Determination of the QCD coupling from the static energy and the free energy

Alexei Bazavov,[1] Nora Brambilla,[2,3] Xavier Garcia i Tormo,[4] Péter Petreczky,[5] Joan Soto,[4,6]
Antonio Vairo,[2] and Johannes Heinrich Weber[1,7]

(TUMQCD Collaboration)

[1]*Department of Computational Mathematics, Science and Engineering and Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824, USA*
[2]*Physik Department, Technische Universität München, James-Franck-Straße 1, D-85748 Garching, Germany*
[3]*Institute for Advanced Study, Technische Universität München,
Lichtenbergstraße 2a, D-85748 Garching, Germany*
[4]*Institut de Ciències del Cosmos, Universitat de Barcelona,
Martí i Franquès 1, E-08028 Barcelona, Catalonia, Spain*
[5]*Physics Department, Brookhaven National Laboratory, Upton, New York 11973, USA*
[6]*Departament de Física Quàntica i Astrofísica, Universitat de Barcelona,
Martí i Franquès 1, E-08028 Barcelona, Catalonia, Spain*
[7]*Excellence Cluster ORIGINS, Boltzmannstrasse 2, D-85748 Garching, Germany*

We present two determinations of the strong coupling α_s. The first one is from the static energy at three-loop accuracy, and may be considered an update of earlier determinations by some of us. The new analysis includes new lattice data at smaller lattice spacings, and reaches distances as short as 0.0237 fm. We present a comprehensive and detailed estimate of the error sources that contribute to the uncertainty of the final result, $\alpha_s(M_Z) = 0.11660^{+0.00110}_{-0.00056}$. The second determination is based on lattice data for the singlet free energy at finite temperature up to distances as small as 0.0081 fm, from which we obtain $\alpha_s(M_Z) = 0.11638^{+0.00095}_{-0.00087}$.

I. INTRODUCTION

A precise determination of the strong coupling α_s is of key importance both for the theory of strong interactions, which is determined by such a fundamental parameter, and for investigations of new physics beyond the standard model. Cross section calculations at the Large Hadron Collider, for example, suffer from the uncertainties related to our limited knowledge of α_s. The last decades have witnessed an impressive effort in the α_s extraction from an ample range of physical observables with a broad sweep of different methods. Notwithstanding all these efforts, the Particle Data Book (PDG) world average value for α_s in 2018, $\alpha_s(M_Z) = 0.1181 \pm 0.0011$, has an overall uncertainty that has almost doubled with respect to the PDG

average in 2013, $\alpha_s(M_Z) = 0.1185 \pm 0.0006$ [1]. This is due to the fact that the uncertainty of several determinations that enter the average is dominated by errors of theoretical origin, which are often difficult to precisely assess. It is important therefore to use a variety of different ways to extract α_s and to validate each extraction at the best of the state of the art.

In this paper we aim at an improved determination of α_s both by updating earlier extractions from the static energy and by proposing a new method that involves the singlet free energy.

The static energy is an observable up to an additive constant, and it is a function of the distance r between the static quark and the static antiquark. In the limit of massless dynamical quarks, it depends only on α_s. It can be calculated on the lattice for any distance $r = \sqrt{x^2 + y^2 + z^2}$, where x, y and z are integer multiples of the lattice spacing. For short distance r it can be calculated in perturbation theory in QCD as a function of α_s using nonrelativistic QCD effective field theories. Notably, perturbation theory is accurate for this quantity at three loops,

and the tree-level result is already sensitive to α_s. The comparison between the perturbative expression and lattice data at distances small enough to be accessible to perturbation theory is a good way to obtain a precise determination of α_s. For this endeavor it is crucial to have lattice data covering an interval of sufficiently small distances in order to be sensitive to the minute details of the curvature of the static energy, namely, with sufficiently fine lattice spacing. At the present day, these lattices still pose a major challenge. In this paper we use lattices [2] with an extraordinarily fine lattice spacing $a = 0.0249$ fm to achieve a systematically improved extraction of α_s.

Additionally we exploit a new idea. One reason for which it is challenging to reach such fine lattice spacings in lattice QCD simulations with dynamical quarks is that one has to simultaneously maintain the control over finite volume effects arising from the propagation of the lightest hadronic modes, namely, the Goldstone bosons, at the pion scale. A lattice simulation at high enough temperature avoids this infrared problem, and thus enables reaching much finer lattice spacings using smaller volumes. We use finite temperature lattices with unprecedentedly fine lattice spacing $a = 0.00848$ fm [3]. The singlet static free energy is again a function of the static quark-antiquark distance and has been calculated on the lattice [3] and perturbatively [4]. The comparison between the two offers a novel and independent method to extract a precise determination of α_s.

The rest of the paper is organized as follows. In Sec. II we give an account of the gauge ensembles and lattice correlators that we use and discuss the relevant systematic effects in the lattice data. In Sec. III after briefly recalling the procedure used in Ref. [5] to extract α_s from the static energy, we compare between the lattice data and the weak-coupling result, discuss our estimates of systematic uncertainties, and finally obtain an updated extraction of α_s. In Sec. IV we discuss the singlet free energy at finite temperature, and outline the relation with the static potential and static energy at zero temperature. We discuss the relevant scale hierarchies, compare with the lattice, and proceed to extract α_s at even shorter distances. Section V contains some discussion of our results, as well as a comparison with previous related works. Finally, in Sec. VI we present a short summary of the main results and conclude. In Appendix A we discuss the gauge ensembles used in this study. In Appendix B we discuss discretization artifacts at short distances in detail. In Appendix C we list the coefficients appearing in the perturbative results used in this paper.

II. LATTICE SETUP

We require the lattice result of the static energy at the smallest available distances in order to compare with the weak-coupling calculations. For this reason we employ the $2 + 1$ flavor ensembles using the highly improved staggered quark (HISQ) action [6] that have been generated for studies of the QCD equation of state [2,7]. The spatial

TABLE I. Parameters for the fine $T = 0$ ensembles. In the seventh column we indicate the number of correlator measurements performed (MEAS).

$m_l = m_s/20$:							
β	a (fm)	N_σ, N_τ	am_s	$m_\pi L$	#TUs	#MEAS	Ref.
7.373	0.060	$64^3 \times 48$	0.0250	2.3	4623	1000	[7]
7.596	0.049	64^4	0.0202	2.6	4757	1000	[7]
7.825	0.040	64^4	0.0164	2.0	4768	1000	[7]

$m_l = m_s/5$:							
β	a [fm]	N_σ, N_τ	am_s	$m_\pi L$	#TUs	#MEAS	Ref.
8.000	0.035	64^4	0.01299	3.6	4616	1000	[2]
8.200	0.029	64^4	0.01071	3.1	4616	1000	[2]
8.400	0.025	64^4	0.00887	2.6	4616	1000	[2]

volume is given by $V_3 = (aN_\sigma)^3$, and the physical length of the Euclidean time direction is aN_τ. We summarize the zero temperature ensembles[1] with $N_\tau = 64$ in Table I, which correspond to two different pion masses in the continuum limit. Namely, these ensembles have different light sea quark masses, $m_l = m_s/20$ or $m_l = m_s/5$. The strange quark mass m_s is at its physical value, namely, such that the mass of the hypothetical $\eta_{s\bar{s}}$ meson is reproduced as $m_{\eta_{s\bar{s}}} = 686$ MeV. The lattice spacing a has been fixed by the r_1 scale, which is defined by the equation

$$r^2 \frac{dE(r)}{dr}\bigg|_{r=r_1} = 1; \qquad (1)$$

see Refs. [2,7] for the details. Here and in the following $E(r)$ denotes the QCD static energy calculated on the lattice. We use the value $r_1 = 0.3106(17)$ fm determined from the pion decay constant. Since the quark mass dependence of r_1/a is small, the $m_l = m_s/20$ and $m_l = m_s/5$ results can be combined to produce a parametrization of r_1/a as a function of the bare lattice gauge coupling β [2]. The tunneling between different topological sectors is suppressed for these gauge ensembles. Nevertheless, no statistically significant difference has been observed for bulk observables other than the light quark condensate [2] in different topological sectors. For the distances we consider in this study the static energy [8] does not vary significantly between the different topological sectors.

In addition, we use the finite temperature results of the singlet free energy with $N_\sigma/N_\tau = 4$ and $N_\tau = 12$, or 16. These ensembles correspond to the thermal QCD medium at temperatures $T = 1/(aN_\tau)$. The finite temperature ensembles

---

[1]The lattice observables (gluon action density, and sea quark condensates) contributing to the equation of state have divergent, additive contributions, which depend only on the lattice spacing. Renormalization group invariant observables are obtained by subtracting the $T = 0$ observables from the $T > 0$ observables at each lattice spacing.

have been generated using lattice parameters that would correspond to the same two pion masses (at zero temperature) in the continuum limit as the zero temperature ensembles [3]. For these we use the same parametrization of r_1/a. We give an account of these ensembles in Appendix A.

The static energy can be calculated using Wilson line correlators in Coulomb gauge or Wilson loops. We use the former because it is more convenient for practical reasons. For additional cross-checks we also studied Wilson loops for $\beta = 7.825$. When extracting the static energy from Wilson loops the spatial lines have to be smeared in order to obtain a reasonable signal-to-noise ratio and suppress excited states. We used one, two, and five levels of spatial smearing using the HYP (hypercubic smearing) algorithm [9] in our calculations. In the studies of the equation of state [2,7], the static energy was extracted from two-exponential fits to the Wilson line correlators using a fixed range of Euclidean time τ/a that was chosen independently from r/a. This previous calculation is not adequate for our analysis. Here we improve the extraction of the static energy in the following way.

(i) We explicitly include the correlation matrix of the static energy in the fits to determine α_s. For this reason we have to recalculate the static energy and preserve its statistical correlations.

(ii) We explicitly distinguish at short distances between data with distinct path geometries instead of averaging over them. This increases the number of data for each β compared to the earlier analysis [5].

(iii) At small τ the correlators of Wilson lines are contaminated by excited states, while at large τ the statistical fluctuations are large, and the effective masses sometimes do not show a clear plateau; see Fig. 1 for some of the worst-case scenarios. Therefore, a careful choice of the lower end, $\tau_{\min}$, and upper end, $\tau_{\max}$, of the fit interval is needed to obtain reliable results of the static energy. The corresponding fit interval $[\tau_{\min} : \tau_{\max}]$ depends on the distance r. We used a data-driven approach to determine suitable fit windows and estimated the systematic errors due to this choice of the fit interval. The procedure of extracting the static energy is demonstrated in Fig. 1, where the solid lines show the static energy and the dashed lines the corresponding total uncertainty. We summarize the fit windows used for obtaining the ground state in Appendix A.

At short distances, namely, for $r \lesssim 0.14$ fm, we observe the nonmonotonic behavior of the effective mass at small τ/a, typically, within $\tau \lesssim 0.14$ fm. This nonmonotonic behavior is not restricted to the Wilson line correlator in Coulomb
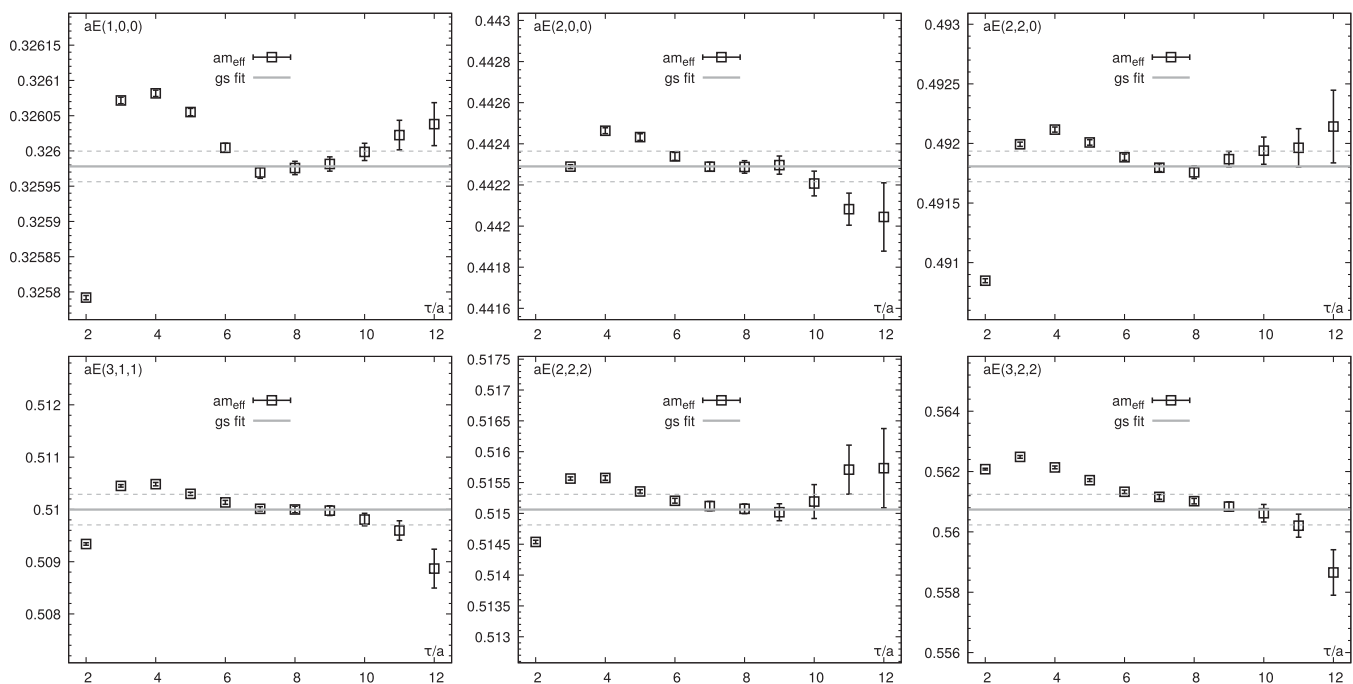

