PROCEEDINGS OF SCIENCE

α_s from non-strange hadronic τ decays

Santiago Peris*
Dept of Physics and IFAE-BIST, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Catalonia (Spain)
E-mail: peris@ifae.es

Diogo Boito
Instituto de Física de São Carlos, Univ. São Paulo CP 369, 13570-970, São Carlos, SP, Brazil
E-mail: boito@ifsc.usp.br

Maarten Golterman
Department of Physics and Astronomy, San Francisco State University, San Francisco, CA 94132, USA
E-mail: maarten@sfsu.edu

Kim Maltman
Department of Mathematics and Statistics, York University, Toronto, ON Canada M3J 1P3
CSSM, University of Adelaide, Adelaide, SA 5005 Australia
E-mail: kmaltman@yorku.ca

We review how the current precision attained in the extraction of $\alpha_s(m_\tau)$ from hadronic τ decays requires the inclusion of Duality Violations (DVs) in the analysis, even though these decays are largely dominated by perturbation theory. A weighted average using the OPAL and ALEPH experimental data yields $\alpha_s(m_Z) = 0.1165 \pm 0.0012$ and $\alpha_s(m_Z) = 0.1185 \pm 0.0015$ in fixed-order and contour-improved perturbation theory, respectively.

α_s(2019): Workshop on precision measurements of the QCD coupling constant
11-15 February, 2019
Trento, Italy

*Speaker.

PoS(ALPHAS2019)018

The α_s determination from hadronic τ decay usually relies on Finite Energy Sum Rules (FESRs). A FESR analysis takes advantage of the analyticity of the current-current correlator

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

where J_μ stands for the non-strange V or A current, $\bar{u}\gamma_\mu d$ or $\bar{u}\gamma_\mu\gamma_5 d$, and the superscripts (0) and (1) label spin, to obtain the following identity [1]

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

which is valid for any $s_0 > 0$ and any weight $w(s)$ analytic inside and on the contour depicted in Fig. 1. The combinations $\Pi^{(1+0)}(q^2)$ and $q^2\Pi^{(0)}(q^2)$ are convenient because they are free of kinematic singularities. In Eq. (2), $\rho^{(1+0)}(s) = \frac{1}{\pi}\,\mathrm{Im}\,\Pi^{(1+0)}(s)$ designates the spectral function and $s = q^2$. From now on, we will suppress the index $(1+0)$.

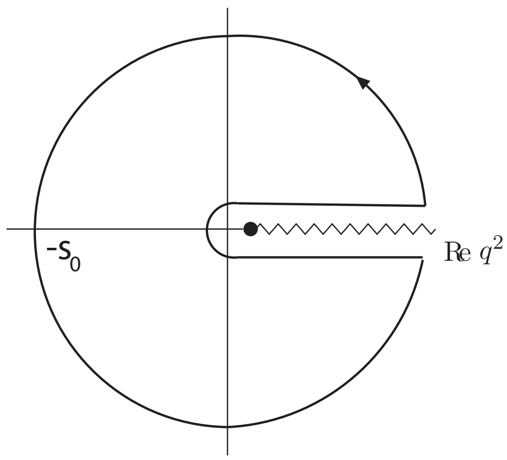

Figure 1: Contour used in the derivation of the FESRs, Eq. (2) .

As it stands, Eq. (2) is exact if for $\Pi(s)$ one is using the exact function. When s_0 is large enough, it begins to make sense to replace this function by its OPE representation from which it may be possible to extract the value of α_s[1]. As the OPE is expected to be asymptotic, and breaks down on the Minkowski axis, there will be a nonvanishing difference between the exact and the OPE representations. We will denote this difference by $\Pi_{DV}(s)$, where DV stands for Duality Violations (DVs). Explicitly,

$$\Pi(s) = \Pi_{OPE}(s) + \Pi_{DV}(s)\ . \tag{3}$$

In a hypothetical world in which the OPE converged, DVs would vanish by definition.

Using Eq. (3), one may rewrite Eq. (2) conveniently as [2]

$$\frac{1}{s_0}\int_0^{s_0} ds\, w(s)\,\rho^{exp}(s) = -\frac{1}{2\pi i s_0}\oint_{|z|=s_0} dz\, w(z)\,\Pi_{OPE}^{(\alpha_s)}(z) - \frac{1}{s_0}\int_{s_0}^{\infty} ds\, w(s)\,\frac{1}{\pi}\,\mathrm{Im}\,\Pi_{DV}(s)\ , \tag{4}$$

[1] Here, we will consider the perturbative series as the contribution from the unit operator to the OPE.

where we have indicated that $\rho^{exp}(s)$ is to be obtained from experimental data and $\Pi^{(\alpha_s)}_{OPE}(s)$ contains the value of α_s to be determined. In practice $w(s)$ will be taken to be a polynomial. At this point several important remarks are in order.

