PHYSICAL REVIEW D **86,** 054018 (2012)

Self-consistency requirements of the renormalization group for setting the renormalization scale

Stanley J. Brodsky*

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, and SLAC National Accelerator Laboratory, Stanford University, Stanford, California 94039, USA

(Received 10 August 2012; published 24 September 2012)

In conventional treatments, predictions from fixed-order perturbative QCD calculations cannot be fixed with certainty because of ambiguities in the choice of the renormalization scale as well as the renormalization scheme. In this paper we present a general discussion of the constraints of the renormalization group (RG) invariance on the choice of the renormalization scale. We adopt the RG-based equations, which incorporate the scheme parameters, for a general exposition of RG invariance, since they simultaneously express the invariance of physical observables under both the variation of the renormalization scale and the renormalization scheme parameters. We then discuss the self-consistency requirements of the RG, such as reflexivity, symmetry, and transitivity, which must be satisfied by the scale-setting method. The principle of minimal sensitivity requires the slope of the approximant of an observable to vanish at the renormalization point. This criterion provides a scheme-independent estimation, but it violates the symmetry and transitivity properties of the RG and does not reproduce the Gell-Mann-Low scale for QED observables. The principle of maximum conformality (PMC) satisfies all of the deductions of the RG invariance—reflectivity, symmetry, and transitivity. Using the PMC, all nonconformal $\{\beta_i^{\mathcal{R}}\}$ terms ($\mathcal{R}$ stands for an arbitrary renormalization scheme) in the perturbative expansion series are summed into the running coupling, and one obtains a unique, scale-fixed, scheme-independent prediction at any finite order. The PMC scales and the resulting finite-order PMC predictions are both to high accuracy independent of the choice of initial renormalization scale, consistent with RG invariance. Moreover, after PMC scale setting, the residual initial scale dependence at fixed order owing to unknown higher-order $\{\beta_i\}$ terms can be substantially suppressed. The PMC thus eliminates a serious systematic scale error in perturbative quantum chromodynamics predictions, greatly improving the precision of tests of the Standard Model and the sensitivity to new physics at collider and other experiments.

DOI: 10.1103/PhysRevD.86.054018 PACS numbers: 12.38.Aw, 11.10.Gh, 11.10.Hi, 12.38.Bx

I. INTRODUCTION

Given the perturbative series for a physical quantity

$$\rho_n = \mathcal{C}_0 \alpha_s^p(\mu) + \sum_{i=1}^{n} \mathcal{C}_i(\mu) \alpha_s^{p+i}(\mu), \tag{1}$$

expanded to nth order in the QCD strong coupling constant $\alpha_s(\mu)$, the renormalization scale μ must be specified in order to obtain a definite prediction. The common practice adopted in the literature is to simply guess a renormalization scale $\mu = Q$, keep it fixed during the calculation (Q is usually assumed to be a typical momentum transfer of the process), and then vary it over an arbitrary range, e.g., $[Q/2, 2Q]$, in order to ascertain the scale uncertainty. However, there are many weak points of this conventional scale-setting method:

(1) Although the infinite perturbative series $\rho_{n\to\infty}$ summed to all orders is renormalization-scale independent, the scale dependence from $\alpha_s(\mu)$ and $\mathcal{C}_i(\mu)$ does not exactly cancel at finite order, leading to a renormalization-scale ambiguity.

(2) The fixed-order estimate in the conventional procedure is also scheme dependent; i.e., different choices of renormalization scheme $\mathcal{R}$ will lead to different theoretical estimates. This is the well-known renormalization scheme ambiguity [1–11].

(3) The conventional scale choice can give unphysical results: For example, for the case of W-boson plus three-jet production at the hadronic colliders, taking μ to be the W-boson transverse energy, the conventional scale-setting method even predicts negative QCD cross section at the next-to-leading order (NLO) [12,13].

(4) As has been shown in Ref. [14], taking an incorrect renormalization scale underestimates the top quark forward-backward asymmetry at the Tevatron.

(5) It should be recalled that there is no ambiguity in setting the renormalization scale in QED.

*sjbth@slac.stanford.edu
†wuxg@cqu.edu.cn

(i) In QED, the coupling $\alpha(q^2)$ is conventionally defined in the Gell-Mann-Low (GM-L) scheme [15] from the potential between heavy charges, and it is normalized at $q^2 = 0$ to the fine-structure constant $\alpha(0) \simeq 1/137.0359\ldots$ [16].

(ii) Because of the Ward-Takahashi identity [17], the divergences in the vertex and fermion wave-function corrections cancel, and the ultraviolet divergence associated with the vacuum polarization defines a natural scale for the coupling constant $\alpha(q^2)$ with q^2 being the squared momentum transfer for the photon propagator. This fact was first observed by Gell-Mann and Low [15]; i.e., in the standard GM-L scheme, the renormalization scale is simply the virtuality of the exchanged photon.
For example, the renormalization scale for the electron-muon elastic scattering based on one-photon exchange is the virtuality of the exchanged photon, i.e., $\mu^2_{\text{GM-L}} = t = q^2$. One can, of course, choose any initial renormalization scale t_0 for calculating the QED amplitude; however, the final result will not depend on the choice of t_0, since

$$\alpha(t) = \frac{\alpha(t_0)}{[1 - \Pi(t, t_0)]}, \tag{2}$$

where

$$\Pi(t, t_0) = \frac{[\Pi(t) - \Pi(t_0)]}{[1 - \Pi(t_0)]},$$

naturally sums all vacuum polarization contributions, both proper and improper, to the dressed photon propagator. [Here $\Pi(t) = \Pi(t, 0)$ is the sum of proper vacuum polarization insertions, subtracted at $t = 0$.] The invariance of the result on the initial scale t_0 is the property used to derive the Callan-Symanzik equations [18,19]. There is, therefore, no reason to vary $\mu_{\text{GM-L}}$ by a factor of $1/2$ or 2, since the photon virtuality t is the unique, optimized scale in the GM-L scheme.

(iii) The renormalization scale in QED is unique in any scheme including dimensional regularization; different schemes can be connected to the GM-L scheme by commensurate scale relations (CSRs) [20], a topic that we discuss below. The computation of higher-order $\{\beta_i^{\mathcal{R}}\}$ functions is thus important for perturbative calculations at the highest orders [21–23].

(6) There are uncanceled large logarithms, as well as *renormalon* terms in higher orders that diverge as $[n!(\beta_i^{\mathcal{R}})^n \alpha_s^n]$ [24]. The convergence of the perturbative series is thus problematic using conventional scale setting. For certain processes such as the top-quark pair production, it is found that the total cross section for the $(q\bar{q})$ channel, $q\bar{q} \to t + \bar{t}$, at the next-to-next-to-leading order (NNLO) is about 50% of the NLO cross section using conventional scale setting [25–27]. Thus, to derive a dependable perturbative estimate, one evidently needs to do even higher order calculations.

(7) The conventional estimate shows a strong dependence on the choice of the renormalization scale μ. It is clearly artificial to guess a renormalization scale $\mu = Q$ and to study its uncertainty by simply varying μ in the arbitrary range $[Q/2, 2Q]$. Why is the scale uncertainty estimated only by varying a factor of $1/2$ or 2, and not, say, $10\,Q$? For example, Ref. [28] shows that after including the first and second order corrections to several deep inelastic sum rules that are attributable to heavy flavor contributions, it is found that the effective scale $\mu \sim 6.5Q$, where the typical scale $Q = m$ with m being the corresponding heavy quark mass.
Moreover, sometimes, there are several choices for the typical momentum transfer of the process, all of which can be taken as the renormalization scale, such as the heavy quark mass and the collision energy of the subprocess. Which invariant provides the correct theoretical estimate?
Using conventional scale setting, there is no definite answer to these questions. One may argue that the correct renormalization scale for the fixed-order prediction can be decided by comparing with the experimental data, but this surely is process dependent and greatly depresses the predictive power of the perturbative quantum chromodynamics (pQCD) theory.

Thus, in summary, the conventional scale setting assigns an arbitrary range and an arbitrary systematic error to fixed-order pQCD predictions. In fact, as we discuss in this article, this *ad hoc* assignment of the range and associated systematic error is unnecessary and can be eliminated.

One may ask: For a general fixed-order calculation, what is the correct *physical* scale or optimized scale? To our understanding, it should provide a prediction independent

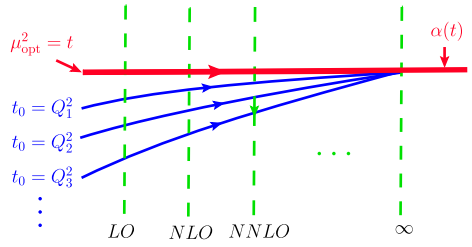

FIG. 1 (color online). Pictorial representation of the optimized renormalization scale μ_{opt}. Taking electron-muon elastic scattering through one-photon exchange as an example: In the GM-L scheme, the optimized scale is $\mu^2_{\text{opt}} = t$, which corresponds to the scale-invariant value $\alpha(t)$. As a comparison, the values of α at fixed orders for different choices of $t_0 = Q_i^2$ $(i = 1, 2, 3, \ldots)$ are shown by thin-and-solid curves.

of the renormalization scheme and the choice of initial scale. A pictorial representation of what is the optimized renormalization scale is shown in Fig. 1, where the electron-muon elastic scattering through one-photon exchange is taken as an illustration. In the GM-L scheme, the optimized scale $\mu^2_{\rm opt} = t$, which corresponds to the scale-invariant value $\alpha(t)$. Moreover, by using the proper scale-setting method, such as the newly suggested principle of maximum conformality (PMC) [25,29–31], the prediction is also scheme independent and the arguments of the coupling in different schemes have the correct displacement. For example, by using the PMC procedure for QED one obtains the correct displacement between the argument of the coupling in the $\overline{\rm MS}$ scheme relative to the GM-L scheme at one loop [32][1]

$$\alpha_{\rm GM\text{-}L}(t) = \alpha_{\overline{\rm MS}}(e^{-5/3}t). \tag{3}$$

As a comparison, the values of α at fixed orders for different choices of $t_0 = Q_i^2$ ($i = 1, 2, 3, \ldots$) are shown by thin-and-solid curves in Fig. 1. A particular choice of t_0 using conventional scale setting may lead to a value of α close to $\alpha(t)$, but this would only be a lucky guess and not the correct answer. As one includes higher-and-higher orders, the guessed scale will lead to a better estimate. In fact, when doing the perturbative calculation up to infinite order, any choice of t_0 will lead to the correct value $\alpha(t)$ as required by the renormalization group (RG) invariance. However, if one chooses $t_0 = t$, the complete all-orders result is obtained from the onset.