FIG. 1. Effective mass of the static energy in units of the inverse lattice spacing using Wilson line correlators in Coulomb gauge. We show $\beta = 8.2, r/a = \sqrt{1}, \beta = 8.4, r/a = \sqrt{4}$, and $\beta = 8.4, r/a = \sqrt{8}$ (top, from left to right), and $\beta = 8.4, r/a = \sqrt{11}, \beta = 8.4, \sqrt{12}$, and $\beta = 8.2, r/a = \sqrt{17}$ (bottom, from left to right) as function of the Euclidean time τ/a. Black open squares indicate the mean and statistical error of the effective mass. Although, we can usually identify an unambiguous plateau in the window $7 \leq \tau/a \leq 11$, we have to restrict the fit to narrower windows in multiple cases that are similar to those shown above. There is no evident pattern regarding β, r/a or r for the appearance of these unstable plateaus. The gray solid line indicates the central value of the ground state fit; the dashed gray lines indicate the one sigma interval for the total error that includes the systematic errors due to the dependence on the fit interval.
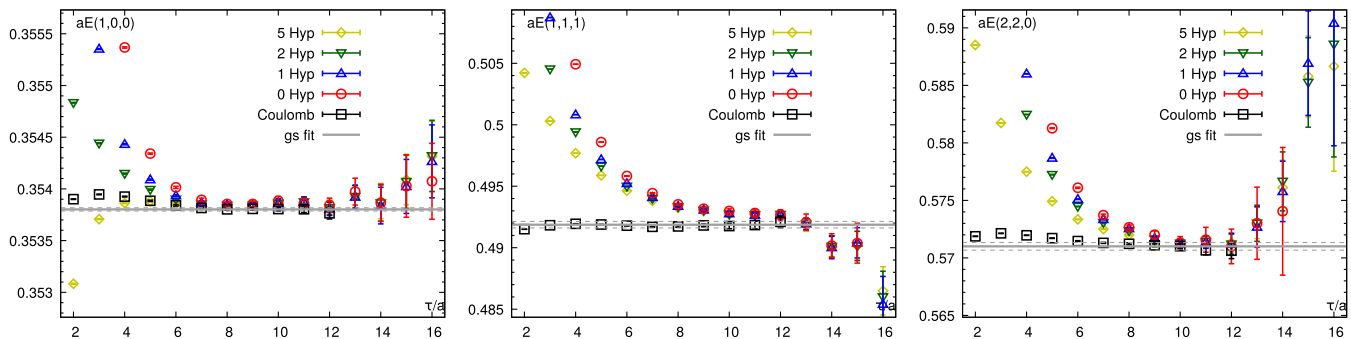
FIG. 2. Effective mass of the static energy in units of the inverse lattice spacing using Wilson loops with or without spatial smearing. We show $\beta = 7.825$ with $r/a = \sqrt{1}$, $\sqrt{3}$, and $\sqrt{8}$ (from left to right) as a function of the Euclidean time τ/a. Black open squares indicate the mean and statistical error of the effective mass for Wilson lines in Coulomb gauge; colored symbols indicate mean and statistical errors of the static energy for Wilson loops. The gray solid line indicates the central value of the ground state fit; the dashed gray lines indicate the one sigma interval. The Wilson loops have the more severe excited state contamination, which is increasingly suppressed by the smearing. More smearing is required for larger r/a, and the signal may drown into statistical noise before the masses of Wilson loops reach the same plateau value as the Wilson line correlators.

gauge, but also appears in Wilson loops with spatially smeared spatial Wilson lines; see Fig. 2. For this reason, these nonmonotonicities are due to the gauge ensembles with improved gauge action instead of being due to the details of the interpolating operator. The effective masses obtained from correlators of Wilson lines in Coulomb gauge and from Wilson loops with HYP smeared spatial lines converge to the same plateau at large τ.

The current and previous results for the static energy are consistent within errors. For the given reasons the new result is more precise at short distances and has a better estimate of the errors. It supersedes the previous calculation [2,7].

The last major source of systematic uncertainties of the lattice data is due to discretization artifacts at short distances, which are caused by the breaking of rotational symmetry from O(3) in the continuum to the cubic group W_3 on the lattice. Even for an improved gauge action these artifacts are significantly larger than any of the other uncertainties of the lattice result. The artifacts express themselves in the form of a variation of the lattice static energy around the continuum static energy at any given distance r. The sign and the size of this variation depends on the geometry of the shortest paths connecting the two sites at distance r on the lattice. This variation is typically large for on-axis distances, i.e., $\boldsymbol{r}/a = (n, 0, 0)$, while being smaller for off-axis distances, and shrinks to the point of numerical irrelevance as the distances become larger than $r/a \gtrsim 5$. These lattice artifacts could be understood qualitatively in leading-order perturbation theory, i.e. at tree level. The analysis presented in Appendix B shows that the large artifacts at $r/a = 1$ are about 8% for the Wilson action and 4% for the Lüscher–Weisz action that we use. The tree-level discretization effects decrease rapidly at larger distances, and become very small for $r/a \geq \sqrt{8}$. One can use the tree-level results to reduce lattice artifacts

also in the interacting theory. This procedure is called the tree-level correction and discussed in Appendix B.

The statistical errors of the static energy obtained on the lattice increase when increasing the separation r. For $r/a \geq \sqrt{8}$ the statistical errors in the static energy are typically larger than the discretization errors after tree-level correction. Thus for these distances discretization errors can be neglected. For $\sqrt{5} \leq r/a \leq \sqrt{8}$ the statistical and discretization errors are comparable. In this region one can deal with the discretization errors by estimating them and combining with the statistical errors. For very short distances $r/a < \sqrt{5}$ the discretization effects are much larger that the statistical errors and should be estimated carefully and corrected for, before the lattice results can be compared to the weak-coupling calculations. The corresponding procedure is described in Appendix B.

III. EXTRACTING α_s FROM THE STATIC ENERGY

The static energy of a quark-antiquark pair is directly accessible in lattice calculations. Our lattice calculation of the static energy for $N_f = 2 + 1$ light flavors has been discussed in Sec. II. In the weak-coupling regime the static energy is known up to three loops, i.e., up to order α_s^4 in the small α_s expansion. The resummation of $\ln \alpha_s$ terms is also known at subleading order. Hence, since realistic lattice calculations for $N_f = 2 + 1$ are available at short enough distances where the small α_s expansion is reliable, α_s can be extracted by comparing the lattice data for the static energy with its perturbative expression for the case of three (approximately) massless quarks.

This program was initially carried out in Ref. [10] and considerably improved in Ref. [5], to a large extent due to the incorporation of shorter distance data on finer lattices. Here, we further improve the α_s extraction by including

three even finer lattices. The abundance of data at short distances allows us to perform the analysis omitting the data with $r/a < \sqrt{5}$ that are susceptible to significant discretization artifacts. It also permits us to quantify the impact of using such data with or without appropriate corrections. Rather than making a separate analysis for each lattice spacing and averaging the final results, here we make a global fit to all lattice data, leaving as a free parameter a normalization constant for each lattice spacing. This approach is well motivated from the earlier observation that no remaining lattice spacing dependence can be resolved from the three lattice spacings in Ref. [5] and that the scale uncertainties for all considered lattice spacings are similar. We also take into account the statistical correlations between the fluctuations of data at different distances on the lattice. Eventually the effect of these correlations turns out to be much smaller than the statistical uncertainty, even though it reduces the uncertainty of the extracted value of α_s.

We focus on data for $r \lesssim 0.14$ fm or $r \lesssim 0.45r_1$; see Fig. 3. The shortest distance that we can access, namely, due to a single lattice spacing[2] on our finest lattice (at zero temperature) is 0.0237 fm. In Ref. [5] only data with 0.057 fm $\lesssim r \lesssim 0.16$ fm were used for the extraction. For that range it was shown that the small α_s expansion was reliable. Since we use only shorter distances in this paper, the perturbative expansion remains reliable. In Ref. [5], separate analyses with data for each gauge ensemble were performed in the range $r/a \geq \sqrt{2}$. In this work, thanks to the finer lattices and the combined analysis, we can afford to vary the minimal distance r/a and quantify the corresponding uncertainty. Data on coarser lattices in the same r range as well as data on finer lattices may be excluded due to restrictions of the r/a range in the analysis. This is shown with more detail in Fig. 3. It was also shown in Ref. [5] that there is no sensitivity to nonperturbative contributions giving rise to powerlike corrections. This is important, since sea quark mass effects, finite volume effects, or effects due to an incorrect expectation value of the topological charge would make themselves noticeable in such contributions,[3] if they were not negligibly small.

To obtain the perturbative result for the static energy we follow Ref. [5]. We use the perturbative result for the force, $F(r) = dE/dr$, and integrate it to obtain the energy. The constant contribution obtained this way is irrelevant and subsumed into the normalization constants that we have to fix anyway when comparing the weak-coupling results to

---

[2]Actually, this is 4% smaller than the lattice spacing itself due to the tree-level correction.

[3]These contributions are also related to the second renormalon of the singlet potential, which needed to be accounted for explicitly in Refs. [11,12] because the range of data used in those references extends to much larger distances.
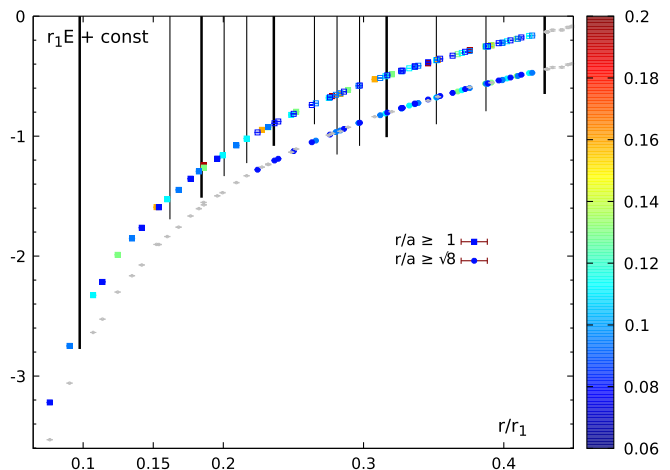

FIG. 3. Lattice data for the static energy in units of r_1. The data are shown in vertically displaced sets corresponding to different minimal distances in units of the lattice spacing r/a that were used in the fits. We restricted the data included in each fit by the maximal distances in physical units r/r_1 (indicated by vertical lines). The thicker vertical lines roughly indicate the maximal distances listed in Tables II–V. The color indicates the lattice spacing in units of r_1. Filled colored symbols in the set of points labeled as $r/a \geq 1$ are not accessible in the fits that use only $r/a \geq \sqrt{8}$. Gray symbols either indicate data that were not used in the fits with a given minimal r/a, but are below the respective maximal r/r_1, or that are above the maximal r/r_1 that we included in the fits.

the lattice data at each β. A choice of the renormalization scale ν (also called the soft scale; see below) proportional to $1/r$ avoids large logarithms of the form $\ln(\nu r)$ [5]. We call the choice $\nu = 1/r$ our standard choice for the renormalization scale. For this choice we have [5]

$$
\begin{aligned}
F(r) = \frac{C_F}{r^2}\alpha_s(1/r) \bigg[&1 + \frac{\alpha_s(1/r)}{4\pi}(\tilde{a}_1 - 2\beta_0) \\
&+ \frac{\alpha_s^2(1/r)}{(4\pi)^2}(\tilde{a}_2 - 4\tilde{a}_1\beta_0 - 2\beta_1) \\
&+ \frac{\alpha_s^3(1/r)}{(4\pi)^3}\bigg(\tilde{a}_3 - 6\tilde{a}_2\beta_0 - 4\tilde{a}_1\beta_1 - 2\beta_2 \\
&+ a_3^L \ln\frac{C_A\alpha_s(1/r)}{2}\bigg) + \mathcal{O}(\alpha_s^4, \alpha_s^4\ln^2\alpha_s) \bigg],
\end{aligned}
\tag{2}
$$

where the first three terms correspond to tree-level, one-loop, and two-loop order, respectively; the fourth term corresponds to three-loop order. The strong coupling, α_s, is understood, here and in the following, as renormalized in the $\overline{\text{MS}}$ scheme. At any time the strong coupling can be traded for the corresponding QCD scale. At three flavors in the $\overline{\text{MS}}$ scheme the QCD scale is $\Lambda_{\overline{\text{MS}}}^{N_f=3}$. The coefficients in Eq. (2) are defined in Ref. [13] and reproduced in

Appendix C. At three loops there is a contribution proportional to $\ln \alpha_s$ as well as a nonlogarithmic term. The origin of these terms can be understood using the effective field theory approach, namely, potential nonrelativistic QCD (pNRQCD) [14,15]. The term, pointed out in Ref. [16], proportional to $\ln \alpha_s$ comes from the energy scale $\sim \alpha_s/r$, which is called the ultrasoft scale to distinguish it from the energy scale $1/r$, which is referred to as the soft scale.

Equation (2) is accurate at next-to-next-to-next-to-leading order (N³LO).

For sufficiently small values of the coupling, i.e., at small enough distances the ultrasoft logarithm $\ln(C_A \alpha_s(1/r)/2)$ eventually becomes large and can be resummed. This can be done using the renormalization group equations in pNRQCD [17–19]. The expression for the force after performing the resummation of the ultrasoft logarithms reads [5]

$$F(r, \nu = 1/r) = \frac{C_F}{r^2} \alpha_s(1/r) \left[1 + \frac{\alpha_s(1/r)}{4\pi}(\tilde{a}_1 - 2\beta_0) + \frac{\alpha_s^2(1/r)}{(4\pi)^2}(\tilde{a}_2 - 4\tilde{a}_1\beta_0 - 2\beta_1) - \frac{\alpha_s^2(1/r)}{(4\pi)^2} \frac{a_3^L}{2\beta_0} \ln \frac{\alpha_s(\mu_{us})}{\alpha_s(1/r)} \right.$$
$$\left. + \frac{\alpha_s^2(1/r)\alpha_s(\mu_{us})}{(4\pi)^3} a_3^L \ln \frac{C_A \alpha_s(1/r)}{2r\mu_{us}} + \frac{\alpha_s^3(1/r)}{(4\pi)^3}(\tilde{a}_3 - 6\tilde{a}_2\beta_0 - 4\tilde{a}_1\beta_1 - 2\beta_2) + \mathcal{O}(\alpha_s^4) \right]. \quad (3)$$

We choose as a standard value for the ultrasoft scale $\mu_{us} = C_A \alpha_s(1/r)/(2r)$ where $C_A = N_c$. The coefficients in Eq. (3) are defined in Ref. [13] and reproduced in Appendix C.

At leading order in $\ln \alpha_s$ we can set μ_{us} to $1/r$ and Eq. (3) reduces to Eq. (2). The first four terms of Eq. (3) include all terms of the form $\alpha_s^{3+n} \ln^n \alpha_s$; i.e., it is accurate at next-to-next-to-leading logarithmic order. Keeping all terms in Eq. (3), we obtain an expression that is accurate at three-loop order with leading ultrasoft resummation.

