A data-driven approach to π^0,η and η' single and double Dalitz decays*

Rafel Escribano[1,2;1)] Sergi Gonzàlez-Solís[2,3;2)]

[1] Grup de Física Teòrica, Departament de Física, Universitat Autònoma de Barcelona, E-08193 Bellaterra (Barcelona), Spain
[2] Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology (BIST), Campus UAB, E-08193 Bellaterra (Barcelona), Spain
[3] CAS Key Laboratory of Theoretical Physics, Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, China

Abstract: The dilepton invariant mass spectra and integrated branching ratios of the single and double Dalitz decays $\mathcal{P}\to l^+l^-\gamma$ and $\mathcal{P}\to l^+l^-l^+l^-$ ($\mathcal{P}=\pi^0,\eta,\eta'$; l=e or μ) are predicted by means of a data-driven approach based on the use of rational approximants applied to π^0,η and η' transition form factor experimental data in the space-like region.

Keywords: electromagnetic processes and properties, phenomenological models, transition form factor

PACS: 13.20.Cz, 12.15.Hh, 12.40.Vv **DOI:** 10.1088/1674-1137/42/2/023109

1 Introduction

Anomalous decays of the neutral pseudoscalar mesons $\mathcal{P}$ ($\mathcal{P}=\pi^0,\eta,\eta'$) are driven through the chiral anomaly of QCD. Of historical importance is the process $\mathcal{P}\to\gamma\gamma$. Apart from being the main decay channel of the π^0 and the η, its experimental discovery confirmed, for the first time, the existence of anomalies. In this case, the two final state photons are real and the transition form factor (TFF) encoding the effects of the strong interactions of the decaying meson is predicted to be simply a constant, the value of the axial anomaly in the chiral and large-N_c limits of QCD, $F_{\pi^0\gamma\gamma}(0)=1/(4\pi^2F_\pi)$ in the case of neutral pion, where $F_\pi\simeq 92$ MeV is the pion decay constant. However, if one of the two photons is virtual, the corresponding TFF is no longer a constant but a function of the transferred momentum to the virtual photon $F_{\mathcal{P}\gamma\gamma^*}(q^2)$, whereas when both photons are virtual the TFF depends on both photon virtualities and is represented by $F_{\mathcal{P}\gamma^*\gamma^*}(q_1^2,q_2^2)$. A single Dalitz decay occurs through the single-virtual TFF after the conversion of the virtual photon into a lepton pair, while double Dalitz decays proceed with the TFF of double virtuality involving two dilepton pairs in the final state. Dalitz decays are hence attractive processes to improve our knowledge of the TFFs of the $\mathcal{P}\gamma^{(*)}\gamma^*$ vertices. This is the main motivation of this work, together with predicting the invariant mass spectra and the branching ratios ($\mathcal{BR}$) of the decays $\mathcal{P}\to l^+l^-\gamma$ and $\mathcal{P}\to l^+l^-l^+l^-$, with $\mathcal{P}=\pi^0,\eta,\eta'$ and l=e or μ.

From the experimental side, the current status is the following. The PDG reported value for the branching ratio of the decay $\pi^0\to e^+e^-\gamma$ is $(1.174\pm0.035)\%$ [1], which is obtained from the PDG fits of the ratio $\Gamma(\pi^0\to e^+e^-\gamma)/\Gamma(\pi^0\to\gamma\gamma)=(1.188\pm0.035)\%$ (the latest measurement of this ratio, $(1.140\pm0.041)\%$, was performed by ALEPH in 2008 [2]) and $B(\pi^0\to\gamma\gamma)=(98.823\pm0.034)\%$. It is worth commenting that this is the second most important decay mode of the π^0. The branching ratio for the decay $\eta\to e^+e^-\gamma$ has recently been measured by the A2 Collaboration at MAMI [3], $(6.6\pm0.4\pm0.4)\times10^{-3}$ (see also the most recent result in Ref. [4]), and the CELSIUS/WASA Collaboration [5], $(7.8\pm0.5\pm0.8)\times10^{-3}$, whilst the PDG quoted value is $(6.9\pm0.4)\times10^{-3}$ [1], in accordance with the result, $(6.72\pm0.07\pm0.31)\times10^{-3}$, from the WASA@COSY Collaboration [6]. The decay $\eta\to\mu^+\mu^-\gamma$ has been studied by the NA60 Collaboration at CERN SPS [7], though they do not provide a value

Received 31 October 2017, Published online 18 January 2018

* Supported by the FPI scholarship BES-2012-055371 (S.G-S), the Secretaria d'Universitats i Recerca del Departament d'Economia i Coneixement de la Generalitat de Catalunya under grant 2014 SGR 1450, the Ministerio de Ciencia e Innovación under grant FPA2011-25948, the Ministerio de Economía y Competitividad under grants CICYT-FEDER-FPA 2014-55613-P and SEV-2012-0234, the Spanish Consolider-Ingenio 2010 Program CPAN (CSD2007-00042), and the European Commission under program FP7-INFRASTRUCTURES-2011-1 (283286) S.G-S also Received Support from the CAS President's International Fellowship Initiative for Young International Scientist (2017PM0031)

1) E-mail: rescriba@ifae.es

2) E-mail: sgonzalez@itp.ac.cn

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI. Article funded by SCOAP[3] and published under licence by Chinese Physical Society and the Institute of High Energy Physics of the Chinese Academy of Sciences and the Institute of Modern Physics of the Chinese Academy of Sciences and IOP Publishing Ltd

for the branching ratio. The PDG fit reports the value $(3.1\pm0.4)\times10^{-4}$ [1]. The branching fraction for the decay $\eta'\to e^+e^-\gamma$ has recently been measured for the first time by the BESIII Collaboration [8], obtaining a value of $(4.69\pm0.20\pm0.23)\times10^{-4}$. To end, the decay $\eta'\to\mu^+\mu^-\gamma$ was measured long ago at the SERPUKHOV-134 experiment with a value of $(1.08\pm0.27)\times10^{-4}$ [9]. Regarding the double Dalitz decays, the KTeV Collaboration measured the branching ratio of the decay $\pi^0\to e^+e^-e^+e^-$, $(3.46\pm0.19)\times10^{-5}$ [10], thus averaging the PDG result to $(3.38\pm0.16)\times10^{-5}$ [1]. The KLOE Collaboration reported the first experimental measurement of $\eta\to e^+e^-e^+e^-$ [11], $(2.4\pm0.2\pm0.1)\times10^{-5}$, which is in agreement with the result, $(3.2\pm0.9\pm0.5)\times10^{-5}$, provided by the WASA@COSY Collaboration [6]. In Ref. [5], upper bounds for the branching ratios of $\eta\to\mu^+\mu^-\mu^+\mu^-$, $<3.6\times10^{-4}$, and $\eta\to e^+e^-\mu^+\mu^-$, $<1.6\times10^{-4}$, both at 90% CL, are reported. Finally, no experimental evidence for $\eta'\to e^+e^-e^+e^-$, $\eta'\to\mu^+\mu^-\mu^+\mu^-$ and $\eta'\to e^+e^-\mu^+\mu^-$ exists.

On the theory side, the effort is focused on encoding the QCD dynamical effects in the anomalous $\mathcal{P}\gamma^{(*)}\gamma^*$ vertices through the corresponding TFF functions[1]. The exact momentum dependence of these TFFs over the whole energy region is not known; we only have theoretical predictions from chiral perturbation theory and perturbative QCD (ChPT, pQCD), thus constraining the low- and space-like large-momentum transfer regions, respectively. The TFF at zero-momentum transfer can be inferred either from the measured two-photon partial width,

$$|F_{\mathcal{P}\gamma\gamma}(0)|^2=\frac{64\pi}{(4\pi\alpha)^2}\frac{\Gamma(\mathcal{P}\to\gamma\gamma)}{M_{\mathcal{P}}^3}\ , \tag{1}$$

or the prediction from the axial anomaly in the chiral and large-N_c limits of QCD, as mentioned before, while the asymptotic behaviour of the TFF at $Q^2\equiv-q^2\to\infty$ should exhibit the right falloff as $1/Q^2$ [12][2]. Furthermore, the operator product expansion (OPE) predicts the behaviour of the double-virtual TFF in the limit $Q_1^2=Q_2^2\equiv Q^2\to\infty$ to be the same as for the single one, that is, $1/Q^2$ [19][3]. For the intermediate-momentum transfer region, the most common parameterisation of the TFF, widely used by experimental analyses, is provided by the vector meson dominance model (VMD). The dispersive representation of the TFF in terms of q^2, where q^2 is the photon virtuality in the time-like momentum region, can be written as

$$F_{\mathcal{P}\gamma\gamma^*}(q^2)=\int_{s_0}^{\infty}\mathrm{d}s\frac{\rho(s)}{s-q^2-i\epsilon}\ , \tag{2}$$

where s_0 is the threshold for the physical intermediate states imposed by unitarity and $\rho(s)=\mathrm{Im}F_{\mathcal{P}\gamma\gamma^*}(s)/\pi$ is the associated spectral function. To approximate this intermediate-energy part of the spectral function, one usually employs one or more single-particle states. As an illustration, the contribution to the spectral function of a narrow-width resonance of mass $M_{\rm eff}$ reduces to $\rho(s)\propto\delta(s-M_{\rm eff}^2)$, which yields

$$F_{\mathcal{P}\gamma\gamma^*}(q^2)=\frac{F_{\mathcal{P}\gamma\gamma}(0)}{1-q^2/\Lambda^2}\ , \tag{3}$$

where $F_{\mathcal{P}\gamma\gamma}(0)$ serves as a normalisation constant and $\Lambda(=M_{\rm eff})$ is a real parameter which fixes the position of the resonance pole on the real axis. However, the simple and successful single-pole approximation given in Eq. (3) breaks down for $q^2=\Lambda^2$. One may cure this limitation by taking into account resonant finite-width effects as proposed by Landsberg in Ref. [70] when considering the transitions $\mathcal{P}\to l^+l^-\gamma$ in a VMD framework. According to this model, these transitions occur through the exchange of the lowest-lying ρ, ω and ϕ vector resonances and their contributions to the TFF are written as

$$\widetilde{F}_{\mathcal{P}\gamma\gamma^*}(q^2)=\left(\sum_{V=\rho,\omega,\phi}\frac{g_{V\mathcal{P}\gamma}}{2g_{V\gamma}}\right)^{-1}\times\sum_{V=\rho,\omega,\phi}\frac{g_{V\mathcal{P}\gamma}}{2g_{V\gamma}}\frac{M_V^2}{M_V^2-q^2-iM_V\Gamma_V(q^2)}, \tag{4}$$

where $\widetilde{F}_{\mathcal{P}\gamma\gamma^*}(q^2)=F_{\mathcal{P}\gamma\gamma^*}(q^2)/F_{\mathcal{P}\gamma\gamma}(0)$ is defined as the normalised TFF, $g_{V\mathcal{P}\gamma}$ and $g_{V\gamma}$ are the $V\mathcal{P}\gamma$ and $V\gamma$ couplings, respectively, M_V the vector masses, and $\Gamma_V(q^2)$ the energy-dependent widths.

Despite the notorious success of VMD in describing lots of phenomena at low and intermediate q^2, particularly useful for the decays we consider in this work, this model can be seen as a first step in a systematic approximation. Padé approximants are used to go beyond VMD in a simple manner, also incorporating information from higher energies, allowing an improved determination of the low-energy constants relative to other methods [20]. For this reason, we make use in our study of the work in Refs. [17, 21], where all current measurements of the space-like TFFs $\gamma^*\gamma\to\mathcal{P}$ [22–27], produced in the reactions $e^+e^-\to e^+e^-\mathcal{P}$, have been accommodated in nice

1) The matrix element of the $\mathcal{P}\gamma^{(*)}\gamma^{(*)}$ transition is given for real or virtual photons by $\mathcal{M}_{\mathcal{P}\to\gamma^{(*)}\gamma^{(*)}}=ie^2\epsilon_{\mu\nu\alpha\beta}q_1^\mu q_2^\nu\varepsilon_1^\alpha\varepsilon_2^\beta F_{\mathcal{P}\gamma^{(*)}\gamma^{(*)}}(q_1^2,q_2^2)$, where q_1 and q_2 are the four-momenta of the photons, ε_1 and ε_2 are their polarization vectors, and $F_{\mathcal{P}\gamma^{(*)}\gamma^{(*)}}(q_1^2,q_2^2)$ is the TFF function [70].

2) Perturbative QCD predicts $\lim_{Q^2\to\infty}Q^2F_{\pi^0\gamma\gamma^*}(Q^2)=2F_\pi$. Alternative values to this result exist, see for instance Refs. [13, 14], though they seem to be disfavoured, as pointed out in Refs. [15, 16]. For η and η', see the asymptotic values obtained in Refs. [17, 18].

