Determination of the QCD $\Lambda$ Parameter and the Accuracy of Perturbation Theory at High Energies

Mattia Dalla Brida,1 Patrick Fritzsch,2 Tomasz Korzec,3 Alberto Ramos,4 Stefan Sint,5 and Rainer Sommer1,6

(ALPHA Collaboration)

1John von Neumann Institute for Computing (NIC), DESY, Platanenallee 6, 15738 Zeuthen, Germany
2Instituto de Física Teórica UAM/CSIC, Universidad Autónoma de Madrid, C/ Nicolás Cabrera 13–15, Cantoblanco, Madrid 28049, Spain
3Department of Physics, Bergische Universität Wuppertal, Gaußstrasse 20, 42119 Wuppertal, Germany
4Theoretical Physics Department, CERN, CH 1211 Geneva 23, Switzerland
5School of Mathematics, Trinity College Dublin, Dublin 2, Ireland
6Institut für Physik, Humboldt-Universität zu Berlin, Newtonstrasse 15, 12489 Berlin, Germany

(Received 22 April 2016; revised manuscript received 28 June 2016; published 24 October 2016)

We discuss the determination of the strong coupling $\alpha_s(m_Z)$ or, equivalently, the QCD $\Lambda$ parameter. Its determination requires the use of perturbation theory in $\alpha_s(\mu)$ in some scheme $s$ and at some energy scale $\mu$. The higher the scale $\mu$, the more accurate perturbation theory becomes, owing to asymptotic freedom. As one step in our computation of the $\Lambda$ parameter in three-flavor QCD, we perform lattice computations in a scheme that allows us to nonperturbatively reach very high energies, corresponding to $\alpha_s = 0.1$ and below. We find that (continuum) perturbation theory is very accurate there, yielding a 3% error in the $\Lambda$ parameter, while data around $\alpha_s \approx 0.2$ are clearly insufficient to quote such a precision. It is important to realize that these findings are expected to be generic, as our scheme has advantageous properties regarding the applicability of perturbation theory.

DOI: 10.1103/PhysRevLett.117.182001

Introduction.—The fundamental parameter of the strong interactions, the coupling $\alpha_s(\mu) = g_s^2(\mu)/(4\pi)$, is an essential input parameter for theory predictions of high-energy processes, in particular for the physics at the LHC [1–3]. Conventionally, the running $\alpha_s(\mu)$ is quoted at the electroweak scale, $\mu = m_Z$. There the coupling is weak, $\alpha = O(1/10)$, and perturbation theory (PT) is usually accurate. In particular, $\alpha_s(m_Z)$ is essentially equivalent to the renormalization-group-invariant $\Lambda$ parameter

$$\Lambda_{\text{MS}} = \varphi_{\text{MS}}(\bar{g}_{\text{MS}}(\mu)) \times \mu,$$

because the function

$$\varphi_s(g_s) = (b_0 g_s^2 - b_1/(2b_0^2)) e^{-1/(2b_0 g_s^2)} \times \exp \left\{ - \int_0^{g_s} dx \left[ \frac{1}{\beta_s(x)} + \frac{1}{b_0 x^2} - \frac{b_1}{b_0 x} \right] \right\}$$

is known precisely by replacing the renormalization group $\beta$ function by its perturbative expansion $\beta_s(\mu) = -g_s^3 \sum_{n=0}^{l_b-1} b_n \mu^n$; in the $\overline{\text{MS}}$ scheme, $\beta_{\text{pert}}(\mu)$ is known up to $l_b = 4$ loops [4,5].

At lower energies, $\mu \ll m_Z$, the perturbative uncertainty in approximating $\beta_s(\mu) \approx \beta_{\text{pert}}(\mu)$ in Eq. (2) is generally not negligible. It is $\Delta \Lambda_s/\Lambda_s = \Delta \varphi_s/\varphi_s = c_1 g_s^{-1} + \cdots$, with coefficients $c_l$, which are, for $l_b \leq 4$, of order one in the $\overline{\text{MS}}$ scheme and expected to be so in “good” schemes in general.