In Fig. 4, we show the two-loop contribution to the force, i.e., the third term of Eq. (2), along with different α_s^4 contributions as a function of the distance r. Here and in the rest of the paper we use, if not differently specified, the standard scales $\nu = 1/r$ and $\mu_{us} = C_A \alpha_s(1/r)/(2r)$. One can see that the $\alpha_s^4 \ln \alpha_s$ contribution is never larger than
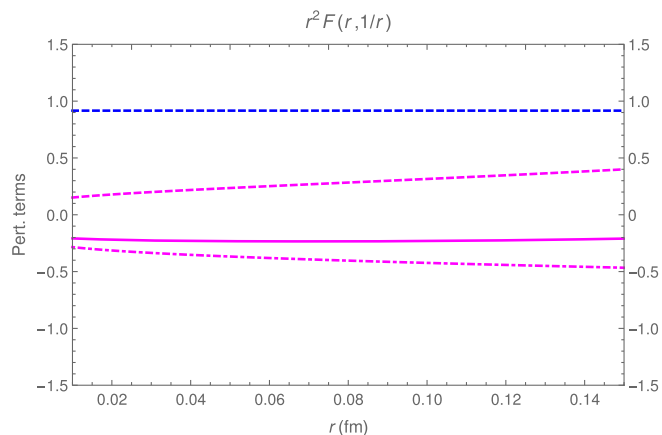

FIG. 4. The magenta lines stand for three-loop contributions: dashed [soft, fourth line of Eq. (2)], solid [ultrasoft, last line of Eq. (2)], and dot dashed [resummed ultrasoft, sum of fourth and fifth lines of Eq. (3)]. The blue dashed line stands for the two-loop contribution [third line of Eq. (2)]. A factor $\alpha_s(1/r)^3$ has been dropped from all terms. We use the standard value $\mu_{us} = C_A \alpha_s(1/r)/(2r)$ for the ultrasoft scale. We have taken $\Lambda_{\overline{MS}}^{N_f=3} = 315$ MeV.

the nonlogarithmic α_s^4 contribution in the entire distance range used in our study, and both contributions are smaller than the two-loop contribution. If we resum the ultrasoft logarithms [i.e., by replacing the last line of Eq. (2) with the fourth and fifth lines of Eq. (3)], the corresponding term becomes slightly larger in absolute value; see the dot-dashed line in Fig. 4, but it is still significantly smaller than the two-loop contribution. Therefore, as already noticed in [5], the leading ultrasoft logarithms can be counted as being of the same order as the other nonlogarithmic three-loop terms, while subleading logarithms may be included in the uncertainty due to the missing four-loop contributions. In Fig. 4, we see also that there is a partial cancellation between the soft and ultrasoft contributions at order α_s^4. This cancellation, however, may be accidental and may not happen at higher orders. Our final result will come from comparing lattice data with perturbation theory at three-loop order with leading ultrasoft resummation, and we use the difference between the N³LO expression of Eq. (2) and the three-loop expression with leading ultrasoft resummation from Eq. (3) to estimate the size of unknown higher-loop contributions. Other ways to estimate the higher-loop contributions are discussed in the following and enter eventually in our final error assessment.

We proceed as follows. On the lattice data side, we put all six data sets (see Table I) together by adding an arbitrary subtraction constant to each set that is determined by the fit. On the continuum side, we generate a grid of $\alpha_s(M_Z, N_f = 5)$ input values ranging from 0.1140 to 0.1200 in steps of 0.0001. We use the perturbative result for the force,[4] given in Eq. (3), restricting ourselves to the three-loop with leading

---

[4]We use perturbative running and decoupling to convert each grid value of $\alpha_s(M_Z, N_f = 5)$ to the corresponding $\Lambda_{\overline{MS}}^{N_f=3}$ value, which we use in the analytical expressions of the force, Eqs. (2) and (3). The details of the running and decoupling are the same as in Ref. [5]. Namely, we use four-loop running, with the charm quark mass equal to 1.6 GeV and the bottom quark mass equal to 4.7 GeV. The effects of higher-order terms in the running are negligible with the current accuracies.

TABLE II. Fits with $r/a \geq \sqrt{8}$ and different treatment of discretization artifacts, here, nonperturbative correction or only tree-level correction. We list uncorrelated and correlated χ^2/d.o.f. in columns 4 and 5. In column 9 we list the perturbative error for the corresponding fit window. The last column displays the outcome for α_s at two-loop order. The tree-level corrected calculations have fewer degrees of freedom, since the data with $r/a = \sqrt{12}$ have been excluded; see the text.

Artifacts	Max(r) [fm]	D.o.f.	Uncorrelated fits	Correlated fits	α_s	δ^{stat}	δ^{lat}	δ^{pert}	α_s^{2L}
Nonperturbative correction	0.097	10	0.33	0.47	0.11658	0.00072	0.00021	$^{+0.00152}_{-0.00045}$	0.11672
Tree-level correction	0.097	9	0.39	0.54	0.11679	0.00050	0.00021	$^{+0.00156}_{-0.00045}$	0.11696
Nonperturbative correction	0.131	42	0.33	0.40	0.11668	0.00051	0.00014	$^{+0.00193}_{-0.00064}$	0.11684
Tree-level correction	0.131	39	0.40	0.48	0.11673	0.00037	0.00014	$^{+0.00190}_{-0.00062}$	0.11697

ultrasoft resummation expression, for the reasons discussed above. For each of the α_s input values we calculate the static energy $E(r)$ by numerically integrating Eq. (3) and fit the lattice data to it, thus determining the subtraction constants. Hence, the subtraction constants for different sets are obtained in independent fits, since the different ensembles are statistically independent. Of course the subtraction constants depend on the input value of α_s. For the determination of the subtraction constants we consider fits with or without the full correlation matrix, which does not lead to statistically distinguishable results. The smallest χ^2/d.o.f. determines the preferred value of $\alpha_s(M_Z)$ by the lattice data. We calculate the χ^2/d.o.f. with or without the full correlation matrix. Typically, the minimum of the χ^2/d.o.f. varies by at most $\delta^{\text{corr}} = \pm 0.0001$ upon inclusion of the full correlation matrix, which is always smaller than the statistical error. We consider this difference as part of the statistical error that is already accounted for. The subtraction constants in the correlated fit are also consistent with the subtraction constants in the uncorrelated fits within fractions of the statistical error. We repeat this procedure for several values of max(r), vary the smallest distances min(r/a) in units of the lattice spacing considered in the analysis, and vary the treatment of discretization artifacts at the shortest distances. Eventually we repeat the entire procedure for expressions with different values of the soft scale, and for the two-loop and three-loop (without resumming the ultrasoft contribution) expressions.

Discretization artifacts are significant at distances $r/a \lesssim 3$, with increasing importance at smaller and smaller distances as detailed in Sec. II. We use different strategies to handle the residual discretization artifacts; see Appendix B for details.

(i) Ignoring them; i.e., we engage in no further corrections beyond the tree-level correction. In this case we restrict the fit to $r/a \geq \sqrt{8}$ and omit $r/a = \sqrt{12}$;

(ii) correcting for them, by following a strategy similar to Ref. [5], i.e., nonperturbative correction (see Appendix B);

(iii) accounting for them, by enlarging the statistical errors to fully cover the discretization artifacts, i.e., tree-level correction and systematic error estimates. This approach is suitable for $r/a \geq \sqrt{5}$.

First, we analyze the data for $r/a \geq \sqrt{8}$ without considering the discretization artifacts beyond the tree-level correction. In these fits we omit the data with $r/a = \sqrt{12}$ due to their exceptionally large discretization artifacts. We generally obtain χ^2/d.o.f. between 0.3 and 0.6 for the best fits, which rises steeply outside a narrow window around the central value of α_s. We estimate the systematic error due to discretization artifacts from the difference to fits with $r/a \geq \sqrt{8}$ with tree-level correction or with nonperturbative correction as being typically about $\delta^{\text{lat}} = \pm 0.00021$. Table II clearly shows that any errors due to discretization artifacts are safely contained in the statistical uncertainty. We point out that the analysis for max(r) ≤ 0.12 fm and $r/a \geq \sqrt{8}$ does not depend on gauge ensembles with the smaller sea quark mass, i.e., uses exclusively ensembles with $m_l = m_s/5$. Second, for $1 < r/a < \sqrt{8}$ the discretization artifacts are too large to be ignored. The data can still be used after correcting for the discretization artifacts through the nonperturbative correction. The χ^2/d.o.f. is about the same, unless data at large distances are discarded. In this case, the χ^2/d.o.f. becomes even smaller, due to the conservative error estimates for the corrections. We do not observe any larger deviation from the results for $r/a \geq \sqrt{8}$, and, hence, estimate that the uncertainty due to the corrections is generally about the same as for $r/a \geq \sqrt{8}$, namely, $\delta^{\text{lat}} = \pm 0.00021$. The discretization errors are safely contained in the statistical uncertainty. The results for fits in four representative fit intervals[5] using nonperturbatively corrected discretization errors are summarized in Table III. In each fit interval all results are consistent. We find slightly lower central values for fits with only ten data or less, which, however, are never statistically significant. We show the tree-level or nonperturbatively corrected lattice data together with the three-loop with leading ultrasoft resummation result with standard scales in Fig. 5. Given the large variation over the r range fine details are difficult to resolve. It can be seen that the fit to the data with large r/a misses the data with $r/a = 1$, i.e., the first data for each β, unless the

---

[5]We considered many more, but restrict the discussion to a representative set of fits.

TABLE III.   Representative fits using nonperturbatively corrected data in different intervals. We list uncorrelated and correlated $\chi^2/\text{d.o.f.}$ in columns 4 and 5. In column 9 we list the aggregate perturbative error for the corresponding fit window. The last column displays the outcome for α_s at two-loop order.

Min$(\frac{r}{a})$	Max(r) [fm]	D.o.f.	Uncorrelated fits	Correlated fits	α_s	δ^{stat}	δ^{lat}	δ^{pert}	α_s^{2L}
$\sqrt{1}$	0.055	6	0.14	0.18	0.11641	0.00051	0.00021	$+0.00081$ -0.00026	0.11638
$\sqrt{1}$	0.073	17	0.22	0.27	0.11660	0.00041	0.00021	$+0.00099$ -0.00031	0.11661
$\sqrt{1}$	0.098	36	0.31	0.34	0.11666	0.00034	0.00021	$+0.00121$ -0.00036	0.11672
$\sqrt{8}$	0.097	10	0.33	0.47	0.11658	0.00072	0.00021	$+0.00152$ -0.00045	0.11672
$\sqrt{1}$	0.131	74	0.35	0.40	0.11668	0.00027	0.00021	$+0.00154$ -0.00049	0.11683
$\sqrt{8}$	0.131	42	0.33	0.40	0.11668	0.00051	0.00021	$+0.00193$ -0.00064	0.11684

discretization artifacts are taken care of with an approach beyond the tree-level correction. For fits to tree-level corrected lattice data restricted to the range $\min(r/a) = \sqrt{8}$ and $\max(r) = 0.131$ fm the $\chi^2/\text{d.o.f.}$ values corresponding to $\alpha_s = 0.1162, 0.1167,$ and 0.1172 are 2.46, 0.49, and 2.12 (cf. fourth row of Table II). For fits to NPC lattice data restricted to the range $\min(r/a) = \sqrt{8}$ and $\max(r) = 0.131$ fm the $\chi^2/\text{d.o.f.}$ values corresponding to $\alpha_s = 0.1162, 0.1167,$ and 0.1172 are 0.97, 0.45, and 1.92 (cf. third row of Table II). Lastly, for fits over the whole range of nonperturbatively corrected lattice data with $\min(r/a) = 1$ and $\max(r) = 0.131$ fm the $\chi^2/\text{d.o.f.}$ values corresponding to $\alpha_s = 0.1162, 0.1167,$ and 0.1172 are 3.48, 0.41, and 4.12 (cf. fifth row of Table III). Thus, we conclude that the nonperturbative corrections have no effect if only data with $r/a \geq \sqrt{8}$ are considered, and the curve obtained only fitting the data with $r/a \geq \sqrt{8}$ also describes the data at shorter distances once the nonperturbative corrections are implemented. On the other hand, without the nonperturbative corrections the short distance data are off.

We show in Fig. 6 the differences between the normalized lattice data and the three-loop with leading ultrasoft resummation result with standard scales, either using nonperturbatively corrected data, or data with only tree-level correction. In the considered range the nonperturbative correction is irrelevant.

We show the differences between the normalized, nonperturbatively corrected lattice data and the three-loop with leading ultrasoft resummation result with standard scales in Fig. 7 for fits with $r/a \geq 1$. The difference is generally smaller than $0.007/r_1$ (about 4 MeV) and fluctuates around 0 with no apparent trend. At distances smaller than $0.15 r_1$, the difference tends to be marginally below 0 for $\max(r) = 0.055$ fm and slightly above 0
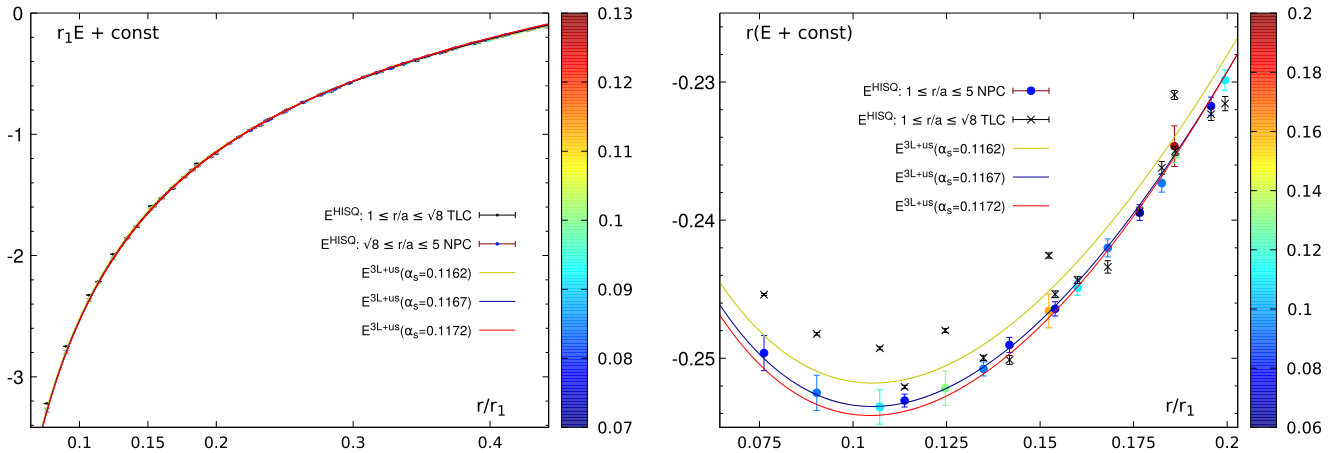