Does there exist such an optimized renormalization scale for a general process in non-Abelian QCD? If it does exist, how can one set it at finite order in a systematic and process-independent way? This is not an easy task. Various scale-setting procedures have been proposed since the 1980s for deriving an optimized scale, such as fastest apparent convergence (FAC) [4–6],[2] the principle of minimum sensitivity (PMS) [7–10], the Brodsky-Lepage-Mackenzie (BLM) [11] procedure, and its extended versions such as the dressed skeleton expansion [34,35], the sequential se-BLM and x-BLM methods [36–38], and the PMC. A short review of FAC, BLM, and PMS can be found in Ref. [39]. In principle, the correctness of a scale-setting method can be judged by the experimental data. However, as we shall discuss, there are self-consistency theoretical requirements that shed light on the reliability of the scale-setting method [40].

Clearly, the prediction for any physical observable must be independent of the choice of renormalization scheme; this is the central property of the RG invariance [18,19,41–43]. As we shall discuss, the RG-based equations [7–10] that incorporate the scheme parameters provide a convenient way for estimating both the scale and scheme dependence of the QCD predictions for a physical process [7–10,30,44]. In this paper, we will utilize such RG-based equations for a general discussion of the RG invariance. We will discuss in detail the self-consistency requirements of the RG [40], such as reflexivity, symmetry, and transitivity, which must be satisfied by a scale-setting method. We will then show whether the scale-setting methods, FAC, BLM/PMC, and PMS, satisfy these requirements.

The remaining parts of this paper are organized as follows: in Sec. II, we give a general demonstration of the RG invariance with the help of the RG-based equations. In Sec. III, we discuss the self-consistency requirements for a scale-setting method, where a graphical explanation of these requirements is also given. In Secs. IV and V, we present a detailed discussion on PMC and PMS scale-setting methods, respectively. Section VI provides a summary.

II. RENORMALIZATION-GROUP-BASED EQUATIONS AND THE RENORMALIZATION-GROUP INVARIANCE

The scale dependence of the running coupling in gauge theory is controlled by the RG equation

$$\beta^{\mathcal{R}} = \frac{\partial}{\partial \ln \mu^2}\left(\frac{\alpha_s^{\mathcal{R}}(\mu)}{4\pi}\right) = -\sum_{i=0}^{\infty} \beta_i^{\mathcal{R}}\left(\frac{\alpha_s^{\mathcal{R}}(\mu)}{4\pi}\right)^{i+2}, \tag{4}$$

where the superscript $\mathcal{R}$ stands for an arbitrary renormalization scheme, such as MS scheme [45], $\overline{\rm MS}$ scheme [32], and MOM scheme [46]. Note that the $\beta_i^{\mathcal{R}}$ functions for the MS and $\overline{\rm MS}$ schemes are the same [47]. Various terms in $\beta_0^{\mathcal{R}}, \beta_1^{\mathcal{R}}, \ldots$, correspond to one-loop, two-loop, …, contributions, respectively. In general, the $\{\beta_i^{\mathcal{R}}\}$ are scheme dependent and depend on the quark mass m_f. According to the decoupling theorem, a quark with mass $m_f \gg \mu$ can be ignored, and we can often neglect m_f terms when $m_f \ll \mu$. Then, for every renormalization scale μ, one can divide the quarks into active ones with $m_f = 0$ and inactive ones that can be ignored. Within this framework, it is well known that the first two coefficients $\beta_{0,1}^{\mathcal{R}}$ are universal; i.e., $\beta_0^{\mathcal{R}} \equiv 11 - 2n_f/3$ and $\beta_1^{\mathcal{R}} \equiv 102 - 38n_f/3$ for n_f-active flavors. Hereafter, we simply write them as β_0 and β_1. It is noted that an analytic extension of $\alpha_s^{\overline{\rm MS}}$ that incorporates the finite-mass quark threshold effects into the running of the coupling has been suggested in Ref. [48]. However, numerically, it is found that taking finite quark mass effects into account analytically in the running, rather than using a fixed n_f between thresholds, leads to effects of

[1]The displacement for higher-order corrections can be obtained by carefully dealing with the differences of the $\{\beta_i\}$ series under different renormalization schemes.

[2]As argued by Grunberg [6] and Krasnikov [33], it is better to be called as the RG-improved effective coupling method. For simplicity, we retain the name as FAC as suggested by Stevenson [8].

the order of 1% for the one-loop running coupling [48]. Here we will work with the conventional $\{\beta_i^{\mathcal{R}}\}$ functions.

It will be convenient to use the first two universal coefficients β_0 and β_1 to rescale the coupling constant and the scale parameter in Eq. (4). By rescaling the coupling constant and the scale parameters as [44]

$$a^{\mathcal{R}} = \frac{\beta_1}{4\pi\beta_0}\alpha_s^{\mathcal{R}} \quad \text{and} \quad \tau_{\mathcal{R}} = \frac{\beta_0^2}{\beta_1}\ln\mu^2,$$

one can express the RG equation (4) in a simpler canonical form

$$\frac{da^{\mathcal{R}}}{d\tau_{\mathcal{R}}} = -(a^{\mathcal{R}})^2[1 + a^{\mathcal{R}} + c_2^{\mathcal{R}}(a^{\mathcal{R}})^2 + c_3^{\mathcal{R}}(a^{\mathcal{R}})^3 + \cdots], \tag{5}$$

where $c_i^{\mathcal{R}} = \beta_i^{\mathcal{R}}\beta_0^{i-1}/\beta_1^i$ for $i = 2, 3, \ldots$.

As an extension of the ordinary coupling constant, one can define a universal coupling constant $a(\tau, \{c_i\})$ to include the dependence on the scheme parameters $\{c_i\}$, which satisfies the following extended RG-based equations [44]:

$$\beta(a, \{c_i\}) = \frac{\partial a}{\partial \tau} = -a^2[1 + a + c_2 a^2 + c_3 a^3 + \cdots], \tag{6}$$

$$\beta_n(a, \{c_i\}) = \frac{\partial a}{\partial c_n} = -\beta(a, \{c_i\})\int_0^a \frac{x^{n+2}dx}{\beta^2(x, \{c_i\})}. \tag{7}$$

The scale equation (6) can be used to evolve the universal coupling function from one scale to another. The scheme equation (7), which was first suggested by Stevenson [8], can be used to relate the coupling functions under different schemes by changing $\{c_i\}$. A solution of the scale equation up to four-loop level has been given in Ref. [30], which agrees with that of the conventional RG equation obtained in the literature; cf., Ref. [49]. By comparing Eq. (5) with Eq. (6), there exists a value of $\tau = \tau_R$ for which

$$a^{\mathcal{R}}(\tau_{\mathcal{R}}) = a(\tau_{\mathcal{R}}, \{c_i^{\mathcal{R}}\}). \tag{8}$$

This shows that any coupling constant $a^{\mathcal{R}}(\tau_{\mathcal{R}})$ can be expressed by the universal coupling constant $a(\tau, \{c_i\})$ under proper correspondence; i.e., the coupling constant $a^{\mathcal{R}}(\tau_{\mathcal{R}})$ can be treated as a special case of the universal coupling constant $a(\tau, \{c_i\})$: Any usual coupling constant $a^{\mathcal{R}}(\tau_{\mathcal{R}})$ is equal to a universal coupling $a(\tau_{\mathcal{R}}, \{c_i\})$ by setting $\{c_i\}$ to be $\{c_i^{\mathcal{R}}\}$, since both coupling constants satisfy the same RG equation by using the same scheme parameters.

Grunberg has pointed out that [4–6] any perturbatively calculable physical quantity can be used to define an effective coupling constant by incorporating the entire radiative corrections into its definition. The effective coupling constant satisfies the same RG equation (and hence the same RG-based equations) as the usual (universal) coupling constant. Thus, the running behavior for both the effective coupling constant and the usual (universal) coupling constant are the same if their RG-based equations are calculated under the same choice of scheme parameters. This idea has later been discussed in detail by Refs. [50,51]. Such an effective coupling constant can be used as a reference to define the renormalization procedure, such as MS scheme and $\overline{\text{MS}}$ scheme.

The RG invariance states that a physical quantity should be independent of the renormalization scale and renormalization scheme [18,19,41–43]. This shows that if the effective coupling constant $a(\tau_{\mathcal{R}}, \{c_i^{\mathcal{R}}\})$ corresponds to a physical observable, it should be independent of any other scale $\tau_{\mathcal{S}}$ and any other scheme parameters $\{c_j^{\mathcal{S}}\}$; i.e.,

$$\frac{\partial a(\tau_{\mathcal{R}}, \{c_i^{\mathcal{R}}\})}{\partial \tau_{\mathcal{S}}} \equiv 0 \quad \text{[scale invariance]}, \tag{9}$$

$$\frac{\partial a(\tau_{\mathcal{R}}, \{c_i^{\mathcal{R}}\})}{\partial c_j^{\mathcal{S}}} \equiv 0 \quad \text{[scheme invariance]}. \tag{10}$$

