New determination of α_{s} from hadronic τ decays

Diogo Boito, Oscar Catà, Maarten Golterman, Matthias Jamin, Kim Maltman, James Osborne, and Santiago Peris

Phys. Rev. D **84**, 113006 — Published 13 December 2011

DOI: 10.1103/PhysRevD.84.113006

LMU-ASC 46/11

A new determination of α_s from hadronic τ decays

Diogo Boito,[a] Oscar Catà,[b,c] Maarten Golterman,[d] Matthias Jamin,[e] Kim Maltman,[f,g]
James Osborne,[d] Santiago Peris[d*]

[a] *Departament de Física and IFAE*
Universitat Autonòma de Barcelona, E-08193 Bellaterra, Barcelona, Spain

[b] *Departament de Física Teòrica and IFIC*
Universitat de València-CSIC, E-46071 València, Spain

[c] *Ludwig-Maximilians-Universität München, Fakultät für Physik*
Arnold Sommerfeld Center for Theoretical Physics, D80333 München, Germany

[d] *Department of Physics and Astronomy*
San Francisco State University, San Francisco, CA 94132, USA

[e] *Institució Catalana de Recerca i Estudis Avançats (ICREA)*
IFAE, Universitat Autonòma de Barcelona
E-08193 Bellaterra, Barcelona, Spain

[f] *Department of Mathematics and Statistics*
York University, Toronto, ON Canada M3J 1P3

[g] *CSSM, University of Adelaide, Adelaide, SA 5005 Australia*

[*] Permanent address: Departament de Física, Universitat Autonòma de Barcelona, E-08193 Bellaterra, Barcelona, Spain

ABSTRACT

We present a new framework for the extraction of the strong coupling from hadronic τ decays through finite-energy sum rules. Our focus is on the small, but still significant non-perturbative effects that, in principle, affect both the central value and the systematic error. We employ a quantitative model in order to accommodate violations of quark-hadron duality, and enforce a consistent treatment of the higher-dimensional contributions of the Operator Product Expansion to our sum rules. Using 1998 OPAL data for the non-strange isovector vector and axial-vector spectral functions, we find the $n_f = 3$ values $\alpha_s(m_\tau^2) = 0.307 \pm 0.019$ in fixed-order perturbation theory, and 0.322 ± 0.026 in contour-improved perturbation theory. For comparison, the original OPAL analysis of the same data led to the values 0.324 ± 0.014 (fixed-order) and 0.348 ± 0.021 (contour-improved).

I. INTRODUCTION

In the past few years there has been a renewed interest in the precision determination of α_s from non-strange hadronic τ decays. One reason for this interest is the recent calculation of the coefficient of the $O(\alpha_s^4)$ term in the perturbative contribution to the Adler function [1]. This contribution dominates the ratio of the hadronic τ decay width and the electronic decay rate [2],

$$R_\tau = \Gamma[\tau^- \to \nu_\tau \,\text{hadrons}\,(\gamma)]/\Gamma[\tau^- \to \nu_\tau e^- \bar{\nu}_e(\gamma)] \; . \tag{1.1}$$

Another reason is the existence of a number of competing analysis methods which lead to results that are not, or only barely, consistent with one another. In fact, the error on α_s from τ decays quoted in a recent (2009) review [3] has gone *up* since its 2006 version, for the simple reason that the result of Ref. [3] was obtained by averaging the central values of all recent τ decay determinations of α_s and the error by considering the spread of these central values. All determinations are based on data from (primarily) the ALEPH (see Ref. [4] for their 1998 analysis and Ref. [5] for their 2005 analysis) and (also) the OPAL (see Ref. [6]) collaborations; the differences in the results are on the theory side.

Clearly, this is an unsatisfactory situation. There are at least three theoretical issues related to the discrepancies between the different determinations. Number one is the long-standing question as to which resummation scheme, fixed-order perturbation theory (FOPT) or contour-improved perturbation theory (CIPT), is best used for evaluating the perturbative contributions to R_τ. Many of the recent reanalyses have focussed on this question [1, 7–11]. Much less attention has been devoted to the two other issues, both of which concern non-perturbative contributions to R_τ. Because of the relatively low value of the τ mass, such contributions cannot be entirely neglected, even if they are expected to be small. Issue number two is the question of whether the Operator Product Expansion (OPE) contributions beyond perturbation theory have been consistently taken into account. Here we are aware of only one systematic investigation [12], in which it was demonstrated that self-consistency problems existed for a number of earlier analyses. Specifically, it was shown that the OPE parameters obtained in those analyses do not provide a good match between data and theory when the upper limit s_0 on the hadronic invariant mass-squared, s, in the weighted integrals over the spectral functions (which enter the finite-energy sum rules (FESRs) employed in

these analyses) is varied away from m_τ^2.[1] Issue number three concerns potential violations of "quark-hadron duality" not taken into account in previous FESR determinations of α_s. In the case of FESRs, the assumption of quark-hadron duality amounts to presuming that all non-perturbative effects are accounted for by higher-dimensional terms in the OPE. To date, this assumption has not seen a systematic investigation. As already explained in Ref. [12], the issues of the OPE and possible violations of quark-hadron duality are intricately connected: without a quantitative analysis of duality violations, it turns out to be difficult to treat the OPE consistently without relying on "external" results. For instance, in the analysis of Ref. [12], the dimension-4 term in the OPE had to be fixed using a result for the gluon condensate from charmonium sum rules.

In this article, we aim to address this situation, focussing on the non-perturbative questions. We use a recently developed model for the duality-violating (DV) part of the *ud*-flavor vector (V) and axial-vector (A) spectral functions [13], which makes it possible to carry out a self-contained FESR analysis in which stability with respect to varying s_0 is checked self-consistently without reliance on external values for any of the OPE parameters.

Since no QCD-based theory of quark-hadron duality exists, we have to resort to a model. This means that our results will be based on the (testable) assumption that this model gives a good description of the DV part of the spectral functions for values of s from $s \to \infty$ down to a minimum value, s_{min}, sufficiently low to lie in the range $s \leq m_\tau^2$ kinematically accessible in hadronic τ decays. We emphasize that the need to make such an assumption has always been a fundamental "shortcoming" of the determination of α_s from τ decays. Assuming quark-hadron duality *a priori*, and therefore neglecting the effect of DVs altogether, also amounts to employing a – probably worse – model. In other words, in order to investigate the systematics related to the assumption of quark-hadron duality, one cannot avoid the adoption of a model of the DV part of the spectral functions. In Ref. [13] it was found that the model we intend to employ gives a reasonable description of the spectral functions in the region 1.1 GeV$^2 \leq s \leq m_\tau^2$.[2] Moreover, the physics of our model is based on a picture of the hadron resonances which are experimentally seen in the spectral function. Resonances

[1] Some of the earlier analyses carried out this test for the FESR based on the kinematic weight (the weight yielding R_τ when $s_0 = m_\tau^2$), for which it works reasonably well. Reference [12] showed that these analyses unambiguously fail the tests for other doubly-pinched FESRs.

[2] In the present, more detailed analysis, we will find that a significantly larger value of s_{min} is preferred.

are not described by perturbation theory or the OPE, and thus should be part of any model aiming to describe violations of quark-hadron duality.

In this work, we do not address the issue of the optimal choice of resummation for the truncated perturbative series in a given FESR. While this systematic, of course, forms a potentially important part of the final theory error on α_s, we have no new elements to add to the discussion of this issue. Moreover, we believe that the non-perturbative part of the systematics should be understood first, in order to get a more reliable picture of the quantitative discrepancy between results based on FOPT and CIPT. We will therefore carry out our whole analysis with both resummation schemes.

To date, two experiments, ALEPH [4, 5] and OPAL [6], have made the non-strange V and A spectral functions from their τ-decay analyses publicly available. The 2005 analysis of ALEPH is more recent, and based on more statistics, and thus would be expected to have smaller experimental errors. Unfortunately, the 2005 ALEPH data cannot be used at present, because correlations due to unfolding have been inadvertently omitted in the original ALEPH analysis and hence from the publicly posted covariance matrices [14].[3] Since the re-analysis of the ALEPH data has yet to be completed, we limit ourselves, in this article, to an analysis employing the OPAL data.

In order to normalize the various exclusive-mode components of the spectral functions, OPAL relied on the branching fractions available in 1998, as well as the then-current values of V_{ud} and the electronic branching fraction B_e. All of these have been updated since then, and this makes it possible to at least partially update the OPAL inclusive spectral distributions as well. Such an update would allow an updated, though still OPAL-based, determination of α_s. Here, since our primary goal is to investigate the impact of the novel features of our treatment of non-perturbative effects on the extracted results for α_s, we choose *not* to perform this update, and instead work with the data in precisely the same form as used by OPAL [6]. We plan to devote a separate article to an adaptation of OPAL data to recent values of the exclusive branching fractions, V_{ud} and B_e, and an investigation of the effect of this adaptation on the value of α_s and other OPE parameters.

This article is organized as follows. In Sec. II we present a brief review of the application

[3] We thank members of the ALEPH collaboration for private communications, in which the existence of this problem has been confirmed.

of FESRs to hadronic τ decays, with emphasis on the issue of quark-hadron duality. In Sec. III we are then able to provide a more thorough discussion of the various systematic errors discussed already above. Preparing for a presentation of our results in Sec. VI, we describe the theory parametrization we will employ in more detail in Sec. IV, and discuss the issue of strong correlations in the integrated data, and our resulting fitting strategies, in Sec. V. Apart from reporting on our fits in Secs. VI A and VI B, we consider also, in Sec. VI D, the $V+A$ channel sum (related to the non-strange part of R_τ) as well as the $V-A$ channel difference. In the latter case, we demonstrate that our fit results satisfy the Weinberg sum rules [15] as well as the DGMLY sum rule for the π electromagnetic mass difference [16]. Section VII contains a summary of our results, including a conversion of α_s to its value at the Z mass; Sec. VIII contains our conclusions.

II. THEORY SUMMARY

Our analysis will involve the correlation functions

$$\begin{aligned}\Pi_{\mu\nu}(q) &= i\int d^4x\, e^{iqx}\langle 0|T\left\{J_\mu(x)J_\nu^\dagger(0)\right\}|0\rangle \\ &= \left(q_\mu q_\nu - q^2 g_{\mu\nu}\right)\Pi^{(1)}(q^2) + q_\mu q_\nu \Pi^{(0)}(q^2) \\ &= \left(q_\mu q_\nu - q^2 g_{\mu\nu}\right)\Pi^{(1+0)}(q^2) + q^2 g_{\mu\nu}\Pi^{(0)}(q^2)\ ,\end{aligned} \tag{2.1}$$

where J_μ is one of the non-strange V or A currents, $\overline{u}\gamma_\mu d$ or $\overline{u}\gamma_\mu\gamma_5 d$, and the superscripts (0) and (1) label the spin. The decomposition in the third line employs the combinations $\Pi^{(1+0)}(q^2)$ and $q^2\Pi^{(0)}(q^2)$, which are free of kinematic singularities. Defining $s=q^2=-Q^2$ and the spectral functions

$$\rho^{(1+0)}(s) = \frac{1}{\pi}\,\mathrm{Im}\,\Pi^{(1+0)}(s)\ , \tag{2.2}$$

Cauchy's theorem and the analytical properties of $\Pi^{(1+0)}(s)$, applied to the contour in Fig. 1, imply the FESR relation

$$I_{V/A}^{(w)}(s_0) \equiv \frac{1}{s_0}\int_0^{s_0} ds\, w(s)\,\rho_{V/A}^{(1+0)}(s) \;=\; -\frac{1}{2\pi i s_0}\oint_{|s|=s_0} ds\, w(s)\,\Pi_{V/A}^{(1+0)}(s)\ , \tag{2.3}$$

valid for any $s_0>0$ and any weight $w(s)$ analytic in the region of the contour [17]. In the present work we will restrict ourselves to polynomial weights. Partial integration allows the right-hand side of Eq. (2.3) to be recast in terms of the Adler function

$$D(s) = -s\,\frac{d\Pi^{(1+0)}(s)}{ds}\ . \tag{2.4}$$

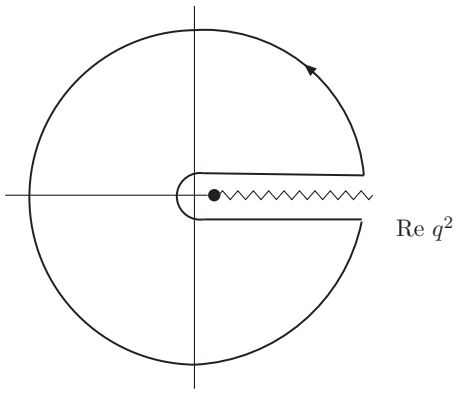

FIG. 1: *Analytic structure of* $\Pi^{(1+0)}(q^2)$ *in the complex* $s = q^2$ *plane. The solid curve shows the contour used in Eq. (2.3).*

The spectral functions $\rho_{V/A}^{(1+0)}(s)$ are measurable in hadronic τ decays. Explicitly, for Standard Model decays induced by the flavor ud (isovector) currents, with $w_T(s;s_0) = (1-s/s_0)^2(1+2s/s_0)$, $w_L(s;s_0) = 2(s/s_0)(1-s/s_0)^2$, and the scaled, non-strange V and A widths

$$R_{V/A;ud} \equiv \frac{\Gamma[\tau^- \to \nu_\tau \,\text{hadrons}_{V/A;ud}\,(\gamma)]}{\Gamma[\tau^- \to \nu_\tau e^- \bar{\nu}_e(\gamma)]} \;, \tag{2.5}$$

one has [18]

$$R_{V/A;ud} = R_{V/A;ud}(s_0 = m_\tau^2) \;, \tag{2.6}$$

in which $R_{V/A;ud}(s_0)$ is defined by

$$R_{V/A;ud}(s_0) = 12\pi^2 |V_{ud}|^2 S_{EW} \, \frac{1}{s_0} \int_0^{s_0} ds \, \left[w_T(s;s_0) \rho_{V/A}^{(1+0)}(s) - w_L(s;s_0) \rho_{V/A}^{(0)}(s) \right] \;, \tag{2.7}$$

where S_{EW} is a short-distance electroweak correction. Since, in the Standard Model,

$$\begin{aligned} \rho_V^{(0)}(s) &= O[(m_d - m_u)^2)] \;, \\ \rho_A^{(0)}(s) &= 2f_\pi^2 \left(\delta(s - m_{\pi^\pm}^2) - \delta(s) \right) + O[(m_d + m_u)^2] \;, \end{aligned} \tag{2.8}$$

the differential distributions proportional to the expression in square brackets in Eq. (2.7) provide a direct measure of $\rho_{V/A}^{(1+0)}(s)$, up to numerically negligible $O\left(m_{u,d}^2\right)$ corrections.[4] The second delta function in Eq. (2.8), which comes from the kinematic singularity present

[4] The π decay constant, $f_\pi = 92.21(14)$ MeV [19], is presently known very accurately. The central value of

in $\Pi^{(0)}$, does not contribute to the integral in Eq. (2.7) as a result of the factor of s in the accompanying weight $w_L(s;s_0)$.

