PHYSICAL REVIEW D **89,** 014027 (2014)

Systematic scale-setting to all orders: The principle of maximum conformality and commensurate scale relations

Stanley J. Brodsky[*]
SLAC National Accelerator Laboratory, Stanford University, Stanford, California 94039, USA

Matin Mojaza[†]
CP3-Origins, Danish Institute for Advanced Studies, University of Southern Denmark, DK-5230, and SLAC National Accelerator Laboratory, Stanford University, Stanford, California 94039, USA

Xing-Gang Wu[‡]
Department of Physics, Chongqing University, Chongqing 401331, People's Republic of China
(Received 29 April 2013; revised manuscript received 6 January 2014; published 28 January 2014)

We present in detail a new systematic method which can be used to automatically eliminate the renormalization scheme and scale ambiguities in perturbative QCD predictions at all orders. We show that all of the nonconformal β-dependent terms in a QCD perturbative series can be readily identified by generalizing the conventional renormalization schemes based on dimensional regularization. We then demonstrate that the nonconformal series of pQCD at any order can be resummed systematically into the scale of the QCD coupling in a unique and unambiguous way due to a special degeneracy of the β terms in the series. The resummation follows from the principal of maximum conformality (PMC) and assigns a unique scale for the running coupling at each perturbative order. The final result is independent of the initial choices of renormalization scheme and scale, in accordance with the principles of the renormalization group, and thus eliminates an unnecessary source of systematic error in physical predictions. We exhibit several examples known to order α_s^4; i.e. (i) the electron-positron annihilation into hadrons, (ii) the tau-lepton decay to hadrons, (iii) the Bjorken and Gross-Llewellyn Smith (GLS) sum rules, and (iv) the static quark potential. We show that the final series of the first three cases are all given in terms of the anomalous dimension of the photon field in $SU(N)$, in accordance with conformality, and with all nonconformal properties encoded in the running coupling. The final expressions for the Bjorken and GLS sum rules directly lead to the generalized Crewther relations, exposing another relevant feature of conformality. The static quark potential shows that PMC scale-setting in the Abelian limit is to all orders consistent with QED scale-setting. Finally, we demonstrate that the method applies to any renormalization scheme and can be used to derive commensurate scale relations between measurable effective charges, which provide nontrivial tests of QCD to high precision. This work extends Brodsky-Lepage-Mackenzie (BLM) scale-setting to any perturbative order, with no ambiguities in identifying β terms in pQCD, demonstrating that BLM scale-setting follows from a principle of maximum conformality.

DOI: 10.1103/PhysRevD.89.014027 PACS numbers: 12.38.Aw, 11.10.Gh, 11.15.Bt, 12.38.Bx

I. INTRODUCTION

An important goal in high energy physics is to make perturbative QCD (pQCD) predictions as precise as possible, not only to test QCD itself, but also to expose new physics beyond the standard model. Recently, we showed a systematic method to determine the argument of the running coupling order by order in pQCD, and in a way that can be readily automatized [1,2]. The new method satisfies all of the principles of the renormalization group [3], and it eliminates an unnecessary source of systematic error. The resulting predictions for physical processes are independent of theoretical conventions such as the choice of renormalization scheme and the initial choice of renormalization scale. The resulting scales also determine the effective number of quark flavors at each order of perturbation theory. The method can be applied to processes with multiple physical scales and is consistent with QED scale-setting.

In this paper we review the method in detail and provide the complete generalization. We show several examples, based on observables recently published in the literature to four-loop order in perturbation theory. Finally, we demonstrate that the method applies to any renormalization scheme and can be used to derive commensurate scale relations between measurable effective charges [4–7]. The method extends the Brodsky-Lepage-Mackenzie (BLM) method [8]

[*]sjbth@slac.stanford.edu
[†]mojaza@cp3-origins.net
[‡]wuxg@cqu.edu.cn

to any perturbative order by following the principle of maximum conformality [9–11], without leaving any ambiguity in identifying β terms at any order in pQCD.

Previous attempts made in the literature to extend BLM scale-setting to higher orders [12–22], have mostly focused on improving convergence of the perturbative series and not removing renormalization scheme and scale ambiguities. Therefore they do not in general satisfy the self-consistency requirements of the renormalization group [3], nor the initial renormalization scale dependence, which must be the prerequisite of any scale-setting method. The main contribution of this work is to provide the systematic method to eliminate the renormalization scale and scheme ambiguities to all orders in pQCD.

Other recent proposals similar in spirit are suggested in Refs. [23,24].

We start our analysis in Sec. II by introducing a generalization of the conventional schemes used in dimensional regularization in which a constant $-\delta$ is subtracted in addition to the standard subtraction $\ln 4\pi - \gamma_E$ of the $\overline{\rm MS}$ scheme. The δ subtraction defines an infinite set of renormalization schemes which we call δ-$\mathcal{R}$enormalization ($\mathcal{R}_\delta$) schemes; since physical results cannot depend on the choice of scheme, predictions must be independent of δ. As will be described in Sec. III, the $\mathcal{R}_\delta$ scheme exposes the general pattern of nonconformal $\{\beta_i\}$ terms, and it reveals a special degeneracy of the terms in the perturbative coefficients which allows us to resum the perturbative series. The resummed series matches the conformal series, which is itself free of any scheme and scale ambiguities as well as being free of a divergent renormalon series. It is the final expression one should use for physical predictions. It also makes it possible to set up an algorithm for automatically computing the conformal series and setting the effective scales for the coupling at each perturbative order.

In Sec. IV we provide several examples, based on observables recently published in the literature to order α_s^4; i.e. (i) the electron-positron annihilation into hadrons, (ii) the tau-lepton decay to hadrons, (iii) the Bjorken and Gross-Llewellyn Smith (GLS) sum rules, and (iv) the static quark potential. We show explicitly that the final series of the first three cases are all given in terms of the anomalous dimension of the photon field in $SU(N)$, in accordance with conformality, and with all nonconformal properties encoded in the running coupling. Moreover, the final expressions for the Bjorken and GLS sum rules directly lead to the generalized Crewther relations, exposing another relevant feature of conformality. The static quark potential furthermore provides an example of how the method can be automatized to give the principal of maximum conformality (PMC) prediction directly from the number of quark flavor dependence of the initial expression. From this example we also demonstrate that the PMC prediction in the Abelian limit is consistent with QED scale-setting. Finally, we demonstrate in Sec. V that the method applies to any renormalization scheme and can be used to derive commensurate scale relations between measurable effective charges, which provide nontrivial tests of QCD to high precision.

II. THE $\delta - \mathcal{R}$ RENORMALIZATION SCHEME

In dimensional regularization logarithmically divergent integrals are regularized by computing them in $d = 4 - 2\epsilon$ dimensions [25–27]. This requires the following transformation of the integration measure and introduction of an arbitrary mass scale μ:

$$\int \mathrm{d}^4 p \to \mu^{2\epsilon} \int \mathrm{d}^{4-2\epsilon} p. \tag{1}$$

Divergences are then separated as $1/\epsilon$ poles and can be absorbed into redefinitions of the couplings. The choice of subtraction procedure is known as the *renormalization scheme* and is chosen at the theorist's convenience. To avoid dealing with coupling constants changing dimensionality as a function of ϵ one rescales the couplings as well with the mass scale μ in the $d = 4 - 2\epsilon$ theory. In particular, for QCD one rewrites the bare gauge coupling $a_0 = \alpha_0/4\pi = g^2/(4\pi)^2$ as

$$a_0 = \mu^{2\epsilon} Z_{a_S} a_S, \tag{2}$$

where a_S is the *renormalized* gauge coupling under a specific renormalization scheme S and Z_{a_S} is the renormalization constant of the coupling. The mass scale μ is now understood as the *renormalization scale*. The bare coupling must be independent of the arbitrary scale μ, thus

$$\mu^2 \frac{da_0}{d\mu^2} = 0. \tag{3}$$

Using this and the expansions

$$\mu^2 \frac{da_S}{d\mu^2} = -\epsilon a_S + \beta(a_S), \tag{4}$$

$$\beta(a) = -a^2 \sum_{i=0}^{\infty} \beta_i a^i, \tag{5}$$

$$Z_a = 1 + \sum_{i=1}^{\infty} z_i a^i, \tag{6}$$

it is easily derived that

$$Z_a = 1 - \frac{\beta_0}{\epsilon} a + \left(\frac{\beta_0^2}{\epsilon^2} - \frac{\beta_1}{2\epsilon}\right) a^2 - \left(\frac{\beta_0^3}{\epsilon^3} - \frac{7}{6}\frac{\beta_0\beta_1}{\epsilon^2} + \frac{\beta_2}{3\epsilon}\right) a^3 + \left(\frac{\beta_0^4}{\epsilon^4} - \frac{23\beta_1\beta_0^2}{12\epsilon^3} + \frac{5\beta_2\beta_0}{6\epsilon^2} + \frac{3\beta_1^2}{8\epsilon^2} - \frac{\beta_3}{4\epsilon}\right) a^4 + \cdots, \tag{7}$$

and the β_i coefficients are known up to β_3, or four loops [28]. The coefficients β_i are renormalization-scheme dependent; however, it is easy to demonstrate by a general scheme transformation that the first two coefficients β_0 and β_1 are universal for all mass-independent renormalization schemes.

In the minimal subtraction (MS) scheme [29], one absorbs the $1/\epsilon$ poles appearing in loop integrals which come in powers of

$$\ln\frac{\mu^2}{\Lambda^2}+\frac{1}{\epsilon}+c, \tag{8}$$

where c is the finite part of the integral. Since anything can be hidden into infinity, one can subtract any finite part as well with the pole. This is equivalent to redefining the arbitrary scale μ in Eq. (1). The $\overline{\rm MS}$ scheme [30] differs from the MS scheme by an additional absorption of the term $\ln(4\pi)-\gamma_E$, which corresponds to redefining μ to

$$\mu^2=\mu^2_{\overline{\rm MS}}\exp(\ln 4\pi-\gamma_E). \tag{9}$$

We will generalize this by defining the *δ-$\mathcal{R}$enormalization scheme*, $\mathcal{R}_\delta$, where one absorbs $\ln(4\pi)-\gamma_E-\delta$, i.e.

$$\mu^2=\mu^2_\delta\exp(\ln 4\pi-\gamma_E-\delta), \tag{10}$$

where δ is an arbitrary finite number, and by appropriate choice will connect all MS-type schemes. In particular,[1]

$$\mathcal{R}_0=\overline{\rm MS}, \tag{11a}$$

$$\mathcal{R}_{\ln 4\pi-\gamma_E}={\rm MS}, \tag{11b}$$

$$\mathcal{R}_{-2}={\rm G}, \tag{11c}$$

where we also provided the connection to the G scheme, which is yet another MS-like scheme proposed in the literature [31].

The scheme transformation between different $\mathcal{R}_\delta$ corresponds simply to a displacement in their corresponding scales, i.e.

$$\mu^2_{\delta_2}=\mu^2_{\delta_1}\exp(\delta_2-\delta_1). \tag{12}$$

In particular,

$$\mu^2_\delta=\mu^2_{\overline{\rm MS}}\exp(\delta). \tag{13}$$

Since all $\mathcal{R}_\delta$'s are connected by scale displacements, the β functions of $a_{\mathcal{R}_\delta}$ defined in Eq. (4) are the same for all $\mathcal{R}_\delta$ to any order. The index δ on $a_{\mathcal{R}_\delta}$ is thus redundant and we denote it instead as $a_{\mathcal{R}}$. Where it will not be ambiguous, we will simply use $a\equiv a_{\mathcal{R}}$.

We can find a power series solution in $1/\ln(\mu/\Lambda)$ for a by solving the renormalization group equation perturbatively. It is simplest to use the extended renormalization group prescription [32,33] where one works with a rescaled coupling and a rescaled logarithm; respectively

$$\hat{a}=\frac{\beta_1}{\beta_0}a,\qquad L_\delta=\frac{\beta_0^2}{\beta_1}\ln(\mu_\delta/\Lambda).$$

The solution up to $\mathcal{O}(L_\delta^{-5})$ reads [2,9]

$$\begin{aligned}\hat{a}(\mu_\delta)=&\frac{1}{L_\delta}+\frac{1}{L_\delta^2}(\mathcal{C}-\ln L_\delta)+\frac{1}{L_\delta^3}[\mathcal{C}^2+\mathcal{C}+c_2\\&-(2\mathcal{C}-\ln L_\delta+1)\ln L_\delta-1]\\&+\frac{1}{L_\delta^4}\left\{\mathcal{C}\left(\mathcal{C}^2+\frac{5}{2}\mathcal{C}+3c_2-2\right)-\frac{1-c_3}{2}\right.\\&-\left[3\mathcal{C}^2+5\mathcal{C}+3c_2-2-\left(3\mathcal{C}-\ln L_\delta+\frac{5}{2}\right)\ln L_\delta\right]\\&\left.\times\ln L_\delta\right\}+\mathcal{O}\left(\frac{1}{L_\delta^5}\right),\end{aligned} \tag{14}$$

where $c_i=\beta_i\beta_0^{i-1}/\beta_1^i$ are the rescaled β-function coefficients and $\mathcal{C}$ is an arbitrary integration constant which in $\mathcal{R}_\delta$ is set to $\mathcal{C}=\ln\beta_0^2/\beta_1$ in order to reproduce the standard $\Lambda_{\overline{\rm MS}}$ scale [2,9,33]. Note that we take the asymptotic scale $\Lambda=\Lambda_{\overline{\rm MS}}$ to be the same for any $\mathcal{R}_\delta$. Alternatively, one can take the scale μ to be the same for any $\mathcal{R}_\delta$, having instead different asymptotic scales Λ_δ.

