value. Terms (34) and (35) contribute to Occam's factor. Indeed, (34) takes large values for the models, which overfit the data. In a typical MLP some hidden units, in the given layer, can be reordered without affecting the values of the network output. This means that for every MLP there exist several equivalent indistinguishable network configurations. This gives rise to an additional normalization factor (35), which must be included to properly define the evidence. The symmetry factor presented above concerns the MLP used in the extraction of the TPE correction from the data [see Sec. III D and Eq. (39)].

5. General scheme

Schematically, the approach discussed above can be summarized as follows:

$$\text{Step 1:} \rightarrow P(\vec{w}|\mathcal{D},\alpha,\{\mathcal{I}_{\rm Phys}\},\mathcal{N}_\beta), \tag{36}$$

$$\text{Step 2:} \rightarrow \mathcal{P}(\alpha|\mathcal{D},\{\mathcal{I}_{\rm Phys}\},\mathcal{N}_\beta), \tag{37}$$

$$\text{Step 3:} \rightarrow \mathcal{P}(\mathcal{D}|\{\mathcal{I}_{\rm Phys}\},\mathcal{N}_\beta). \tag{38}$$

We see that the evidence and the other posterior probabilities may depend on physical assumptions. Obviously, their impact on the final results must be carefully discussed.

D. Extraction of the TPE correction

The formalism discussed above was applied to extract the proton FFs and TPE correction [33] from the world elastic e^-p and e^+p scattering data. We utilized the unpolarized cross-section ratio, $R_{+/-}$, and PT data. The first two types of observables depend on two input variables Q^2 and ϵ, while the last one depends only on Q^2.

On the other hand the TPE correction is a function of two input variables, but the FFs depend on Q^2 only. This property was encoded in the network configuration by dividing the MLP into two sectors (see Fig. 3). In the first there are g units connected with only the Q^2 input, while in the other there are t units connected with both input units. We denote this network as

$$\mathcal{N}_{g,t}\left(\begin{pmatrix}Q^2\\ \varepsilon\end{pmatrix};\vec{w}\right) = \begin{pmatrix} G_E^{\mathcal{N}} \\ G_M^{\mathcal{N}} \\ \Delta\tilde{C}_{2\gamma}^{\mathcal{N}}\end{pmatrix}. \tag{39}$$

The BNN formalism seems to be well suited for performing a model-independent analysis but because the utilized data turned out to be not informative enough some model assumptions had to be made.

The main constraint was induced by the following assumption:

(A) The PT data are less sensitive to the TPE correction than the cross-section measurements [3]; hence the TPE contribution to $\mathcal{R}_{1\gamma}$ can be neglected.

As a consequence the TPE correcting term was considered only in the case of the unpolarized cross-section and $R_{+/-}$ data. Its extraction was induced by the presence of the PT measurements in the fit. Certainly, it is an approximation; therefore we distinguish between $\Delta C_{2\gamma}$ as is defined by theory and $\Delta\tilde{C}_{2\gamma}$ as given by the BNN analysis. Both quantities enter the reduced cross-section formula in the same way [see Eq. (16)], but the latter is needed to get a consistent fit of the CS, $R_{+/-}$, and PT.

To get the FFs properly behaved at $Q^2 = 0$ and the TPE term at $\epsilon = 0$ (as suggested by C invariance [40–42]), we introduced $S_{\rm Phys}$ [Eq. (28)]. This is a χ^2 function containing three artificial points (for details see Sec. III of Ref. [33]).

In order to find the optimal MLP configuration 45 different configurations[4] of MLPs were trained. The largest evidence was obtained for the model $\mathcal{N}_{5,6}$.

In general, the optimal fit should be given by an average (weighted by evidence) over all hypothetical models. In this case the physical observable $\mathcal{F}$, which is a function of the FFs and the TPE, reads

$$\langle\mathcal{F}(G_E,G_M,\Delta\tilde{C}_{2\gamma})\rangle = \int_{\mathcal{S}} D\mathcal{N}\,\mathcal{F}\big(G_E^{\mathcal{N}},G_M^{\mathcal{N}},\Delta\tilde{C}_{2\gamma}^{\mathcal{N}}\big)\mathcal{P}(\mathcal{N}|\mathcal{D}). \tag{40}$$

In reality, the above integral can be written as a discrete series,

$$\langle\mathcal{F}(G_E,G_M,\Delta\tilde{C}_{2\gamma})\rangle = \sum_{m=1}^{M}\sum_{g=1,t=1}^{g+t=m}\mathcal{F}\big(G_E^{\mathcal{N}_{g,t}},G_M^{\mathcal{N}_{g,t}},\Delta\tilde{C}_{2\gamma}^{\mathcal{N}_{g,t}}\big)\mathcal{P}(\mathcal{N}_{g,t}|\mathcal{D}), \tag{41}$$

where $M \in \mathbb{N}$.

It turned out that the evidence for the $\mathcal{N}_{5,6}$ model was much larger than for the other analyzed configurations of networks. Hence expression (41) contains only one dominant term,

$$\langle\mathcal{F}(G_E,G_M,\Delta\tilde{C}_{2\gamma})\rangle \approx \mathcal{F}\big(G_E^{\mathcal{N}_{5,6}},G_M^{\mathcal{N}_{5,6}},\Delta\tilde{C}_{2\gamma}^{\mathcal{N}_{5,6}}\big). \tag{42}$$

[4]The number of units in the hidden layer (g and t) was varied.

IV. HADRONIC CALCULATIONS

A. Box diagrams

The TPE correction is computed in a way similar to that in Refs. [2,7–9,16,44]. Four box diagrams contribute to the 2γ amplitude (see Fig. 1): two with the nucleon intermediate hadronic state [denoted as $\Box(N)$] and two with the $P_{33}(1232)$ hadronic intermediate state [denoted as $\Box(P_{33})$]. The TPE contribution (13) reads

$$\mathcal{I}_{2\gamma} = 2\frac{e^2}{Q^2}\mathrm{Im}\{w_N^{\parallel} + w_N^{\times} + w_\Delta^{\parallel} + w_\Delta^{\times}\}, \tag{43}$$

where $w_{N,\Delta}^{\parallel}$ and $w_{N,\Delta}^{\times}$ are the one-loop integrals represented by direct and exchange $\Box(N)$ and $\Box(P_{33})$ diagrams, respectively:

$$w_{N,\Delta}^{\parallel} = e^4\int\frac{d^4l}{(2\pi)^4}\frac{L_{\parallel}^{\alpha\mu\nu}\mathcal{H}_{\alpha\mu\nu}^{N,\Delta}}{D(-k')}, \tag{44}$$

$$w_{N,\Delta}^{\times} = e^4\int\frac{d^4l}{(2\pi)^4}\frac{L_{\times}^{\alpha\mu\nu}\mathcal{H}_{\alpha\mu\nu}^{N,\Delta}}{D(k)}, \tag{45}$$

where

$$D(x) = [(q+l)^2 + i\epsilon][l^2 + i\epsilon][(l+x)^2 - m^2 + i\epsilon] \times \left[(p'+l)^2 - M_{p,\Delta}^2 + i\epsilon\right]. \tag{46}$$

We keep a nonzero electron mass $m = 0.510$ MeV/c^2. $M_\Delta = 1232$ MeV/c^2 denotes the P_{33} resonance mass. The numerators of the integrals in Eqs. (44) and (45) are given by the contraction of three-dimensional leptonic and hadronic tensors.

The leptonic tensor is defined as follows:

$$L_{\parallel,\times}^{\alpha\mu\nu} \equiv \sum_{\text{spin}} j^{\alpha *} j_{\parallel,\times}^{\mu\nu}, \tag{47}$$

where

$$j_{\parallel}^{\mu\nu} = \overline{u}(k')\gamma^\mu(\hat{k}' - \hat{l} + m)\gamma^\nu u(k), \tag{48}$$

$$j_{\times}^{\mu\nu} = \overline{u}(k')\gamma^\mu(\hat{k} + \hat{l} + m)\gamma^\nu u(k), \tag{49}$$

$$\hat{x} = x_\mu\gamma^\mu. \tag{50}$$

Four-vector j^α is given by Eq. (5).

