THE EUROPEAN
PHYSICAL JOURNAL C

Extending the predictive power of perturbative QCD using the principle of maximum conformality and the Bayesian analysis

Jian-Ming Shen[1], **Zhi-Jian Zhou**[1], **Sheng-Quan Wang**[2], **Jiang Yan**[3], **Zhi-Fei Wu**[3], **Xing-Gang Wu**[3,a],
Stanley J. Brodsky[4,b]

[1] School of Physics and Electronics, Hunan Provincial Key Laboratory of High-Energy Scale Physics and Applications, Hunan University, Changsha 410082, People's Republic of China
[2] Department of Physics, Guizhou Minzu University, Guiyang 550025, People's Republic of China
[3] Department of Physics, Chongqing Key Laboratory for Strongly Coupled Physics, Chongqing University, Chongqing 401331, People's Republic of China
[4] SLAC National Accelerator Laboratory, Stanford University, Stanford, CA 94309, USA

Abstract In addition to the evaluation of high-order loop contributions, the precision and predictive power of perturbative QCD (pQCD) predictions depends on two important issues: (1) how to achieve a reliable, convergent fixed-order series, and (2) how to reliably estimate the contributions of unknown higher-order terms. The recursive use of renormalization group equation, together with the Principle of Maximum Conformality (PMC), eliminates the renormalization scheme-and-scale ambiguities of the conventional pQCD series. The result is a conformal, scale-invariant series of finite order which also satisfies all of the principles of the renormalization group. In this paper we propose a novel Bayesian-based approach to estimate the size of the unknown higher order contributions based on an optimized analysis of probability distributions. We show that by using the PMC conformal series, in combination with the Bayesian analysis, one can consistently achieve high degree of reliability estimates for the unknown high order terms. Thus the predictive power of pQCD can be greatly improved. We illustrate this procedure for two pQCD observables: $R_{e^+e^-}$ and R_τ, which are each known up to four loops in pQCD. Numerical analyses confirm that by using the scale-independent and more convergent PMC conformal series, one can achieve reliable Bayesian probability estimates for the unknown higher-order contributions.

1 Introduction

Quantum chromodynamics (QCD) is the fundamental non-Abelian gauge theory of the strong interactions. Because of its property of asymptotic freedom, the QCD couplings between quarks and gluons become weak at short distances, allowing systematic perturbative calculations of physical observables involving large momentum transfer [1,2]. A physical observable must satisfy "renormalization group invariance" (RGI) [3–7]; i.e., the infinite-order perturbative QCD (pQCD) approximant of a physical observable must be independent of artificially introduced parameters, such as the choice of the renormalization scheme or the renormalization scale μ_r. A fixed-order pQCD prediction can violate RGI due to the mismatching of the scale of the perturbative coefficients with the corresponding scale of the strong coupling at each order. For example, invalid scheme-dependent predictions can be caused by an incorrect criteria for setting the renormalization scale; e.g., by simply choosing the scale to eliminate large logarithmic contributions. The error caused by the incorrect choice of the renormalization scale can be reduced to a certain degree by including enough higher-order terms and by the mutual cancellation of contributions from different orders. However, the complexity of high-order loop calculations in pQCD makes the available perturbative series terminate at a finite order, and thus the sought-after cancellations among different orders can fail. Clearly, as the precision of the experimental measurements is increased, it becomes critically important to eliminate theoretical uncertainties from the renormalization scale and scheme ambigui-

ties and to also obtain reliable estimates of the contributions from unknown higher-order (UHO) terms.

Recall that in the case of high precision calculations in quantum electrodynamics (QED), the renormalization scale is chosen to sum all vacuum polarization contributions, e.g.

$$\alpha(q^2) = \frac{\alpha(q_0^2)}{1 - \Pi(q^2, q_0^2)}, \qquad (1)$$

where $\Pi(q^2, q_0^2) = (\Pi(q^2, 0) - \Pi(q_0^2, 0))/(1 - \Pi(q_0^2, 0))$ sums all vacuum polarization contributions, both proper and improper, into the dressed photon propagator. This is the standard Gell Mann-Low renormalization scale-setting for perturbative QED series [4].

Similarly, in non-Abelian QCD, the *Principle of Maximum Conformality* (PMC) [8–12] provides a rigorous method for obtaining a correct fixed-order pQCD series consistent with the principles of renormalization group [13–15]. The evolution of the running QCD coupling is governed by the renormalization group equation (RGE),

$$\beta(\alpha_s) = \mu_r^2 \frac{d\alpha_s(\mu_r)}{d\mu_r^2} = -\alpha_s^2(\mu_r) \sum_{i=0}^{\infty} \beta_i \alpha_s^i(\mu_r). \qquad (2)$$

where the $\{\beta_i\}$-functions are now known up to five-loop level in the $\overline{\text{MS}}$-scheme [16–24]. As in QCD, all β terms are summed into the running coupling by the PMC. After PMC scale setting, the resulting pQCD series is then identical to the corresponding conformal theory with $\beta = 0$. The PMC thus fixes the renormalization scale consistent with the RGE. It extends the Brodsky-Lepage-Mackenzie method [25] for scale-setting in pQCD to all orders, and it reduces analytically to the standard scale-setting procedure of Gell-Mann and Low in the QED Abelian limit (small number of colors, $N_C \rightarrow 0$ [26]). The resulting relations between the predictions for different physical observables, called *commensurate scale relations* [27,28], ensure that the PMC predictions are independent of the theorist's choice of the renormalization scheme. The PMC thus eliminates both renormalization scale and scheme ambiguities. As an important byproduct, because the RG-involved factorially divergent renormalon-like terms such as $n!\beta_0^n \alpha_s^n$ [29–31] are eliminated, the convergence of the PMC perturbative series is automatically improved. In contrast, if one guesses the renormalization scale such as choosing it to match the factorization scale, one will obtain incorrect, scheme-dependent, factorially divergent results for the pQCD approximant, as well as violating the analytic $N_C \rightarrow 0$ Abelian limit. Such an *ad hoc* procedure will also contradict the unification of the electroweak and strong interactions in a grand unified theory.

In practice, it has been conventional to take μ_r as the typical momentum flow (Q) of the process in order to obtain the central value of the pQCD series and to then vary μ_r within a certain range, such as $[Q/2, 2Q]$, as a measure of a combined effect of scale uncertainties and the contributions from uncalculated higher-order (UHO) terms. The shortcomings of this *ad hoc* treatment are apparent: each term in the perturbative series is scale-dependent, and thus the prediction will not satisfy the requirement of RGI. Furthermore, an estimate of the UHO contributions cannot be characterized in a statistically meaningful way; one can only obtain information for the β-dependent terms in the uncalculated higher-order terms which control the running of α_s, and there are no constraints on the contribution from the higher-order conformal β-independent terms.

Since the exact pQCD result is unknown, it would be helpful to quantify the UHO's contribution in terms of a probability distribution. The *Bayesian analysis* is a powerful method to construct probability distributions in which Bayes' theorem is used to iteratively update the probability as new information becomes available. In this article, we will show how one can apply the Bayesian analysis to predict the uncertainty of the UHO contributions as a weighted probability distribution. This idea was pioneered by Cacciari and Houdeau [32], and has been developed more recently in Refs. [33–35]. As illustrations of the power of this method, we will apply the Bayesian analysis to estimate the UHO contributions to several hadronic QCD observables.

Previous applications of Bayesian-based approach have been based on highly scale-dependent pQCD series. As discussed above, it clearly is important to instead use a renormalization scale-invariant series as the basis in order to show the predictive power of the Bayesian-based approach. In the paper, we shall adopt the PMC scale-invariant conformal series as the starting point for estimating the magnitude of unknown higher order contributions using the Bayesian-based approach.

The remaining parts of the paper are organized as follows: In Sect. 2, we show how the Bayesian analysis can be applied to estimate the contributions of the UHO terms. In Sect. 3, we will give a mini-review of how high precision predictions can be achieved by using the PMC single scale-setting approach (PMCs). In Sect. 4, we will apply the PMCs and the Bayesian-based approach to give predictions with constrained high order uncertainties for two observables, $R_{e^+e^-}$ and R_τ, Sect. 5 is reserved for a summary. For convenience and as a useful reference, we provide a general introduction to probability and the Bayesian analysis [36], together with a useful glossary of the terminology in the Appendix.

2 The Bayesian-based approach

In this section we will show how one can apply a Bayesian-based approach in order to give a realistic estimate of the size

of the unknown higher order pQCD contributions to predictions for physical observables. We shall show that by using the PMC conformal series, in combination with the Bayesian analysis of probability distributions, one can consistently achieve a high degree of reliability estimate for the UHO-terms. Thus the predictive power of pQCD can be greatly improved.

We will explain the Bayesian-based approach by applying it to the series of a perturbatively calculable physical observable (ρ). If the perturbative approximant of the physical observable starts at order $\mathcal{O}(\alpha_s^l)$ and stops at the k_{th} order $\mathcal{O}(\alpha_s^k)$, one has

$$\rho_k = \sum_{i=l}^{k} c_i \alpha_s^i, \tag{3}$$

which represents the partial sum consisting of the first several terms in the perturbative expansion. c_i are the coefficients of the perturbative expansion.

For conventional pQCD series, the limit, $k \to \infty$, does not exist, as perturbative expansions are divergent [37,38]. The typical divergent contributions are renormalons (see e.g. Ref. [29]). The divergent nature of the pQCD series is related to the fact that ρ is a non-analytic function of the coupling α_s in $\alpha_s = 0$. The conventional pQCD series are believed to be asymptotic expansions of the physical observable.

The asymptotic nature of the divergent perturbative expansion implies that up to some order N adding terms to the expansion improves the accuracy of the prediction, but beyond N the divergent contributions to the series dominate and the sum explodes. The truncated expansion at order N gives the optimally accurate approximation available for the observable (ρ), and represents the *optimal truncation* of the asymptotic series.

With these statements, following we will give a brief introduction on how to apply the Bayesian-based approach to a fixed-order pQCD series, and estimate the size of its unknown higher orders in terms of the properties of a probability distribution. Other applications and developments of the Bayesian-based method can be found in Refs. [32–35].