This solution for $\hat{a}(\mu_\delta)$ holds for massless QCD or as long as the active quark masses are below Λ. In the case where there are active quarks with masses higher than Λ, one must take the quark threshold effects into account when running the coupling from Λ to μ_δ. This can be done by e.g. solving the renormalization group equation with an analytic β function that takes quark masses into account [34] or by using matching equations at each quark mass threshold when running the coupling to μ_δ [35].

III. OBSERVABLES IN $\mathcal{R}_\delta$

Consider an observable in pQCD in some scheme which we put as the reference scheme $\mathcal{R}_0$ with the following expansion:

$$\rho_0(Q^2)=a(\mu_0)^n\sum_{k=0}^{\infty}r_{k+1}(Q^2/\mu_0^2)a(\mu_0)^k, \tag{15}$$

where μ_0 stands for some initial renormalization scale and Q is the kinematic scale of the process. The full pQCD series is formally independent of the choice of the initial renormalization scale μ_0, if it were possible to sum the entire series. However, this goal is not feasible in practice, especially because of the $k!\beta^k\alpha_s^k$ renormalon growth of the nonconformal series. When a perturbative expansion is truncated at any finite order, it generally becomes renormalization scale and scheme dependent; i.e., dependent on theoretical conventions. This can be exposed by using the $\mathcal{R}_\delta$ scheme. Since results in any $\mathcal{R}_\delta$ are related by scale displacements, we can derive the general expression for ρ in $\mathcal{R}_\delta$ by using the displacement relation between couplings in any $\mathcal{R}_\delta$ scheme:

[1]Note that we have chosen $\overline{\rm MS}$ as the reference scheme for $\mathcal{R}_0$. This is done since most results today are known in this scheme; however, there is nothing special about $\overline{\rm MS}$, and $\mathcal{R}_0$ can be redefined to be any other MS-like scheme.

$$a(\mu_0) = a(\mu_\delta) + \sum_{n=1}^{\infty} \frac{1}{n!} \frac{\mathrm{d}^n a(\mu)}{(\mathrm{d}\ln \mu_0^2)^n}\bigg|_{\mu=\mu_\delta} (-\delta)^n, \quad (16)$$

where we used $\ln \mu_0^2/\mu_\delta^2 = -\delta$. It is useful to derive the general displacement relation for $a(\mu_0)^k$ for any k as an expansion in a up to order a^{k+3}:

$$a(\mu_0)^k = a(\mu_\delta)^k + k\beta_0\delta a(\mu_\delta)^{k+1} + k\left[\beta_1\delta + \frac{k+1}{2}\beta_0^2\delta^2\right] a(\mu_\delta)^{k+2} + k\left[\beta_2\delta + \frac{2k+3}{2}\beta_0\beta_1\delta^2 + \frac{(k+1)(k+2)}{3!}\beta_0^3\delta^3\right] a(\mu_\delta)^{k+3}. \quad (17)$$

Inserting this expression into Eq. (15) at each order $a(\mu_0)^k$ we find the expression for ρ for an arbitrary δ to order a^4, that is in any $\mathcal{R}_\delta$ scheme, to be

$$\begin{aligned}\rho_\delta(Q^2) = {} & r_1 a_1(\mu_\delta)^n + [r_2 + n\beta_0 r_1\delta_1] a_2(\mu_\delta)^{n+1} + \left[r_3 + n\beta_1 r_1\delta_1 + (n+1)\beta_0 r_2\delta_2 + \frac{n(n+1)}{2}\beta_0^2 r_1\delta_1^2\right] a_3(\mu_\delta)^{n+2} \\ & + \left[r_4 + n\beta_2 r_1\delta_1 + (n+1)\beta_1 r_2\delta_2 + (n+2)\beta_0 r_3\delta_3 + \frac{n(3+2n)}{2}\beta_0\beta_1 r_1\delta_1^2 + \frac{(n+1)(n+2)}{2}\beta_0^2 r_2\delta_2^2 \right. \\ & \left. + \frac{n(n+1)(n+2)}{3!}\beta_0^3 r_1\delta_1^2\right] a_4(\mu_\delta)^{n+3} + \mathcal{O}(a^5), \end{aligned} \quad (18)$$

where $\mu_\delta^2 = Q^2 e^\delta$, the initial scale is for simplicity set to $\mu_0^2 = Q^2$, and we have defined in Eq. (15) $r_i(1) = r_i$. An artificial index was introduced on each a and δ to keep track of which coupling each δ term is associated with. They are not an indication of different variables; i.e. $a_1 = a_2 = \ldots$ and $\delta_1 = \delta_2 = \ldots$. The use of the artificial indices will be made clear in a moment.

The above expression shows the scheme dependence explicitly; e.g. if $\mathcal{R}_0 = \overline{\mathrm{MS}}$, then choosing $\delta = \delta_i = \ln 4\pi - \gamma_E$ will give the result in the MS scheme. The initial scale choice is arbitrary and is not the final argument of the running coupling; the final scales will be independent of the initial renormalization scale.

In a conformal (or scale-invariant) theory, where $\{\beta_i\} = \{0\}$, the δ dependence vanishes in Eq. (18). Therefore by absorbing all $\{\beta_i\}$ dependence into the running coupling at each order, we obtain a final result independent of the initial choice of scale and scheme. The coefficients in the final expression will thus be equal to those of the conformal theory. The use of $\mathcal{R}_\delta$ allows us to put this on rigorous grounds. From the explicit expression in Eq. (18) it is easy to confirm that

$$\frac{\partial\rho_\delta}{\partial\delta} = -\beta(a)\frac{\partial\rho_\delta}{\partial a}. \quad (19)$$

The scheme invariance of the physical prediction requires that $\partial\rho_\delta/\partial\delta = 0$. Therefore the scales in the running coupling must be shifted and set such that the conformal terms associated with the β function are removed; the remaining conformal terms are by definition renormalization scheme independent. The numerical value for the prediction at finite order is then scheme independent as required by the renormalization group. The scheme-invariance criterion is a theoretical requirement of the renormalization group; it must be satisfied at any truncated order of the perturbative series, and is different from the formal statement that the all-orders expression for a physical observable is renormalization scale and scheme invariant; i.e. $\mathrm{d}\rho/\mathrm{d}\mu_0 = 0$. The final series obtained corresponds to the theory for which $\beta(a) = 0$; i.e. the conformal series. This demonstrates to any order the concept of the *principal of maximum conformality* (PMC) [9], which states that all nonconformal terms in the perturbative series must be resummed into the running coupling.

The expression in Eq. (18) exposes the pattern of $\{\beta_i\}$ terms in the coefficients at each order. It is possible to infer more from Eq. (18); since there is nothing special about a particular value of δ, we conclude that some of the coefficients of the $\{\beta_i\}$ terms are degenerate; e.g. the coefficient of $\beta_0 a(Q)^2$ and $\beta_1 a(Q)^3$ can be set equal. Thus for any scheme, the expression for ρ can be put to the form

$$\begin{aligned}\rho(Q^2) = {} & r_{1,0} a(Q)^n + [r_{2,0} + n\beta_0 r_{2,1}] a(Q)^{n+1} + \left[r_{3,0} + n\beta_1 r_{2,1} + (n+1)\beta_0 r_{3,1} + \frac{n(n+1)}{2}\beta_0^2 r_{3,2}\right] a(Q)^{n+2} \\ & + \left[r_{4,0} + n\beta_2 r_{2,1} + (n+1)\beta_1 r_{3,1} + (n+2)\beta_0 r_{4,1} + \frac{n(3+2n)}{2}\beta_0\beta_1 r_{3,2} + \frac{(n+1)(n+2)}{2}\beta_0^2 r_{4,2} \right. \\ & \left. + \frac{n(n+1)(n+2)}{3!}\beta_0^3 r_{4,3}\right] a(Q)^{n+3} + \mathcal{O}(a^{n+4}), \end{aligned} \quad (20)$$

where the $r_{i,0}$ are the conformal parts of the perturbative coefficients; i.e. $r_i = r_{i,0} + \mathcal{O}(\{\beta_i\})$.

The $\mathcal{R}_\delta$ scheme not only illuminates the $\{\beta_i\}$ pattern, but it also exposes a *special degeneracy* of coefficients at different orders. The degenerate coefficients can themselves be functions of $\{\beta_i\}$, hence Eq. (20) is not to be understood as an expansion in $\{\beta_i\}$, but at a pattern of $\{\beta_i\}$ with degenerate coefficients that must be matched.

The artificial indices in the expansion in Eq. (18) reveals how the $\{\beta_i\}$ terms must be absorbed into the running coupling: The different δ_k's keep track of the power of the $1/\epsilon$ divergence of the associated diagram at each loop order in the following way; the $\delta_k^p a^m$ term indicates the term associated to a diagram with $1/\epsilon^{1+m-n-k}$ divergence for any power p of δ. Grouping the different δ_k terms, one recovers in the $N_c \to 0$ Abelian limit [36] the dressed skeleton expansion. Resumming the series according to this expansion thus correctly reproduces the QED limit of the observable and matches the conformal series with the running coupling evaluated at effective scales at each order. Using this information we can rearrange the expression in Eq. (20) in the skeleton-like expansion:

$$\begin{aligned}\rho(Q^2) = a(Q)^n\Big[r_{1,0} + n(\beta_0 a(Q) + \beta_1 a(Q)^2 + \beta_2 a(Q)^3) r_{2,1} + \frac{n}{2}((n+1)\beta_0^2 a(Q)^2 + (3+2n)\beta_0\beta_1 a(Q)^3) r_{3,2} \\ + \frac{n(n+1)(n+2)}{3!}\beta_0^3 r_{4,3} a(Q)^3\Big] + a(Q)^{n+1}\Big[r_{2,0} + (n+1)(\beta_0 a(Q) + \beta_1 a(Q)^2) r_{3,1} + \frac{(n+1)(n+2)}{2}\beta_0^2 r_{4,2} a(Q)^2\Big] \\ + a(Q)^{n+2}[r_{3,0} + (n+2)\beta_0 r_{4,1} a(Q)] + a(Q)^{n+3}[r_{4,0}] + \mathcal{O}(a^{n+4}).\end{aligned} \tag{21}$$

A. Systematic all-orders PMC scale-setting

It is easy to see from Eq. (21) that we can resum all $r_{i,1}$ terms, which come with a linear factor of β_j, *to all orders* by defining new scales Q_i at each order as follows (for simplicity, we treat the higher-power β_j terms later):

$$\begin{aligned} r_{1,0} a(Q_1)^n &= r_{1,0} a(Q)^n - n a(Q)^{n-1}\beta(a) r_{2,1}, \\ r_{2,0} a(Q_2)^{n+1} &= r_{2,0} a(Q)^{n+1} - (n+1) a(Q)^n \beta(a) r_{3,1}, \\ r_{3,0} a(Q_3)^{n+2} &= r_{3,0} a(Q)^{n+2} - (n+2) a(Q)^{n+1}\beta(a) r_{4,1}, \\ &\vdots \\ r_{k,0} a(Q_k)^k &= r_{k,0} a(Q)^k - k a(Q)^{k-1}\beta(a) r_{k+1,1}. \end{aligned} \tag{22}$$

From the scale displacement equation (16) for a it is straightforward to see that

$$\begin{aligned} a(Q_k)^k = a(Q)^k + k a(Q)^{k-1}\beta(a)\ln\frac{Q_k^2}{Q^2} \\ + \frac{k}{2} a(Q)^{k-2}\left[\beta\frac{\partial\beta}{\partial a} a(Q) + (k-1)\beta(a)^2\right]\ln^2\frac{Q_k^2}{Q^2} \\ + \cdots. \end{aligned} \tag{23}$$

It follows from identifying Eq. (22) with (23) that to absorb all linear β_j terms, the scales Q_k must satisfy

$$-\frac{r_{k+1,1}}{r_{k,0}} = \ln\frac{Q_k^2}{Q^2} + \frac{1}{2}\left[\frac{\partial\beta}{\partial a} + (k-1)\frac{\beta}{a}\right]\ln^2\frac{Q_k^2}{Q^2} + \cdots, \tag{24}$$

where $r_{k,0}$ are the conformal coefficient and $r_{k+1,1}$ are the degenerate coefficients of linear β_j terms. This leads to the self-consistency equation for Q_k:

$$\ln\frac{Q_k^2}{Q^2} = \frac{-r_{k+1,1}/r_{k,0}}{1 + \frac{1}{2}\left[\frac{\partial\beta}{\partial a} + (k-1)\frac{\beta}{a}\right]\ln\frac{Q_k^2}{Q^2} + \cdots}. \tag{25}$$