We distinguish two types of hadronic tensors, one for the nucleon and another for the P_{33} intermediate state:

$$\mathcal{H}_{\alpha\mu\nu}^{N,\Delta} \equiv \sum_{\text{spin}} h_\alpha^* h_{\mu\nu}^{N,\Delta}, \tag{51}$$

where h_α is given by Eq. (6) and

$$h_{\mu\nu}^{N} = \overline{u}(p')\Gamma_\nu(-l)(\hat{p}' + \hat{l} + M_p)\Gamma_\mu(q+l)u(p). \tag{52}$$

The proton electromagnetic vertex Γ_μ is defined by Eq. (7). The hadronic tensor for the $\Box(P_{33})$ diagrams has the form

$$h_{\mu\nu}^{\Delta} = \overline{u}(p')\Gamma_{\mu\xi}^{\Delta,\mathrm{in}}(-l, p'+l)(\hat{p}' + \hat{l} + M_\Delta) \times \Lambda^{\xi\eta}(p'+l)\Gamma_{\nu\eta}^{\Delta,\mathrm{out}}(q+l, p'+l)u(p). \tag{53}$$

$\Gamma_{\nu\mu}^{\Delta,\mathrm{out}}(q_\Delta, P)$ and $\Gamma_{\nu\mu}^{\Delta,\mathrm{in}}(q_\Delta, P)$ denote the vertex for the $\gamma^* N \to \Delta$ and $\Delta \to N\gamma^*$ transitions. For more detailed definitions see Sec. IV C.

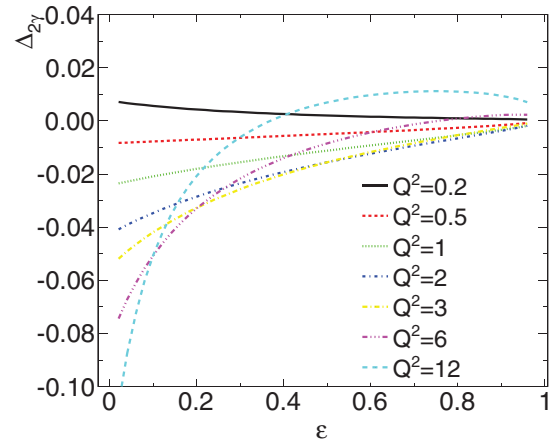

FIG. 4. (Color online) $\Delta_{2\gamma}$ [Eq. (14)] computed for the $\Box(N)$ TPE contribution. The form factors from Ref. [7] are used. The values of Q^2 are in units of GeV^2.

For the Rarita-Schwinger 3/2 spin field propagator we take

$$S_{\mu\nu}^{\Delta} = -\frac{i(\hat{p} + M_\Delta)}{p^2 - M_\Delta^2 + i\Gamma_\Delta M_\Delta}\Lambda_{\mu\nu}(p). \tag{54}$$

Similarly as in Kondratyuk *et al.* [9] we set $\Gamma_\Delta \to 0$.[5] With this simplification the dominant contribution to the loop integrals comes from the P_{33} resonance mass pole. Hence the choice of on-shell projection operator

$$\Lambda_{\mu\nu}(p) = g_{\mu\nu} - \frac{1}{3}\gamma_\mu\gamma_\nu - \frac{2p_\mu p_\nu}{3M_\Delta^2} + \frac{p_\mu\gamma_\nu - p_\nu\gamma_\mu}{3M_\Delta} \tag{55}$$

leads to the same results as the off-shell projection operator discussed by Kondratyuk *et al.* Taking into consideration this approximation for the projector operator simplifies and also accelerates the algebraic decomposition of the integrals in Eqs. (44) and (45). The procedure for computing the integrals in Eqs. (44) and (45) is described in Appendix A.

B. Nucleon form factors

For the nucleon FFs we consider two parametrizations:

(i) parametrization I, the sum of three monopoles,

$$F_k(Q^2) = \sum_{i=1}^{3}\frac{f_i^k}{m_i^k + Q^2}, \tag{56}$$

where $f_3^1 = m_3^1(1 - f_1^1/m_1^1 - f_2^1/m_2^1)$ and $f_3^2 = m_3^2(\kappa_p - f_1^2/m_1^2 - f_2^2/m_2^2)$;

(ii) parametrization II, the sum of three dipoles,

$$F_k(Q^2) = \sum_{i=1}^{3} f_i^k\left(1 + \frac{Q^2}{\left(m_i^k\right)^2}\right)^{-2}, \tag{57}$$

where $f_3^1 = 1 - f_1^1 - f_2^1$ and $f_3^2 = \kappa_p - f_1^2 - f_2^2$.

[5] Recently, Borisyuk and Kobushkin [18] performed calculations in which the impact of the nonzero value of Γ_Δ is discussed.

Parametrization I was previously discussed by Blunden *et al*. [7] (BMT05). In order to cross-check our algebraic and numerical procedures we repeat and check the calculations done in Ref. [7]. In Fig. 4 we present predictions for $\Delta_{2\gamma}$ obtained for the same kinematics and form factors as in BMT05 (for comparison see the plots in Figs. 2 and 3(a) of Ref. [7]). We notice the excellent agreement between our hadronic model and the BMT05 predictions.

The $\Delta_{2\gamma}$ (or $D_{2\gamma}$) function depends weakly on the proton form factor parametrization. Small differences between TPE predictions based on parametrizations I and II appear for larger values of Q^2 (see Figs. 5 and 6). But this is the region where the validity of the theoretical approach can be questionable.

C. $P_{33}(1232)$ form factors

The hadronic vertex $\Gamma^{\Delta,\text{out}}_{\mu\nu}(q,P_\Delta)$ for the $\gamma^* p \to \Delta^{++}$ transition is obtained by assuming that the P_{33} resonance is described by the Rarita-Schwinger 3/2 spin field,

$$\overline{\Psi}^{\nu}(P)\Gamma^{\Delta,\text{out}}_{\mu\nu}(q,P)u(p), \quad q = P - p. \tag{58}$$

One of the commonly discussed vertex parametrizations is the following [55]:

$$\Gamma^{\Delta,\text{out}}_{\mu\nu}(q \equiv P - q, P) = \left[\frac{C_5^V}{M^2}(g_{\mu\nu} p\cdot q - p_\mu q_\nu) + \frac{C_4^V}{M^2}(g_{\mu\nu} q\cdot P - P_\mu q_\nu) + \frac{C_3^V}{M}(g_{\mu\nu}\hat{q} - \gamma_\mu q_\nu)\right]\gamma_5, \tag{59}$$

where P is the four-momentum of the outgoing P_{33} resonance, while p denotes the four-momentum of the incoming proton.

The $\Delta^{++} \to \gamma^* p$ vertex reads [9]

$$\Gamma^{\Delta,\text{in}}_{\mu\nu}(p,P_\Delta) = \gamma_0\left(\Gamma^{\Delta,\text{out}}_{\mu\nu}(p,P_\Delta)\right)^\dagger\gamma_0. \tag{60}$$

For the $N \to P_{33}$ transition form factors we consider two scenarios:

(i) P_{33}[SU(6)] model: There is only one vector form factor C_3^V; two others are obtained by assuming SU(6) quark model relations [56], namely,