FIG. 5.   Left: Normalized lattice data and the weak-coupling result for the static energy in units of r_1. The colored or gray bullets show the nonperturbatively corrected (NPC) lattice data; the black crosses show the tree-level corrected (TLC) lattice data for $r/a < \sqrt{8}$. Only the colored symbols are used in the fit with $r/a \geq \sqrt{8}$. The three-loop with leading ultrasoft resummation result with standard scales is shown as lines for the $\alpha_s(M_Z)$ grid values corresponding to the best fit (0.1167), and the values which are one standard deviation lower or higher ($\alpha_s = 0.1162$ or 0.1172). The color palette corresponds to different values of a/r_1. Right: The normalized lattice data and the weak-coupling result for the static energy multiplied by r. Here the black symbols correspond to the uncorrected data, while the color symbols correspond to nonperturbatively corrected data. The color palette corresponds to different values of a/r_1.
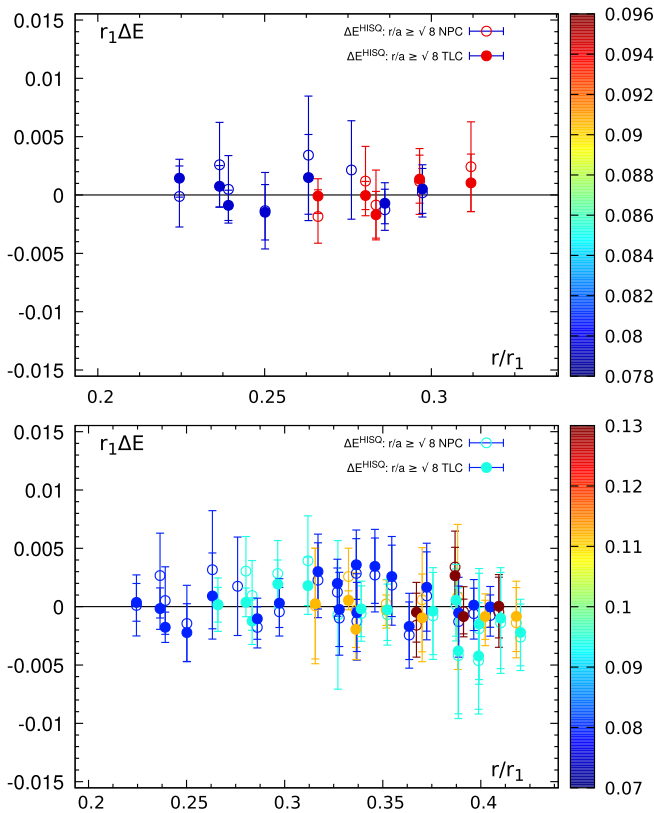
FIG. 6.   Difference between NPC lattice data, or the only TLC lattice data, and the weak-coupling result for the static energy in units of r_1 (after removal of the additive constant). The open circles show the nonperturbatively corrected lattice data, the filled bullets the only tree-level corrected lattice data. For the fits we have used $r/a \geq \sqrt{8}$. Differences between the central values of α_s are 0.00021 or smaller for the fits using different treatment of discretization artifacts. The two panels correspond to $\max(r) = 0.098$, and 0.131 fm.

otherwise. The central value of α_s is 0.00027 lower for fits restricted to the shortest distance interval ($\max(r) = 0.055$ fm) than for $\max(r) = 0.131$ fm. This difference may be indicative of the imperfect removal of the discretization artifacts, but is still compatible with the estimate of the discretization error using $\min(r/a) = \sqrt{8}$, and is more than covered by statistical uncertainties.

We perform another set of tests to investigate the systematic uncertainties of the lattice result. Namely, we test for a bias due to combining multiple lattice spacings into a single analysis that assumes that the data are within uncertainties consistent with the continuum limit. We separately analyze the data at the finest lattice spacing, $\beta = 8.4$. Here we also explore whether using data with $r/a \geq \sqrt{5}$ or $\sqrt{8}$ with errors that are inflated by estimates of the discretization artifacts has a significant impact on our results. We use the same uncertainty estimate for the influence of the discretization artifacts that we obtained previously, since it is larger than any estimates that we
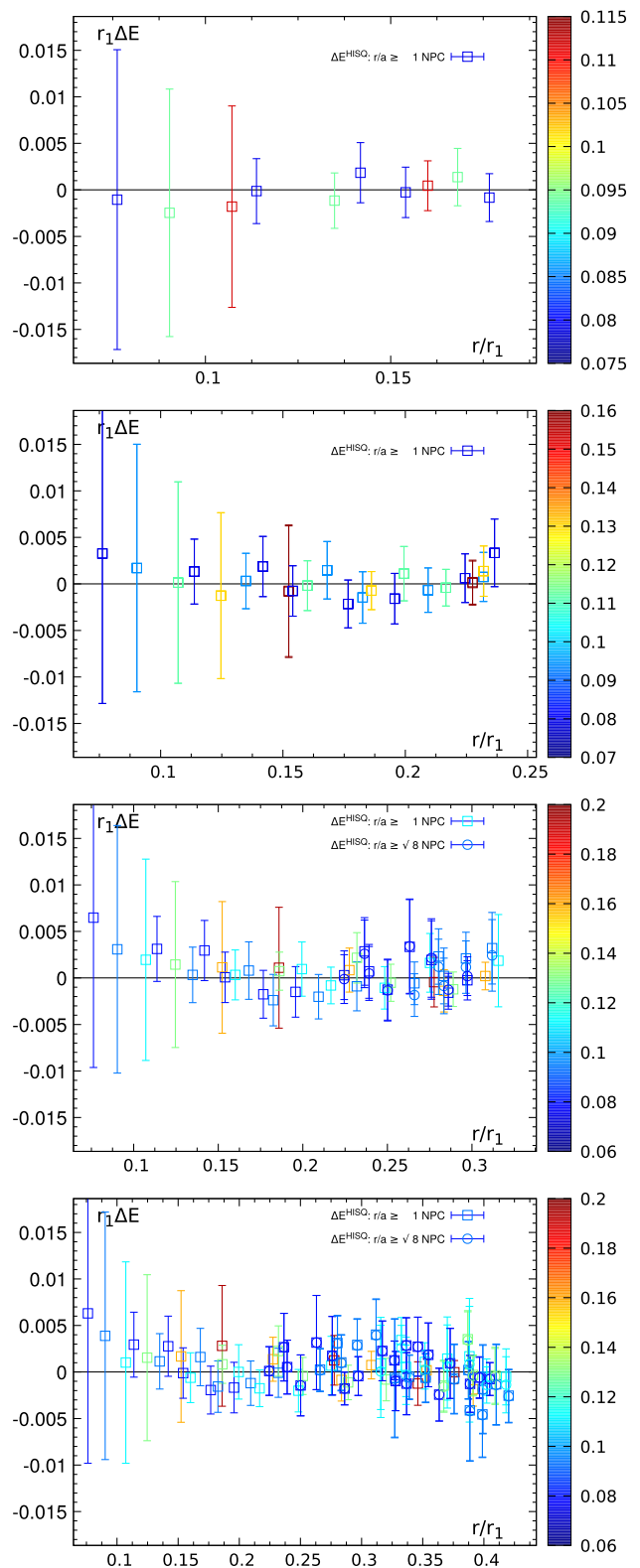

FIG. 7.   Difference between normalized, NPC lattice data and the weak-coupling result for the static energy in units of r_1. For $\max(r) < 0.056$ fm, inaccuracy of the nonperturbative correction appears to be responsible for a slight decrease of the central value of α_s. The four panels correspond to $\max(r) = 0.055$, 0.073, 0.098 and 0.131 fm (from top to bottom).

TABLE IV. Fits using only the finest lattice spacing, $\beta = 8.4$, and different treatment of discretization artifacts, i.e., with or without systematic error estimates and the tree-level correction. We list uncorrelated and correlated $\chi^2/\text{d.o.f.}$ in columns 5 and 6. In column 10 we list the perturbative error for the corresponding fit window. For the lattice error due to discretization artifacts we use the results of the analysis with multiple lattice spacings. The tree-level corrected calculations have fewer degrees of freedom, since the data with $r/a = \sqrt{12}$ have been excluded; see the text.

Artifacts	Min(r/a)	Max(r) [fm]	D.o.f.	Uncorrelated fits	Correlated fits	α_s	δ^{stat}	δ^{lat}	δ^{pert}
Tree-level correction with syst. errors	$\sqrt{5}$	0.098	9	0.29	0.36	0.11687	0.00040	0.00021	$^{+0.00146}_{-0.00043}$
Tree-level correction with syst. errors	$\sqrt{8}$	0.098	7	0.28	0.35	0.11685	0.00049	0.00021	$^{+0.00157}_{-0.00046}$
Tree-level correction	$\sqrt{8}$	0.098	6	0.50	0.60	0.11683	0.00035	0.00021	$^{+0.00155}_{-0.00043}$
Tree-level correction with syst. errors	$\sqrt{5}$	0.121	18	0.37	0.45	0.11687	0.00038	0.00021	$^{+0.00164}_{-0.00051}$
Tree-level correction with syst. errors	$\sqrt{8}$	0.121	16	0.38	0.46	0.11687	0.00044	0.00021	$^{+0.00176}_{-0.00055}$
Tree-level correction	$\sqrt{8}$	0.121	15	0.48	0.58	0.11687	0.00033	0.00021	$^{+0.00172}_{-0.00053}$

would obtain by comparing the results in this set of tests. Table IV shows that the separate analysis for the finest lattice spacing is consistent within statistical errors or within the even smaller estimate of the uncertainties due to the treatment of discretization artifacts. While the preferred central value in this analysis appears to be marginally higher than in the analysis that combines multiple spacings, this is not relevant for our final result.

To rule out the possibility that the nonperturbative correction of the lattice data at short distances introduces a systematic bias in the determination of α_s we also performed fits without implementing these corrections, but rather using these corrections as systematic error on the short distance data. This fit resulted in $\alpha_s = 0.1157(11)$ with $\chi^2/\text{d.o.f.} = 0.71$ or $\alpha_s = 0.1163(7)$ with $\chi^2/\text{d.o.f.} = 0.87$ for the first or second fit ranges [max(r) = 0.55 or 0.073 fm] in Table III, which is compatible with the other values within errors. An even better agreement is obtained for the larger values of max(r). Thus, we do not find any indication for a possible bias in α_s toward smaller values due to the nonperturbative correction, because the result for α_s without the nonperturbative corrections is never larger than with the nonperturbative correction.

In order to estimate the perturbative error of the small α_s expansion, i.e., the error made by neglecting contributions beyond three loops, we proceed similarly to Ref. [5] using three methods to quantify the uncertainty of the perturbative calculation. Namely, we add and subtract a term with the typical size of the soft four-loop contribution, we analyze the dependence on the soft scale ν, and we use the difference between the three-loop order with leading ultrasoft resummation and the three-loop result. We estimate the perturbative error in each method as the difference between $\alpha_s(M_Z)$ calculated in that way and the $\alpha_s(M_Z)$ value obtained from the three-loop with leading ultrasoft resummation expression at standard scales. In Ref. [5], the error due to the soft scale variation was estimated by taking $\nu = \sqrt{2}/r$ and $\nu = 1/\sqrt{2}r$. In this study, we enlarge the range of scale variation by considering $\nu = 2/r$ and

$\nu = 1/(2r)$. In doing so, we realize that the ν dependence of the extracted α_s is not monotonous: it has a minimum close to $\nu = 1/\sqrt{2}r$; see Fig. 8. For the lowest soft scale $\nu = 1/(2r)$ the use of perturbation theory becomes inadequate much earlier than for other choices. Namely, for max(r) $\gtrsim 0.1$ fm, the fits do not describe the data in a satisfying manner, and $\chi^2/\text{d.o.f.}$ increases much more rapidly for $\nu = 1/(2r)$ than for the other choices of the soft scale. This is expected on the grounds of the observation that the 2014 analysis showed that the scale $\nu = 1/(\sqrt{2}r)$ was adequate up to about $r = r_1/2$. If we exclude those larger distances the upper perturbative error is always given in terms of the soft scale variation to $\nu = 2/r$. We observe that for max(r) $\gtrsim 0.09$ fm the soft scale variation to $\nu = 1/(\sqrt{2}r)$ contributes to a larger estimate of the lower perturbative error than the soft higher-order terms, in line with the 2014 results. However, for smaller distances max(r) $\lesssim 0.09$ fm the estimate of the lower perturbative error due to inclusion of soft higher-order terms exceeds the estimate of the lower perturbative error due to soft scale
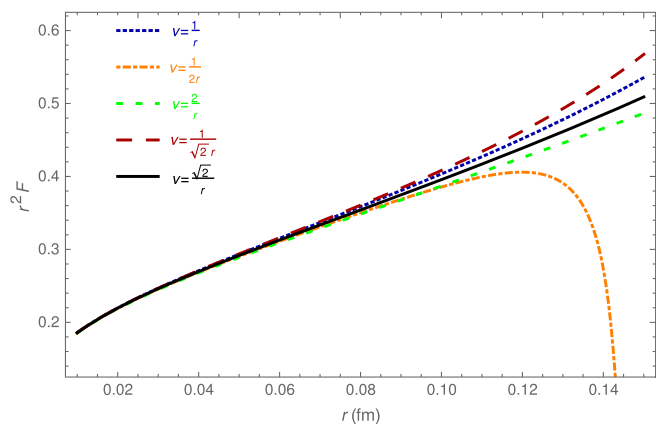

FIG. 8. The force for different values of the soft scale ν. The ultrasoft scale is kept fixed to $\mu_{us} = C_A \alpha_s(1/r)/(2r)$. We have taken $\Lambda^{N_f=3}_{\overline{\text{MS}}} = 315$ MeV.

variation. We take the most extreme among the error estimates from the soft contribution as the upper and lower perturbative errors due to the soft contribution. We also estimate the perturbative error by evaluating the difference between the three-loop with leading ultrasoft resummation result and the fixed-order three-loop result. The fixed-order three-loop result always yields a lower central value than the *three-loop with leading ultrasoft resummation* result. In order to be conservative, we apply the difference as a symmetric error. This has an even larger effect on the lower perturbative error than the soft scale variation or the soft higher-order terms. Moreover, the soft contribution to the lower perturbative error decreases for smaller values of $\max(r)$ more rapidly than the ultrasoft contribution. We stress that the lower perturbative error is always smaller than the statistical error for $r \lesssim 0.1$ fm. We add this symmetric perturbative error due to (inclusion or not of) ultrasoft resummation to the perturbative error due to the soft contribution in quadrature. The reasoning for adding these errors in quadrature is that it is unknown whether there is a similar partial compensation between soft and ultrasoft contributions at higher orders as it happens at three loops, and as such, we have to treat the corresponding uncertainties as independent. Lastly, we compare the three-loop with leading ultrasoft resummation and the two-loop results [first three lines of Eq. (2)]. As can be read off from Tables II and III, the difference to the two-loop result never exceeds +0.00025 and decreases for smaller values of $\max(r)$; i.e., it is smaller than the statistical errors and smaller than the other effects due variation of soft scale, soft higher-order terms, or variation of the ultrasoft resummation. The χ^2/d.o.f. does not change significantly between using the two-loop or three-loop with leading ultrasoft resummation results. Hence, we confirm the criterion for having the lattice data in the perturbative regime. We observe that the smaller $\max(r)$ is the smaller the variation of the central value of α_s between fits with different forms of the weak-coupling results becomes. Table III shows clearly that, for a given $\min(r/a)$, the perturbative errors are dramatically reduced at smaller distances, as expected, while the statistical error increases as fewer data are used to constrain the fits.