First, as a consequence of the residue theorem, a monomial $w(s) = (s/s_0)^N$ produces an OPE contribution $(-1)^N C_{2N+2}/s_0^{N+1}$ to the right-hand side of Eq. (4), where the coefficients C_{2N+2} are related to the condensates of dimension $2N+2$ (see also the discussion below). The C_{2N+2} are typically not known *a priori* but can be determined using Eq. (4) if α_s and $\mathrm{Im}\,\Pi_{DV}(s)$ have been previously determined. For a given set of experimental data $\rho^{exp}(s)$, the presence of DVs affects any C_{2N+2} determined in this way, except in the case $N = 0$ ($w(s) = 1$) where C_2 vanishes for the V and A correlators[2]. Second, although $\mathrm{Im}\,\Pi_{DV}(s)$ is certainly non-zero as a result of the non-convergence of the OPE on the Minkowski axis, its precise form is in principle unknown. In early τ-decay analyses this problem was dealt with by assuming that the use of polynomials $w(s)$ with zeros at $s = s_0$ of sufficiently high order ("pinching") would provide sufficient suppression of contributions from the region near the Minkowski axis to allow DVs to be safely neglected. This assumption is predicated on the expectation that Π_{DV} will be maximal in the vicinity of the Minkowski axis,[3] where, given that it represents a contribution missed by the asymptotic OPE, one expects $\mathrm{Im}\,\Pi_{DV}(s) \sim e^{-\gamma s} \times$ (oscillation), in analogy to the way the asymptotic renormalon series misses a non-perturbative term of order e^{-b/α_s}. Third, the use of pinching, regrettably, poses a problem: a polynomial with a high-order zero necessarily also has a high degree, and a high-degree polynomial produces contributions from C_{2N+2} with large N to the right-hand side of the FESR (4). Such contributions are not known (unless DVs and α_s have somehow already been determined, as remarked above). This leads to a "no-go" theorem [4, 5]:

"It is not possible to simultaneously suppress DV and high-dimension condensate contributions."

In order to avoid the contribution from C_{2N+2} with large N, one could use a low-degree polynomial, but this could then fail to provide enough pinching to be able to safely neglect DVs. In summary, one way or the other, the inclusion of DVs in Eq. (2) is unavoidable. This requires a concrete parametrization of $\mathrm{Im}\,\Pi_{DV}(s)$ which then allows its parameters to be determined with the help of Eq. (4), through a fit in an appropriate window of large-enough s_0.

Recently, such parametrization has been obtained [6]. The assumptions needed to derive it are rather mild: First, an asymptotic Regge spectrum for mesons at $N_c = \infty$ and, second, a constant width-over-mass ratio in the limit that the radial excitation number $\to \infty$, for $N_c = 3$. Both these assumptions are true in QCD in two dimensions (where all these properties can actually be computed), are supported by the string picture of hadrons [7], and are in agreement with phenomenology [8]. The resulting expression for $\mathrm{Im}\,\Pi_{DV}(s)$ then reads

$$\frac{1}{\pi}\mathrm{Im}\,\Pi_{DV}(q^2) \sim \mathrm{e}^{-2\pi\frac{a}{N_c}\,\frac{q^2}{\Lambda^2_{\mathrm{QCD}}}} \sin\left[\frac{2\pi}{\Lambda^2_{\mathrm{QCD}}}\left(q^2 - c - b\log\frac{q^2}{\Lambda^2_{\mathrm{QCD}}}\right)\right]\left(1+\mathscr{O}\left(\frac{1}{N_c};\frac{1}{q^2};\frac{1}{\log q^2}\right)\right)\,, \tag{5}$$

where $\Lambda_{\mathrm{QCD}} \sim 1\,\mathrm{GeV}$ is the characteristic QCD scale, related to the string tension. The result (5) is in accord with our expectations for an asymptotic OPE described above. Apart from a mild

[2] The u and d quark masses are very small and, consequently, neglected.

logarithmic dependence, modulated by the constant b and subleading at large q^2, this form can be conveniently expressed as [9]

$$\frac{1}{\pi}\mathrm{Im}\,\Pi_{DV}(q^2) = e^{-\delta-\gamma q^2}\,\sin\left(\alpha+\beta\, q^2\right), \qquad q^2 \gg \Lambda^2_{\mathrm{QCD}}\,, \tag{6}$$

and this is, in fact, the parametrization we have used in our analyses. We emphasize that, in principle, a different set of parameters $\delta_{V,A}, \gamma_{V,A}, \alpha_{V,A}$ and $\beta_{V,A}$ should be used for the V and A channels since they are related to the resonance spectrum.

This expression (6) was not available for use in the first τ decay determinations of α_s [10, 11, 12]. These pioneering analyses employed a strategy, which we will refer to as the truncated-OPE strategy (tOPE), in which pinched weights were used and both DVs and high-dimension OPE contributions were neglected. Recent examples of the continued use of this strategy may be found in Refs. [13, 14].