$$C_5^V(Q^2) = 0, \quad C_4^V(Q^2) = -\frac{M}{M_\Delta}C_3^V(Q^2). \tag{61}$$

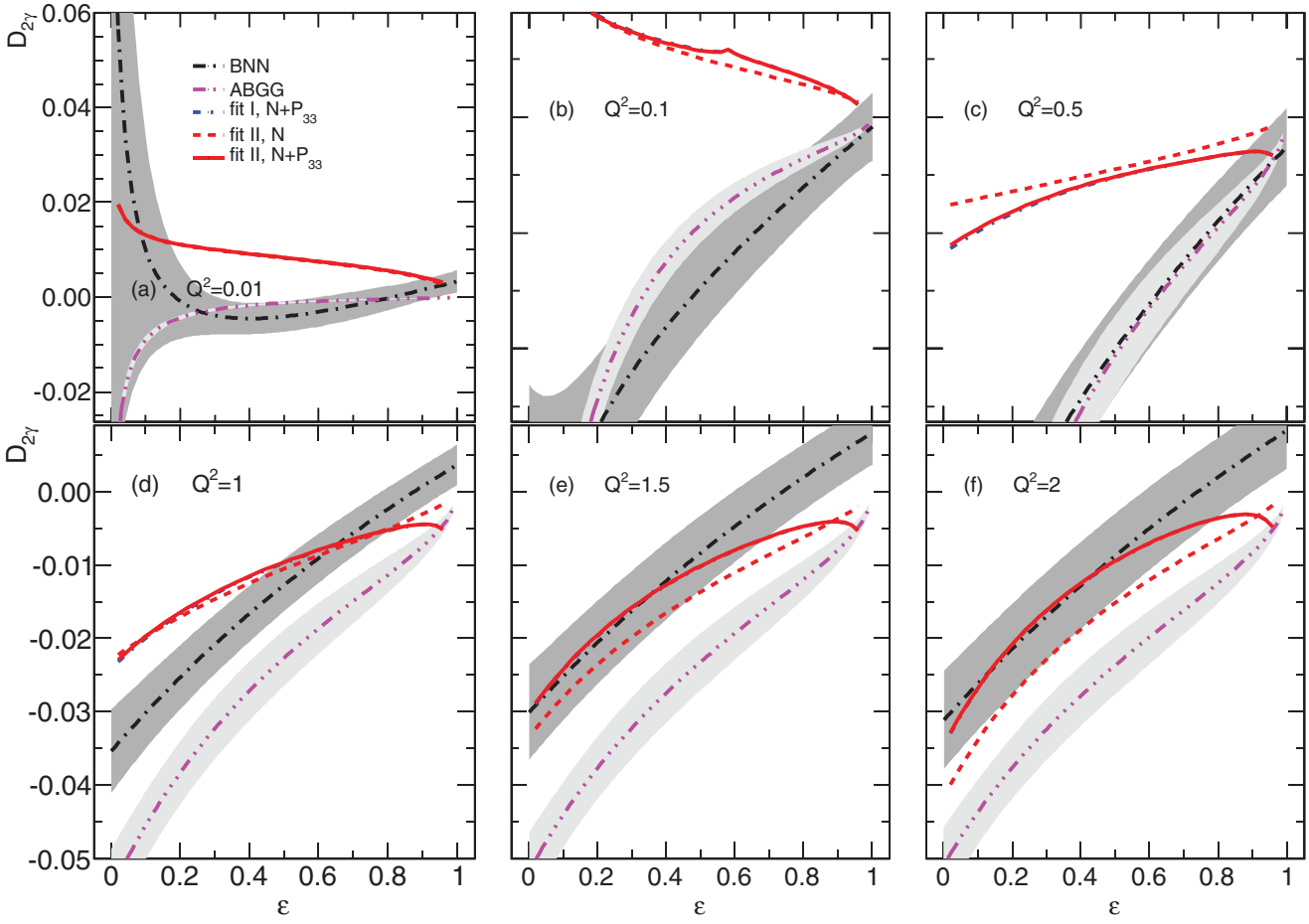

FIG. 5. (Color online) Predictions of $D_{2\gamma}$ [Eq. (65)] based on the BNN and HM (fits I and II) as well as ABGG approaches for $Q^2 \leq 2$. The TPE correction includes either elastic (N) or elastic and P_{33} resonance (full-model) contributions. The values of Q^2 are in units of GeV^2. The shaded areas show 1σ error computed from the covariance matrix.

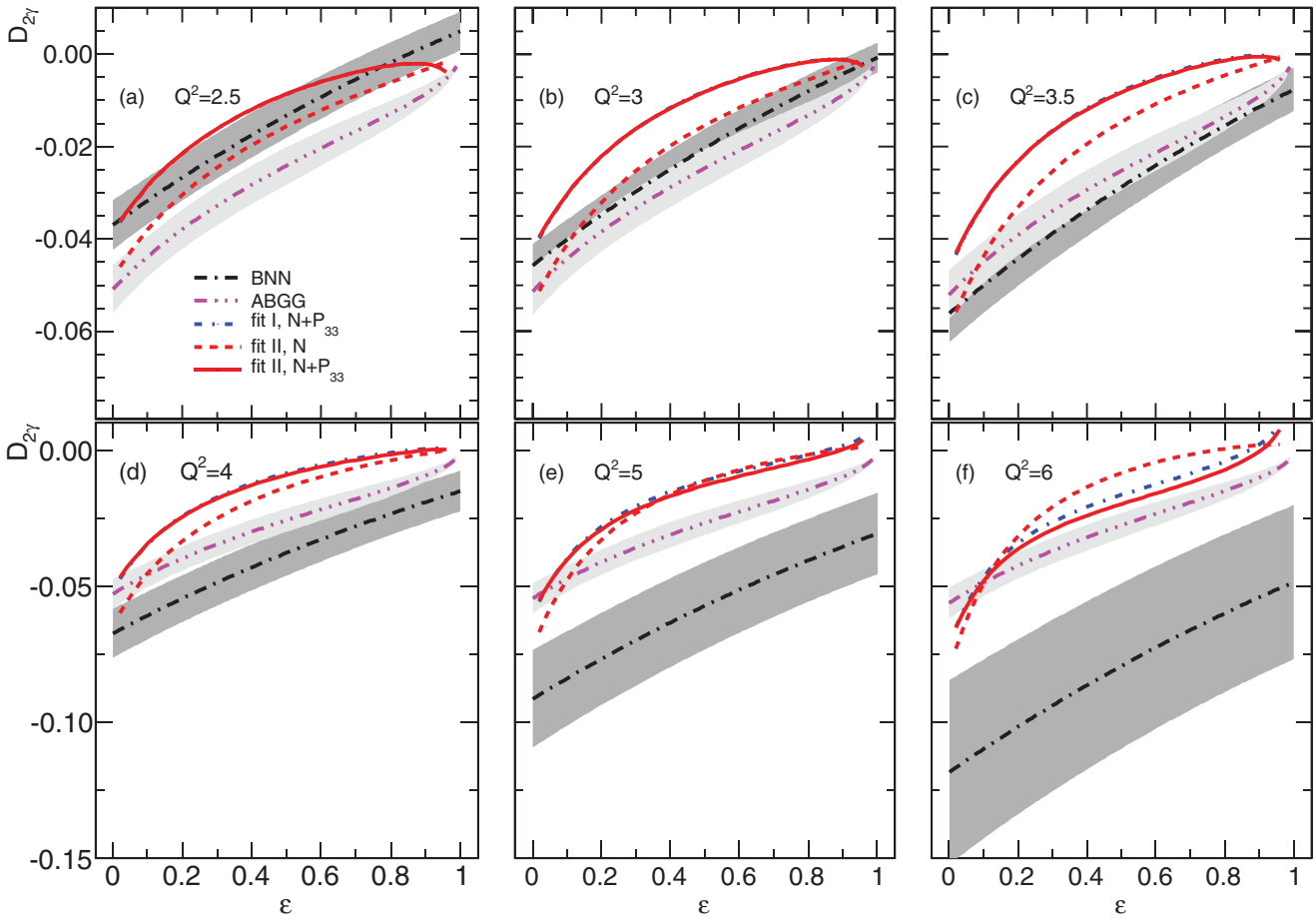

FIG. 6. (Color online) Predictions of $D_{2\gamma}$ [Eq. (65)] based on the BNN and HM (fits I and II) as well as ABGG approaches for $Q^2 > 2$. The TPE correction includes either elastic (N) or elastic and P_{33} resonance (full-model) contributions. The values of Q^2 are in units of GeV^2. The shaded areas show 1σ error computed from the covariance matrix.

In this case we parametrize the $C_3^V(Q^2)$ form factor as follows [57]:

$$C_3^V(Q^2) = \frac{2.05}{\left(1 + \frac{Q^2}{0.54\,\mathrm{GeV}^2}\right)^2}. \tag{62}$$

(ii) P_{33}(full) model: We apply the form factors from Ref. [58], namely,

$$C_i^V(Q^2) = c_i^V \left(1 + \frac{Q^2}{a_i M_V^2}\right)^{-1} G_D(Q^2), \tag{63}$$

where $a_3 = a_4 = 4$, $a_5 = 0.776$, $c_3^V = 2.13$, $c_4^V = -1.51$, $c_5^V = 0.48$, and

$$G_D(Q^2) = \left(1 + \frac{Q^2}{M_V^2}\right)^{-2}, \quad \text{with} \quad M_V = 0.84\ \mathrm{GeV}. \tag{64}$$

Over a wide Q^2 range (for $Q^2 > 0.1\,\mathrm{GeV}^2$) the form factors given by Eq. (63) take values similar to the MAID07 form factors [59].

The P_{33}(full) model is different from the one applied by Kondratyuk *et al.* [9] (denoted as KBMT05). However, in the intermediate Q^2 range the predictions are comparable (as seen by comparing our Fig. 7 with Fig. 2 from Ref. [9]). We notice

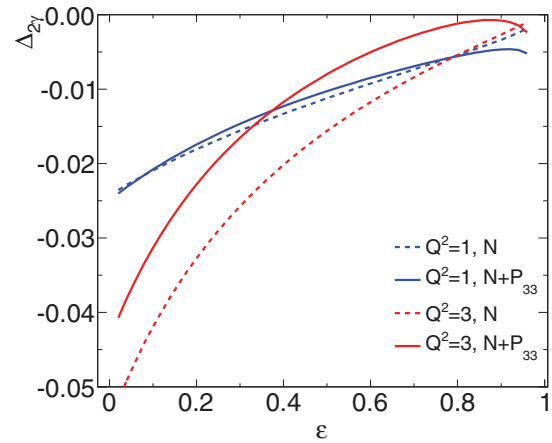

FIG. 7. (Color online) $\Delta_{2\gamma}$ [Eq. (14)] computed for either $\square(N)$ or $\square(N + P_{33})$ TPE contributions. The form factors from Ref. [7] are applied. The inelastic TPE correction, given by $\square(P_{33})$, is computed within the P_{33}(full) model. The values of Q^2 are in units of GeV^2.