Let us summarize the considerations of the preceding paragraphs. We have to use $\max(r) \lesssim 0.1$ fm to perform the full scale variation and keep the perturbative uncertainties fully under control. We should ideally use significantly more than ten data point to limit the impact of the imperfectly treated discretization artifacts. Given the considerations of the preceding paragraphs, we take the result for $1 \le r/a \le 5$ and $\max(r) = 0.073$ fm, namely, $\alpha_s(M_Z) = 0.11660$ as our final result, which corresponds to $r_1 \Lambda_{\overline{\text{MS}}}^{N_f=3} = 0.4943$. The uncertainty of the scale r_1 is ± 0.0017 fm, which yields an error of $\delta^{\text{scale}} = \pm 1.7$ MeV for $\Lambda_{\overline{\text{MS}}}^{N_f=3}$, and $\delta^{\text{scale}} = \pm 0.00010$ for $\alpha_s(M_Z, N_f = 5)$. The

value $\alpha_s(1/r, N_f = 3)$ corresponding to the largest distance $\max(r) = 0.073$ fm used in this analysis yields $\alpha_s(1/\max(r), N_f = 3) = 0.2499$ and safely permits the scale variation by the factor 2 or even more. Therefore, the final result and full error budget of our zero temperature lattice calculation are given as

$$\alpha_s(M_Z, N_f = 5) = 0.11660^{+0.00110}_{-0.00056}, \qquad (4)$$

$$\delta\alpha_s(M_Z, N_f = 5) = (41)^{\text{stat}}(21)^{\text{lat}}(10)^{r_1}\left(^{+95}_{-13}\right)^{\text{soft}}(28)^{\text{us}}, \qquad (5)$$

or in terms of $\Lambda_{\overline{\text{MS}}}^{N_f=3}$ as

$$\Lambda_{\overline{\text{MS}}}^{N_f=3} = 314.0^{+15.5}_{-8.0} \text{ MeV}, \qquad (6)$$

$$\delta\Lambda_{\overline{\text{MS}}}^{N_f=3} = (5.8)^{\text{stat}}(3.0)^{\text{lat}}(1.7)^{r_1}\left(^{+13.4}_{-1.8}\right)^{\text{soft}}(4.0)^{\text{us}} \text{ MeV}. \qquad (7)$$

We have added the statistical error and the lattice discretization error of the static energy, the total error of the r_1 scale, and the perturbative error in quadrature.

In order to compare the current analysis to the previous analysis [5], we use the smaller window $[1/(\sqrt{2}r), \sqrt{2}/r]$ for the variation of the soft scale ν, and do not account for the uncertainty arising from the difference between resumming or not the leading ultrasoft logarithms to obtain

$$\delta^{\sqrt{2}}\alpha_s(M_Z, N_f = 5) = (41)^{\text{stat}}(21)^{\text{lat}}(10)^{r_1}\left(^{+37}_{-13}\right)^{\text{pert}}, \qquad (8)$$

$$\delta^{\sqrt{2}}\Lambda_{\overline{\text{MS}}}^{N_f=3} = (5.8)^{\text{stat}}(3.0)^{\text{lat}}(1.7)^{r_1}\left(^{+5.2}_{-1.8}\right)^{\text{pert}} \text{ MeV}. \qquad (9)$$

In this case, the perturbative error is not dominant anymore. Thus, the presented analysis has approximately halved the uncertainties of Ref. [5]. Nevertheless our final errors are only 10% (upper error) and 30% (lower error) smaller than the ones in [5], since we have accounted for the other possible sources of uncertainty listed above. The central values of Eq. (4) and of the final result in Ref. [5] coincide.

IV. EXTRACTING α_s FROM THE SINGLET FREE ENERGY

In this section, we consider the extraction of the strong coupling from the singlet free energy at nonzero temperature, as it is expected that at small distances medium effects are small. We define the singlet free energy in terms of the correlation function of two thermal Wilson lines in Coulomb gauge

$$F_S(r, T) = -T \ln\left(\frac{1}{N_c}\langle \text{Tr}[W(r)W^\dagger(0)]\rangle\right). \qquad (10)$$

At distances much smaller than the inverse temperature $rT \ll 1$, we can write using pNRQCD [4]

$$F_S(r,T) = V_s(r,\mu_{us}) + \delta F_S(r,T,\mu_{us}), \qquad (11)$$

where μ_{us} is the ultrasoft scale. The form of the thermal correction depends on the scale hierarchy. One could consider the case $1/r \gg \alpha_s/r \gg T \gg m_D \sim gT$ or the case $1/r \gg T \gg m_D \sim gT \gg \alpha_s/r$. In the former case $\mu_{us} \sim \alpha_s/r$ and $\delta F_S(r,T,\mu_{us}) = \delta E_{US}(\mu_{us}) + \Delta F_S(r,T)$ with $E_{US}(\mu_{us})$ being the ultrasoft contribution to the static energy in the vacuum. In the latter case $\mu_{us} \sim T$ and $\delta F_s(r,T,\mu_{us})$ has been calculated to order g^5; i.e., see Eqs. (16)–(19) in Ref. [3]. In the latter case the cancellation of the ultrasoft factorization scale dependence cannot be verified because of the unknown g^6 contribution to $\delta F_s(r,T,\mu_{us})$. Since $V_s(r,\mu_{us})$ has a term $\sim \alpha_s^3 \ln(\mu_{us}r)/r$, however, the difference between the $T = 0$ static energy and singlet free energy, $E(r) - F_S(r,T)$ should have a term $\sim \alpha_s^3 \ln(rT)/r$. This complicates the extraction of the strong coupling from the singlet free energy. The matching between NRQCD and pNRQCD also induces a term $\sim g^6 T$ in $F_S(r,T)$ for both scale hierarchies [4], which also needs to be considered.

The singlet free energy has been studied on the lattice in Ref. [3] using a wide temperature range and several lattice spacings, i.e., several temporal extents N_τ. The shortest distance that we can access, due to a single lattice spacing on our finest lattice at $T > 0$, is 0.00814 fm. For our analysis the relevant data correspond to $N_\tau = 10$, 12 and 16, since rT has to be small. From the analysis of Ref. [3] we know that thermal effects are small for $rT \lesssim 0.3$. To understand the temperature dependence of the singlet free energy in more detail we show the lattice results for the difference[6] between the singlet free energy at $T > 0$ and the static energy at $T = 0$ for $\beta = 8.4$ (corresponding to the finest zero temperature lattice) in Fig. 9. For other β values the results are similar.

From the figure we see that for $r/a \leq \sqrt{6}$ the difference approaches a constant proportional to the temperature according to the above expectations and no temperature effects beyond this constant can be seen in this range within the errors of the lattice results. Therefore, we conclude that in this regime the scale hierarchy $1/r \gg \alpha_s/r \gg T$ is appropriate. We treat the finite temperature data in this range as if at zero temperature and fit them with the three-loop with leading ultrasoft resummation result for the static

---

[6]The discretization artifacts between both quantities cancel exactly at tree level. Nonetheless, we still use the tree-level corrected distance in Fig. 9, in order to permit the visual distinction between data that are inequivalent in the full QCD result. Moreover, we matched the perturbative result to the lattice data at $r/a = 1$ using a constant that mimics the effect of the unknown $g^6 T$ term.
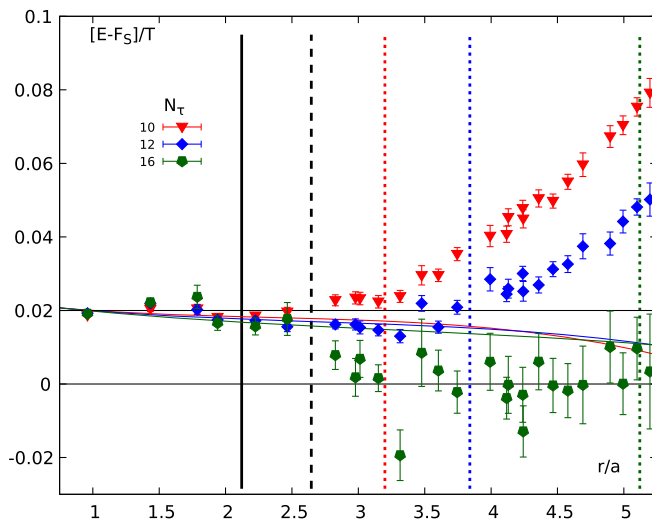

FIG. 9.   The difference between the static energy at $T = 0$ and the singlet free energy for $\beta = 8.4$ calculated with $N_\tau = 10$, 12 and 16 in units of the temperature as function of distance in units of the lattice spacing. The lines correspond to the $\delta F(r,T)$ calculated at order g^5. The dotted vertical lines show the boundary $rT = 0.3$ for different N_τ. The solid or dashed vertical lines show the boundaries between $r/a < \sqrt{5}$ and $r/a \geq \sqrt{5}$, or between $r/a < \sqrt{8}$ and $r/a \geq \sqrt{8}$, respectively.

energy with standard scales. Alternatively one could use the $N_\tau = 12$ data in the range where thermal effects appear to be small together with the two-loop result for the static energy, where the problem of the ultrasoft scale dependence does not enter at all. This, however, would lead to a larger theoretical uncertainty.

For distances $r/a > \sqrt{6}$ we see some temperature dependence. For the lattice data with $N_\tau = 12$ this temperature dependence is to some extent captured by the known g^5 result for $\delta F_S(r,T)$, but for $N_\tau = 10$ and 16 it is off. On the one hand, for $N_\tau = 16$, i.e., for a lower temperature, the temperature effects are larger than what is expected based on the g^5 result. Since the deviation of the temperature effects from a constant or from the g^5 result is quite similar to the size of the typical discretization artifacts or the statistical errors of the data for $r/a > \sqrt{6}$, it is unclear whether this is truly an effect due to the finite temperature. On the other hand, for $N_\tau = 10$, where the same temperature corresponds to a coarser lattice, the temperature effects are in the opposite direction and may be caused by temperature-dependent discretization errors [3].

The known g^5 result consists of two contributions with opposite sign but similar magnitude for the temperatures under consideration, and, therefore, the overall temperature effect is small due to cancellations. As these two contributions are due to either nonstatic or static Matsubara modes at orders g^4 or g^5, respectively, it is clear that this is an accidental cancellation in the temperature window of our lattice simulations. Furthermore, there is no reason to

expect that discretization artifacts in both contributions are similar enough to achieve a cancellation between them. In fact, it has been shown that the $\mathcal{O}(a^2)$ discretization errors appear to be more pronounced in the difference $E - F_S$ than in E or F_S individually [3].

For these reasons, we conclude that for $r/a > \sqrt{6}$ we do not have sufficient understanding of thermal effects to use F_S for extraction of α_s with fully controlled uncertainties. Nonetheless, the cancellations that appear to be at work both in the g^5 result and in the lattice data suggest that a fit of the lattice data with the zero temperature result may still be possible as a cross-check. Hence, we again treat the finite temperature data in this range as if at zero temperature and fit them with the three-loop with leading ultrasoft resummation result with standard scales.

In order to determine α_s from the singlet free energy at $T > 0$ we proceed as follows. We analyze the $T = 0$ static energy result ($N_\tau = 64$) in the same r/a intervals for which we expect that temperature effects in the singlet free energy are under control or are small due to accidental cancellations. We use the same weak-coupling result, namely, the three-loop with leading ultrasoft resummation result with

standard scales to obtain α_s from the singlet free energy. We estimate the uncertainty due to discretization artifacts to be the same as in the zero temperature analysis, i.e., $\delta^{\text{lat}} = \pm 0.00021$. Similarly, we estimate the uncertainty due to temperature effects from the difference of the central values of α_s for any combination of N_τ values in each fit interval.

First, we perform fits in the range where temperature effects are expected to be constant and the scale hierarchy $1/r \gg \alpha_s/r \gg T \gg m_D \sim gT$ applies. We report the result of representative fits with $r/a \leq 2$ in Table V, which corresponds to $rT \lesssim 0.17$ or 0.13 for $N_\tau = 12$ or 16, respectively. The central values of α_s in fits to the singlet free energy with $N_\tau = 12$ or 16 bracket the result for the static energy at zero temperature, with deviations being smaller than any other uncertainties. Hence, we conclude that we do not resolve nonconstant $T > 0$ effects as expected. As for the $T = 0$ analysis, we see a marginal decrease of the central value of α_s upon restriction to shorter distances. This decrease is about the same magnitude as in the $T = 0$ analysis, but is delayed to significantly shorter distances, which supports our interpretation that it

TABLE V. Fits with $r/a \leq 2$, or $r/a \leq 3$. We list uncorrelated and correlated $\chi^2/\text{d.o.f.}$ in columns 7 and 8. In column 13 we list the perturbative error for the corresponding fit window. The last column displays the outcome for α_s at two-loop order. δ^{pert} or α_s^{2L} for $N_\tau = 16$ were not computed because the $N_\tau = 16$ data are not used for the final α_s result. Note that the max(rT) and $\delta^{T>0}$ columns are not applicable for the $N_\tau = 64$ results.