The tOPE strategy proceeds as follows. A set of five polynomials,

$$w_{kl}(y) = (1+2y)\,(1-y)^{2+k}\,y^l \quad , \quad y = s/s_0, \quad s_0 = m_\tau^2\,, \tag{7}$$

with $(k,l) \in \{(0,0),(1,0),(1,1),(1,2),(1,3)\}$, is chosen, and the corresponding set of weighted spectral integrals, evaluated at $s_0 = m_\tau^2$ only, used to extract four parameters: α_s and the coefficients $C_{D=4,6,8}$. Since these polynomials reach up to degree 7 in s, the FESR (4), in principle, receives contributions also from the C_D with $D = 10$, 12, 14 and 16, which, because they are unknown, are neglected. This neglect is predicated on an assumed $O\left(\Lambda^D_{\mathrm{QCD}}/m_\tau^D\right)$ suppression of dimension D OPE contributions. In other words, the OPE is effectively treated as if it were convergent at the scale $s = m_\tau^2$. The term with DVs in Eq. (4) is also neglected. While this strategy may have been reasonable in the early work of Refs. [10, 11, 12], when the error in the extracted α_s was $\sim 10-15\%$, it is clear that, as time goes by, and errors decrease, the assumptions underpinning this approach need to be checked and, if necessary, the method needs to be revised.

Partly with this idea in mind, Ref. [14] has recently generalized the tOPE strategy by investigating a variety of alternate polynomial combinations, obtaining, in all cases, consistent results with good-quality fits. However, although the analysis of Ref. [14] showed no obvious sign that the results obtained might be unreliable and the value of α_s extracted might be polluted by a systematic error that the variations studied might not be capable of identifying, all these results, as shown in Ref. [15], *do* contain a hidden $\sim +6\%$ systematic error in the extracted value of $\alpha_s(m_\tau^2)$. For a full account of this systematic error, we refer to Ref. [15]. Here, we will just report on one particularly clean test that illustrates the point.

The test works as follows. We consider a model designed to closely match the actual experimental spectral data, but constructed to have an input value of $\alpha_s(m_\tau^2)$, $\alpha_s(m_\tau^2)^{fake} = 0.312$, and corresponding chosen values for the DV parameters[3]. We then generate a set of fake data for the $V+A$ spectral function[4], using exactly the same binning and the same correlations as in the actual experiment, by letting the data points fluctuate according to a multivariant Gaussian distribution defined with the experimental covariance matrix.[5] An example of the resulting spectral distribution

[3]See Ref.[15], for more details.

[4]This is the channel that Ref. [14] considers to be optimal for the reliability of the tOPE strategy.

[5]The fake data is generated only for $s_0 \geq 1.55\,\mathrm{GeV}^2$, which is the value we obtained in our true-data fits for the onset of the asymptotic DV expression (6). Below this s_0, the two data sets, fake and true, are identical.

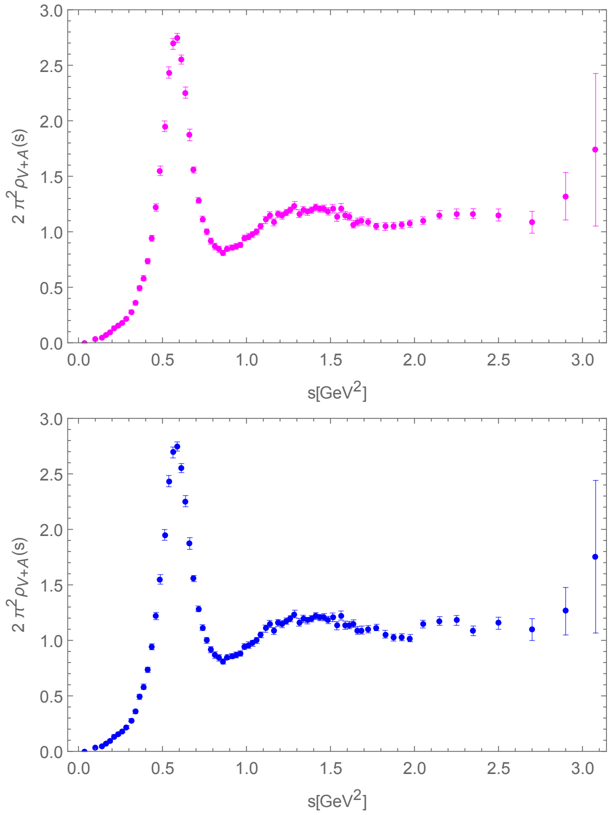

Figure 2: Left: $V+A$ fake data, generated as described in the text, as a function of s. Right: True ALEPH data [13] as a function of s. The fake data has been generated for $s \geq 1.55\,\mathrm{GeV}^2$; below this value the two sets of data are identical.

is shown in Fig. 2. The point is that, if the tOPE strategy were reliable, it should be able to reproduce the input value of $\alpha_s(m_\tau^2)^{fake}$ from the fake data set, within errors, and in spite of the neglect of higher-dimension OPE contributions and the absence of a representation of integrated DVs in the theoretical form it assumes.