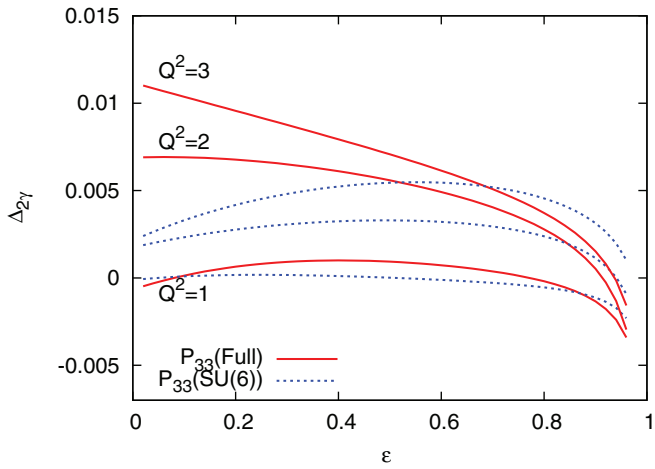

FIG. 8. (Color online) $\Delta_{2\gamma}$ [Eq. (14)] given by the resonance P_{33} contribution only. Calculations are done for the P_{33}[SU(6)] and P_{33}(full) models. The values of Q^2 are in units of GeV^2.

also a qualitative agreement with predictions of the Borisyuk and Kobushkin model [13]. In both approaches the TPE $\Box(P_{33})$ correction is positive in the low and intermediate Q^2 range and it reduces the total TPE correction.

In contrast to $\Delta_{2\gamma}(N)$, the function $\Delta_{2\gamma}(P_{33})$ depends on the details of the hadronic model. Indeed, there are small but noticeable differences between the predictions of KBMT05 and the P_{33}(full) model. To illustrate the model dependence of $\Delta_{2\gamma}(P_{33})$ the predictions of the TPE correction obtained within P_{33}[SU(6)] and P_{33}(full) models are plotted in Fig. 8. There is a clear discrepancy between predictions of the two approaches.

V. NEURAL NETWORK VERSUS HADRONIC MODEL

The electromagnetic FFs of the proton are the input of the hadronic model used in this paper. For comparison of the self-consistency the proton FFs of HM are obtained from the fit of the HM to the same unpolarized cross-section data as in the BNN (for details see Appendix B). The PT and $R_{+/-}$ data are not taken into consideration, and the constraint coming from assumption (A) does not affect the results. The TPE correction contains either $\Box(N)$ or $\Box(N + P_{33})$ contributions. The obtained FF parameters are given in Tables I (fit I) and II (fit II), while the values of $\chi^2_{\rm min}$/NDF (where NDF is the number of degrees of freedom) are reported in Table III.

It is interesting to notice that the mass parameters of fit I are not well spaced. For instance, the parameters m_1^1 and m_3^1 take quite similar values. The same feature characterizes the fits from Ref. [7], where parametrization I was also discussed but it was fitted to the FFs from Ref. [60]. Indeed, this parametrization at large Q^2 behaves as $1/Q^2$, while it is expected (based on theoretical arguments [61,62]) that $G_{E,M} \sim 1/Q^4$.

The above observations may suggest that parametrization I is too simple to describe accurately the FFs over a wide Q^2 range. In order to verify this statement we make two fits. In the first we consider the data below $Q^2 = 1$ GeV^2, while in the other we use the data below $Q^2 = 0.5$ GeV^2. For the first case we get the mass parameters $m_1^1 = 2.59$, $m_2^1 = 0.95$, and $m_3^1 = 0.18$ and for the other we get $m_1^1 = 1.70$, $m_2^1 = 0.21$, and $m_3^1 = 9.16$. We see that for low-Q^2 data fits the mass parameters are well separated. However, because of the problems noted above, in further discussion the HM with the FFs given by fit I is treated as a toy model, discussed to present the systematic properties of the hadronic approach.

At low Q^2 there is a visible discrepancy between the BNN and the hadronic model FF fits. This is illustrated in Fig. 9, where the ratio $\mu_p G_E/G_M$ is plotted. There is a satisfactory agreement between fits I and II. In contrast, the ratio $\mu_p G_E/G_M$ predicted by the BNN approach is more consistent with the recent PT measurements [54] (whose data were not included in the BNN fit).

The low-Q^2 discrepancy between the HM and BNN approaches is the result of different treatments of the TPE corrections. This is illustrated in Fig. 5, where we plot the function

$$D_{2\gamma} = \frac{\Delta_{2\gamma}}{1+\Delta_{2\gamma}} = \frac{\Delta C_{2\gamma}}{\sigma_{R,1\gamma+2\gamma}}. \tag{65}$$

It can be seen that the BNN and HM predictions are inconsistent for $Q^2 \in (0.02, 1)$ GeV^2. In contrast, below $Q^2 < 0.02$ GeV^2 and at low ε there is a good agreement between TPE predictions obtained within the two methodologies as well as other theoretical calculations [63]. This low-ε and low-Q^2 behavior of the BNN fit seems to be a systematic property of all BNN-based parametrizations. This is illustrated in Fig. 10, where the $R_{+/-}$ values predicted by the BNN models, rejected due to too small values of the evidence (see Table V), are plotted. In the limit of $\varepsilon \to 0$, with Q^2 very low but fixed, $\sigma_{R,1\gamma}$

TABLE I. Form factor parameters for fit I (56) for the hadronic model with elastic (N) (left panel) and elastic and resonance P_{33} (right panel) TPE contributions. Mass parameters are in the units of GeV^2.

	$k=1$	$k=2$		$k=1$	$k=2$
m_1^k	1.234	0.321	m_1^k	1.221	0.327
m_2^k	0.181	4.298	m_2^k	0.173	4.019
m_3^k	1.085	4.641	m_3^k	1.097	4.450
f_1^k	−6.569	0.694	f_1^k	−7.934	0.713
f_2^k	0.055	−13.44	f_2^k	0.051	−10.16

TABLE II. Form factor parameters for the II (57) for the hadronic model with elastic (N) (left panel) and elastic and resonance P_{33} (right panel) TPE contributions. Mass parameters are in the units of GeV.

	$k=1$	$k=2$		$k=1$	$k=2$
m_1^k	0.7732	1.0595	m_1^k	0.7866	1.0247
m_2^k	0.9489	1.5629	m_2^k	0.9641	1.4914
m_3^k	0.8457	0.5474	m_3^k	0.8550	0.5082
f_1^k	3.9833	1.2645	f_1^k	4.3360	1.4592
f_2^k	3.9334	−0.269	f_2^k	3.7328	−0.3175

TABLE III. $\chi^2_{\min}$/NDF values.

FF	(N)	$(N + P_{33})$
fit I	389/403	397/403
fit II	386/403	395/403

[Eq. (10)] is dominated by the magnetic contribution, and the main constraint comes from the fact that $G_M(Q^2 = 0) = \mu_p$. As a result the TPE fit is affected by several data points present in the low-Q^2 and low-ε domain.

For completeness of the low-Q^2 comparison we report the values of the proton radius obtained from the BNN and HM fits in Table IV.

The value of $\sqrt{\langle r_E^2 \rangle}$ computed from the BNN fit is consistent with fit II and the recent atomic measurement [43] [0.84087(39) fm]. However, it disagrees with the prediction based on fit I. The latter is inconsistent with fit II as well.

There are two major reasons for the above inconsistency. The first one is induced by the systematic differences between the predictions of the TPE by the BNN and HM approaches in the low-Q^2 range. The discrepancy between the values of the proton radius based on fits I and II is the result of the problem of parametrization I (mentioned already above) with the proper description of the FFs over a wide Q^2 range. In general, the low number of parameters in fits I and II limits the flexibility of the FF parametrizations and their ability for simultaneous description of the low- and high-Q^2 data. Therefore the low- and high-Q^2 fit dependence can be affected by the high- and low-Q^2 data.