Demonstration: We provide an intuitive demonstration for the RG invariance from the above RG-based equations. Given two effective coupling constants $a(\tau_{\mathcal{R}}, \{c_i^{\mathcal{R}}\})$ and $a(\tau_{\mathcal{S}}, \{c_i^{\mathcal{S}}\})$ defined under two different schemes $\mathcal{R}$ and $\mathcal{S}$, one can expand $a(\tau_{\mathcal{R}}, \{c_i^{\mathcal{R}}\})$ in a power series of $a(\tau_{\mathcal{S}}, \{c_i^{\mathcal{S}}\})$ through a Taylor expansion:

$$\begin{aligned} &a(\tau_{\mathcal{R}}, \{c_i^{\mathcal{R}}\}) \\ &= a(\tau_{\mathcal{S}} + \bar{\tau}, \{c_i^{\mathcal{S}} + \bar{c}_i\}) \\ &= a(\tau_{\mathcal{S}}, \{c_i^{\mathcal{S}}\}) + \left(\frac{\partial a}{\partial \tau}\right)_{\mathcal{S}}\bar{\tau} + \sum_i\left(\frac{\partial a}{\partial c_i}\right)_{\mathcal{S}}\bar{c}_i \\ &\quad + \frac{1}{2!}\left[\left(\frac{\partial^2 a}{\partial \tau^2}\right)_{\mathcal{S}}\bar{\tau}^2 + 2\left(\frac{\partial^2 a}{\partial \tau \partial c_i}\right)_{\mathcal{S}}\bar{\tau}\bar{c}_i + \sum_{i,j}\left(\frac{\partial^2 a}{\partial c_i \partial c_j}\right)_{\mathcal{S}}\bar{c}_i\bar{c}_j\right] \\ &\quad + \frac{1}{3!}\left[\left(\frac{\partial^3 a}{\partial \tau^3}\right)_{\mathcal{S}}\bar{\tau}^3 + \cdots\right] + \cdots, \end{aligned} \tag{11}$$

where $\bar{\tau} = \tau_{\mathcal{R}} - \tau_{\mathcal{S}}$, $\bar{c}_i = c_i^{\mathcal{R}} - c_i^{\mathcal{S}}$, and the subscript $\mathcal{S}$ next to the partial derivatives means they are evaluated at the point $(\tau_{\mathcal{S}}, \{c_i^{\mathcal{S}}\})$.

The right-hand side of Eq. (11) can be regrouped according to the different orders of scheme parameters $\{\bar{c}_i\}$. After differentiating both sides of Eq. (11) over $\tau_{\mathcal{S}}$, we obtain

$$\begin{aligned} \frac{\partial a(\tau_{\mathcal{R}}, \{c_i^{\mathcal{R}}\})}{\partial \tau_{\mathcal{S}}} &= \frac{\partial^{(n+1)} a(\tau_{\mathcal{S}}, \{c_i^{\mathcal{S}}\})}{\partial \tau_{\mathcal{S}}^{(n+1)}}\frac{\bar{\tau}^n}{n!} \\ &\quad + \sum_i \frac{\partial^{(n+1)} a(\tau_{\mathcal{S}}, \{c_i^{\mathcal{S}}\})}{\partial c_i^{\mathcal{S}} \partial \tau_{\mathcal{S}}^{(n)}}\frac{\bar{\tau}^{n-1}\bar{c}_i}{(n-1)!} + \cdots, \end{aligned} \tag{12}$$

where n stands for the highest perturbative order for a fixed-order calculation. It is noted that Eq. (12) can be further simplified with the help of RG equations

(6) and (7). If we set $n \to \infty$, the right-hand side of Eq. (12) tends to zero, and we obtain the scale-invariance equation (9). This shows that if $a(\tau_{\mathcal{R}}, \{c_i^{\mathcal{R}}\})$ corresponds to a physical observable (corresponding to the case of $n \to \infty$), it will be independent of any other scale τ_S. Similarly, doing the first derivative of $a(\tau_{\mathcal{R}}, \{c_i^{\mathcal{R}}\})$ with respect to the scheme parameter $c_j^{\mathcal{S}}$, one can obtain the scheme-invariance equation (10).

In other words, if one uses an effective coupling constant $a(\tau_{\mathcal{S}}, \{c_i^{\mathcal{S}}\})$ under the renormalization scheme $\mathcal{S}$ and with an initial renormalization scale $\{\tau_{\mathcal{S}}\}$ to predict the value of $a(\tau_{\mathcal{R}}, \{c_i^{\mathcal{R}}\})$, the RG invariances (9) and (10) tell us the following:

(i) If we have summed all types of $c_i^{\mathcal{S}}$ terms (or equivalently the $\{\beta_i^{\mathcal{S}}\}$ terms) into the coupling constant, as is the case of an infinite-order calculation, then our final prediction of $a(\tau_{\mathcal{R}}, \{c_i^{\mathcal{R}}\})$ will be independent of any choice of initial scale $\tau_{\mathcal{S}}$ and renormalization scheme $\mathcal{S}$.

(ii) According to Eq. (12), for a fixed-order estimation (i.e., $n \neq \infty$), there is some residual initial-scale dependence. This is reasonable: As shown by Eq. (11), for a fixed-order calculation, the unknown $\{\beta_i^{\mathcal{S}}\}$ terms in the higher orders are necessary to cancel the scale dependence from the lower-order terms.

If we can find a proper way to sum up all the known types of $\{\beta_i^{\mathcal{S}}\}$ terms into the coupling constant, and at the same time suppress the contributions from those unknown types of $\{\beta_i^{\mathcal{S}}\}$ terms effectively, such residual initial scale dependence can be greatly suppressed. The PMC has been designed for such a purpose [29–31], whose properties will be discussed in more detail in the following sections.

(iii) If setting all the differences of the renormalization scheme parameters, $\bar{c}_i \equiv 0$ $(i = 1, 2, \ldots)$, Eq. (11) returns to a scale-expansion series for the coupling constant expanding over itself but specified at another scale; i.e.,

$$a(\tau_{\mathcal{R}}, \{c_i^{\mathcal{R}}\}) = a(\tau_{\mathcal{S}}, \{c_i^{\mathcal{R}}\}) + \left(\frac{\partial a(\tau_{\mathcal{S}}, \{c_i^{\mathcal{R}}\})}{\partial \tau_{\mathcal{S}}}\right)\bar{\tau} + \frac{1}{2!}\left(\frac{\partial^2 a(\tau_{\mathcal{S}}, \{c_i^{\mathcal{R}}\})}{\partial \tau_{\mathcal{S}}^2}\right)\bar{\tau}^2 + \frac{1}{3!}\left(\frac{\partial^3 a(\tau_{\mathcal{S}}, \{c_i^{\mathcal{R}}\})}{\partial \tau_{\mathcal{S}}^3}\right)\bar{\tau}^3 + \cdots. \quad (13)$$

Using the RG scale equation (6), the right-hand side of the above equation can be rewritten as a perturbative series of $a(\tau_{\mathcal{S}}, \{c_i^{\mathcal{R}}\})$, whose coefficient at each order is a $\{\beta_i^{\mathcal{R}}\}$ series.

If one considers N_c to be an analytic variable, then the scale setting known from the non-Abelian theory $SU(N_c)$ must agree with the Abelian QED theory at $N_c \to 0$. This shows that the above discussions are also suitable for QED; i.e., by taking the limit $N_c \to 0$ at fixed $\alpha = C_F \alpha_s$ with $C_F = (N_c^2 - 1)/2N_c$, we effectively return to the QED case [52,53].

III. SELF-CONSISTENCY REQUIREMENTS FOR A SCALE-SETTING METHOD

It has been noted that if one knows how to set the optimal scale in all cases, then one can translate the result freely from one scheme to another scheme through proper scale relations [34,35]. This observation has later been emphasized in Ref. [20], where the scale transformation among different schemes are called CSRs. It shows that even though the expansion coefficients could be different under different renormalization schemes, after a proper scale setting, one can find a relation between the effective renormalization scales, which ensures the total result remains the same under any renormalization schemes. For simplicity, following the suggestion of Ref. [40], we omit the scheme parameters in the coupling constant in discussing the self-consistent requirements for a scale-setting method, but will retrieve them when necessary.

In principle, the correctness of a scale-setting method can be judged by experimental data. However, it has been suggested that some self-consistency requirements can shed light on the reliability of the scale-setting method [40], in which some initial discussions have been presented. These requirements together with their explanations are listed in the following:

(1) *Existence* and *uniqueness* of the renormalization scale μ. Any scale-setting method must satisfy these two requirements.

(2) *Reflexivity*. Given an effective coupling $\alpha_s(\mu)$ specified at a renormalization scale μ, we can express it in terms of itself but specified at another renormalization scale μ',

$$\alpha_s(\mu) = \alpha_s(\mu') + f_1(\mu, \mu')\alpha_s^2(\mu') + \cdots, \quad (14)$$

where $f_1(\mu, \mu') \propto \ln(\mu^2/\mu'^2)$. When the scale μ' is chosen to be μ, the above equation reduces to a trivial identity.

From the scale invariance (9), up to infinite orders, we have

$$\frac{\partial \alpha_s(\mu)}{\partial \ln \mu'^2} \equiv 0. \quad (15)$$

This, inversely, means that if $\alpha_s(\mu)$ is known (say, an experimentally measured effective coupling), and we try to use the above perturbative equation to *predict* $\alpha_s(\mu)$ from itself, then any deviation of μ' from μ would lead to an inaccurate result owing to the truncation of expansion series.

More explicitly, for a fixed-order expansion with the highest perturbative order n, from Eq. (12), we obtain

$$\frac{\partial \alpha_s(\mu)}{\partial \ln \mu'^2} \propto \frac{(\ln \mu^2/\mu'^2)^n}{n!} \frac{\partial^{(n+1)} \alpha_s(\mu')}{\partial (\ln \mu'^2)^{(n+1)}}.$$

This shows, generally, the right-hand side of Eq. (14) depends on μ' at any fixed order.

Thus, to get a correct fixed-order estimate for $\alpha_s(\mu)$, a self-consistency scale setting must take the unique value $\mu' = \mu$ on the right-hand side of Eq. (14). If a scale setting satisfies such property, we say it is *reflexive*.

It is found that the reflexivity is a basic requirement for a self-consistency scale-setting method and for the physical (effective) coupling constant $\alpha_s(\mu)$, which provides the necessary condition for the following two properties: *symmetry* and *transitivity*; i.e., if a scale setting does not satisfy the reflexivity, it cannot satisfy the following two properties: symmetry and transitivity either.