For sufficiently large s_0, and ignoring for a moment that the OPE is only valid for large euclidean Q^2, the right-hand side of Eq. (2.3) can be approximated using the OPE for $\Pi_{V/A}^{(1+0)}(s)$. Experimental spectral data can then be used to fit the OPE, and extract parameters such as α_s [2]. In what follows, we will denote the experimental version of the spectral integral on the left-hand side of Eq. (2.3) by $I_{V/A;ex}^{(w)}(s_0)$ (generically, $I_{ex}^{(w)}(s_0)$) and the theoretical representation of the contour integral on the right-hand side by $I_{V/A;th}^{(w)}(s_0)$ (generically, $I_{th}^{(w)}(s_0)$).

In the upper part of the energy region allowed by τ-decay kinematics $\left[\Pi_{V/A}^{(1+0)}(s)\right]_{\rm OPE}$ is dominated by its dimension $D=0$ contribution, *i.e.*, the perturbative contribution in the chiral limit.[5] The perturbative expression for the Adler-function (2.4), which is known to order α_s^4 [1], can be written as[6]

$$[D(s)]_{\rm OPE}^{D=0} = \frac{1}{4\pi^2}\sum_{n=0}^{\infty} a_s^n(\mu^2)\sum_{k=1}^{n+1} k\, c_{nk}\left(\log\frac{-s}{\mu^2}\right)^{k-1} \quad , \tag{2.9}$$

with $a_s(\mu^2)=\alpha_s(\mu^2)/\pi$. Since $D(s)$ is independent of μ^2, we can choose $\mu^2=-s$, indicating that only the coefficients c_{n1} are independent; all other c_{nk} can be expressed in terms of the c_{n1} through the renormalization group. In the $\overline{MS}$ scheme, $c_{01}=c_{11}=1$, $c_{21}=1.63982$, $c_{31}=6.37101$ and $c_{41}=49.07570$ [1]. We will use the guess $c_{51}=283$ of Ref. [8] for the next coefficient, assigning an uncertainty of ± 283 in order to estimate the error due to truncating perturbation theory (*cf.* Sec. VI C).

The freedom to choose μ^2 in Eq. (2.9) is at the heart of the different prescriptions employed for evaluating the perturbative contribution to the right-hand side of Eq. (2.3): in FOPT, $\mu^2=s_0$ is used in Eq. (2.9), whereas in CIPT $\mu^2=-s$ is employed inside the contour integral on the right-hand side of Eq. (2.3) [21].

---

the $\tau\to\pi\nu_\tau$ branching fraction, B_π, employed in the 1998 OPAL analysis corresponds to the somewhat larger value 94.0 MeV. In order to match exactly the OPAL treatment of the V and A spectral functions we employ the latter value in our analysis. Note that, since B_π was obtained by the PDG in a combined fit to the full set of τ basis modes, it would, in fact, be inconsistent to change just B_π without simultaneously changing all other branching fractions. We will revert to the updated value $f_\pi=92.21(14)$ MeV in our later analysis employing updated OPAL data.

[5] Perturbative contributions proportional to powers of the quark masses are included in $D\geq 2$ OPE terms.

[6] See, for instance, Ref. [20].

Beyond perturbation theory, one may improve the approximation to the right-hand side of Eq. (2.3) by including higher-dimension contributions to $\Pi^{(1+0)}_{\rm OPE}(s)$. Explicitly,

$$\Pi^{(1+0)}_{\rm OPE}(s) = \sum_{k=0}^{\infty} \frac{C_{2k}(s)}{(-s)^k} \ , \tag{2.10}$$

with the OPE coefficients C_{2k} logarithmically dependent on s through perturbative corrections. The term with $k = 0$ corresponds to the purely perturbative, mass-independent contributions, represented by Eq. (2.9). The C_{2k}, with $k > 1$, contain non-perturbative $D = 2k$ condensate contributions, and are, in principle, different for the V and A channels.

We will neglect C_2, which is purely perturbative and quadratic in the light quark masses.[7] It has been suggested that a non-perturbative $D = 2$ term should be added in order to account for the truncation of the perturbative series for the $D = 0$ term [22]. We postpone an investigation of this issue to future work, and here set $C_2 = 0$.

Neglecting contributions of $O(m^4_{u,d})$ or proportional to $\langle \bar{u}u\rangle - \langle \bar{d}d\rangle$, both of which are numerically very small, the coefficient C_4 is a linear combination of the "gluon condensate" $\langle a_s G_{\mu\nu} G^{\mu\nu}\rangle$ (with $G_{\mu\nu}$ the gluon field strength), and the chiral condensates $m_i\langle\overline{\psi}_i\psi_i\rangle$, $i = u,\ d,\ s$. To leading order in α_s both contributions to C_4 are the same in the V and A channels. Differences in the $D = 4$ V and A light quark condensate contributions enter beginning at $O(\alpha_s)$ or $O(m^4_{u,d})$ [23].

The coefficient C_6 is assumed to be dominated by four-quark condensates, because the contribution from $\langle g^3 f_{abc} G^{a\nu}_{\mu} G^{b\kappa}_{\nu} G^{c\mu}_{\kappa}\rangle$ vanishes at leading order in α_s [24], and the contributions from lower-dimensional operators are suppressed by powers of the quark mass. Coefficient functions for the four-quark condensates were calculated to next-to-leading order in Ref. [25]. At $D = 8$ there is a proliferation of operators, and very little detailed information is available. As we will explain in Sec. IV below, we will not need to consider terms with $D > 8$.

The OPE is valid when the euclidean distance $|x|$ in Eq. (2.1) is small compared to Λ^{-1}_{QCD}, or, equivalently, when euclidean Q^2 is positive and large. However, both perturbation theory and the OPE are expected to break down near the positive real q^2 axis [26]. We may account

[7] Since in the present study we are only dealing with the light up- and down-quark correlators, the $D = 2$ mass-squared corrections are tiny. Still, one version of our analysis code has implemented all known m^2 corrections up to $\mathcal{O}(\alpha_s^3)$, and we have verified that they are indeed negligible.

for this additional non-perturbative effect by writing the right-hand side of Eq. (2.3) as [13]

$$-\frac{1}{2\pi i s_0}\oint_{|s|=s_0} ds\, w(s)\,\left(\Pi^{(1+0)}_{\rm OPE}(s)+\Delta(s)\right)\,, \tag{2.11}$$

with

$$\Delta(s)\equiv\Pi^{(1+0)}(s)-\Pi^{(1+0)}_{\rm OPE}(s)\,. \tag{2.12}$$

The difference $\Delta(s)$ defines the duality violating contribution to $\Pi^{(1+0)}(s)$.

All previous determinations of α_s from hadronic τ decays have assumed, implicitly or explicitly, that integrated DV contributions are small enough to be neglected for the weights employed in the analysis. While this assumption has sometimes been checked for self-consistency (see, *e.g.*, Ref. [12]), a comprehensive quantitative estimate of the impact of DVs on the precision with which α_s can be determined has not been provided. One of the aims of the present work is to provide a comprehensive analysis which takes DVs into account, and hence remedies this shortcoming.

As shown in Ref. [13], if $\Delta(s)$ is assumed to decay fast enough as $s\to\infty$, the right-hand side of the FESR relation (2.3) can be rewritten as

$$I^{(w)}_{th}(s_0)=-\frac{1}{2\pi i s_0}\oint_{|s|=s_0} ds\, w(s)\,\Pi^{(1+0)}_{\rm OPE}(s)+\mathcal{D}_w(s_0)\,, \tag{2.13}$$

with

$$\mathcal{D}_w(s_0)=-\frac{1}{s_0}\int_{s_0}^{\infty} ds\, w(s)\,\frac{1}{\pi}\,{\rm Im}\,\Delta(s)\,. \tag{2.14}$$

The imaginary parts $\frac{1}{\pi}\,{\rm Im}\,\Delta_{V/A}(s)$ can be interpreted as the DV parts, $\rho^{\rm DV}_{V/A}(s)$, of the V/A spectral functions. Following Ref. [13], we will parametrize $\rho^{\rm DV}_{V/A}(s)$ as

$$\rho^{\rm DV}_{V/A}(s)=\kappa_{V/A}\,e^{-\gamma_{V/A}s}\sin\left(\alpha_{V/A}+\beta_{V/A}s\right)\,. \tag{2.15}$$

This introduces, in addition to α_s and the $D\geq 4$ OPE condensates, four new parameters in each channel. The OPE, supplemented by the *ansatz* (2.15), will be assumed to hold for $s\geq s_{min}$, where s_{min} will have to be inferred from fits to the experimental data. The extended analysis, including DVs, will of course only be possible if s_{min} lies significantly below m_τ^2.

The *ansatz* (2.15) is modeled on the asymptotic behavior for large s of a semi-realistic model for the QCD spectrum in a given channel. This model was developed in Ref. [27], based on earlier ideas described in Ref. [28]. It incorporates a combination of large-N_c

insights (narrow resonances with widths increasing with mass) as well as the Regge picture for the spacing between resonances. These ingredients lead quite naturally to the exponential decay in Eq. (2.15) with the decay parameter $\gamma \sim 1/N_c$, as well as the oscillatory behavior represented by the sine function, both with arguments (approximately) linear in s. We favor this model over other attempts to model DVs because of its natural connection to the resonance structure of the spectral distributions, something that is not evident in other models (such as those based on instantons). For detailed discussions of the model, see Refs. [27, 29, 30].

III. SYSTEMATIC ERRORS

There are three sources of systematic error affecting, to various extents, existing FESR determinations of α_s. Since the investigation of two of these sources is the central aim of this work, we briefly describe each of the three sources here, before embarking on the details of our analysis.

1. There are (at least) two ways of partially resumming the perturbative contribution to $I^{(w)}_{th}(s_0)$, CIPT and FOPT (*cf.* Sec. II). The relative merits of the two methods have been the subject of a number of investigations [1, 7–11]. While no particular preference is given to either scheme in Refs. [1, 11], CIPT is favored in [7, 9], whereas Refs. [8, 10] give arguments in favor of FOPT, in the latter work through a new CI expansion in a conformally mapped coupling.

   We will not attempt to resolve the associated systematic uncertainty in this work, but instead report on the results of our fits using both CIPT and FOPT. In fact, it is interesting to see what discrepancy remains between CIPT and FOPT after other systematic errors, described below, have been properly taken into account.

2. With the exception of Ref. [12], the OPE has not been treated consistently in previous extractions of α_s from τ decays, in the sense we now explain. Consider a term of order $1/s^k$ in the OPE of Eq. (2.10). The dominant term in the expansion in α_s of the corresponding coefficient C_D, $D = 2k$, is a constant of order one (times the relevant condensate). In the sum rule (2.3), this term in the OPE is picked out by the term of degree $k-1$ in the weight $w(s)$. Other terms in $w(s)$ will also pick up

contributions from $C_D(s)$, because of the logarithmic dependence of $C_D(s)$ on s, but such contributions will not be dominant as they are suppressed by at least one extra power of α_s. Thus, if weight functions up to degree n are used in the fits, it follows that terms up to at least order $k = n + 1$ must be kept in the OPE in order to retain all potentially relevant contributions not suppressed by at least one extra power of α_s.

In the conventional analysis of Refs. [5–7, 31] this was not done: while weights up to degree $n = 7$ in s were employed (which would generally require keeping terms in the OPE up to $D = 16$ ($k = 8$)), only terms up to $D = 8$ ($k = 4$) were retained. As noted earlier, it was found in Ref. [12] that this is not self-consistent: with the parameter values found in those fits, the s_0 dependence of the theory curves does not match that of the data for the majority of the weights employed, as well as for alternate degree 2 and 3 weights which explicitly test the $D \leq 8$ parameters obtained from the original fits.[8] In this work, we will restrict ourselves to weight functions of degree $n \leq 3$, corresponding to keeping terms up to $D = 8$ in the OPE. As we will explain below, this implies we will have to vary s_0 in the FESR (2.3); it is not possible to restrict a consistent analysis to only $s_0 = m_\tau^2$ without additional uncontrollable assumptions.

3. Typically, previous analyses neglected the presence of duality violations.[9] While in some cases this assumption was checked for self-consistency [12], such a check does not provide a quantitative assessment of the impact of residual DV effects. It has long been known that FESRs with a simple weight like $w = 1$ have sizable DVs, even at scales $\sim 2 - 3$ GeV^2, but that switching to weights "pinched" (having a zero) near $s = s_0$ significantly reduces this effect [34]. It has become a standard assumption that using only weights which are at least doubly pinched will suppress DVs sufficiently so as not to affect the value of, or error on, α_s. Clearly, in view of the quite small errors on α_s reported in the recent literature, this issue is in need of further investigation.

   The use of weights which are at least doubly pinched, *i.e.*, which contain at least two powers of $s_0 - s$, forces us to vary s_0 in Eq. (2.3), if the OPE is to be treated consistently

[8] See also Ref. [27], where it was shown in a model study that adding a $D = 10$ term to the OPE in an analysis like that of Refs. [5–7] can make a significant difference.