However, when we use the sets of polynomials suggested in Ref. [14] in fits to the fake data, we always find the value of $\alpha_s(m_\tau^2)$ to be overestimated by $\sim +0.02$, a systematic error of $+6\%$, which, in terms of the statistical errors of these fits, amounts to 5–7σ. Therefore, the tOPE strategy clearly fails. One might think that the tOPE could have also reproduced the right result, had the fake data set been generated without DVs. Such fake data, however, would not be able to reproduce the residual oscillations present even in the $V+A$ spectral distribution (see below). And the fact remains that the tOPE strategy, in ignoring higher-dimension terms in the OPE without justification, and failing to detect the presence of residual DVs in the fake data case, can produce a systematic shift in the extracted value of α_s whose presence cannot be exposed by looking at the variation in the output α_s produced when the tOPE analysis is performed using the various polynomial set choices considered in Ref. [14].

It has been argued [14], referring to the left panel of Fig. 3, that the spectral function in the $V+A$ channel is so flat at high s as to be free from DVs. This argument, however, is rather

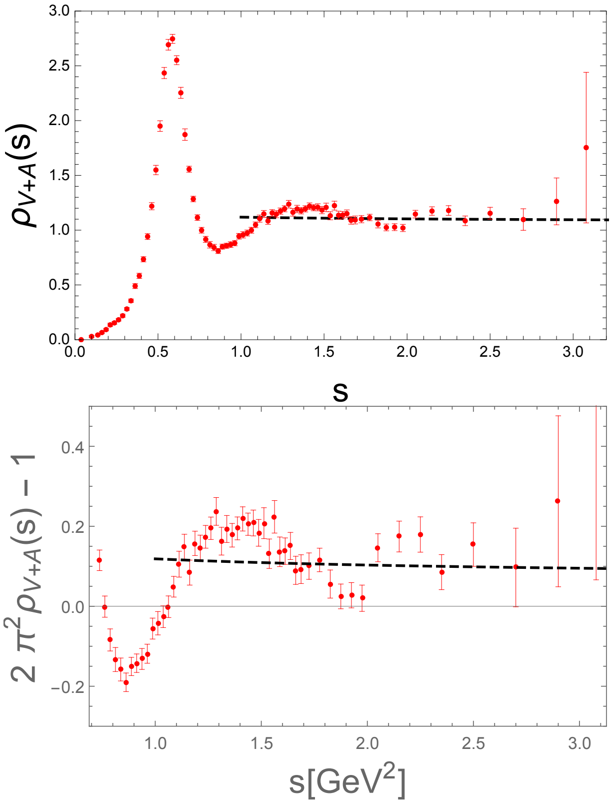

Figure 3: Left: $V+A$ spectral function. Right: $V+A$ spectral function, after the parton model contribution has been subtracted. The black dashed curve is the result of perturbation theory.

misleading. The right panel in Fig. 3 shows the same spectral function, but now with the (α_s-independent) parton-model contribution subtracted. The black dashed curve in this panel shows the corresponding result from perturbation theory. One sees that, even at $s = 2\,\mathrm{GeV}^2$, the data points agree, within errors, with the parton model. In other words, the α_s-dependent part of the perturbative contribution cancels against the DV oscillation at that point. There is no sense in which the DVs are small *relative* to the α_s-dependent perturbative contributions, from which the value of α_s is extracted, and, therefore, there is no sense in which the DVs may be reliably neglected. A similar effect is seen at $s \simeq 2.2\,\mathrm{GeV}^2$ but, this time, DVs and the α_s-dependent part of perturbation theory add up, rather than cancel each other. Again, the size of DVs is comparable to that of the α_s-dependent perturbative contributions. Notice that these data are very correlated, so the fact that a group of three data points, with central values very close together at $s \simeq 2.2\,\mathrm{GeV}^2$, are above the perturbative curve while another group of three data points, again very close together at $s \simeq 2\,\mathrm{GeV}^2$, are below the perturbative curve is difficult to explain as a fluctuation in the data, and not as the sign of a true residual oscillation. Above $s = 2.5\,\mathrm{GeV}^2$ the errors are too large to tell. Furthermore, there is no doubt that both V and A separately contain DV oscillations, so the safest assumption is that $V+A$ also has them, even if they are smaller for $V+A$ than for the individual V and A cases. At any rate, smaller or not, we have seen that they can easily affect the extraction of α_s, as illustrated in the fake-data test discussed above. Reference [15] contains more details of the different tests one may carry out, all of them leading to the conclusion that the tOPE strategy is unreliable, with an associated systematic error of $\sim +0.02$ in the value of $\alpha_s(m_\tau^2)$.

Given this state of affairs, we have recently proposed [5, 16] a different strategy that takes DVs into account explicitly, parametrized as in Eq. (6), and employs the 3 polynomials (to be considered together or separately)

$$w_0 = 1\ , \quad w_2 = 1 - y^2 \quad \text{and} \quad w_3 = (1-y)^2(1+2y)\ . \tag{8}$$

The α_s value and the two coefficients $C_{6,8}$ (which are the only OPE coefficients contributing to the w_2 and w_3 FESRs) are then fit using the integrated spectral data in a window of s_0 extending from m_τ^2 down to a lower value determined by the fit itself. The choice of polynomials is dictated by a desire to avoid, first, contributions from high-dimension terms in the OPE and, second, the use polynomials with a term linear in y (which receive a contribution from C_4, associated with the gluon condensate) since model studies suggest that perturbation theory behaves poorly for such weights, whether with the FOPT or the CIPT choice for the scale μ [17][6].