3) The OPE predicts for the case of the pion $\lim_{Q^2\to\infty}Q^2F_{\pi^0\gamma^*\gamma^*}(Q^2,Q^2)=2F_\pi/3$.

agreement with experimental data using these rational approximants. We benefit from these parameterisations valid in the space-like region to predict the transitions $\mathcal{P}\gamma^{(*)}\gamma^*$ in the time-like region for the Dalitz decays we are interested in. Despite some shortcomings Padé approximants have in entirely describing the TFF analytical structure in the time-like region, our primary aim is to achieve reasonable results for these decays which might serve as a guideline for the experimental collaborations. Different parameterizations existing in the literature are based on interpolation formulas [28, 29], resonance chiral theory [30, 31] and dispersive techniques [32, 33], among others [34–43].

This paper is organised as follows. In Section **2**, we introduce our description of the π^0, η and η' transition form factors using the mathematical method of Padé approximants. Sections **3** and **4** are devoted to the analysis of single and double Dalitz decays, respectively, and predictions for several invariant mass spectra and branching ratios are given. Finally, in Section **5**, we present our conclusions.

2 Transition form factors

The usefulness of Padé approximants (PAs) as fitting functions for different form factors has been illustrated, for example, in Refs. [17, 18, 20, 21, 44, 45]. It is not our purpose to provide here formal details of the method but rather to cover some important aspects for consistency. The PAs to a given function are ratios of two polynomials (with degree L and M, respectively)[1)],

$$P_M^L(q^2) = \frac{\sum_{j=0}^{L} a_j (q^2)^j}{\sum_{k=0}^{M} b_k (q^2)^k} = \frac{a_0+a_1q^2+\cdots+a_L(q^2)^L}{1+b_1q^2+\cdots+b_M(q^2)^M}, \quad (5)$$

constructed such that the Taylor expansion around the origin exactly coincides with that of $f(q^2)$ up to the highest possible order, i.e., $\mathcal{O}(q^2)^{L+M+1}$. We would like to point out that the previous VMD ansatz for the form factor, Eq. (3), can be viewed as the first element in a sequence of PAs which can be constructed in a systematic way. By considering higher-order terms in the sequence, one may be able to describe the experimental space-like data with an increasing level of accuracy. The important difference with respect to the traditional VMD approach is that, as a Padé sequence, the approximation is well-defined and can be systematically improved. Although polynomial fitting is more common, in general, rational approximants are able to approximate the original function in a much broader range in momentum than a polynomial [46]. Once a Padé sequence is chosen, the highest element belonging to that particular sequence will be fixed by experimental data. To be specific, the sequence will be stopped as soon as the additional coefficients of the next order PA are statistically compatible with zero.

Despite the success of PAs as fitting functions for space-like TFFs, some important remarks are in order. First, there is no *a priori* mathematical proof ensuring the convergence of a Padé sequence to the unknown TFF function, though a pattern of convergence may be inferred from the data analysis *a posteriori*[2)]. For instance, the excellent performance of PAs in Ref. [18] (see Figs. 2 and 7 there) seems to indicate that the convergence of the η TFF normalisation and low-energy constants is assured (see also Ref. [45] for the η' case). Second, unlike the space-like TFF data analyses [17, 21], one should not expect to reproduce the time-like TFF data, since a Padé approximant contains only isolated poles and cannot reproduce a time-like cut[3)]. However, if this right-hand cut is approximated by one or more single-particle states in the form of one or several narrow-width resonances, as stated before, then the Padé method may be still used up to the first resonance pole, indeed, up to neighbourhoods of the pole. A detailed explanation of the mathematical reasons for this being the case, upon discussion, can be found in Section II of Ref. [45][4)]. Other approaches, such as the so-called z-expasion, incorporate *ab initio* the unitary cut and are shown to be convenient for describing form factors [72]. The size of the region which is affected by the presence of the pole, a disk of radius ε, is not known but, as we will see later, may be deduced, thus fixing the range of application of the PAs for time-like data. Third and last, the poles found in the PAs fitting the TFFs cannot be directly associated with physical resonance poles in the second Riemann sheet of the complex plane. These, in turn, may be obtained following the prescriptions of Refs. [47–49], which is beyond the scope of the present work.

We would like to emphasise that the use of PAs as fitting functions for some set of experimental data can be viewed as an effective mathematical method which intrinsically contains relevant physical information of the function represented by the data set. In this work, we benefit from the findings of Refs. [17, 21], where the π^0, η and η' TFFs were fixed in the space-like region from the analysis of the intermediate process $\gamma\gamma^*\to\mathcal{P}$ by several

1) Without any loss of generality, we take $b_0=1$ for definiteness.

2) For a detailed discussion of Padé convergence applied to form factors, see for instance Ref. [20].

3) The TFF function is unknown but expected to be analytical in the entire q^2-complex plane, except for a branch cut along the real axis for $q^2\geq 4M_\pi^2$.

4) These reasons are based on the fact that the imaginary part of the TFF at the $\pi\pi$ threshold and beyond is smooth and thus acceptable up to the first resonance pole where the threshold expansion in powers of the pion momentum would break down.

experimental collaborations, to predict the time-like region of the same TFFs needed for the description of the reaction $\mathcal{P}\to\gamma\gamma^*$ and therefore for the single and double Dalitz decays studied here. The extrapolated version of the TFFs used in this analysis, from the space-like region to the time-like one, are discussed case by case in the following.

2.1 $\pi^0\to\gamma\gamma^*$

Given the small phase-space available in the $\pi^0\to e^+e^-\gamma$ transition, $4m_l^2\leqslant q^2\leqslant M_\pi^2$, the π^0 TFF can be expressed in terms of its Taylor expansion[1)],

$$F_{\pi^0\gamma\gamma^*}(q^2)=F_{\pi^0\gamma\gamma}(0)\left(1+b_\pi\frac{q^2}{M_\pi^2}+c_\pi\frac{q^4}{M_\pi^4}+\cdots\right), \quad (6)$$

where $F_{\pi^0\gamma\gamma}(0)$ is fixed from Eq. (1) and the values of the low-energy constants (LECs), slope and curvature, b_π and c_π, respectively, are borrowed from Eqs. (12,13) in Ref. [21][2)],

$$\begin{aligned} b_\pi&=3.24(12)_{\rm stat}(19)_{\rm sys}\times10^{-2}\ ,\\ c_\pi&=1.06(9)_{\rm stat}(25)_{\rm sys}\times10^{-3}\ , \end{aligned} \quad (7)$$

where the statistical error is the result of a weighted average of several fits using different types of PAs to the same joint set of π^0 TFF space-like data (see Fig. 1(a) for a graphical example) and the systematic error is attributed to the model dependence of the PA method. In this way, the values obtained for the LECs can be considered as model independent. It is worth mentioning that the systematic errors ascribed to the LECs are quite conservative, in the sense that they are obtained from a comparison of the constants predicted by several well-established phenomenological models for the TFF and their counterparts extracted using various types of PAs from fits to pseudo-data sets generated by the different models. For each LEC, the systematic error is chosen to be the largest difference among these comparisons, making the whole approach reliable and model independent [21].

As one can see in Fig. 1(b), the Taylor expansion in Eq. (6) nicely describes the $\pi^0\to e^+e^-\gamma$ time-like experimental data [71]. Thus, it can be safely used for the description of the π^0 TFF in this region within the range of available phase-space, since the first pole seems to appear, for all types of PAs considered, inside the region of ρ-dominance [21], thus well beyond the phase-space end point. Finally, we would like to remark that this expansion of the single-virtual π^0 TFF will be used for predicting both $\pi^0\to e^+e^-\gamma$ and $\pi^0\to e^+e^-e^+e^-$ decays, the latter by means of a factorisation of the double-virtual π^0 TFF in terms of a product of the single-virtual one (see Subsection **2.4** for details).

2.2 $\eta\to\gamma\gamma^*$

In order to describe the time-like region of the η TFF from the space-like data analysis in Ref. [17], we will employ two PAs, the $P_1^5(q^2)$ and the $P_2^2(q^2)$. These are the highest-order PAs one can achieve when confronted with the joint sets of space-like experimental data which, for the convenience of the reader, we represent in Fig. 2(a). The sequence $P_1^L(q^2)$ is used when the TFF is believed to be dominated by a single resonance, while the $P_N^N(q^2)$

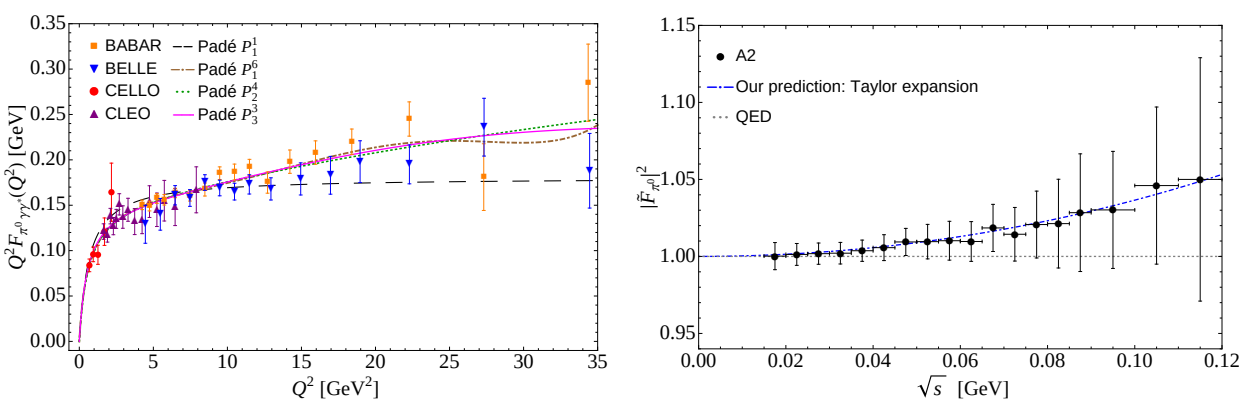

Fig. 1. (color online) Representation of the π^0 TFF as obtained from fits to the space-like experimental data using Padé approximants [21] (a) and our prediction for its modulus square normalised time-like counterpart, $\widetilde{F}_{\pi^0\gamma\gamma^*}(q^2)$, as a function of the invariant dilelectron mass, $\sqrt{s}\equiv m_{ee}$ (b). For the π^0 time-like TFF, the prediction coming from the Taylor expansion in Eq. (6) (dot-dashed blue line) is compared to the experimental data from $\pi^0\to e^+e^-\gamma$ [71] (black circles). The QED prediction (dotted grey line) is also displayed.

1) The form of the Taylor expansion is written in this way to make the slope and curvature parameters non-dimensional.

2) In Ref. [21], the sign of the slope parameter in the Taylor expansion of the π^0 TFF in Eq. (1) should be negative in order to agree with Eq. (4) of the same reference.

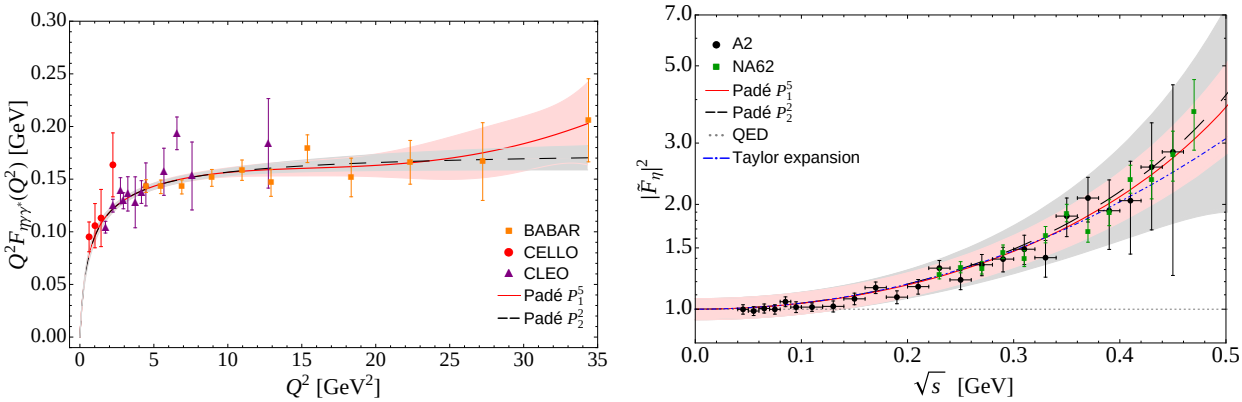

Fig. 2. (color online) Representation of the η TFF as obtained from fits to the space-like experimental data using Padé approximants [17] (a) and our predictions for its modulus square normalised time-like counterpart, $\widetilde{F}_{\eta\gamma\gamma^*}(q^2)$, as a function of the invariant dilepton mass, $\sqrt{s}\equiv m_{ll}$ (b). For the η time-like TFF, the predictions coming from the $P_1^5(q^2)$ (red solid line) and $P_2^2(q^2)$ (long-dashed black line) PAs, and the Taylor expansion (dot-dashed blue line) are compared with the experimental data from $\eta\to e^+e^-\gamma$ [4] (black circles) and $\eta\to\mu^+\mu^-\gamma$ [7] (green squares). The one-sigma error bands associated with $P_1^5(q^2)$ (light-red) and $P_2^2(q^2)$ (light-grey) PAs are also shown. The QED prediction (dotted grey line) is also displayed.

one is appropriate for the case where the TFF fulfils the asymptotic behaviour[1)]. A Taylor expansion equivalent to Eq. (6), with $b_\eta = 0.60(6)_{\rm stat}(3)_{\rm sys}$ and $c_\eta = 0.37(10)_{\rm stat}(7)_{\rm sys}$ for the slope and curvature parameters, respectively, is better not to be used in this case because of the larger phase-space available, $4m_l^2 \leqslant q^2 \leqslant M_\eta^2$. From the analysis in Ref. [17], we also obtained that the fitted poles for the $P_1^L(q^2)$ sequence are seen in the range (0.71,0.77) GeV, beyond the phase-space end point, thus again making our approach applicable and the predictions reliable.