While the $\overline{\text{MS}}$ scheme makes sense only perturbatively, physical schemes defined beyond the perturbative expansion are easily derived from short-distance QCD observables $\varphi_s(\mu) = c_1 g_{\overline{\text{MS}}}^2(\mu) + O(g_{\overline{\text{MS}}}^4(\mu))$ via

$$g_s^2(\mu) = \varphi_s(\mu)/c_1 = g_{\overline{\text{MS}}}^2(\mu) + O(g_{\overline{\text{MS}}}^4(\mu)).$$

It is clear that high energies $\mu$ (small $\alpha_s$) and at least $l_b = 3$ are needed if one aims for a precision determination of $\alpha_s(m_Z)$. Replacing high energy by just a larger $l_b$ is dangerous because the perturbative expansion is only asymptotic, not convergent, and nonperturbative “corrections” can be large. In particular, whether one has lost control is difficult to detect because our knowledge of nonperturbative physics is very incomplete. Thus, it is a challenge to reach an accuracy of a few percent in $\Lambda_{\text{MS}}$ equivalent to subpercent accuracy in $\alpha_s(m_Z)$. 

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.
Unfortunately, the determinations that quote the smallest uncertainties do typically not come from observables at large $\mu$, and uncertainties are dominated by systematics such as unknown higher-order perturbative and nonperturbative terms. Both the Particle Data Group [6] and the Flavour Lattice Averaging Group [7] are, therefore, not just taking weighted averages of the individual determinations to arrive at their world averages.

Here we consider a family of observables (schemes) where lattice simulations allow one simultaneously to reach high precision and high energy before using PT. Then, PT at $\mu = O(m_Z)$ can be employed with confidence. In addition, one can check its applicability at lower scales. The crucial feature enabling the study of PT at high energy with continuum extrapolated nonperturbative lattice results is that we use a finite volume renormalization scheme [8,9]. QCD is considered inside a small volume of linear extent $L$ with boundary conditions and observables that do not contain any other scale. Details will be presented below. The renormalization scale then is

$$\mu = 1/L,$$

and the continuum limit of lattice simulation results is easily taken for $L/\alpha \gg 1$, with modestly sized lattices. This is the strategy of the ALPHA Collaboration, but so far it has been mostly restricted to unphysical models with an insufficient number of quark flavors [9–11]. For the interesting case of $N_f = 3$ QCD, the strategy was applied by the CP-PACS Collaboration [12]. We now have very precise results for $N_f = 3$ that allow us to see important details previously hidden by uncertainties (see also [13]).

In this Letter we discuss the most essential step: the accuracy of PT for couplings $\alpha \lesssim 0.2$ and our resulting precision for $\Lambda$. We will see that it is crucial to nonperturbatively reach $\alpha \approx 0.1$ to have confidence in PT at the 3%–4% level in $\Lambda$. On the other hand, at $\alpha \geq 0.15$ and using the three-loop beta-function, one of our schemes $\nu = -1/2$ shows a 10% systematic error in $\Lambda$. This is not a statistical fluctuation, as we will demonstrate by Eq. (21).

Given that a priori our scheme has favorable properties for PT and that other tests of perturbation theory with similar precision and similarly small $\alpha$ are presently not available, our result gives reason for concern in determinations of $\sigma_{\text{PT}}(m_Z)$ from $\mu$ values of a few GeV. This kind of lack of accuracy of PT may be one of the sources of the spread of results reviewed in [6].