A large variety of different fits using Eq. (4) and the three polynomials above, and employing the V channel alone, or the V and A channels combined, were carried out in Ref. [16], to which we refer for more details. The results obtained were consistent in all cases, not only for α_s but also for the $C_{6,8}$ coefficients.

In Fig. (4) we show how the associated V and A spectral functions are described by our parametrization in Eq. (6) at high s, where the asymptotic DV form is expected to apply. A number of additional consistency checks were also carried out; for example, the first Weinberg sum rule. In Fig. 5 we show the result of this sum rule, i.e.,

$$\int_0^\infty ds\, (\rho_V(s) - \rho_A(s)) - 2f_\pi^2 = 0\ , \tag{9}$$

as a function of the point s_{sw} at which one switches from the experimental data to the corresponding theoretical description. The left panel shows the case without DVs in the theoretical description; in this case, the experimental data switches to zero since the perturbative contribution cancels in the $V-A$ difference. The right panel shows the case where the DV parametrization (6) is employed in the theoretical description, which, through the second term on the righthand side of Eq. (4), allows us to extend the upper limit in the sum rule to infinity. It is clear that taking DVs into account constitutes an improvement.

Other tests were also considered. Two tests we find particularly interesting probe the idea of truncating the OPE. Using again the polynomials w_{kl} in Eq. (7), the left panels of Fig. 6 show the example of w_{11} and w_{13} as a function of s_0, as obtained with the tOPE strategy. We emphasize that, within this strategy, fits are being done solely at $s_0 = m_\tau^2$. Therefore, it is not surprising that the data agree rather well with the theory curve at this s_0. However, the theory description quickly departs from the data as soon as s_0 is lowered, which is a clear sign that the s_0 scaling on the theory side of the corresponding FESR is not correct. This is a consequence of neglecting the higher-dimension terms in the OPE that contribute to these sum rules. For comparison, we also show the same result once DVs are taken into account, and the corresponding OPE coefficients have been determined

[6]Fixed-order perturbation theory (FOPT) refers to the choice of the scale $\mu^2 = s_0$, where s_0 is the radius of the contour in Eq. (4). Contour-improved perturbation theory (CIPT) refers to the choice $\mu^2 = z$, where z is the complex integration variable along the contour in Eq. (4).

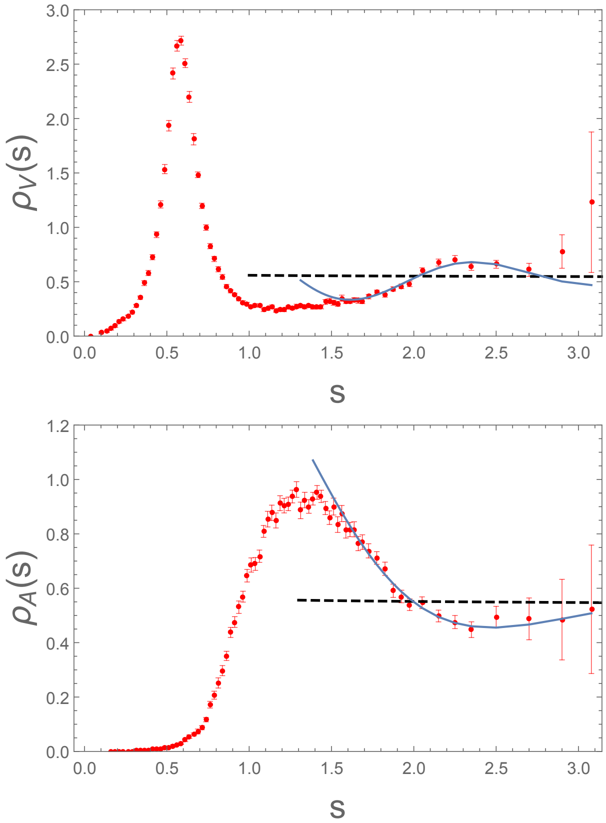

Figure 4: Left: V spectral function, together with the parton model result (dashed black curve) and the result from the DV parametrization (6) obtained from Eq. (4) (blue curve). Right: The same for the A spectral function.

with the help of Eqs. (7.3) of Ref. [16] (which are versions of the FESR (4)). We emphasize that what this figure shows is that, as a result of the absence of higher-dimension terms assumed negligible in the tOPE strategy, the tOPE version of the theory side of the FESRs (4) fails to scale correctly with s_0 as s_0 decreases below m_τ^2. In other words, the argument that the scale m_τ is large enough to effectively suppress the contributions from the higher-dimension C_D to the FESR (4), based on an assumed naive $\Lambda_{\mathrm{QCD}}^D/m_\tau^D$ scaling, turns out to be incorrect. This is compatible with the known asymptotic character of the OPE, which implies that these coefficients must eventually become significantly larger than implied by this naive scaling for sufficiently large dimension D.