To leading logarithmic order (LLO) we have

$$\ln\frac{Q_{k,\mathrm{LLO}}^2}{Q^2} = -\frac{r_{k+1,1}}{r_{k,0}}. \tag{26}$$

This resums all linear β_j terms, but introduces higher-power β_j terms beyond the order a^{k+1}. For example, suppose that we are computing an observable to order a^p. The scales Q_k must resum all $\beta_j r_{k+1,1}$ terms without introducing higher order ones up to order a^p. This means that Q_k must be computed to the $p-(k+1)$ logarithmic order $(\mathrm{N}^{p-(k+1)})$LLO. Let us explicitly perform the resummation up to a^4 for the first scale Q_1, that is, up to next-to-next-to-leading logarithmic order (NNLLO). The general expression for the NLLO scale reads

$$\ln\frac{Q_{k,\mathrm{NLLO}}^2}{Q^2} = \frac{-r_{k+1,1}/r_{k,0}}{1 + \frac{1}{2}\left[\frac{\partial\beta}{\partial a} + (k-1)\frac{\beta}{a}\right]\left(-\frac{r_{k+1,1}}{r_{k,0}}\right)}. \tag{27}$$

To find the NNLLO scale, we first write the self-consistency equation (exposing one higher logarithmic order in the denominator):

$$\ln\frac{Q_k^2}{Q^2} = \frac{-r_{k+1,1}/r_{k,0}}{1 + \frac{1}{2}\left[\frac{\partial\beta}{\partial a} + (k-1)\frac{\beta}{a}\right]\ln\frac{Q_k^2}{Q^2} + \frac{1}{3!}\left[\beta\frac{\partial^2\beta}{\partial a^2} + \left(\frac{\partial\beta}{\partial a}\right)^2 + 3(k-1)\frac{\beta}{a}\frac{\partial\beta}{\partial a} + (k-1)(k-2)\frac{\beta^2}{a^2}\right]\ln^2\frac{Q_k^2}{Q^2} + \cdots}. \tag{28}$$

Then we replace the logarithms in the denominator with the expansion of its NLLO expression in Eq. (27):

$$\ln\frac{Q^2_{k,\mathrm{NLLO}}}{Q^2} = -\frac{r_{k+1,1}}{r_{k,0}}\left(1+\frac{1}{2}\left[\frac{\partial\beta}{\partial a}+(k-1)\frac{\beta}{a}\right]\frac{r_{k+1,1}}{r_{k,0}}+\cdots\right). \tag{29}$$

We thus get

$$\ln\frac{Q^2_{k,\mathrm{NNLLO}}}{Q^2} = \frac{-r_{k+1,1}/r_{k,0}}{1+\frac{1}{2}[\frac{\partial\beta}{\partial a}+(k-1)\frac{\beta}{a}](-\frac{r_{k+1,1}}{r_{k,0}})+\frac{1}{3!}[\beta\frac{\partial^2\beta}{\partial a^2}-\frac{1}{2}(\frac{\partial\beta}{\partial a})^2-\frac{(k-1)(k+1)}{2}\frac{\beta^2}{a^2}](\frac{r_{k+1,1}}{r_{k,0}})^2}. \tag{30}$$

This procedure iterates to any desired order.

For observables, where the higher-power β_j coefficients vanish, i.e. $r_{3,2}=r_{4,2}=r_{4,3}=0$ (this is the case in e.g. the Adler D function), these scales give the final PMC expression for the observable, which is invariant under any scheme transformation:

$$\rho(Q^2) = r_{1,0}a(Q_{n,\mathrm{NNLLO}})^n + r_{2,0}a(Q_{n+1,\mathrm{NLLO}})^{n+1} + r_{3,0}a(Q_{n+2,\mathrm{LLO}})^{n+2} + r_{4,0}a(Q)^{n+3} + \mathcal{O}(a^{n+4}). \tag{31}$$

This is the conformal series with coefficients that are independent of the renormalization scheme. Note that the last scale remains ambiguous. This ambiguity only affects the highest order term. The final expression and coefficients are therefore not affected by the ambiguity of the last scale and thus the renormalization scheme dependence has been eliminated and the renormalization scale dependence only resides in the highest power coupling of the perturbative series. We note that one does not need the full expression of the a^5 coefficient to set the last scale, Q_{n+3}, but only the coefficient $r_{5,1}$.

Let us now generalize to observables that do depend on higher powers in β_j. This is for example the case in $R_{e^+e^-\to\mathrm{hadrons}}(s)$. One can use the procedure just described, but instead of Eq. (22) we use its generalization:

$$r_{k,0}a(Q_k)^k = r_{k,0}a(Q)^k - ka(Q)^{k-1}\beta(a)r_{k+1,1} + \frac{k}{2}\left[a(Q)^{k-1}\frac{\mathrm{d}\beta}{\mathrm{d}\ln\mu^2}+(k-1)a(Q)^{k-2}\beta(a)^2\right]\times r_{k+2,2}+\cdots. \tag{32}$$

It is easy to verify that these expressions, which define the PMC scales Q_k, correctly resum all $\{\beta_i\}$ terms in ρ. Equation (32) is systematically derived by replacing the $\ln^j Q_1^2/Q^2$ by $r_{k,j}$ in the logarithmic expansion of $a(Q_k)^k$ in Eq. (23) up to the highest known $r_{k,n}$ coefficient in pQCD. We introduce a short-hand notation of Eq. (32):

$$a(Q_k)^k = a(Q)^k + ka(Q)^{k-1}\beta(a)\{R_{k,1}+\Delta^{(1)}_k(a)R_{k,2} + \Delta^{(2)}_k(a)R_{k,3}+\cdots+\Delta^{(n)}_k(a)R_{k,n+1}\}, \tag{33}$$

where

$$R_{k,j} = (-1)^j\frac{r_{k+j,j}}{r_{k,0}}, \tag{34a}$$

$$\Delta^{(1)}_k(a) = \frac{1}{2}\left[\frac{\partial\beta}{\partial a}+(k-1)\frac{\beta}{a}\right], \tag{34b}$$

$$\Delta^{(2)}_k(a) = \frac{1}{3!}\left[\beta\frac{\partial^2\beta}{\partial a^2}+\left(\frac{\partial\beta}{\partial a}\right)^2+3(k-1)\frac{\beta}{a}\frac{\partial\beta}{\partial a} + (k-1)(k-2)\frac{\beta^2}{a^2}\right],\ldots. \tag{34c}$$

Following the same procedure as before, one finds the final expressions for $Q_{k,\mathrm{LLO}}$, $Q_{k,\mathrm{NLLO}}$, and $Q_{k,\mathrm{NNLLO}}$ to be

$$\ln\frac{Q^2_{k,\mathrm{LLO}}}{Q^2} = R_{k,1}, \tag{35a}$$

$$\ln\frac{Q^2_{k,\mathrm{NLLO}}}{Q^2} = \frac{R_{k,1}+\Delta^{(1)}_k(a)R_{k,2}}{1+\Delta^{(1)}_k(a)R_{k,1}}, \tag{35b}$$

$$\ln\frac{Q^2_{k,\mathrm{NNLLO}}}{Q^2} = \frac{R_{k,1}+\Delta^{(1)}_k(a)R_{k,2}+\Delta^{(2)}_k(a)R_{k,3}}{1+\Delta^{(1)}_k(a)R_{k,1}+(\Delta^{(1)}_k(a))^2(R_{k,2}-R^2_{k,1})+\Delta^{(2)}_k(a)R^2_{k,1}},\ldots. \tag{35c}$$

These final expression are generic and can be used directly. We thus have a procedure which systematically sets the PMC scales to all orders.It is easy to see that the leading order values of the effective scales are independent of the initial renormalization scale μ_0. This follows since taking $\mu_0 \neq Q$ we must replace $R_{k,1} \to R_{k,1} + \ln Q^2/\mu_0^2$ and thus the leading order effective scales read $\ln Q^2_{k,LO}/\mu_0^2 = R_{k,1} + \ln Q^2/\mu_0^2$, where μ_0 cancels and Eq. (35a) at LLO is recovered. This generalizes to any order. Since the β function is not known to all orders, a higher order residual renormalization-scale dependence will enter through the running coupling. This residual renormalization-scale dependence is strongly suppressed in the perturbative regime of the coupling [37,38].

The effective scales contain all the information of the nonconformal parts of the initial pQCD expression for ρ in Eq. (20); this is exactly the purpose of the running coupling. The quotient form of Eq. (35c) sums up an infinite set of terms related to the known $r_{j,k\neq 0}$ which appear at every higher order due to the special degeneracy of Eq. (20). The method systematically sums up all known nonconformal terms, in principle to all orders, but is in practice truncated due to the limited knowledge of the β function.

In earlier PMC scale-setting [9–11], and its predecessor, the Brodsky-Lepage-Mackenzie (BLM) method [7,8,14], the PMC/BLM scales have been set by using a perturbative expansion in a and only approximate conformal series have been obtained. Here, we have been able to obtain the conformal series as revealed in dimensional regularization schemes. The final scales in Eq. (35) have naturally become functions of the coupling through the β function, in principle, to all orders.

B. Automation

In many cases the coefficients in a pQCD expression for an observable are computed numerically, and the $\{\beta_i\}$ dependence is not known explicitly. It is, however, easy to extract the dependence on the number of quark flavors N_f, since N_f enters analytically in any loop diagram computation. To use the systematic method presented in this paper, one puts the pQCD expression into the form of Eq. (20). Due to the special degeneracy in the coefficient of the $\{\beta_i\}$ terms, the N_f series can be matched to the $r_{j,k}$ coefficients in a unique way.[2] This allows one to automate the scale-setting process algorithmically.

The nth order coefficient in pQCD has an expansion in N_f which reads

$$r_n = c_{n,0} + c_{n,1}N_f + \cdots + c_{n,n-1}N_f^{n-1}. \tag{36}$$

By inspection of Eq. (20) it is seen that there are exactly as many unknown coefficients in the $\{\beta_i\}$ expansion at the order a^n as the N_f coefficients, $c_{n,j}$. This is realized due to the special degeneracy found in (20). The $r_{i,j}$ coefficients in Eq. (20) can thus be expressed in terms of the $c_{n,j}$ coefficients. The highest power in N_f at any order should always be associated with the same power in β_0. The first β_0 appears at order a^{n+1}. We derive the relations between $c_{n,j}$ and $r_{i,j}$ for a general gauge group, where we define C_A and C_F as the quadratic Casimir coefficients of the adjoint and quark representations and T as the generator trace normalization. For QCD these coefficients read $C_A = N_c$, $C_F = (N_c^2 - 1)/2N_c$, and $T = 1/2$. Using that $\beta_0 = 11/3C_A - 4/3TN_f$ we can find $r_{2,0}$ and $r_{2,1}$:

$$r_2 = r_{2,0} + n\beta_0 r_{2,1} = \left[r_{2,0} + r_{2,1}\frac{11nC_A}{3}\right] - r_{2,1}\frac{4n}{3}TN_f.$$

This leads to

$$r_{2,1} = -\frac{3}{4T}\frac{c_{2,1}}{n}, \qquad r_{2,0} = c_{2,0} + \frac{11C_A}{4T}c_{2,1}. \tag{37}$$

At the next order we have

$$r_3 = r_{3,0} + n\beta_1 r_{2,1} + (n+1)\beta_0 r_{3,1} + \frac{n(n+1)}{2}\beta_0^2 r_{3,2},$$

where $r_{2,1}$ is already known. Expanding as before in terms of N_f (with the higher order β_i coefficient given in [28]) we find the matching:

$$r_{3,2} = \frac{9}{8T^2}\frac{c_{3,2}}{n(n+1)}, \tag{38a}$$

$$r_{3,1} = \frac{1}{8(n+1)T^2}[6Tc_{2,1}(5C_A + 3C_F) - 33c_{3,2}C_A - 6Tc_{3,1}], \tag{38b}$$

$$r_{3,0} = c_{3,0} + \frac{1}{16T^2}[11C_A(11c_{3,2}C_A + 4Tc_{3,1}) - 12Tc_{2,1}C_A(7C_A + 11C_F)]. \tag{38c}$$

Similarly, we can find the $r_{4,j}$ coefficients:

$$r_{4,3} = \left(-\frac{3}{4T}\right)^3 \frac{3!}{n(n+1)(n+2)}c_{4,3}, \tag{39a}$$

$$r_{4,2} = \frac{1}{32(n+1)(n+2)T^3}\left[2nT^2c_{2,1}(79C_A + 66C_F) - 9\left(\frac{4(3+2n)}{n+1}Tc_{3,2}(5C_A + 3C_F) - 33c_{4,3}C_A - 4Tc_{4,2}\right)\right], \tag{39b}$$

[2] In principle, one must treat the N_f terms unrelated to renormalization of the gauge coupling as part of the conformal coefficient; e.g., the N_f terms coming from light-by-light scattering in QED and the N_f terms unrelated to the renormalization of the tri-gluon and quartic-four-gluon vertices belongs to the conformal series.

$$r_{4,1} = \frac{1}{64(n+2)T^3}\Big[4T^2c_{2,1}(-(37n+360)C_AC_F + 2(91n-150)C_A^2 - 18(n+6)C_F^2) + 48T^2c_{3,1}(5C_A+3C_F) + \frac{12Tc_{3,2}}{n+1}C_A((152n+173)C_A + 33(4n+5)C_F) - 33C_A(33c_{4,3}C_A + 8Tc_{4,2}) - 48T^2c_{4,1}\Big], \tag{39c}$$

$$r_{4,0} = c_{4,0} + \frac{1}{64T^3}[2T^2c_{2,1}C_A(8(228-77n)C_AC_F + (840-1127n)C_A^2 + 132(n+6)C_F^2) - 48T^2c_{3,1}C_A(7C_A+11C_F) - 2904Tc_{3,2}C_A^2C_F + 176T^2c_{4,1}C_A - 1848Tc_{3,2}C_A^3 + 484Tc_{4,2}C_A^2 + 1331c_{4,3}C_A^3]. \tag{39d}$$

Using these relations automatically gives the effective scales in Eq. (35c).