Our predictions for the modulus squared of the normalised time-like η TFF, $\widetilde{F}_{\eta\gamma\gamma^*}(q^2)$, as a function of the invariant dilepton mass, $\sqrt{s}\equiv m_{ll}$, are shown in Fig. 2(b), together with the experimental data points from the A2 Collaboration for the decay $\eta\to e^+e^-\gamma$ [4] (black circles) and the NA60 experiment for $\eta\to\mu^+\mu^-\gamma$ [7][2)] (green squares). The predictions from the $P_1^5(q^2)$ (solid red line) and $P_2^2(q^2)$ (long-dashed black line) are almost identical and in nice agreement with the experimental data, whereas the Taylor expansion (dot-dashed blue line) is not so precise in the upper part of the spectrum. For this reason, we will use both PAs indiscriminately in our analysis. The one-sigma error bands associated with $P_1^5(q^2)$ and $P_2^2(q^2)$ PAs are displayed in light-red and light-grey, respectively. These error bands are built from the uncertainty in the coefficients of the PAs[3)] and the normalisation factor extracted from the two-photon decay width. These bands also include a (tiny) source attributed to the systematic uncertainty coming from the differing results, as obtained from the element where we stop the sequence and the preceding ones.

2.3 $\eta'\to\gamma\gamma^*$

The description of the whole time-like η′ TFF by means of PAs is elaborate. The available phase space, $4m_l^2 \leqslant q^2 \leqslant M_{\eta'}^2$, now includes an energy region where poles associated with these PAs can emerge. The analysis of the η′ TFF space-like data performed in Ref. [17] (see Fig. 3(a) for the corresponding space-like fit results) revealed the appearance of a pole in the range (0.83,0.86) GeV for the cases of a $P_1^L(q^2)$ sequence. Consequently, we cannot employ the method of PAs for describing the time-like TFF in the entire phase-space region, and a complementary approach must be used. Therefore, we propose to match the description based on PAs to that given by Eq. (4) at a certain energy point[4)]. Given the mass and the width of the ρ meson, the first of the

1) In Ref. [17], the fit to space-like data is done for $Q^2|F(Q^2)|$ and not for the TFF itself. As a consequence, PAs satisfying the correct asymptotic limit, that is, $\lim_{Q^2\to\infty} Q^2F(Q^2)=$const., are represented by the sequence $P_N^N(q^2)$.

2) We thank S. Damjanovic from the NA60 experiment for providing us with the time-like TFF data points obtained from $\eta\to\mu^+\mu^-\gamma$.

3) The coefficients of the PAs along with their errors and the correlation matrix can be obtained from the authors upon request.

4) To proceed with the matching, we have considered an energy-dependent width for the ρ resonance,

$$\Gamma_\rho(q^2)=\Gamma_\rho\frac{q^2}{M_\rho^2}\frac{\sigma^3(q^2)}{\sigma^3(M_\rho^2)}\ ,$$

with $\sigma(q^2)=\sqrt{1-4M_\pi^2/q^2}$, and a constant width for the ω and ϕ narrow resonances. Input values for the masses and widths as well as for the rest of the couplings entering Eq. (4) are taken from Ref. [1].

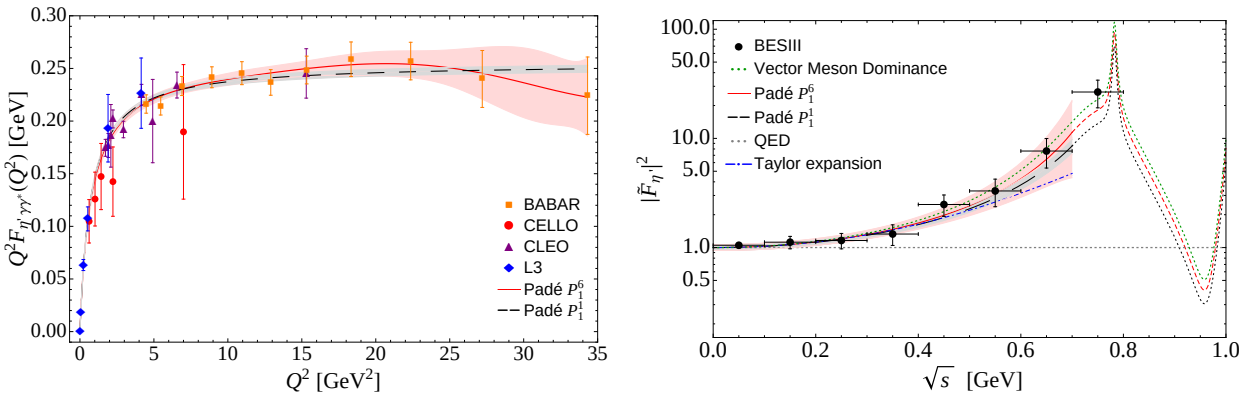

Fig. 3. (color online) Representation of the η' TFF as obtained from fits to the space-like experimental data using Padé approximants [17] (a) and our predictions for its modulus squared normalised time-like counterpart, $\widetilde{F}_{\eta'\gamma\gamma^*}(q^2)$, as a function of the invariant dilepton mass, $\sqrt{s}\equiv m_{ll}$ (b). For the η' time-like TFF, the predictions up to the matching point located at $\sqrt{s}$=0.70 GeV coming from the $P_1^6(q^2)$ (solid red line) and $P_1^1(q^2)$ (long-dashed black line) PAs, and the Taylor expansion (dot-dashed blue line) are compared to the experimental data from $\eta'\to e^+e^-\gamma$ [8] (black circles). From the matching point on, rescaled versions of the VMD description in Eq. (4) are used. The one-sigma error bands associated to $P_1^6(q^2)$ (light-red) and $P_1^1(q^2)$ (light-grey) PAs are also shown. The QED (dotted grey line) and VMD (dotted green line) predictions are also displayed.

resonances included in the VMD description, the region of influence due to its presence may be defined using the half-width rule as $M_\rho\pm\Gamma_\rho/2$ [50], thus deducing the value of the radius ε mentioned earlier. The particular energy point located at $\sqrt{s}\simeq0.70$ GeV, the lowest value of the former region for $M_\rho\simeq775$ MeV and $\Gamma_\rho\simeq150$ MeV, fixes the optimal matching point[1)] With this value fixed, a representation valid in the whole phase-space domain is that given by the PA below the matching point and Eq. (4) above it. In order to match both descriptions of the form factor at the matching point we have to rescale the shape of the VMD spectrum accordingly. In this manner, we keep track of the resonant behaviour in the upper part of the spectrum where PAs cannot be applied, while the low-energy region is predicted in a more systematic way as compared to VMD by PAs established uniquely from space-like data. This will allow us to integrate the whole spectrum and predict the branching ratio of the several η' Dalitz decays considered here.

Our predictions for the time-like η' TFF, together with the experimental data points from the BESIII Collaboration for the decay $\eta'\to e^+e^-\gamma$ [8] (black circles), are displayed in Fig. 3(b). The results from the $P_1^6(q^2)$ (red solid line) and $P_1^1(q^2)$ (black long-dashed line) are shown up to the matching point. The corresponding error bands are in light-red and light-grey, respectively. From the matching point on, our predictions are replaced by a rescaled VMD representation based on the three lowest-lying vector resonances. Our PA-based predictions are again in fine agreement with experiment. A Taylor expansion with $b_{\eta'}=1.30(15)_{\rm stat}(7)_{\rm sys}$ and $c_{\eta'}=1.72(47)_{\rm stat}(34)_{\rm sys}$ [17] and the VMD parameterization, Eq. (4), are also included for comparison. It is worth mentioning that an extrapolation of the $P_1^6(q^2)$ PA beyond the matching point nicely passes through the last experimental point. This can be understood in the following terms. The VMD description includes three resonances whose poles are located at different places, while the $P_1^L(q^2)$ PAs include only one. Making use of the single-pole approximation in Eq. (3) and the Taylor expansion in Eq. (6) for the case of the η', the slope parameter is identified as $b_{\eta'}=m_{\eta'}^2/\Lambda^2$. Using the $b_{\eta'}$ value deduced from Eq. (4) one gets $\Lambda=M_{\rm eff}=0.822(58)$ GeV, where the error is due to the half-width rule and can be utilised as a measure of the region of influence of the pole. The former value is very similar to the one obtained from the pole position of the $P_1^6(q^2)$ PA, located at $\sqrt{s}$=0.833 GeV. Therefore, the region of influence of this pole can be estimated to be in the interval (0.77,0.89) GeV. It is for this reason that the last experimental point would also be in agreement with the $P_1^6(q^2)$ prediction [45]. This is not so for the $P_1^1(q^2)$ PA, thus showing that increasing the Padé order allows for a better description of the data. In any case, for the numerical analysis of the different decays involving the η', we keep both PAs for the sake of comparison.

2.4 $\mathcal{P}\to\gamma^*\gamma^*$

The double-virtual TFF, $F_{\mathcal{P}\gamma^*\gamma^*}(q_1^2,q_2^2)$, depends on both photon virtualities, q_1 and q_2. Due to Bose symmetry, it must satisfy $F_{\mathcal{P}\gamma^*\gamma^*}(q_1^2,q_2^2)=F_{\mathcal{P}\gamma^*\gamma^*}(q_2^2,q_1^2)$.

1) The region of influence attributed to the ω and ϕ poles is negligible, since these are narrow resonances and are placed far from the matching point.

Its normalisation is obviously the same as the single-virtual TFF, $F_{\mathcal{P}\gamma^*\gamma^*}(0,0) = F_{\mathcal{P}\gamma\gamma^*}(0)$, and can be extracted either from the two-photon partial width by means of Eq. (1) or from the axial anomaly. It must also satisfy that when one of the photons is put on-shell the double-virtual TFF becomes the single-virtual one, i.e. $\lim_{q_i^2\to 0} F_{\mathcal{P}\gamma^*\gamma^*}(q_1^2,q_2^2)=F_{\mathcal{P}\gamma\gamma^*}(q^2)$ for $i=1,2$. In addition, the double-virtual TFF can fulfil the asymptotic space-like constraints, $\lim_{Q^2\to\infty} F_{\mathcal{P}\gamma^*\gamma^*}(-Q^2,0) \propto 1/Q^2$ [51] and $\lim_{Q^2\to\infty} F_{\mathcal{P}\gamma^*\gamma^*}(-Q^2,-Q^2)\propto 1/Q^2$ [19].

Due to the lack of experimental information in the case of double-virtual TFFs, our initial ansatz will be to use the standard factorisation approach, which in terms of normalised form factors reads $\widetilde{F}_{\mathcal{P}\gamma^*\gamma^*}(q_1^2,q_2^2)= \widetilde{F}_{\mathcal{P}\gamma\gamma^*}(q_1^2,0)\widetilde{F}_{\mathcal{P}\gamma\gamma^*}(0,q_2^2)$ [52–54]. This double-virtual TFF description may or may not satisfy the high-energy constraints above. As an example, the PA $P_1^0(q^2) = a_0/(1-a_1q^2)$, corresponding to the single-pole approximation in Eq. (3) motivated by VMD, would induce a $1/q^4$ term in the double-virtual TFF, which violates the last of the asymptotic constraints mentioned before (OPE prediction) [19, 55–57]. For this reason, we also use for our study the lowest order bivariate approximant

$$P_1^0(q_1^2,q_2^2)=\frac{a_{0,0}}{1-\frac{b_{1,0}}{M_{\mathcal{P}}^2}(q_1^2+q_2^2)+\frac{b_{1,1}}{M_{\mathcal{P}}^4}q_1^2q_2^2}\ , \tag{8}$$

which consists of a generalisation of the univariate PAs named Chisholm approximants (CAs) [46]. The analysis of the $\pi^0\to e^+e^-$ decay is a recent example that illustrates the application of these CAs [58] (see also P. Masjuan's contribution in Ref. [59]). In Eq. (8), $a_{0,0}$ is identified as the normalisation $F_{\mathcal{P}\gamma^*\gamma^*}(0,0)$ and then fixed from Eq. (1), $b_{1,0}$ is the slope of the single-virtual TFF obtained in Refs. [17, 21], that is, b_π from Eq. (7) for the pion and $b_{\eta^{(\prime)}}$ from Eq. (5) in Ref. [17] for the η and η', respectively, and $b_{1,1}$ corresponds to the double-virtual slope which may be extracted in the future as soon as experimental data for the double-virtual TFFs become available. For the numerical analysis, we consider, as a conservative estimate, varying $b_{1,1}$ from the value respecting the OPE prediction, $b_{1,1}=0$, to $b_{1,1}=2b_{1,0}^2$, far from the factorisation result $b_{1,1} = b_{1,0}^2$. In this manner, we can test the sensitivity of our predictions to the double-virtual slope. We also encourage experimental groups to perform double-virtual TFF measurements in order to fix this parameter. In this work, we employ both descriptions indiscriminately, the factorisation ansatz and the bivariate approximant in Eq. (8). See also Ref. [60] for a recent approach to the double-virtual TFF of the η meson based on the standard factorisation approach.