It would be very instructive to be able to test this assumption about the simple $C_D/s_0^{D/2}$ suppression for scales $s_0 \geq m_\tau^2$. Clearly, if the higher-dimension terms in the OPE are suppressed at the scale m_τ^2, they should be even more suppressed at scales larger than m_τ^2. Although, regrettably, it is not possible to test this with the τ data, with some mild assumptions it is possible to do so using data for $e^+e^- \to$ hadrons [18, 19].

Using the so-called "optimal" weights proposed in Ref. [14]:

$$w^{(2,n)}(y) = 1-(n+2)y^{n+1}+(n+1)y^{n+2}\,, \tag{10}$$

with $n = 1,\ldots,5$, which are doubly pinched, one may determine α_s and $C_{6,8,10}$ at $s_0 = m_\tau^2$, provided

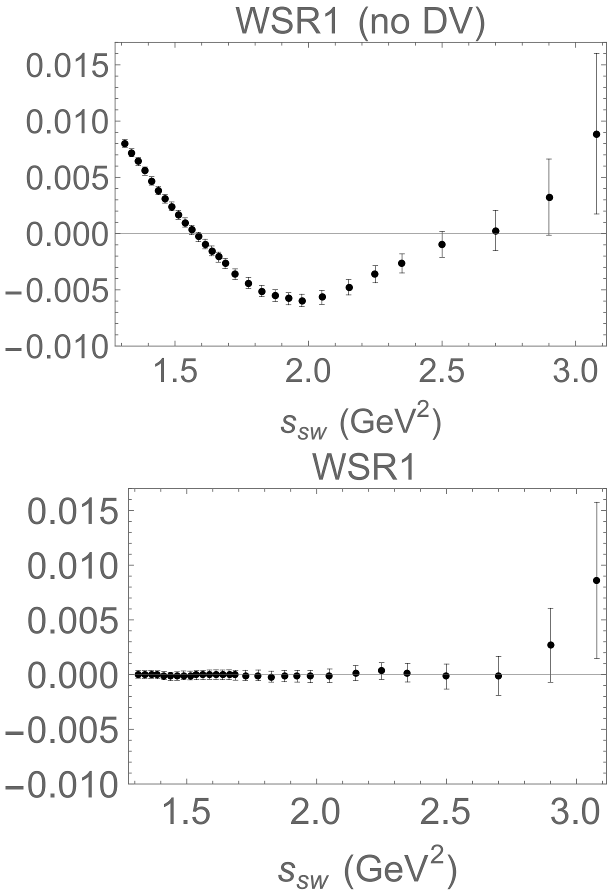

Figure 5: Left: First Weinberg sum rule without DVs. Right: First Weinberg sum rule with DVs taken into account.

one neglects $C_{12,14,16}$ in the FESRs (4). In the SU(3) limit, one finds that the correlator of two electromagnetic currents is $2/3$ times the correlator of two isospin currents, as they would appear in the V channel in τ decay. Consequently, the physics of these two situations cannot be very different. In Ref. [19] we presented a preliminary version of this type of analysis. The result is shown in Fig. 7. In this figure we plot the result for the difference between the contribution of the OPE to the FESR (i.e., the righthand side of Eq. (4) without the DV term) at a variable s_0 minus the same for $s_0 = m_\tau^2$, as a function of s_0[7]. The result is represented by the two black curves (dashed for CIPT and solid for FOPT). We also plot the same difference, but now computed with the e^+e^- data as the red points. The fact that they both agree, and vanish at $s = s_0$, is nothing but a consequence of our definition. What is more interesting is that, not only for $s_0 < m_\tau^2$ but also for $s_0 > m_\tau^2$, the two descriptions clearly disagree. This is, again, a rather clear sign that the assumption that higher-dimension terms in the OPE are negligible is not supported by the data. Similar conclusions follow from using the weights of Eq. (7) instead of those of Eq. (10).

In summary, we have presented conclusive evidence that the neglect of higher-dimension terms in the OPE and DVs at the core of the truncated OPE strategy leads to an irreducible systematic

[7]We do this to account for correlations.