The SF scheme.—Our scheme is based on the so-called Schrödinger functional (SF) [14]. There are several introductory texts on the topic with emphasis on different aspects, including the general field-theoretic concept [15], detailed descriptions [16,17], and a review of concepts, history, and recent results [18]. Here we just summarize those aspects which are needed to judge our findings below. Dirichlet boundary conditions are imposed in Euclidean time,

$$A_k(x)|_{x_0=0} = C_k, \quad A_k(x)|_{x_0=L} = C'_k,$$

for $k = 1, 2, 3$. The gauge potentials $A_k$ are taken periodic in space with period $L$ while quark fields are included as described in [19]. (In particular, the periodicity angle $\theta$ introduced in [20] is set to $\theta = \pi/5$.) The six dimensionless matrices

$$LC_k = \text{diag} \left[ \eta - \frac{\pi}{3}, \eta \left( \nu - \frac{1}{2} \right), -\eta \left( \nu + \frac{1}{2} \right), -\frac{\pi}{3} \right],$$

$$LC'_k = \text{diag} \left[ -\left( \eta + \pi \right), \eta \left( \nu + \frac{1}{2} \right), \frac{\pi}{3}, -\eta \left( \nu - \frac{1}{2} \right), -\frac{2\pi}{3} \right],$$

just depend on the two real parameters $\eta, \nu$, which multiply the Abelian generators of SU(3).

With these boundary conditions, the field which minimizes the action is unique up to gauge equivalence [9] and denoted by $A_\mu = B_\mu^{\text{class}}$. In the temporal gauge, $B_0 = 0$, it is given by $B_\mu^{\text{class}}(x) = C_k + (C'_k - C_k) x_0/L$ and corresponds to a constant color electric field. A family of couplings [21], $g_s$, is then obtained by taking $1/C_0$ in Eq. (3) to be the $\eta$ derivative of the effective action. This yields a simple path integral expectation value,

$$\langle \partial_\eta S |_{\eta=0} \rangle = \frac{12\pi}{g_s^2},$$

which is well suited for a Monte Carlo evaluation in the latticized theory. Small fluctuations around the background field generate the nontrivial orders in PT. It is worth pointing out that the whole one-parameter family of couplings can be obtained from numerical simulations at $\nu = 0$, as the $\nu$ dependence is analytically known,

$$\frac{1}{g_s^2} = \frac{1}{g_s^2} - \nu \bar{v},$$

in terms of the $\nu = 0$ observables $\bar{g}_s^2 = \bar{g}_s^2(0)$ and $\bar{v}$.

Advantageous properties of these couplings are: 1. $\Delta_{\text{stat}}\bar{g}_s^2 = s(a/L)\bar{g}_s^2 + O(\bar{g}_s^4)$, for $\Delta_{\text{stat}}$, the statistical error, at a given length of the Monte Carlo sample. This property makes it possible to maintain high precision at high energy. 2. The typical $-\mu^{-1}, \mu^{-2}$ renormalon contributions [22] are absent because the finite volume provides an infrared momentum cutoff. Instead, the leading known nonperturbative contribution is due to a secondary stationary point of the action [23] at $\bar{g}_s^2 S(B^{\text{class}}) - S(B^{\text{class}}) = 32.9$. It generates corrections to PT of order

$$\exp(-2.62/\alpha) \sim (\Lambda/\mu)^{3.8},$$

which evaluates to $O(10^{-6})$ for $\alpha = 0.2$. At such values of $\alpha$, fields with nonzero topology are even further suppressed.
given that $g_0^2[S(B) - S(B^{\text{class}})] \geq 6a^2$ \cite{9,14}. The $\beta$ function is known, including its three-loop term,

$$(4\pi)^3 \times b_{2,\nu} = -0.06(3) - \nu \times 1.26, \quad (N_f = 3), \quad (9)$$

and for reasonable values of $\nu$ the three-loop term is of order one as it is in the \overline{MS} scheme. As we will see, discretization effects are very small; at tree level of perturbation theory they are $O((a/L)^3)$. They are known to two-loop order in PT \cite{24} and we can subtract those pieces \cite{25}.