3 Single Dalitz decays

The generic amplitude for the single Dalitz decays involves the corresponding single-virtual TFF, as described in Section **2**, and reads

$$\begin{aligned}\mathcal{A}(\mathcal{P}\to l^+l^-\gamma) &= -ie^2F_{\mathcal{P}\gamma\gamma^*}(q^2)\varepsilon^{\alpha\beta\mu\nu}k_\alpha q_\beta\epsilon_\mu^*(k)\\ &\quad\times\frac{-ig_{\nu\rho}}{q^2}\bar{u}(p_{l^-})(-ie)\gamma^\rho v(p_{l^+}).\end{aligned} \tag{9}$$

The corresponding differential decay rate is given by

$$\begin{aligned}\frac{d\Gamma_{\mathcal{P}\to l^+l^-\gamma}}{d\sqrt{s}\,\Gamma^{\rm exp}_{\mathcal{P}\to\gamma\gamma}} &= \frac{4\alpha}{3\pi\sqrt{s}}\sqrt{1-\frac{4m_l^2}{s}}\left(1+\frac{2m_l^2}{s}\right)\\ &\quad\times\left(1-\frac{s}{M_{\mathcal{P}}^3}\right)^3|\widetilde{F}_{\mathcal{P}\gamma\gamma^*}(s)|^2,\end{aligned} \tag{10}$$

where $s\equiv q^2=(p_++p_-)^2\equiv M^2_{l^+l^-}$ is the dilepton invariant mass and the TFF appears to be normalised in order to avoid misunderstandings due to the different conventions for $F_{\mathcal{P}\gamma\gamma^*}(0)$ existing in the literature. For the numerical computations, we have employed the PrimEx Coll. resulting value $\Gamma_{\pi^0\to\gamma\gamma}=7.82(14)(17)$ eV [61] and the PDG fitted values $\Gamma_{\eta\to\gamma\gamma}=0.516(18)$ keV and $\Gamma_{\eta'\to\gamma\gamma}=4.35(14)$ keV [1].

3.1 $\pi^0\to e^+e^-\gamma$

The decay process $\pi^0\to e^+e^-\gamma$ was suggested for the first time by Dalitz in 1951 [62]. The first branching ratio prediction arose from a pure QED radiative corrections calculation neglecting the momentum dependence of the TFF [63]. These radiative corrections have been revisited recently in Ref. [64]. When our description for the differential decay rate distribution in terms of the dilepton invariant mass is compared to the QED point-like prediction, the agreement is found to be almost perfect. The reason is that the main contribution to this rate comes from the very low-energy part of the mass spectrum, where the effect of the TFF is negligible[1)]. Tiny differences appear solely in the high-energy part of the spectrum. Our result for the integrated branching ratio is 1.174(12)%, as compared to the QED point-like value 1.172%. Our prediction is in excellent agreement with the PDG reported value (1.174±0.035)% [1]. The main source of error we have quoted arises from the uncertainty associated with the low energy constants in Eq. (7). The error related to the measured two-photon decay width [61] is also included. Our result is in accord with several theoretical predictions existing in the literature [33, 35, 36, 70] and the QED estimates of Refs. [33, 65, 66]. Needless to say, the dimuon mode in the final state is kinematically forbidden in this case.

1) We arbitrarily define the very low-, low- and high-energy part of the spectrum as the part contributing 50%, 30% and 20%, respectively, to the integrated decay rate.

Table 1. Branching ratio predictions for $\eta\to l^+l^-\gamma$ ($l=e,\mu$) and their comparisons with experimental measurements and previous theoretical calculations.

source	$\mathcal{BR}(\eta\to e^+e^-\gamma)\times10^3$	$\mathcal{BR}(\eta\to\mu^+\mu^-\gamma)\times10^4$
QED	6.38	2.17
this work (P_1^5)	$6.60^{+0.50}_{-0.47}$	$3.25^{+0.40}_{-0.36}$
this work (P_2^2)	$6.61^{+0.53}_{-0.49}$	$3.30^{+0.65}_{-0.56}$
PDG [1]	6.9±0.4	3.1±0.4
H. Berghauser et al [3]	$6.6\pm0.4_{\rm stat}\pm0.4_{\rm syst}$	
WASA-at-COSY Coll. [6]	$6.72\pm0.07_{\rm stat}\pm0.31_{\rm syst}$	
QED [66]	6.38	2.18
FF 2,3 [34]	6.57	3.05
FF 4 [34]	6.53	2.87
LFQM [35]	6.95	2.94
hidden gauge [36]	6.57±0.03	3.05±0.04
modified VMD [36]	6.55±0.03	2.97±0.05

3.2 $\eta\to l^+l^-\gamma$ ($l=e,\mu$)

Due to the larger mass of the η as compared to the π^0, the dimuon mode in the final state is also accessible in this case. For that reason, these decays are more challenging for testing the momentum dependence of the TFF, and higher deviations from the QED point-like estimates are expected. This is precisely what our predictions reflect in Fig. 4, where the dilepton invariant mass distributions of $\eta\to e^+e^-\gamma$ (solid blue curve) and $\eta\to\mu^+\mu^-\gamma$ (solid black curve) are compared to the QED predictions (dotted and dashed grey curves, respectively). For the sake of comparison, we have employed a single-pole $P_1^5(q^2)$ PA in our prediction. The diagonal (asymptotically well-behaved) PA $P_2^2(q^2)$ produces a very similar description, as expected from the small differences between the two PA versions of the TFF in the whole kinematical range, as discussed in Section **2.2**. With one version or the other, the impact of the TFF is much bigger, in absolute terms, in the dimuon case, the reason being again a peaked distribution in the very-low energy region of the dielectron spectrum, which gives the most important contribution to the branching ratio, where the effect of the TFF is negligible. Notice that the high-energy parts of the spectra overlap, since the only difference between them is the dilepton production threshold. Once the distributions are integrated, we see from Table 1 that the branching ratio of the dimuon mode has been increased by 50% as compared to its prediction in QED, while the effect in the dielectron case is not sizeable (compatible with the QED calculation, within errors). The asymmetric errors shown in Table 1 arise from the error bands displayed in Fig. 2. Our predictions are in agreement with present experimental measurements. They also agree with previous theoretical calculations, such as the QED predictions in Ref. [66], the results in Ref. [34] using different form factors (FF) based on constraints from QCD and on experiments, the values in Ref. [35] from the light-front quark model (LFQM), and those obtained in Ref. [36] within the hidden gauge and the modified VMD models.

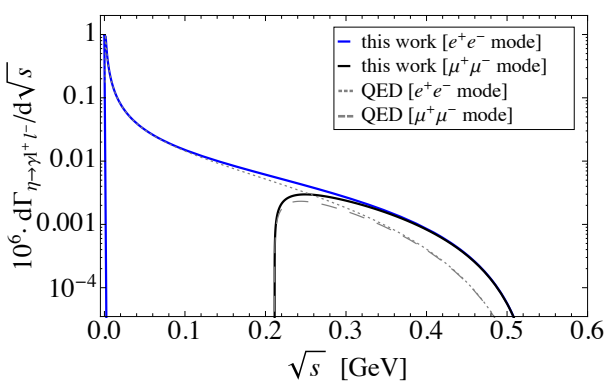

Fig. 4. (color online) Decay rate distribution for $\eta\to e^+e^-\gamma$ (solid blue curve) and $\eta\to\mu^+\mu^-\gamma$ (solid black curve). The corresponding QED estimates are also displayed (dotted and long-dashed grey curves, respectively).

3.3 $\eta'\to l^+l^-\gamma$ ($l=e,\mu$)

Due to the even larger mass of the η', the phase space is now enlarged by around 400 MeV and the associated TFF could be explored up to virtualities of the order of 1 GeV^2 if decay distributions of these processes were available. Unfortunately, this is not the case and predictions for these distributions must be provided. Our proposal for the η' TFF is the one discussed in Section **2.3**, which consists in using the PA results obtained from space-like data up to the matching point located at $\sqrt{s}=0.70$ GeV, then supplemented by a VMD description including the ρ, ω and ϕ vector resonances beyond that point. In Fig. 5, the dilepton invariant mass distributions of $\eta'\to e^+e^-\gamma$ (solid blue curve) and $\eta'\to\mu^+\mu^-\gamma$ (solid black curve) are compared to the corresponding

QED predictions (dotted and dashed grey curves, respectively). For such a comparison, we have employed for our prediction a single-pole $P_1^6(q^2)$ PA. The description achieved from the diagonal PA $P_1^1(q^2)$ is quite similar, as anticipated from the observation of Fig. 3. For the dielectron case, the decay distribution evidences again a marked peak at low energy which, despite the contribution coming from the resonance region, dominates the branching ratio, as occurred in $\pi^0(\eta)\to e^+e^-\gamma$. However, the effect of the TFF on the $\eta'\to\mu^+\mu^-\gamma$ decay distribution is larger than in $\eta\to\mu^+\mu^-\gamma$, increasing the branching ratio by a factor of 2. This is because of both a larger phase space and the effect of passing through a q^2 region where resonances can be produced on-shell. Interestingly, the contribution of the ρ resonance bends the distribution, while the inclusion of the ω resonance accounts for the sharp peak around 0.8 GeV. Numerical predictions for the branching ratios are presented in Table 2, where the errors shown come from the error

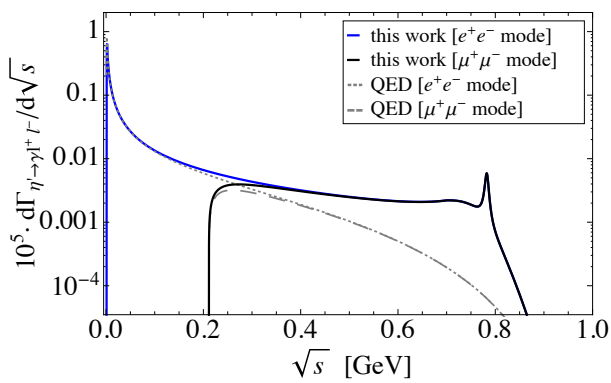

Fig. 5. (color online) Decay rate distribution for $\eta'\to e^+e^-\gamma$ (solid blue curve) and $\eta'\to\mu^+\mu^-\gamma$ (solid black curve). The corresponding QED estimates are also displayed (dotted and long-dashed grey curves, respectively).

Table 2. Branching ratio predictions for $\eta'\to l^+l^-\gamma$ ($l=e,\mu$) and their comparisons with experimental measurements and previous theoretical calculations

source	$\mathcal{BR}(\eta'\to e^+e^-\gamma)\times10^4$	$\mathcal{BR}(\eta'\to\mu^+\mu^-\gamma)\times10^4$
QED	3.94	0.38
this work (P_1^6)	$4.42^{+0.39}_{-0.35}$	$0.81^{+0.16}_{-0.13}$
this work (P_1^1)	$4.35^{+0.29}_{-0.27}$	0.74±0.06
PDG [1, 9]		1.08±0.27
BESIII Coll. [8]	$4.69\pm0.20_{\rm stat}\pm0.23_{\rm syst}$	
hidden gauge [36]	4.62±0.17	0.98±0.05
modified VMD [36]	4.53±0.17	0.90±0.05

bands linked to the TFF. Our predictions are in accordance with the theoretical calculations in Ref. [36], while they are slightly below the recent experimental measurement of $\eta'\to e^+e^-\gamma$ [8] and the old measurement of $\eta'\to\mu^+\mu^-\gamma$ [9], though in agreement within errors in both cases.