[9] Exceptions are Refs. [7, 32]; however, as is clear from the results of the present work, those investigations of DVs involved additional assumptions which we are able to avoid in the present analysis.

in the sense described above. Suppose one wants to consider only $s_0 = m_\tau^2$, and fit α_s as well as $C_{4,\ldots,D=2k}$, *i.e.*, k parameters.[10] To fit k parameters, one requires at least $k+1$ data points, and therefore at least $k+1$ linearly independent weights if one uses only $s_0 = m_\tau^2$. If all weights contain the factor $(s_0 - s)^2$, the minimally required highest degree will be $n = k+2$. This is inconsistent with our criterion of point 2 above, which would require terms up to order $k+3$ to be kept in the OPE. The only way out is to vary s_0, and/or to consider weights that are less than doubly-pinched. This makes the need to take DVs into account more urgent, because it is certainly not justified to assume that integrated DVs are negligible for weights which are less than doubly pinched, over any sizable interval in s_0 below m_τ^2.

IV. PARAMETRIZATION USED IN FITS

In this section, we describe in detail the parametrization of the theory that we will use in our analysis of the data.

A. Selection of moments

We wish to consider terms in the OPE only up to $D = 8$, for several reasons. First, we expect that small contributions to typical OPE integrals associated with these condensates will be potentially sensitive to residual integrated DV contributions, making it possible to check the impact of DVs on earlier determinations of the condensates. Second, it appears unlikely that we can reliably determine condensates with $D > 8$ from existing data. Hence we focus on FESRs where such contributions will be strongly suppressed. Finally, little is known about the OPE, but it is almost certainly not a convergent expansion. One would thus expect it to break down at some sufficiently high order, making it prudent to limit ourselves to a relatively low maximum order.

From the arguments in Sec. III, this restricts us to weights with degree ≤ 3, *i.e.*, to at most four linearly independent weights. Since we will be fitting up to four OPE parameters (α_s as well as the $D = 4,\ 6$ and 8 condensates), in addition to the DV parameters in each

[10] Recall that we set $C_2 = 0$ in this article.

channel, this already forces us to consider the sum rules (2.3) with more than one value of s_0 for at least one weight. We will vary s_0 over the interval $[s_{min}, m_\tau^2]$, and explore the stability of the fits as a function of s_{min}.

In this work, we choose to consider the weights[11]

$$\begin{aligned} \hat{w}_0(x) &= 1 \ , \\ \hat{w}_2(x) &= 1 - x^2 \ , \\ \hat{w}_3(x) &= (1-x)^2(1+2x) = w_T(s;s_0) \ , \\ x &\equiv s/s_0 \ . \end{aligned} \tag{4.1}$$

A key point is that we explicitly incorporate DVs in our fits, and therefore need to use at least one weight sensitive not only to α_s and the OPE coefficients, but also to the DV parameters. This stands in contrast to other work to date, where a desire to neglect DVs motivated the use of (at least) doubly-pinched weights, which are known to suppress such contributions. Such doubly-pinched weights are, in fact, too insensitive to the DV parameters to allow for reliable fits of these parameters. In order to maximize our sensitivity to DVs and hence improve our ability to fit DV parameters we include the unpinched weight $\hat{w}_0$ in all our fits. The other weights have been chosen by requiring them to be of degree ≤ 3, to have no term linear in s, and to be singly pinched ($\hat{w}_2$) or doubly pinched ($\hat{w}_3$). Alternative sets of weights satisfying the same requirements are obtained by replacing either $\hat{w}_2$ or $\hat{w}_3$ with $1-x^3$. Our results with these alternative sets are completely consistent with those obtained from the set (4.1).

For constant C_D, $I_{th}^{(\hat{w}_2)}$ picks out the $D = 6$ term, while $I_{th}^{(\hat{w}_3)}$ picks out the $D = 6$ and $D = 8$ terms. In practice, the logarithmic dependence of C_D on s beyond leading order in α_s implies that all terms in the OPE contribute for all choices of $w(s)$. However, C_6 and C_8 will be primarily determined by the $\hat{w}_2$ and $\hat{w}_3$ FESRs. In the present work, we will represent the $D \geq 4$ contributions using effective values C_4, C_6 and C_8 independent of s. This implies that C_4 does not contribute to fits involving any of the moments of Eq. (4.1). For the case of C_4, we have checked, in the case of fits to $I_{ex}^{\hat{w}_0}(s_0)$, that the numerical effect of this approximation is tiny, *cf.* Sec. VI A.

As we will show in Sec. VI, fits to moments with weights $\hat{w}_0$, $\hat{w}_2$ and $\hat{w}_3$ lead to stable and

[11] We use hats in order to distinguish our set of weights from a different set of weights considered in Ref. [12].

self-consistent results. One might also consider including moments with weights containing a linear term in s, such as the weight $w_1(x) = 1-x$. Such moments are sensitive to the gluon condensate and would thus allow us to estimate its size, although due care would have to be exercised with the interpretation of such a result since this condensate mixes with the unit operator.[12] However, renormalon-inspired model studies of higher orders in perturbation theory along the lines of Refs. [8, 20] appear to indicate that perturbation theory, truncated at currently known orders, may be less well converged to the full resummed result for $I_{th}^{(w_1)}(s_0)$ than for $I_{th}^{(\hat{w}_{0,2,3})}(s_0)$. This may be related to the observation that the $D = 4$ term, present in the OPE representation of $I_{th}^{(w_1)}(s_0)$, is affected by the leading infrared-renormalon ambiguity in the Borel resummation of the associated perturbative series. In contrast, non-perturbative contributions to the OPE representations of $I_{th}^{(\hat{w}_{2,3})}(s_0)$ depend most significantly on the C_D with $D > 4$, and are affected primarily only by subleading ambiguities, associated with more distant infrared renormalon poles. This is one reason we restrict ourselves to the weights (4.1) in this article.

We have also studied weights with a term linear in x, such as $w_1(x)$, added to our set of weights (4.1), but find that, with analyses of the type we present in Sec. VI, the uncertainty on the resulting determination of the gluon condensate remains large compared to that of other determinations in the literature. We intend to further investigate such fits, and in particular the determination of the gluon condensate, in a forthcoming article. Here we just note that fits similar to those presented in Sec. VI but also including the weight w_1 yield results consistent with the fits presented in Secs VI A and VI B. Finally, we observe that $\hat{w}_{0,2,3}(x)$ are a complete, linearly-independent set of weights of degree ≤ 3 without a term linear in s.

B. Duality violations

We will parametrize the duality-violating part of ρ_V and ρ_A as in Eq. (2.15). This introduces four new parameters in each channel, forcing us to consider values of s_0 over an interval $[s_{min}, m_\tau^2]$. Since our *ansatz* (2.15) is assumed to hold only for sufficiently large s, s_{min} must be large enough to lie in the region of validity of this assumption, but low enough

[12] A similar observation holds for some of the condensates in C_6 and C_8.

to be kinematically accessible in hadronic τ decay. Our *ansatz* is therefore only practical if we also assume that such an s_{min} exists. We are then interested in a value of s_{min} that is small enough to maximize the data available for use in our fits, requiring at the same time that the DV *ansatz* with that choice of s_{min} gives a good description of the data.

A priori, there is no reason for s_{min} to be equal in the V and A channels. We therefore present two types of analysis: one for the V channel alone, and one for the combined V and A channels ($V\&A$). In the latter case, we will employ an s_{min} common to both channels. This is equivalent to assuming an $s_{min} < m_\tau^2$ exists such that the asymptotic behavior has set in for both the V and A channels for all $s > s_{min}$. Since it seems unlikely for the asymptotic behavior in a given channel to set in below the lowest resonance in that channel, we expect to find $s_{min} \sim m_{a_1}^2$ or higher for the combined $V\&A$ fits. In practice, we find an optimal choice $s_{min} \sim 1.4 - 1.5$ GeV^2.

We have also considered fits to only the A channel, but find that the data in that channel lead to a poor determination of the DV parameters, and thus also of α_s. We believe this is due to the lower quality of the A-channel data, rather than the absence of a sufficiently low A-channel s_{min}, but it is impossible to decide this from the data alone. We therefore do not discuss purely A-channel fits, and restrict our analysis of this channel to combined fits involving the V channel as well.

V. CORRELATIONS AND FITTING STRATEGIES

Values of the left-hand side, $I_{ex}^{(w)}(s_0)$, of Eq. (2.3) for nearby values of s_0 are very strongly correlated, and this has repercussions for the choice of fitting strategies. We describe the strong correlations in Sec. V A, and our strategies in Sec. V B below. As already explained in the introduction, we limit ourselves to an analysis of the OPAL data [6].

A. Correlations and errors

The data we will use are the OPAL compilation of the non-strange V and A spectral functions.[13] These data appear in our fits through the weighted integrals, $I_{ex}^{(w)}(s_0)$, appearing

[13] We would like to thank Sven Menke for making the data files available to us.

on the left-hand sides of the FESRs (2.3), for the various weight functions we consider. These integrals are, of course, represented numerically by sums over the appropriate sets of experimental bins. Since OPAL's bin width is 0.032 GeV^2, varying s_0 between approximately 1.5 GeV^2 and 3.120 GeV^2 (OPAL's highest bin in the V channel, and almost equal to $m_\tau^2 = 3.157$ GeV^2) or 3.088 GeV^2 (OPAL's highest bin in the A channel) provides about 50 data points for each integral.

The integrals $I_{ex}^{(w)}(s_0)$ are, however, highly correlated. For instance, if we consider $I_{ex}^{\hat{w}_0}(s_0)$ on the interval $s_0 \in [1.504, 3.136]$ GeV^2,[14] the corresponding correlation matrix has sub- or super-diagonal elements as large as 0.998, a largest eigenvalue ~ 33, and a smallest eigenvalue ~ 0.00019. Nonetheless, as we will show in Sec. VI A, it turns out that reliable, standard χ^2 fits to the data for $\hat{w}_0$, using our parametrization of the right-hand side of Eq. (2.3), are possible.

The situation changes if we consider two or more weight functions simultaneously. Focussing on our primary fits, with moments constructed with the weight functions $\hat{w}_{0,2,3}$, not only the correlations for each moment have to be taken into account, but also the cross-correlations between these different moments, because they are not independent of each other. In fact, if we consider the full correlation matrix for a combination of moments, for a range of s_0 values, it turns out to have zero eigenvalues at machine precision because of the strong cross-correlations. This means that standard χ^2 fits, employing the full correlation matrix in constructing the function to be minimized, are not possible in this case. This remains true if we "thin out" the data, *i.e.*, if we use fewer values of s_0 on a given interval, by a factor two to four. This puts standard χ^2 fits out of reach for simultaneous fits to multiple moments, forcing us to either use a different fitting strategy, or to drop such fits from consideration.

Because of this problem, we will perform fits to multiple moments using a different "fit quality" $\mathcal{Q}^2$. $\mathcal{Q}^2$ will be a positive-definite quadratic form in the differences between data and theory; for a description of some possible choices, see Sec. V B. Any such $\mathcal{Q}^2$ can be minimized to give an estimate of the fit parameters, as long as we have a reasonable way to estimate the parameter errors and covariances associated with the fit. In this article, we

[14] These numbers are at the right edges of the bins centered at $s = 1.488$ GeV^2 and $s = 3.120$ GeV^2, respectively.

will estimate the parameter error matrix by propagating errors through a linear fluctuation analysis starting from the full data covariance matrix; for details, see App. A.

B. Fitting strategies

In this subsection, we explain three different fitting strategies we have used, in view of the problem of strong correlations described in Sec. V A.

1. The simplest fit we can perform is a standard χ^2 fit to a single moment. We will choose our single-moment fit to be the one with weight function $\hat{w}_0(x) = 1$, since it does not suppress contributions from any part of the spectrum, and is sensitive enough to the DV part of our fitting function to give reasonably good fits for the DV parameters. It should be noted here that our main goal is to minimize the fit error on α_s. The only reason one cares about the "nuisance" parameters $\kappa_{V,A}$, $\gamma_{V,A}$, $\alpha_{V,A}$, and $\beta_{V,A}$ is that they describe part of the physics, and as such have to be taken into account in any fit. We emphasize again that, under the assumption that our DV *ansatz* (2.15) gives a good description of the DVs, there are no reasons to limit ourselves to pinched weights and, in fact, strong arguments not to do so.

   We have also considered fitting the spectral function directly, since it is maximally sensitive to DVs in the kinematically allowed region. Such a fit can be cast in terms of an FESR obtained by choosing $w(s) = 1$ and replacing $\rho^{(1+0)}(s)$ with its derivative on the left-hand side of Eq. (2.3). It turns out that it is not possible to determine α_s from such a fit, basically because the spectral function is much less sensitive to α_s than it is to the DV parameters.[15] In contrast, pinched weights suppress the contribution from DVs more than $\hat{w}_0$ does. We find, in fact, that fits with a single doubly-pinched weight are not stable if one tries to fit both the OPE and (strongly suppressed) DV parameters. In our experience, the most stable and precise results from a single-moment fit are obtained using $\hat{w}_0(x) = 1$. Fits to singly-pinched weights also appear to work well: we have checked standard χ^2 fits with weights $\hat{w}_2(x)$ or $1 - x^3$ and find

[15] In Ref. [13] fits to the spectral function were presented, but there α_s was kept fixed. Note that contributions to $\rho_{V/A}$ from the higher D terms in the OPE are negligibly small.

results in excellent agreement with those reported in Tables 1 and 4 below, though with somewhat larger errors on α_s.

2. It is of course interesting to see whether simultaneous fits to multiple moments can be used to reduce errors, in particular the error on α_s. However, in this case, we run into the problem of strong correlations described above in Sec. V A. The simplest solution is to omit correlations in constructing the fit quality $\mathcal{Q}^2$, and choose a $\mathcal{Q}^2$ which is diagonal in the differences between data and theory. Working with a set of s_0, $\{s_0^k\}$ in some fitting window, and letting $\delta I_{ex}^{(w)}(s_0)$ be the error on the weighted spectral integral $I_{ex}^{(w)}(s_0)$, obtained using the full data covariance matrix, such a fit quality has the form

$$\mathcal{Q}^2_{diag} = \sum_w \sum_{s_0^k} \left(\frac{I_{ex}^{(w)}(s_0^k) - I_{th}^{(w)}(s_0^k;\vec{p})}{\delta I_{ex}^{(w)}(s_0^k)} \right)^2 , \tag{5.1}$$

where we have made the dependence of the weighted theory integral on the set of fit parameters $\vec{p}$ explicit and the outer sum runs over the set of weights included in the analysis.