The downside of the SF scheme is that the coefficient $s(a/L)$ diverges like $(L/a)^{1/2+\epsilon}$ for large $L/a$ and is not that small in general. Here $\epsilon$ is the dynamical critical exponent of the algorithm while the $1/2$ in the exponent is due to the variance of the observable \cite{25}. High statistics is needed and our computation is limited to $L/a \leq 24$. A second issue is the acceleration of the approach to the continuum limit through Symanzik improvement. With our Dirichlet boundary conditions the Symanzik effective Lagrangian contains terms located at the time-boundaries. These are responsible for $O(a)$ effects. We cancel them by corresponding improvement terms with coefficients $c_l$ and $\bar{c}_l$ known only in PT, see below.

**Step scaling functions and $\Lambda$ parameter.**—The non-perturbative energy dependence of finite volume couplings is constructed from the step scaling function \cite{8}

$$\sigma(u) = g_0^2(1/(2L))|_{g_0^2(1/L) = u, m = 0}, \quad (10)$$

where $m = 0$ ensures the quark-mass independence of the scheme \cite{26}. The step scaling function corresponds to a discrete version of the $\beta$ function and is computed as the continuum limit $a/L \rightarrow 0$ of its lattice approximants $\Sigma_{\nu}(u, a/L)$. The conditions $g_0^2(1/L) = u$ and $m = 0$ then refer to an $(L/a)^2$ lattice, and fix the bare coupling and bare quark mass of the theory. $g_0^2(1/(2L))$ is to be evaluated for the same bare parameters on a $(2L/a)^4$ lattice.

We will use the $\nu = 0$ scheme for a reference, dropping the index $\nu$. The scale $L_0$ is defined by a value $u_0$ and the condition

$$\tilde{g}_0^2(1/L_0) = u_0. \quad (11)$$

The solution of the implicit equation

$$u_k = \sigma(u_{k+1}), \quad (12)$$

for $u_{k+1}, \ k = 0, 1, \ldots$, gives a series of couplings $u_k = \tilde{g}_0^2(2^k/L_0)$. With a few steps, one reaches $\mu = 1/L_n = 2^n/L_0 = O(m_z)$, and the perturbative $\varphi$ at this high scale will give a good approximation to $L_0\Lambda$,

$$L_0\Lambda = 2^n\varphi(\sqrt{u_n}). \quad (13)$$

Note that, thanks to Eq. (7) and the exact relation between $\Lambda$ parameters \cite{9,20}

$$r_\nu = \Lambda/\Lambda_\nu = e^{-\nu \times 1.25516}, \quad (14)$$

the same combination $L_0\Lambda$ can be obtained in any scheme with $\nu \neq 0$. Whatever different values of $\nu$, numbers of steps $(n)$, and perturbative orders $(l_\nu)$ give consistent results is an excellent way to test the reliability of perturbation theory.

**Simulations.**—We used the standard Wilson plaquette action and three massless $O(a)$-improved \cite{27,28} quarks simulated by a variant of the OpenQCD code \cite{29,30}. At eight couplings $g_0^2(1/L)$ in the range $1.11–2.02$, we simulated pairs of lattices $L/a, 2L/a$ with $L/a = 4, 6, 8$ and at three couplings we also included $L/a = 12$.

Between 80 000 and 300 000 independent Monte Carlo measurements were taken on each lattice. As we have already noted, nontrivial topology is very suppressed in these small volumes \cite{31}. Therefore, topology freezing \cite{32,33} is irrelevant here.

A critical issue for any lattice computation is the removal of discretization effects. In preparation of our continuum extrapolations, we apply both Symanzik improvement of the action and perturbative improvement of the step scaling function \cite{25}. In comparison to earlier work, we here propagate the estimated uncertainty of those $O(a)$ improvement coefficients that are only known perturbatively into the errors of the step scaling functions. They can then be assumed to be free of linear $a$ effects within their errors. Details are found in \cite{34}.

**Continuum extrapolations and results.**—As the residual linear $a$ effects are treated as an uncertainty, we can proceed with continuum extrapolations linear in $a^2$. First, we look at the data in Fig. 1. They are statistically compatible with having no $a$ effects for $L/a \geq 6$; for $N_f = 0$, this was found with similar precision for $L/a \geq 5$ (see Fig. 3 of \cite{35}).