Min($\frac{r}{a}$)	Max($\frac{r}{a}$)	N_τ	Max(rT)	Max(r) [fm]	D.o.f.	Uncorrelated fits	Correlated fits	α_s	δ^{stat}	δ^{syst}	$\delta^{T>0}$	δ^{pert}	α_s^{2L}
1	2	16	0.125	0.030	17	0.12	0.14	0.11621	0.00098	0.00021	0.00017	$\cdots$	$\cdots$
1	2	12	0.167	0.030	17	0.18	0.22	0.11638	0.00080	0.00021	0.00017	$+0.00043 \atop -0.00016$	0.11629
1	2	64	n/a	0.057	6	0.12	0.13	0.11646	0.00057	0.00021	n/a	$+0.00081 \atop -0.00026$	0.11643
1	2	16	0.125	0.057	30	0.19	0.20	0.11640	0.00089	0.00021	0.00006	$\cdots$	$\cdots$
1	2	12	0.167	0.057	30	0.16	0.18	0.11651	0.00074	0.00021	0.00005	$+0.00060 \atop -0.00021$	0.11644
1	2	64	n/a	0.078	12	0.15	0.16	0.11657	0.00051	0.00021	n/a	$+0.00100 \atop -0.00030$	0.11658
1	2	16	0.125	0.078	35	0.22	0.24	0.11651	0.00082	0.00021	0.00006	$\cdots$	$\cdots$
1	2	12	0.167	0.078	37	0.19	0.22	0.11663	0.00063	0.00021	0.00006	$+0.00081 \atop -0.00025$	0.11661
1	2	64	n/a	0.096	15	0.15	0.16	0.11659	0.00048	0.00021	n/a	$+0.00112 \atop -0.00031$	0.11663
1	2	12	0.167	0.091	41	0.22	0.25	0.11667	0.00063	0.00021	0.00008	$+0.00088 \atop -0.00027$	0.11667
1	3	16	0.1875	0.030	30	0.34	0.39	0.11658	0.00073	0.00021	0.00003	$\cdots$	$\cdots$
1	3	12	0.25	0.030	30	0.20	0.28	0.11661	0.00058	0.00021	0.00003	$+0.00046 \atop -0.00017$	0.11652
1	3	64	n/a	0.055	6	0.14	0.18	0.11641	0.00051	0.00021	n/a	$+0.00081 \atop -0.00026$	0.11638
1	3	16	0.1875	0.058	69	0.42	0.46	0.11672	0.00068	0.00021	0.00031	$\cdots$	$\cdots$
1	3	12	0.25	0.057	68	0.20	0.26	0.11671	0.00054	0.00021	0.00030	$+0.00062 \atop -0.00021$	0.11665
1	3	64	n/a	0.073	17	0.22	0.27	0.11660	0.00041	0.00021	n/a	$+0.00099 \atop -0.00031$	0.11661
1	3	16	0.1875	0.077	82	0.56	0.61	0.11682	0.00064	0.00021	0.00022	$\cdots$	$\cdots$
1	3	12	0.25	0.077	84	0.24	0.30	0.11680	0.00047	0.00021	0.00020	$+0.00078 \atop -0.00026$	0.11677
1	3	64	n/a	0.096	28	0.28	0.30	0.11665	0.00035	0.00021	n/a	$+0.00115 \atop -0.00035$	0.11670
1	3	16	0.1875	0.098	89	0.66	0.71	0.11686	0.00063	0.00021	0.00021	$\cdots$	$\cdots$
1	3	12	0.25	0.096	95	0.14	0.17	0.11682	0.00045	0.00021	0.00017	$+0.00090 \atop -0.00028$	0.11682
1	3	64	n/a	0.134	40	0.39	0.44	0.11668	0.00031	0.00021	n/a	$+0.00142 \atop -0.00045$	0.11680
1	3	12	0.25	0.133	109	0.29	0.32	0.11684	0.00040	0.00021	0.00016	$+0.00115 \atop -0.00037$	0.11690

is due to imperfections of the nonperturbative correction procedure. Within the smallest statistical uncertainties or within the even smaller estimates of the lattice discretization uncertainties, all of these results are consistent with the zero temperature analysis of the previous section, while the individual estimates of perturbative errors are systematically smaller; compare with Table III.

Next, we perform fits in the range where temperature effects beyond a constant appear to be small due to accidental cancellations. We assume the same scale hierarchy $1/r \gg \alpha_s/r \gg T \gg m_D \sim gT$, noting that it is at present not verifiable if the ultrasoft factorization scale dependence cancels in this case. Namely, we perform the fits with $r/a \leq 3$ and report the results in Table V as well.[7] This corresponds to $rT \lesssim 0.25$ or 0.19 for $N_\tau = 12$ or 16 respectively. $\chi^2/\text{d.o.f.}$ is practically unchanged for $N_\tau = 12$, but increases slightly for $N_\tau = 16$. Remarkably, the agreement with the previous $T = 0$ analysis, using up to $r/a \leq 5$, does not change significantly as $\max(r)$ becomes larger. This suggests that the differences between the results are not caused by effects of the finite temperature, but rather by the more severe influence of the imperfections of the nonperturbative correction procedure in the $T = 0$ analysis with $r/a \leq 2$ or $r/a \leq 3$. Hence, the estimate of the $\delta^{T>0}$ errors for $r > 0.05$ fm cannot be separated from the δ^{lat} errors due to lattice discretization uncertainties.

Let us summarize the considerations of the preceding paragraphs. For $T > 0$ we have to use $r/a \lesssim 2$ to guarantee the cancellation of ultrasoft factorization scale dependence, although we do not see any indication that it does not work for somewhat larger distances, i.e., $r/a \leq 3$. The weak-coupling picture suggests that cancellation between different medium effects is responsible for the small thermal modification of the singlet free energy. While the medium effects according to the weak-coupling picture are not unambiguously resolved in the data, the data seem to be affected by an accidental cancellation in the temperature window under consideration. Perturbative uncertainties are dramatically reduced at smaller distances $\max(r)$. In Fig. 10 we show our final $T = 0$ result with the reported $\chi^2/\text{d.o.f.}$ for two different fits to the singlet free energy data. The data show no nontrivial thermal effects in the ranges considered and are consistent with the central value of the $T = 0$ result.

Given the considerations of the preceding paragraphs, we take the $N_\tau = 12$ result for $1 \leq r/a \leq 2$ and $\max(r) = 0.030$ fm, namely, $\alpha_s(M_Z) = 0.11638$ as our final result, which corresponds to $r_1 \Lambda^{N_f=3}_{\overline{\text{MS}}} = 0.4900$. Furthermore, it has a largely independent error budget. We note that the finite temperature result for these distances does not depend on gauge ensembles corresponding to the larger sea quark

---

[7]Using an even larger fit range up to $r/a \leq \sqrt{12}$ leads to the same conclusions; i.e., the nonconstant thermal effects are still numerically irrelevant.
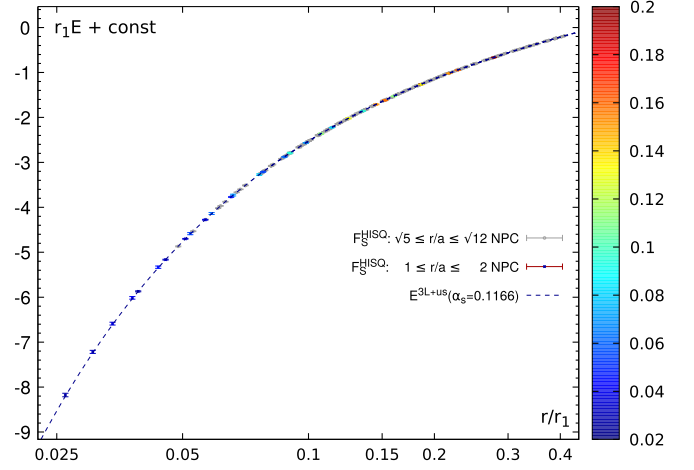

FIG. 10. Normalized lattice data and weak-coupling result for the static energy in units of r_1. We use a logarithmic scale for the coordinate axis. The colored or gray bullets show the NPC lattice data for $r/a \leq 2$ or $2 < r/a \leq \sqrt{12}$. The three-loop with leading ultrasoft resummation with standard scales is shown for the $\alpha_s(M_Z)$ grid values corresponding to the best fits for the r and rT intervals as indicated. The vertical lines of the same color indicate $\max(r)$ of the fits.

mass, i.e., uses exclusively ensembles with $m_l = m_s/20$. The scale uncertainty is the same as in the $T = 0$ analysis, $\delta^{\text{scale}} = \pm 1.7$ MeV for $\Lambda^{N_f=3}_{\overline{\text{MS}}}$, and $\delta^{\text{scale}} = \pm 0.00010$ for $\alpha_s(M_Z, N_f = 5)$. The value $\alpha_s(1/r, N_f = 3)$ corresponding to the largest distance $\max(r) = 0.030$ fm used in this analysis yields $\alpha_s(1/\max(r), N_f = 3) = 0.1820$ and safely permits the scale variation by a factor of 2 or even more. Therefore, the final result and full error budget of our finite temperature lattice calculation are given as

$$\alpha_s(M_Z, N_f = 5) = 0.11638^{+0.00095}_{-0.00087}, \qquad (12)$$

$$\delta\alpha_s(M_Z, N_f = 5)$$
$$= (80)^{\text{stat}}(21)^{\text{lat}}(17)^{T>0}(10)^{r_1}\left(^{+40}_{-06}\right)^{\text{soft}}(15)^{\text{us}}, \qquad (13)$$

or in terms of $\Lambda^{N_f=3}_{\overline{\text{MS}}}$ as

$$\Lambda^{N_f=3}_{\overline{\text{MS}}} = 310.9^{+13.5}_{-12.3} \text{ MeV}, \qquad (14)$$

$$\delta\Lambda^{N_f=3}_{\overline{\text{MS}}}$$
$$= (11.3)^{\text{stat}}(3.0)^{\text{lat}}(2.4)^{T>0}(1.7)^{r_1}\left(^{+5.6}_{-0.8}\right)^{\text{soft}}(2.1)^{\text{us}} \text{ MeV}. \qquad (15)$$

We have added the statistical error, the lattice discretization error of the static energy, the finite temperature error due to using the singlet free energy, the total error of the r_1 scale, and the perturbative error in quadrature.

V. DISCUSSION

In Secs. III and IV, we have presented two extractions of the strong coupling constant from two different observables, the static energy on zero temperature lattices, and the singlet free energy on finite temperature lattices.

Both results, i.e., $\alpha_s(M_Z, N_f = 5) = 0.11660^{+0.00110}_{-0.00056}$, and $\alpha_s(M_Z, N_f = 5) = 0.11638^{+0.00095}_{-0.00087}$, are in excellent agreement. This is despite the fact that they are based on several statistically independent sets of gauge ensembles corresponding to the QCD vacuum or the quark-gluon plasma and have been obtained using two different sea quark masses. The systematic uncertainties due to the lattice discretization and due to the lattice scale are not independent, but do not play a significant role in either of the two error budgets. Moreover, the two results are obtained in different r ranges, $[0.0237 \text{ fm}, 0.0734 \text{ fm}]$ or $[0.0081 \text{ fm}, 0.0301 \text{ fm}]$, which overlap only in a narrow window. As such, we may even consider the perturbative uncertainties of the two analyses as practically independent. We stress that every single fit in our analysis is contained already within the purely statistical errors of either the final $T = 0$ or the final $T > 0$ results as long as the soft scale is at $\nu = 1/r$.

There has been an observation in pure Yang-Mills theory that the approach to the continuum limit of $r_0 \Lambda^{N_f=0}_{\overline{\text{MS}}}$ extracted in the $Q\bar{Q}$-scheme[8] appears to change if data with $\alpha^3_{Q\bar{Q}} \lesssim 0.01$ are considered [20]. Namely, shorter distances corresponding to smaller $\alpha_{Q\bar{Q}}$ seemed to be related to smaller $r_0 \Lambda^{N_f=0}_{\overline{\text{MS}}}$ in the continuum limit, with about 10% lower continuum value in the pure Yang-Mills theory analysis. We show such a comparison for our data in Fig. 11. The fits with $\max(r) \lesssim 0.05$ fm are safely in the $\alpha^3_{Q\bar{Q}} \lesssim 0.01$ range, and are consistent with either a stable value or a modest drop. To account for possible artifacts in our analysis due to the use of small r/a, we have included the systematic error estimate $\delta^{\text{lat}} = \pm 0.00021$.

Our result from the static energy supersedes the result of the previous analysis [5], $\alpha_s(M_Z, N_f = 5) = 0.1166^{+0.0012}_{-0.0008}$, obtained using three gauge ensembles with light quark masses $m_l = m_s/20$. Although it reproduces the previous central value, it includes more conservative estimates of the known sources of systematic error. Nevertheless it succeeds in reducing the uncertainties through the use of finer lattice spacings.

In Ref. [12] the strong coupling was calculated from the static energy using two different analyses of spatially smeared Wilson loop data on three gauge ensembles with a pion mass of about 300 MeV, which are both consistent with our results. The presented two-step analysis (obtaining
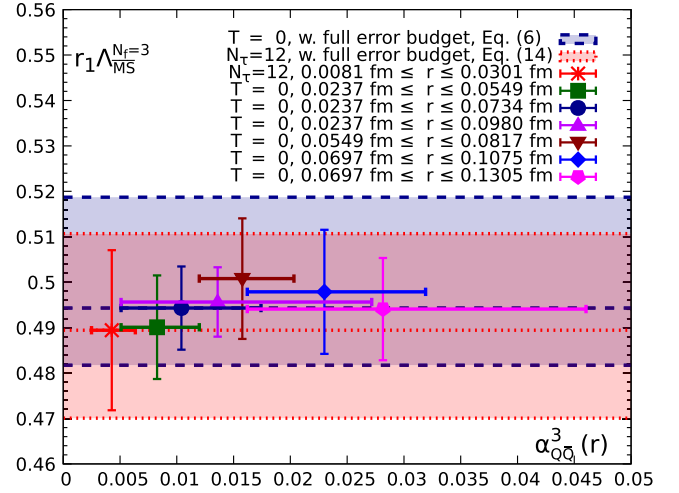

FIG. 11. The QCD scale $r_1 \Lambda^{N_f=3}_{\overline{\text{MS}}}$ as a function of $\alpha^3_{Q\bar{Q}}(r)$. The error bars of $r_1 \Lambda^{N_f=3}_{\overline{\text{MS}}}$ correspond to statistical errors only. For $\alpha^3_{Q\bar{Q}}(r)$ we use the value at $r = (\min(r) + \max(r))/2$ as the central value, and the error bars of $\alpha^3_{Q\bar{Q}}$ correspond to varying r from $\min(r)$ to $\max(r)$ (statistical errors are negligible). The data with $0.0549 \text{ fm} \leq r$ are selected from our $T = 0$ analysis with $r/a \geq \sqrt{5}$ using only $\beta = 8.4$; compare with Table IV. The data with $0.0697 \text{ fm} \leq r$ are selected from our $T = 0$ analysis with $r/a \geq \sqrt{8}$; compare with Table II. All other data sets are with $r/a \geq 1$. All individual results are in agreement with our final results, Eqs. (6) and (14), which are displayed as overlapping bands with the corresponding full error budgets.

the continuum limit first and then extracting the strong coupling) yielded $\alpha_s(M_Z, N_f = 5) = 0.1166^{+0.0021}_{-0.0020}$, while the second analysis using a single global fit yielded a higher value $\alpha_s(M_Z, N_f = 5) = 0.1179^{+0.0015}_{-0.0014}$.

Our result is slightly smaller than the PDG average [1] $\alpha_s(M_Z, N_f = 5) = 0.1181(11)$ as well as the recent FLAG average [21] $\alpha_s(M_Z, N_f = 5) = 0.11823(81)$. It is also smaller than the result of the ALPHA collaboration [22] $\alpha_s(M_Z, N_f = 5) = 0.11852(84)$, which uses step scaling and the Schrödinger functional method. These three values are, however, correlated. Our result is smaller than the results of the HPQCD collaboration using pseudoscalar quarkonium correlators or small Wilson loops [23,24] with three or four sea quark flavors, namely, $\alpha_s(M_Z, N_f = 5) = 0.1183(7)$, or $\alpha_s(M_Z, N_f = 5) = 0.11822(74)$, respectively. Our result agrees with the very recent lattice determination in terms of pseudoscalar quarkonium correlators [25], $\alpha_s(M_Z, N_f = 5) = 0.1159(12)$, or the preceding analysis of charmonium correlators [26], $\alpha_s(M_Z, N_f = 5) = 0.11622(83)$, or the analyses of higher quarkonium moments [27] $\alpha_s(M_Z, N_f = 5) = 0.1176(26)$. Moreover, our result is consistent within errors with the recent lattice determinations from the hadronic vacuum polarization [28] $\alpha_s(M_Z, N_f = 5) = 0.1181(27)^{+0.0008}_{-0.0022}$, or

---

[8]The gauge coupling in the $Q\bar{Q}$-scheme is defined as $\alpha_{Q\bar{Q}}(1/r) = r^2 F(r)/C_F$.

from the gauge-fixed gluon propagator in Landau gauge [29], $\alpha_s(M_Z, N_f = 5) = 0.1172(11)$.