Often, such a fit is referred to as "uncorrelated." Indeed, if $\mathcal{Q}^2$ would be interpreted as a χ^2, the standard χ^2 errors obtained from such a fit would miss the effect of correlations and be (significantly) underestimated. However, we emphasize that we will not compute parameter errors in this way; instead we will propagate errors using the linear fluctuation analysis of App. A, thus taking into account all correlations explicitly.[16]

While we will not report on fits to Eq. (5.1), but instead rely on fits described under items 1 and 3, we have carried out many such fits. They yield results fully consistent with those we do report, but always lead to larger parameter errors.

3. A third type of fit we will consider is one in which $\mathcal{Q}^2$ incorporates the correlation sub-matrix corresponding to each individual moment employed in the fit, but not the cross-correlations between different moments. Full correlations are again to be

[16] We find indeed that this leads to much larger errors than naive "χ^2" errors would suggest.

included via the linear fluctuation analysis described in App. A. We choose

$$\mathcal{Q}^2_{block} = \sum_w \sum_{s_0^i,\, s_0^j} \left(I^{(w)}_{ex}(s_0^i) - I^{(w)}_{th}(s_0^i;\vec{p}) \right) \left(C^{(w)} \right)^{-1}_{ij} \left(I^{(w)}_{ex}(s_0^j) - I^{(w)}_{th}(s_0^j;\vec{p}) \right) \ , \qquad (5.2)$$

with C_w the covariance matrix of the set of moments with fixed weight w and s_0 running over the chosen fit window range. The motivation for this form is that the cross-correlations between two moments arise mainly because the weight functions used in multiple-moment fits are in practice close to being linearly dependent (even though, as a set of polynomials, of course they are not). This dependency might be reinforced by the relatively large errors on the data for values of s toward m_τ^2, because it is primarily in this region that the weights $\hat{w}_0$, $\hat{w}_2$ and $\hat{w}_3$ differ from each other.

A key observation is that it does not matter which fit quality $\mathcal{Q}^2$ one chooses,[17] as long as errors are propagated appropriately. Whatever the motivation for a particular choice, such a choice is useful if it turns out to allow a reliable fit, and to reduce errors on the fit parameters. We note that, of course, it is not possible to use the minimum value of diagonal or block diagonal $\mathcal{Q}^2$ of Eqs. (5.1) and (5.2) obtained in such a fit in order to derive a confidence level; only the relative size of minimum values compared between different fits with the same choice of $\mathcal{Q}^2$ is meaningful.

VI. FITS

In this section, we will present the results from our fits, using the parametrization of the theory explained in Sec. IV and employing the strategies of Sec. V B. All fits are based on the original, unmodified OPAL data, including the OPAL normalization for the π-pole contribution, which corresponds to a central value of 94.0 MeV for f_π.

Section VI A contains our "benchmark" fit, which is a standard χ^2 fit of $I^{(\hat{w}_0)}_{ex}(s_0)$ for the V channel. In this case, the fit quality is the standard χ^2 function, which of course employs the full $I^{(\hat{w}_0)}_{ex}(s_0)$ covariance matrix, generated from the covariance matrix of the original OPAL data. We will also consider a combined fit of the same moment to the V and A channels.

[17] As long as $\mathcal{Q}^2$ is a positive-definite quadratic form in the differences $I^{(w)}_{ex}(s_0^k) - I^{(w)}_{th}(s_0^k;\vec{p})$.

In Sec. VI B we consider simultaneous fits to $I_{ex}^{(\hat{w}_0)}$, $I_{ex}^{(\hat{w}_2)}$ and $I_{ex}^{(\hat{w}_3)}$, again for both the pure V channel and combined $V\&A$ channel cases. As already discussed in Sec. V A, we find that standard χ^2 fits are not possible. Our best results in the V channel originate from minimizing the fit quality (5.2). Our main conclusion is that, presumably because of the strong correlations between different moments, these type of fits do not help reduce the error on α_s significantly. They do, however, provide cross-checks, verifying that our *ansatz* (2.15) also describes moments other than just $I_{ex}^{(\hat{w}_0)}$, including the non-strange component of R_τ,

$$R_{V+A;ud}(s_0) = R_{V;ud}(s_0) + R_{A;ud}(s_0) \ , \tag{6.1}$$

which is proportional to $I_{ex,V+A}^{(\hat{w}_3)}(s_0)$. They also give access to the V- and A-channel OPE coefficients $C_{6,V}$, $C_{6,A}$, $C_{8,V}$ and $C_{8,A}$.

In Sec. VI C we consider the additional errors originating from the truncation of perturbation theory, and in Sec. VI D we show that our fits both give a good description of $R_{V+A;ud}$, and satisfy, within errors, the Weinberg and DGMLY $V-A$ sum-rule constraints.

A. Fits with the weight $\hat{w}_0(x)$

We begin with a standard χ^2 fit to the V-channel $w(s) = \hat{w}_0(x) = 1$ FESR. Fit results are presented in Table 1, which shows all fit parameters, as well as the number of degrees of freedom (dof), and the χ^2 per degree of freedom. Errors are standard χ^2 errors; errors computed with Eq. (A5) are typically somewhat larger, but similar in size.

We observe that there is excellent stability for the results with $s_{min} = 1.4$, 1.5 and 1.6 GeV^2, with the errors getting somewhat larger at 1.6 GeV^2. At values of $s_{min} \geq 1.7$ GeV^2, γ_V becomes very small and tends to go negative, which is clearly unphysical. However, within errors such fits are always consistent with those shown in the table. We choose the results obtained at $s_{min} = 1.5$ GeV^2 to fix our central values, and treat the spread of the fit results for s_{min} ranging from 1.4 to 1.6 GeV^2 as an error. From this simple fit, we obtain for α_s at the τ mass the values

$$\begin{aligned} \alpha_s(m_\tau^2) &= 0.307 \pm 0.018 \pm 0.004 \qquad \text{(FOPT)} \ , \\ \alpha_s(m_\tau^2) &= 0.322 \pm 0.025 \pm 0.004 \qquad \text{(CIPT)} \ , \end{aligned} \tag{6.2}$$

where the second error represents the variation with s_{min} discussed above. We note that there is good stability over the full range of s_{min} values covered in Table 1. Since our fitting

s_{min}	dof	χ^2/dof	α_s	κ_V	γ_V	α_V	β_V
1.3	53	0.44	0.320(23)	0.026(18)	0.42(50)	0.54(54)	2.85(30)
1.4	50	0.35	0.311(19)	0.019(13)	0.23(44)	−0.29(64)	3.27(33)
1.5	47	0.36	0.307(18)	0.017(11)	0.16(42)	−0.52(74)	3.38(38)
1.6	44	0.38	0.308(20)	0.018(15)	0.22(51)	−0.47(82)	3.36(41)
1.7	41	0.38	0.305(19)	0.012(12)	0.03(51)	−0.61(86)	3.41(41)
1.3	53	0.45	0.332(37)	0.037(27)	0.64(53)	0.57(58)	2.80(33)
1.4	50	0.36	0.326(27)	0.023(16)	0.35(48)	−0.32(64)	3.27(33)
1.5	47	0.37	0.322(25)	0.020(13)	0.25(44)	−0.57(73)	3.39(38)
1.6	44	0.38	0.323(27)	0.022(20)	0.31(57)	−0.53(81)	3.38(41)
1.7	41	0.39	0.320(25)	0.014(13)	0.08(53)	−0.68(85)	3.43(40)

TABLE 1: *Standard χ^2 fits to Eq. (2.3) with $w(s) = 1$, V channel. FOPT results are shown above the double horizontal line, CIPT results below. Errors are standard χ^2 errors; γ_V and β_V in* GeV^{-2}.

function is non-linear, in general χ^2 errors are expected to be asymmetric. We have therefore also computed asymmetric errors for all the fit parameters. We find that the error on α_s is nearly symmetric, and that only errors on κ_V and γ_V show a significant asymmetry. For instance, we find, at $s_{min} = 1.5$ GeV^2, that, for FOPT,

$$\begin{aligned} \alpha_s(m_\tau^2) &= 0.307^{+0.018}_{-0.021} \,, \\ \kappa_V &= 0.017^{+0.027}_{-0.007} \,, \\ \gamma_V &= 0.16^{+0.63}_{-0.34}\ \mathrm{GeV}^{-2} \,, \\ \alpha_V &= -0.52^{+0.71}_{-0.78} \,, \\ \beta_V &= 3.38^{+0.40}_{-0.36}\ \mathrm{GeV}^{-2} \,. \end{aligned} \tag{6.3}$$

This shows that omitting DVs from the fit, which is equivalent to setting $\kappa_V = 0$, would lead to a poor fit. For CIPT the asymmetries show the same pattern.

In Fig. 2, we show the quality of the match between the fitted $I_{th}^{(\hat{w}_0)}(s_0)$ and $I_{ex}^{(\hat{w}_0)}(s_0)$, as well as of the match between the experimental spectral function and its theoretical counterpart in which parameter values are obtained from the FESR fit for $s_{min} = 1.5$ GeV^2.

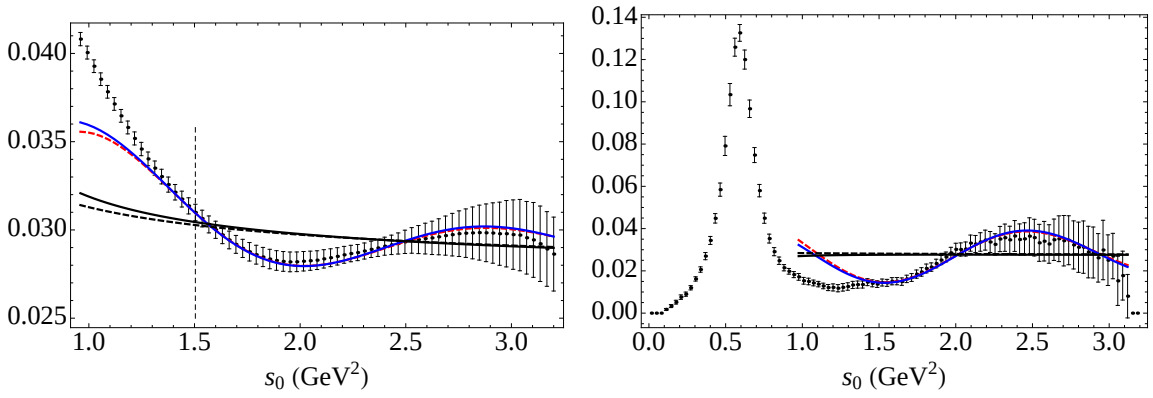

FIG. 2: *Left panel: comparison of* $I_{ex}^{(\hat{w}_0)}(s_0)$ *and* $I_{th}^{(\hat{w}_0)}(s_0)$ *for the* $s_{min} = 1.5$ GeV^2 *V-channel fits of Table 1. Right panel: comparison of the theoretical spectral function resulting from this fit with the experimental results. CIPT fits are shown in red (dashed) and FOPT in blue (solid). The (much flatter) black curves display only the OPE parts of the FOPT (solid) and CIPT (dashed) fit results. The vertical dashed line indicates the location of* s_{min}.

	α_s	κ_V	γ_V	α_V	β_V
α_s	1	-0.69	-0.67	0.70	-0.62
κ_V	-0.69	1	0.99	-0.47	0.43
γ_V	-0.67	0.99	1	-0.48	0.43
α_V	0.70	-0.47	-0.68	1	-0.98
β_V	-0.62	0.43	0.43	-0.98	1

TABLE 2: *Parameter correlation matrix for the FOPT fit with* $s_{min} = 1.5$ GeV^2 *shown in Table 1.*

The left panel shows the FESR fit itself, while the right panel shows the results of this comparison for the spectral function case. We emphasize that the spectral function was not part of the fit; the agreement between the theoretical and experimental versions of $\rho_V(s)$ is an output. Agreement with data is good in the full fit window $s_0 \geq s_{min} = 1.5$ GeV^2. The black curves, in contrast, show the OPE parts of the theoretical curves, *i.e.*, the curves obtained by removing the DV contributions from the blue and red curves. It is clear that DVs are needed to give a good description of the data for $I_{ex}^{(\hat{w}_0)}$ and the spectral function

itself, in agreement with our conclusion based on Eq. (6.3).

There are strong correlations between the parameters shown in Table 1 and Eq. (6.3). Such strong correlations are present in all our fits, and are unavoidable, given the number of fit parameters. We emphasize that this cannot be resolved by simply omitting the duality-violating part from the theory – one cannot "improve" fits by throwing out physics that is known to have an impact on those fits! In Table 2 we show the full parameter correlation matrix corresponding to the FOPT result quoted in Eq. (6.2). The analogous matrix for the CIPT result looks very similar.

We have also considered fits like those shown in Table 1, but including the contribution coming from the logarithmic dependence of C_4 on s. For the latter, we estimated the quark-condensate contribution from the Gell-Mann–Oakes-Renner relation [35], and we took $\langle a_s G_{\mu\nu} G^{\mu\nu} \rangle = 0.021$ GeV^4. We find that corresponding changes in the numbers in Table 1 are at most a tiny fraction of the fitting errors.

We performed a similar type of fit to the combination of V and A channels. If we use all possible s_0 values, we find that the standard χ^2 fit function (involving the very strongly correlated spectral integral covariance matrix) is very flat, admitting not just "physical" solutions (consistent with those in Table 1), but also solutions that are clearly unphysical (with, for instance, values for α_s drifting down to unacceptably low values as a function of s_{min}). This happens for both CIPT and FOPT. Parameter errors for these solutions can be very large, consistent with the flatness of the χ^2 landscape, and there is a strong sensitivity of central values to initial guesses for the parameter values.

However, if we thin out the s_0 values *i.e.*, use only every nth value of $I_{ex}^{\hat{w}_0}(s_0)$, for some value of $n > 1$, we find that fits with $n = 2$, 3 or 4 are much more stable than the one with $n = 1$ (no thinning) described above. In Table 3 we show our results for $n = 3$, which is the choice leading to the most stable fits.[18] The problem disappears when we use fit quality (5.1), computing errors with Eq. (A5), but this method leads to significantly larger errors.