To summarize the low-Q^2 discussion we would like to emphasize that in both the present and the BNN data analyses

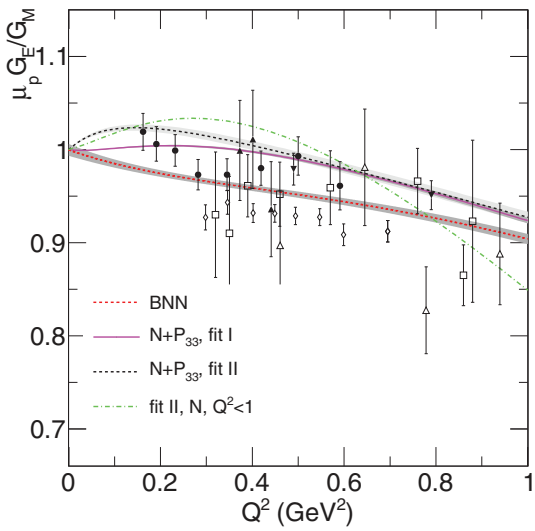

FIG. 9. (Color online) The ratio $\mu_p G_E/G_M$ calculated based on fits I and II (N and resonance P_{33} contributions) as well as the BNN fit. Additionally, the ratio $\mu_p G_E/G_M$ obtained based on the fit (parametrization II) to the unpolarized cross-section data below $Q^2 = 1$ GeV^2 is also plotted. The PT data (points with error bars) are taken from Ref. [53] and Zhan *et al.* [54] (open diamonds). The shaded areas show 1σ error computed from the covariance matrix.

TABLE IV. Values of the proton radius $\sqrt{\langle r_E^2 \rangle}$ obtained from the BNN and HM fits in femtometers.

BNN	fit I	fit II
0.85 ± 0.01	0.898 ± 0.001	0.867 ± 0.002

our attention was not particularly focused on the $Q^2 \to 0$ limit. Certainly, accurate calculations of the proton radius require more careful discussion, as is reported in Refs. [54,64–68].

Above $Q^2 = 1$ GeV^2 the BNN FF ratio $\mu_p G_E/G_M$, on the qualitative level, is comparable with the hadronic model predictions (Fig. 11). All fits agree well with the PT measurements [53]. As could be expected, the inclusion of the $\square(P_{33})$ contribution into the hadronic model increases the value of the electric form factor at larger values of Q^2 (see Fig. 11).

Excellent consistency between predictions of the TPE effect by the BNN and HM approaches appears for $Q^2 \in (1,3)$ GeV^2 (see Figs. 10 as well as 5 and 6). Above $Q^2 = 3$ GeV^2 the agreement is at the 2σ level only.

In order to show the strength of the BNN approach we confront its predictions of the TPE effect with our previous global analysis (ABGG) [29] made in the conventional way (see Figs. 5, 6, and 10). In this approach, following the proposal of Ref. [28], some functional form of the TPE term was postulated. But the same cross-section and PT data as in the case of the BNN were analyzed. The constraint (A) was also imposed. Although both the electric and the magnetic FF fits of the ABGG analysis are very similar to those obtained within the BNN and the other phenomenological approaches [34], the predictions of the TPE correction agree with the HM only for Q^2 around 3 GeV^2 (see Fig. 10). In the ABGG approach the model dependence of the final fits was not discussed. The successful fits were characterized by a reasonable value of $\chi^2_{\min}$/NDF. However, in the BNN analysis the models rejected, due to too low evidence, were characterized also by a reasonable $\chi^2_{\min}$/NDF (see Table V). But the TPE corrections predicted based on these fits, similarly to those for the ABGG analysis, are inconsistent with the best BNN fit and the hadronic model calculations.

The results of our paper are complementary to the conclusions of Ref. [27] (AMT), where a global analysis of the world ep data was also performed. The TPE correction was given by the sum of elastic $\square(N)$ and inelastic contributions. The latter was described by a phenomenological function, which

TABLE V. The minimum of χ^2 and the maximum of the evidence obtained for the best BNN model (in bold) and the fits rejected because of too small values of the evidence. The total number of points in the fit is 529.

	$\mathcal{N}_{4,2}$	$\mathcal{N}_{4,3}$	$\mathcal{N}_{6,2}$	$\mathcal{N}_{6,3}$	$\mathcal{N}_{6,4}$	$\mathcal{N}_{5,7}$	$\mathcal{N}_{5,6}$
$\chi^2_{\min}$	507	511	497	493	486	539	**478**
ln(evidence)	−633	−630	−635	−624	−639	−699	**−611**

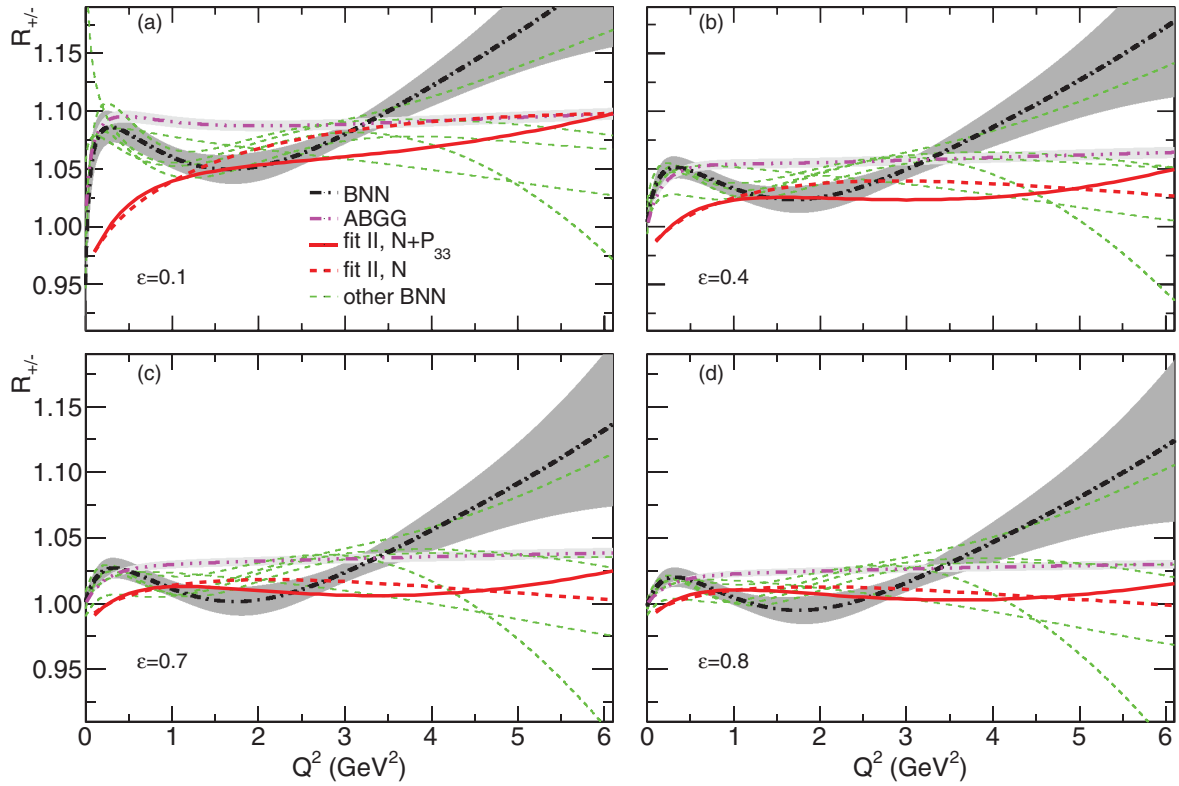

FIG. 10. (Color online) The ratio $R_{+/-}$ computed based on the BNN (lines) and HM (squares and triangles) approaches as well as the preliminary VEPP-3 measurements [37] (filled circles). The HM predictions are computed for the model (fit II) which contains either elastic (N) or elastic and resonance P_{33} TPE contributions. The triangle points are right-shifted by 0.02 GeV^2. The shaded areas show 1σ error computed from the covariance matrix.

approximates the resonance [10] and Generalized Parton Distributions-based [14] fraction of the TPE effect. To compute the elastic contribution the FFs (parametrization I) were fitted to the electromagnetic FFs from [60]. The FFs (parametrization from [61]) were fitted to the unpolarized cross-section

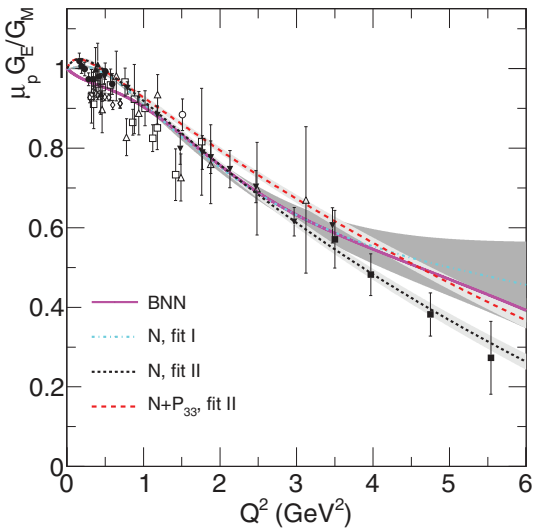

FIG. 11. (Color online) The ratio $\mu_p G_E/G_M$ computed based on fits I and II as well as for the BNN. The HM fits include either elastic (N) or elastic and P_{33} resonance contributions. The PT data are taken from Refs. [53,54]. The shaded areas show 1σ error computed from the covariance matrix.