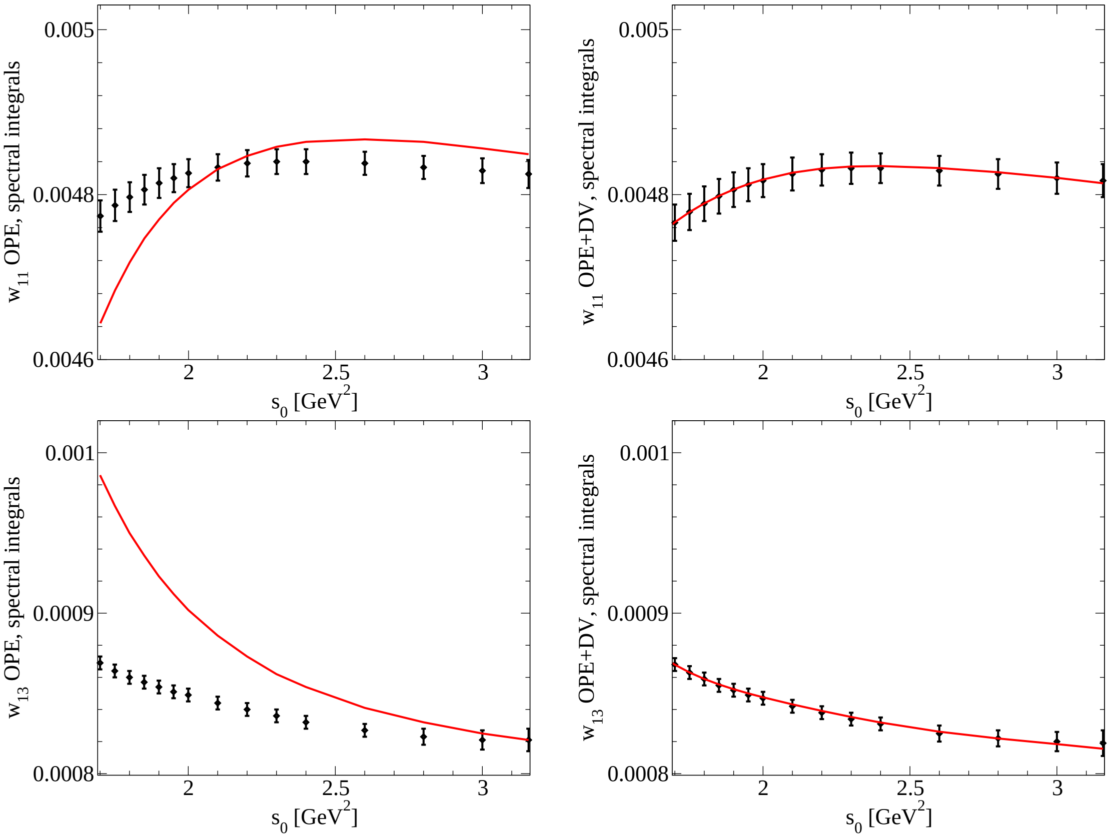

Figure 6: Comparison of the agreement between the lefthand and righthand sides of the FESR (4) for the weights w_{11} and w_{13} within the tOPE strategy, with DVs and high-dimension condensates neglected (left), and with DVs taken into account and condensates determined via Eqs. (7.3) of Ref. [16] (right).

error of the order of $+0.02$ in the value of $\alpha_s(m_\tau^2)$. This method should therefore be considered unreliable and, consequently, no longer be used, at least not without adding a potential $\sim +0.02$ systematic error to the tOPE results. As an alternative, we have proposed a strategy that parametrizes DVs as in Eq. (6) and includes them in the analysis from the start, and which makes no *a priori* assumptions about the values of the relevant OPE coefficients, which are to be determined by the data through fits employing Eq. (4) in a window of values of s_0 ranging up to m_τ^2. The result of these fits to ALEPH data leads to [16]

$$\begin{aligned}
\alpha_s(m_\tau) &= 0.296 \pm 0.010 \longrightarrow \alpha_s(m_Z) = 0.1155 \pm 0.0014 \qquad (\text{FOPT}) , \\
\alpha_s(m_\tau) &= 0.310 \pm 0.014 \longrightarrow \alpha_s(m_Z) = 0.1174 \pm 0.0019 \qquad (\text{CIPT}) .
\end{aligned} \tag{11}$$

Combining these results with those based on the OPAL data, we obtain as our final result

$$\alpha_s(m_Z) = 0.1165 \pm 0.0012 \quad (\text{FOPT}) , \qquad \alpha_s(m_Z) = 0.1185 \pm 0.0015 \quad (\text{CIPT}) . \tag{12}$$

These results are in very good agreement with the value for α_s obtained from the same type of FESRs using the e^+e^- data below the charm threshold [20]. We emphasize that, in this case, the s_0 values being used are sufficiently larger than m_τ^2 to make the contribution from DVs marginal,

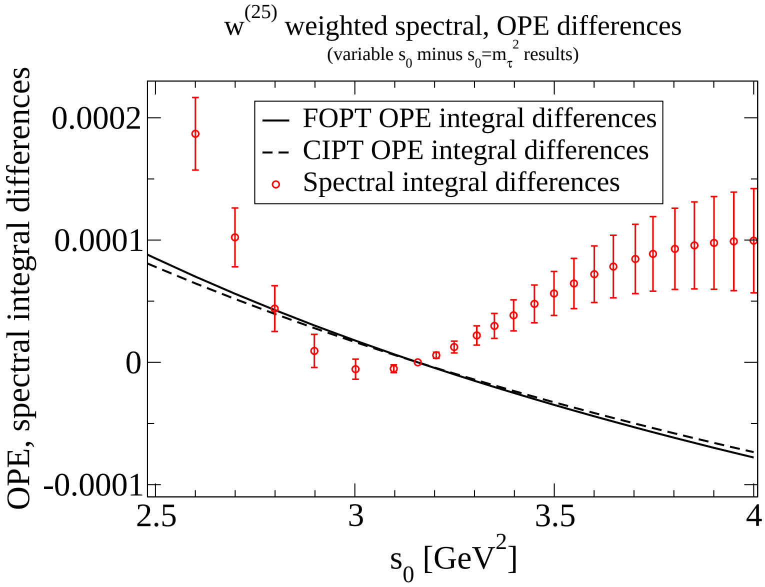

Figure 7: Electromagnetic FESR tests of the tOPE strategy using the set of weights of Eq. (10). Comparisons of differences between a variable s_0 and $s_0 = m_\tau^2$ versions of the OPE and spectral integrals. The OPE parameter values are obtained from the implementation of the tOPE strategy using the weights of Eq. (10) and $s_0 = m_\tau^2$ only in the fits.