(3) *Symmetry.* Given two different effective coupling constants $\alpha_{s1}(\mu_1)$ and $\alpha_{s2}(\mu_2)$ under two different renormalization schemes, we can expand any one of them in terms of the other:

$$\alpha_{s1}(\mu_1) = \alpha_{s2}(\mu_2) + r_{12}(\mu_1, \mu_2)\alpha_{s2}^2(\mu_2) + \cdots,$$
$$\alpha_{s2}(\mu_2) = \alpha_{s1}(\mu_1) + r_{21}(\mu_2, \mu_1)\alpha_{s1}^2(\mu_1) + \cdots.$$

After a general scale setting, we have

$$\alpha_{s1}(\mu_1) = \alpha_{s2}(\mu_2^*) + \tilde{r}_{12}(\mu_1, \mu_2^*)\alpha_{s2}^2(\mu_2^*) + \cdots,$$
$$\alpha_{s2}(\mu_2) = \alpha_{s1}(\mu_1^*) + \tilde{r}_{21}(\mu_2, \mu_1^*)\alpha_{s1}^2(\mu_1^*) + \cdots.$$

Note the following:

(i) The new effective scales $\mu_{1,2}^*$ may or may not be equal to $\mu_{1,2}$, depending on the choice of the scale-setting method. The coefficients $\tilde{r}_{12}$ and $\tilde{r}_{21}$ are changed accordingly in order to obtain a consistent result.

(ii) We have implicitly set the effective scales at NLO level to be equal to the LO ones. We will adopt this choice throughout the paper. The effective scales for the highest-order terms are usually taken as the same effective scales at the one-lower order, since they are the scales strictly set by the known terms [30,44].

Setting $\mu_2^* = \lambda_{21}\mu_1$ and $\mu_1^* = \lambda_{12}\mu_2$, if

$$\lambda_{12}\lambda_{21} = 1, \tag{16}$$

we say that the scale setting is symmetric.

Explanation: If $\mu_2^* = \lambda_{21}\mu_1$ and $\mu_1^* = \lambda_{12}\mu_2$, we obtain

$$\alpha_{s1}(\mu_1) = \alpha_{s2}(\lambda_{21}\mu_1) + \tilde{r}_{12}(\mu_1, \lambda_{21}\mu_1)\alpha_{s2}^2(\lambda_{21}\mu_1) + \cdots \tag{17}$$

and

$$\alpha_{s2}(\mu_2) = \alpha_{s1}(\lambda_{12}\mu_2) + \tilde{r}_{21}(\mu_2, \lambda_{12}\mu_2)\alpha_{s1}^2(\lambda_{12}\mu_2) + \cdots. \tag{18}$$

As a combination of Eqs. (17) and (18), we obtain

$$\alpha_{s1}(\mu_1) = \alpha_{s1}(\lambda_{12}\lambda_{21}\mu_1) + [\tilde{r}_{12}(\mu_1, \lambda_{21}\mu_1) + \tilde{r}_{21}(\lambda_{21}\mu_1, \lambda_{12}\lambda_{21}\mu_1)]\alpha_{s1}^2(\lambda_{12}\lambda_{21}\mu_1) + \cdots. \tag{19}$$

From the reflexivity property, if a scale setting is symmetric, i.e., satisfying Eq. (16), we will obtain

$$\tilde{r}_{12}(\mu_1, \mu_2^*) + \tilde{r}_{21}(\mu_2, \mu_1^*) = 0, \tag{20}$$

and vice versa. This shows that the symmetry property (16) and the relation (20) are mutually necessary and sufficient conditions.

The symmetry feature is necessary since it further gives us a unique relation for the scales before and after the scale setting,

$$\mu_1\mu_2 = \mu_1^*\mu_2^*.$$

(4) *Transitivity.* Given three effective coupling constants $\alpha_{s1}(\mu_1)$, $\alpha_{s2}(\mu_2)$, and $\alpha_{s3}(\mu_3)$ under three renormalization schemes, we can expand any one of them in terms of the other; i.e.,

$$\alpha_{s1}(\mu_1) = \alpha_{s2}(\mu_2) + r_{12}(\mu_1, \mu_2)\alpha_{s2}^2(\mu_2) + \cdots,$$
$$\alpha_{s2}(\mu_2) = \alpha_{s3}(\mu_3) + r_{23}(\mu_2, \mu_3)\alpha_{s3}^2(\mu_3) + \cdots,$$
$$\alpha_{s3}(\mu_3) = \alpha_{s1}(\mu_1) + r_{31}(\mu_3, \mu_1)\alpha_{s1}^2(\mu_1) + \cdots.$$

After a general scale setting, we obtain

$$\alpha_{s1}(\mu_1) = \alpha_{s2}(\mu_2^*) + \tilde{r}_{12}(\mu_1, \mu_2^*)\alpha_{s2}^2(\mu_2^*) + \cdots,$$
$$\alpha_{s2}(\mu_2) = \alpha_{s3}(\mu_3^*) + \tilde{r}_{23}(\mu_2, \mu_3^*)\alpha_{s3}^2(\mu_3^*) + \cdots,$$
$$\alpha_{s3}(\mu_3) = \alpha_{s1}(\mu_1^*) + \tilde{r}_{13}(\mu_3, \mu_1^*)\alpha_{s1}^2(\mu_1^*) + \cdots.$$

Setting $\mu_2^* = \lambda_{21}\mu_1$, $\mu_3^* = \lambda_{32}\mu_2$, and $\mu_1^* = \lambda_{13}\mu_3$, if

$$\lambda_{13}\lambda_{32}\lambda_{21} = 1, \tag{21}$$

we say that the scale setting is *transitive*.

Explanation: If $\mu_2^* = \lambda_{21}\mu_1$, $\mu_3^* = \lambda_{32}\mu_2$, and $\mu_1^* = \lambda_{13}\mu_3$, we obtain

$$\alpha_{s1}(\mu_1) = \alpha_{s2}(\lambda_{21}\mu_1) + \tilde{r}_{12}(\mu_1, \lambda_{21}\mu_1)\alpha_{s2}^2(\lambda_{21}\mu_1) + \cdots, \tag{22}$$

$$\alpha_{s2}(\mu_2) = \alpha_{s3}(\lambda_{32}\mu_2) + \tilde{r}_{23}(\mu_2, \lambda_{32}\mu_2)\alpha_{s3}^2(\lambda_{32}\mu_2) + \cdots, \tag{23}$$

$$\alpha_{s3}(\mu_3) = \alpha_{s1}(\lambda_{13}\mu_3) + \tilde{r}_{31}(\mu_3, \lambda_{13}\mu_3)\alpha_{s1}^2(\lambda_{13}\mu_3) + \cdots. \quad (24)$$

As a combination of Eqs. (22)–(24), we obtain

$$\alpha_{s1}(\mu_1) = \alpha_{s1}(\lambda_{13}\lambda_{32}\lambda_{21}\mu_1) + \alpha_{s1}^2(\lambda_{13}\lambda_{32}\lambda_{21}\mu_1) \times [\tilde{r}_{31}(\lambda_{32}\lambda_{21}\mu_1, \lambda_{13}\lambda_{32}\lambda_{21}\mu_1) + \tilde{r}_{23}(\lambda_{21}\mu_1, \lambda_{32}\lambda_{21}\mu_1) + \tilde{r}_{12}(\mu_1, \lambda_{21}\mu_1)] + \cdots. \quad (25)$$

From the reflexivity property, if a scale setting is transitive, i.e., satisfying Eq. (21), we will obtain

$$\tilde{r}_{12}(\mu_1, \mu_2^*) + \tilde{r}_{23}(\mu_2^*, \mu_3^*) + \tilde{r}_{31}(\mu_3^*, \mu_1) = 0, \quad (26)$$

and vice versa. This shows that the transitivity property (21) and the relation (26) are mutually necessary and sufficient conditions.

The transitivity property shows that under a proper scale-setting method, we have $\lambda_{21} \equiv \lambda_{23}\lambda_{31}$, which means that the scale ratio λ_{21} for any two effective couplings α_{s1} and α_{s2} is independent of the choice of an intermediate effective coupling α_{s3} under any renormalization scheme. Thus the relation between any two observables is independent of the choice of renormalization scheme. In fact, the transitivity property provides the theoretical foundation for the existence of CSRs among different physical observables [20].

The transitivity feature gives us a unique relation for all the scales before and after the scale setting,

$$\mu_1\mu_2\mu_3 = \mu_1^*\mu_2^*\mu_3^*.$$

The transitivity property is very important for a self-consistency scale setting, which is a natural requirement from the RG invariance. It has already been pointed out that why the renormalization group is called a *group* is mainly because of such a transitivity property [41–43].

The transitivity property (21) can be extended to an arbitrary number of effective coupling constants; i.e., if we have n_{th} effective coupling constants, which are related with a similar manner as above, then their transitivity relation is

$$\lambda_{1n}\lambda_{n(n-1)}\cdots\lambda_{32}\lambda_{21} = 1. \quad (27)$$

One may observe that the symmetry is a special case of transitivity, since if setting $\alpha_{s3}(\mu_3) \equiv \alpha_{s1}(\mu_1)$, we have $\lambda_{11} \equiv 1$ and $\tilde{r}_{11}(\mu_1, \mu_1) \equiv 0$ owing to the reflexivity, which thus changes the transitive relation $\lambda_{13}\lambda_{32}\lambda_{21} = 1$ into the symmetric relation $\lambda_{12}\lambda_{21} = 1$.

As a summary, a scale-setting method that satisfies the existence and uniqueness of the renormalization scale, reflexivity, symmetry, and transitivity effectively establishes equivalent relations among all the effective coupling constants and, thus, among all the physical observables.

A Graphic explanation of these requirements

In this subsection, we present a more intuitive explanation of these requirements based on the universal coupling $a(\tau, \{c_i\})$ and the RG-based equations (6) and (7).

In the RG-based equations (6) and (7), there is no explicit reference to the QCD parameters, such as the number of colors or the number of active flavors. Therefore, aside from its infinite dimensional character, $a(\tau, \{c_i\})$ is just a mathematical function like, say, Bessel functions or any other special functions [44]. In practice, because of the unknown higher-order scheme parameters $\{c_i\}$, we need to truncate the beta function $\beta(a, \{c_i\})$ and solve the universal coupling constant $a(\tau, \{c_i\})$ in a finite-dimensional subspace; i.e., we need to evaluate $a(\tau, \{c_i\})$ in a subspace where higher order $\{c_i\}$ terms are zero. In principle, this function can be computed to an arbitrary degree of precision, limited only by the truncation of the fundamental beta function.