2.1 Basic definitions and assumptions

Consider a generic measure of "credibility", applicable to any possible perturbative series such as Eq. (3), varied over the space of a set of *priori* unknown perturbative coefficients $c_l, c_{l+1}, \ldots$. These coefficients are regarded as random variables in Bayesian statistics. One can define a *probability density function* (p.d.f.), $f(c_l, c_{l+1}, \ldots)$, which satisfies the following normalization condition

$$\int f(c_l, c_{l+1}, \ldots) \, \mathrm{d}c_l \, \mathrm{d}c_{l+1} \cdots = 1, \tag{4}$$

and the parameters can be marginalized according to

$$f(c_l, \ldots, c_k) = \int f(c_l, \ldots, c_k, c_{k+1}, \ldots) \, \mathrm{d}c_{k+1} \ldots . \tag{5}$$

If not specified, here and following, the ranges of integration for the variables are all from $-\infty$ to $+\infty$. The conditional p.d.f. of a generic (uncalculated) coefficient c_n with given coefficients $c_l, \ldots, c_k$, is then by definition,

$$f_c(c_n|c_l, \ldots, c_k) = \frac{f(c_l, \ldots, c_k, c_n)}{f(c_l, \ldots, c_k)}, \quad (n > k). \tag{6}$$

A key point of the Bayesian-based approach is to make the reasonable assumption that all the coefficients c_i ($i = l, l+1, \ldots$) are finite and bounded by the absolute value of a common number $\bar{c}$ ($\bar{c} > 0$) [32], namely

$$|c_i| \le \bar{c}, \quad \forall \, i. \tag{7}$$

If none of the coefficients have been calculated, one can only say that $\bar{c}$ is a positive real number where its order of magnitude is *priori* unknown. If the first several coefficients such as $c_l, \ldots, c_k$ have been calculated, one may use them to give an estimate of $\bar{c}$, which in turn restricts the possible values for the unknown coefficient c_n ($n > k$). The value of $\bar{c}$ is thus a (hidden) parameter which will disappear (through marginalization) in the final results. The set of uncertain variables that defines the space is thus the set constituted by the parameter $\bar{c}$ and all of the coefficients $c_l, c_{l+1}, \ldots$. Three reasonable hypotheses then follow from the above assumption (7); i.e.

- The order of magnitude of $\bar{c}$ is equally probable for all values. This can be encoded by defining a p.d.f. for $\ln \bar{c}$, denoted by $g(\ln \bar{c})$, as the limit of a flat distribution within the region of $-|\ln \epsilon| \le \ln \bar{c} \le |\ln \epsilon|$, where ϵ is a small parameter tends to 0,

$$g(\ln \bar{c}) = \frac{1}{2|\ln \epsilon|} \theta(|\ln \epsilon| - |\ln \bar{c}|). \tag{8}$$

Equivalently, a p.d.f. for $\bar{c}$, which is denoted by $g_0(\bar{c})$, satisfies

$$g_0(\bar{c}) = \frac{1}{2|\ln \epsilon|} \frac{1}{\bar{c}} \theta \left(\frac{1}{\epsilon} - \bar{c} \right) \theta(\bar{c} - \epsilon), \tag{9}$$

where $\theta(x)$ is the Heaviside step function. In practice we will perform all calculations (both analytical and numerical) with $\epsilon \ne 0$, and take the limit $\epsilon \to 0$ for the final result.

- The conditional p.d.f. of an unknown coefficient c_i given $\bar{c}$, which is denoted by $h_0(c_i|\bar{c})$, is assumed in the form of a uniform distribution, i.e.,

$$h_0(c_i|\bar{c}) = \frac{1}{2\bar{c}} \theta(\bar{c} - |c_i|), \tag{10}$$

which implies that the condition (7) must be strictly satisfied. The p.d.f. $h_0(c_i|\bar{c})$ will act as the *likelihood function* for $\bar{c}$ in later calculations.

- All the coefficients c_i ($i = l, l + 1, \ldots$) are mutually *independent*, with the exception for the common bound, i.e. $|c_i| \leq \bar{c}$, which implies the joint conditional p.d.f., denoted by $h(c_i, c_j|\bar{c})$,

$$h(c_i, c_j|\bar{c}) = h_0(c_i|\bar{c})h_0(c_j|\bar{c}), \quad \forall \ i \neq j. \quad (11)$$

The hypotheses (9), (10) and (11) completely define the *credibility measure* over the whole space of a priori uncertain variables $\{\bar{c}, c_l, c_{l+1}, \ldots\}$. They then define every possible inherited measure on a subspace associated with the pQCD approximate of a physical observable whose first several coefficients are known.

One may question the reasonability of the original assumption (7) due to the fact that the full pQCD series is divergent. However, in practice, of all the unknown higher orders, we shall concentrate on the terms before the optimal truncation. For all the terms before the optimal truncation N, it is reasonable to give a finite common boundary, $\bar{c}$, for their coefficients. For definiteness, we modify the assumption (7) as,

$$|c_i| \leq \bar{c}, \quad \forall \ i \leq N. \quad (12)$$

This modification will not change the above three hypotheses (9, 10, 11).

2.2 Bayesian analysis

In this subsection, we calculate the conditional p.d.f. of a generic (uncalculated) coefficient c_n ($n > k$) with given coefficients $c_l, \ldots, c_k$, denoted as $f_c(c_n|c_l, \ldots, c_k)$, based on the Bayes' theorem.

Schematically, we first reformulate the conditional p.d.f. $f_c(c_n|c_l, \ldots, c_k)$ as,

$$f_c(c_n|c_l, \ldots, c_k) = \int h_0(c_n|\bar{c}) f_{\bar{c}}(\bar{c}|c_l, \ldots, c_k) \mathrm{d}\bar{c}, \quad (13)$$

where $f_{\bar{c}}(\bar{c}|c_l, \ldots, c_k)$ is the conditional p.d.f. of $\bar{c}$ given $c_l, \ldots, c_k$. Applying Bayes' theorem, we have

$$f_{\bar{c}}(\bar{c}|c_l, \ldots, c_k) = \frac{h(c_l, \ldots, c_k|\bar{c})g_0(\bar{c})}{\int h(c_l, \ldots, c_k|\bar{c})g_0(\bar{c})\mathrm{d}\bar{c}}, \quad (14)$$

where $h(c_l, \ldots, c_k|\bar{c}) = \prod_{i=l}^{k} h_0(c_i|\bar{c})$ according to (11) is the *likelihood function* for $\bar{c}$. Inserting the Bayes' formula (14) and the factorization property (11) into (13), and taking the limit $\epsilon \to 0$ for the final result, one obtains

$$f_c(c_n|c_l, \ldots, c_k) = \lim_{\epsilon \to 0} \frac{\int h_0(c_n|\bar{c}) \prod_{i=l}^{k} h_0(c_i|\bar{c})g_0(\bar{c})\mathrm{d}\bar{c}}{\int \prod_{i=l}^{k} h_0(c_i|\bar{c})g_0(\bar{c})\mathrm{d}\bar{c}}$$

$$= \frac{1}{2} \frac{n_c}{n_c + 1} \frac{\bar{c}_{(k)}^{n_c}}{(\max\{|c_n|, \bar{c}_{(k)}\})^{n_c+1}}$$

$$= \begin{cases} \frac{n_c}{2(n_c+1)\bar{c}_{(k)}}, & |c_n| \leq \bar{c}_{(k)} \\ \frac{n_c \bar{c}_{(k)}^{n_c}}{2(n_c+1)|c_n|^{n_c+1}}, & |c_n| > \bar{c}_{(k)} \end{cases}. \quad (15)$$

where $\bar{c}_{(k)} = \max\{|c_l|, \ldots, |c_k|\}$, and $n_c = k - l + 1$ represents the number of known perturbative coefficients, $c_l, \ldots, c_k$. It is easy to confirm the normalization condition, $\int_{-\infty}^{\infty} f_c(c_n|c_l, \ldots, c_k)\mathrm{d}c_n = 1$. Equation (15) indicates the conditional p.d.f. $f_c(c_n|c_l, \ldots, c_k)$ depends on the entire set of the calculated coefficients via $\bar{c}_{(k)} = \max\{|c_l|, \ldots, |c_k|\}$. The existence of such a probability density distribution within the uncertainty interval represents the main difference with other approaches, such as the conventional scale variation approach, which only gives an interval without a probabilistic interpretation. Equation (15) also shows a symmetric probability distribution for negative and positive c_n, predicts a uniform probability density in the interval $[-\bar{c}_{(k)}, \bar{c}_{(k)}]$ and decreases monotonically from $\bar{c}_{(k)}$ to infinity. The knowledge of probability density $f_c(c_n|c_l, \ldots, c_k)$ allows one to calculate the degree-of-belief (DoB, also called "Bayesian probability" or "subjective probability" or "credibility") that the value of c_n is constrained within some interval. The smallest credible interval (CI) of fixed $p\%$ DoB for c_n ($n > k$) turns out to be centered at zero, and thus we denote it by $[-c_n^{(p)}, c_n^{(p)}]$. It is defined implicitly by

$$p\% = \int_{-c_n^{(p)}}^{c_n^{(p)}} f_c(c_n|c_l, \ldots, c_k) \, \mathrm{d}c_n, \quad (16)$$

and further by using the analytical expression in Eq. (15), we obtain

$$c_n^{(p)} = \begin{cases} \bar{c}_{(k)} \frac{n_c+1}{n_c} p\%, & p\% \leq \frac{n_c}{n_c+1} \\ \bar{c}_{(k)} [(n_c+1)(1-p\%)]^{-\frac{1}{n_c}}, & p\% > \frac{n_c}{n_c+1} \end{cases}. \quad (17)$$