Our result is also consistent with phenomenological estimates based on the bottomonium spectrum in NRQCD [30], $\alpha_s(M_Z, N_f = 5) = 0.1178(51)$, or quarkonia spectra in pNQRCD [31], $\alpha_s(M_Z, N_f = 5) = 0.1195(53)$, which have errors comparable to an earlier extraction from radiative bottomonium decays [32], $\alpha_s(M_Z, N_f = 5) = 0.119^{+0.006}_{-0.005}$. Finally, a recent study of the static energy using a subset of our data in a much larger r window, with an assumption similar to the one of Ref. [12] and approximations to the weak-coupling result that prevent the quantification of the uncertainties, finds $\alpha_s(M_Z, N_f = 5) = 0.1168$ [33], which agrees with our result.

VI. CONCLUSIONS

In this paper, we have calculated the static energy at zero temperature and the singlet free energy at finite temperature in $2 + 1$ flavor QCD using the HISQ action, and determined the strong coupling. We improved and extended the previous $2 + 1$ flavor HISQ analysis published in Ref. [5]. We included three finer lattice spacings, overhauled the analysis strategy, treated systematic uncertainties more conservatively, and restricted the analysis to much shorter distances that are deeper in the perturbative regime. Our main results are given in Eqs. (4)–(7). Moreover, for the first time we used the singlet free energy on finite temperature lattices to reach much smaller distances than ever used for the analysis of the static energy, obtaining an independent result with a complementary error budget given in Eqs. (12)–(15). Both results agree without even having to invoke the systematic errors of the weak-coupling calculation.

Both of our results for the central value of α_s are lower than a few lattice QCD determinations as well as the PDG average. However, the result from the static energy is compatible within errors. Furthermore, our results agree with some other lattice determinations of α_s based on moments of quarkonium correlators [25]. Finally, we stress that the extraction of α_s from the static energy is the only lattice extraction based, in the continuum part, on an observable known up to three loops and accurate at order α_s^4. This is due to the fact that the static energy is sensitive to α_s already at leading order.

ACKNOWLEDGMENTS

APPENDIX A: GAUGE ENSEMBLES AND CORRELATORS

In this appendix, we discuss additional ensembles that have not been summarized in Table I, since they are used only for auxiliary calculations, and fits to the primary lattice correlators. Further zero temperature ensembles that have only been used for developing a procedure to treat the discretization artifacts are summarized in Table VI. They also have been generated by the HotQCD collaboration [7,36] using the rational hybrid Monte Carlo (RHMC) algorithm and use $2 + 1$ flavors with the strange quark mass at its physical value and the mass of the two degenerate light quarks at 5% of the strange quark mass.

We have also used the gauge configurations generated for the study of quark number susceptibilities at high temperatures [37] and for the study of color screening [3] by the TUMQCD collaboration. The RHMC algorithm was used for $2 + 1$ quark flavors with the strange quark mass at its physical value and the mass of the two

TABLE VI. Parameters for the coarse $T = 0$ ensembles. In the seventh column we indicate the number of correlator measurements performed.

β	a (fm)	N_σ, N_τ	am_s	$m_\pi L$	#TUs	#MEAS	Ref.
$m_l = m_s/20$:							
6.664	0.117	32^4	0.0514	3.0	4000	400	[36]
6.740	0.109	48^4	0.0476	4.2	4000	800	[36]
6.800	0.103	32^4	0.0448	2.7	4000	400	[36]
6.880	0.095	48^4	0.0412	3.7	8200	820	[36]
6.950	0.089	32^4	0.0386	2.3	19400	1940	[36]
7.030	0.082	48^4	0.0356	3.2	7800	780	[36]
7.150	0.070	$64^3 \times 48$	0.0320	2.7	4047	667	[36]
7.280	0.065	$64^3 \times 48$	0.0284	2.5	3978	650	[36]

β	a [fm]	N_σ, N_τ	am_s	$m_\pi L$	#TUs	#MEAS	Ref.
$m_l = m_s/5$:							
7.030	0.082	48^4	0.0712	6.4	2990	598	[2]
7.825	0.040	64^4	0.01542	4.0	2980	298	[2]

TABLE VII. Parameters of $N_\tau = 12$ ensembles with aspect ratio 4 and $m_l = m_s/20$. Adjacent correlators are separated by 10 TUs.

β	am_s	T (MeV)	#TUs	Ref.
	$m_l = m_s/20$:			
7.373	0.0250	273	85120	[7]
7.596	0.0202	334	98010	[7]
7.650	0.0202	350	3230	[37]
7.825	0.0164	408	134600	[7]
8.570	0.008376	770	6320	[3]
8.710	0.007394	866	6490	[3]
8.850	0.006528	974	6340	[3]
9.060	0.004834	1162	7430	[3]
9.230	0.004148	1340	7280	[3]
9.360	0.003691	1495	7910	[3]
9.490	0.003285	1667	9780	[3]
9.670	0.002798	1938	7650	[3]

β	am_s	T (MeV)	#TUs	Ref.
	$m_l = m_s/5$:			
8.000	0.01299	474	71670	[3]
8.200	0.01071	562	71390	[3]
8.400	0.00887	667	71170	[3]

degenerate light quarks at 5% or 20% of the strange quark mass, i.e., corresponding to the zero temperature ensembles. We summarize the nonzero temperature gauge ensembles used in the determination of the strong coupling in Tables VII and VIII. We have increased the statistics of some $N_\tau = 16$ gauge ensembles specifically for this study and updated the Polyakov loop in Table VIII.

TABLE VIII. Parameters of the new $N_\tau = 16$ ensembles with aspect ratio 4 and expectation values of bare Polyakov loops. Adjacent correlators are separated by 10 TUs.

β	am_s	T (MeV)	#TUs	L^{bare}, Ref.
	$m_l = m_s/20$:			
7.825	0.00164	306	67960	0.002962(06)
8.570	0.008376	577	10400	[3]
8.710	0.007394	650	10190	[3]
8.850	0.006528	731	4480	[3]
9.060	0.004834	872	41870	0.015993(20)
9.230	0.004148	1005	3610	[3]
9.360	0.003691	1121	3530	[3]
9.490	0.003285	1250	6790	[3]
9.670	0.002798	1454	42060	0.025530(29)

β	am_s	T (MeV)	#TUs	L^{bare}
	$m_l = m_s/5$:			
8.000	0.001299	356	11460	[3]
8.200	0.001071	422	10660	[3]
8.400	0.000887	500	64370	0.007819(10)

TABLE IX. For most data sets the best fits in the ground state extraction were obtained with restricted fit windows due to instabilities at large τ/a.

β	7.373	7.596	7.825	8.000	8.200	8.400
r/a			Min(τ/a), Max(τ/a)			
(1,0,0)	9, 12	9, 11	7, 12	8, 11	7, 9	7, 12
(1,1,0)	6, 12	7, 12	8, 11	10, 12	8, 12	6, 12
(1,1,1)	6, 12	8, 12	8, 12	8, 12	7, 9	8, 12
(2,0,0)	7, 12	6, 12	10, 12	9, 12	10, 12	6, 9
(2,1,0)	6, 10	8, 10	7, 12	10, 12	7, 9	9, 12
(2,1,1)	7, 12	6, 12	6, 12	10, 12	8, 11	9, 12
(2,2,0)	6, 12	7, 10	8, 11	10, 12	8, 12	7, 9
(3,0,0)	8, 12	8, 12	7, 9	7, 12	6, 12	9, 12
(2,2,1)	8, 12	7, 10	7, 12	9, 12	6, 10	7, 12
(3,1,0)	6, 12	8, 10	9, 11	9, 12	8, 10	10, 12
(3,1,1)	6, 12	8, 12	7, 12	8, 11	7, 9	7, 10
(2,2,2)	7, 10	7, 9	8, 12	7, 12	8, 12	6, 9
(3,2,0)	5, 12	9, 12	7, 12	8, 12	8, 10	10, 12
(3,2,1)	6, 12	6, 12	9, 12	7, 10	8, 10	8, 12
(4,0,0)	7, 10	7, 12	7, 9	7, 10	7, 12	7, 9
(4,1,0)	6, 10	7, 12	9, 12	8, 12	8, 10	9, 12
(3,2,2)	8, 12	8, 10	6, 12	7, 10	8, 10	8, 10
(4,1,1)	7, 12	9, 12	8, 11	10, 12	8, 10	8, 12
(3,3,0)	9, 12	8, 12	10, 12	9, 12	8, 10	8, 12
(3,3,1)	8, 10	8, 11	8, 12	8, 12	8, 10	9, 12
(4,2,0)	7, 10	8, 10	7, 12	8, 12	8, 12	10, 12
(4,2,1)	9, 12	9, 12	7, 12	8, 12	9, 12	10, 12
(3,3,2)	9, 12	9, 12	7, 12	7, 12	9, 12	7, 12
(4,2,2)	8, 12	8, 12	10, 12	10, 12	10, 12	7, 12
(4,3,0)	7, 12	7, 10	7, 12	8, 10	10, 12	9, 12

In Table IX we summarize the necessary cuts to the fit ranges for the extraction of the static energy from the lattice correlators.

APPENDIX B: DISCRETIZATION ARTIFACTS

In this appendix, we briefly discuss the discretization artifacts of the static energy calculated on the lattice. We follow the treatment outlined in Ref. [3] with a few refinements.

It has been known for a long time that the static energy is affected by discretization artifacts at distances that are only slightly larger than the lattice spacing, i.e., $r/a \gtrsim 1$. The ratio of the static energy on the lattice and in the continuum exhibits a distinct pattern of a few percent to permill level variation around 1, which dies out quickly for larger distances; see Fig. 12. This pattern is quite similar in the tree-level calculation and in the QCD result (with or without sea quarks). Hence, one may partly account for these discretization artifacts by dividing the lattice QCD static energy by the ratio of the lattice and continuum static energies from a tree-level calculation, i.e., assuming one-gluon exchange and no running coupling. This procedure is

called tree-level correction. At tree level the static energy is written as

$$E_{\text{tree}}^{\text{lat}}(r) = -\frac{4}{3} g_0^2 \int \frac{d^3 k}{(2\pi)^3} e^{i\mathbf{k}\mathbf{r}} D_{00}(k), \qquad \text{(B1)}$$

with $D_{00}(k)$ being the temporal component of the lattice gluon propagator on the lattice. The results on the lattice static energy at tree level calculated for Wilson and Lüscher-Weisz action are shown in Fig. 12. Writing

$$E_{\text{tree}}^{\text{lat}}(r) = -\frac{4}{3} g_0^2 \frac{1}{4\pi r_I} \qquad \text{(B2)}$$

defines the so-called improved distance r_I. In practice, the tree-level correction is included in the comparison between continuum and lattice results either as

$$E_{\text{imp}}(r_I) = E^{\text{lat}}(r) \qquad \text{(B3)}$$

or—by introducing a fit parameter $\kappa_I(r/a, \beta)$ for each distinctive path geometry $\mathbf{r}/a$ between the quark and antiquark at each lattice spacing—as

$$E_{\text{imp}}(r) = E^{\text{lat}}(r) + \kappa_I\left(\frac{\mathbf{r}}{a}, \beta\right) \frac{r_I - r}{r r_I}. \qquad \text{(B4)}$$

Here we use the first approach.

However, this is not sufficient to remove the discretization artifacts entirely. Ideally, one would replace the tree-level correction by a one- or two-loop correction, i.e., calculating the correction factors using lattice perturbation theory at one- or two-loop level, and including the tadpole factors. We do not pursue this approach here as it requires extensive calculation beyond the scope of this study. As one can see from Fig. 12 discretization errors are at few percent level for $r/a \leq 2$. For $r/a = \sqrt{5}$ and $\sqrt{6}$ these errors are

below a percent and are even smaller for $r/a \geq \sqrt{8}$. Based on previous studies we expect the same pattern of discretization effects in the interacting theory. We show in Fig. 13 that a similar pattern is observed for the HISQ action after the tree-level correction. For distances $r/a \geq \sqrt{8}$ the statistical errors of the static energy are larger than the discretization errors.

Hence, there are three options to cope with the situation: The first option is to completely avoid the data with significant residual discretization artifacts. This corresponds to the first analysis strategy discussed in Sec. III.

The second option is to estimate the size of the residual discretization artifacts in the data at small r/a and combine a systematic error estimate of similar magnitude with the statistical errors to reduce the statistical weight of these data.

The first quark-antiquark distance that we can access with distinctive path geometries (ignoring permutations) is $r/a = 3$, specifically, either through $\mathbf{r}/a = (3, 0, 0)$ or $(2, 2, 1)$. We average these two results as $E^{\text{lat}}(r = 3a)$, and use a Cornell fit to interpolate between the data for $r_I(\sqrt{8}a), 3a$, and $r_I(\sqrt{10}a)$ to the distances $r_I((3, 0, 0)a) = 2.9788a$ and $r_I((2, 2, 1)a) = 3.0123a$. The difference $\Delta^{\text{lat}}(\mathbf{r}/a) = E^{\text{lat}}(\mathbf{r}/a) - E^{\text{int}}(\mathbf{r}/a)$ between the data and the interpolation at $\mathbf{r}/a = (3, 0, 0)$, or $(2, 2, 1)$ already includes the tree-level correction, i.e., $\Delta^{\text{lat}}((3, 0, 0)) - \Delta^{\text{lat}}((2, 2, 1))$ is an estimate for the residual discretization artifacts after the tree-level correction. On the other hand, $E^{\text{lat}}((3, 0, 0)) - E^{\text{lat}}((2, 2, 1))$ is an estimate for the discretization artifacts without the tree-level correction. The latter is usually about 0.1%, the former is usually 1 order of magnitude smaller. Since the latter certainly overestimates the residual discretization artifacts after tree-level correction and the former may be accidentally small due to statistical fluctuations, we proceed using the geometric average of the two. We smooth out fluctuations between
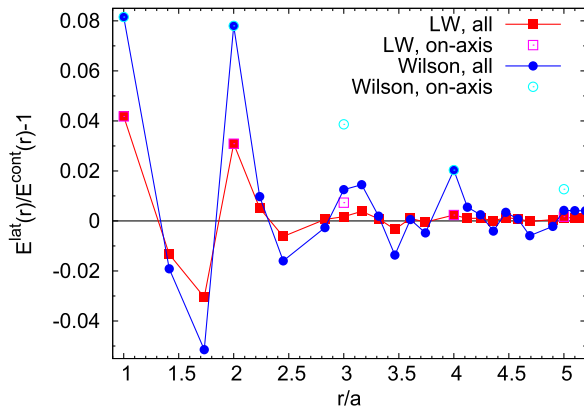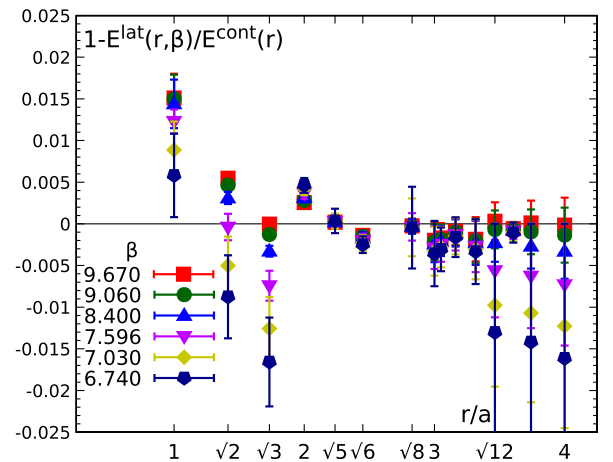

FIG. 12.   Cutoff effects in the tree-level lattice static energy with Lüscher-Weisz action are usually much smaller than for the Wilson gauge action. Both results are normalized by the corresponding continuum result.