In this formalism, any two effective coupling constants can be related by some evolution path on the hypersurface defined by $a(\tau, \{c_i\})$. In Fig. 2 we illustrate the paths that represent the operations of reflexivity, symmetry, and transitivity. We can pictorially visualize that the evolution paths satisfy all three self-consistency properties. A closed path starting and ending at the same point A represents the operation of identity. Since the predicted value does not depend on the chosen path, if the effective coupling constant at A is a_A, after completing the path we will also end up with an effective coupling a_A. Similarly, if we

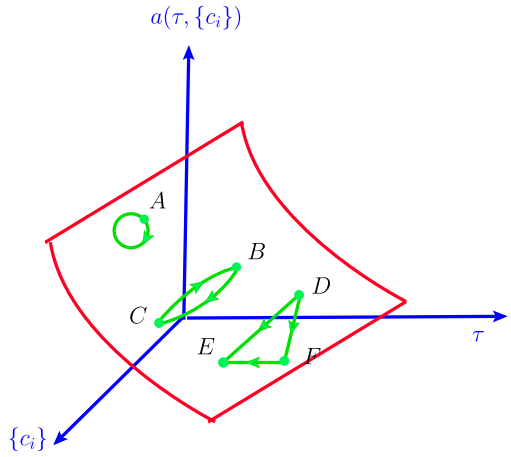

FIG. 2 (color online). Pictorial representation of the self-consistency of the scale-setting method through the universal coupling function $a(\tau, \{c_i\})$. Point A with a closed path represents the operation of reflexivity. The paths $\overline{BC}$ and $\overline{CB}$ represent the operation of symmetry, and the paths $\overline{DF}$, $\overline{FE}$, and $\overline{DE}$ represent the operation of transitivity.

evolve a_B at B to a value a_C at C, we are guaranteed that when we evolve a_C at C back to point B, the result will be a_B. Hence, the evolution equations also satisfy symmetry. Transitivity follows in a similar manner; i.e., going directly from D to E gives the same result as going from D to E through a third point F.

In the following two sections, we will make a detailed discussion on how these self-consistency conditions are satisfied or broken by the two frequently adopted scale-setting methods: BLM/PMC and PMS. As for FAC, its FAC scale is determined by requiring all higher-order corrections to be zero.[3] FAC satisfies all the above mentioned self-consistent requirements, whose demonstration is similar to that of BLM/PMC and is simpler [40], so we will not repeat it here.

IV. THE PMC SCALE SETTING

The PMC provides the principle underlying BLM scale setting, so if not specially stated, we usually treat them on equal footing.

A. What is PMC?

In the original BLM paper [11], the physical observable is expanded as

$$\rho = C_0 \alpha_{s,\overline{\rm MS}}(\mu)\left[1 + (An_f + B)\frac{\alpha_{s,\overline{\rm MS}}(\mu)}{\pi}\right], \qquad (28)$$

where μ is the renormalization scale, and the n_f term is attributable to the quark vacuum polarization. For clarity, we have taken the familiar $\overline{\rm MS}$ scheme. When absorbing all the NLO terms involving n_f into the running coupling, we obtain [11]

$$\rho = C_0 \alpha_{s,\overline{\rm MS}}(\mu^*)\left[1 + C_1^*\frac{\alpha_{s,\overline{\rm MS}}(\mu^*)}{\pi}\right], \qquad (29)$$

where

$$\mu^* = \mu \exp(3A) \quad \text{and} \quad C_1^* = \frac{33}{2}A + B. \qquad (30)$$

The new scale μ^* and the coefficient C_1^* are n_f independent. The term $33A/2$ in C_1^* serves to remove that part of the constant B that renormalizes the NLO coupling constant.

Through these procedures, it was suggested that the pQCD convergence can be greatly improved [11]. However, after a proper extension of BLM, it can do much more than that.

In deriving Eq. (29), Brodsky, Lepage, and Mackenzie already observed that to derive the correct scheme-independent LO QED/QCD scale, one should deal with the β_0 term rather than the n_f term. This point has lately been emphasized in Refs. [54,55], where an interesting feature for the NLO Balitsky-Fadin-Kuraev-Lipatov Pomeron intercept function $\omega(Q^2, 0)$ has been found; i.e., after using BLM scale setting, the function $\omega(Q^2, 0)$ has a very weak dependence on the gluon virtuality Q^2 in comparison with that derived from the conventional scale setting under the MOM scheme and $\overline{\rm MS}$ scheme [54]. The BLM has also been applied with some modifications for determining the effective scale in lattice perturbative theory by Lepage and Mackenzie [56], which greatly enhances the predictive power of lattice perturbative theory. However, BLM in its original form is difficult to apply to higher-order calculations because of the emergence of higher-order n_f terms as n_f^2 term, n_f^3 term, etc.

As an extension of BLM scale setting, a program to deal with higher order n_f terms associated with renormalization has been raised in Ref. [36], which suggests that one can expand the effective scale as a perturbative series. Later on, an enhanced discussion of this suggestion up to NNLO level has been presented in Ref. [20], where the perturbative series of the effective scale is exponentiated, which is consistent with PMC procedure. In that work it is pointed out that the n_f^2 term at the NNLO should be first identified with the β_0^2 term and then be absorbed into the coupling constant.[4]

The pioneering work for PMC is done in Ref. [29], which shows that a single global PMC scale, valid at LO, can be derived from basic properties of the perturbative QCD cross section. Later on, explicit formulas for setting PMC scales up to NNLO has been presented in Ref. [30]. It has also been pointed out that by introducing the PMC-BLM correspondence principle, we can improve the previous BLM procedure to deal with the process up to all orders, whose estimation is the same as PMC. In this sense, we say that PMC and BLM are equivalent to each other. Recently, by applying PMC to the top-quark pair hadroproduction up to the NNLO level at the Tevatron and LHC colliders, the most striking feature of PMC has been observed, which shows that the PMC scales and the resulting finite-order PMC predictions are both to high accuracy independent of the choice of an initial renormalization scale, consistent with RG invariance [14,25,31]. This implies that the serious systematic renormalization scale error introduced by using conventional scale setting can be eliminated by PMC through a self-consistency way.

A ''flow chart'' that illustrates the PMC procedure is presented in Fig. 3, where $\mathcal{R}$ stands for an arbitrary renor-

[3]This method itself is useful to define an effective coupling constant for a physical process [4–6]. However, it will give wrong results when applied to QED processes. The FAC forces all higher-order corrections to vanish and runs the risk of the better approximation being *dragged down* by the poorer one [8].

[4]Strictly, together with the n_f term at the same order, it should be arranged into a proper linear combination of the β_1 term and the β_0^2 term; the β_0^2 term will be absorbed into the LO PMC scale, and the β_1 term will be absorbed into the NLO PMC scale [30].

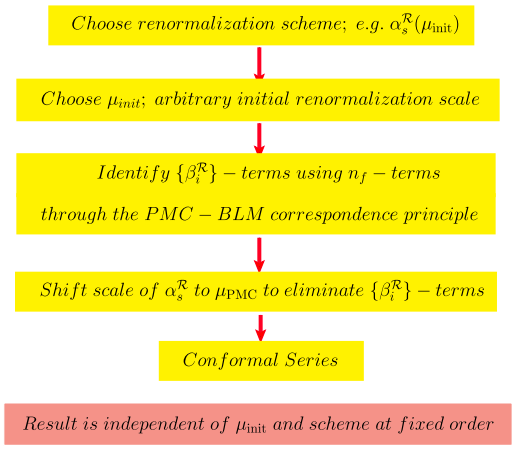

FIG. 3 (color online). A "flow chart" that illustrates the PMC procedure, where $\mathcal{R}$ stands for an arbitrary renormalization scheme.

malization scheme. The PMC provides an unambiguous and systematic way to set the optimized renormalization scale up to all orders; i.e., we first arrange all the coefficients, which usually are given as a series in n_f, for each perturbative order into $\{\beta_i^{\mathcal{R}}\}$ terms or non-$\{\beta_i^{\mathcal{R}}\}$ terms, and absorb different types of the $\{\beta_i^{\mathcal{R}}\}$ term into the running coupling constant, order by order.[5] Different types of the $\{\beta_i^{\mathcal{R}}\}$ term are absorbed into different PMC scales. Different skeleton graphs can have different PMC scales. The PMC scales themselves will be a perturbative expansion series in α_s. After this procedure, all nonconformal $\{\beta_i^{\mathcal{R}}\}$ terms in the perturbative expansion are resummed into the running couplings so that the remaining terms in the perturbative series are identical to that of a conformal theory; i.e., the corresponding theory with $\{\beta_i^{\mathcal{R}}\} \equiv \{0\}$.

As a simple explanation of PMC, for the coefficient $\mathcal{C}_1(\mu)$ at the NLO level, we have

$$\mathcal{C}_1(\mu) = \mathcal{C}_{10}(\mu) + \mathcal{C}_{11}(\mu) n_f, \tag{31}$$

$$= \tilde{\mathcal{C}}_{10}(\mu) + \tilde{\mathcal{C}}_{11}(\mu)\beta_0, \tag{32}$$

where μ stands for an arbitrary initial renormalization scale, the coefficients $\mathcal{C}_{10}(\mu)$ and $\mathcal{C}_{11}(\mu)$ are n_f independent, $\tilde{\mathcal{C}}_{10} = \mathcal{C}_{10} + \frac{33}{2}\mathcal{C}_{11}$, and $\tilde{\mathcal{C}}_{11} = -\frac{3}{2}\mathcal{C}_{11}$. The LO PMC scale $\mu_{\rm PMC}$ is then set by the condition

$$\tilde{\mathcal{C}}_{11}(\mu_{\rm PMC}) = 0. \tag{33}$$

This prescription ensures that, as in QED, vacuum polarization contributions attributable to the light-fermion pairs are absorbed into the coupling constant. Note that because $\mathcal{C}_{11} \propto \tilde{\mathcal{C}}_{11}$, one can practically obtain the PMC scale by using the equation $\mathcal{C}_{11}(\mu_{\rm PMC}) \equiv 0$, which is usually adopted in the literature.[6] However, one should keep in mind that Eq. (33) is exact.