With the help of Eq. (15), one can then derive the conditional p.d.f. for the uncalculated higher order term $\delta_n = c_n \alpha_s^n$, ($n > k$), and the smallest $p\%$-CI for δ_n, namely, $[-c_n^{(p)} \alpha_s^n, c_n^{(p)} \alpha_s^n]$. For the next UHO, i.e. $n = k+1$, the conditional p.d.f. of δ_{k+1} given coefficients $c_l, \ldots, c_k$, denoted by $f_\delta(\delta_{k+1}|c_l, \ldots, c_k)$, reads,

$$f_\delta(\delta_{k+1}|c_l, \ldots, c_k)$$
$$= \left(\frac{n_c}{n_c+1}\right) \frac{1}{2\alpha_s^{k+1}\bar{c}_{(k)}}$$
$$\begin{cases} 1, & |\delta_{k+1}| \leq \alpha_s^{k+1}\bar{c}_{(k)} \\ \left(\frac{\alpha_s^{k+1}\bar{c}_{(k)}}{|\delta_{k+1}|}\right)^{n_c+1}, & |\delta_{k+1}| > \alpha_s^{k+1}\bar{c}_{(k)} \end{cases}, \quad (18)$$

Equation (18) indicates an important characteristic of the posterior distribution: a central plateau with power suppressed tails. The distributions for ρ_{k+1} and δ_{k+1} are the same, up to a trivial shift given by the perturbative result (3). Thus the conditional p.d.f. of ρ_{k+1} for given coefficients $c_l, \ldots, c_k$, denoted by $f_\rho(\rho_{k+1}|c_l, \ldots, c_k)$, can be obtain directly,

$$
\begin{aligned}
& f_\rho(\rho_{k+1}|c_l, \ldots, c_k) \\
& = \left(\frac{n_c}{n_c + 1}\right) \frac{1}{2\alpha_s^{k+1}\bar{c}_{(k)}} \\
& \begin{cases} 1, & |\rho_{k+1} - \rho_k| \leq \alpha_s^{k+1}\bar{c}_{(k)} \\ \left(\frac{\alpha_s^{k+1}\bar{c}_{(k)}}{|\rho_{k+1} - \rho_k|}\right)^{n_c+1}, & |\rho_{k+1} - \rho_k| > \alpha_s^{k+1}\bar{c}_{(k)} \end{cases}.
\end{aligned} \quad (19)
$$

We can also estimate more UHOs of the perturbative series (3), e.g. the sum from the next UHO to the optimal truncation, $\Delta_k = \sum_{i=k+1}^N c_i \alpha_s^i$. The detail p.d.f. formulas of Δ_k are given in the appendix. In this work we shall concentrate on estimating the next UHO, c_{k+1}, for given coefficients $c_l, \ldots, c_k$.

In the case of the conventional pQCD series, where the coefficients $\{c_l, c_{l+1}, \ldots, c_k\}$ are renormalization scale dependent, the smallest CI, e.g. $[-c_n^{(p)}, c_n^{(p)}]$, for the DoB of the coefficient c_n under the fixed probability $p\%$ is also scale dependent. In order to achieve the goal of the Bayesian Optimization suggested by Refs. [39,40], i.e., to achieve the optimal smallest CI for the UHO by using the least possible number of given terms, it is clearly better to use a perturbative series with scale-invariant coefficients; i.e.,

$$
\frac{\partial}{\partial \mu_r^2} c_i = 0, \quad \forall \ i. \quad (20)
$$

For a general pQCD approximant ρ_k, such as Eq. (3), it is easy to confirm that

$$
\mu_r^2 \frac{\partial \rho_k}{\partial \mu_r^2}\bigg|_{c_i} = -\beta(\alpha_s)\frac{\partial}{\partial \alpha_s}\rho_k, \quad (21)
$$

where the subscript c_i means the partial derivative is done with respect to the perturbative coefficients only. It shows that if a perturbative series satisfies $\beta(\alpha_s) = 0$, its coefficients will be scale-invariant. The PMC series satisfies this requirement by definition, and thus is well matched to achieve the goal of Bayesian Optimization. Our numerical results given in the following Sect. 4 shall confirm this point.

2.3 Consistent estimate for the contribution of unknown high order pQCD contributions

One can calculate the expectation value and the standard deviation for c_n, δ_{k+1}, and ρ_{k+1} according to the p.d.f.s (15), (18) and (19), respectively. The expectation value and the standard deviation are the essential parameters. In the following, we shall adopt the determination of ρ_{k+1} as an illustration.

It is conventional to estimate the central value of ρ_{k+1} as its expectation value $E(\rho_{k+1})$ and estimate the theoretical uncertainty of ρ_{k+1} as its standard deviation, σ_{k+1}. The expectation value $E(\rho_{k+1})$ can be related to the expectation value of δ_{k+1}, i.e. $E(\rho_{k+1}) = E(\delta_{k+1}) + \rho_k$. For the present prior distribution, $E(\delta_{k+1}) = 0$, due to the fact that the symmetric probability distribution (18) is centered at zero. To predict the next UHO, δ_{k+1}, of ρ_k consistently, it is useful to define a critical DoB, $p_c\%$, which equals to the least value of $p\%$ that satisfies the following equations,

$$
\rho_{i-1} + c_i^{(p)}\alpha_s^i \geq \rho_i + c_{i+1}^{(p)}\alpha_s^{i+1}, \ (i = l+1, \ldots, k), \quad (22)
$$

$$
\rho_{i-1} - c_i^{(p)}\alpha_s^i \leq \rho_i - c_{i+1}^{(p)}\alpha_s^{i+1}, \ (i = l+1, \ldots, k). \quad (23)
$$

Thus, for any $p \geq p_c$, the error bars determined by the $p\%$-CIs provide consistent estimates for the next UHO, i.e. the smallest $p\%$-CIs ($p \geq p_c$) of ρ_{i+1} predicted from ρ_i are well within the smallest $p\%$-CIs of the one-order lower ρ_i predicted from ρ_{i-1}, ($i = l+1, l+2, \ldots, k$). The value of p_c is nondecreasing when k increases. In practice, in order to obtain a consistent and high DoB estimation, we will adopt the smallest $p_s\%$-CI; i.e.

$$
[E(\rho_{k+1}) - c_{k+1}^{(p_s)}\alpha_s^{k+1}, E(\rho_{k+1}) + c_{k+1}^{(p_s)}\alpha_s^{k+1}], \quad (24)
$$

as the final estimate for ρ_{k+1}, where $p_s = \max\{p_c, p_\sigma\}$. Here $p_\sigma\%$ represents the DoB for the 1σ-interval, and $\rho_{k+1} \in [E(\rho_{k+1}) - \sigma_{k+1}, E(\rho_{k+1}) + \sigma_{k+1}]$.

3 The principle of maximum conformality

The PMC was originally introduced as a multi-scale-setting approach (PMCm) [9–12], in which distinct PMC scales at each order are systematically determined in order to absorb specific categories of $\{\beta_i\}$-terms into the corresponding running coupling α_s at different orders. Since the same type of $\{\beta_i\}$-terms emerge at different orders, the PMC scales at each order can be expressed in perturbative form. The PMCm has two kinds of residual scale dependence due to the unknown perturbative terms [41]; i.e., the last terms of the PMC scales are unknown (*first kind of residual scale dependence*), and the last terms in the pQCD approximant are not fixed since its PMC scale cannot be determined (*second kind of residual scale dependence*). Detailed discussions of the residual scale dependence can be found in the reviews [42,43]. The PMC single-scale-setting approach (PMCs) [44] has been recently suggested in order to suppress the residual scale dependence and to make the scale-setting procedures much simpler. The PMCs procedure determines a single overall

effective α_s with the help of RGE; the resulting PMC renormalization scale represents the overall effective momentum flow of the process. The PMCs is equivalent to PMCm in the sense of perturbative theory, and the PMCs prediction is also free of renormalization scale-and-scheme ambiguities up to any fixed order [45]. The PMCs is also equivalent to the very recently suggested single-scale-setting method [46], which follows the idea of "Intrinsic Conformality" [47]. By using the PMCs, the *first kind of residual scale dependence* will be greatly suppressed due to its α_s-power suppression and the exponential suppression; the overall PMC scale has the same precision for all orders, and thus the *second kind of residual scale dependence* is exactly removed. Moreover, due to the independence on the renormalization scheme and scale, the resulting conformal series with an overall single value of $\alpha_s(Q_*)$ provides not only precise pQCD predictions for the known fixed order, but also a reliable basis for estimating the contributions from the unknown higher-order terms.

Within the framework of the pQCD, the perturbative approximant for physical observable ϱ can be written in the following form:

$$\varrho_n = \sum_{i=1}^{n} r_i(\mu_r^2/Q^2)\alpha_s^{p+i-1}(\mu_r), \tag{25}$$

where Q represents the kinematic scale and the index p ($p \geq 1$) indicates the α_s-order of the leading-order (LO) contribution. For the perturbative series (25), its perturbative coefficients r_i can be divided into the conformal parts ($r_{i,0}$) and non-conformal parts (proportional to β_i), i.e. $r_i = r_{i,0} + \mathcal{O}(\{\beta_i\})$. The $\{\beta_i\}$-pattern at different orders exist a special degeneracy [9,10,48], i.e.

$$r_1 = r_{1,0},$$
$$r_2 = r_{2,0} + p\beta_0 r_{2,1},$$
$$r_3 = r_{3,0} + p\beta_1 r_{2,1} + (p+1)\beta_0 r_{3,1} + \frac{p(p+1)}{2}\beta_0^2 r_{3,2},$$
$$r_4 = r_{4,0} + p\beta_2 r_{2,1} + (p+1)\beta_1 r_{3,1} + \frac{p(3+2p)}{2}\beta_1\beta_0 r_{3,2}$$
$$+ (p+2)\beta_0 r_{4,1} + \frac{(p+1)(p+2)}{2}\beta_0^2 r_{4,2}$$
$$+ \frac{p(p+1)(p+2)}{3!}\beta_0^3 r_{4,3},$$
$$\cdots$$

The coefficients $r_{i,j}$ are general functions of the renormalization scale μ_r, which can be redefined as

$$r_{i,j} = \sum_{k=0}^{j} C_j^k \hat{r}_{i-k,j-k} \ln^k(\mu_r^2/Q^2), \tag{26}$$

where the reduced coefficients $\hat{r}_{i,j} = r_{i,j}|_{\mu_r=Q}$, and the combination coefficients $C_j^k = j!/(k!(j-k)!)$.