![](image.png)

FIG. 1. Continuum limit of the step scaling function $\Sigma^{(i)}(u, a/L)/u$ with $i = 2$ loop improvement and for $\nu = 0$. As an illustration, a constant ($n_\nu = 0$, dashed, fit $G$) and a linear ($n_\nu = 2$, fit $C$) continuum extrapolation are shown. Continuum extrapolated results include the errors due to $c_l$ and $\bar{c}_l$ (cf. text). The star symbols show the perturbative $\sigma$ computed from the three-loop $\beta_{\text{pert}}$. 

182001-3
Both the continuum limit of the step scaling function and its cutoff effects are smooth functions of the coupling. This motivates global fits of the form
\[ \Sigma_c^{(i)}(u, a/L) = \sigma_c(u) + \rho_c^{(i)}(u)(a/L)^2, \]
where \( i \) is the order of PT to which cutoff effects are removed [34]. We performed various such fits in order to assess the systematic errors that result from the assumptions made in the fit functions. We parameterize the cutoff effects by a polynomial in \( u \), with the correct asymptotics for small \( u \),
\[ \rho_c^{(i)}(u) = \sum_{k=1}^{n} \rho_{c,k}^{(i)} u^{i+1+k}, \]
where the case of neglecting cutoff effects is covered by \( n_i^{(i)} = 0 \). The continuum step scaling function is naturally parameterized by a polynomial in \( u \),
\[ \sigma_c(u) = u + u^2 \sum_{k=0}^{3} s_k u^k. \]

Lower-order coefficients are fixed to their known perturbative values while \( s_3 ("n_c = 1") \) or \( s_2, s_3 ("n_c = 2") \) are fit parameters. A selection of such fits are illustrated in Table I. Instead of the parameters of the continuum step scaling function, the table directly shows the extracted \( L_0^\Lambda \), where \( L_0 \) is defined through Eq. (11) and the value \( u_0 = 2.012 \). Recalling Eq. (7) and using \( \bar{v} = 0.1199(10) \) (see next section), we have
\[ g_0^2/(1/L_0) = 2.012[1 - 0.1199(10) \times 2.012 u]^{-1}. \]

Apart from the form of the fit, \( L_0^\Lambda \) depends on the value of \( n \), where Eq. (13) with \( \beta_c = \beta_{c}^{\text{pert}} \) is used. Because we insert \( \beta_{c}^{\text{pert}} \) at three loop, the residual dependence on the coupling is \( O(\alpha^2(1/L_n)) \).

The observed behavior, Fig. 2, is consistent with a dominantly linear dependence of \( L_0^\Lambda \) on \( \alpha^2(1/L_n) \). For \( \nu = 0 \) the slope is not very significant and disappears around \( \nu = 0.3 \), but for \( \nu = -0.5 \) it is quite large and outside errors.

This result suggests performing alternative fits, where the continuum step scaling function is parameterized by an effective four-loop \( \beta \) function, adding a term \( b_3^{\text{eff}} u^3 \) to the perturbative known ones. The determined \( L_0^\Lambda \) are then automatically independent of \( n \) and we include \( b_3^{\text{eff}} \) instead of \( u_{n=4} \) in the table. For \( \nu = -0.5 \) the effective fit value is larger than it should be in a well-behaved perturbative expansion.