The automation process can be outlined as follows:

(1) Choose any δ-$\mathcal{R}$enormalization scheme and scale.
(2) Compute the physical observable in pQCD and extract the $N_{\rm f}$ coefficients, $c_{k,j}$.
(3) Find the β_i coefficients, $r_{k,j}$ from the $c_{k,j}$ coefficients and compute the PMC scales, Q_k.
(4) The final pQCD expression for the observable reads

$$\rho_{\rm final}(Q) = \sum_{k=0} r_{k+1,0}a(Q_{k+1})^{n+k}. \tag{40}$$

This procedure demonstrates that the N_f terms can be unambiguously associated to the $\{\beta_i\}$ terms to all orders. It also shows that PMC is the underlying principle of BLM scale-setting.

The PMC method can be used to set separate scales for different skeleton diagrams; this is particularly important for multiscale processes. In general, the $\{\beta_i\}$ coefficients multiply terms involving logarithms in each of the invariants [11]. For instance, in the case of $q\bar{q} \to Q\bar{Q}$ near the heavy quark threshold in pQCD, the PMC assigns different scales to the annihilation process and the rescattering corrections involving the heavy quarks' relative velocity [39]. It also can be used to set the scale for the "lensing" gluon-exchange corrections that appear in the Sivers, Collins, and Boer-Mulders effects. Moreover, for the cases when the process involves several energy regions, e.g. hard, soft, etc., one may adopt methods such as the nonrelativistic QCD effective theory [40] and the soft-collinear effective theory [41,42] to set the PMC scales; i.e., one first sets the PMC scales for the higher energy region, then integrate it out to form a lower energy effective theory and sets the PMC scales for this softer energy region, etc. In this way one obtains different effective PMC scales for each energy region, at which all the PMC properties also apply.

As an important remark, one should keep in mind that the determination of the factorization scale is a separate issue from renormalization scale-setting since it is present even in a conformal theory when $\beta = 0$. Nevertheless, in the literature the factorization scale in hadronic processes is often set to be equal to the renormalization scale. In principle, the factorization scale can be determined if one has knowledge of the nonperturbative light-front wave functions of the initial or final state hadrons. However, the PMC can also be used to set the scale of the coupling that appears in the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) or Efremov-Radyushkin-Brodsky-Lepage (ERBL) evolution equations, and it is consistent with the usual factorization properties for hard-process cross sections in QCD. It is therefore important to separate the renormalization and factorization scales in hadronic processes [2].

IV. EXAMPLES

We now consider three examples based on the Adler function [43], D, which can be measured indirectly through the dispersion relation:

$$D(Q^2) = Q^2\int_{4m_\pi^2}^{\infty}\frac{R_{e^+e^-}(s)}{(s+Q^2)^2}\mathrm{d}s, \tag{41}$$

where $R_{e^+e^-}$ is the ratio for electron-positron annihilation into hadrons.

The Adler function is particularly instructive to consider, since its conformal and nonconformal parts can be separated by using renormalization group (RG) arguments. Explicitly, the Adler function can be written in terms of the photon field anomalous dimension, γ, and the vacuum polarization function, Π, as follows [44,45]:

$$\bar{D}(Q^2) = \kappa^{-1}D(Q^2) = \gamma(a) - \beta(a)\frac{d}{da}\Pi(Q^2,a), \tag{42}$$

where $\beta(a)$ is the β function of the running coupling and we have defined the *normalized* Adler function $\bar{D}$ where $\kappa = d_F\sum_f Q_f^2$ and d_F is the dimension of the quark color representation, which in QCD reads $d_F = N_c$. We will work with this normalization throughout the related examples. In perturbation theory we define

$$\gamma(a) = \kappa\sum_{n=0}^{\infty}\gamma_n a(Q)^n, \tag{43}$$

$$\Pi(a) = \kappa\sum_{n=0}^{\infty}\Pi_n a(Q)^n, \tag{44}$$

which are now known to four-loop order [45–54]. The PMC procedure then follows by absorbing all β-dependent terms, which following Sec. III A becomes a trivial exercise once the degenerate coefficients $r_{i,j}$ have been identified.

As a fourth example, we consider a case where the explicit conformal and nonconformal parts are not known. Here we make explicit use of the automation procedure to derive the special degeneracy as described in Sec. III B.

A. $e^+e^- \to$ hadrons

The ratio for electron-positron annihilation into hadrons, $R_{e^+e^-}$ can inversely to Eq. (41) be computed from the Adler function, D, as follows:

$$\bar{R}_{e^+e^-}(s) = \frac{1}{2\pi i}\int_{-s-i\epsilon}^{-s+i\epsilon}\frac{\bar{D}(Q^2)}{Q^2}dQ^2. \tag{45}$$

It is easy to show that to order a^4 the perturbative expression for $\bar{R}_{e^+e^-}$ in terms of γ_n and Π_n reads

$$\begin{aligned}\bar{R}_{e^+e^-}(Q) = {}& \gamma_0 + \gamma_1 a(Q) + [\gamma_2 + \beta_0\Pi_1]a(Q)^2 \\ & + \left[\gamma_3 + \beta_1\Pi_1 + 2\beta_0\Pi_2 - \beta_0^2\frac{\pi^2\gamma_1}{3}\right]a(Q)^3 \\ & + \left[\gamma_4 + \beta_2\Pi_1 + 2\beta_1\Pi_2 + 3\beta_0\Pi_3 \right. \\ & \left. - \frac{5}{2}\beta_0\beta_1\frac{\pi^2\gamma_1}{3} - 3\beta_0^2\frac{\pi^2\gamma_2}{3} - \beta_0^3\pi^2\Pi_1\right]a(Q)^4. \end{aligned} \tag{46}$$

As expected, this expression has exactly the form of Eq. (20), with the following identification of the coefficients $r_{i,j}$:

$$r_{i,0} = \gamma_i, \tag{47a}$$

$$r_{i,1} = \Pi_{i-1}, \qquad i \geq 2, \tag{47b}$$

$$r_{i,2} = -\frac{\pi^2}{3}\gamma_{i-2}, \qquad i \geq, 3 \tag{47c}$$

$$r_{i,3} = -\pi^2\Pi_{i-3}, \qquad i \geq 4. \tag{47d}$$

The expressions for the coefficients γ_i and Π_i can be found in Refs. [45,54], and the four-loops β function is given in Ref. [28]. The γ_i contain N_f terms, but since they are independent of δ to any order, they are kept fixed in the scale-setting procedure. Notice that this is a feature in dimensional regularization.

Now it is easy to set the exact PMC scales from Eq. (35c) using that $R_{k,j} = (-1)^j r_{k+j,j}/\gamma_k$,

$$\ln\frac{Q_3^2}{Q^2} = -\frac{\Pi_3}{\gamma_3}, \tag{48a}$$

$$\ln\frac{Q_2^2}{Q^2} = -\frac{\Pi_2 + \frac{1}{2}[\frac{\partial\beta}{\partial a} + \frac{\beta}{a}]\frac{\pi^2}{3}\gamma_2}{\gamma_2 - \frac{1}{2}[\frac{\partial\beta}{\partial a} + \frac{\beta}{a}]\Pi_2}, \tag{48b}$$

$$\ln\frac{Q_1^2}{Q^2} = \frac{-(\Pi_1 + \frac{1}{2}\frac{\partial\beta}{\partial a}\frac{\pi^2}{3}\gamma_1 - \frac{1}{3!}[\beta\frac{\partial^2\beta}{\partial a^2} + (\frac{\partial\beta}{\partial a})^2]\pi^2\Pi_1)}{\gamma_1 - \frac{1}{2}\frac{\partial\beta}{\partial a}\Pi_1 + \frac{1}{3!}[\beta\frac{\partial^2\beta}{\partial a^2} - \frac{1}{2}(\frac{\partial\beta}{\partial a})^2]\frac{\Pi_1^2}{\gamma_1} - \frac{1}{4}(\frac{\partial\beta}{\partial a})^2\frac{\pi^2}{3}\gamma_1}. \tag{48c}$$

The final resummed expression for $\bar{R}_{e^+e^-}$ reads

$$\begin{aligned}\bar{R}_{e^+e^-}(Q) = {}& \gamma_0 + \gamma_1 a(Q_1) + \gamma_2 a(Q_2)^2 \\ & + \gamma_3 a(Q_3)^3 + \gamma_4 a(Q_4)^4. \end{aligned} \tag{49}$$

The scale Q_4 is unknown since it requires the knowledge of the order a^4 coefficient of Π; to leading order it reads

$$\ln\frac{Q_{4,\mathrm{LLO}}^2}{Q^2} = -\frac{\Pi_4}{\gamma_4}, \tag{50}$$

however, it is possible to estimate this value. This is so, since Π_4 can be written as

$$\begin{aligned}\Pi_4 = {}& -\frac{3}{4}\beta_0^3\Pi_{2,3} + \frac{9}{4}\beta_0^2\Pi_{3,2} + \frac{7}{8}\beta_1\beta_0\Pi_{2,2} \\ & - \frac{9}{4}\beta_0\Pi_{4,1} - \frac{1}{4}\beta_2\Pi_{2,1} - \frac{3}{4}\beta_1\Pi_{3,1} + \frac{3\Pi_{5,0}}{4}, \end{aligned} \tag{51}$$

where $\Pi_{i,j}$ are the coefficients of the *bare* vacuum polarization function Π_0:

$$\Pi_0(Q, a_0) = \sum_{l=1}^{\infty} a_0^{l-1}\left(\frac{\mu^2}{Q^2}\right)^{l\epsilon}\sum_{k=-l}^{\infty}\epsilon^k\Pi_{l,k}, \tag{52}$$

and where ϵ is the dimensional regularization parameter, $d = 4 - 2\epsilon$. In Eq. (51) only $\Pi_{5,0}$ is unknown . We can thus compute $Q_{4,\mathrm{LLO}}$ as a function of $\Pi_{5,0}$ and for five active flavors we find

$$Q_{4,\mathrm{LLO}} = 0.9e^{0.00013\times\Pi_{5,0}}Q. \tag{53}$$

Because of the small partner it is reasonable to set $Q_4 = Q$. The final result in numerical form in terms of $\alpha = \alpha_s/\pi$ for QCD with five active flavors reads

$$\begin{aligned}\bar{R}_{e^+e^-}(Q) = \frac{3}{11}R_{e^+e^-}(Q) = {}& 1 + \alpha(Q_1) + 1.84\alpha(Q_2)^2 \\ & - 1.00\alpha(Q_3)^3 - 11.31\alpha(Q_4)^4. \end{aligned} \tag{54}$$

This is a more convergent result compared to previous estimates, and it is free of any scheme and scale ambiguities (up to strongly suppressed residual ones).

From this expression we can determine the asymptotic scale Λ from the empirical data [55]:

$$\frac{3}{11}R^{\rm exp}_{e^+e^-}(\sqrt{s}=31.6\ {\rm GeV})=1.0527\pm0.0050.$$

To this end we use the logarithmic expansion for a in Eq. (14) and the known expressions for the γ_i and Π_i coefficients. For five active flavors we find

$$\Lambda_\delta=\Lambda_{\overline{\rm MS}}=419^{+222}_{-168}\ {\rm MeV}, \tag{55}$$

which gives us the numerical values for the PMC scales: $Q_1=1.3Q$, $Q_2=1.2Q$, $Q_3=5.3Q$. These final scales determine the effective number of quark flavors at each order of perturbation theory [34].

Finally, the QCD coupling at the M_Z scale, $\alpha_s(M_Z)$ can be computed using again the power expansion for a in terms of $1/\ln(\mu/\Lambda)$. We find

$$\alpha_s(M_Z)=0.132^{+0.010}_{-0.011}. \tag{56}$$

The error on this result is a reflection of the experimental uncertainty on $R^{\rm exp}_{e^+e^-}$, which cannot be eliminated. This value is somewhat larger than the present world average $\alpha_s(M_Z)=0.1184\pm0.0007$, which is a global fit of all types of experiments. However, it is consistent with the values obtained from e^+e^- colliders, i.e. $\alpha_s(M_Z)=0.13\pm0.005\pm0.03$ by the CLEO collaboration [56] and $\alpha_s(M_Z)=0.1224\pm0.0039$ from the jet shape analysis [57]. Moreover, in computing $\alpha_s(M_Z)$ we have assumed massless quarks. The estimate will decrease when taking threshold effects properly into account[3] as shown in [35].