4 Double Dalitz decays

The double Dalitz decays involve the double-virtual TFF, as described in Section **2**. They involve four leptons in the final state, which makes the phase-space integration much more tedious. In the case of having two pairs of non-identical leptons, that is $\eta^{(\prime)}\to e^+e^-\mu^+\mu^-$, the required diagram is shown in Fig. 6(a) and the decay amplitude reads

$$\mathcal{A}_{\eta^{(\prime)}\to e^+e^-\mu^+\mu^-}=-i\frac{e^4}{q^2k^2}F_{\eta^{(\prime)}\gamma^*\gamma^*}(q^2,k^2)\times\varepsilon^{\alpha\beta\mu\nu}k_\alpha q_\beta\bar{u}(q_{e^-})\gamma_\mu v(q_{e^+})\bar{u}(q_{\mu^-})\gamma_\nu v(q_{\mu^+}),\quad(11)$$

where $q^2=(q_{e^+}+q_{e^-})^2\equiv M^2_{e^+e^-}$ and $k^2=(q_{\mu^+}+q_{\mu^-})^2\equiv M^2_{\mu^+\mu^-}$ are the dielectron and dimuon invariant masses, respectively. The corresponding differential decay rate can be reduced to[1)]

$$\frac{d^2\Gamma_{\eta^{(\prime)}\to e^+e^-\mu^+\mu^-}}{dM^2_{e^+e^-}M^2_{\mu^+\mu^-}\Gamma^{\rm exp}_{\eta^{(\prime)}\to\gamma\gamma}}=\mathcal{S}\frac{8\alpha^2}{9\pi^2m^6_{\eta^{(\prime)}}}\frac{1}{k^2q^2}\times\sqrt{1-\frac{4m_e^2}{k^2}}\sqrt{1-\frac{4m_\mu^2}{q^2}}\left(1+\frac{2m_e^2}{k^2}\right)\left(1+\frac{2m_\mu^2}{q^2}\right)\times\left\{\frac{1}{4}\left[m^2_{\eta^{(\prime)}}-(k^2+q^2)\right]^2-k^2q^2\right\}^{3/2}|\widetilde{F}(k^2,q^2)|^2\ ,\quad(12)$$

where, in this case, $\mathcal{S}=2$, in agreement with the expression given in Ref. [36]. In the case of having two pairs of identical leptons in the final state, that is $\mathcal{P}\to e^+e^-e^+e^-$ or $\eta^{(\prime)}\to\mu^+\mu^-\mu^+\mu^-$, one must instead consider both the direct (a) and exchange (b) diagrams in Fig. 6. In this case, the total amplitude of the process then reads

$$\mathcal{A}=\mathcal{A}_{\rm dir}-\mathcal{A}_{\rm exch}\ ,\quad(13)$$

where the minus sign appears due to the exchange of the two identical leptons. Squaring the former amplitude,

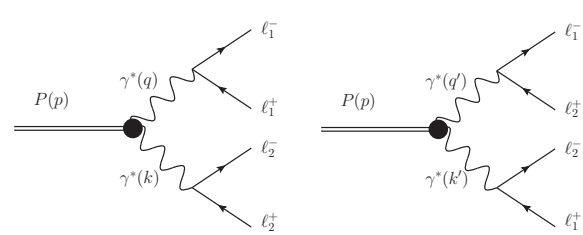

Fig. 6. Direct (a) and exchange (b) diagrams associated with double Dalitz decays.

1) See, for instance, Ref. [67] for reducing four-body final-state distributions into two invariant masses.

Table 3. Branching ratio predictions for $\pi^0\to e^+e^-e^+e^-$ and their comparisons with experimental measurements.

source	double-virtual TFF		$\mathcal{BR}(\pi^0\to e^+e^-e^+e^-)\times10^5$	
			direct+exchange	interference
this work	chisholm approximants	$b_{1,1}=0$	3.42937(9)	-0.03599
		$b_{1,1}=b_{1,0}^2$	3.42937(9)	-0.03599
		$b_{1,1}=2b_{1,0}^2$	3.42936(9)	-0.03599
	factorisation ansatz	Eq. (6)	3.42945(9)	-0.03599
	QED		3.41607	-0.03484
experimental measurements			3.38(16) [1]	
			3.46(19) [10]	

one arrives at

$$|\mathcal{A}|^2=|\mathcal{A}_{\rm dir}|^2+|\mathcal{A}_{\rm exch}|^2-2\Re(\mathcal{A}_{\rm dir}\mathcal{A}^*_{\rm exch}) \ , \qquad (14)$$

where not only the contributions from both the direct and exchange diagrams appear, but also that of the interference term. The contributions to the total partial decay width from the first and second terms of Eq. (14) are obviously the same, that is, $\Gamma_{\rm dir}=\Gamma_{\rm exch}$. In this way, the decay width $\Gamma_{\rm dir+exch}$ is obtained to be that of Eq. (12), now with $\mathcal{S}=1$, in accordance with Ref. [41], once the factor $\frac{1}{2!2!}$ accounting for the two pairs of identical particles in the final state and the sum over the two contributions have been taken into account. Regarding the interference term, a detailed expression in terms of five invariant masses, which requires a Monte Carlo (MC) simulation to be integrated, is relegated to the appendix.

4.1 $\pi^0\to e^+e^-e^+e^-$

The only possible double Dalitz decay of the neutral pion is $\pi^0\to e^+e^-e^+e^-$. Other possibilities are not kinematically allowed. In view of the results from $\pi^0\to e^+e^-\gamma$, one may expect that the overall effect of the TFF will be again small. In Fig. 7, we show our results for the different contributions to the decay rate distribution as a function of the invariant mass of one dielectron pair of the direct diagram. Concretely, we display the curve corresponding to the direct diagram (green solid line), the curve of the contribution of the exchange diagram expressed in terms of the former dielectron invariant mass of the direct diagram[1)] (dotted red line), the interference term (dotted blue line) and finally the total distribution (dotted black line). We want to note that the contribution from both direct and exchange diagrams integrates in the same way and that the interference is small and destructive. Our $\mathcal{BR}$ predictions are shown in Table 3, from which we corroborate that the effect of the TFF is small because the main contribution to the $\mathcal{BR}$ proceeds from the very low-momentum transferred region where a peak emerges, as already occurred in $\pi^0\to e^+e^-\gamma$. Our results are in good agreement with current experimental measurements. The source of the associated error comes from the uncertainty on the low-energy parameters in Eq. (7). Notice that the sensitivity of this decay to the variations of the double-virtual slope parameter, $b_{1,1}$, is at the fifth decimal place. In this sense, the high level of accuracy demanded to infer its value is unthinkable at the current level of precision. It is also interesting to compare with other authors' results. We are in good agreement with the results given in Refs. [36, 41, 53] for the direct and exchange contributions, while the result of Ref. [35] is about 5% lower than our predictions. Regarding the interference term, we have a perfect agreement with Refs. [36, 53] and a value about 30% higher than Ref. [41], while Ref. [35] did not consider this term. Comparing with previous QED estimates, we agree with Refs. [65, 66] for the direct and exchange contributions. For the interference term the former did not consider it and the latter gave a result 5 times larger than ours.

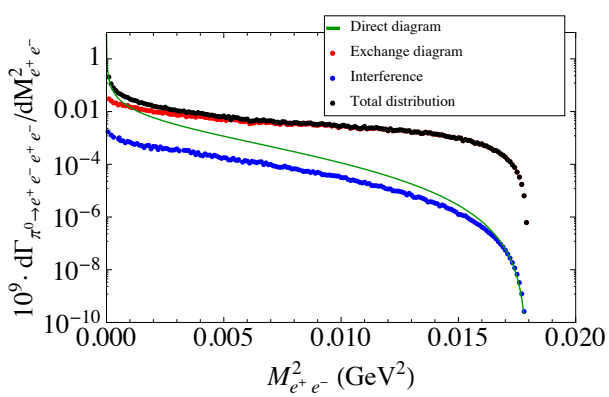

Fig. 7. (color online) Different contributions to the $\pi^0\to e^+e^-e^+e^-$ decay rate distribution as a function of one dielectron invariant mass of the direct diagram: direct diagram (solid green curve), exchange diagram (dotted red curve), interference term (dotted blue curve), and total distribution (dotted black curve).

1) The curve of the exchange diagram expressed in its own variables would look equal to the solid green line of Fig. 7. In this work, we have opted to show, in just one figure, all the contributions as a function of one dielectron invariant mass of the direct diagram. In this convention, the exchange diagram as expressed in Fig. 7 has also required a MC integration.

Table 4. Branching ratio predictions for $\eta\to e^+e^-\mu^+\mu^-$ and their comparison with the current experimental upper bound.

source	double-virtual TFF		$\mathcal{BR}(\eta\to e^+e^-\mu^+\mu^-)\times10^6$
this work	chisholm approximants	$b_{1,1}=0$	2.39(12)
		$b_{1,1}=b_{1,0}^2$	2.39(12)
		$b_{1,1}=2b_{1,0}^2$	2.38(12)
	factorisation ansatz	P_1^5	$2.35^{+0.47}_{-0.40}$
		P_2^2	$2.39^{+0.66}_{-0.53}$
	QED		1.57
experimental measurement			$<1.6\times10^{-4}$ (90% CL) [5]

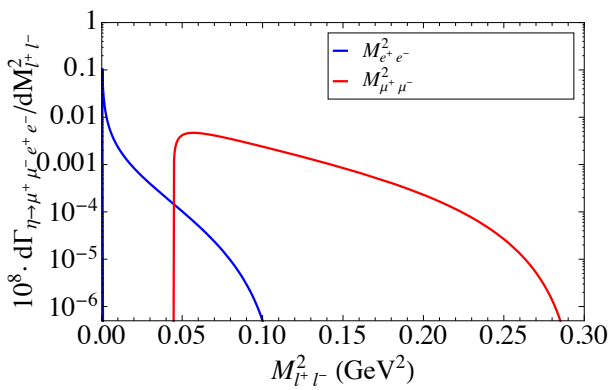

Fig. 8. (color online) Decay rate distribution for $\eta\to e^+e^-\mu^+\mu^-$ as a function of dielectron (blue curve) and dimuon (red curve) invariant mass.

4.2 $\eta\to l^+l^-l^+l^-$ ($l=e,\mu$)

The double Dalitz decays of the η, $\eta\to e^+e^-e^+e^-$, $\eta\to\mu^+\mu^-\mu^+\mu^-$ and $\eta\to e^+e^-\mu^+\mu^-$ processes are now kinematically allowed. Let us first analyse the latter, for simplicity. In this case, the two dilepton pairs are different and consequently there is no interference phenomenon. Hence, the distribution rate is just given by Eq. (12) and shown in Fig. 8 in two different manners, one expressed in terms of the dielectron invariant mass and the other in the dimuon variable (blue and red solid lines respectively), where, of course, both curves integrate the same. Our predictions are shown in Table 4, where the source of the associated errors comes from the error bands associated with the TFF for the case of the factorisation approach, and from the uncertainty on the single-virtual slope for the description employing CAs. From the experimental side, we respect the current upper limit, while from the theory side, because of the appearance of a dimuon pair in the final state, the effect of the TFF increases the $\mathcal{BR}$ about 50%, for the same arguments as explained for $\eta\to l^+l^-\gamma$. This decay, though much more sensitive than $\pi^0\to e^+e^-e^+e^-$ to the double-virtual slope, $b_{1,1}$, would require accurate measurements as well as demanding a very precise description of the TFF, in order to diminish its associated error, for deducing $b_{1,1}$, far from the present situation. Comparing with other authors, we agree with the predictions of Ref. [36], while we have found discrepancies with the value 5.83×10^{-7} of Ref. [35], with the prediction 2×10^{-7} of Ref. [34] and with the estimate 7.84×10^{-7} of Ref. [66]. In the latter case, the reason seems to be a typographical fault of a factor of 2 missing, as pointed out in both Refs. [36, 53]. In such a case, it would reproduce the QED result of Table 4 as it should be, because they did not consider the momentum dependence of the TFF.

The decays involving two identical dilepton pairs in the final state, $\eta\to e^+e^-e^+e^-$ and $\eta\to\mu^+\mu^-\mu^+\mu^-$, require us to consider Eq. (14). Their distributions are given in Fig. 9 ((a) and (b) respectively) as a function of one dilepton invariant mass of the direct diagram. We explicitly show the contribution from the direct diagram (solid green curve), the curve of the exchange diagram expressed in terms of the former dielectron (dimuon) invariant mass of the direct diagram (dotted red curve), the interference term (dotted blue curve) and the total decay rate distribution (dotted black line). The integrated $\mathcal{BR}$ results are shown in Table 5, where the error comes again from the error bands of the TFF description as given in Fig. 2, for the factorisation approach, and from the uncertainty associated with the slope, b_η, for the description using CAs. Comparing with present experimental status, our prediction for $\eta\to e^+e^-e^+e^-$ is compatible at less than 1σ with the KLOE measurement [11], as well as with the recent measurement value of the WASA@COSY collaboration [6], while our estimate for $\eta\to\mu^+\mu^-\mu^+\mu^-$ respects the current upper bound of Ref. [5]. We have found the same trend as in $\eta\to l^+l^-\gamma$, that is, while the overall effect of the TFF on the electronic mode is small, increasing the $\mathcal{BR}$ of $\eta\to e^+e^-e^+e^-$ by 6% with respect to the QED estimate, the impact on the muonic channel, $\eta\to\mu^+\mu^-\mu^+\mu^-$, becomes important, increasing the $\mathcal{BR}$ by a factor ranging (1.6–1.7) with respect to the QED calculation. As a consequence, the sizeable sensitivity of $\eta\to\mu^+\mu^-\mu^+\mu^-$ to the TFF of double virtuality makes it a good candidate to improve our knowledge of it. Interestingly, a precise experimental measurement of this mode at the percent level of precision leaves us in a position to estimate the value of $b_{1,1}$. For that purpose, it is also required to diminish the associated uncertainty with

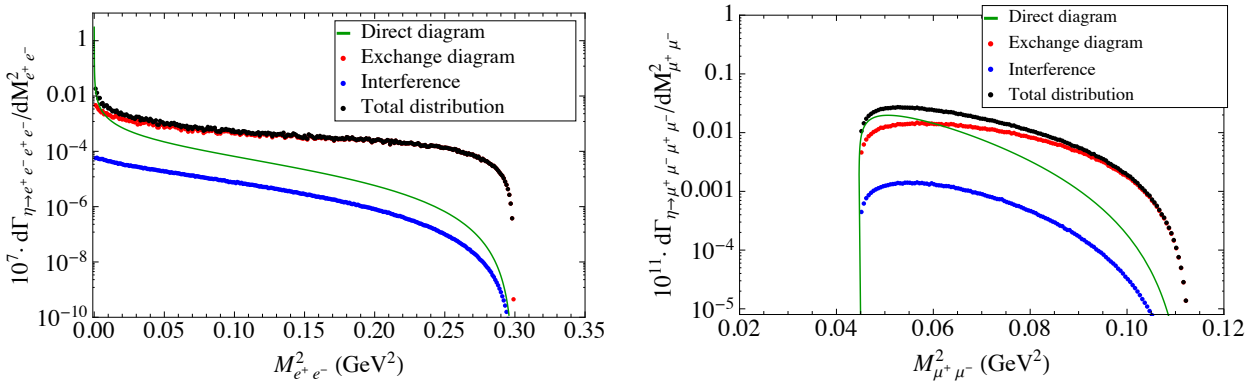

Fig. 9. (color online) Different contributions to the $\eta\to e^+e^-e^+e^-$ (a) and $\eta\to\mu^+\mu^-\mu^+\mu^-$ (b) decay rate distributions as a function of one dielectron invariant mass of the direct diagram: direct diagram (solid green curve), exchange diagram (dotted red curve), interference term (dotted blue curve), and total distribution (dotted black curve).