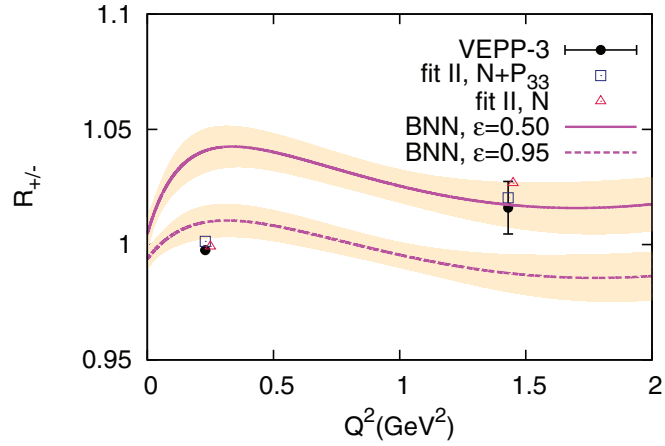

FIG. 12. (Color online) Predictions of $R_{+/-}$ [Eq. (2)] computed based on the BNN, HM (fit II), and ABGG approaches. The TPE correction includes either elastic (N) or elastic and P_{33} resonance (full-model) contributions. Additionally, the plots of $R_{+/-}$ predicted based on the BNN fits rejected due to too small values of the evidence (see Table V) are presented (other BNNs). The shaded areas show 1σ error computed from the covariance matrix.

TABLE VI. $R_{+/-}$ predictions for the VEPP-3 experiment [37] computed within the BNN and HM approaches. In the latter the model contains either elastic (N) or elastic and resonance P_{33} TPE contributions (fit II).

Q^2 (GeV^2)	0.23	1.43	0.82	0.96
ε	0.95	0.50	0.42	0.29
	1.009	1.017	1.037	1.043
BNN	±0.007	± 0.010	± 0.007	± 0.008
$\square(N)$	0.999	1.026	1.020	1.028
$\square(N + P_{33})$	1.001	1.020	1.019	1.027

data (corrected by the TPE) and the PT measurements. It was shown that the cross-section data modified due to the TPE effect are consistent with the PT measurements. An effort was made to estimate uncertainties of the theoretical model for the TPE effect.

In the AMT as well as in the ABGG approaches the analyses were performed in the spirit of frequentist statistics (using the least-squares method), while the neural network analysis was done within Bayesian statistics (for the short review see [69]). In both statistical approaches to find the best fit some error function is minimized. But in the BNN approach the procedure for finding the optimal model is more complicated. In the first stage of the approach a large population of MLPs (more than 1000 networks of given architectures) is trained to find the configuration of the weight parameters for which the error function is at the local minimum. The best model favored by the data maximizes the evidence. It is the probability distribution, which only partially depends on the error function. It contains Occam's contribution[6] [Eqs. (34) and (35)], which penalizes too complex models and allows one to choose the fit with the best predictive power.

The idea of the BNN formalism is to distinguish the statistical model, describing the data, that is characterized by good predictive power. To verify this property we make an estimate of $R_{+/-}$ for the new measurements done in the VEPP-3 experiment [37], which were not included in the BNN analysis. It can be noticed that the BNN and the HM predictions are consistent with the new data (see Fig. 12). For future tests we provide reader with our predictions of $R_{+/-}$ for two other kinematics regions, which are going to be explored in the Novosibirsk experiment (see Table VI).

VI. SUMMARY

The TPE correction was computed within the hadronic model. For the hadronic intermediate states the proton and $P_{33}(1232)$ resonance were considered. The electromagnetic proton form factor parameters were obtained from a global fit to the cross-section data only. Two FF parametrizations were discussed: the sum of monopoles and dipoles. The TPE $\square(P_{33})$ contribution was computed by taking different forms of transition vertex and form factors than discussed previously. In particular, two parametrizations of the FFs for the $N \to P_{33}$ transition were considered.

The main goal of this paper was to confront the predictions of the TPE effect coming from the hadronic model and the Bayesian analysis of the ep scattering data. The latter was performed by applying the neural network framework. The BNN response was constrained by the assumption that the PT data are not sensitive to the TPE effect. Hence this comparison provides also a quantitative verification of this assumption.

It was demonstrated that the BNN and the hadronic model predictions agree on a quantitative level over a wide Q^2 range. In particular, for Q^2 between 1 and 3 GeV^2 the TPE corrections resulting from the two approaches are very similar. In the intermediate Q^2 range ($Q^2 > 3$ GeV^2) the agreement is at the 2σ level. This is the kinematical domain where the data are limited. Obviously, the limited amount of data affect the BNN predictions. On the other hand, in this kinematical limit the applied hadronic model can be questionable.

For Q^2 between 0.01 and 0.8 GeV^2 the BNN and hadronic model predictions are inconsistent. In this Q^2 range assumption (A) does not work effectively. A similar inconsistency appears when one compares with the ABGG predictions of the TPE. Indeed, in the ABGG approach assumption (A) also played a crucial role in the analysis.

A next step to improve the BNN approach would be to replace assumption (A) by a weaker statement and enlarge the number of independent TPE functions from one to four. The main problem is that the PT data seem to be not informative enough about the ε dependence of the TPE effect.

ACKNOWLEDGMENTS

We thank J. Sobczyk, J. Żmuda, and D. Prorok for reading the manuscript as well as C. Juszczak for useful hints about software development and his comments on the paper. We thank also J. Arrington for his remarks on the previous version of the paper. Part of the calculations has been carried out at the Wroclaw Centre for Networking and Supercomputing (http://www.wcss.wroc.pl), supported by Grant No. 268.

APPENDIX A: EVALUATION OF THE TPE TERM

We consider the proton electromagnetic FFs of the form

$$F_i(t) = \sum_{k=1}^{L_i} \sum_{n_k^i=1}^{N_k^i} \frac{f_i^{n_k^i}}{\left(t - M_{i,n_k}^2\right)^k}, \quad i = 1,2. \tag{A1}$$

Every kth pole function can be written as a derivative,

$$\frac{1}{(n-1)!}\partial_x^{n-1}\frac{1}{t-x}\bigg|_{x=M^2} = \frac{1}{(t-M^2)^n}. \tag{A2}$$

We introduce the notation

$$\mathcal{D}_M^n f(M^2) \equiv \frac{1}{(n-1)!}\frac{\partial^n f(x^2)}{\partial (x^2)^n}\bigg|_{x^2=M^2}, \quad n = 1,2,\ldots, \tag{A3}$$

where $\mathcal{D}_M^0 f(M^2) \equiv f(M^2)$. Then the form factor is written in the form

$$F_i(t) = \sum_{k=1}^{L_i} \sum_{n_k^i=1}^{N_k^i} \mathcal{D}_{M_{n_k^i}}^{k-1} \frac{f_i^{n_k^i}}{t - M_{n_k^i}^2}. \tag{A4}$$

[6]It may happen that for the model with the highest evidence the error function is not at a global minimum.

We decompose both $w_{N,\parallel}$ and $w_{N,\times}$ functions into four components,

$$w_N = \sum_{i=1}^{2}\sum_{j=1}^{2} w_N^{ij}. \tag{A5}$$

The ij component, due to its form factor, reads

$$w_N^{ij} = \sum_{k=1}^{L_i}\sum_{n_k^i=1}^{N_k^i}\sum_{l=1}^{L_j}\sum_{n_l^j=1}^{N_l^j} f_i^{n_k^i} f_j^{n_l^j} \mathcal{D}^{k-1}_{M_{n_k^i}} \mathcal{D}^{l-1}_{M_{n_l^j}} \mathcal{E} \times \left(L^{\alpha\mu\nu}\mathcal{H}^{N,ij}_{\alpha\mu\nu}, M_{n_k^i}, M_{n_l^j}, M_p\right), \tag{A6}$$

where the leptonic tensor $L^{\alpha\mu\nu}$ is given by either

$$L_\parallel^{\alpha\mu\nu} = \mathrm{Tr}((\hat{k}+m)\gamma^\alpha(\hat{k}'+m)\gamma^\mu(\hat{k}'-\hat{l}+m)\gamma^\nu) \tag{A7}$$

or

$$L_\times^{\alpha\mu\nu} = \mathrm{Tr}((\hat{k}+m)\gamma^\alpha(\hat{k}'+m)\gamma^\mu(\hat{k}'-\hat{l}+m)\gamma^\nu). \tag{A8}$$