We will come back to this issue shortly, but first we give our result for \( L_0^\Lambda \). We take the standard polynomial fit to \( \sigma \) (for \( \nu = 0 \)) with \( \alpha_6 \approx 0.1 \) \( (u_6 \approx 1.2) \). A typical perturbative error of size \( \Delta(L^\Lambda) = \alpha_2^2 L^\Lambda \) is then a factor of 3 or more below our statistical errors. We quote [with \( \bar{g}_0^2/(1/L_0) = 2.012 \)]
\[ L_0^\Lambda = 0.0303(8) \rightarrow L_0^\Lambda(3)_{\text{MS}} = 0.0791(21). \]

with the known [9,20] \( L^\Lambda /\Lambda_{\text{MS}} \). This is the result of fit \( C \). It is in perfect agreement with all variations of the global fit, even with fit \( G \), which neglects all cutoff effects but uses only data with \( L/a \geq 8 \). It has a rather conservative error. If an even more conservative result is desired, one may take the one of fit \( D \), \( L_0^\Lambda = 0.0303(13) \).

Accuracy of perturbation theory.—While \( b_{3,\text{eff}} \) is large for \( \nu = -0.5 \), it does have an error of around 50%. A much better precision can be achieved by directly considering the observable
\[ \omega(u) = \bar{v}|_{\bar{g}_0^2/(1/L_0)=a,u,m=0} = v_1 + v_2 u + O(u^2), \]
with coefficients \( v_1 = 0.14307, v_2 = -0.004693 \) [11]. In contrast to the step scaling function, \( \omega(u) \) does not require

### Table I. Results for \( \nu \) in the upper part.

<table>
<thead>
<tr>
<th>Fit</th>
<th>( u_3 )</th>
<th>( i )</th>
<th>( (L/a)_{\text{min}} )</th>
<th>( n_i^{(i)} )</th>
<th>( n_c )</th>
<th>( L_0^\Lambda \times 100 )</th>
<th>( b_3^{\text{eff}} \times (4\pi)^4 )</th>
<th>( \chi^2 )</th>
<th>Degrees of freedom</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1.193(4)</td>
<td>0</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>3.04(8)</td>
<td>14.7</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>1.194(4)</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>3.07(8)</td>
<td>14.2</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>1.193(5)</td>
<td>2</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>3.03(8)</td>
<td>14.5</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>1.192(7)</td>
<td>2</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>3.03(13)</td>
<td>14.5</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>1.191(11)</td>
<td>2</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>3.00(11)</td>
<td>4(3)</td>
<td>14.6</td>
<td>16</td>
</tr>
<tr>
<td>F</td>
<td>1.191(11)</td>
<td>2</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>3.01(11)</td>
<td>4(3)</td>
<td>12.7</td>
<td>9</td>
</tr>
<tr>
<td>G</td>
<td>1.191(11)</td>
<td>2</td>
<td>8</td>
<td>0</td>
<td>2</td>
<td>3.02(20)</td>
<td>13.0</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>1.191(11)</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>3.04(10)</td>
<td>3(3)</td>
<td>14.1</td>
<td>16</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fit</th>
<th>( \nu )</th>
<th>( i )</th>
<th>( (L/a)_{\text{min}} )</th>
<th>( n_i^{(i)} )</th>
<th>( n_c )</th>
<th>( L_0^\Lambda \times 100 )</th>
<th>( b_3^{\text{eff}} \times (4\pi)^4 )</th>
<th>( \chi^2 )</th>
<th>Degrees of freedom</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>-0.5</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>3.03(15)</td>
<td>11(5)</td>
<td>10.4</td>
<td>16</td>
</tr>
<tr>
<td>H</td>
<td>0.3</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>3.04(10)</td>
<td>0(3)</td>
<td>20.0</td>
<td>16</td>
</tr>
</tbody>
</table>
pairs of lattices, so that the continuum extrapolation can be performed using data for the entire range of lattice sizes \( L/a = 6, 8, 10, 12, 16, 24 \). Improvement and fits for obtaining the continuum limit are carried out in analogy to those of \( \Sigma \). Figure 3 shows the result of two different fits with fit parameters \( d_k \) in \( \omega(u) = v_1 + v_2 u + d_1 u^2 + d_2 u^3 + d_3 u^4 \) and in \( \omega(u) = v_1 + d_1 u^1 + d_2 u^2 + d_3 u^3 + d_4 u^4 \). The overall band of the two fits may be taken as a safe estimate of the continuum limit. As an example we find \( \omega(2.012) = 0.1199(10) \) for both fits, leading to Eq. (18). In the above analysis we did not use data with \( L/a = 6 \). Including them yields only tiny changes and excellent \( \chi^2 \) values.