Following the standard PMCs procedures [44], the overall effective scale can be determined by requiring all the non-conformal $\{\beta_i\}$-terms to vanish; the pQCD approximant (25) then changes to the following conformal series,

$$\varrho_n|_{\text{PMCs}} = \sum_{i=1}^{n} \hat{r}_{i,0}\alpha_s^{p+i-1}(Q_*), \tag{27}$$

where the PMC scale Q_* can be fixed up to N^{2}LL-accuracy for $n = 4$, i.e. $\ln Q_*^2/Q^2$ can be expanded as a power series over $\alpha_s(Q)$,

$$\ln \frac{Q_*^2}{Q^2} = T_0 + T_1\alpha_s(Q) + T_2\alpha_s^2(Q) + \mathcal{O}(\alpha_s^3), \tag{28}$$

where the coefficients T_i ($i = 0, 1, 2$) are all functions of the reduced coefficients $\hat{r}_{i,j}$, whose expressions can be found in Ref. [44]. Equation (28) shows that the PMC scale Q_* is also represented as power series in α_s, which resums all the known $\{\beta_i\}$-terms, and is explicitly independent of μ_r at any fixed order. It represents the physical momentum flow of the process and determines an overall effective value of α_s. Together with the μ_r-independent conformal coefficients, the resulting pQCD series is exactly scheme-and-scale independent [45], thus providing a reliable basis for estimating the contributions of the unknown terms.

4 Numerical results

In this section, we apply the PMCs approach to scale setting in combination with the Bayesian method for estimating uncertainties from the uncalculated higher order terms, for two physical observables $R_{e^+e^-}$ and R_τ, all of which are now known up to four loops in pQCD. We will show how the magnitude of the "unknown" terms predicted by the Bayesian-based approach varies as more-and-more loop terms are determined.

The ratio $R_{e^+e^-}$ for e^+e^- annihilation is defined as

$$R_{e^+e^-}(Q) = \frac{\sigma\left(e^+e^- \to \text{hadrons}\right)}{\sigma\left(e^+e^- \to \mu^+\mu^-\right)}$$
$$= 3\sum_q e_q^2\left[1 + R(Q)\right], \tag{29}$$

where $Q = \sqrt{s}$ is the e^+e^- center-pf-mass collision energy at which the ratio is measured. The pQCD approximant of $R(Q)$, denoted by $R_n(Q)$, reads, $R_n(Q) = \sum_{i=1}^{n} r_i(\mu_r/Q)\alpha_s^i(\mu_r)$. The pQCD coefficients at $\mu_r = Q$ have been calculated in the $\overline{\text{MS}}$-scheme in Refs. [49–52]. The coefficients at any other scales can then be obtained via RGE evolution. For illustration, we shall take $Q \equiv 31.6$ GeV [53] throughout this paper to illustrate the numerical predictions.

The ratio R_τ for hadronic τ decays is defined as

$$R_\tau(M_\tau) = \frac{\sigma(\tau \to \nu_\tau + \text{hadrons})}{\sigma(\tau \to \nu_\tau + \bar{\nu}_e + e^-)}$$

Table 1 The known coefficients for $R_n(Q)$. The coefficients of conventional scale setting, $r_i(\mu_r)$, are for $\mu_r = Q$. The conformal coefficients, $r_{i,0}$, are scale-independent

	i=1	i=2	i=3	i=4
$r_i(\mu_r = Q)$	0.3183	0.1428	–0.4130	0.8257
$r_{i,0}$	0.3183	0.1865	–0.0324	–0.1128

Table 2 The known coefficients for $\tilde{R}_n(M_\tau)$. The coefficients of conventional scale setting, $r_i(\mu_r)$, are for $\mu_r = M_\tau$. The conformal coefficients, $r_{i,0}$, are scale-independent

	$i = 1$	$i = 2$	$i = 3$	$i = 4$
$r_i(\mu_r = M_\tau)$	0.3183	0.5271	0.8503	1.3046
$r_{i,0}$	0.3183	0.2174	0.1108	0.0698

$$= 3 \sum |V_{ff'}|^2 \left(1 + \tilde{R}(M_\tau)\right), \tag{30}$$

where $V_{ff'}$ are Cabbibo-Kobayashi-Maskawa matrix elements, $\sum |V_{ff'}|^2 = |V_{ud}|^2 + |V_{us}|^2 \approx 1$ and $M_\tau = 1.77686$ GeV [36]. The pQCD approximant of $\tilde{R}(M_\tau)$, denoted by $\tilde{R}_n(M_\tau)$, reads, $\tilde{R}_n(M_\tau) = \sum_{i=1}^{n} r_i(\mu_r/M_\tau)\alpha_s^i(\mu_r)$; the coefficients can be obtained using the known relation of $R_\tau(M_\tau)$ to $R_{e^+e^-}(Q)$ [54].

In order to do the numerical evaluation, the RunDec program [55,56] is adopted to calculate the value of α_s. For self-consistency, the four-loop α_s-running behavior will be used. The world average $\alpha_s(M_z) = 0.1179 \pm 0.0009$ [36] is adopted as a reference.

4.1 Single-scale PMCs predictions

After applying the PMCs approach, the overall renormalization scale for each process can be determined. If the pQCD approximants are known up to two-loop, three-loop, and four-loop level, respectively, the corresponding overall scales are

$$Q_*|_{e^+e^-} = \{35.36, 39.67, 40.28\} \text{ GeV}, \tag{31}$$
$$Q_*|_\tau = \{0.90, 1.01, 1.05\} \text{ GeV}. \tag{32}$$

The PMC scales Q_* are independent of the initial choice of the renormalization scale μ_r. In the case of the leading-order ratios with $n = 1$, one has no information to set the effective scale, and thus for definiteness, we will set it to be Q, or M_τ, respectively, which gives $R_1 = 0.04428$, and $\tilde{R}_1 = 0.0891$.

We present the first four conformal coefficients $r_{i,0}$ ($i = 1, 2, 3, 4$) in Tables 1, and 2, in which the conventional coefficients r_i ($i = 1, 2, 3, 4$) at a specified scale are also presented in comparison. Because the coefficients r_i ($i \geq 2$) of the conventional pQCD series are scale-dependent at every orders, the Bayesian-based approach can only be applied after one specifies the choices for the renormalization scale, thus introducing extra uncertainties for the Bayesian-based approach.

Table 3 The predicted smallest 95.5% credible intervals (CI) for the scale-dependent conventional coefficients $r_i(\mu_r)$ ($i = 3, 4, 5$) at the scale $\mu_r = Q$ and the scale-invariant coefficients $r_{i,0}(i = 3, 4, 5)$ of $R_n(Q = 31.6 \text{ GeV})$ via the Bayesian-based approach. The exact values ("EC") are presented as comparisons

	$r_3(\mu_r = Q)$	$r_4(\mu_r = Q)$	$r_5(\mu_r = Q)$
CI	[−0.8663, 0.8663]	[−0.7314, 0.7314]	[−1.1989, 1.1989]
EC	−0.4130	0.8257	–

	$r_{3,0}$	$r_{4,0}$	$r_{5,0}$
CI	[−0.8663, 0.8663]	[−0.5638, 0.5638]	[−0.4622, 0.4622]
EC	−0.0324	−0.1128	–

Table 4 The predicted smallest 95.5% credible intervals (CI) for the scale-dependent conventional coefficients $r_i(\mu_r)$ ($i = 3, 4, 5$) at the scale $\mu_r = M_\tau$ and the scale-invariant coefficients $r_{i,0}(i = 3, 4, 5)$ of $\tilde{R}_n(M_\tau)$ via the Bayesian-based approach. The exact values ("EC") are presented for comparison

	$r_3(\mu_r = M_\tau)$	$r_4(\mu_r = M_\tau)$	$r_5(\mu_r = M_\tau)$
CI	[−1.4346, 1.4346]	[−1.5060, 1.5060]	[−1.8942, 1.8942]
EC	0.8503	1.3046	–

	$r_{3,0}$	$r_{4,0}$	$r_{5,0}$
CI	[−0.8663, 0.8663]	[−0.5638, 0.5638]	[−0.4622, 0.4622]
EC	0.1108	0.0698	–

On the other hand, the PMCs series is a scale-independent conformal series in powers of the effective coupling $\alpha_s(Q_*)$; the PMCs thus provides a reliable basis for obtaining constraints on the predictions for the unknown higher-order contributions.

4.2 Estimation of UHOs using the Bayesian-based approach

In this subsection, we give estimates for the UHOs of the pQCD series $R_n(Q = 31.6 \text{ GeV})$ and $\tilde{R}_n(M_\tau)$. More explicitly, we will predict the magnitude of the unknown coefficient c_{i+1} from the known ones $\{c_1, \ldots, c_i\}$ by using the Bayesian-based approach.

First, we present the predicted smallest 95.5% CIs and the exact values [1] ("EC") of the scale-invariant conformal coefficients $c_i = r_{i,0}$ ($i = 3, 4, 5$) of the PMCs series of $R_n(Q = 31.6 \text{ GeV})$ and $\tilde{R}_n(M_\tau)$ in Tables 3 and 4, respectively. For comparison, the similarly predicted scale-dependent conventional coefficients $c_i = r_i(\mu_r)$ ($i = 3, 4, 5$) of the conventional series of $R_n(Q = 31.6 \text{ GeV})$ and $\tilde{R}_n(M_\tau)$ at the specific scales $\mu_r = Q$ and M_τ are also presented. It is noted

---

[1] The "exact" value means that it is obtained by directly using the known perturbative series.