B. $\tau\to\nu_\tau+$ hadrons

It is straightforward to apply our results to the τ decay into hadrons ratio $R_\tau=\sigma_{\tau\to\nu_\tau+{\rm hadrons}}/\sigma_{\tau\to\nu_\tau+\bar{\nu}_e+e^-}$, which can be computed from $R_{e^+e^-}$ [58]:

$$R_\tau(M_\tau)=2\int_0^{M_\tau^2}\frac{ds}{M_\tau^2}\left(1-\frac{s}{M_\tau^2}\right)^2\left(1+\frac{2s}{M_\tau^2}\right)\tilde{R}_{e^+e^-}(s), \tag{57}$$

where $\tilde{R}_{e^+e^-}$ is equal to $R_{e^+e^-}$ but with $\kappa=d_F\sum Q_f^2$ replaced by $\kappa'=d_F\sum|V_{ff'}|^2$, where $V_{ff'}$ are the Cabbibo-Kobayashi-Maskawa (CKM) matrix elements and $(\sum Q_f)^2=0$, since light-by-light scattering does not contribute. We define in the same way $\tilde{\gamma}$ and $\tilde{\Pi}$ (i.e. with no light-by-light contributions). In terms of Eq. (20), the coefficients for the normalized $\bar{R}_\tau=R_\tau/\kappa'$ read

$$r_{i,0}=\tilde{\gamma}_i, \tag{58a}$$

$$r_{i,1}=\tilde{\Pi}_{i-1}+\frac{19}{12}\tilde{\gamma}_{i-1},\qquad i\geq2, \tag{58b}$$

$$r_{i,2}=\left(\frac{265}{72}-\frac{\pi^2}{3}\right)\tilde{\gamma}_{i-2}+\frac{19\tilde{\Pi}_{i-2}}{6},\qquad i\geq3, \tag{58c}$$

$$r_{i,3}=\left(\frac{265}{24}-\pi^2\right)\tilde{\Pi}_{i-3}+\left(\frac{3355}{288}-\frac{19\pi^2}{12}\right)\tilde{\gamma}_{i-3},\qquad i\geq4. \tag{58d}$$

The final expression reads

$$\bar{R}_\tau(M_\tau)=\tilde{\gamma}_0+\tilde{\gamma}_1a(Q_1)+\tilde{\gamma}_2a(Q_2)^2+\tilde{\gamma}_3a(Q_3)^3+\tilde{\gamma}_4a(Q_4)^4. \tag{59}$$

Since there are three active quark flavors for $M_\tau\approx1.777$ GeV), we find from the CKM matrix that $\kappa=3(|V_{ud}|^2+|V_{us}|^2)\approx3$. The effective scales read $Q_1=0.67Q$, $Q_2=0.71Q$, $Q_3=582Q$ and for the same reason as in the case of $R_{e^+e^-}$ we set $Q_4=Q$. The scale Q_3 has been computed to NLLO since its LLO value is smaller than the asymptotic scale Λ. The final result in numerical form for three active quark flavors reads

$$\frac{1}{3}R_\tau(M_\tau)=1+\alpha(Q_1)+2.15\alpha(Q_2)^2+3.44\alpha(Q_3)^3+6.64\alpha(Q_4)^4, \tag{60}$$

with $\alpha=\alpha_s/\pi$. Using the asymptotic scale found from $R_{e^+e^-}$ we estimate the QCD contribution to the τ decay to be

$$R_\tau(M_\tau)=3.66^{+0.15}_{-0.22}. \tag{61}$$

This prediction is in good agreement with the experimental result from the OPAL collaboration [59]; $R^{\rm exp}_\tau(M_\tau)=3.593\pm0.008$.

C. Bjorken and GLS sum rules

The Bjorken sum rule [60] and the Gross-Llewellyn Smith (GLS) sum rule [61] obey well-known identities in conformal field theory, known as the Crewther relations [53,54,62–65], which through the Adler function can be used to expose the conformal terms. In this example, we show that both sum rules after PMC scale-setting have perturbative expansions that match exactly the inverse of the anomalous dimension, γ^{-1}, and is what one expects in a conformal field theory.

The Bjorken sum rule expresses the integral over the spin distributions of quarks inside of the nucleon in terms of its axial charge times a coefficient function C^{Bjp}:

$$\Gamma_1^{p-n}(Q^2)=\int_0^1[g_1^{ep}(x,Q^2)-g_1^{en}(x,Q^2)]dx=\frac{g_A}{6}C^{Bjp}(a)+\sum_{i=2}^{\infty}\frac{\mu_{2i}^{p-n}(Q^2)}{Q^{2i-2}}, \tag{62}$$

[3]We thank Ali N. Khorramian for comments on this point.

where g_1^{ep} and g_1^{en} are the spin-dependent proton and neutron structure functions, g_A is the nucleon axial charge as measured in neutron β decay. The sum in the second line of Eq. (62) describes for the nonperturbative power corrections (higher twist) which are inaccessible for pQCD. Focusing on the perturbative part, we define

$$C^{Bjp}(Q^2) = 1 - 3C_F a(Q^2) + \sum_{n=2}^{\infty} \tilde{C}_n a(Q^2)^n. \quad (63)$$

The Gross-Llewellyn Smith (GLS) sum rule,

$$\frac{1}{2}\int_0^1 F_3^{\nu p+\bar{\nu}p}(x,Q^2)dx = 3C^{\mathrm{GLS}}(a), \quad (64)$$

relates the lowest moment of the isospin singlet structure function $F_3^{\nu p+\bar{\nu}p}(x,Q^2)$ to a coefficient $C^{\mathrm{CLS}}(a_s)$, which appears in the operator product expansion of the axial and vector nonsinglet currents. We are again only considering the perturbative contribution and define

$$C^{\mathrm{GLS}}(Q^2) = 1 - 3C_F a(Q^2) + \sum_{n=2}^{\infty} C_n a(Q^2)^n. \quad (65)$$

The (extended) Crewther relation [62–65] states that there exists a relation between the two sum rules through the Adler function $D(Q^2)$ given in Eq. (A1) as follows:

$$\tilde{\bar{D}}(Q^2)C^{Bjp}(a) = 1 + \frac{\beta(a)}{a}\tilde{K}(a),$$
$$\tilde{K}(a) = a\tilde{K}_1 + a^2\tilde{K}_2 + a^3\tilde{K}_3 + \cdots, \quad (66)$$

and

$$\bar{D}(Q^2)C^{\mathrm{GLS}}(a) = 1 + \frac{\beta(a)}{a}K(a),$$
$$K(a) = aK_1 + a^2K_2 + a^3K_3 + \cdots. \quad (67)$$

The tilde on $\bar{D}$ and K indicates the corresponding expressions without the light-by-light type terms, i.e.

$$\bar{D} = \tilde{\bar{D}} + \bar{D}_{lbl}, \quad (68a)$$

$$K = \tilde{K} + K_{lbl}. \quad (68b)$$

The term proportional to the β function describes the deviation from the limit of exact conformal invariance, with the deviations starting at order a^2. Both sum rules have been explicitly computed to four loops [30,53,54,66–68] and shown to obey the extended Crewther relations [53,54].[4]

[4]There is a recent claim [69] that the existing four-loop coefficient of the Bjorken sum rule [53,54] is missing some singlet-diagram contributions. This is relevant only for the explicit evaluation of $\tilde{K}_3$, and does not change the results of this section.

We can use the Crewther relations to extract the conformal and nonconformal parts of C^{Bjp} and C^{GLS}. Denoting the power expansion of $\bar{D}$ by

$$\bar{D}(Q^2) = 1 + \sum_{n=1}^{\infty} d_n a(Q^2)^n, \quad (69)$$

and expanding its inverse perturbatively gives us

$$\begin{aligned} C^{\mathrm{GLS}}(a) = 1 &- d_1 a + a^2[d_1^2 - d_2 - \beta_0 K_1] \\ &+ a^3[2d_1d_2 - d_1^3 - d_3 + \beta_0(d_1K_1 - K_2) - \beta_1 K_1] \\ &+ a^4[d_1^4 + d_2^2 - d_4 - 3d_1^2d_2 + 2d_1d_3 + \beta_1(d_1K_1 - K_2) \\ &+ \beta_0(-d_1^2K_1 + d_1K_2 + d_2K_1 - K_3) - \beta_2 K_1]. \end{aligned} \quad (70)$$

The expression for C^{Bjp} is the same after putting tildes on the coefficients. The d_i are given in terms of γ_i, Π_i, and β_i as follows:

$$d_1 = \gamma_1 = 3C_F, \quad (71a)$$

$$d_{i\geq 2} = \gamma_i + \sum_{k=0}^{i-2}(i-1-k)\beta_k\Pi_{i-1-k}. \quad (71b)$$

We use this to find the degenerate $r_{i,j}$ coefficients of Eq. (20):

$$r_{2,1} = -K_1 - \Pi_1, \quad (72a)$$

$$r_{3,1} = -\frac{K_2}{2} - \Pi_2 + \left(\frac{K_1}{2} + \Pi_1\right)\gamma_1, \quad (72b)$$

$$\begin{aligned} r_{4,1} = &-\frac{K_3}{3} - \Pi_3 + (K_2 + 4\Pi_2)\frac{\gamma_1}{3} \\ &- \left(\frac{K_1}{3} + \Pi_1\right)\gamma_1^2 + (K_1 + 2\Pi_1)\frac{\gamma_2}{3}, \end{aligned} \quad (72c)$$

$$r_{4,2} = \frac{1}{3}(K_1\Pi_1 + \Pi_1^2), \quad (72d)$$

$$r_{3,2} = 0, \qquad r_{4,3} = 0. \quad (72e)$$

The degeneracy allows us to resum the series as described earlier. The final result is

$$\begin{aligned} C^{\mathrm{GLS}}(a) = 1 &- a(Q_1)\gamma_1 + a(Q_2)^2(\gamma_1^2 - \gamma_2) \\ &+ a(Q_3)^3(-\gamma_1^3 + 2\gamma_2\gamma_1 - \gamma_3) + a(Q_4)^4(\gamma_1^4 \\ &- 3\gamma_2\gamma_1^2 + 2\gamma_3\gamma_1 + \gamma_2^2 - \gamma_4) + \mathcal{O}(a^5), \end{aligned} \quad (73)$$

exposing the $r_{i,0}$ coefficients. This expression is simply the inverse of the anomalous dimension:

$$C^{\rm GLS}(a) = \gamma^{-1}(Q_1, Q_2, Q_3, \ldots), \tag{74}$$

where we used that $\gamma_0 = 1$. The arguments of γ^{-1} on the right-hand side indicate the effective scales at each order in perturbation theory, once the inverse is Taylor expanded. All the above expressions also apply to the Bjorken sum rules, with the coefficients replaced by the ones with tilde. In particular, $C^{Bjp}(a) = \tilde{\gamma}^{-1}(\tilde{Q}_1, \tilde{Q}_2, \tilde{Q}_3, \ldots)$.

Since, the Adler function itself after PMC scale-setting is simply given by the anomalous dimension:

$$D(Q) = \gamma(Q_1, Q_2, Q_3, \ldots), \tag{75}$$

and correspondingly for $\tilde{D}$, the Crewther relations can be expressed as

$$\tilde{\tilde{D}}(\tilde{Q})C^{Bjp}(\mu) = \frac{\tilde{\gamma}(\tilde{Q}_1, \tilde{Q}_2, \ldots)}{\tilde{\gamma}(\tilde{\mu}_1, \tilde{\mu}_2, \ldots)} = 1, \tag{76a}$$

$$\bar{D}(Q)C^{\rm GLS}(\mu) = \frac{\gamma(Q_1, Q_2, \ldots)}{\gamma(\mu_1, \mu_2, \ldots)} = 1, \tag{76b}$$

where the last equality follows due to conformality. These are the generalized Crewther relations, which set the commensurate scale relations between the scale of the Adler function and those of the sum rules.

D. Static quark potential

As a last example we consider the potential between two static quarks, where the degeneracy is not explicitly apparent in the literature. The static quark potential is known to order a^4 in the $\overline{\rm MS}$ scheme as an expansion in the number of massless flavors, N_f [70–75]:

$$\begin{aligned} V(Q^2) = -\frac{(4\pi)^2 C_F}{Q^2} a(Q^2)\Big[1 &+ (c_{2,0} + c_{2,1}N_f)a(Q^2) \\ &+ (c_{3,0} + c_{3,1}N_f + c_{3,2}N_f^2)a(Q^2)^2 \\ &+ (c_{4,0} + c_{4,1}N_f + c_{4,2}N_f^2 + c_{4,3}N_f^3)a(Q^2)^3 \\ &+ 8\pi^2 C_A^3 \ln\frac{\mu_{\rm IR}^2}{Q^2} a(Q^2)^3\Big] + \mathcal{O}(a^5), \end{aligned} \tag{77}$$

where we have chosen the initial scale of the running coupling $\mu^2 = Q^2$, but kept the explicit IR divergent logarithm $\ln \mu_{\rm IR}^2/Q^2$, which is not related to coupling constant renormalization, but is coming from the non-Abelian gluon "H-diagram". This IR divergence is a feature of pQCD and is canceled by nonperturbative contributions from the "ultrasoft" region [75,76], which is controlled by the domain of color confinement. The regularization comes from the energy difference between color-singlet and octet intermediate states [76].