Choosing again the fit with $s_{min} = 1.5$ GeV^2, we find for our combined V and A channel

[18] Negative values for $\gamma_{V,A}$ can appear because we use the analytic form of the integral in Eq. (2.14). Note, however, that all values we find in the fits are consistent with a positive value, as physically required.

fit the results

$$\alpha_s(m_\tau^2) \;=\; 0.308 \pm 0.016 \pm 0.009 \qquad \text{(FOPT)} \;, \tag{6.4}$$
$$\alpha_s(m_\tau^2) \;=\; 0.325 \pm 0.022 \pm 0.011 \qquad \text{(CIPT)} \;,$$

with the second error representing, as above, the variation over the neighboring s_{min} values, 1.4 and 1.6 GeV^2. The results reported in Tables 1 and 3 are in good agreement. We note that the central values for γ_V shown in Table 3 are very small, compared with those in Table 1, but given the errors there is no inconsistency. In Fig. 3 we show the quality of the fit in the panels on the left, for V (top) and A (bottom), for $s_{min} = 1.5$ GeV^2. In the panels on the right we show again the match between the experimental spectral functions and their theoretical counterparts with parameter values obtained from the $s_{min} = 1.5$ GeV^2 FESR fit. As before, black curves show only the OPE parts of the theoretical curves. Again, we see that the data clearly confirm the presence of DVs, which are well described by our *ansatz*.

s_{min}	dof	χ^2/dof	α_s	$\kappa_{V,A}$	$\gamma_{V,A}$	$\alpha_{V,A}$	$\beta_{V,A}$
1.3	30	0.81	0.326(13)	0.0186(88)	0.18(35)	0.35(46)	2.95(27)
				0.094(51)	1.15(35)	0.21(79)	−3.42(45)
1.4	28	0.69	0.317(15)	0.0140(68)	0.01(33)	−0.31(58)	3.29(31)
				0.085(39)	1.06(30)	−0.5(1.1)	−3.06(61)
1.5	26	0.69	0.308(16)	0.0134(69)	−0.01(34)	−0.67(70)	3.46(36)
				0.110(71)	1.15(35)	−1.1(1.1)	−2.73(59)
1.6	24	0.73	0.308(17)	0.0150(98)	0.06(41)	−0.64(74)	3.45(38)
				0.15(13)	1.29(45)	−1.2(1.2)	−2.67(65)
1.7	22	0.68	0.304(18)	0.0131(99)	0.00(42)	−0.80(77)	3.51(38)
				1.0(2.0)	2.14(84)	−2.3(2.1)	−2.1(1.1)
1.3	30	0.85	0.346(19)	0.026(12)	0.37(34)	0.40(51)	2.89(30)
				0.072(35)	1.01(31)	0.10(85)	−3.38(48)
1.4	28	0.72	0.336(21)	0.0170(86)	0.10(35)	−0.36(56)	3.30(30)
				0.075(33)	0.99(28)	−0.5(1.1)	−3.04(61)
1.5	26	0.71	0.325(22)	0.0152(80)	0.05(35)	−0.73(67)	3.48(35)
				0.101(66)	1.11(35)	−1.2(1.0)	−2.74(58)
1.6	24	0.75	0.324(23)	0.017(12)	0.11(44)	−0.72(71)	3.47(37)
				0.14(12)	1.25(45)	−1.3(1.1)	−2.66(62)
1.7	22	0.70	0.318(23)	0.014(11)	0.03(44)	−0.88(75)	3.54(37)
				1.0(1.9)	2.11(86)	−2.3(2.1)	−2.1(1.1)

TABLE 3: *Standard χ^2 fits to Eq. (2.3) for $w(s) = 1$, combined V&A channels. FOPT results are shown above the double horizontal line, CIPT results below. The first line for each s_{min} gives the V DV parameters; the second line the A ones. Every third value of s_0 starting at s_{min} is included in the fits. Errors are standard χ^2 errors; $\gamma_{V,A}$ and $\beta_{V,A}$ in* GeV^{-2}.

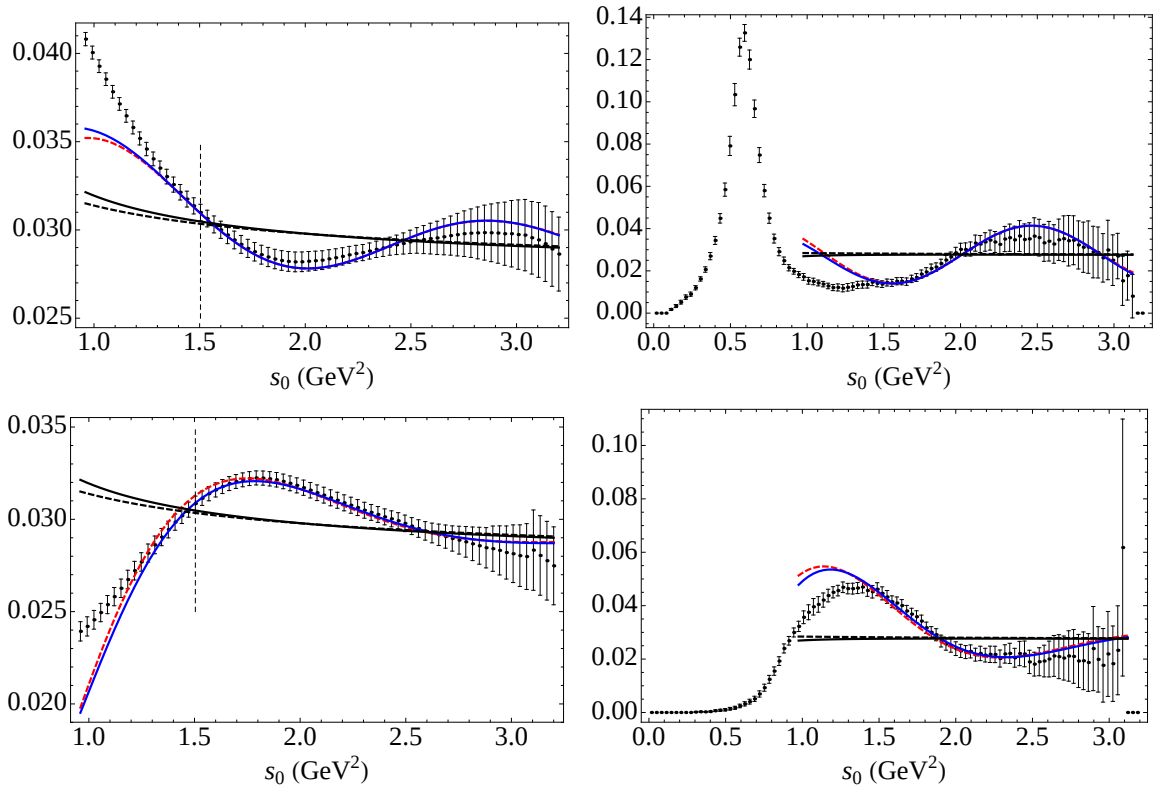

FIG. 3: *Left panels: comparison of* $I_{ex}^{(\hat{w}_0)}(s_0)$ *and* $I_{th}^{(\hat{w}_0)}(s_0)$ *for the* $s_{min} = 1.5$ GeV^2 *combined V& A fits of Table 3 (top: V channel, bottom: A channel). Right panels: comparison of the theoretical spectral function resulting from this fit with the experimental results. CIPT fits are shown in red (dashed) and FOPT in blue (solid). The (much flatter) black curves display only the OPE parts of the FOPT (solid) and CIPT (dashed) fit results. The vertical dashed lines indicate the location of* s_{min}.

B. Multiple-weight fits with the weights $\hat{w}_0$, $\hat{w}_2$ and $\hat{w}_3$

One would naively expect that by using more moments, more information could be extracted from the data. This would help reducing the errors reported in Tables 1 and 3. Higher-degree weights, however, also require the introduction of additional OPE fit parameters. This, in combination with the very strong correlations, may turn out to reduce the extra constraints placed on the parameters (α_s and the DV parameters) entering the $\hat{w}_0$ FESR by the additional moments. In addition, as we already pointed out in Sec. V A, it appears to be impossible to perform standard χ^2 fits to multiple moments.

Table 4 shows the results of simultaneous fits to moments with weights $\hat{w}_0$, $\hat{w}_2$, and $\hat{w}_3$

for the V channel using the fit quality (5.2), with errors computed from Eq. (A5); see Fig. 4 for a visual representation of the qualities of the resulting fits for the case $s_{min} = 1.5$ GeV^2. Results are consistent with those presented in Sec. VI A, but errors are slightly larger.

s_{min}	dof	$\mathcal{Q}^2$/dof	α_s	κ_V	γ_V	α_V	β_V	$10^2 C_{6,V}$	$10^2 C_{8,V}$
1.3	167	0.42	0.300(18)	0.050(35)	0.87(48)	0.38(77)	2.87(44)	−0.39(40)	0.45(68)
1.4	158	0.33	0.304(17)	0.027(18)	0.46(43)	−0.48(88)	3.35(48)	−0.43(31)	0.67(47)
1.5	149	0.33	0.304(19)	0.021(12)	0.31(38)	−0.7(1.1)	3.46(58)	−0.46(33)	0.76(51)
1.6	140	0.33	0.305(23)	0.025(17)	0.41(43)	−0.6(1.4)	3.41(74)	−0.43(46)	0.68(76)
1.7	131	0.34	0.303(25)	0.0136(95)	0.10(39)	−0.8(1.5)	3.47(73)	−0.50(45)	0.88(71)
1.3	167	0.40	0.332(47)	0.035(32)	0.60(64)	0.5(1.0)	2.84(52)	−0.27(59)	0.19(95)
1.4	158	0.32	0.327(31)	0.023(16)	0.34(46)	−0.3(1.0)	3.26(54)	−0.43(36)	0.58(58)
1.5	149	0.32	0.322(31)	0.020(13)	0.26(42)	−0.6(1.3)	3.39(66)	−0.50(37)	0.73(62)
1.6	140	0.33	0.323(42)	0.025(17)	0.37(47)	−0.5(1.7)	3.35(89)	−0.48(54)	0.66(98)
1.7	131	0.34	0.319(39)	0.014(10)	0.11(41)	−0.7(1.7)	3.43(84)	−0.57(48)	0.89(85)

TABLE 4: *Fits to Eq. (2.3) with weights $\hat{w}_{0,2,3}$, V channel, using fit quality (5.2). FOPT results are shown above the double horizontal line, CIPT fits below. Errors have been computed using Eq. (A5); γ_V and β_V in* GeV^{-2}, *$C_{6,V}$ in* GeV^6 *and $C_{8,V}$ in* GeV^8.

Again, the black curves in Fig. 4 show the OPE parts of the theoretical curves, *i.e.*, the results obtained by removing the DV contributions from the blue and red curve results. For the spectral function, $I_{ex}^{(\hat{w}_0)}$ and $I_{ex}^{(\hat{w}_2)}$, it is again clear that no good description of the data can be obtained without a model for DVs. This is not the case for the doubly-pinched moment $I_{ex}^{(\hat{w}_3)}$. In this case, one would expect that a reasonably good fit can be obtained without DVs, for values of s_{min} down to somewhere below ~ 2 GeV^2. This is consistent with the results of Ref. [12] for various doubly pinched weights, and, for the doubly pinched kinematic weight, also with the results of Refs. [5, 6]; in those cases, reasonably good matches for the sum of the vector and axial channels were obtained without the inclusion of DVs. However, one would also expect that a best fit can only be obtained by *shifting* the OPE parameters relative to those reported in Table 4.

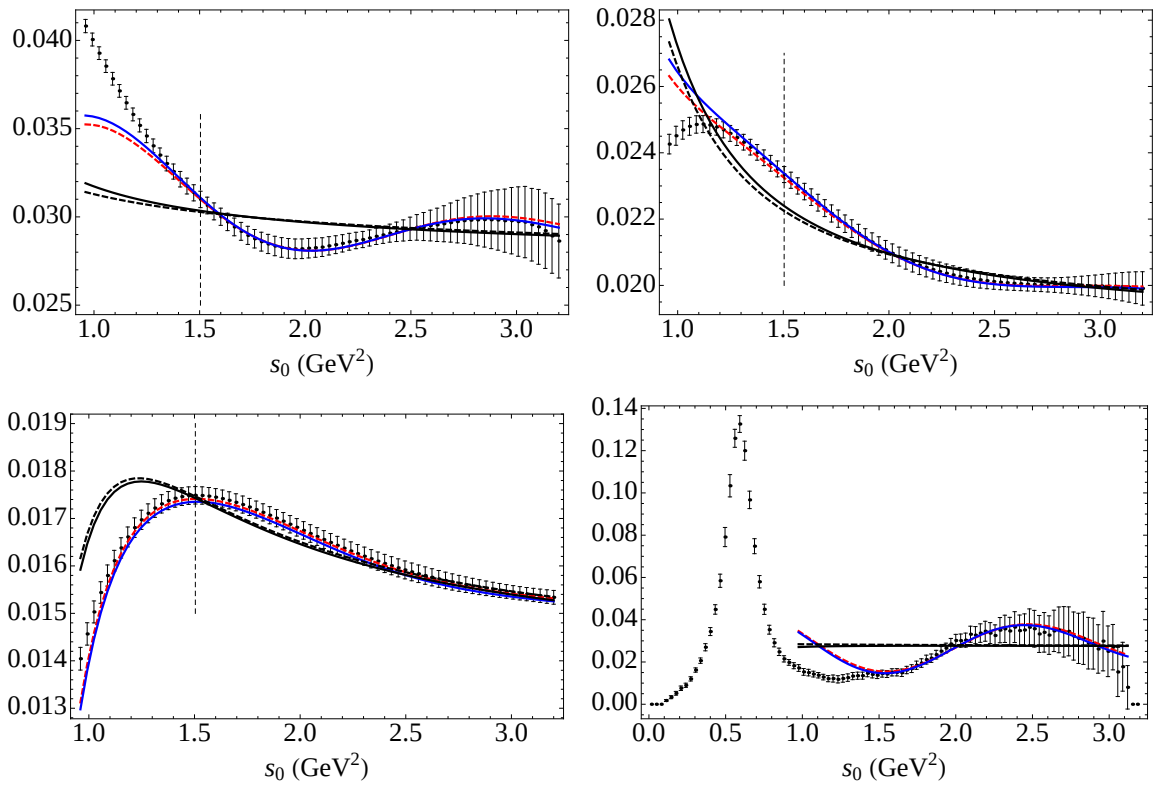

FIG. 4: *V-channel fits of Table 4, showing the theoretical and experimental versions of the moments $I^{(\hat{w}_0)}$ (top left), $I^{(\hat{w}_2)}$ (top right) and $I^{(\hat{w}_3)}$ (bottom left), for $s_{min} = 1.5$ GeV^2. Bottom-right panel: comparison of the theoretical spectral function resulting from this fit with the experimental results. CIPT fits are shown in red (dashed) and FOPT in blue (solid). The (flatter) black curves display only the OPE parts of the FOPT (solid) and CIPT (dashed) fit results. The vertical dashed lines indicate the location of s_{min}.*

An estimate similar to Eq. (6.2), using the fit with $s_{min} = 1.5$ GeV^2, leads to

$$\begin{aligned} \alpha_s(m_\tau^2) &= 0.304 \pm 0.019 \pm 0.001 \qquad \text{(FOPT)} \,, \\ \alpha_s(m_\tau^2) &= 0.322 \pm 0.031 \pm 0.005 \qquad \text{(CIPT)} \,, \end{aligned} \tag{6.5}$$

which is in excellent agreement with Eq. (6.2). We see, however, that the simultaneous fit to multiple moments does not help reduce the error. It is an interesting question whether fit qualities other than those of Eqs. (5.1) and (5.2) exist that would lead to smaller errors. Fits like those reported in Table 4 using fit quality (5.1) lead to consistent results, but with significantly larger errors than those shown in the table.