Table 5 The predicted smallest 95.5% credible intervals (CI) for the scale-dependent coefficients $r_3(\mu_r)$ and $r_4(\mu_r)$ of $R_n(Q = 31.6 \text{ GeV})$ by using the Bayesian-based approach at three scales $\mu_r = Q$, $Q/2$ and $2Q$, respectively. The exact values ("EC") are present for comparison

		$r_3(\mu_r)$	$r_4(\mu_r)$
$\mu_r = Q$	CI	$[-0.8663, 0.8663]$	$[-0.7314, 0.7314]$
	EC	-0.4130	-0.8257
$\mu_r = 2Q$	CI	$[-1.1213, 1.1213]$	$[-0.7297, 0.7297]$
	EC	0.1643	-1.0089
$\mu_r = Q/2$	CI	$[-0.8663, 0.8663]$	$[-0.9473, 0.9473]$
	EC	-0.5348	0.4272

that the exact values of $r_{3,0}$ and $r_{4,0}$ lie within the predicted 95.5% CIs. In contrast, for the conventional coefficients, most of the exact values of r_3 and r_4 are lying within the predicted 95.5% CIs; However, there are one exception for r_4, i.e. for $R_n(Q = 31.6 \text{ GeV})$, the exact values of r_4 are outside the region of the 95.5% CIs. These exceptions may be removed by taking a different choice of renormalization scale; e.g., for the case of $R_n(Q = 31.6 \text{ GeV})$, as shown by Table 5, the exact value of r_4 will lie within the predicted 95.5% CI if setting $\mu_r = Q/2$. Table 5 also confirms that the CIs predicted from the conventional series are also scale dependent. Thus, in comparison with the renormalon-divergent and scale-dependent conventional series, it is essential to use the renormalon-free and scale-invariant PMCs series in order to estimate the unknown higher-order coefficients. More explicitly, we present more predicted CIs with three typical DoBs in Figs. 1 and 2, respectively.

Second, we present the probability density distributions for the two observables $R(Q = 31.6 \text{ GeV})$ and $\tilde{R}(M_\tau)$ with different states of knowledge predicted by PMCs and the

Bayesian-based approach in Fig. 3. The four lines in each figure correspond to different degrees of knowledge: given LO (dotted), given NLO (dotdashed), given N^2LO (solid) and given N^3LO (dashed). These figures illustrate the characteristics of the posterior distribution: a symmetric plateau with two suppressed tails. The posterior distribution given by the Bayesian-based approach depends on the prior distribution, and as more and more loop terms become known, the probability is updated with less and less dependence on the prior; i.e., the probability density becomes increasingly concentrated (the plateau becomes narrower and narrower and the tail becomes shorter and shorter) as more and more loop terms for the distribution are determined.

Third, we present the $p_s\%$ CIs of $R_n(Q = 31.6 \text{ GeV})$ with $n = (2, 3, 4, 5)$ predicted from the one-order lower $R_{n-1}(Q = 31.6 \text{ GeV})$ based on the Bayesian-based approach in Fig. 4, where $p_s\% = 95.4\%$ for the scale-independent PMCs series, and $p_s\% = 98.4\%$ for the scale-dependent conventional series. The calculated values ("CV") of the pQCD approximants $R_n(Q = 31.6 \text{ GeV})$ with $n = (1, 2, 3, 4)$ are also presented as a comparison. The triangles and the quadrates are for the PMCs series and the conventional (conv.) scale-dependent series, respectively. Analogous results for $\tilde{R}_n(M_\tau)$ are also given in Fig. 4. Both the center values and the error bars (or CIs) are scale-independent for the PMCs series. The results for conventional series of R_n and $\tilde{R}_n$ are for $\mu_r = Q$ and M_τ, respectively. Figure 4 shows that the error bars (or CIs) predicted by using the Bayesian-based approach quickly approach their steady points for both the PMCs and conventional series. As expected, the error bars provide consistent and high DoB estimates for the UHOs for both the PMCs and conventional series; e.g., the error bars of $R_{n+1}(Q)$ ($n = 2, 3, 4$) predicted from $R_n(Q)$ are well within the error bars of the one-order lower $R_n(Q)$ predicted from
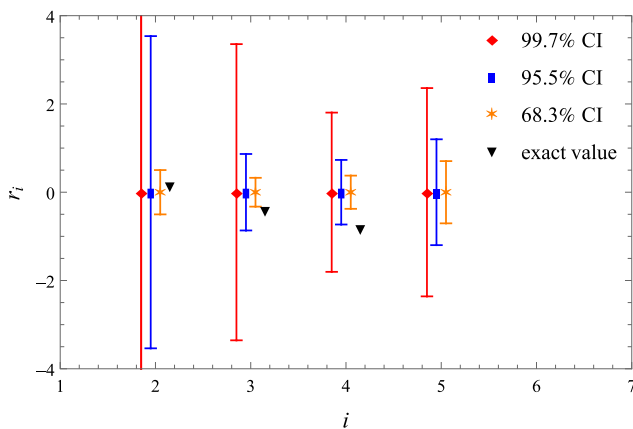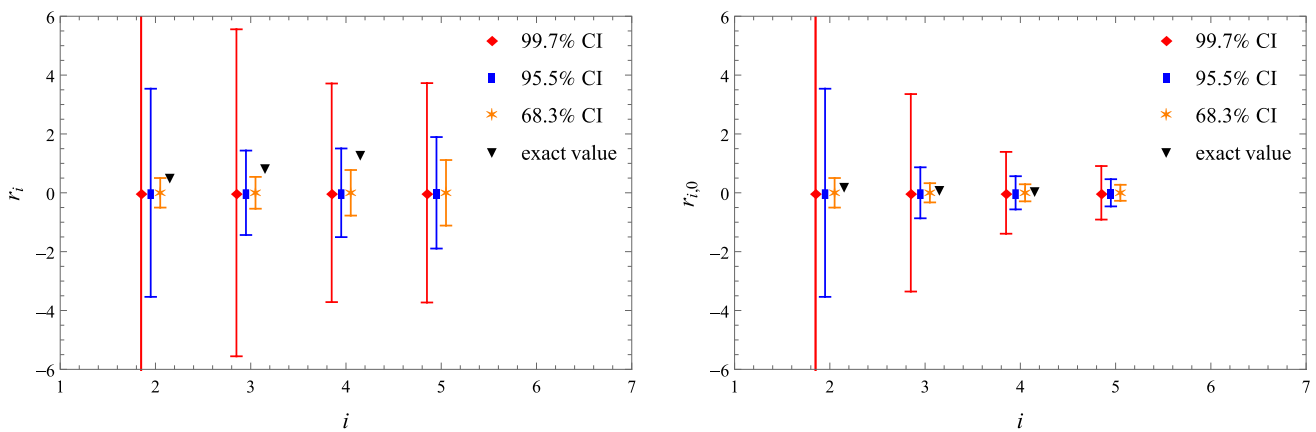

Fig. 1 The predicted credible intervals (CI) with three typical DoBs for the scale-dependent coefficients $r_i(\mu_r)$ at the scale $\mu_r = Q$ and the scale-invariant $r_{i,0}$ of $R_n(Q = 31.6\text{GeV})$ under the Bayesian-based approach, respectively. The red diamonds, the blue rectangles,
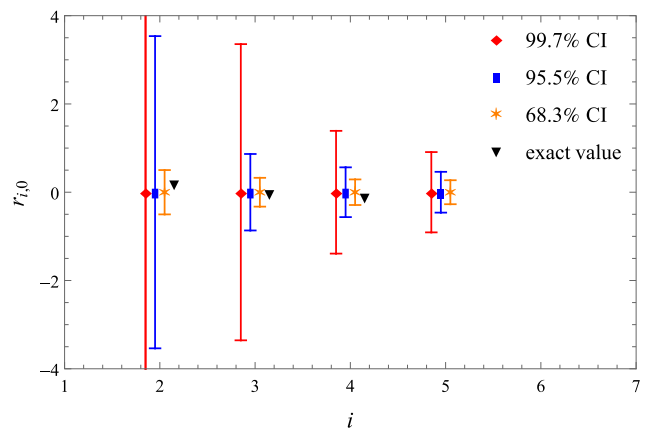
the golden yellow stars and the black inverted triangles together with their error bars, are for 99.7% CI, 95.5% CI, 68.3% CI, and the exact values of the coefficients at different orders, respectively

Fig. 2 The predicted credible intervals (CI) with three typical DoBs for the scale-dependent coefficients $r_i(\mu_r)$ at the scale $\mu_r = M_\tau$ and the scale-invariant $r_{i,0}$ of $\tilde{R}_n(M_\tau)$ under the Bayesian-based approach, respectively. The red diamonds, the blue rectangles, the golden yellow stars and the black inverted triangles together with their error bars, are for 99.7% CI, 95.5% CI, 68.3% CI, and the exact values of the coefficients at different orders, respectively
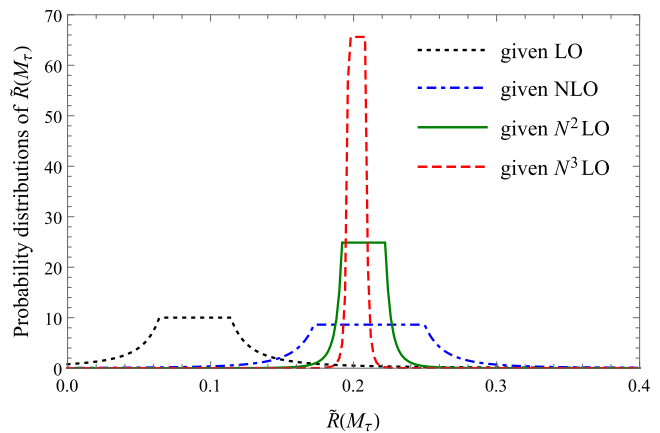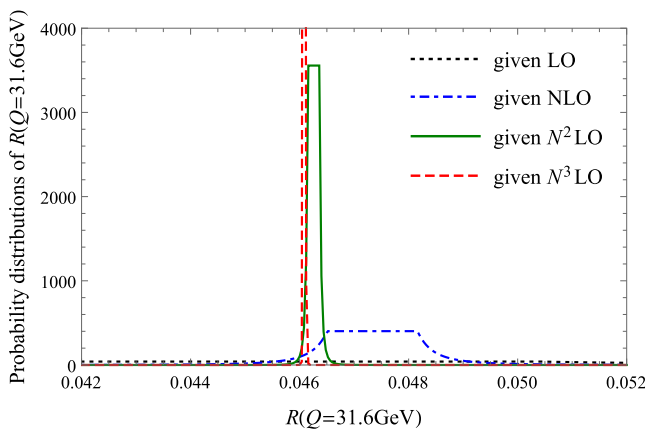
Fig. 3 The probability density distributions of $R(Q = 31.6$ GeV$)$ and $\tilde{R}(M_\tau)$ with different states of knowledge predicted by PMCs and the Bayesian-based approach, respectively. The black dotted, the blue dash-dotted, the green solid and the red dashed lines are results for the given LO, NLO, N^2LO and N^3LO series, respectively