The degenerate coefficients $r_{i,j}$ are determined from the $c_{i,j}$ coefficients as given in Sec. III B for $n = 1$. From the explicit expression for $c_{i,j}$ given in Refs. [73–75] we find the degenerate coefficients of the static potential to be

$$r^V_{2,0} = -\frac{8}{3}C_A, \tag{78a}$$

$$r^V_{2,1} = \frac{5}{3}, \qquad r^V_{3,2} = \left(\frac{5}{3}\right)^2, \qquad r^V_{4,3} = \left(\frac{5}{3}\right)^3, \ldots, \qquad r^V_{n+1,n} = \left(\frac{5}{3}\right)^n, \tag{78b}$$

$$r^V_{3,0} = \frac{C_A}{36}[(532 - 1584\zeta_3 - 9\pi^4 + 144\pi^2)C_A + 33(48\zeta_3 - 35)C_F], \tag{79a}$$

$$r^V_{3,1} = \left(7\zeta_3 - \frac{217}{36}\right)C_A + \left(\frac{35}{8} - 6\zeta_3\right)C_F, \tag{79b}$$

$$\begin{aligned} r^V_{4,0} = \frac{11}{36}C_A\Bigg((456\zeta_3 - 1440\zeta_5 + 571)C_F^2 &- 9\cdot 56.83(1)\frac{d_F^{abcd}d_F^{abcd}}{TN_A}\Bigg) + \Bigg(-\frac{758\zeta_3}{3} + 220\zeta_5 \\ &+ \frac{3709}{54}\Bigg)C_A^2 C_F + C_A^3\left(\frac{3077\zeta_3}{3} - 1293.54(1) + \frac{484\pi^4}{135}\right) \\ &- 136.39(12)\frac{d_F^{abcd}d_A^{abcd}}{N_A}, \end{aligned} \tag{79c}$$

$$\begin{aligned} r^V_{4,1} = &\left(\frac{392\zeta_3}{3} - 20\zeta_5 - \frac{66769}{648}\right)C_A C_F \\ &+ C_A^2\left(196.58 - 192\zeta_3 - \frac{88\pi^4}{135}\right) \\ &+ \left(40\zeta_5 - \frac{38\zeta_3}{3} - \frac{571}{36}\right)C_F^2 + 56.83(1)\frac{d_F^{abcd}d_F^{abcd}}{4TN_A}, \end{aligned} \tag{79d}$$

$$r^V_{4,2} = \left(23\zeta_3 + \frac{4\pi^4}{45} - \frac{2981}{144}\right)C_A + \left(\frac{5171}{216} - 26\zeta_3\right)C_F. \tag{79e}$$

The static potential can then be written without any explicit dependence on N_f:

$$\begin{aligned} V(Q^2) = -\frac{(4\pi)^2 C_F}{Q^2} a(Q^2)\Big[1 &+ (r^V_{2,0} + r^V_{2,1}\beta_0)a(Q^2) \\ &+ (r^V_{3,0} + \beta_1 r^V_{2,1} + 2\beta_0 r^V_{3,1} + \beta_0^2 r^V_{3,2})a(Q^2)^2 \\ &+ \left(r^V_{4,0} + \beta_2 r^V_{2,1} + 2\beta_1 r^V_{3,1} + \frac{5}{2}\beta_1\beta_0 r^V_{3,2} + 3\beta_0 r^V_{4,1}\right. \\ &\left. + 3\beta_0^2 r^V_{4,2} + \beta_0^3 r^V_{4,3}\right)a(Q^2)^3 + 8\pi^2 C_A^3 \ln\frac{\mu^2}{Q^2} a(Q^2)^3\Big] \\ &+ \mathcal{O}(a^5). \end{aligned} \tag{80}$$

Next, we reduce the expression to the conformal series by using the PMC scales, which are read off from Eq. (35):

$$V(Q^2) = -\frac{(4\pi)^2 C_F}{Q^2}\left[a(Q_1^2) + r_{2,0}^V a(Q_2^2)^2 + r_{3,0}^V a(Q_3^2)^3 + r_{4,0}^V a(Q_4^2)^4 + 8\pi^2 C_A^3 \ln\frac{\mu^2}{Q^2} a(Q^2)^4\right] + \mathcal{O}(a^5). \tag{81}$$

The expression in the bracket defines the effective charge $a_V(Q^2)$. Its IR divergence can be removed by adding the ultrasoft contributions to the static potential. The final expression for the effective charge $\alpha_V = 4\pi a_V$ in a $SU(3)$ gauge theory with N_f light quarks read

$$\alpha_V(Q^2) = \alpha(Q_1) - 0.64\alpha(Q_2)^2 - 0.78\alpha(Q_3)^3 + \left(3.49 + \frac{27}{8\pi}\ln\frac{\mu^2}{Q^2}\right)\alpha(Q_4^2)^4, \tag{82}$$

with

$$Q^2_{1,\text{all orders}} = Q^2 \exp(-5/3), \tag{83a}$$

$$Q^2_{2,\text{NLLO}} = Q^2 \exp(0.42 + 0.57\beta_0\alpha(Q^2)), \tag{83b}$$

$$Q^2_{3,\text{LLO}} = Q^2 \exp(-5.87), \tag{83c}$$

where $\beta_0 = 11 - \frac{2}{3}N_f$. Note that the PMC scale for Q_1 holds to all orders, which follows since $r_{n+1,n} = (5/3)^n$ at any order n. The scales for Q_2 and Q_3 have been expanded in α to the order consistent with the pQCD truncation. We note that the third PMC scale Q_3 is greatly suppressed compared to the kinematic scale Q. This is to be expected physically, since the kinematically accessible region shrinks with the loop order. This indicates the breakdown of perturbation theory in higher order QCD, which is also explicitly evident by the IR divergent term appearing in the a^4 coefficient. At these higher orders nonperturbative effects must be taken into account.

Finally, we can now explicitly show that PMC scale-setting is consistent with the Gell Mann-Low (GM-L) scheme and the effective coupling in QED. It is well known that the effective QED coupling in the massless limit, is to leading order related to the $\overline{\text{MS}}$ coupling by a scale displacement; $\alpha_{\text{GM-L}}(Q^2) = \alpha_{\overline{\text{MS}}}(Q^2 e^{-5/3})$. The effective QED coupling is precisely defined as the effective charge of the QED static potential between two (formally) infinitely charged particles. By inspection of Eqs. (78a), (79a), and (79c) it can be seen all the higher order conformal coefficients, $r_{i\geq 2,0}$ are proportional to non-Abelian group invariants, which vanish in the Abelian limit [36]; e.g. $C_A \to 0$. This generalizes to any order. This means that the PMC expression for the effective charge of the static potential is given to all orders in perturbation theory by

$$\alpha_{V,\text{QED}}(Q^2) = \alpha(Q^2 e^{-5/3}). \tag{84}$$

Thus PMC reproduces the correct result in the Abelian limit.

V. COMMENSURATE SCALE RELATIONS

We demonstrate that the generic expression in Eq. (20) extends to any scheme, that is, the special degeneracy in the $\mathcal{R}_\delta$ scheme of an observable is inherited in all physical schemes. This is done by relating different observables in pQCD using the effective charge method [4–7]. These commensurate scale relations must be independent of the choice of scheme. The scales are given by the systematic scale-setting method just described.

Any observable ρ can be used to define an effective charge a_ρ. Considering the case where the Born level result for the observable is just a constant such as $R_{e^+e^-}$; i.e. $n = 0$ in Eq. (15), the effective charge is defined by the relation

$$\rho(Q^2) = \rho_0(Q^2)[1 + a_\rho(Q^2)], \tag{85}$$

where ρ_0 is the Born (tree-level) result and Q^2 is the measured scale. Thus, a_ρ can be understood in perturbation theory as summing up the entire perturbative series into one effective coupling; the effective charge of the process.

It follows that the effective charge has an expansion in the $\mathcal{R}_\delta$ coupling $a_\mathcal{R}$ similar to the expansion in Eq. (20) (we put back the index $\mathcal{R}$ on the coupling in this section to avoid confusion). By normalizing running coupling such that at leading order the running coupling is equal to the effective charge, i.e.

$$a_\rho = \hat{a}_\mathcal{R} + \frac{r_2}{r_1^2}\hat{a}_\mathcal{R}^2 + \frac{r_3}{r_1^3}\hat{a}_\mathcal{R}^3 + \cdots, \tag{86}$$

where $\hat{a}_\mathcal{R} = r_1 a_\mathcal{R}$, the effective charge itself can be considered as a running coupling of a physical scheme related to the corresponding observable. The above expansion then defines the scheme transformation from the $\mathcal{R}_\delta$ scheme to the ρ scheme. Since any two effective charges a_A and a_B can be computed in the $\mathcal{R}_\delta$ scheme, it follows that a_A can be written as an expansion in a_B by scheme transformations. Thus, any effective charge defines a physical renormalization scheme. The β functions of such schemes are different from the β function of the $\mathcal{R}_\delta$ schemes, but are related by the identity

$$\beta_A(a_A) = \frac{\partial a_A}{\partial \hat{a}_\mathcal{R}}\beta_\mathcal{R}(\hat{a}_\mathcal{R}), \tag{87}$$

where A is some physical scheme corresponding to the effective charge a_A. From this identity it follows that the first two coefficients of the β function are universal [77].

The expansion of a_A in $\hat{a}_{\mathcal{R}}$ can be put to the form

$$a_A(Q^2) = \hat{a}_{\mathcal{R}}(Q^2) + [r^A_{2,0} + \hat{\beta}_0 r^A_{2,1}]\hat{a}_{\mathcal{R}}(Q^2)^2 + [r^A_{3,0} + \hat{\beta}_1 r^A_{2,1} + 2\hat{\beta}_0 r^A_{3,1} + \hat{\beta}_0^2 r^A_{3,2}]\hat{a}_{\mathcal{R}}(Q^2)^3 + \left[r^A_{4,0} + \hat{\beta}_2^{\mathcal{R}} r^A_{2,1} + 2\hat{\beta}_1 r^A_{3,1} + \frac{5}{2}\hat{\beta}_1\hat{\beta}_0 r^A_{3,2} + 3\hat{\beta}_0 r^A_{4,1} + 3\hat{\beta}_0^2 r^A_{4,2} + \hat{\beta}_0^3 r^A_{4,3}\right]\hat{a}_{\mathcal{R}}(Q^2)^4 + \mathcal{O}(\hat{a}^5_{\mathcal{R}}), \quad (88)$$

where $r^A_{i,j}$ are related to the coefficients $r_{i,j}$ of the observable A, by

$$r^A_{i,j} = \frac{r_{i,j}}{r_1^{i-j}}, \quad (89)$$

and $\hat{\beta}_i$ are here the coefficients of the beta function of $\hat{a}_{\mathcal{R}}$, which are related to the usual β_i in Eq. (5) by

$$\hat{\beta}_i = \frac{\beta_i}{r_1^{i+1}}. \quad (90)$$

The renormalization scheme dependence of $\hat{\beta}_2$ is denoted with a superscript. The coefficient β^A_2 can be found in terms of $\hat{\beta}_0$, $\hat{\beta}_1$ and $\hat{\beta}^{\mathcal{R}}_2$ from Eq. (88) using Eq. (87):

$$\beta^A_2 = \hat{\beta}^{\mathcal{R}}_2 - \hat{\beta}_1 r^A_{2,0} + \hat{\beta}_0^3(r^A_{3,2} - {r^A_{2,1}}^2) + 2\hat{\beta}_0^2(r^A_{3,1} - r^A_{2,0}r^A_{2,1}) - \hat{\beta}_0({r^A_{2,0}}^2 + r^A_{3,0}). \quad (91)$$

Using this, it can be shown that the special degeneracy in Eq. (88) is preserved when relating the effective charge a_A with another effective charge a_B, i.e.