A good measure of the deviation from two-loop perturbation theory is

\[
[\omega(\bar{\alpha}) - v_1 - v_2 \bar{\alpha}^2] / v_1 = -3.7(2) \alpha^2
\]

at \( \alpha = 0.19 \). It is quite large and statistically significant beyond any doubt. If one attempts to describe this by perturbation theory, the three-loop coefficient \( v_3 \) has to be too large for perturbation theory to be trustworthy at \( \alpha = 0.2 \). Again, we come to the conclusion that \( \alpha \approx 0.1 \) needs to be reached nonperturbatively before perturbation theory becomes accurate.

**Summary and conclusions.**—Our chosen definition of \( \alpha_s(\mu) \) allows us to compute it with very good precision through lattice Monte Carlo simulations. In particular, we have controlled the errors due to the discretization of the theory at large \( \mu \). Known nonperturbative corrections are parametrically very small, \( O(e^{-2.6/\alpha}) \). In other words, we have an excellent scheme to test the accuracy of PT in a given region of \( \alpha \).

In fact, we have a family of schemes, depending on \( \nu \). For small positive \( \nu \), the couplings follow perturbation theory very closely in the full investigated range \( 0.1 \leq \alpha \leq 0.2 \), as illustrated by the flatness of \( \Lambda \) in Fig. 2 extracted from Eq. (13) with the three-loop \( \beta \) function.

However, for negative \( \nu \), e.g., \( \nu = -0.5 \), values of \( \alpha \) just below 0.2 are not small enough to confirm perturbative behavior. The observable \( \tilde{\nu} \), Fig. 3, shows that the \( \alpha \) dependence seen in Fig. 2 is not just a statistical fluctuation. We could take the continuum limit of \( \tilde{\nu} \) with very high precision and Eq. (21) shows a clear deviation from the known perturbative terms, corresponding to \( l_b = 3 \), at \( \alpha = 0.19 \).

We conclude that it is essential to reach \( \alpha = 0.1 \) in order to be able to achieve a precision around 3% for the \( \Lambda \) parameter. Fortunately, we have access to that region and can quote such an accuracy in Eq. (19). While of course schemes exist where three-loop running holds accurately down to smaller energies—for example, as far as we can tell, \( \nu = 0.3 \) produces flatness in Fig. 2—knowing whether a chosen scheme possesses this property is difficult unless one also has control over the \( \alpha \approx 0.1 \) region. Once this is achieved, data at larger \( \alpha \) become irrelevant.

We reported in this Letter a part of our determination of a precise value for \( \Lambda_{\overline{\text{MS}}} \). As our next step, we will soon connect \( L_0 \) to the decay constants of pion and kaon, as explained above and in [36].

We thank our colleagues in the ALPHA Collaboration, in particular M. Bruno, C. Pena, S. Schaefer, H. Simma, and U. Wolff, for many useful discussions. We thank U. Wolff for a critical reading of the manuscript. We would also like to show our gratitude to S. Schaefer and H. Simma for their invaluable contribution in adapting the OpenQCD code. We thank the computer centers at HLRN (bep00040) and NIC at DESY, Zeuthen for providing computing resources and support. S.S. acknowledges support by Science Foundation Ireland (SFI) under Grant No. 11/RFP/PHYS3218. P.F. acknowledges financial support from the Spanish Ministerio de Economía y Competitividad (MINECO’s) “Centro de Excelencia Severo Ochoa” Programme under Grant No. SEV-2012-0249, as well as from the Ministerio de Ciencia e Innovación (MCINN) Grant No. FPA2012-31686. This work is based on previous work supported by the Deutsche Forschungsgemeinschaft in the SFB/TR 09.