The hadronic tensor reads

$$\mathcal{H}^{N,ij}_{\alpha\mu\nu} = \mathrm{Tr}\big((\hat{p}+M_p)\Gamma_\alpha(-q)(\hat{p}'+M_p) \times \Gamma^i_\mu(-l)(\hat{p}'+\hat{l}+M_p)\Gamma^j_\nu(q+l)\big), \tag{A9}$$

$$\Gamma^1_\mu(l) \equiv \gamma_\mu, \quad \Gamma^2_\mu(l) = \frac{i\sigma_{\mu\nu}q^\nu}{2M_p}. \tag{A10}$$

$\mathcal{E}(\mathcal{N},m_a,m_b,m_h)$ is the one-loop integral defined as

$$\mathcal{E}_\parallel(\mathcal{N},m_a,m_b,m_h) = \int \frac{d^4l}{(2\pi)^4} \frac{\mathcal{N}}{\left[l^2-m_a^2\right]\left[(q+l)^2-m_b^2\right][(q+l)^2+i\epsilon][l^2+i\epsilon][(k'-l)^2-m^2+i\epsilon]\left[(p'+l)^2-m_h^2+i\epsilon\right]} \tag{A11}$$

for direct box diagram and

$$\mathcal{E}_\times(\mathcal{N},m_a,m_b,m_h) = \int \frac{d^4l}{(2\pi)^4} \frac{\mathcal{N}}{\left[l^2-m_a^2\right]\left[(q+l)^2-m_b^2\right][(q+l)^2+i\epsilon][l^2+i\epsilon][(k+l)^2-m^2+i\epsilon]\left[(p'+l)^2-m_h^2+i\epsilon\right]} \tag{A12}$$

for exchange box diagram.

In the case of the P_{33} intermediate state we proceed in a similar manner. For instance, for the P_{33}(full) model we have

$$w_\Delta = \sum_{i=1}^{3}\sum_{j=1}^{3} w_\Delta^{ij}. \tag{A13}$$

The components of the hadronic tensor read

$$\mathcal{H}^{\Delta,ij}_{\alpha\mu\nu} = \mathrm{Tr}\big((\hat{p}+M_p)\Gamma_\alpha(-q)(\hat{p}'+M_p)\Gamma^{\Delta,\mathrm{in}}_{\mu\xi,i}(-l,p'+l)[\hat{p}'+\hat{l}+M_\Delta]\Lambda^{\xi\eta}(p'+l)\Gamma^{\Delta,\mathrm{out}}_{\eta\nu,j}(q+l,p'+l)\big), \tag{A14}$$

where

$$\Gamma^{\Delta,\mathrm{in(out)}}_{\mu\nu,i} \equiv \Gamma^{\Delta,\mathrm{in(out)}}_{\mu\nu}\big(C_k^V \to \delta_{ki},\ k=3,4,5\big). \tag{A15}$$

In the case of the P_{33}(full) model, the resonance form factors have the general form

$$C_i^V(t) = -c_i^V \frac{a_i M_V^2}{t-a_i M_V^2}\frac{M_V^4}{\left(t-M_V^2\right)^2}. \tag{A16}$$

They can be written in the form

$$C_i^V(t) = \frac{c_i^{1,1}}{\left[t-\left(M_i^{1,1}\right)^2\right]} + \frac{c_i^{1,2}}{\left[t-\left(M_i^{1,2}\right)^2\right]} + \frac{c_i^{2,1}}{\left[t-\left(M_i^{2,1}\right)^2\right]^2}, \tag{A17}$$

where

$$C_i^{1,1} = -c_i^V \frac{a_i M_V^2}{(a_i-1)^2}, \quad M_i^{1,1} = \sqrt{a_i}M_V, \tag{A18}$$

$$C_i^{1,2} = c_i^V \frac{a_i M_V^2}{(a_i-1)^2}, \quad M_i^{1,2} = M_V, \tag{A19}$$

$$C_i^{2,1} = c_i^V \frac{a_i M_V^4}{(a_i-1)}, \quad M_i^{2,1} = M_V. \tag{A20}$$

$$w_\Delta^{ij} = \sum_{k=1}^{2}\sum_{n_k=1}^{N_k}\sum_{l=1}^{2}\sum_{n_l=1}^{N_l} c_i^{k,n_k} c_j^{l,n_l} \mathcal{D}^{k-1}_{M_i^{k,n_k}} \mathcal{D}^{l-1}_{M_j^{l,n_l}} \times \mathcal{I}\left(L^{\alpha\mu\nu}\mathcal{H}^{\Delta,ij}_{\alpha\mu\nu}, M_i^{k,n_k}, M_j^{l,n_l}, M_\Delta\right), \tag{A21}$$

where $N_1 = 2$, and $N_2 = 1$, $c_i^{2,2} = 0$.

The algebraic calculations, such as computing the leptonic (A7) and (A8) and hadronic (A9) and (A14) tensors and their contractions, are done with the help of the FeynCalc package [70,71].

The integrals in Eqs. (A11) and (A12) are expressed (also with the help of routines in FeynCalc) in terms of Veltman-Passariono scalar loop integrals [72,73]. Because of

the complex structure of the the numerators of the integrals in Eqs. (A11) and (A12) some prereduction of the numerator with the denominator is necessary.

Once the analytic expressions for all integral components are obtained, their numerical values are computed with the LoopTool library [74,75].

APPENDIX B: χ^2

In order to get the FF parameters of parametrizations I (56) and II (57) we analyze the same unpolarized cross-section data as in Ref. [33].

We consider the following χ^2 function:

$$\chi^2 = \sum_{k=1}^{N} \left[\sum_{i=1}^{n_k} \left(\frac{\lambda_k \sigma_{ki}^{\rm th} - \sigma_{ki}^{ex}}{\Delta\sigma_{ki}} \right)^2 + \left(\frac{\lambda_k - 1}{\Delta\lambda_k} \right)^2 \right]. \quad \text{(B1)}$$

Here $N = 28$ is the number of independent data sets in the fit, n_k is a number of points in the kth data set, $\sigma_{ki}^{\rm th}$ is the reduced cross section given by Eq. (16), while σ_{ki}^{ex} and $\Delta\sigma_{ki}^{ex}$ denote the experimental measurement and its error. By $\Delta\lambda_k$'s the systematic normalization errors are introduced. For every data set the normalization parameter λ_k is established from the fit. Our treatment of the systematic normalization errors is the same as in Refs. [27,29,33,76]. For a statistical explanation of this procedure see Ref. [77].