$R_{n-1}(Q)$; The conclusions for $\tilde{R}_n(M_\tau)$ are similar. Detailed numerical results are presented in Table 6, where the 2nd, 4th, and 6th columns show the calculated central values ("CV") of the fixed-order pQCD approximants $R_n(Q = 31.6\text{GeV})$ and $\tilde{R}_n(M_\tau)$ for $n = 2, 3, 4$ respectively, and the 3rd, 5th, and 7th columns show the predicted $p_s\%$ credible intervals ("CI") of those approximants for $n = 3, 4, 5$ respectively. The predicted CIs for $R_2(Q)$ and $\tilde{R}_2(M_\tau)$ are sufficiently conservative and thus are not presented in the table. The DoB $(p_s\%)$ is given in the last column. For the present prior distributions, $p_\sigma\% = 65.3\%$ for $l = 1$ and $k = 4$. Thus the DoB $p_s\%$, given in Table 6, is also the critical DoB, i.e. $p_s = p_c$.

Our final predictions for the five-loop predictions of $R_5(Q)$ and $\tilde{R}_5(M_\tau)$ based on the PMCs and the Bayesian-based approach read,

$$R_5(Q = 31.6\text{GeV}) = 0.04609 \pm 0.00042 \pm 0.00002, \quad (33)$$

$$\tilde{R}_5(M_\tau) = 0.2032^{+0.0092}_{-0.0086} \pm 0.0083, \quad (34)$$

where the first error is for $\Delta\alpha_s(M_Z) = \pm 0.0009$ and the second error represents high DoBs $p_s\%$ which are consistent with the estimates for the UHOs. Note that the very small uncertainty ± 0.00002 for $R_5(Q = 31.6\text{GeV})$ is determined by the 95.4% CI according to the Bayesian-based approach, $[-r^{(95.4)}_{5,0}\alpha_s^5(Q_*), r^{(95.4)}_{5,0}\alpha_s^5(Q_*)]$, where $\alpha_s^5(Q_*) \simeq 0.00004$ and the predicted $r^{(95.4)}_{5,0} = 0.4596$ are all small. Our prediction for hadronic τ decays, $\tilde{R}_5(M_\tau)$, can be compared with those given in Refs. [57–60].

5 Summary

The PMC provides a rigorous first-principle method to eliminate conventional renormalization scheme and scale ambi-
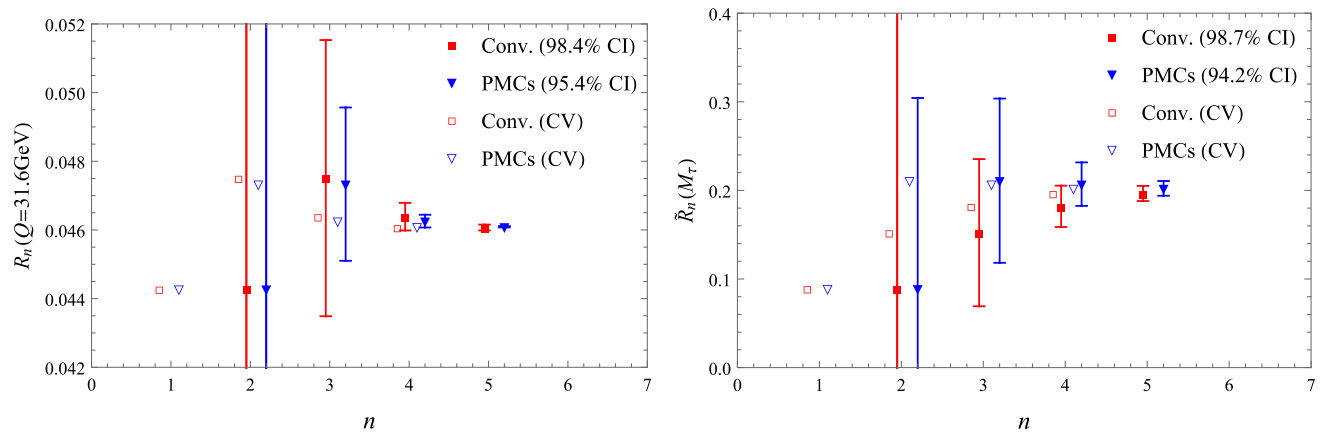
Fig. 4 Comparison of the calculated central values ("CV") of the pQCD approximants $R_n(Q = 31.6$ GeV) and $\tilde{R}_n(M_\tau)$ for $n = (1, 2, 3, 4)$ with the predicted $p_s\%$ credible intervals (CI) of those approximants for $n = (2, 3, 4, 5)$. The blue hollow triangles and red hollow quadrates represent the calculated central values of the fixed-order pQCD predictions using PMCs and conventional (Conv.) scale-setting, respectively. The blue solid triangles and red solid quadrates with error bars represent the predicted $p_s\%$ CIs using the Bayesian-based approach based on the PMCs conformal series ($p_s = 95.4$ for $R_n(Q = 31.6$ GeV), $p_s = 94.2$ for $\tilde{R}_n(M_\tau)$) and the conventional (Conv.) scale-dependent series ($p_s = 98.4$ for $R_n(Q = 31.6$ GeV), $p_s = 98.7$ for $\tilde{R}_n(M_\tau)$), respectively

Table 6 Comparison of the calculated central values ("CV") with the predicted $p_s\%$ credible intervals ("CI") of the pQCD approximants $R_n(Q = 31.6$GeV) and $\tilde{R}_n(M_\tau)$ based on the PMC conformal series and the conventional (conv.) scale-dependent series up to n_{th}-order level, respectively. The DoB of the CI ($p_s\%$) is given in the last column. The results for the PMC conformal series are scale-independent. The results for the conventional scale-dependent series are calculated at $\mu_r = Q$ and M_τ for $R_n(Q = 31.6$GeV) and $\tilde{R}_n(M_\tau)$, respectively

	CV, $n = 2$	CI, $n = 3$	CV, $n = 3$	CI, $n = 4$	CV, $n = 4$	CI, $n = 5$	$p_s\%$
$R_n(Q = 31.6$GeV$)\|_{\text{PMCs}}$	0.04733	[0.04510, 0.04956]	0.04626	[0.04607, 0.04645]	0.04609	[0.04607, 0.04611]	95.4%
$\tilde{R}_n(M_\tau)\|_{\text{PMCs}}$	0.2110	[0.1183, 0.3037]	0.2071	[0.1825, 0.2317]	0.2023	[0.1940, 0.2106]	94.2%
$R_n(Q = 31.6$GeV$)\|_{\text{conv.}}$	0.04751	[0.04349, 0.05153]	0.04639	[0.04599, 0.04679]	0.04607	[0.04598, 0.04616]	98.4%
$\tilde{R}_n(M_\tau)\|_{\text{conv.}}$	0.1523	[0.0692, 0.2354]	0.1820	[0.1586, 0.2054]	0.1965	[0.1879, 0.2051]	98.7%

guities for high-momentum transfer processes in pQCD up to any fixed order. Its predictions have a solid theoretical foundation, satisfying renormalization group invariance and all other self-consistency conditions derived from the renormalization group. The PMCs is a single-scale-setting approach, which determines a single overall effective/correct $\alpha_s(Q_*)$ by using all of the RG-involved nonconformal $\{\beta_i\}$-terms. The resulting PMCs series is a renormalon-free and scale-invariant conformal series; it thus achieves precise fixed-order pQCD predictions and provides a reliable basis for predicting unknown higher-order contributions.

The Bayesian analysis provides a compelling approach for estimating the UHOs from the known fixed-order series by adopting a probabilistic interpretation. The conditional probability of the unknown perturbative coefficient is first given by a subjective prior distribution, which is then updated iteratively according to the Bayes' theorem as more and more information is included. The posterior distribution given by the Bayesian-based approach depends on the subjective prior distribution (or the assumptions), and as more-and-more information updates the probability, less-and-less

dependence on the prior distribution (or the assumptions) can be achieved, as confirmed in Fig. 3.

We have defined an objective measure which characterizes the uncertainty due to the uncalculated higher order (UHO) contributions of a perturbative QCD series using the Bayesian analysis. This uncertainty is given as a credible interval (CI) with a degree of belief (DoB, also called Bayesian probability). The numerical value for the uncertainty, the critical DoB, is given as a percentage $p_c\%$. When $p_c\% = 95\%$, it means that there is a 95% probability that the exact answer is within this range. The CI with DoB $p_s\%$ in Fig. 4 and Table 6 takes into account the uncertainties in the values of the input physics parameters, such as the value of α_s, which will become very small at high order due to the α_s^n-power suppression. Detailed numerical results are presented in Table 6, where the 2nd, 4th, and 6th columns show the calculated central values of the fixed-order pQCD approximants $R_n(Q = 31.6$GeV) and $\tilde{R}_n(M_\tau)$ for $n = 2, 3, 4$, respectively, and the 3rd, 5th, 7th columns show the predicted $p_s\%$ credible interval of those approximants for $n = 3, 4, 5$, respectively. The 8th column shows the DoB ($p_s\%$) of the

credible interval presented in the 3rd, 5th, 7th columns. The calculated p_s value, $p_s = \max\{p_c, p_\sigma\}$, equals p_c since the DoB of the 1σ-interval $p_\sigma\%$ equals 65.3% for the present prior distributions.