The PMC-BLM correspondence principle suggested in Ref. [30] is based on the fact that the purpose of the running coupling in any gauge theory is to sum up all the terms involving the $\{\beta_i^{\mathcal{R}}\}$ functions; conversely, one can find all the needed $\{\beta_i^{\mathcal{R}}\}$ terms at any relevant order by identifying terms arising from the order-by-order expansion of the running coupling. This principle provides a one-to-one correspondence between the n_f series and the $\beta_i^{\mathcal{R}}$ series, and it provides a practical way of identifying the terms in the n_f series in the required $\beta_i^{\mathcal{R}}$ series. The $\{\beta_i^{\mathcal{R}}\}$ series derived from Eq. (13) provides the foundation for the PMC-BLM correspondence principle, since it shows which $\{\beta_i^{\mathcal{R}}\}$ terms should be kept at a specific perturbative order. This procedure provides a convenient and consistent way of treating the $\{\beta_i^{\mathcal{R}}\}$ terms in the perturbative series. Its advantages will be shown in the next subsection. Such a choice of $\{\beta_i^{\mathcal{R}}\}$ series is not completely identical to the suggestion of Refs. [36–38,57]. In Refs. [36–38], as an extension of BLM scale setting to all orders, the large β_0 approximation is adopted with some modifications to simplify the calculation (called the seBLM and the xBLM approaches [37]).[7]

B. The properties of PMC

It is straightforward to verify that PMC satisfies all the self-consistency requirements outlined above.

(1) The existence and uniqueness of the renormalization scale μ are guaranteed, since the scale-setting conditions for PMC are often linear equations in $\ln\mu^2$.

As a simpler explanation, if the NLO coefficient $\mathcal{C}_1(\mu)$ in Eq. (1) has the form

$$\mathcal{C}_1(\mu) = (a + bn_f) + (c + dn_f)\ln\mu^2, \tag{34}$$

where a, b, c, and d are constants free of n_f. The LO PMC scale can be set as

$$\ln\mu_{\rm PMC}^{\rm LO} = -\frac{b}{2d} + \mathcal{O}(\alpha_s), \tag{35}$$

where the higher-order α_s terms will be determined by n_f terms at the NLO level or even higher levels.

[5] In practice, we can directly deal with n_f terms of the coefficients without changing them into $\{\beta_i^{\mathcal{R}}\}$ terms and eliminate the n_f terms from the highest power to none also in an order-by-order manner. The results are the same because of the PMC-BLM correspondence [30].

[6] This should be used with care, since if $\mathcal{C}_{10}$ is a constant free of scale, then such a practical way will give the wrong NLO coefficient rather than the correct one $\tilde{\mathcal{C}}_{10}$.

[7] Theoretical differences for different treatments will be discussed in more detail and will be presented elsewhere.

(2) Reflexivity is satisfied. The PMC requires all $\ln(\mu^2/\mu'^2)$ terms in Eq. (14) to vanish; thus we obtain

$$\mu' = \mu.$$

(3) Symmetry is trivial, because after PMC scale setting, we always have

$$\tilde{r}_{12}(\mu_1, \mu_2^*) = -\tilde{r}_{21}(\mu_2, \mu_1^*).$$

That is, the two NLO coefficients differ only by a sign. Thus, requiring one of them to be $\{\beta_i^{\mathcal{R}}\}$ independent is equivalent to requiring the other one also to be $\{\beta_i^{\mathcal{R}}\}$ independent. This argument ensures the symmetric relation, $\lambda_{12}\lambda_{21} = 1$, is satisfied after PMC scale setting.

(4) Transitivity is also satisfied by PMC. After PMC scale setting, the two coefficients $\tilde{r}_{12}(\mu_1, \mu_2^*)$ and $\tilde{r}_{23}(\mu_2^*, \mu_3^*)$ in the following two series:

$$\alpha_{s1}(\mu_1) = \alpha_{s2}(\mu_2^*) + \tilde{r}_{12}(\mu_1, \mu_2^*)\alpha_{s2}^2(\mu_2^*) + \mathcal{O}(\alpha_{s2}^3) \tag{36}$$

and

$$\alpha_{s2}(\mu_2^*) = \alpha_{s3}(\mu_3^*) + \tilde{r}_{23}(\mu_2^*, \mu_3^*)\alpha_{s3}^2(\mu_3^*) + \mathcal{O}(\alpha_{s3}^3), \tag{37}$$

should be independent of $\{\beta_i\}$. After substituting Eq. (37) into Eq. (36), we obtain

$$\alpha_{s1}(\mu_1) = \alpha_{s3}(\mu_3^*) + [\tilde{r}_{12}(\mu_1, \mu_2^*) + \tilde{r}_{23}(\mu_2^*, \mu_3^*)]\alpha_{s3}^2(\mu_3) + \mathcal{O}(\alpha_{s3}^3). \tag{38}$$

We see that the new NLO coefficient $[\tilde{r}_{12}(\mu_1, \mu_2^*) + \tilde{r}_{23}(\mu_2^*, \mu_3^*)]$ will also be $\{\beta_i^{\mathcal{R}}\}$ independent, since it is the sum of two $\{\beta_i^{\mathcal{R}}\}$-independent quantities. These arguments ensure the transitive relation, $\lambda_{31} = \lambda_{32}\lambda_{21}$, will be satisfied after PMC scale setting.

As a combination of all the above mentioned PMC features, the advantages of PMC are clear[8]:

(1) It keeps the information of the higher-order corrections but in a more convergent perturbative series. After PMC scale setting, the divergent *renormalon* series with $n!$ growth disappear in the perturbative series, so that a more convergent perturbative series is obtained. Such better convergence has already been found in the original BLM paper [11] and the following BLM literature even at the NLO level.

(2) After PMC scale setting, the renormalization scale dependence is transformed to the initial renormalization scale dependence, and it is found that such initial renormalization scale dependence can be highly suppressed or even eliminated:

(i) The resulting expressions are conformally invariant and thus do not depend on the choice of renormalization scheme.

(ii) One can obtain proper scale displacements among the PMC scales that are derived under different schemes or conventions.

(iii) One can obtain the CSR between any two physical observables such as the generalized Crewther relation connecting the Bjorken sum rule to the e^+e^- annihilation cross section. Many leading order CSRs have been derived in Ref. [20]. The CSRs have no scale ambiguity and are independent of the choice of renormalization scheme. The relative scales in the CSR ensure that two observables pass through each quark threshold in synchrony. The coefficients in the CSR can be identified with those obtained in conformally invariant gauge theory [38,58–61].

(iv) There can be some residual scheme dependence for a fixed-order calculation owing to unknown higher-order terms. However, this scheme dependence can be highly suppressed in a similar way as that of the residual initial scale dependence; such effects can be estimated by using the RG-based scheme equations [30,44].

(3) The PMC provides a fundamental and systematic way to set the optimized renormalization scale for the fixed-order calculation. In principle, PMC needs an initial renormalization scale to initialize it. However, it is found that the estimates after PMC scale setting are independent of any choice of the initial renormalization scale—even the PMC scales themselves are independent of any choice of initial scale and are ''physical'' at any fixed order. This is because the PMC scale itself is a perturbative series and the unknown higher-order $\{\beta_i^{\mathcal{R}}\}$ terms are to be absorbed into the higher-order term of PMC scale and will be strongly power suppressed. One example of this behavior is shown in Refs. [14,25,31], where the top-quark pair total cross section and the top-quark pair forward-backward asymmetry are almost free from the choice of initial renormalization scale even at the NNLO level.

(4) Moreover, it is found that the PMC scale setting can also be adopted for the QED case. The variable N_C can be taken as an analytic variable. In the Abelian

[8]In the PMC, the same procedure is valid for both spacelike and timelike arguments; in particular, this leads to well-behaved perturbative expansion, since all the large $\{\beta_i^{\mathcal{R}}\}$-dependent terms on the timelike side involving π^2 terms are fully absorbed into the coupling. The PMC does not change the spacelike or timelike nature of the initial renormalization scale Q_0, since, in general, all the PMC scales are equal to Q_0 times an exponential factor [30].

limit $N_C \to 0$ at fixed $\alpha = C_F\alpha_s$ with $C_F = (N_c^2-1)/2N_c$ [52], the PMC method agrees with the standard Gell-Mann-Low procedure for setting the renormalization scale in QED, a consistency requirement of analyticity of Yang-Mills gauge theories.

(5) After PMC scale setting, the number of active flavors n_f is correctly determined [48]. Using the PMC ensures that the expansion is unchanged as one passes each quark threshold, since all vacuum-polarization effects attributable to each new quark are automatically absorbed into the effective coupling constant.

(6) The argument of the running coupling has timelike or spacelike values appropriate to the physics of the PMC scale; for example, the scale of the QED coupling that sums all vacuum polarization corrections in the lowest order $e^+e^- \to \mu^+\mu^-$ amplitude is $\alpha(t)$ in the Gell-Mann-Low scheme. As in QED, the running QCD coupling is complex in the timelike domain, reflecting the contribution of diagrams with physical unitarity cuts.

V. THE PMS SCALE SETTING

A. What is PMS?

The PMS states that [7–10] if an estimate depends on some *unphysical* parameters,[9] then their values should be chosen so as to minimize the sensitivity of the estimate to small variations in these parameters; i.e., this method chooses $\mu_{\rm PMS}$ at the stationary point of ρ_N:

$$\frac{\partial \rho_N}{\partial \mu}\bigg|_{\mu=\mu_{\rm PMS}} \equiv 0, \tag{39}$$

or

$$\frac{\partial \rho_N}{\partial \ln(\mu^2)}\bigg|_{\mu=\mu_{\rm PMS}} \equiv 0. \tag{40}$$

Here Eq. (40) can be solved with the help of the usual renormalization group equation (4).

B. The properties of PMS

Unlike the case of PMC, in general, there are no known theorems that guarantee the existence or the uniqueness of the PMS solution. Although for practical cases, PMS does provide solutions, and when there are more than one solution usually only one of them lies in the physically reasonable region [7–10], these observations alone do not guarantee that PMS will be trouble-free for new processes.