Table 5. Branching ratio predictions for $\eta\to e^+e^-e^+e^-$ and $\eta\to\mu^+\mu^-\mu^+\mu^-$ and their comparisons with experimental measurements.

source	double-virtual TFF		$\mathcal{BR}(\eta\to e^+e^-e^+e^-)\times10^5$		$\mathcal{BR}(\eta\to\mu^+\mu^-\mu^+\mu^-)\times10^9$	
			dir+exch	interf	dir+exch	interf
this work	chisholm approximants	$b_{1,1}=0$	2.74(3)	-0.02	4.47(33)	-0.32
		$b_{1,1}=b^2_{1,0}$	2.73(3)	-0.03	4.31(31)	-0.32
		$b_{1,1}=2b^2_{1,0}$	2.73(3)	-0.03	4.15(30)	-0.32
	factorisation ansatz	P^5_1	$2.72^{+0.43}_{-0.37}$	-0.03	$4.23^{+0.83}_{-0.70}$	-0.43
		P^2_2	$2.73^{+0.45}_{-0.39}$	-0.03	$4.30^{+1.12}_{-0.91}$	-0.47
	QED		2.56	-0.02	2.59	-0.19
experimental measurements			$3.2(9)_{\rm stat}(5)_{\rm sys}$ [6] $2.4(2)_{\rm stat}(1)_{\rm sys}$ [11]		$<3.6\times10^{-4}$ (90% CL) [5]	

the TFF. Here enters the ability of the Padé method we use for accommodating new experimental data as soon as released by experimental groups. However, the same exercise for $\eta\to e^+e^-e^+e^-$ would demand accurate measurements at the per mille level to unveil this quantity, far from the present situation. Our predictions are in good agreement with the results of Ref. [36] for the electronic mode, while (10–15)% over the muonic prediction. Comparing with Ref. [35] (which did not considered the interference term) we are 10% over for the electronic case, while his result for $\eta\to\mu^+\mu^-\mu^+\mu^-$ is 60% smaller. We are also in accordance with the estimate of Ref. [34] for $\eta\to e^+e^-e^+e^-$. Regarding the pure QED calculation of Ref. [66], we are in perfect agreement for the electronic channel while tiny differences are found in the muonic decay, probably caused by the updated values of our input values.

4.3 $\eta'\to l^+l^-l^+l^-$ ($l=e,\mu$)

Regarding the double Dalitz decays of the η', we have the same three possible final states as for the η. However, in this case we have only adopted the factorisation approach ansatz for describing the double-virtual TFF of the η'. The reason is because the use of Chisholm approximants, which may respect the appropriate asymptotic behaviour q^{-2}, would only apply at low energies, concretely up to the matching point where PAs are applicable, while beyond, we are somehow forced to employ the factorisation approximation, through a VMD description, which induces a q^{-4} term. So, there is no gain respecting the high-energy behaviour in the low-energy region if we violate it at high energies. We compute first $\eta'\to e^+e^-\mu^+\mu^-$, again through Eq. (12). Noticeably, it follows the same trend as $\eta\to e^+e^-\mu^+\mu^-$, with the difference that this case is sensitive to the resonance region, as can be read off from Fig. 10. Once more, the low-momentum region basically dominates the distribution when working with the electronic variable (solid blue curve), while it is a smooth falling function of the dimuonic momentum with a small bump and a sharp peak accounting for the effect of the ρ and the ω, respectively (solid red curve). Both curves integrate to the same $\mathcal{BR}$. Our predictions are presented in Table 6 without, in this case, any experimental reference to compare with. The effect of the TFF increases by a factor of about 2 the $\mathcal{BR}$ respect to the QED estimate, which is much notorious than in $\eta\to e^+e^-\mu^+\mu^-$.

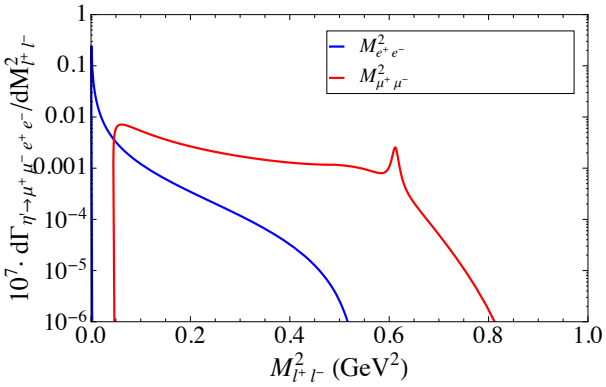

Fig. 10. (color online) Decay rate distribution for $\eta'\to e^+e^-\mu^+\mu^-$ as a function of dielectron (blue curve) and dimuon (red curve) invariant mass.

Table 6. Branching ratio predictions for $\eta'\to e^+e^-\mu^+\mu^-$.

source	$\mathcal{BR}(\eta'\to e^+e^-\mu^+\mu^-)\times10^7$
this work (P_1^6)	$6.80^{+1.39}_{-1.17}$
this work (P_1^1)	$6.25^{+0.83}_{-0.72}$
QED	3.21
Exp. measurements	not seen

The decay spectra for $\eta'\to e^+e^-e^+e^-$ and $\eta'\to\mu^+\mu^-\mu^+\mu^-$ shown in Fig. 11 ((a) and (b) respectively) have been computed by taking Eq. (14) into account. We have represented the contributions of the direct diagram (solid green curve), the exchange diagram expressed in terms of the variable of the direct diagram (dotted red curve), the interference term (dotted blue curve) and lastly the total distribution (dotted black line). One interesting feature concerning phase space is that the electronic mode (Fig. 11(a)) is clearly sensitive to the intermediate vector resonances, while the muonic (Fig. 11(b)) is basically not. Our predictions are presented in Table 7, and also reflect the tendency that the effect of the TFF is sizeable and larger than for the case of the η. In particular, the $\mathcal{BR}$ of $\eta'\to e^+e^-e^+e^-$ and $\eta'\to\mu^+\mu^-\mu^+\mu^-$ have increased by 20% and by a factor of 2, respectively. On the experimental side, we have no observations to compare with, while on the theory side we have only found the predictions given in Ref. [36] with which we are in good agreement for the cases with two identical dilepton pairs in the final state, while we are slightly below for $\eta'\to\mu^+\mu^-e^+e^-$.

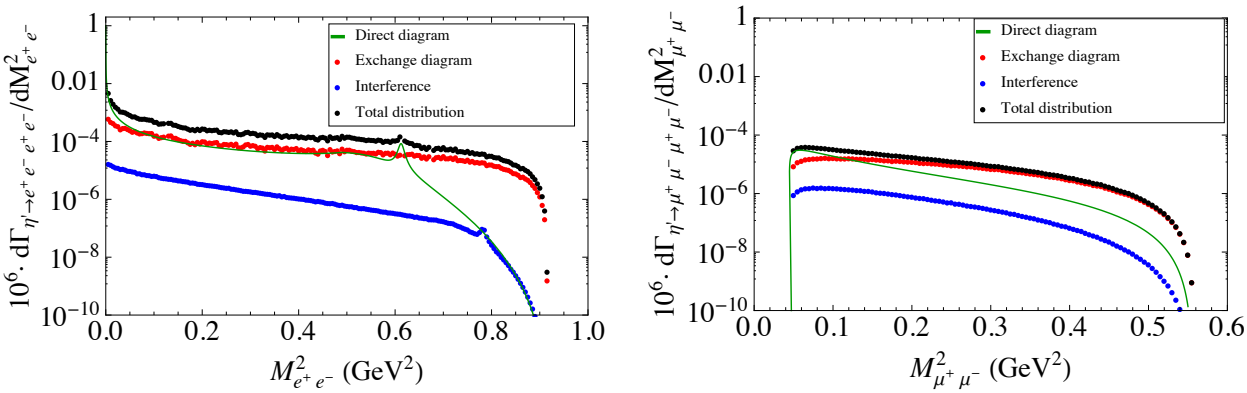

Fig. 11. (color online) Different contributions to the $\eta'\to e^+e^-e^+e^-$ (a) and $\eta'\to\mu^+\mu^-\mu^+\mu^-$ (b) decay rate distributions as a function of one dielectron invariant mass of the direct diagram: direct diagram (solid green curve), exchange diagram (dotted red curve), interference term (dotted blue curve), and total distribution (dotted black curve).

Table 7. Branching ratio predictions for $\eta'\to e^+e^-e^+e^-$ and $\eta'\to\mu^+\mu^-\mu^+\mu^-$.

source	double-virtual TFF		$\mathcal{BR}(\eta'\to e^+e^-e^+e^-)\times10^6$		$\mathcal{BR}(\eta'\to\mu^+\mu^-\mu^+\mu^-)\times10^8$	
			dir+exch	interf	dir+exch	interf
this work	factorisation ansatz	P_1^6	$2.15^{+0.35}_{-0.29}$	-0.03	$2.19^{+0.23}_{-0.19}$	-0.44
		P_1^1	$2.09^{+0.28}_{-0.24}$	-0.01	$2.06^{+0.17}_{-0.15}$	-0.41
	QED		1.75	-0.01	0.98	-0.11
experimental measurements			not seen		not seen	

5 Conclusions

The single and double Dalitz decays $\mathcal{P}\to l^+l^-\gamma$ and $\mathcal{P}\to l^+l^-l^+l^-$ ($\mathcal{P}=\pi^0,\eta,\eta'$; l=e or μ) have been analysed by means of a data-driven model-independent description of the transition $\mathcal{P}\to\gamma^{(*)}\gamma^*$. We have benefited from (our) previous findings on the space-like single-virtual TFF $\gamma\gamma^*\to\mathcal{P}$ obtained through the use of Padé approximants to represent these transitions in the time-like energy region where they are applicable. We have shown that this extrapolation from the space-like to the time-like is supported by current experimental data η and η' TFFs obtained from $\eta^{(\prime)}\to e^+e^-\gamma$ and $\eta\to\mu^+\mu^-\gamma$ decays. This nice behaviour proves that these TFFs are well approximated by meromorphic functions. Regarding the TFF of double virtuality, besides the standard factorisation approach, we have motivated the use of bivariate approximants, which would satisfy the high-energy constraints and whose coefficients may be determined as soon as experimental data become available. From the phenomenological point of view, we have found that the invariant mass distributions involving electrons in the final state show strong peaks in the very low-momentum transfer region, which mainly dominate the contribution to the branching ratios, hence suppressing the effect of the TFFs. However, distributions involving muons in the final state are much more homogeneously distributed and clearly manifest the neat effect of the TFF, which, in particular, is enhanced for the η' decays due to phase-space considerations. Our final branching ratio predictions are summarised in Table 8, where a combined weighted average[1)] of the results given in the different tables has been considered and the uncertainties symmetrised. The values of n_σ shown in the table give the number of standard deviations the experimental measurements are from our predictions. All these predictions are seen to be in accordance with present experimental measurements; only $\eta'\to\mu^+\mu^-\gamma$ and $\eta\to e^+e^-e^+e^-$ appear slightly in tension. It is worth stating that our results for the η and η' decays are independent of η-η' mixing effects, the reason being that all the PAs used to fit the space-like TFFs are fixed at zero momentum transfer by the corresponding experimental two-photon decay widths and not by the axial-anomaly predictions in terms of the mixing parameters. Similarly, when diagonal PAs were used, a constant behaviour for $Q^2F(Q^2)$ was imposed at $Q^2\to\infty$, as predicted by perturbative QCD, but without fixing the associated constants in terms of these mixing parameters. To finish, we would like once more to encourage experimental groups to measure these TFFs.