Finally, in Table 5, we show the results of combined $V\&A$ channel fits, similar to those

of Table 3, but for the weights $\hat{w}_0$, $\hat{w}_2$ and $\hat{w}_3$. From this table, we obtain

$$\begin{aligned} \alpha_s(m_\tau^2) &= 0.302 \pm 0.015 \pm 0.001 \qquad \text{(FOPT)} \ , \\ \alpha_s(m_\tau^2) &= 0.322 \pm 0.024 \pm 0.008 \qquad \text{(CIPT)} \ , \end{aligned} \tag{6.6}$$

with the second error again reflecting the variation over the range $s_{min} = 1.4$ to 1.6 GeV^2. Comparing Table 5 with Table 4 we see that adding the A channel gives little extra information (apart from estimates of the axial OPE and DV parameters). Adding the A channel increases the central values of γ_V, but all parameter values are consistent between these two tables within (sometimes substantial) errors. The fit with $s_{min} = 1.7$ GeV^2 is clearly not meaningful (in particular for FOPT), and the errors indicate that at this value of s_{min}, $\mathcal{Q}^2$ is very flat in some directions in parameter space. Indeed, restricting the value of α_s to the value obtained at $s_{min} = 1.6$ GeV^2 leads to a good fit for the remaining parameters, consistent with the unrestricted $s_{min} = 1.6$ GeV^2 fit, and with a value for $\mathcal{Q}^2$ almost equal to the value reported in the table.

Let us next deduce some implications of the above fits for the breaking of the factorization hypothesis in the $D = 6$ condensates. To begin with, we will assume that the $D = 6$ condensates are dominated by their leading-order contribution. The corresponding contribution to the V/A correlators is given by [2]:

$$\begin{aligned} C_{6,V/A} = &-8\,\pi^2 a_s \left\langle (\bar{u}\gamma_\mu \begin{pmatrix} \gamma_5 \\ 1 \end{pmatrix} t^a d)(\bar{d}\gamma^\mu \begin{pmatrix} \gamma_5 \\ 1 \end{pmatrix} t^a u) \right\rangle \\ &- \frac{8}{9}\,\pi^2 a_s \sum_{q=u,d,s} \left\langle (\bar{u}\gamma_\mu t^a u + \bar{d}\gamma_\mu t^a d)(\bar{q}\gamma^\mu t^a q) \right\rangle + \mathcal{O}(a_s^2) \ , \end{aligned} \tag{6.7}$$

where, in the first line, the upper Dirac structure γ_5 corresponds to the V channel and the lower 1 to the A channel. The two four-quark condensates can be parametrized in terms of their factorization values by

$$\begin{aligned} \langle (\bar{q}_i \gamma_\mu t^a q_j)(\bar{q}_j \gamma^\mu t^a q_i) \rangle &\equiv -\frac{4}{9} \langle \bar{q}_i q_i \rangle \langle \bar{q}_j q_j \rangle \cdot \rho_1 \ , \\ \langle (\bar{q}_i \gamma_\mu \gamma_5 t^a q_j)(\bar{q}_j \gamma^\mu \gamma_5 t^a q_i) \rangle &\equiv \frac{4}{9} \langle \bar{q}_i q_i \rangle \langle \bar{q}_j q_j \rangle \cdot \rho_5 \ , \end{aligned} \tag{6.8}$$

where the parameters ρ_1 and ρ_5 would be equal to one if the vacuum-saturation approximation were exact. Further assuming that isospin breaking is small in the light u- and d-quark

s_{min}	dof	$\mathcal{Q}^2$/dof	α_s	$\kappa_{V,A}$	$\gamma_{V,A}$	$\alpha_{V,A}$	$\beta_{V,A}$	$10^2 C_{6,V/A}$	$10^2 C_{8,V/A}$
1.3	104	0.66	0.305(11)	0.033(18)	0.57(37)	−0.11(63)	3.15(37)	−0.36(24)	0.51(38)
				0.053(20)	0.78(23)	−0.93(63)	−2.81(36)	0.08(27)	0.26(52)
1.4	98	0.48	0.303(13)	0.023(12)	0.32(37)	−0.76(77)	3.51(43)	−0.45(23)	0.74(35)
				0.082(41)	0.98(29)	−1.30(75)	−2.61(42)	−0.12(41)	0.78(91)
1.5	92	0.47	0.302(15)	0.020(10)	0.24(35)	−0.97(94)	3.61(51)	−0.49(25)	0.82(40)
				0.109(84)	1.11(40)	−1.50(89)	−2.51(49)	−0.24(53)	1.1(1.3)
1.6	86	0.46	0.302(19)	0.024(16)	0.35(43)	−0.9(1.2)	3.58(66)	−0.48(36)	0.78(60)
				0.19(20)	1.37(52)	−1.7(1.2)	−2.44(63)	−0.36(74)	1.6(2.0)
1.7	80	0.41	0.277(35)	0.6(1.9)	2.3(2.1)	−1.6(5.9)	3.9(3.1)	−1.00(85)	1.7(1.9)
				1.0(2.3)	2.02(84)	−3.1(2.6)	−1.7(1.4)	−1.6(1.9)	5.7(7.0)
1.3	104	0.55	0.348(19)	0.0206(86)	0.22(30)	0.26(57)	2.98(32)	−0.20(23)	0.19(34)
				0.073(33)	1.02(30)	−0.14(73)	−3.41(41)	0.41(23)	−0.49(43)
1.4	98	0.45	0.330(22)	0.0188(96)	0.18(35)	−0.48(82)	3.36(45)	−0.43(25)	0.61(40)
				0.085(42)	1.03(29)	−0.84(93)	−2.87(52)	0.04(40)	0.30(86)
1.5	92	0.45	0.322(24)	0.0185(95)	0.18(34)	−0.8(1.0)	3.51(55)	−0.52(28)	0.77(47)
				0.105(80)	1.11(40)	−1.2(1.0)	−2.66(58)	−0.15(55)	0.8(1.3)
1.6	86	0.41	0.321(31)	0.023(15)	0.31(41)	−0.7(1.4)	3.49(73)	−0.51(40)	0.73(73)
				0.18(19)	1.36(52)	−1.4(1.4)	−2.58(76)	−0.28(80)	1.2(2.1)
1.7	80	0.42	0.317(37)	0.022(17)	0.29(43)	−0.8(1.7)	3.54(86)	−0.55(51)	0.81(96)
				1.0(2.2)	2.13(88)	−2.2(2.7)	−2.1(1.4)	−0.7(1.3)	2.7(4.4)

TABLE 5: *Fits to Eq. (2.3) with weights $\hat{w}_{0,2,3}$, combined V and A channels, using fit quality (5.2). FOPT results are shown above the double horizontal line, CIPT fits below. Errors have been computed with Eq. (A5); $\gamma_{V,A}$ and $\beta_{V,A}$ in GeV^{-2}, $C_{6,V}$ and $C_{6,A}$ in GeV^6 and $C_{8,V}$ and $C_{8,A}$ in GeV^8. The first line for each s_{min} gives the V channel OPE and DV parameters; the second line the A channel ones. Every third value of s_0 starting at s_{min} is included in the fits.*

sector, that is $\langle \bar{u}u \rangle = \langle \bar{d}d \rangle \equiv \langle \bar{q}q \rangle$, Eqs. (6.7) and 6.8 imply

$$C_{6,V/A} \,=\, \frac{32}{81}\,\pi^2 a_s\,\langle \bar{q}q \rangle^2 \begin{pmatrix} 2\,\rho_1 - 9\,\rho_5 \\ 11\,\rho_1 \end{pmatrix} . \tag{6.9}$$

Inverting Eq. (6.9), on the basis of the results of Table 5, estimates of the parameters $\rho_{1,5}$ can be deduced. As representative examples, for the central fits with $s_{min} = 1.5$ GeV^2, one obtains:

$$\begin{aligned} &\rho_1 = -\,1.4 \pm 3.2\,, \quad \rho_5 \,=\, 3.3 \pm 1.5 \qquad \text{(FOPT)}\,, \\ &\rho_1 = -\,0.9 \pm 3.1\,, \quad \rho_5 \,=\, 3.4 \pm 1.5 \qquad \text{(CIPT)}\,, \end{aligned} \tag{6.10}$$

where $\langle \bar{q}q \rangle(m_\tau^2) \,=\, -\,(272\,\text{MeV})^3$ [33], together with our results for $\alpha_s(m_\tau^2)$, has been employed. The central results of Eq. (6.10) display sizable deviations from the factorization values $\rho_1 = \rho_5 = 1$, though, given the large uncertainties, the significance is not very high. The employed perturbative resummation scheme does not seem to play a big role for the condensate or DV parameters, suggesting that this choice is mostly compensated for by the differing α_s values.

C. Errors from truncating perturbation theory

One of the uncertainties afflicting any determination of α_s is that perturbation theory needs to be truncated, irrespective of whether the CIPT or FOPT resummation scheme is used for the truncated series. As already mentioned in Sec. II, we use the known values of c_{n1} up to $n = 4$, together with the estimate $c_{51} = 283$. We assign a 100% error, ± 283, to this estimate, using this as a measure of the truncation uncertainty.

For the fits presented in Eqs. (6.2), (6.4), (6.5) and (6.6) we find that this variation of c_{51} leads to a shift of at most ± 0.006 in $\alpha_s(m_\tau^2)$; for Eq. (6.2) the shift is ± 0.005. The shifts in other parameters are also small (well within fitting errors).

D. Consistency with the $V + A$ and $V - A$ chiral sum rules

While Figs. 2–4 show that the parameter values obtained from the fits corresponding to these figures give a good description of the data, with those parameter values in hand one

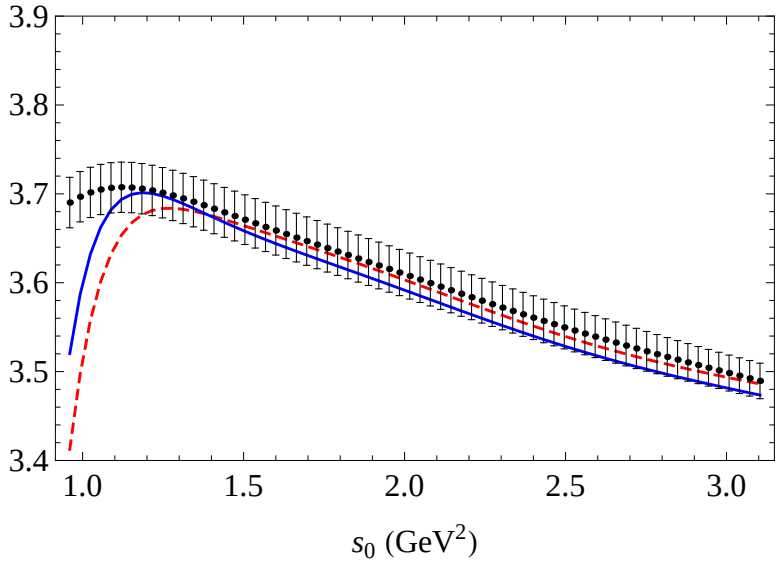

FIG. 5: $R_{V+A;ud}(s_0)$ *as a function of* s_0*, with* $s_{min} = 1.5$ GeV^2 *theory curves from Table 5 for CIPT (dashed red curve) and FOPT (solid blue curve).*

may also consider other quantities. The total non-strange scaled $V + A$ branching fraction $R_{V+A;ud}$ (*cf.* Eqs. (2.5), (2.7) and (6.1)) has always played a central role in the study of hadronic τ decays. In particular, following Refs. [5, 6], we may consider $R_{V+A;ud}(s_0)$ for a hypothetical τ of mass-squared $m_\tau^2 = s_0$, as a function of s_0. We show our version of this quantity in Fig. 5, using the parameter values for the fit with $s_{min} = 1.5$ GeV^2 of Table 5, the only fit reported that simultaneously yields all parameters needed to evaluate $R_{V+A;ud}(s_0)$.[19] Clearly, our result compares well with the fits shown in Fig. 10 of Ref. [6],[20] especially keeping in mind that in Fig. 5 we only show errors on the experimental spectral integrals, and not on the theory curves. Both CIPT and FOPT describe the data well down to $s_0 = 1.5$ GeV^2. We used the same values for $S_{EW} = 1.0194$ and $|V_{ud}|^2 = 0.9512$ as Ref. [6] in order to plot $R_{V+A;ud}(s_0)$. In view of the discussion in Ref. [12] of the analysis of Refs. [5, 6], we conclude that our fits pass the test of $R_{V+A;ud}(s_0)$ much better than the original analyses of Refs. [5, 6], and over a wider range of s_0 than the alternate, self-consistent fits obtained ignoring DVs in Ref. [12]. We emphasize, though, that it is not sufficient to find a satisfactory description of this quantity only – at the very least all FESRs used in the

[19] We included the pion-pole contribution to the longitudinal part in Eq. (2.7) in both the data points and the theory curves in Fig. 5.