To discuss PMS properties in a renormalization scheme-independent way, following the suggestion of Ref. [40], we adopt the 't Hooft scheme [62] to define the running behavior of the effective coupling constant. Under the 't Hooft scheme, all the scheme parameters $\{c_i\}$ are set to zero, and Eq. (6) simplifies to

$$\frac{da}{d\tau} = -a^2(1+a), \tag{41}$$

whose solution can be written as

$$\tau = \frac{1}{a} + \ln\left(\frac{a}{1+a}\right). \tag{42}$$

In the above solution, for convenience, we have redefined τ as $\frac{\beta_0^2}{\beta_1}\ln(\frac{\mu^2}{\Lambda^{\prime\rm tH2}_{\rm QCD}})$, where $\Lambda^{\prime\rm tH}_{\rm QCD}$ is the asymptotic scale under the 't Hooft scheme. The 't Hooft coupling constant has a formal singularity, $a(\tau,\{c_i\}) \equiv a(0,\{0\}) = \infty$, which provides a precise definition for the asymptotic scale $\Lambda^{\prime\rm tH}_{\rm QCD}$ [62]; i.e., it is defined to be the pole of the coupling function.

Given two effective coupling constants a_1 and a_2 under the 't Hooft scheme, they are related by the perturbative series

$$a_1(\tau_1) = a_2(\tau_2) + (\tau_2-\tau_1)a_2^2(\tau_2) + \cdots. \tag{43}$$

PMS proposes the choice of μ_2 (or equivalently, τ_2) at the stationary point, i.e.,

$$\frac{da_1}{d\tau_2} = 0 = \frac{d}{d\tau_2}[a_2(\tau_2) + (\tau_2-\tau_1)a_2^2(\tau_2)]. \tag{44}$$

Then, we obtain the condition

$$1 + a_2 = \frac{1}{2(\tau_1-\tau_2)}. \tag{45}$$

In order to obtain τ_2 in terms of τ_1, one must solve the last equation in conjunction with

$$\frac{1}{a_2} + \log\left(\frac{a_2}{1+a_2}\right) = \tau_2. \tag{46}$$

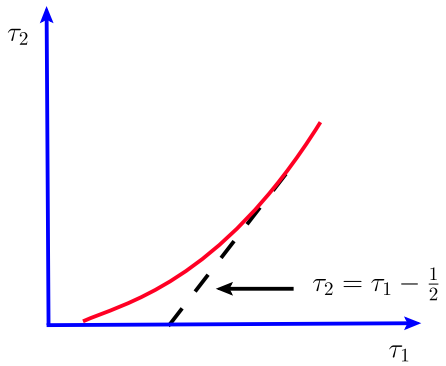

FIG. 4 (color online). The dependence of the PMS scale parameter τ_2 as a function of the external scale parameter τ_1.

[9]Here the *unphysical* parameter means that which is known not to affect the true result.

In Fig. 4 we present the graphical solution of the PMS scale parameter τ_2 as a function of the external scale parameter τ_1. One may observe two points:

(i) $\tau_2 \geq \tau_1 - \frac{1}{2}$. Since $\tau_2 \neq \tau_1$ in any cases, so PMS explicitly violates the reflexivity. For a fixed-order estimation, when one uses an effective coupling constant to predict itself, the application of PMS would lead to an inaccurate result.

(ii) In the large momentum region ($\tau_1 \gg 1$), we obtain $a_2(\tau_2) \to 0$, and

$$\tau_2 \simeq \tau_1 - \frac{1}{2}. \tag{47}$$

Under the same renormalization scheme R, we have the same asymptotic parameter $\Lambda_{\rm QCD}^{\prime {\rm tH}\text{-}R}$ for both a_1 and a_2. Here $\Lambda_{\rm QCD}^{\prime {\rm tH}\text{-}R}$ is the 't Hooft scale associated with the R scheme, where the word *associated* means we are choosing the particular 't Hooft scheme that shares the same 't Hooft scale with the R scheme. Then the relation (47) in terms of μ_1 and μ_2 becomes

$$\mu_2 \simeq \mu_1 \exp\left(-\frac{\beta_1}{4\beta_0^2}\right). \tag{48}$$

More generally, it is found that after PMS scale setting, the scale displacement between any two scales μ_i and μ_j in the large momentum region is

$$\lambda_{ij} = \frac{\mu_i}{\mu_j} \simeq \exp\left(-\frac{\beta_1}{4\beta_0^2}\right). \tag{49}$$

This would mean that

$$\lambda_{12}\lambda_{21} \simeq \exp\left(-\frac{\beta_1}{2\beta_0^2}\right) \neq 1, \tag{50}$$

$$\lambda_{13}\lambda_{32}\lambda_{21} \simeq \exp\left(-\frac{3\beta_1}{4\beta_0^2}\right) \neq 1. \tag{51}$$

This shows that the PMS does not satisfy the symmetry and transitivity requirements. Let us point out that adding the scheme-parameter optimization in PMS does not change any of the above conclusions. It only makes the solution much more complicated [63]. The inability of PMS to meet these self-consistency requirements resides in that the derivative operations in general do not commute with the operations of reflexivity, symmetry, and transitivity.

As argued in Sec. III, any truncated perturbative series will explicitly break RG invariance (9); i.e., Eq. (9) can be only approximately satisfied for any fixed-order estimation. The precision depends on which perturbative order we have calculated, the convergence of the perturbative series, and how we set the renormalization scale. As shown by Eq. (40), the PMS requires the truncated series, i.e., the approximant of a physical observable, to satisfy the RG invariance near $\mu = \mu_{\rm PMS}$. This provides the underlying reason for why PMS does not satisfy the reflexivity, symmetry, and transitivity properties. This shows the necessity of further careful studies of theoretical principles lying below PMS.

The PMC and PMS scale-setting methods each gives specific predictions for physical observables at finite order; however, their predictions are very different:

(i) The PMC sums all $\{\beta_i^{\mathcal{R}}\}$ terms in an arbitrary renormalization scheme $\mathcal{R}$ in the fixed-order prediction into the running coupling, leaving the conformal series. It satisfies all of the RG properties, reflexivity, symmetry, and transitivity. The PMC prediction is thus scheme independent, and it automatically assigns the correct displacement of the intrinsic scales between schemes. The variation of the prediction away from the PMC scale exposes the nonzero $\{\beta_i^{\mathcal{R}}\}$-dependent terms. The PMC prediction does have small residual dependence on the initial choice of scale owing to the truncated unknown higher-order $\{\beta_i^{\mathcal{R}}\}$ terms, which will be highly suppressed by the proper choice of PMC scales.

(ii) The PMS chooses the renormalization scale such that the first derivative of the fixed-order calculation with respect to the scale vanishes. This criterion of minimal sensitivity gives predictions that are not the same as the conformal prediction, and the PMS prediction depends on the choice of renormalization scheme,[10] and it disagrees with QED scale setting in the Abelian limit. For example, in the case of $e^+e^- \to gq\bar{q}$, the PMS scale decreases with increasing gluon jet mass and increasing flavor number, opposite to the correct physical behavior [64]. The PMS does not satisfy the RG properties of symmetry, reflexivity, and transitivity, so that relations between observables depend on the choice of the intermediate renormalization scheme.

VI. SUMMARY

The conventional scale-setting procedure assigns an arbitrary range and an arbitrary systematic error to fixed-order pQCD predictions. As we have discussed in this article, this *ad hoc* assignment of the range and associated systematic error is unnecessary and can be eliminated by a proper scale-setting method.

Renormalization group invariance (9) states that a physical quantity should be independent of the renormalization scale and renormalization scheme. With the help of the RG-based equations that incorporate the scheme parameters, we have presented a general demonstration for the RG invariance by setting the perturbative series up to infinite orders.

[10] As shown in Ref. [63], by using the PMS together with the scheme equations (7) and the scheme-independent equation (10), such renormalization scheme dependence can be reduced to a certain degree through an order-by-order procedure.

We have discussed the necessary self-consistency conditions for a scale-setting method, such as the existence and uniqueness of the renormalization scale, reflexivity, symmetry, and transitivity. There properties are natural deductions of RG invariance. We have shown that PMC satisfies these requirements, whereas the PMS does not. We have also pictorially argued that the formalism based on the RG-based equations satisfies all these requirements for scale and scheme variation.

The PMS requires that the slope of the approximant of an observable vanishes at the renormalization point. With the help of the RG-based equations, it has been argued that PMS can provide renormalization-scheme-dependent estimates [7–10]. We have shown that the PMS violates the symmetry and transitivity properties of the renormalization group, and it does not reproduce the Gell-Mann-Low scale for QED observables. Equation (51) shows that the relation between any two physical observables after PMS scale setting depends on which renormalization scheme is chosen for the calculation, which explicitly breaks the *group properties* of the RG equations. In addition, the application of PMS to jet production gives unphysical results [64], since it sums physics into the running coupling not associated with renormalization. This implies the necessity of further careful studies of theoretical principles lying below PMS.

The PMC provides a fundamental and systematic way to set the optimized renormalization scale at fixed order in pQCD. The PMC has a solid theoretical background [29,30], it provides the underlying principle for BLM, and many PMC features have already been noted in the BLM literature. Most important, it is found after standard PMC scale setting, the theoretical prediction is essentially independent of the choice of initial renormalization scale, and the theorist's choice of renormalization scheme is consistent with the RG invariance.

The most important goal for a scale-setting method is to eliminate the renormalization scheme and initial scale dependences—more fundamental requirements than improving convergence of the pQCD series. In the literature, however, some extensions of BLM scale setting have concentrated on how to improve the pQCD convergence, such as the large β_0 expansion [65] and the sequential BLM (seBLM) and xBLM [36–38]. In fact, once one sets the scales properly, as PMC does, much better pQCD convergence than the conventional scale-setting method is automatic, since the divergent renormalon series with $n!$ growth has been absorbed into the effective scales and disappears in the perturbative series. An example of this improved convergence can be found in our analysis for the top-quark pair production at the NNLO level [14,25,31].

Two more subtle points for PMC scale setting:

(i) In some specific kinematical regions, such as for the heavy quark pair production in the threshold region, Coulomb-type corrections will lead to sizable contributions that are enhanced by factors of π/v, and the PMC scale can be relatively soft for $v \to 0$ (v, the heavy quark velocity). Thus the terms that are proportional to (π/v) or $(\pi/v)^2$ should be treated separately in that different PMC scales are adopted in the estimation [25,66].