Table 8. Central final branching ratio predictions as a combined weighted average of the results presented. Errors are symmetrised. n_σ stands for the number of standard deviations the measured results are from our predictions.

decay	this work	experimental value [1]	n_σ
$\pi^0\to e^+e^-\gamma$	1.169(1)%	1.174(35)%	0.15
$\eta\to e^+e^-\gamma$	$6.61(50)\times10^{-3}$	$6.90(40)\times10^{-3}$	0.45
$\eta\to\mu^+\mu^-\gamma$	$3.26(46)\times10^{-4}$	$3.1(4)\times10^{-4}$	0.26
$\eta'\to e^+e^-\gamma$	$4.38(32)\times10^{-4}$	$4.69(20)(23)\times10^{-4}$	0.70
$\eta'\to\mu^+\mu^-\gamma$	$0.75(6)\times10^{-4}$	$1.08(27)\times10^{-4}$	1.19
$\pi^0\to e^+e^-e^+e^-$	$3.36689(5)\times10^{-5}$	$3.34(16)\times10^{-5}$	0.17
$\eta\to e^+e^-e^+e^-$	$2.71(2)\times10^{-5}$	$2.4(2)(1)\times10^{-5}$	1.38
$\eta\to\mu^+\mu^-\mu^+\mu^-$	$3.98(15)\times10^{-9}$	$<3.6\times10^{-4}$	
$\eta\to e^+e^-\mu^+\mu^-$	$2.39(7)\times10^{-6}$	$<1.6\times10^{-4}$	
$\eta'\to e^+e^-e^+e^-$	$2.10(45)\times10^{-6}$	not seen	
$\eta'\to\mu^+\mu^-\mu^+\mu^-$	$1.69(36)\times10^{-8}$	not seen	
$\eta'\to e^+e^-\mu^+\mu^-$	$6.39(91)\times10^{-7}$	not seen	

We thank very much Pere Masjuan, Santi Peris, Pablo Roig and Pablo Sánchez Puertas for fruitful discussions on this topic. We are grateful to the NA60 Collaboration, in particular to Sanja Damjanovic, for providing us with the transition form factor data obtained from $\eta\to\mu^+\mu^-\gamma$. S. González-Solís would also like to thank the Institut für Kernphysik of the Johannes Gutenberg-Universität in Mainz for hospitality.

Appendix: Interference term

Four-body decay width in invariant variables

The partial decay width of a particle P of mass M_P decaying into four particles $p_1p_2p_3p_4$ reads [1]

$$\Gamma(P\to p_1p_2p_3p_4) = \int \mathrm{d}\Phi(p_P;q_1,q_2,q_3,q_4) \times\frac{(2\pi)^4}{2M_P}\overline{\left|\mathcal{M}(P\to p_1p_2p_3p_4)\right|^2}, \quad \text{(A1)}$$

where $\mathrm{d}\Phi(p_P;q_1,q_2,q_3,q_4)$ is the four-body phase-space element given by

$$\mathrm{d}\Phi(p_P;q_1,q_2,q_3,q_4)=\delta^4\left(p_P-\sum_{i=1}^4 q_i\right)\prod_{i=1}^4\frac{d^3\boldsymbol{q}_i}{(2\pi)^3 2E_i}. \quad \text{(A2)}$$

Following Refs. [68, 69], the phase space is expressed in terms of independent invariant masses (instead of using three-

1) For the combined weighted average we follow the standard procedure for unconstrained averaging described in Ref. [1], that is to say, $\overline{\mathcal{BR}}\pm\delta\overline{\mathcal{BR}}=\left(\sum_i\mathcal{BR}_i/(\delta\mathcal{BR}_i)^2/\sum_i 1/(\delta\mathcal{BR}_i)^2\right)\pm\left(\sum_i 1/(\delta\mathcal{BR}_i)^2\right)^{-1/2}$, where $\mathcal{BR}_i$ and $\delta\mathcal{BR}_i$ are the values and errors displayed in the different tables for the same particular branching ratio.

momenta and angles) as

$$\begin{aligned}&\mathrm{d}\Phi(p_P;q_1,q_2,q_3,q_4)\\&=\frac{1}{8\pi^{10}M_P^2}(-B)^{-1/2}\mathrm{d}M_{12}^2\mathrm{d}M_{34}^2\mathrm{d}M_{14}^2\mathrm{d}M_{124}^2\mathrm{d}M_{134}^2,\end{aligned}\tag{A3}$$

where $M_{ij}^2=(q_i+q_j)^2$ and $M_{ijk}^2=(q_i+q_j+q_k)^2$. In the cases that concern us, that is, the interference term for $\pi^0\to e^+e^-e^+e^-$ and $\eta^{(\prime)}\to l^+l^-l^+l^-$ (l=e,μ), B reads

$$\begin{aligned}B\;=\;&m_l^8+\Big[M_{124}^2\left(M_{134}^2-M_{34}^2\right)+M_{14}^2\left(-M_P^2+M_{34}^2\right)\Big]^2\\&+M_{12}^2\Big\{M_{12}^2\left(M_{134}^2-M_{14}^2\right)^2\\&-2\left(M_{134}^2-M_{14}^2\right)\left(M_{124}^2M_{134}^2-M_P^2M_{14}^2\right)\\&+2\Big[\left(-2M_P^2+M_{134}^2-M_{14}^2\right)M_{14}^2\\&+M_{124}^2\left(M_{134}^2+M_{14}^2\right)\Big]M_{34}^2\Big\}\\&+m_l^4\Big\{M_{12}^4+M_{124}^4+4M_{124}^2M_{134}^2+M_{134}^4\\&-2M_P^2\Big[2\left(M_{124}^2+M_{134}^2\right)+M_{14}^2\Big]\\&+2\left(4M_{124}^2+M_{134}^2+M_{14}^2\right)M_{34}^2+M_{34}^4\\&+2M_{12}^2\left(M_{124}^2+4M_{134}^2+M_{14}^2+M_{34}^2\right)\Big\}\\&+m_l^6\Big[4M_P^2-2\left(3M_{12}^2+M_{124}^2+M_{134}^2+3M_{34}^2\right)\Big]\\&-2m_l^2\Big\{M_{12}^4\left(M_{134}^2+M_{14}^2\right)\\&+\left(2M_P^2-M_{124}^2-M_{134}^2\right)\left(-M_{124}^2M_{134}^2+M_P^2M_{14}^2\right)\\&+\Big[M_{124}^4+\left(-3M_P^2+M_{124}^2+M_{134}^2\right)M_{14}^2\Big]M_{34}^2\\&+\left(M_{124}^2+M_{14}^2\right)M_{34}^4\\&+M_{12}^2\Big[M_{34}^4-3M_P^2M_{14}^2+M_{124}^2M_{14}^2+M_{134}^2M_{14}^2\\&+\left(M_{124}^2+M_{134}^2-2M_{14}^2\right)M_{34}^2\Big]\Big\},\end{aligned}\tag{A4}$$

where m_l is the lepton mass and the boundary of the physical allowed region fulfils $B=0$. Reference [68] points out that the choice of variables $M_{12}^2,M_{34}^2,M_{14}^2,M_{124}^2$, and M_{134}^2 is convenient to facilitate the finding of the integration limits of B, since it only depends quadratically on each of the variables. Other choices can lead to quartics.

Integration limits

In order to find the physical region of one variable, for instance M_{14}^2, one must solve $B=0$, obtaining

$$\begin{aligned}M_{14}^{2\pm}\;=\;&\frac{1}{\lambda(M_{12}^2,M_{34}^2,M_P^2)}\\&\times\Big[-b\pm2\sqrt{G(M_{124}^2,M_{34}^2,M_{12}^2,m_l^2,M_P^2,m_l^2)}\\&\times\sqrt{G(M_{134}^2,M_{34}^2,M_{12}^2,m_l^2,M_P^2,m_l^2)}\Big],\end{aligned}\tag{A5}$$

where $\lambda(a,b,c)=a^2+b^2+c^2-2ab-2ac-2bc$ is the basic two-particle kinematical function, the Källen function, b is given by

$$\begin{aligned}b\;=\;&(M_{34}^2-M_{12}^2)\Big[M_{34}^2(M_{124}^2+m_l^2)\\&-(M_{124}^2-m_l^2)(M_{134}^2-m_l^2)-2m_l^4\Big]\\&-M_{12}^2\Big[(M_{124}^2-m_l^2)(M_{134}^2-m_l^2)\\&+M_{34}^2(M_{124}^2+M_{134}^2+2m_l^2-2M_P^2)\\&+(M_{134}^2+3m_l^2)M_P^2\Big]+(M_{12}^2)^2(M_{134}^2+m_l^2),\end{aligned}\tag{A6}$$

and

$$\begin{aligned}G(x,y,z,u,v,w)\;=\;&u^2z-uvw\\&+uvx-uvz+uwy-uwz-uxy-uxz-uyz\\&+uz^2+v^2w+vw^2-vwx-vwy-vwz-vxy\\&+vyz-wxy+wxz+x^2y+xy^2-xyz,\end{aligned}\tag{A7}$$

is the basic four-particle kinematic function. As argued in Ref. [68], the integration limits of the remaining variables, $M_{12}^2,M_{34}^2,M_{124}^2$, and M_{134}^2 are obtained after solving

$$\begin{aligned}G(M_{124}^2,M_{34}^2,M_{12}^2,m_l^2,M_P^2,m_l^2)\;&=\;0,\\G(M_{134}^2,M_{34}^2,M_{12}^2,m_l^2,M_P^2,m_l^2)\;&=\;0,\end{aligned}\tag{A8}$$

while the dilepton invariant masses M_{12}^2 and M_{34}^2 range from threshold $4m_l^2$ to $(M_P-2m_l)^2$ and $4m_l^2$ to $(M_P-M_{12})^2$, respectively.

Matrix element of the interference term

The last term in Eq. (14) reads

$$\begin{aligned}\mathcal{A}_1\mathcal{A}_2\;=\;&\frac{e^8|F(q^2,k^2)||F(q'^2,k'^2)|}{q^2k^2q'^2k'^2}\varepsilon^{\mu\nu\alpha\beta}\varepsilon^{\mu'\nu'\alpha'\beta'}\\&\times(q_1+q_2)_\mu(q_3+q_4)_\nu(q_1+q_4)_{\mu'}(q_2+q_3)_{\nu'}\\&\times\mathrm{Tr}[(\not{q}_1+m_l)\gamma_\alpha(\not{q}_2-m_l)\gamma_{\beta'}(\not{q}_3+m_l)\gamma_\beta(\not{q}_4-m_l)\gamma_{\alpha'}].\end{aligned}\tag{A9}$$

The trace and the corresponding contractions with both the product of Levi-Civita tensors and the different diphoton four momenta in Eq. (A9) have been computed with FormCalc. To give a result in the chosen variables, one needs to perform some replacements in the former equation:

i) $q_i^2=m_l^2$;

ii) $M_{23}^2=2(m_l^2+q_2\cdot q_3)$;

iii) $q_2\cdot q_3=\frac{1}{2}(M_P^2-4m_l^2)-q_1\cdot q_2-q_1\cdot q_3-q_1\cdot q_4-q_2\cdot q_4-q_3\cdot q_4$;

iv) $q_1\cdot q_3=\frac{1}{2}(M_{134}^2-3m_l^2)-q_1\cdot q_4-q_3\cdot q_4$;

v) $q_2\cdot q_4=\frac{1}{2}(M_{124}^2-3m_l^2)-q_1\cdot q_2-q_1\cdot q_4$;

vi) $q_1\cdot q_2=\frac{1}{2}M_{12}^2-m_l^2,\;\;q_1\cdot q_4=\frac{1}{2}M_{14}^2-m_l^2$,