[1] G. Ron *et al.*, Phys. Rev. C **84**, 055204 (2011); A. J. R. Puckett *et al.*, *ibid.* **85**, 045203 (2012).
[2] P. G. Blunden, W. Melnitchouk, and J. A. Tjon, Phys. Rev. Lett. **91**, 142304 (2003).
[3] P. A. M. Guichon and M. Vanderhaeghen, Phys. Rev. Lett. **91**, 142303 (2003).
[4] Y. C. Chen, A. Afanasev, S. J. Brodsky, C. E. Carlson, and M. Vanderhaeghen, Phys. Rev. Lett. **93**, 122301 (2004).
[5] Y.-S. Tsai, Phys. Rev. **122**, 1898 (1961); L. W. Mo and Y. S. Tsai, Rev. Mod. Phys. **41**, 205 (1969); Y.-S. Tsai, Lectures given at NATO Advanced Institute on Electron Scattering and Nuclear Structure at Cagliari, Italy, September 1970, SLAC-PUB-0848.
[6] Y. M. Bystritskiy, E. A. Kuraev, and E. Tomasi-Gustafsson, Phys. Rev. C **75**, 015207 (2007).
[7] P. G. Blunden, W. Melnitchouk, and J. A. Tjon, Phys. Rev. C **72**, 034612 (2005).
[8] H. Q. Zhou, C. W. Kao, S. N. Yang, and K. Nagata, Phys. Rev. C **81**, 035208 (2010).
[9] S. Kondratyuk, P. G. Blunden, W. Melnitchouk, and J. A. Tjon, Phys. Rev. Lett. **95**, 172503 (2005).
[10] S. Kondratyuk and P. G. Blunden, Nucl. Phys. A **778**, 44 (2006).
[11] D. Borisyuk and A. Kobushkin, Phys. Rev. C **74**, 065203 (2006).
[12] D. Borisyuk and A. Kobushkin, Phys. Rev. C **78**, 025208 (2008).
[13] D. Borisyuk and A. Kobushkin, Phys. Rev. C **86**, 055204 (2012).
[14] A. V. Afanasev, S. J. Brodsky, C. E. Carlson, Y.-C. Chen, and M. Vanderhaeghen, Phys. Rev. D **72**, 013008 (2005).
[15] D. Borisyuk and A. Kobushkin, Phys. Rev. D **79**, 034001 (2009).
[16] J. A. Tjon, P. G. Blunden, and W. Melnitchouk, Phys. Rev. C **79**, 055201 (2009).
[17] N. Kivel and M. Vanderhaeghen, J. High Energy Phys. 04 (2013) 029.
[18] D. Borisyuk and A. Kobushkin, arXiv:1306.4951 [hep-ph].
[19] J. Arrington, P. G. Blunden, and W. Melnitchouk, Prog. Part. Nucl. Phys. **66**, 782 (2011).
[20] M. Vanderhaeghen, Few Body Syst. **41**, 103 (2007).
[21] C. E. Carlson and M. Vanderhaeghen, Annu. Rev. Nucl. Part. Sci. **57**, 171 (2007).
[22] J. Arrington, Phys. Rev. C **69**, 022201 (2004).
[23] E. Tomasi-Gustafsson and G. I. Gakh, Phys. Rev. C **72**, 015209 (2005).
[24] V. Tvaskis, J. Arrington, M. E. Christy, R. Ent, C. E. Keppel, Y. Liang, and G. Vittorini, Phys. Rev. C **73**, 025206 (2006).
[25] M. A. Belushkin, H.-W. Hammer, and U.-G. Meissner, Phys. Lett. B **658**, 138 (2008).
[26] D. Borisyuk and A. Kobushkin, Phys. Rev. C **76**, 022201 (2007).
[27] J. Arrington, W. Melnitchouk, and J. A. Tjon, Phys. Rev. C **76**, 035205 (2007).
[28] Y. C. Chen, C. W. Kao, and S. N. Yang, Phys. Lett. B **652**, 269 (2007).
[29] W. M. Alberico, S. M. Bilenky, C. Giunti, and K. M. Graczyk, Phys. Rev. C **79**, 065204 (2009).
[30] E. Tomasi-Gustafsson, M. Osipenko, E. A. Kuraev, and Yu. Bystritsky, arXiv:0909.4736 [hep-ph].
[31] J. Guttmann, N. Kivel, M. Meziane, and M. Vanderhaeghen, Eur. Phys. J. A **47**, 77 (2011).
[32] S. Venkat, J. Arrington, G. A. Miller, and X. Zhan, Phys. Rev. C **83**, 015203 (2011).
[33] K. M. Graczyk, Phys. Rev. C **84**, 034314 (2011).
[34] I. A. Qattan, A. Alsaad, and J. Arrington, Phys. Rev. C **84**, 054317 (2011).
[35] I. A. Qattan and A. Alsaad, Phys. Rev. C **83**, 054307 (2011); **84**, 029905(E) (2011).
[36] R. P. Bennett, AIP Conf. Proc. **1441**, 156 (2012).
[37] A. V. Gramolin *et al.*, Nucl. Phys. Proc. Suppl. **225–227**, 216 (2012).
[38] L. Ice *et al.* (OLYMPUS Collaboration), AIP Conf. Proc. **1423**, 206 (2012).
[39] K. M. Graczyk, P. Plonski, and R. Sulej, J. High Energy Phys. 09 (2010) 053.
[40] M. P. Rekalo and E. Tomasi-Gustafsson, Eur. Phys. J. A **22**, 331 (2004).
[41] M. P. Rekalo and E. Tomasi-Gustafsson, Nucl. Phys. A **740**, 271 (2004).
[42] M. P. Rekalo and E. Tomasi-Gustafsson, Nucl. Phys. A **742**, 322 (2004).
[43] R. Pohl *et al.*, Nature (London) **466**, 213 (2010).
[44] L. C. Maximon and J. A. Tjon, Phys. Rev. C **62**, 054320 (2000).
[45] H. W. L. Naus and J. H. Koch, Phys. Rev. C **36**, 2459 (1987); P. C. Tiemeijer and J. A. Tjon, *ibid.* **42**, 599 (1990).
[46] B. Denby, Comput. Phys. Commun. **49**, 429 (1988); B. Mellado *et al.*, Phys. Lett. B **611**, 60 (2005); K. Kurek, E. Rondio, R. Sulej, and K. Zaremba, Meas. Sci. Technol. **18**, 2486 (2007); J. Damgov and L. Litov, Nucl. Instrum. Methods A **482**, 776 (2002); T. Bayram, S. Akkoyun, and S. O. Kara, Ann. Nucl. Energy **63**, 172 (2014); S. Akkoyun, T. Bayram, S. O. Kara, and A. Sinan, J. Phys. G **40**, 055106 (2013).
[47] J. Hertz, A. Krogh, and R. G. Palmer, *Introduction to the Theory of Neural Computation*, Santa Fe Institute Studies in the Sciences of Complexity (Westview, Boulder, CO, 1991).

[48] R. D. Ball *et al*. (The NNPDF Collaboration), Nucl. Phys. B **874**, 36 (2013).
[49] G. Cybenko, Math. Control Signals Syst. **2**, 303 (1989).
[50] C. M. Bishop, *Neural Networks for Pattern Recognition* (Oxford University Press, New York, 2008).
[51] D. J. C. MacKay, Neural Comput. **4**, 415 (1992); **4**, 448 (1992); in *Models of Neural Networks III*, edited by E. Domany, J. L. van Hemmen, and K. Schulten, Sec. 6 (Springer-Verlag, New York, 1994).
[52] D. J. C. MacKay, Ph.D. thesis, California Institute of Technology, 1991, http://www.inference.phy.cam.ac.uk/mackay/PhD.html.
[53] T. Pospischil *et al*. (A1 Collaboration), Eur. Phys. J. A **12**, 125 (2001); O. Gayou *et al*., Phys. Rev. C **64**, 038202 (2001); (Jefferson Lab Hall A. Collaboration), Phys. Rev. Lett. **88**, 092301 (2002); V. Punjabi *et al*., Phys. Rev. C **71**, 055202 (2005); **71**, 069902(E) (2005); C. B. Crawford *et al*., Phys. Rev. Lett. **98**, 052301 (2007); M. K. Jones *et al*. (Resonance Spin Structure Collaboration), Phys. Rev. C **74**, 035201 (2006).
[54] X. Zhan *et al*., Phys. Lett. B **705**, 59 (2011).
[55] H. F. Jones and M. D. Scadron, Ann. Phys. (NY) **81**, 1 (1973).
[56] J. Liu, N. C. Mukhopadhyay, and L. Zhang, Phys. Rev. C **52**, 1630 (1995).
[57] L. Alvarez-Ruso, S. K. Singh, and M. J. Vicente Vacas, Phys. Rev. C **59**, 3386 (1999).
[58] O. Lalakulich, E. A. Paschos, and G. Piranishvili, Phys. Rev. D **74**, 014009 (2006).
[59] D. Drechsel, S. S. Kamalov, and L. Tiator, Eur. Phys. J. A **34**, 69 (2007).
[60] P. Mergell, U. G. Meissner, and D. Drechsel, Nucl. Phys. A **596**, 367 (1996).
[61] J. J. Kelly, Phys. Rev. C **70**, 068202 (2004).
[62] S. J. Brodsky and G. R. Farrar, Phys. Rev. Lett. **31**, 1153 (1973).
[63] W. A. McKinley and H. Feshbach, Phys. Rev. **74**, 1759 (1948); R. H. Dalitz, Proc. R. Soc. London A **206**, 509 (1951); **206**, 521 (1951).
[64] J. Arrington, J. Phys. G **40**, 115003 (2013).
[65] J. C. Bernauer *et al*. (A1 Collaboration), Phys. Rev. Lett. **105**, 242001 (2010).
[66] J. Arrington and I. Sick, Phys. Rev. C **76**, 035201 (2007).
[67] J. Arrington, Phys. Rev. Lett. **107**, 119101 (2011).
[68] R. Pohl, R. Gilman, G. A. Miller, and K. Pachucki, Annu. Rev. Nucl. Part. Sci. **63**, 175 (2013).
[69] J. Beringer *et al*. (Particle Data Group), Phys. Rev. D **86**, 010001 (2012); G. D'Agostini, *Bayesian Reasoning in Data Analysis: A Critical Introduction* (World Scientific, Singapore, 2003).
[70] R. Mertig, M. Bohm, and A. Denner, Comput. Phys. Commun. **64**, 345 (1991).
[71] R. Mertig and R. Scharf, Comput. Phys. Commun. **111**, 265 (1998).
[72] G. 't Hooft and M. Veltman, Nucl. Phys. B **153**, 365 (1979).
[73] G. Passariono and M. Veltman, Nucl. Phys. B **160**, 151 (1979).
[74] G. J. van Oldenborgh and J. A. M. Vermaseren, Z. Phys. C **46**, 425 (1990).
[75] T. Hahn and M. Perez-Victoria, Comput. Phys. Commun. **118**, 153 (1999).
[76] J. Arrington, Phys. Rev. C **68**, 034325 (2003).
[77] G. D'Agostini, Nucl. Instrum. Meth. A **346**, 306 (1994).