if not negligible. We should also recall that the τ-based results rely on the assumption that our theory representation, which is expected to be valid for asymptotically large s_0, holds in a region of s extending down to below the τ mass. The good agreement shown in Fig. 4 and the consistency with the value obtained from e^+e^- is evidence for the validity of this assumption.

Acknowledgments. The work of D.B. is supported by the São Paulo Research Foundation (Fapesp) grant No. 2015/20689-9 and by the Brazilian National Council for Scientific and Technological Development (CNPq), grant No. 305431/2015-3. This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under Award Number DE-SC0013682 (MG). KM is supported by a grant from the Natural Sciences and Engineering Research Council of Canada, and SP by CICYTFEDER-FPA2014-55613-P, 2014-SGR-1450 and the CERCA Program/Generalitat de Catalunya.

References

[1] R. Shankar, Phys. Rev. D **15**, 755 (1977); R. G. Moorhouse, M. R. Pennington and G. G. Ross, Nucl. Phys. B **124**, 285 (1977); K. G. Chetyrkin and N. V. Krasnikov, Nucl. Phys. B **119**, 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. C **19**, 301 (1983); E. G. Floratos, S. Narison and E. de Rafael, Nucl. Phys. B **155**, 115 (1979); R. A. Bertlmann, G. Launer and E. de Rafael, Nucl. Phys. B **250**, 61 (1985).

[2] O. Catà, M. Golterman and S. Peris, Phys. Rev. D **79**, 053002 (2009) [arXiv:0812.2285 [hep-ph]].

[3] E. C. Poggio, H. R. Quinn and S. Weinberg, Phys. Rev. D **13**, 1958 (1976)

[4] S. Peris, D. Boito, M. Golterman and K. Maltman, Mod. Phys. Lett. A **31**, 1630031 (2016) [arXiv:1606.08898 [hep-ph]].

[5] D. Boito, O. Catà, M. Golterman, M. Jamin, K. Maltman, J. Osborne and S. Peris, Phys. Rev. D **84**, 113006 (2011) [arXiv:1110.1127 [hep-ph]]; D. Boito, M. Golterman, M. Jamin, A. Mahdavi, K. Maltman, J. Osborne and S. Peris, Phys. Rev. D **85**, 093015 (2012) [arXiv:1203.3146 [hep-ph]].

[6] D. Boito, I. Caprini, M. Golterman, K. Maltman and S. Peris, Phys. Rev. D **97**, 054007 (2018) [arXiv:1711.10316 [hep-ph]].

[7] M. Shifman and A. Vainshtein, Phys. Rev. D **77**, 034002 (2008) [arXiv:0710.0863 [hep-ph]].

[8] P. Masjuan, E. Ruiz Arriola and W. Broniowski, Phys. Rev. D **85**, 094006 (2012) [arXiv:1203.4782 [hep-ph]].

[9] O. Catà, M. Golterman and S. Peris, JHEP **08**, 076 (2005) [hep-ph/0506004].

[10] F. Le Diberder and A. Pich, Phys. Lett. B **289**, 165 (1992).

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

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

[13] M. Davier, A. Höcker, B. Malaescu, C. Z. Yuan and Z. Zhang, Eur. Phys. J. C **74**, 2803 (2014) [arXiv:1312.1501 [hep-ex]].

[14] A. Pich and A. Rodríguez-Sánchez, Phys. Rev. D **94**, 034027 (2016) [arXiv:1605.06830 [hep-ph]].

[15] D. Boito, M. Golterman, K. Maltman and S. Peris, Phys. Rev. D **95**, 034024 (2017) [arXiv:1611.03457 [hep-ph]].

[16] D. Boito, M. Golterman, K. Maltman, J. Osborne and S. Peris, Phys. Rev. D **91**, 034003 (2015) [arXiv:1410.3528 [hep-ph]].

[17] M. Beneke, D. Boito and M. Jamin, JHEP **01**, 125 (2013) [arXiv:1210.8038 [hep-ph]].

[18] See, e.g., A. Keshavarzi, D. Nomura and T. Teubner, Phys. Rev. D **97**, 114025 (2018) [arXiv:1802.02995 [hep-ph]].

[19] D. Boito, M. Golterman, K. Maltman and S. Peris, SciPost Phys. Proc. **1**, 053 (2019) [arXiv:1811.01581 [hep-ph]].

[20] D. Boito, M. Golterman, A. Keshavarzi, K. Maltman, D. Nomura, S. Peris and T. Teubner, Phys. Rev. D **98**, 074030 (2018) [arXiv:1805.08176 [hep-ph]].