$q_3\cdot q_4=\frac{1}{2}M_{34}^2-m_l^2$. Finally, the expression for the interference term in Eq. (A9) reads

$$
\begin{aligned}
&\frac{e^8|F(M_{12}^2,M_{34}^2)||F(M_{14}^2,M_{23}^2)|}{M_{12}^2M_{34}^2M_{14}^2(2m_l^2+M_P^2-M_{124}^2-M_{134}^2+M_{14}^2)^2}\Big\{-2m_l^8+M_{12}^6\left(M_{134}^2-M_{14}^2\right)+4m_l^6\left(-M_{12}^2+M_{124}^2+M_{134}^2-M_{34}^2\right)\\
&+m_l^4\Big[M_{12}^4-3\left(M_{124}^2+M_{134}^2\right)^2+4M_P^2M_{14}^2+2M_{12}^2\left(M_{124}^2+5M_{134}^2-2M_{14}^2-5M_{34}^2\right)+2\left(5M_{124}^2+M_{134}^2-2M_{14}^2\right)M_{34}^2+M_{34}^4\Big]\\
&-\left(M_{124}^4+M_{134}^4-2M_P^2M_{14}^2-2M_{134}^2M_{34}^2+2M_{14}^2M_{34}^2+M_{34}^4\right)\times\Big[M_{124}^2\left(M_{134}^2-M_{34}^2\right)+M_{14}^2\left(-M_P^2+M_{34}^2\right)\Big]\\
&+M_{12}^4\Big[\left(M_P^2+2M_{134}^2-2M_{14}^2\right)M_{14}^2+\left(-2M_{134}^2+M_{14}^2\right)M_{34}^2+M_{124}^2\left(-3M_{134}^2+2M_{14}^2+M_{34}^2\right)\Big]\\
&+m_l^2\Big[-M_{12}^6+\left(M_{124}^2+M_{134}^2\right)^3-\left(M_{124}^2+M_{134}^2\right)\times\left(5M_{124}^2+M_{134}^2-4M_{14}^2\right)M_{34}^2+\left(M_{124}^2+M_{134}^2-4M_{14}^2\right)M_{34}^4\\
&-M_{34}^6+M_{12}^4\left(M_{124}^2+M_{134}^2-4M_{14}^2+M_{34}^2\right)+M_{12}^2\Big\{-\left(M_{124}^2+M_{134}^2\right)\left(M_{124}^2+5M_{134}^2-4M_{14}^2\right)+2M_P^2\left(M_{124}^2-M_{134}^2+2M_{14}^2\right)\\
&+2\Big[3\left(M_{124}^2+M_{134}^2\right)+4M_{14}^2\Big]M_{34}^2+M_{34}^4\Big\}-2M_P^2\Big[M_{124}^4+\left(M_{134}^2+2M_{14}^2\right)\left(M_{134}^2-M_{34}^2\right)+M_{124}^2\left(-2M_{134}^2+2M_{14}^2+M_{34}^2\right)\Big]\Big]\\
&+M_{12}^2\Big\{M_{134}^6+M_{124}^4\left(3M_{134}^2-M_{14}^2-2M_{34}^2\right)-M_{134}^4\left(M_{14}^2+2M_{34}^2\right)-2M_{124}^2\left(M_P^2M_{14}^2+M_{134}^2M_{14}^2-M_{134}^2M_{34}^2+2M_{14}^2M_{34}^2+M_{34}^4\right)\\
&+M_{134}^2\Big[-2M_P^2M_{14}^2+M_{34}^2\left(-4M_{14}^2+M_{34}^2\right)\Big]+M_{14}^2\Big[M_{34}^2\left(4M_{14}^2+M_{34}^2\right)+M_P^2\left(4M_{14}^2+6M_{34}^2\right)\Big]\Big\}\Big\}. \qquad \text{(A10)}
\end{aligned}
$$

References

1 C. Patrignani et al (Particle Data Group), Chin. Phys. C, **40**: 100001 (2016)
2 A. Beddall and A. Beddall, Eur. Phys. J. C, **54**: 365 (2008)
3 H. Berghauser et al, Phys. Lett. B, **701**: 562 (2011)
4 P. Aguar-Bartolome et al (A2 Collaboration), Phys. Rev. C, **89**: 044608 (2014) [arXiv:1309.5648 [hep-ex]]
5 M. Berlowski et al (CELSIUS/WASA Collaboration), Phys. Rev. D, **77**: 032004 (2008) [arXiv:0711.3531 [hep-ex]]
6 P. Adlarson et al, Phys. Rev. C, **94**: 065206 (2016) [arXiv:1509.06588 [nucl-ex]]
7 R. Arnaldi et al (NA60 Collaboration), Phys. Lett. B, **677**: 260 (2009) [arXiv:0902.2547 [hep-ph]]
8 M. Ablikim et al (BESIII Collaboration), Phys. Rev. D, **92**: 012001 (2015) [arXiv:1504.06016 [hep-ex]]
9 R. I. Dzhelyadin et al, Sov. J. Nucl. Phys., **32**: 520 (1980) [Yad. Fiz., **32**: 1005 (1980)]
10 E. Abouzaid et al (KTeV Collaboration), Phys. Rev. Lett., **100**: 182001 (2008) [arXiv:0802.2064 [hep-ex]]
11 F. Ambrosino et al (KLOE and KLOE-2 Collaborations), Phys. Lett. B, **702**: 324 (2014) [arXiv:1105.6067 [hep-ex]]
12 G. P. Lepage and S. J. Brodsky, Phys. Rev. D, **22**: 2157 (1980)
13 A. V. Manohar, Phys. Lett. B, **244**: 101 (1990)
14 J. M. Gerard and T. Lahna, Phys. Lett. B, **356**: 381 (1995) [hep-ph/9506255]
15 Z. H. Guo and P. Roig, Phys. Rev. D, **82**: 113016 (2010) [arXiv:1009.2542 [hep-ph]]
16 P. Roig and J. J. Sanz-Cillero, Nucl. Part. Phys. Proc., **258-259**: 98 (2015) [arXiv:1407.3192 [hep-ph]]
17 R. Escribano, P. Masjuan, and P. Sanchez-Puertas, Phys. Rev. D, **89**: 034014 (2014) [arXiv:1307.2061 [hep-ph]]
18 R. Escribano, P. Masjuan, and P. Sanchez-Puertas, Eur. Phys. J. C, **75**: 414 (2015) [arXiv:1504.07742 [hep-ph]]
19 V. A. Novikov, M. A. Shifman, A. I. Vainshtein, M. B. Voloshin, and V. I. Zakharov, Nucl. Phys. B, **237**: 525 (1984)
20 P. Masjuan, S. Peris, and J. J. Sanz-Cillero, Phys. Rev. D, **78**: 074028 (2008) [arXiv:0807.4893 [hep-ph]]
21 P. Masjuan, Phys. Rev. D, **86**: 094021 (2012) [arXiv:1206.2549 [hep-ph]]
22 H. J. Behrend et al (CELLO Collaboration), Z. Phys. C, **49**: 401 (1991)
23 J. Gronberg et al (CLEO Collaboration), Phys. Rev. D, **57**: 33 (1998) [hep-ex/9707031]
24 M. Acciarri et al (L3 Collaboration), Phys. Lett. B, **418**: 399 (1998)
25 B. Aubert et al (BaBar Collaboration), Phys. Rev. D, **80**: 052002 (2009) [arXiv:0905.4778 [hep-ex]]
26 P. del Amo Sanchez et al (BaBar Collaboration), Phys. Rev. D, **84**: 052001 (2011) [arXiv:1101.1142 [hep-ex]]
27 S. Uehara et al (Belle Collaboration), Phys. Rev. D, **86**: 092007 (2012) [arXiv:1205.3249 [hep-ex]]
28 G. P. Lepage and S. J. Brodsky, Phys. Lett. B, **87**: 359 (1979)
29 T. Feldmann and P. Kroll, Phys. Rev. D, **58**: 057501 (1998) [hep-ph/9805294]
30 H. Czyz, S. Ivashyn, A. Korchin, and O. Shekhovtsova, Phys. Rev. D, **85**: 094010 (2012) [arXiv:1202.1171 [hep-ph]]
31 P. Roig, A. Guevara, and G. López Castro, Phys. Rev. D, **89**: 073016 (2014) [arXiv:1401.4099 [hep-ph]]
32 C. Hanhart, A. Kupśc, U.-G. Meissner, F. Stollenwerk, and A. Wirzba, Eur. Phys. J. C, **73**: 2668 (2013) [Eur. Phys. J. C, **75**: 242 (2015)] [arXiv:1307.5654 [hep-ph]].
33 M. Hoferichter, B. Kubis, S. Leupold, F. Niecknig, and S. P. Schneider, Eur. Phys. J. C, **74**: 3180 (2014) [arXiv:1410.4691 [hep-ph]]
34 F. Persson, hep-ph/0106130
35 C. C. Lih, J. Phys. G, **38**: 065001 (2011) [arXiv:0912.2147 [hep-ph]]
36 T. Petri, arXiv:1010.2378 [nucl-th]
37 S. S. Agaev, V. M. Braun, N. Offen, and F. A. Porkert, Phys. Rev. D, **83**: 054020 (2011) [arXiv:1012.4671 [hep-ph]]
38 I. Balakireva, W. Lucha, and D. Melikhov, Phys. Rev. D, **85**: 036006 (2012) [arXiv:1110.6904 [hep-ph]]
39 C. Q. Geng and C. C. Lih, Phys. Rev. C, **86**: 038201 (2012) [Phys. Rev. C, **87**: 039901 (2013)] [arXiv:1209.0174 [hep-ph]]
40 Y. Klopot, A. Oganesian, and O. Teryaev, Phys. Rev. D, **87**: 036013 (2013) [Phys. Rev. D, **88**: 059902 (2013)] [arXiv:1211.0874 [hep-ph]]

41 C. Terschlüsen, B. Strandberg, S. Leupold, and F. Eichstädt, Eur. Phys. J. A, **49**: 116 (2013) [arXiv:1305.1181 [hep-ph]]
42 Y. Klopot, A. Oganesian, and O. Teryaev, JETP Lett., **99**: 679 (2014) [arXiv:1312.1226 [hep-ph]]
43 S. S. Agaev, V. M. Braun, N. Offen, F. A. Porkert, and A. Schäfer, Phys. Rev. D, **90**: 074019 (2014) [arXiv:1409.4311 [hep-ph]]
44 P. Masjuan Queralt, arXiv:1005.5683 [hep-ph]
45 R. Escribano, S. Gonzalez-Solis, P. Masjuan, and P. Sanchez-Puertas, Phys. Rev. D, **94**: 054033 (2016) [arXiv:1512.07520 [hep-ph]]
46 G. A. Baker and P. Graves-Morris, *Encyclopedia of Mathematics and Its Applications*, (Cambridge: Cambridge University Press, U.K.) (1996)
47 P. Masjuan and J. J. Sanz-Cillero, Eur. Phys. J. C, **73**: 2594 (2013) [arXiv:1306.6308 [hep-ph]]
48 P. Masjuan, J. Ruiz de Elvira, and J. J. Sanz-Cillero, Phys. Rev. D, **90**: 097901 (2014) [arXiv:1410.2397 [hep-ph]]
49 S. Gonzàlez-Solís, P. Masjuan, and P. Sánchez-Puertas, in preparation
50 P. Masjuan, E. Ruiz Arriola, and W. Broniowski, Phys. Rev. D, **87**: 014005 (2013) [arXiv:1210.0760 [hep-ph]]
51 M. Hayakawa and T. Kinoshita, Phys. Rev. D, **57**: 465 (1998) [Phys. Rev. D, **66**: 019902 (2002)] [hep-ph/9708227]
52 M. Knecht and A. Nyffeler, Phys. Rev. D, **65**: 073034 (2002) [hep-ph/0111058]
53 A. R. Barker, H. Huang, P. A. Toale, and J. Engle, Phys. Rev. D, **67**: 033008 (2003) [hep-ph/0210174]
54 F. Jegerlehner and A. Nyffeler, Phys. Rept., **477**: 1 (2009) [arXiv:0902.3360 [hep-ph]]
55 H. Suura, T. F. Walsh, and B. L. Young, Lett. Nuovo Cim., **4S2**: 505 (1972) [Lett. Nuovo Cim., **4**: 505 (1972)]
56 G. Kopp, T. F. Walsh, and P. M. Zerwas, Nucl. Phys. B, **70**: 461 (1974)
57 K. S. Babu and E. Ma, Phys. Lett. B, **119**: 449 (1982)
58 P. Masjuan and P. Sanchez-Puertas, arXiv:1504.07001 [hep-ph]
59 P. Adlarson et al, arXiv:1412.5451 [nucl-ex]
60 C. W. Xiao, T. Dato, C. Hanhart, B. Kubis, U.-G. Meissner, and A. Wirzba, arXiv:1509.02194 [hep-ph]
61 I. Larin et al (PrimEx Collaboration), Phys. Rev. Lett., **106**: 162303 (2011) [arXiv:1009.1681 [nucl-ex]]
62 R. H. Dalitz, Proc. Phys. Soc. A, **64**: 667 (1951)
63 D. W. Joseph, Nuovo Cim., **16**: 997 (1960)
64 T. Husek, K. Kampf, and J. Novotny, Phys. Rev. D, **92**: 054027 (2015) [arXiv:1504.06178 [hep-ph]]
65 N. M. Kroll and W. Wada, Phys. Rev., **98**: 1355 (1955)
66 T. Miyazaki and E. Takasugi, Phys. Rev. D, **8**: 2051 (1973)
67 E. Byckling and K. Kajantie, Particle Kinematics, Wiley (1973)
68 P. Nyborg, H. S. Song, W. Kernan, and R. H. Good, Phys. Rev. B, **140**: 914 (1965)
69 N. Byers and C. N. Yang, Rev. Mod. Phys., **36**: 595 (1964)
70 L. G. Landsberg, Phys. Rept., **128**: 301 (1985)
71 P. Adlarson et al (A2 Collaboration), Phys. Rev. C, **95**: 025202 (2017) [arXiv:1611.04739 [hep-ex]]
72 B. Bhattacharya, R. J. Hill, and G. Paz, Phys. Rev. D, **84**: 073006 (2011) [arXiv:1108.0423 [hep-ph]]