(ii) The factorization scale μ_f that enters into the predictions for QCD inclusive reactions is introduced to match nonperturbative and perturbative aspects of the parton distributions in hadrons. The factorization scale occurs even for a conformal theory with $\{\beta_i^{\mathcal{R}}\} = 0$, where α_s is constant. The factorization scale should be chosen to match the nonperturbative bound state dynamics with perturbative Dokshitzer-Gribov-Lipatov-Altarelli-Parisi evolution. This can be done explicitly for electron-atom or atom-atom inelastic scattering processes in QED using the known bound state dynamics of atoms. This could also be done in hadron physics using nonperturbative models such as anti-de Sitter/QCD and light-front holography; recent reviews can be found in Refs. [67,68]. There is clearly no reason to equate the factorization scale to the renormalization scale [69]. We expect that the factorization scale ambiguity will also be reduced by applying the PMC scale setting to the kernels of Dokshitzer-Gribov-Lipatov-Altarelli-Parisi evolution equations.

In summary, the systematic application of the PMC can eliminate a major ambiguity of pQCD predictions from scale and scheme ambiguities, thus greatly improving the precision of tests of the Standard Model and the sensitivity to new physics at colliders and other experiments.

ACKNOWLEDGMENTS

We thank Hung-Jung Lu, Leonardo di Giustino, George Sterman, Andrei L. Kataev, Sergey V. Mikhailov, Zvi Bern, Stefan Hoeche, and Dmitry V. Shirkov for helpful conversations. This work was supported in part by the Program for New Century Excellent Talents in University under Grant No. NCET-10-0882, Natural Science Foundation of China under Grants Nos. 11075225 and 11275280, and the Department of Energy Contract No. DE-AC02-76SF00515, SLAC-PUB-15133.

[1] W. Celmaster and R. J. Gonsalves, Phys. Rev. D **20**, 1420 (1979); Phys. Rev. Lett. **42**, 1435 (1979).
[2] L. F. Abbott, Phys. Rev. Lett. **44**, 1569 (1980).
[3] A. J. Buras, Rev. Mod. Phys. **52**, 199 (1980).
[4] G. Grunberg, Phys. Lett. **95B**, 70 (1980).
[5] G. Grunberg, Phys. Lett. **110B**, 501 (1982).
[6] G. Grunberg, Phys. Rev. D **29**, 2315 (1984).
[7] P. M. Stevenson, Phys. Lett. **100B**, 61 (1981).
[8] P. M. Stevenson, Phys. Rev. D **23**, 2916 (1981).
[9] P. M. Stevenson, Nucl. Phys. **B203**, 472 (1982).
[10] P. M. Stevenson, Nucl. Phys. **B231**, 65 (1984).
[11] S. J. Brodsky, G. P. Lepage, and P. B. Mackenzie, Phys. Rev. D **28**, 228 (1983).
[12] C. F. Berger, Z. Bern, L. Dixon, F. Cordero, D. Forde, T. Gleisberg, H. Ita, D. Kosower, and D. Maître, Phys. Rev. D **80**, 074036 (2009).
[13] D. Maitre *et al.*, Proc. Sci., E PS-HEP (2009) 367.
[14] S. J. Brodsky and X. G. Wu, Phys. Rev. D **85**, 114040 (2012).
[15] M. Gell-Mann and F. E. Low, Phys. Rev. **95**, 1300 (1954).
[16] J. Beringer *et al.* (Particle Data Group), Phys. Rev. D **86**, 010001 (2012).
[17] J. C. Ward, Phys. Rev. **78**, 182 (1950); Y. Takahashi, Nuovo Cimento **6**, 371 (1957).
[18] C. G. Callan, Phys. Rev. D **2**, 1541 (1970).
[19] K. Symanzik, Commun. Math. Phys. **18**, 227 (1970).
[20] S. J. Brodsky and H. J. Lu, Phys. Rev. D **51**, 3652 (1995).
[21] A. L. Kataev and S. A. Larin, Pis'ma Zh. Eksp. Teor. Fiz. **96**, 64 (2012); arXiv:1205.2810.
[22] P. A. Baikov, K. G. Chetyrkin, J. H. Kuhn, and J. Rittinger, J. High Energy Phys. 07 (2012) 017.
[23] P. A. Baikov, K. G. Chetyrkin, J. H. Kuhn, and C. Sturm, arXiv:1207.2199.
[24] M. Beneke, Phys. Rep. **317**, 1 (1999).
[25] S. J. Brodsky and X. G. Wu, Phys. Rev. D **86**, 014021 (2012).
[26] U. Langenfeld, S. Moch, and P. Uwer, Phys. Rev. D **80**, 054009 (2009).
[27] M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer, and M. Wiedermann, Comput. Phys. Commun. **182**, 1034 (2011).
[28] J. Blumlein and W. L. van Neerven, Phys. Lett. B **450**, 417 (1999).
[29] S. J. Brodsky and L. D. Giustino, arXiv:1107.0338.
[30] S. J. Brodsky and X. G. Wu, Phys. Rev. D **85**, 034038 (2012).
[31] S. J. Brodsky and X. G. Wu, Phys. Rev. Lett. **109**, 042002 (2012).
[32] W. A. Bardeen, A. J. Buras, D. W. Duke, and T. Muta, Phys. Rev. D **18**, 3998 (1978).
[33] N. V. Krasnikov, Nucl. Phys. **B192**, 497 (1981).
[34] H. J. Lu and C. A. R. Sa de Melo, Phys. Lett. B **273**, 260 (1991).
[35] H. J. Lu, Phys. Rev. D **45**, 1217 (1992).
[36] G. Grunberg and A. L. Kataev, Phys. Lett. B **279**, 352 (1992).
[37] S. V. Mikhailov, J. High Energy Phys. 06 (2007) 009.
[38] A. L. Kataev and S. V. Mikhailov, Teor. Mat. Fiz. **170**, 174 (2012).
[39] B. L. Ioffe, V. S. Fadin, and L. N. Lipatov, *Quantum Chromodynamics Perturbative and Nonperturbative Aspects* (Cambridge University, Cambridge, England, 2010).
[40] S. J. Brodsky and H. J. Lu, arXiv:hep-ph/9211308; S. J. Brodsky, Report No. SLAC-PUB-6304.
[41] E. C. G. Stuckelberg and A. Peterman, Helv. Phys. Acta **26**, 499 (1953).
[42] N. N. Bogoliubov and D. V. Shirkov, Dokl. Akad. Nauk SSSR **103**, 391 (1955); *Introduction to the Theory of Quantized Fields* (Interscience, New York, 1959).
[43] A. Peterman, Phys. Rep. **53**, 157 (1979).
[44] H. J. Lu and S. J. Brodsky, Phys. Rev. D **48**, 3310 (1993).
[45] G. 't Hooft, Nucl. Phys. **B61**, 455 (1973); S. Weinberg, Phys. Rev. D **8**, 3497 (1973).
[46] H. Georgi and H. D. Politzer, Phys. Rev. D **14**, 1829 (1976).
[47] A. L. Kataev and M. D. Vardiashvili, Phys. Lett. B **221**, 377 (1989); **241**, 644(E) (1990).
[48] S. J. Brodsky, M. S. Gill, M. Melles, and J. Rathsman, Phys. Rev. D **58**, 116006 (1998).
[49] K. G. Chetyrkin, B. A. Kniehl, and M. Steinhauser, Phys. Rev. Lett. **79**, 2184 (1997).
[50] A. Dhar and V. Gupta, Phys. Rev. D **29**, 2822 (1984).
[51] V. Gupta, D. V. Shirkov, and O. V. Tarasov, Int. J. Mod. Phys. A **06**, 3381 (1991).
[52] S. J. Brodsky and P. Huet, Phys. Lett. B **417**, 145 (1998).
[53] A. L. Kataev, Phys. Lett. B **691**, 82 (2010).
[54] S. J. Brodsky, V. S. Fadin, V. T. Kim, L. N. Lipatov, and G. B. Pivovarov, JETP Lett. **70**, 155 (1999).
[55] S. J. Brodsky, V. S. Fadin, V. T. Kim, L. N. Lipatov, and G. B. Pivovarov, JETP Lett. **76**, 249 (2002).
[56] G. P. Lepage and P. B. Mackenzie, Phys. Rev. D **48**, 2250 (1993).
[57] A. I. Vainshtein and V. I. Zakharov, Phys. Rev. Lett. **73**, 1207 (1994); **75**, 3588(E) (1995).
[58] R. J. Crewther, Phys. Rev. Lett. **28**, 1421 (1972).
[59] D. J. Broadhurst and A. L. Kataev, Phys. Lett. B **315**, 179 (1993).
[60] S. J. Brodsky, G. T. Gabadadze, A. L. Kataev, and H. J. Lu, Phys. Lett. B **372**, 133 (1996).
[61] P. A. Baikov, K. G. Chetyrkin, and J. H. Kuhn, Phys. Rev. Lett. **104**, 132004 (2010).
[62] G. 't Hooft, in *Proceedings of the International School of Subnuclear Physics, Erice, Italy, 1977*, edited by A. Zichichi, Subnuclear series Vol. 15 (Plenum, New York, 1979), p. 943.
[63] A. C. Mattingly and P. M. Stevenson, Phys. Rev. D **49**, 437 (1994).
[64] G. Kramer and B. Lampe, Z. Phys. C **39**, 101 (1988).
[65] M. Neubert, Phys. Rev. D **51**, 5924 (1995); C. N. Lovett-Turner and C. J. Maxwell, Nucl. Phys. B **432**, 147 (1994); **452**, 188 (1995); A. G. Grozin, arXiv:hep-ph/0311050.
[66] S. J. Brodsky, A. H. Hoang, J. H. Kühn, and T. Teubner, Phys. Lett. B **359**, 355 (1995).
[67] S. J. Brodsky and G. F. de Teramond, Proc. Sci., QCD-TNT-II (2011) 008.
[68] G. F. de Teramond and S. J. Brodsky, arXiv:1203.4025.
[69] X. G. Wu and S. J. Brodsky, Prog. Part. Nucl. Phys. (to be published).