[20] For an early investigation of this type, see Ref. [37].

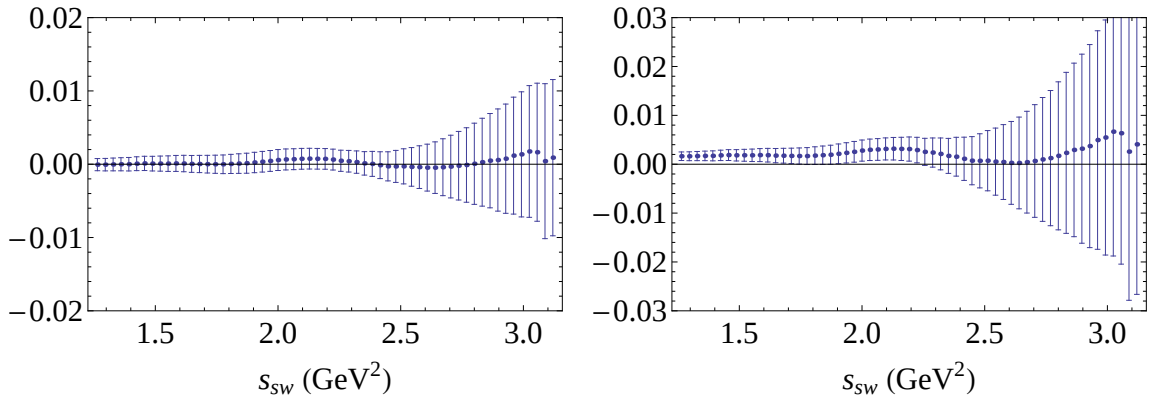

FIG. 6: *First (left, in* GeV^2*) and second (right, in* GeV^4*) Weinberg sum rules. Data are used for* $s \leq s_{sw}$*, while the DV ansatz (2.15) with values from Table 5 for* $s_{min} = 1.5$ GeV^2 *has been used for* $s \geq s_{sw}$*. FOPT figures are shown; CIPT figures look essentially identical.*

fits should show a similarly good match between experiment and theory as a function of s_0, as was shown to be the case for our fits in Secs. VI A and VI B above.

We may perform a similar test on the fits reported in Table 4. Of course, in that case only V-channel parameters are available, so one should consider the corresponding ratio $R_{V;ud}$ as a function of s_0. For the V channel, $R_{V;ud}(s_0)$ coincides (up to the overall factor $12\pi^2 S_{EW}|V_{ud}|^2$) with $I_{ex}^{(\hat{w}_3)}(s_0)$. The corresponding match between data and theory for $s_{min} = 1.5$ GeV^2 is shown in the left bottom panel of Fig. 4.

We can also test our results by considering how well they satisfy the classical chiral $V-A$ sum rule constraints represented by the two Weinberg sum rules [15] and the DGMLY sum rule for the π electromagnetic mass splitting [16]. These tests focus specifically on the DV part of our spectral functions, since the OPE for the $V-A$ correlator (and hence the OPE part of the spectral *ansatz*) has no $D = 0$ contribution and $D \geq 2$ OPE contributions to $\rho_{V/A}(s)$ are tiny, and can be ignored.

Weinberg's sum rules can be written as

$$\begin{aligned} \int_0^\infty ds \left(\rho_V^{(1+0)}(s) - \rho_A^{(1+0)}(s)\right) &= \int_0^\infty ds \left(\rho_V^{(1)}(s) - \rho_A^{(1)}(s)\right) - 2f_\pi^2 = 0 \,, \\ \int_0^\infty ds\, s \left(\rho_V^{(1+0)}(s) - \rho_A^{(1+0)}(s)\right) &= \int_0^\infty ds\, s \left(\rho_V^{(1)}(s) - \rho_A^{(1)}(s)\right) - 2m_\pi^2 f_\pi^2 = 0 \,, \end{aligned} \tag{6.11}$$

where we assumed, as before, that we can neglect terms of order $m_i m_j$ with $i, j = u, d$, even though there is a term of order $m_i m_j \alpha_s^2$ linearly divergent in s_0 in the second sum rule. The fact that this term is still very small at $s_0 = m_\tau^2$ amounts to the observation that the chiral

symmetry breaking terms in the second of Eq. (6.11) are not visible in the data. This means that in our test of the second sum rule we can assume ourselves to be effectively in the chiral limit, in which the divergence does not appear. In Fig. 6 we show the integrals in Eq. (6.11) as a function of the "switch point" s_{sw} below which experimental spectral data is used and above which the difference of the V and A DV *ansätze* of Eq. (2.15) is employed for $\rho_{V-A}(s)$. The DV contributions were obtained using the DV parameter values of Table 5. If the DV *ansatz* is, as we have assumed, reliable in the window of s_0 employed in the FESR fits which produce these DV parameter values, the $V-A$ sum rules should be satisfied for all values of s_{sw} lying in this s_0 fit window. We see that this condition is well satisfied for both of the Weinberg sum rules. The errors shown are those from the experimental ($s < s_{sw}$) part of the integral on the left-hand side of the sum rules only.

The first Weinberg sum rule is, of course, closely related to the difference of the V and A $\hat{w}_0(x) = 1$ FESRs. Therefore, the very good quality of the matches between the experimental spectral integrals and our theoretical representations thereof precludes the first Weinberg sum rule being badly broken by our fits. The second Weinberg sum rule (as well as the DGMLY sum rule discussed below) can be viewed as a prediction, because the fits of Table 5 do not involve any weight with a term linear in s. We also note that if indeed we take the second sum rule in the chiral limit, we should omit the term $-2m_\pi^2 f_\pi^2$. The value of this term is -0.00034 GeV^4, and it can thus indeed safely be dropped from the sum rule — the difference would not be visible in the figure.

Finally, we consider the DGMLY sum rule for the π electromagnetic mass difference. To leading order in the chiral expansion, one has that

$$\int_0^\infty ds\, s \log\left(s/\mu^2\right) \left(\rho_V^{(1)}(s) - \rho_A^{(1)}(s)\right) = -\frac{8\pi f_0^2}{3\alpha}\left(m_{\pi^\pm}^2 - m_{\pi^0}^2\right) \ , \tag{6.12}$$

where α is the fine-structure constant, and f_0 the π decay constant in the two-flavor chiral limit. On the right-hand side we take as input the values $f_0 = 87.0 \pm 0.6$ MeV [38] and $m_{\pi^+}^2 - m_{\pi^0}^2 = 0.00126 \pm 0.00008$ GeV^2 [19].[21] Because of the second Weinberg sum rule, the left-hand side of Eq. (6.12) is, in fact, independent of the scale μ. In Fig. 7 we show the left-hand and the right-hand sides of Eq. (6.12), as a function of s_{sw}. The left-hand side is represented by the data points, while the gray band represents the right-hand side,

[21] Our uncertainty on the electromagnetic contribution to the difference $m_{\pi^+}^2 - m_{\pi^0}^2$ comes from an estimate of the contribution of $m_u - m_d$ to the pion mass difference [39].

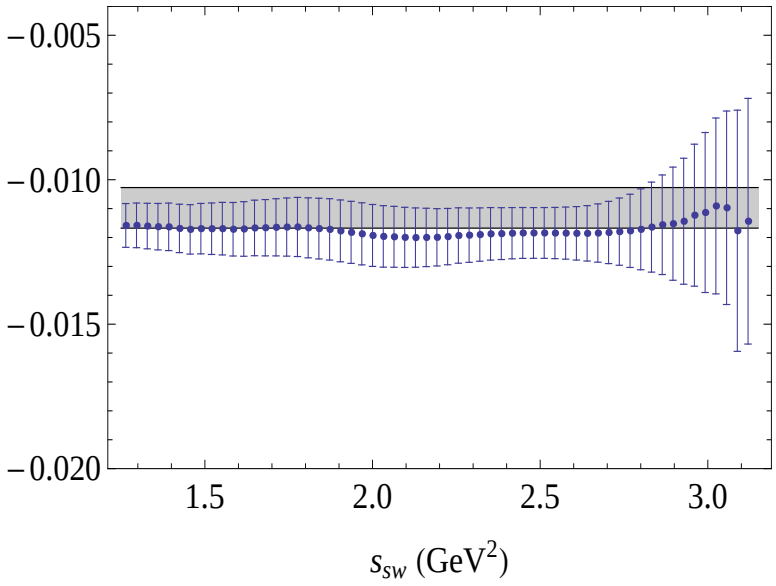

FIG. 7: *The DGMLY π electromagnetic mass difference sum rule: left-hand side (data points) and right-hand side (gray band) of Eq. (6.12), as a function of s_{sw}, in* GeV^4. *Data are used for $s \leq s_{sw}$, the DV ansatz (2.15), with fit parameter values from the $s_{min} = 1.5$ GeV^2 entries of Table 5, for $s \geq s_{sw}$. The FOPT figure is shown; the CIPT figure is essentially identical. The gray band represents an estimate of the error on the right-hand side of Eq. (6.12).*

including the error resulting from the uncertainties in f_0 and the pion electromagnetic mass difference. As for the Weinberg sum rules, the integral on the left-hand side is computed using experimental data for $s \leq s_{sw}$, and the DV *ansatz*, with DV parameters from the $s_{min} = 1.5$ GeV^2 entries of Table 5, for $s \geq s_{sw}$. The errors shown again come from the experimental data part of the integral only. We chose $\mu^2 = 2.5$ GeV^2, which makes the errors in the figure relatively small. For lower values of μ^2 the errors are larger for larger s_{sw}, but the results are consistent with Eq. (6.12) being satisfied for all μ^2 between 1.5 GeV^2 and m_τ^2.

One might consider using these sum rules as a further constraint on the DV parameters.[22] Of course, this can only be done for combined $V\&A$ fits, which requires us to assume that the larger of the two s_{min} values for the V and A channels still lies sufficiently far below m_τ^2

[22] See, for instance, Ref. [36], where, however, the *difference* $V - A$ was modeled with an expression of the form (2.15). Given the results of Tables 3 and 5, such a parametrization does not seem to be favored by the data. In the present work we avoid any additional assumptions about the relations between the DV parameters in the V and A channels by introducing separate DV parameters for each channel.

to make such a combined fit, using our DV *ansatz* in both channels, reliable. Since in this work our primary results are obtained from purely V-channel fits, which do not require this assumption, we postpone this possibility to a future investigation.

VII. SUMMARY OF RESULTS

Our most stable results come from fits to only the V channel. Moreover, since we need to include a model for DVs, using only the V channel avoids the additional assumption that our *ansatz* already adequately describes the A-channel data in the interval between 1.5 GeV^2 and m_τ^2. We emphasize, however, that all the results from our combined V and A channel fits are consistent with those from our V-channel analyses. In addition, they pass several further $V+A$ and $V-A$ channel tests, as demonstrated in Sec. VI D.

In view of these observations, we will choose the fit to $I_{ex}^{(\hat{w}_0)}$ in the vector channel as our central result. Adding the errors from the fit, the variation of s_{min} and the variation of c_{51} in quadrature, we find

$$\begin{aligned} \alpha_s(m_\tau^2) &= 0.307 \pm 0.019 \qquad (\overline{MS},\ n_f = 3,\ \text{FOPT}) \ , \\ \alpha_s(m_\tau^2) &= 0.322 \pm 0.026 \qquad (\overline{MS},\ n_f = 3,\ \text{CIPT}) \ . \end{aligned} \tag{7.1}$$

We do not average over CIPT and FOPT results, because we believe that it is useful to see the difference between values for α_s obtained with the two resummation schemes after all non-perturbative effects have been consistently taken into account. Running these values up to the Z mass M_Z yields [40][23]

$$\begin{aligned} \alpha_s(M_Z^2) &= 0.1169 \pm 0.0025 \qquad (\overline{MS},\ n_f = 5,\ \text{FOPT}) \ , \\ \alpha_s(M_Z^2) &= 0.1187 \pm 0.0032 \qquad (\overline{MS},\ n_f = 5,\ \text{CIPT}) \ , \end{aligned} \tag{7.2}$$

where we symmetrized the slightly asymmetric errors one obtains after running up to the Z mass.

These values can be compared to those obtained by OPAL from the same data [6]. The OPAL values at the τ mass are

$$\alpha_s(m_\tau^2) = 0.324 \pm 0.014 \qquad (\overline{MS},\ n_f = 3,\ \text{FOPT, OPAL}) \ , \tag{7.3}$$

[23] The specifics of the evolution to the Z mass are as discussed in Ref. [8].

$$\alpha_s(m_\tau^2) \;=\; 0.348 \pm 0.021 \qquad (\overline{MS},\; n_f = 3,\; \text{CIPT, OPAL})\;,$$

where we added the experimental and theoretical errors quoted by OPAL in quadrature. We observe the shift to lower central values, with somewhat larger errors that follow from using our new framework for analyzing the data. We also note that our CIPT and FOPT values are somewhat closer, and that, because of the larger errors, the difference between our two values for α_s is less significant.

VIII. CONCLUSION

In this article, we provided a new framework for the extraction of α_s and other OPE parameters from hadronic τ decays. This new framework combines two elements that have not been taken into account in the "traditional" analysis of hadronic τ decays. One is a consistent treatment of the OPE, as discussed in Ref. [12], and the other is a detailed quantitative estimate of the effect of violations of quark-hadron duality, using a parametrization proposed in Refs. [27, 29]. As explained in detail in Secs. III and IV, these two elements are intricately intertwined.

Our new framework comes with the price of introducing four new fit parameters for each channel, the vector and axial-vector DV parameters. Nevertheless, we demonstrated that our method is feasible by presenting a rather complete analysis based on the OPAL data for the V and A spectral functions [6]. With the larger number of parameters, and the corresponding need to vary s_0 away from m_τ^2, it should come as no surprise that our errors are typically larger than those found in earlier analyses, which simply did not take DVs into account quantitatively. We emphasize that this means that the systematic errors of those earlier analyses were understimated – we believe significantly so in some cases.