In contrast, each term in a conventional perturbative series is highly scale-dependent, thus the Bayesian-based approach can only be applied after one assumes choices for the perturbative scale. What's more, the $n!$ renormalon series leads to divergent behavior especially at high order ;[2] e.g., the exact value of the conventional coefficient r_4 is even outside the 95.5% CI predicted from $\{r_1, r_2, r_3\}$ for $R(Q)$, which can be found in Table 3. Thus, it is critical to use the more convergent and scale-independent PMC conformal series as the basis for estimating the unknown higher-order coefficients.

As we have shown, by using the PMCs approach in combination with the Bayesian analysis, one can obtain highly precise fixed-order pQCD predictions and achieve consistent estimates with high DoB for the unknown higher-order contributions. In the present paper, we have illustrated this procedure for two important hadronic observables, R_{e+e^-} and R_τ, which have been calculated up to four-loops in pQCD. The elimination of the uncertainty in setting the renormalization scale for fixed-order pQCD predictions using the PMCs, together with the reliable estimate for the uncalculated higher-order contributions obtained using the Bayesian analysis, greatly increases the precision of collider tests of the Standard Model and thus the sensitivity to new phenomena.

Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Authors' comment: All the numerical results can be achieved by using the formulas listed in the body of the text, so there is no data to be deposited.]

---

[2] Such renormalon divergence also makes the hidden parameter $\bar{c}$ to be much larger than the PMC one, thus if choosing the same degree-of-belief, the PMC credible interval shall be much smaller.

Appendix A: Theorems and laws for probability distributions

An abstract definition of probability can be given by considering a set S, called the sample space, and possible subsets A, B, ..., the interpretation of which is left open. The *probability P* is a real-valued function defined by the following axioms due to Kolmogorov [61]:

1. For every subset A in S, $P(A) \geq 0$;
2. For disjoint subsets (i.e., $A \cap B = \emptyset$),
   $P(A \cup B) = P(A) + P(B)$;
3. $P(S) = 1$.

In addition, the *conditional probability $P(A|B)$* (read as P of A given B) is defined as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{A1}$$

From this definition and using the fact that $A \cap B$ and $B \cap A$ are the same, one obtains *Bayes' theorem*,

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}. \tag{A2}$$

From the three axioms of probability and the definition of conditional probability, one obtains the *law of total probability*,

$$P(B) = \sum_i P(B|A_i)P(A_i), \tag{A3}$$

for any subset B and for disjoint A_i with $\cup_i A_i = S$. This can be combined with Bayes' theorem (A2) to give

$$P(A|B) = \frac{P(B|A)P(A)}{\sum_i P(B|A_i)P(A_i)}, \tag{A4}$$

where the subset A could, for example, be one of the A_i.

Appendix B: The Bayesian analysis

In Bayesian statistics, the subjective interpretation of probability is used to quantify one's *degree of belief* in a *hypothesis*. The hypothesis is often characterized by one or more parameters. This allows one to define a *probability density function* (p.d.f.) for a parameter, which reflects one's knowledge about where its true value lies.

Consider an experiment whose outcome is characterized by a vector of data $\boldsymbol{x}$. A hypothesis H is a statement about the probability for the data, often written $P(\boldsymbol{x}|H)$. This could, for example, completely define the p.d.f. for the data (a simple hypothesis), or it could specify only the functional form of the p.d.f., with the values of one or more parameters not determined (a composite hypothesis). If the probability $P(\boldsymbol{x}|H)$ for data $\boldsymbol{x}$ is regarded as a function of the hypothesis H, then it

is called the *likelihood* of H, usually written $L(H)$. Consider the hypothesis H is characterized by one or more continuous parameters $\boldsymbol{\theta}$, in which case $L(\boldsymbol{\theta}) = P(\boldsymbol{x}|\boldsymbol{\theta})$ is called the *likelihood function*. Note that the likelihood function itself is not a p.d.f. for $\boldsymbol{\theta}$.

In the Bayesian analysis, inference is based on the posterior p.d.f. $p(\boldsymbol{\theta}|\boldsymbol{x})$, whose integral over any given region gives the degree of belief for $\boldsymbol{\theta}$ to take on values in that region, given the data $\boldsymbol{x}$. This is obtained from Bayes' theorem (A4), which can be written

$$p(\boldsymbol{\theta}|\boldsymbol{x}) = \frac{P(\boldsymbol{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int P(\boldsymbol{x}|\boldsymbol{\theta}')\pi(\boldsymbol{\theta}')d\boldsymbol{\theta}'}, \tag{B1}$$

where $P(\boldsymbol{x}|\boldsymbol{\theta})$ is the likelihood function for $\boldsymbol{\theta}$; i.e., the joint p.d.f. for the data viewed as a function of $\boldsymbol{\theta}$, evaluated with data actually obtained in the experiment. The function $\pi(\boldsymbol{\theta})$ is the prior p.d.f. for $\boldsymbol{\theta}$. Note that the denominator in Eq. (B1) serves to normalize the posterior p.d.f. to unity. The likelihood function, prior, and posterior p.d.f.s all depend on $\boldsymbol{\theta}$, and are related by Bayes' theorem, as usual.

Bayesian statistics does not supply a rule for determining the prior $\pi(\boldsymbol{\theta})$; this reflects the analyst's subjective degree of belief (or state of knowledge) about $\boldsymbol{\theta}$ before the measurement was carried out.

Appendix C: The p.d.f. for more UHOs

The sum from the next UHO to the optimal truncation, $\Delta_k = \sum_{i=k+1}^{N} c_i \alpha_s^i$, depends on the values of the unknown coefficients $c_{k+1}, c_{k+2}, \ldots, c_N$. It's conditional p.d.f. $f_\Delta(\Delta_k|c_l, \ldots, c_k)$ can be written as

$$f_\Delta(\Delta_k|c_l, \ldots, c_k) = \int \left[\delta \left(\Delta_k - \sum_{n=k+1}^{N} \alpha_s^n c_n \right) \right]$$
$$\times f_{rc}(c_{k+1}, c_{k+2}, \ldots, c_N|c_l, \ldots, c_k)$$
$$\times \mathrm{d}c_{k+1}\mathrm{d}c_{k+2} \cdots \mathrm{d}c_N, \tag{C1}$$

where $f_{rc}(c_{k+1}, c_{k+2}, \ldots, c_N|c_l, \cdots, c_k)$ is the conditional p.d.f. of $c_{k+1}, c_{k+2}, \ldots, c_N$ given $c_l, \ldots, c_k$. This expression is too complicated to be handled analytically. In order to perform a numerical integration of Eq. (C1), we can rewrite it as

$$f_\Delta(\Delta_k|c_l, \ldots, c_k) = \int \left[\delta \left(\Delta_k - \sum_{n=k+1}^{N} \alpha_s^n c_n \right) \right]$$
$$\times \left[\prod_{n=k+1}^{N} h_0(c_n|\bar{c}) \right] f_{\bar{c}}(\bar{c}|c_l, \ldots, c_k)$$
$$\times \mathrm{d}\bar{c}\, \mathrm{d}c_{k+1}\mathrm{d}c_{k+2} \cdots \mathrm{d}c_N. \tag{C2}$$

where $f_{\bar{c}}(\bar{c}|c_l, \ldots, c_k)$ is the conditional p.d.f. of $\bar{c}$ given $c_l, \ldots, c_k$, which can be obtained by using the Bayes' for-

mula (14) and taking the limit $\epsilon \to 0$ for the final result,

$$f_{\bar{c}}(\bar{c}|c_l, \ldots, c_k) = n_c \frac{\bar{c}_{(k)}^{n_c}}{\bar{c}^{n_c+1}} \theta(\bar{c} - \bar{c}_{(k)}). \tag{C3}$$

Appendix D: A glossary

Priori probability: the probability estimate prior to receiving new information.

Posterior probability: the revised probability that takes into account new available information.

Probability density function: a non-negative function which describes the distribution of a continuous random variable.

Random variable: a variable that takes different real values as a result of the outcomes of a random event or experiment.

Credibility: also called "degree of belief", "Bayesian probability", or "subjective probability".

Credibility measure: the credibility measure plays a similar role as the probability measure but applies to Bayesian probability.

UHO: Unknown Higher-Order

PMC: Principle of Maximum Conformality

PMCs: PMC single scale-setting approach

PMCm: PMC multi scale-setting approach

RGE: Renormalization Group Equation

RGI: Renormalization Group Invariance

p.d.f.: probability density function

CI: Credible interval

DoB: Degree-of-Belief

EC: Exact value

References

1. D.J. Gross, F. Wilczek, Ultraviolet Behavior of Nonabelian Gauge Theories. Phys. Rev. Lett. **30**, 1343 (1973)
2. H.D. Politzer, Reliable Perturbative Results for Strong Interactions? Phys. Rev. Lett. **30**, 1346 (1973)
3. A. Petermann, Normalization of constants in the quanta theory. Helv. Phys. Acta **26**, 499 (1953)
4. M. Gell-Mann, F.E. Low, Quantum electrodynamics at small distances. Phys. Rev. **95**, 1300 (1954)
5. A. Peterman, Renormalization Group and the Deep Structure of the Proton. Phys. Rept. **53**, 157 (1979)
6. C.G. Callan Jr., Broken scale invariance in scalar field theory. Phys. Rev. D **2**, 1541 (1970)
7. K. Symanzik, Small distance behavior in field theory and power counting. Commun. Math. Phys. **18**, 227 (1970)
8. S.J. Brodsky, L. Di Giustino, Setting the Renormalization Scale in QCD: The Principle of Maximum Conformality. Phys. Rev. D **86**, 085026 (2012)
9. M. Mojaza, S.J. Brodsky, X.G. Wu, Systematic All-Orders Method to Eliminate Renormalization-Scale and Scheme Ambiguities in Perturbative QCD. Phys. Rev. Lett. **110**, 192001 (2013)

10. S.J. Brodsky, M. Mojaza, X.G. Wu, Systematic Scale-Setting to All Orders: The Principle of Maximum Conformality and Commensurate Scale Relations. Phys. Rev. D **89**, 014027 (2014)