$$a_A(Q^2) = a_B(Q^2) + [r^{AB}_{2,0} + \hat{\beta}_0 r^{AB}_{2,1}]a_B(Q^2)^2 + [r^{AB}_{3,0} + \hat{\beta}_1 r^{AB}_{2,1} + 2\hat{\beta}_0 r^{AB}_{3,1} + \hat{\beta}_0^2 r^{AB}_{3,2}]a_B(Q^2)^3 + [r^{AB}_{4,0} + \beta^B_2 r^{AB}_{2,1} + 2\hat{\beta}_1 r^{AB}_{3,1} + \frac{5}{2}\hat{\beta}_1\hat{\beta}_0 r^{AB}_{3,2} + 3\hat{\beta}_0 r^{AB}_{4,1} + 3\hat{\beta}_0^2 r^{AB}_{4,2} + \hat{\beta}_0^3 r^{AB}_{4,3}]a_B(Q^2)^4, \quad (92)$$

where the coefficients $r^{AB}_{i,j}$ are related to the $\mathcal{R}_\delta$ coefficients as follows:

$$r^{AB}_{2,0} = r^A_{2,0} - r^B_{2,0}, \quad (93a)$$

$$r^{AB}_{2,1} = r^A_{2,1} - r^B_{2,1}, \quad (93b)$$

$$r^{AB}_{3,0} = r^A_{3,0} - r^B_{3,0} - 2r^B_{2,0}r^{AB}_{2,0}, \quad (93c)$$

$$r^{AB}_{3,1} = r^A_{3,1} - r^B_{3,1} - r^B_{2,0}r^{AB}_{2,1} - r^B_{2,1}r^{AB}_{2,0}, \quad (93d)$$

$$r^{AB}_{3,2} = r^A_{3,2} - r^B_{3,2} - 2r^B_{2,1}r^{AB}_{2,1}, \quad (93e)$$

$$r^{AB}_{4,0} = r^A_{4,0} - r^B_{4,0} - 3r^B_{2,0}r^{AB}_{3,0} - ({r^B_{2,0}}^2 + 2r^B_{3,0})r^{AB}_{2,0}, \quad (93f)$$

$$r^{AB}_{4,1} = r^A_{4,1} - r^B_{4,1} - 2r^B_{2,0}r^{AB}_{3,1} - 2r^B_{2,1}r^{AB}_{3,0} + r^A_{3,0}r^B_{2,1} - r^A_{2,1}r^B_{3,0} - \frac{4}{3}(r^B_{3,1} + 2r^B_{2,0}r^B_{2,1})r^{AB}_{2,0}, \quad (93g)$$

$$r^{AB}_{4,2} = r^A_{4,2} - r^B_{4,2} - 2r^B_{2,1}r^{AB}_{3,1} - 2r^B_{3,1}r^{AB}_{2,1} - r^B_{2,0}r^{AB}_{3,2} - \frac{1}{3}({r^B_{2,1}}^2 + 2r^B_{3,2})r^{AB}_{2,0}, \quad (93h)$$

$$r^{AB}_{4,3} = r^A_{4,3} - r^B_{4,3} - 3r^B_{2,1}r^{AB}_{3,2} - 3r^B_{3,2}r^{AB}_{2,0}. \quad (93i)$$

This demonstrates that the special degeneracy of the $\{\beta_i\}$ coefficients is not a prerogative of the $\mathcal{R}_\delta$ schemes, but is a general feature of perturbation theory. Since any two effective charges are related by the same perturbative pattern of Eq. (20), we can directly use the systematic scale-setting method presented in the previous section to eliminate the initial scale ambiguity in Eq. (92). This explicitly shows the renormalization scheme invariance of the scale-setting method to all orders in perturbation theory. The final relation between any two effective charges is thus

$$a_A(Q^2) = a_B(Q_1^2) + r^{AB}_{2,0}a_B(Q_2^2)^2 + r^{AB}_{3,0}a_B(Q_3^2)^3 + r^{AB}_{4,0}a_B(Q_4^2)^4 + \mathcal{O}(a_B^5), \quad (94)$$

where the commensurate scale relations between the two charges are exactly the PMC scales Q_i as given by Eq. (35) with the β function being that of a_B. For completeness we provide also the expression for the four-loop β-function coefficient of an effective charge A:

$$\beta^A_3 = \hat{\beta}^{\mathcal{R}}_3 - 2\hat{\beta}^{\mathcal{R}}_2 r^A_{2,0} + 5\hat{\beta}_1\hat{\beta}_0^2(r^A_{3,2} - {r^A_{2,1}}^2) + 4\hat{\beta}_1\hat{\beta}_0(r^A_{3,1} - r^A_{2,0}r^A_{2,1}) + \hat{\beta}_1{r^A_{2,0}}^2 + 2\hat{\beta}_0^4(2{r^A_{2,1}}^3 - 3r^A_{3,2}r^A_{2,1} + r^A_{4,3}) + 6\hat{\beta}_0^3(2r^A_{2,0}{r^A_{2,1}}^2 - 2r^A_{3,1}r^A_{2,1} - r^A_{2,0}r^A_{3,2} + r^A_{4,2}) + 6\hat{\beta}_0^2(2r^A_{2,1}{r^A_{2,0}}^2 - 2r^A_{3,1}r^A_{2,0} - r^A_{2,1}r^A_{3,0} + r^A_{4,1}) + 2\hat{\beta}_0({r^A_{2,0}}^3 - 3r^A_{3,0}r^A_{2,0} + r^A_{4,0}). \quad (95)$$

As a particularly simple example, we relate the effective charge of R_τ; a_τ, to that of $R_{e^+e^-}$; a_R and apply the systematic scale-setting method to derive commensurate scale relations between the two effective charges. The final result is completely independent of the intermediate renormalization scheme and scale used to compute a_τ and a_R.

The degenerate coefficients of a_τ and a_R in the $\mathcal{R}_\delta$ scheme can be read off from Eqs. (58) and (47), from which we compute $r^{\tau,R}_{i,j}$. Using Eqs. (93) and (94) we can readily express a_τ as a perturbative series in a_R:

$$a_\tau(Q^2) = a_R(Q^2_{R,1}) - \frac{\gamma_{3,lbl}}{\gamma_1^3} a_R(Q^2_{R,3})^3 - \frac{\gamma_1\gamma_{4,lbl} - 3\gamma_2\gamma_{3,lbl}}{\gamma_1^5} a_R(Q^2_{R,4})^4, \quad (96)$$

where the PMC scales $Q_{R,i}$ are given by the systematic method and where $\gamma_{i,\mathrm{lbl}}$ is the light-by-light part of γ_i; i.e. the two effective charges are equivalent up to light-by-light terms. The corrections start at order a^3, since there is no light-by-light contribution to γ_2. The PMC scales expanded in $a_\tau(Q)$ read

$$\ln \frac{Q^2_{R,1}}{Q^2} = -\frac{19}{12} - \frac{169}{144}\hat{\beta}_0 a_R(Q^2) - \left(\frac{761\hat{\beta}_0^2}{192} + \frac{169\hat{\beta}_1}{96}\right) a_R(Q^2)^2, \quad (97a)$$

$$\ln \frac{Q^2_{R,3}}{Q^2} = \left(\frac{19}{12} + \frac{\Pi_1}{\gamma_1}\right)\frac{\gamma_{3,lbl}}{\gamma_3} - \frac{\Pi_{3,lbl}}{\gamma_3}. \quad (97b)$$

By definition the scale where the expression for the effective charge a_τ applies is $Q^2 = M_\tau^2$. At this scale, the number of light flavors is $N_f = 3$. Light-by-light diagrams are proportional to $(\sum_f Q_f)^2$, which vanishes exactly when summing over the three light quarks. Therefore in three-flavor QCD the two effective charges are identical to all orders; i.e. using $a_{R/\tau} = \gamma_1\alpha_{R/\tau}/(4\pi) = \alpha_{R/\tau}/\pi$, we have

$$\frac{\alpha_\tau(M_\tau^2)}{\pi} = \frac{\alpha_R(Q^2_{R,1})}{\pi}, \quad (98)$$

where the commensurate scale for a_R up to four-loop order is given by

$$\ln \frac{Q^2_{R,1}}{M_\tau^2} = -\frac{19}{12} - \frac{169}{64}\frac{\alpha_R(M_\tau^2)}{\pi} - \frac{83273}{3072}\frac{\alpha_R(M_\tau^2)^2}{\pi^2}. \quad (99)$$

This relation (at one lower order) has been shown to be in very good agreement with experiment [78] demonstrating a highly nontrivial consistency check of QCD, free of any scheme and scale ambiguities.

VI. CONCLUSION

In this paper we have shown that a generalization of the conventional $\overline{\mathrm{MS}}$ scheme is illuminating. It enables one to determine the general (and degenerate) pattern of nonconformal $\{\beta_i\}$ terms and to systematically determine the argument of the running coupling order by order in pQCD, in a way which is readily automatized. The resummed series matches the conformal series, in which no factorially divergent $n!\beta^n\alpha_s^n$ "renormalon" series appear and which is free of any scheme and scale ambiguities. Thus using the PMC/BLM procedure, all nonconformal contributions in the perturbative expansion series are summed into the running coupling by shifting the renormalization scale in α_s from its initial value, and one obtains unique, scale-fixed, scheme-independent predictions at any finite order. The resulting PMC scales and finite-order PMC predictions are both to high accuracy independent of the choice of initial renormalization scale. The PMC procedure also provides scale-fixed, scheme-independent commensurate scale relations, relations between observables which are based on the underlying conformal behavior of QCD such as the generalized Crewther relation. Furthermore, we have shown that PMC is consistent with QED scale-setting, where there is no ambiguity in choosing the final scale of the effective coupling. The PMC satisfies all of the principles of the renormalization group: reflectivity, symmetry, and transitivity, and it thus eliminates an unnecessary source of systematic error in pQCD predictions.

ACKNOWLEDGMENTS

We thank Konstantin Chetyrkin, Joseph Day, Leonardo Di Giustino, Stefan Höche, Andrei L. Kataev, and Ali N. Khorramian for useful comments. We are grateful to Sheng-Quan Wang for carefully proofreading the manuscript. M. M. thanks SLAC theory group for kind hospitality. This work was supported in part by the Department of Energy Contract No. DE-AC02-76SF00515, the Natural Science Foundation of China under Grant No. 11275280, and the Danish National Research Foundation, Grant No. DNRF90.

APPENDIX: ON THE DISCREPANCY BETWEEN THE NEW PMC RESULTS AND PREVIOUS BLM BASED RESULTS

In this Appendix we address the question raised by A. L. Kataev in [79][5] about the discrepancy of the new PMC result for the Adler function in this work and Refs. [1,2] and previous results based on BLM. Let us immediately stress that the discrepancy occurs due to the still unsettled question of which n_f terms in the perturbative coefficients should be treated as conformal terms and which should be related to renormalization (nonconformal terms). This statement applies to perturbative QCD in general and therefore also applies to the PMC results given in this work for $R_{e^+e^-}$, $R_{\tau\to\nu+\mathbf{h}}$, the Bjorken sum rule and the GLS sum rule.

Let us first remind that the purpose of BLM/PMC scale-setting is to set the scales of perturbative QCD in such a way that the scheme and scale ambiguities of the final expression are essentially eliminated. This goal can be reached by understanding the separation of the conformal and nonconformal contributions to the perturbative series.

[5]We note that in the second version of this paper, the criticism on our work was withdrawn. Nevertheless, this Appendix might still be helpful to avoid future confusion.

Note that the definition we are using for the conformal series is the one in which all $\{\beta_i\} = 0$. The definition used in [79] seems to be different. On the other hand, the same definition is used in some of the previous BLM literature, so why is there a discrepancy? The reason is that in the previous literature the explicit n_f series has been used to identify the nonconformal terms, while in our recent papers we have used the $\mathcal{R}_\delta$ scheme. We believe that the latter is more correct for the following reason: Using the explicit n_f series, one is forced to relate all n_f terms to renormalization of the coupling which is strictly incorrect, but a good approximation, at least to the orders known in pQCD. Instead, one should treat the n_f terms unrelated to renormalization of the coupling as part of the conformal coefficient; e.g., the n_f terms coming from light-by-light scattering in QED and the n_f terms unrelated to the renormalization of the trigluon and quartic-four-gluon vertices belongs to the conformal series (see e.g. [80]).

Using the $\mathcal{R}_\delta$ scheme, one can instead derive the $\{\beta_i\}$-series expansion of the perturbative coefficients, thereby avoiding the use of the n_f-series expansion. So far no known pQCD results have been computed in the $\mathcal{R}_\delta$ scheme. However, for the particular cases of the $R_{e^+e^-}$, $R_{\tau\to\nu+\mathbf{h}}$, the Bjorken sum rule and the GLS sum rule we have been able to derive the $\{\beta_i\}$-series expansion. This is possible because they are all related to the Adler function, which can be explicitly written as a sum of conformal and nonconformal contributions:

$$D(Q^2) = \gamma(a) + \beta(a)\frac{d}{da}\Pi(Q^2, a). \tag{A1}$$

Thus, we have suggested that the PMC result for the Adler function reads

$$D(Q^2) = \gamma_0 + \gamma_1\alpha(Q_1) + \gamma_2\alpha(Q_2)^2 + \gamma_3\alpha(Q_3)^3 + \cdots, \tag{A2}$$

which in numerical form, known up to four loops [45], reads (we neglect for simplicity the light-by-light or singlet type terms)

$$\begin{aligned} D(Q^2) = \sum_{f=1}^{n_f} Q_f^2[1 &+ \alpha(Q_1) + (2.60 - 0.15n_f)\alpha(Q_2)^2 \\ &+ (9.74 - 2.04n_f - 0.02n_f^2)\alpha(Q_3)^3 \\ &+ (41.09 - 13.00n_f - 0.49n_f^2 + 0.005n_f^3)\alpha(Q_4)^4] \\ &+ \mathcal{O}(\alpha^5), \end{aligned} \tag{A3}$$

where $\alpha = \alpha_s/\pi$ and the PMC scales Q_i are functions of Q and are derived as given in Sec. III A. Their values are unimportant for the purpose of this Appendix. Notice that n_f terms enter already from NLO and beyond. We propose that these n_f terms should not be absorbed into the running coupling since they are independent of which renormalization scheme is used in dimensional regularization. The appearance of these n_f terms is the reason behind the discrepancy.