Our analysis leads us to a new estimate for both the central value of α_s and the error, which, in our opinion, should be interpreted as superseding previous estimates in the literature; for our result, based on OPAL data, see Sec. VII. We found our most reliable fits to be those of the V channel, although fits including also the A channel lead to results consistent with our most precise V-channel fits. As our primary concern in this article is with previously underestimated non-perturbative effects, we presented results for both contour-improved as well as fixed-order perturbation theory. Our analysis was based on the original OPAL data, unmodified for subsequent changes in the various exclusive branching

fractions. This choice was made in order to facilitate interpretation of the differences in our results from those obtained by OPAL. With this choice, these differences are solely the result of differences in the analysis method. We plan to present an analysis of OPAL data with updated normalizations in the near future.

The accuracy of the results presented here depends in part on our ability to correctly model the physics present in duality violations. Conservatively, our results can be seen as providing a lower bound on the error introduced by ignoring duality violations. However, the stability of α_s, a purely perturbative parameter, across the range of moments analyzed here provides strong support for the validity of our *ansatz*. We therefore surmise that not only is the *ansatz* able to accurately describe the data, but also that it provides a reasonable quantitative description of the physics of duality violations in the light-quark V and A channels.

In cases with multiple weights, standard χ^2 fits were not possible, and we performed alternate fits, propagating errors as described in App. A. We point out, however, that our final result is based on a standard χ^2 fit to the moment $I_{ex}^{(\hat{w}_0)}$. All other fits yield results completely consistent with this result, including the error on α_s. The χ^2 error on α_s obtained from our fit to $I_{ex}^{(\hat{w}_0)}$ is very close to that obtained with the method of App. A. Because our parameter covariance matrix obtained with that method scales linearly with the data covariance matrix, the error on α_s will be reduced once the improved spectral function data expected from BaBar and/or Belle becomes available.

We observe that a difference remains between the central values for α_s obtained using CIPT and FOPT, though this is less significant than the difference found previously by OPAL, *cf.* Eq. (7.3). In this context, we note that, as mentioned already in Sec. IV A, a term linear in s in any of the weights employed in Eq. (2.3) picks out the $D = 4$ term in the OPE, which parametrizes the leading renormalon ambiguity in the perturbative expansion. This, then, raises the question whether differences between the behavior of CIPT and FOPT fits, including those associated with any dependence on the choice of weight, might be used to constrain renormalon pole models that have been used previously to investigate the resummation of the perturbative series for the various sum rules employed in the study of hadronic τ decays [8, 11]. We plan to pursue such an investigation in a future work.

Acknowledgments

We would like to thank Martin Beneke, Claude Bernard, Andreas Höcker, Manel Martinez, and Ramon Miquel for useful discussions. We also would like to thank Sven Menke for significant help with understanding the OPAL spectral-function data. MG thanks IFAE and the Department of Physics at UAB, and OC and KM thank the Department of Physics and Astronomy at SFSU for hospitality. DB, MJ and SP are supported by CICYTFEDER-FPA2008-01430, SGR2005-00916, the Spanish Consolider-Ingenio 2010 Program CPAN (CSD2007-00042). SP is also supported by a fellowship from the Programa de Movilidad PR2010-0284. OC is supported in part by MICINN (Spain) under Grant FPA2007-60323, by the Spanish Consolider Ingenio 2010 Program CPAN (CSD2007-00042) and by the DFG cluster of excellence "Origin and Structure of the Universe." MG and JO are supported in part by the US Department of Energy, and KM is supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

Appendix A: Error propagation

Consider a fit quality

$$\mathcal{Q}^2 = [d_i - t_i(\vec{p})]\, C_{0,ij}^{-1}\, [d_j - t_j(\vec{p})] \ , \tag{A1}$$

in which d_i are the binned data, $t_i(\vec{p})$ is a function that describes this data set for a set of parameters $\vec{p}$, and C_0 is a positive-definite, symmetric, but otherwise arbitrary matrix. In this Appendix, we use the summation convention for repeated indices. The parameters $\vec{p}$ are determined by finding the global minimum of $\mathcal{Q}^2$, which satisfies

$$\frac{\partial \mathcal{Q}^2}{\partial p_\alpha} = -2\, \frac{\partial t_i(\vec{p})}{\partial p_\alpha}\, C_{0,ij}^{-1}\, [d_j - t_j(\vec{p})] = 0 \ . \tag{A2}$$

Varying, in this equation, the data by an amount δd_i, and the parameters by δp_α leads to

$$\frac{\partial^2 t_i(\vec{p})}{\partial p_\alpha \partial p_\beta}\, C_{0,ij}^{-1}\, [d_j - t_j(\vec{p})]\, \delta p_\beta + \frac{\partial t_i(\vec{p})}{\partial p_\alpha}\, C_{0,ij}^{-1} \left[\delta d_j - \frac{\partial t_j(\vec{p})}{\partial p_\beta}\, \delta p_\beta \right] = 0 \ . \tag{A3}$$

If the fit is good, so that the deviations $d_j - t_j(\vec{p})$ are small, we may ignore the term with the second derivative, leading to

$$\delta p_\alpha = A_{\alpha\beta}^{-1} \frac{\partial t_i(\vec{p})}{\partial p_\beta}\, C_{0,ij}^{-1}\, \delta d_j \ , \tag{A4}$$

or, for the covariance matrix $\langle \delta p_\alpha \delta p_\beta \rangle$,

$$\langle \delta p_\alpha \delta p_\beta \rangle = A_{\alpha\alpha'}^{-1} A_{\beta\beta'}^{-1} \frac{\partial t_i(\vec{p})}{\partial p_{\alpha'}} \frac{\partial t_j(\vec{p})}{\partial p_{\beta'}}\, C_{0,ik}^{-1} C_{0,j\ell}^{-1}\, C_{k\ell} \ , \tag{A5}$$

in which

$$A_{\alpha\beta} = \frac{\partial t_i(\vec{p})}{\partial p_\alpha}\, C^{-1}_{0,ij} \frac{\partial t_j(\vec{p})}{\partial p_\beta} \ , \tag{A6}$$

and

$$C_{k\ell} = \langle \delta d_k \delta d_\ell \rangle \tag{A7}$$

is the data covariance matrix. This provides us with an estimate for the full correlation matrix for the parameter set $\vec{p}$. We note that, if $C_{0,ij}$ is chosen to be equal to the data covariance matrix C_{ij}, this expression simplifies to

$$\langle \delta p_\alpha \delta p_\beta \rangle = A^{-1}_{\alpha\beta} \ . \tag{A8}$$

This is equal to the usual χ^2 error matrix estimate, given by the inverse of one-half times the second derivative of $\mathcal{Q}^2$ at its minimum, if, again, the fit is good enough to ignore terms proportional to $d_i - t_i(\vec{p})$.

---

[1] P. A. Baikov, K. G. Chetyrkin and J. H. Kühn, Phys. Rev. Lett. **101**, 012002 (2008) [arXiv:0801.1821 [hep-ph]].

[2] E. Braaten, S. Narison, and A. Pich, Nucl. Phys. B **373** 581 (1992).

[3] S. Bethke, Eur. Phys. J. **C64**, 689 (2009) [arXiv:0908.1135 [hep-ph]].

[4] R. Barate *et al.* [ALEPH Collaboration], Eur. Phys. J. **C4**, 409 (1998).

[5] S. Schael *et al.* [ALEPH Collaboration], Phys. Rept. **421**, 191 (2005) [arXiv:hep-ex/0506072].

[6] K. Ackerstaff *et al.* [OPAL Collaboration], Eur. Phys. J. C **7**, 571 (1999) [arXiv:hep-ex/9808019].

[7] M. Davier *et al.*, Eur. Phys. J. **C56**, 305 (2008) [arXiv:0803.0979 [hep-ph]].

[8] M. Beneke and M. Jamin, JHEP **0809**, 044 (2008) [arXiv:0806.3156 [hep-ph]].

[9] S. Menke, arXiv:0904.1796 [hep-ph].

[10] I. Caprini and J. Fischer, Eur. Phys. J. C **64**, 35 (2009) [arXiv:0906.5211 [hep-ph]]; I. Caprini and J. Fischer, Phys. Rev. **D84**, 054019 (2011), [arXiv:1106.5336 [hep-ph]].

[11] S. Descotes-Genon, B. Malaescu, arXiv:1002.2968 [hep-ph].

[12] K. Maltman, T. Yavin, Phys. Rev. **D78**, 094020 (2008) [arXiv:0807.0650 [hep-ph]].

[13] O. Catà, M. Golterman, S. Peris, Phys. Rev. **D79**, 053002 (2009) [arXiv:0812.2285 [hep-ph]].

[14] D. R. Boito, O. Catà, M. Golterman, M. Jamin, K. Maltman, J. Osborne, S. Peris, Nucl. Phys. Proc. Suppl. **218**, 104 (2011) [arXiv:1011.4426 [hep-ph]].

[15] S. Weinberg, Phys. Rev. Lett. **18**, 507 (1967).

[16] T. Das, G. S. Guralnik, V. S. Mathur, F. E. Low, J. E. Young, Phys. Rev. Lett. **18**, 759 (1967).

[17] R. Shankar, Phys. Rev. **D15**, 755 (1977); R. G. Moorhose, M. R. Pennington and G. G. Ross, Nucl. Phys. **B124**, 285 (1977); K. G. Chetyrkin and N. V. Krasnikov, Nucl. Phys. **B119**, 174 (1977); K. G. Chetyrkin, N. V. Krasnikov and A. N. Tavkhelidze, Phys. Lett. **76B**, 83 (1978); N. V. Krasnikov, A. A. Pivovarov and N. N. Tavkhelidze, Z. Phys. **C19**, 301 (1983); E. G. Floratos, S. Narison and E. de Rafael, Nucl. Phys. **B155**, 115 (1979); R. A. Bertlmann, G. Launer and E. de Rafael, Nucl. Phys. **B250**, 61 (1985).

[18] Y. -S. Tsai, Phys. Rev. **D4**, 2821 (1971).

[19] K. Nakamura *et al.* (Particle Data Group), J. Phys. **G37**, 075021 (2010).

[20] M. Jamin, JHEP **0509**, 058 (2005) [hep-ph/0509001].

[21] A. A. Pivovarov, Z. Phys. C **53**, 461 (1992) [Sov. J. Nucl. Phys. **54**, 676 (1991)] [Yad. Fiz. **54** (1991) 1114] [arXiv:hep-ph/0302003]; F. Le Diberder, A. Pich, Phys. Lett. **B286**, 147 (1992).

[22] S. Narison, V. I. Zakharov, Phys. Lett. **B679**, 355 (2009) [arXiv:0906.4312 [hep-ph]].

[23] K. G. Chetyrkin, V. P. Spiridonov, S. G. Gorishnii, Phys. Lett. **B160**, 149 (1985); A. Pich, J. Prades, JHEP **9910**, 004 (1999) [hep-ph/9909244].

[24] M. S. Dubovikov, A. V. Smilga, Nucl. Phys. **B185**, 109 (1981); W. Hubschmid, S. Mallik, Nucl. Phys. **B207**, 29 (1982).

[25] L. V. Lanin, V. P. Spiridonov, K. G. Chetyrkin, Sov. J. Nucl. Phys. **44**, 812 (1986) [Yad. Fiz. **44**, 1372 (1986)]; L. E. Adam, K. G. Chetyrkin, Phys. Lett. **B329**, 129 (1994) [hep-ph/9404331].

[26] E. C. Poggio, H. R. Quinn, S. Weinberg, Phys. Rev. **D13**, 1958 (1976).

[27] O. Catà, M. Golterman, S. Peris, Phys. Rev. **D77**, 093006 (2008) [arXiv:0803.0246 [hep-ph]].

[28] B. Blok, M. A. Shifman and D. X. Zhang, Phys. Rev. D **57**, 2691 (1998) [Erratum-ibid. D **59**, 019901 (1999)] [arXiv:hep-ph/9709333]; I. I. Y. Bigi, M. A. Shifman, N. Uraltsev, A. I. Vainshtein, Phys. Rev. **D59**, 054011 (1999) [hep-ph/9805241]; M. A. Shifman, [hep-ph/0009131]; M. Golterman, S. Peris, B. Phily, E. de Rafael, JHEP **0201**, 024 (2002) [hep-ph/0112042].

[29] O. Catà, M. Golterman, S. Peris, JHEP **0508**, 076 (2005) [hep-ph/0506004].

[30] M. Jamin, JHEP **1109**, 141 (2011), [arXiv:1103.2718 [hep-ph]].

[31] F. Le Diberder, A. Pich, Phys. Lett. **B289**, 165 (1992).

[32] S. Narison, Phys. Lett. **B673**, 30 (2009) [arXiv:0901.3823 [hep-ph]].

[33] M. Jamin, Phys. Lett. **B538**, 71 (2002) [hep-ph/0201174].

[34] K. Maltman, Phys. Lett. **B440**, 367 (1998) [hep-ph/9901239]; C. A. Dominguez, K. Schilcher, Phys. Lett. **B448**, 93 (1999) [hep-ph/9811261].

[35] M. Gell-Mann, R. J. Oakes, B. Renner, Phys. Rev. **175**, 2195 (1968).

[36] M. Gonzalez-Alonso, A. Pich, J. Prades, Phys. Rev. **D82**, 014019 (2010) [arXiv:1004.4987 [hep-ph]]; Phys. Rev. **D81**, 074007 (2010) [arXiv:1001.2269 [hep-ph]].

[37] M. Girone, M. Neubert, Phys. Rev. Lett. **76**, 3061 (1996) [hep-ph/9511392].

[38] A. Bazavov, *et al.* [MILC Collaboration], PoS **LATTICE2010**, 074 (2010) [arXiv:1012.0868 [hep-lat]].

[39] G. Amoros, J. Bijnens, P. Talavera, Nucl. Phys. **B602**, 87 (2001) [hep-ph/0101127].

[40] K. G. Chetyrkin, B. A. Kniehl, M. Steinhauser, Phys. Rev. Lett. **79**, 2184 (1997) [hep-ph/9706430].