11. S.J. Brodsky, X.G. Wu, Scale Setting Using the Extended Renormalization Group and the Principle of Maximum Conformality: the QCD Coupling Constant at Four Loops. Phys. Rev. D **85**, 034038 (2012)

12. S.J. Brodsky, X.G. Wu, Eliminating the renormalization scale ambiguity for top-pair production using the principle of maximum conformality. Phys. Rev. Lett. **109**, 042002 (2012)

13. S.J. Brodsky, X.G. Wu, Self-Consistency Requirements of the Renormalization Group for Setting the Renormalization Scale. Phys. Rev. D **86**, 054018 (2012)

14. X.G. Wu, Y. Ma, S.Q. Wang, H.B. Fu, H.H. Ma, S.J. Brodsky, M. Mojaza, Renormalization Group Invariance and Optimal QCD Renormalization Scale-Setting. Rept. Prog. Phys. **78**, 126201 (2015)

15. X.G. Wu, S.J. Brodsky, M. Mojaza, The Renormalization Scale-Setting Problem in QCD. Prog. Part. Nucl. Phys. **72**, 44 (2013)

16. D.J. Gross, F. Wilczek, Asymptotically Free Gauge Theories - I. Phys. Rev. D **8**, 3633 (1973)

17. H.D. Politzer, Asymptotic Freedom: An Approach to Strong Interactions. Phys. Rept. **14**, 129 (1974)

18. W.E. Caswell, Asymptotic Behavior of Nonabelian Gauge Theories to Two Loop Order. Phys. Rev. Lett. **33**, 244 (1974)

19. O.V. Tarasov, A.A. Vladimirov, A.Y. Zharkov, The Gell-Mann-Low Function of QCD in the Three Loop Approximation. Phys. Lett. B **93**, 429 (1980)

20. S.A. Larin, J.A.M. Vermaseren, The Three loop QCD Beta function and anomalous dimensions. Phys. Lett. B **303**, 334 (1993)

21. T. van Ritbergen, J.A.M. Vermaseren, S.A. Larin, The Four loop beta function in quantum chromodynamics. Phys. Lett. B **400**, 379 (1997)

22. K.G. Chetyrkin, Four-loop renormalization of QCD: Full set of renormalization constants and anomalous dimensions. Nucl. Phys. B **710**, 499 (2005)

23. M. Czakon, The Four-loop QCD beta-function and anomalous dimensions. Nucl. Phys. B **710**, 485 (2005)

24. P.A. Baikov, K.G. Chetyrkin, J.H. Kuhn, Five-Loop Running of the QCD coupling constant. Phys. Rev. Lett. **118**, 082002 (2017)

25. S.J. Brodsky, G.P. Lepage, P.B. Mackenzie, On the Elimination of Scale Ambiguities in Perturbative Quantum Chromodynamics. Phys. Rev. D **28**, 228 (1983)

26. S.J. Brodsky, P. Huet, Aspects of SU(N(c)) gauge theories in the limit of small number of colors. Phys. Lett. B **417**, 145 (1998)

27. S.J. Brodsky, H.J. Lu, Commensurate scale relations in quantum chromodynamics. Phys. Rev. D **51**, 3652 (1995)

28. X.D. Huang, X.G. Wu, Q. Yu, X.C. Zheng, J. Zeng, J.M. Shen, Generalized Crewther relation and a novel demonstration of the scheme independence of commensurate scale relations up to all orders. Chin. Phys. C **45**, 103104 (2021)

29. M. Beneke, Renormalons. Phys. Rept. **317**, 1 (1999)

30. M. Beneke, V.M. Braun, Naive nonAbelianization and resummation of fermion bubble chains. Phys. Lett. B **348**, 513 (1995)

31. M. Neubert, Scale setting in QCD and the momentum flow in Feynman diagrams. Phys. Rev. D **51**, 5924 (1995)

32. M. Cacciari, N. Houdeau, Meaningful characterization of perturbative theoretical uncertainties. JHEP **09**, 039 (2011)

33. E. Bagnaschi, M. Cacciari, A. Guffanti, L. Jenniches, An extensive survey of the estimation of uncertainties from missing higher orders in perturbative calculations. JHEP **02**, 133 (2015)

34. M. Bonvini, Probabilistic definition of the perturbative theoretical uncertainty from missing higher orders. Eur. Phys. J. C **80**, 989 (2020)

35. C. Duhr, A. Huss, A. Mazeliauskas, R. Szafron, An analysis of Bayesian estimates for missing higher orders in perturbative calculations. JHEP **09**, 122 (2021)

36. R. L. Workman [Particle Data Group], "Review of Particle Physics," PTEP **2022**, 083C01 (2022)

37. F.J. Dyson, Divergence of perturbation theory in quantum electrodynamics. Phys. Rev. **85**, 631–632 (1952)

38. G. 't Hooft, Can We Make Sense Out of Quantum Chromodynamics? Subnucl. Ser. **15**, 943 (1979)

39. D.R. Jones, M. Schonlau, W.J. Welch, Efficient Global Optimization of Expensive Black-Box Functions. J. Global Optim. **13**, 455 (1998)

40. E. Brochu, V. M. Cora, N. de Freitas, A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning, arXiv:1012.2599

41. X.C. Zheng, X.G. Wu, S.Q. Wang, J.M. Shen, Q.L. Zhang, Reanalysis of the BFKL Pomeron at the next-to-leading logarithmic accuracy. J. High Energy Phys. **10**, 117 (2013)

42. X.G. Wu, J.M. Shen, B.L. Du, X.D. Huang, S.Q. Wang, S.J. Brodsky, The QCD Renormalization Group Equation and the Elimination of Fixed-Order Scheme-and-Scale Ambiguities Using the Principle of Maximum Conformality. Prog. Part. Nucl. Phys. **108**, 103706 (2019)

43. X. D. Huang, J. Yan, H. H. Ma, L. Di Giustino, J. M. Shen, X. G. Wu, S. J. Brodsky, Detailed Comparison of Renormalization Scale-Setting Procedures based on the Principle of Maximum Conformality. arXiv:2109.12356 [hep-ph]

44. J.M. Shen, X.G. Wu, B.L. Du, S.J. Brodsky, Novel All-Orders Single-Scale Approach to QCD Renormalization Scale-Setting. Phys. Rev. D **95**, 094006 (2017)

45. X.G. Wu, J.M. Shen, B.L. Du, S.J. Brodsky, Novel demonstration of the renormalization group invariance of the fixed-order predictions using the principle of maximum conformality and the *C*-scheme coupling. Phys. Rev. D **97**, 094030 (2018)

46. J. Yan, Z.F. Wu, J.M. Shen, X.G. Wu, Precise perturbative predictions from fixed-order calculations. J. Phys. G **50**, 045001 (2023)

47. L. Di Giustino, S.J. Brodsky, S.Q. Wang, X.G. Wu, Infinite-order scale-setting using the principle of maximum conformality: A remarkably efficient method for eliminating renormalization scale ambiguities for perturbative QCD. Phys. Rev. D **102**, 014015 (2020)

48. H.Y. Bi, X.G. Wu, Y. Ma, H.H. Ma, S.J. Brodsky, M. Mojaza, Degeneracy Relations in QCD and the Equivalence of Two Systematic All-Orders Methods for Setting the Renormalization Scale. Phys. Lett. B **748**, 13 (2015)

49. P.A. Baikov, K.G. Chetyrkin, J.H. Kuhn, Order $\alpha^4(s)$ QCD Corrections to Z and tau Decays. Phys. Rev. Lett. **101**, 012002 (2008)

50. P.A. Baikov, K.G. Chetyrkin, J.H. Kuhn, Adler Function, Bjorken Sum Rule, and the Crewther Relation to Order α_s^4 in a General Gauge Theory. Phys. Rev. Lett. **104**, 132004 (2010)

51. P.A. Baikov, K.G. Chetyrkin, J.H. Kuhn, J. Rittinger, Adler Function, Sum Rules and Crewther Relation of Order $\mathcal{O}(\alpha_s^4)$: the Singlet Case. Phys. Lett. B **714**, 62 (2012)

52. P.A. Baikov, K.G. Chetyrkin, J.H. Kuhn, J. Rittinger, Vector Correlator in Massless QCD at Order $\mathcal{O}(\alpha_s^4)$ and the QED beta-function at Five Loop. JHEP **1207**, 017 (2012)

53. R. Marshall, A Determination of the Strong Coupling Constant α^-s From e^+e^- Total Cross-section Data. Z. Phys. C **43**, 595 (1989)

54. C.S. Lam, T.-M. Yan, Decays of Heavy Lepton and Intermediate Weak Boson in Quantum Chromodynamics. Phys. Rev. D **16**, 703 (1977)

55. K.G. Chetyrkin, J.H. Kuhn, M. Steinhauser, RunDec: A Mathematica package for running and decoupling of the strong coupling and quark masses. Comput. Phys. Commun. **133**, 43 (2000)

56. F. Herren, M. Steinhauser, Version 3 of RunDec and CRunDec. Comput. Phys. Commun. **224**, 333 (2018)
57. M. Beneke, M. Jamin, alpha(s) and the tau hadronic width: fixed-order, contour-improved and higher-order perturbation theory. JHEP **09**, 044 (2008)
58. D. Boito, M. Jamin, R. Miravitllas, Scheme Variations of the QCD Coupling and Hadronic τ Decays. Phys. Rev. Lett. **117**, 152001 (2016)
59. D. Boito, P. Masjuan, F. Oliani, Higher-order QCD corrections to hadronic τ decays from Padé approximants. JHEP **08**, 075 (2018)
60. I. Caprini, Renormalization-scheme variation of a QCD perturbation expansion with tamed large-order behavior. Phys. Rev. D **98**, 056016 (2018)
61. A.N. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung* (Springer, Berlin, 1933)
62. A.N. Kolmogorov, *Foundations of the Theory of Probability*, 2nd edn. (Chelsea, New York, 1956)