On the other hand, as also described in Sec. III B, when one is confronted with not knowing the $\{\beta_i\}$-series expansion, the explicit n_f series can be used as an approximation to get the PMC result.

Suppose we did not have Eq. (A1) in hand and instead only knew the Adler function in perturbation theory with its n_f series for each coefficient. This reads

$$\begin{aligned} D(Q^2) = \sum_{f=1}^{n_f} Q_f^2[1 &+ \alpha(Q) + (1.99 - 0.12n_f)\alpha(Q)^2 \\ &+ (18.24 - 4.22n_f + 0.086n_f^2)\alpha(Q)^3 \\ &+ (135.87 - 34.52n_f + 1.88n_f^2 + 0.01n_f^3)\alpha(Q)^4], \\ &+ O(\alpha^5), \end{aligned} \tag{A4}$$

where we have set the initial renormalization scale $\mu = Q$. From Sec. III B we then find the PMC result to be

$$\begin{aligned} D(Q^2) = \sum_{f=1}^{n_f} Q_f^2[1 &+ \alpha(\hat{Q}_1) + 0.083\alpha(\hat{Q}_2)^2 \\ &- 23.22\alpha(\hat{Q}_3)^3 + 81.24\alpha(\hat{Q}_4)^4] + O(\alpha^5). \end{aligned} \tag{A5}$$

Notice that up to NLO this, as expected, agrees with all the previous BLM literature including the recent paper of A. L. Kataev [79]. For the higher order terms a careful treatment of the light-by-light terms must be made. We stress again that in this result *the effective PMC scales $\hat{Q}_i$ include absorption of n_f terms that are unrelated to renormalizing the running coupling*, and the expression is therefore only an approximation.

For $R_{e^+e^-}$ both methods were used in Ref. [2], where it was found that the final numerical results of the two methods are consistent with each other. Furthermore, a recent analysis on the Higgs boson inclusive decay channels $H \to b\bar{b}$ and $H \to gg$ up to four loops [81] shows that also here the two treatments are consistent with each other and even with the BLM scale-setting approach proposed in Ref. [23], which also resorts to the explicit use of the n_f series. Thus, the explicit n_f-series treatment seems to be a good approximation for phenomenological applications.

This explicit treatment should explain the discrepancies between the conformal coefficients in this work and the previous ones [1,2] (based on the $\mathcal{R}_\delta$ scheme) and those found previously in the BLM literature (based on the n_f-series expansion). Here, we have explained why we believe the former treatment is formally more correct, while in some applications the results are effectively the same.

[1] M. Mojaza, S. J. Brodsky, and X.-G. Wu, Phys. Rev. Lett. **110**, 192001 (2013).
[2] X.-G. Wu, S. J. Brodsky, and M. Mojaza, Prog. Part. Nucl. Phys. **72**, 44 (2013).
[3] S. J. Brodsky and X.-G. Wu, Phys. Rev. D **86**, 054018 (2012).
[4] G. Grunberg, Phys. Lett. **95B**, 70 (1980); **110B**, 501(E) (1982).
[5] G. Grunberg, Phys. Rev. D **29**, 2315 (1984).
[6] A. Dhar and V. Gupta, Phys. Rev. D **29**, 2822 (1984).
[7] S. J. Brodsky and H. J. Lu, Phys. Rev. D **51**, 3652 (1995).
[8] S. J. Brodsky, G. P. Lepage, and P. B. Mackenzie, Phys. Rev. D **28**, 228 (1983).
[9] S. J. Brodsky and X.-G. Wu, Phys. Rev. D **85**, 034038 (2012); **86**, 079903(E) (2012).
[10] S. J. Brodsky and X.-G. Wu, Phys. Rev. Lett. **109**, 042002 (2012).
[11] S. J. Brodsky and L. Di Giustino, Phys. Rev. D **86**, 085026 (2012).
[12] H. J. Lu and C. A. R. Sa de Melo, Phys. Lett. B **273**, 260 (1991); **285**, 399(E) (1992).
[13] H. J. Lu, Phys. Rev. D **45**, 1217 (1992).
[14] G. Grunberg and A. L. Kataev, Phys. Lett. B **279**, 352 (1992).
[15] G. Grunberg, Phys. Rev. D **46**, 2228 (1992).
[16] M. Beneke and V. M. Braun, Phys. Lett. B **348**, 513 (1995).
[17] M. Neubert, Phys. Rev. D **51**, 5924 (1995).
[18] C. N. Lovett-Turner and C. J. Maxwell, Nucl. Phys. **B452**, 188 (1995).
[19] P. Ball, M. Beneke, and V. M. Braun, Nucl. Phys. **B452**, 563 (1995).
[20] S. J. Brodsky, J. R. Ellis, E. Gardi, M. Karliner, and M. A. Samuel, Phys. Rev. D **56**, 6980 (1997).
[21] S. J. Brodsky, E. Gardi, G. Grunberg, and J. Rathsman, Phys. Rev. D **63**, 094017 (2001).
[22] K. Hornbostel, G. P. Lepage, and C. Morningstar, Phys. Rev. D **67**, 034023 (2003).
[23] S. V. Mikhailov, J. High Energy Phys. 06 (**2007**) 009.
[24] A. L. Kataev and S. V. Mikhailov, Teor. Mat. Fiz. **170**, 174 (2012)
[25] C. G. Bollini and J. J. Giambiagi, Nuovo Cimento Soc. Ital. Fis. B **12**, 20 (1972).
[26] G.'t Hooft and M. J. G. Veltman, Nucl. Phys. **B44**, 189 (1972).
[27] C. G. Bollini and J. J. Giambiagi, Phys. Lett. **40B**, 566 (1972).
[28] T. van Ritbergen, J. A. M. Vermaseren, and S. A. Larin, Phys. Lett. B **400**, 379 (1997). For a generalization of β_3, see the Appendix of M. Mojaza, C. Pica, and F. Sannino, Phys. Rev. D **82**, 116009 (2010).
[29] G.'t Hooft, Nucl. Phys. **B61**, 455 (1973).
[30] W. A. Bardeen, A. J. Buras, D. W. Duke, and T. Muta, Phys. Rev. D **18**, 3998 (1978).
[31] K. G. Chetyrkin, A. L. Kataev, and F. V. Tkachov, Nucl. Phys. **B174**, 345 (1980).
[32] P. M. Stevenson, Phys. Rev. D **23**, 2916 (1981).
[33] H.-J. Lu and S. J. Brodsky, Phys. Rev. D **48**, 3310 (1993).
[34] S. J. Brodsky, M. S. Gill, M. Melles, and J. Rathsman, Phys. Rev. D **58**, 116006 (1998).
[35] K. G. Chetyrkin, B. A. Kniehl, and M. Steinhauser, Phys. Rev. Lett. **79**, 2184 (1997).
[36] S. J. Brodsky and P. Huet, Phys. Lett. B **417**, 145 (1998).
[37] S. J. Brodsky and X.-G. Wu, Phys. Rev. D **86**, 014021 (2012).
[38] S. J. Brodsky and X.-G. Wu, Phys. Rev. D **85**, 114040 (2012).
[39] S. J. Brodsky, A. H. Hoang, J. H. Kuhn, and T. Teubner, Phys. Lett. B **359**, 355 (1995).
[40] G. T. Bodwin, E. Braaten, and G. P. Lepage, Phys. Rev. D **51**, 1125 (1995); **55**, 5853(E) (1997).
[41] C. W. Bauer, S. Fleming, D. Pirjol, and I. W. Stewart, Phys. Rev. D **63**, 114020 (2001).
[42] C. W. Bauer, D. Pirjol, and I. W. Stewart, Phys. Rev. D **65**, 054022 (2002).
[43] S. L. Adler, Phys. Rev. D **10**, 3714 (1974).
[44] K. G. Chetyrkin, J. H. Kuhn, and A. Kwiatkowski, Phys. Rep. **277**, 189 (1996).
[45] P. A. Baikov, K. G. Chetyrkin, J. H. Kuhn, and J. Rittinger, J. High Energy Phys. 07 (**2012**) 017.
[46] K. G. Chetyrkin, A. L. Kataev, and F. V. Tkachov, Phys. Lett. **85B**, 277 (1979).
[47] M. Dine and J. R. Sapirstein, Phys. Rev. Lett. **43**, 668 (1979).
[48] W. Celmaster and R. J. Gonsalves, Phys. Rev. Lett. **44**, 560 (1980).
[49] W. Celmaster and R. J. Gonsalves, Phys. Rev. D **21**, 3112 (1980).
[50] S. G. Gorishnii, A. L. Kataev, and S. A. Larin, Phys. Lett. B **259**, 144 (1991).
[51] L. R. Surguladze and M. A. Samuel, Phys. Rev. Lett. **66**, 560 (1991); **66**, 2416(E) (1991).
[52] P. A. Baikov, K. G. Chetyrkin, and J. H. Kuhn, Phys. Rev. Lett. **101**, 012002 (2008).
[53] P. A. Baikov, K. G. Chetyrkin, and J. H. Kuhn, Phys. Rev. Lett. **104**, 132004 (2010).
[54] P. A. Baikov, K. G. Chetyrkin, J. H. Kuhn, and J. Rittinger, Phys. Lett. B **714**, 62 (2012).
[55] R. Marshall, Z. Phys. C **43**, 595 (1989).
[56] R. Ammar *et al.* (CLEO Collaboration), Phys. Rev. D **57**, 1350 (1998).
[57] G. Dissertori, A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, G. Heinrich, and H. Stenzel, J. High Energy Phys. 02 (**2008**) 040.
[58] C. S. Lam and T.-M. Yan, Phys. Rev. D **16**, 703 (1977).
[59] K. Ackerstaff *et al.* (OPAL Collaboration), Eur. Phys. J. C **7**, 571 (1999).
[60] J. D. Bjorken, Phys. Rev. D **1**, 1376 (1970).
[61] D. J. Gross and C. H. Llewellyn Smith, Nucl. Phys. **B14**, 337 (1969).
[62] D. J. Broadhurst and A. L. Kataev, Phys. Lett. B **315**, 179 (1993).
[63] G. T. Gabadadze and A. L. Kataev, JETP Lett. **61**, 448 (1995)
[64] R. J. Crewther, Phys. Lett. B **397**, 137 (1997).
[65] V. M. Braun, G. P. Korchemsky, and D. Mueller, Prog. Part. Nucl. Phys. **51**, 311 (2003).
[66] G. Altarelli, R. K. Ellis, and G. Martinelli, Nucl. Phys. **B143**, 521 (1978); **B146**, 544(E) (1978).
[67] S. G. Gorishnii and S. A. Larin, Phys. Lett. B **172**, 109 (1986).

[68] S. A. Larin and J. A. M. Vermaseren, Phys. Lett. B **259**, 345 (1991).
[69] S. A. Larin, Phys. Lett. B **723**, 348 (2013).
[70] T. Appelquist, M. Dine, and I. J. Muzinich, Phys. Lett. **69B**, 231 (1977).
[71] M. Peter, Phys. Rev. Lett. **78**, 602 (1997).
[72] Y. Schroder, Phys. Lett. B **447**, 321 (1999).
[73] A. V. Smirnov, V. A. Smirnov, and M. Steinhauser, Phys. Lett. B **668**, 293 (2008).
[74] A. V. Smirnov, V. A. Smirnov, and M. Steinhauser, Phys. Rev. Lett. **104**, 112002 (2010).
[75] C. Anzai, Y. Kiyo, and Y. Sumino, Phys. Rev. Lett. **104**, 112003 (2010).
[76] N. Brambilla, A. Pineda, J. Soto, and A. Vairo, Rev. Mod. Phys. **77**, 1423 (2005).
[77] G. t Hooft, *Recent Developments in Gauge Theories* (Plenum Press, New York, 1980), p. 135; , reprint in *Unity of Forces in the Universe Vol. II*, edited by (World Scientific, Singapore, 1982), p. 1004.
[78] S. J. Brodsky, S. Menke, C. Merino, and J. Rathsman, Phys. Rev. D **67**, 055008 (2003).
[79] A. L. Kataev, arXiv:1305.4605v1.
[80] M. Binger and S. J. Brodsky, Phys. Rev. D **74**, 054016 (2006).
[81] S.-Q. Wang, X.-G. Wu, X.-C. Zheng, J.-M. Shen, and Q.-L. Zhang, arXiv:1308.6364.