FIG. 13.   Cutoff effects in the lattice static energy in $(2+1)$-flavor QCD obtained in this work after using tree-level correction.

TABLE X. Systematic uncertainty estimates for the discretization artifacts from the static energy at $T = 0$ or the singlet free energy with $N_\tau = 16$ or 12 at $r/a = 3$. The estimates have been calculated individually for each data set. The fluctuations within a factor 2 or 3 between the uncertainty estimates for the different data sets are due to statistical fluctuations of the data.

β	N_τ	$r_1 \delta_{\rm abs}^{\rm syst}$	$\delta_{\rm rel}^{\rm syst}(\frac{r}{a} = 3)$
7.596	64	0.0041	0.00103
7.825	64	0.0041	0.00093
8.000	64	0.0038	0.00076
8.200	64	0.0037	0.00066
8.400	64	0.0034	0.00054
7.825	16	0.0161	0.00360
8.000	16	0.0179	0.00362
8.200	16	0.0161	0.00290
8.400	16	0.0108	0.00172
8.570	16	0.0055	0.00079
8.710	16	0.0057	0.00074
8.850	16	0.0071	0.00086
9.060	16	0.0078	0.00082
9.230	16	0.0084	0.00079
9.360	16	0.0098	0.00085
9.490	16	0.0111	0.00089
9.670	16	0.0054	0.00038
7.596	12	0.0052	0.00128
7.650	12	0.0065	0.00156
7.825	12	0.0052	0.00116
8.000	12	0.0038	0.00078
8.200	12	0.0033	0.00060
8.400	12	0.0028	0.00045
8.570	12	0.0030	0.00044
8.710	12	0.0041	0.00054
8.850	12	0.0079	0.00095
9.060	12	0.0097	0.00102
9.230	12	0.0114	0.00108
9.360	12	0.0066	0.00057
9.490	12	0.0058	0.00046
9.670	12	0.0049	0.00035

these geometric averages for different lattice spacings by demanding that the systematic error estimate has to be at least as large as the mean of the systematic error for the two neighboring lattice spacings. We summarize the systematic error estimates for the different lattices in Table X. The absolute size of the estimate of the residual discretization error is typically less than 0.5% of $1/r_1$, namely, less than 3 MeV. This option is suitable for $r/a \geq \sqrt{5}$ at reasonable accuracy and should be applied for all distances with at most two steps along any of the lattice axes.

Finally, we may estimate the residual discretization artifacts in the data for each individual r/a, and correct for these errors before comparing lattice data with the weak-coupling results. This is the procedure that has been used in Refs. [3,5]. In this procedure, we define the correction factors

$$K\left(\frac{r}{a}, \beta\right) = \frac{E^{\rm lat}(r, a(\beta))}{E^{\rm cont}(r_I)} - 1, \qquad (B5)$$

and the nonperturbatively corrected static energy

$$E_{\rm imp}^{\rm lat}(r, \beta) = \frac{E^{\rm lat}(\frac{r}{a}, a(\beta))}{1 + K(\frac{r}{a}, \beta)}, \qquad (B6)$$

which is compared to the continuum static energy $E^{\rm cont}(r_I)$. In the following, we spell out how we estimate the residual discretization artifacts, and how we correct for these. We use two different estimates for the continuum $E^{\rm cont}(r_I)$.

(a) For the lattices under consideration and $r/a \geq \sqrt{5}$ the relative discretization artifacts are less than 0.1% after the tree-level correction. We interpolate the static energy on the finest lattices at such distances to obtain a continuum estimate[9] that can be used to take ratios of the QCD static energy on coarser lattices and in the continuum.

Our updated interpolation procedure differs in three ways from the procedure outlined in Ref. [3]. Before the interpolation we add a systematic error estimate of 1.5%, 0.5%, 0.5%, 0.5%, 0.04%, 0.02%, and 0.3% to the data on fine lattices at $r/a = 1, \sqrt{2}, \sqrt{3}, 2, \sqrt{5}, \sqrt{6}$, and $\sqrt{12}$, respectively. Furthermore, we modify the data before the interpolation by -1.0%, $+0.5\%$, -0.5%, and -0.3% for $r/a = 1, \sqrt{3}, 2$, and $\sqrt{12}$, respectively. These estimates of systematic errors are motivated by our previous analysis of the discretization artifacts at short distances [3]. We split the interpolation with a modified Cornell potential into two intervals, a shorter distance (usually seven points) and a larger distance interval, whose details are decided dynamically depending on the χ^2/d.o.f. of the fit, and switch between the two interpolations in an overlap region. In each interval we require fewer higher-order terms than for a single fit of all data. We use a spline interpolation to estimate the model dependence. In the next step, we calculate a weighted average $\langle K(r/a, \beta) \rangle$ of the ratios $K(r/a, \beta, \beta^{\rm ref})$ obtained with continuum estimates based on different reference lattices (indicated by $\beta^{\rm ref}$) for each r/a and β as previously.

Using the static energy at zero temperature restricted to $r/a \leq 8$ we vary the reference lattice down to $\beta^{\rm ref} \geq 7.28$ and obtain correction factors for all but the finest lattice ($\beta = 8.4$) down to $\beta \geq 6.664$. Using the singlet free energy with $N_\tau = 12$ restricted to $r/a \leq \sqrt{12}$, namely, $rT \leq 1/\sqrt{12} \approx 0.289$, we obtain correction factors for all but the finest lattice ($\beta = 9.67$) down to $\beta \geq 7.373$.

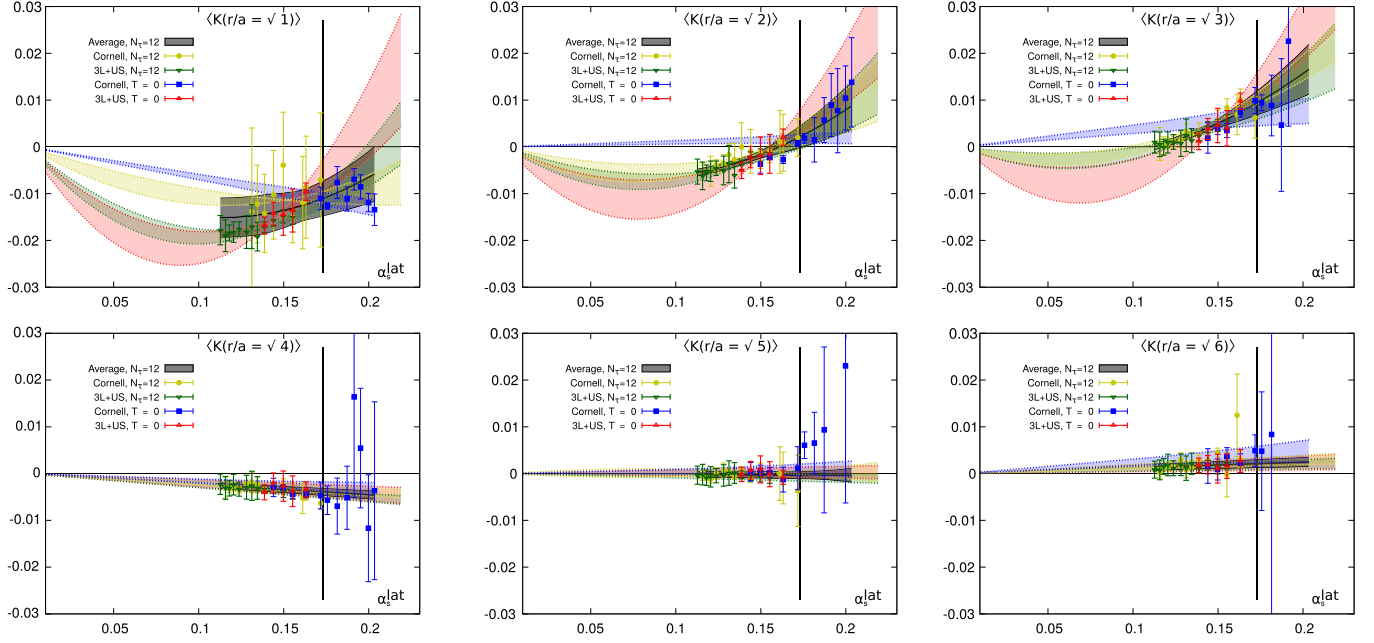[9]We extrapolate to 10% shorter distances; see Ref. [3].

FIG. 14.   The correction factors $\langle K \rangle$ for the nonperturbative correction of the static energy as a function of the boosted coupling $\alpha_s^{\text{lat}} = 10/(4\pi\beta u_0^4)$ for $r/a = \sqrt{1}$, $\sqrt{2}$, and $\sqrt{3}$ (top row), and $r/a = \sqrt{4}$, $\sqrt{5}$, and $\sqrt{6}$ (bottom row). The $\langle K \rangle$ are obtained for Cornell-type fits with $\min(r/a) = \sqrt{5}$ and for weak-coupling type fits with $\min(r/a) = \sqrt{1}$. $\langle K \rangle$ obtained through extrapolation as function of α_s^{lat} is indicated through the colored bands, while the final result is shown as the gray band. Only data to the left of the vertical line correspond to $\beta \geq 7.373$ and have been used in this study.

(b) We use the three loop result with the leading ultrasoft resummation as a fit form for a continuum estimate up to $r \leq 0.13$ fm. We restrict to this r range to ascertain that the fit is performed in the perturbative regime, which limits this analysis to fine lattice spacings, i.e., to $\beta \geq 7.596$ in practical terms. This fit form is physically more motivated but needs an input value for α_s. Before we fit a constant to the difference of the weak-coupling result and the lattice data, we add a systematic error estimate of 1.5%, 0.5%, 0.5%, 0.5%, 0.04%, 0.02%, and 0.3% to the data at $r/a = 1, \sqrt{2}, \sqrt{3}, 2, \sqrt{5}, \sqrt{6}$, and $\sqrt{12}$, respectively. In order to minimize the bias towards a final determination of α_s, we vary the range of the input value α_s between 0.1155 and 0.1175, and determine for each input the normalization constant through a weighted average. Out of the set of inputs we accept either all of the fits with $\chi^2/\text{d.o.f.} \leq 1$ or the three fits with smallest $\chi^2/\text{d.o.f.}$. We estimate the uncertainty of the correction factor by adding the statistical error and the variation of the continuum estimate with the input value of α_s in quadrature.

Using the static energy at zero temperature restricted to $r/a \leq 8$ we obtain correction factors for all lattices down to $\beta \geq 7.596$. Using the singlet free energy with $N_\tau = 12$ restricted to $r/a \leq \sqrt{12}$, namely, $rT \leq 1/\sqrt{12} \approx 0.289$, we obtain correction factors for all lattices down to $\beta \geq 7.825$.

All four estimates of the correction factors are complementary, since they have different systematic uncertainties. In Fig. 14 we show the corresponding correction factors from Cornell fits using lattice data at $r/a \geq \sqrt{5}$ on the reference lattice, or from fits with the weak-coupling result using $r/a \geq \sqrt{1}$. All four results are in very good agreement.

We extrapolate the $\langle K(r/a, \beta) \rangle$ towards the continuum limit using a boosted coupling $\alpha_s^{\text{lat}} = 10/(4\pi\beta u_0^4)$. We list the corresponding tadpole factors u_0 in Table XI, which

TABLE XI.   Tadpole factors u_0 for the lattices used in the correction procedure. Tadpole factors have been obtained using zero temperature or finite temperature lattices. Temperature effects are always at most 0.00001.

β	u_0	β	u_0	β	u_0
6.608	0.86823	7.030	0.88173	8.200	0.90604
6.664	0.87026	7.100	0.88363	8.400	0.90909
6.700	0.87152	7.150	0.88493	8.570	0.91152
6.740	0.87288	7.200	0.88621	8.710	0.91341
6.770	0.87388	7.280	0.88817	8.850	0.91522
6.800	0.87485	7.373	0.89035	9.060	0.91778
6.840	0.87613	7.596	0.89517	9.230	0.91973
6.880	0.87736	7.650	0.89627	9.360	0.92116
6.910	0.87827	7.825	0.89962	9.490	0.92253
6.950	0.87945	8.000	0.90274	9.670	0.92435

have been obtained from the plaquette at $T = 0$, or with $N_\tau = 12$ or 16.

For the extrapolation we use

$$\left\langle K\left(\frac{r}{a}, \alpha_s^{\text{lat}}\right)\right\rangle = \sum_{i=1}^{N} \kappa_i\left(\frac{r}{a}\right)(\alpha_s^{\text{lat}})^i. \quad \text{(B7)}$$

We try extrapolations with up to $N = 3$ terms, but never require the third-order term; i.e., we cannot fix its coefficient with small uncertainties. While data are quite insensitive to the second-order term for $r/a \geq 2$, it cannot be avoided for $r/a < 2$. The coefficients of the first- and second-order term are anticorrelated. Only in the case of the Cornell fit to lattice data at zero temperature we cannot resolve the second-order term successfully. We also tested fits which include the zeroth-order term, but it turned out that this term was never necessary.

For each estimate (Cornell vs weak-coupling fits and zero vs finite temperature) we take the median of the correction factors from fits including data starting at $\min(r/a) = \sqrt{1}$ up to $\min(r/a) = \sqrt{5}$. We estimate the overall error from typical uncertainties, or the spread between results for different $\min(r/a)$, whichever is larger. We show the median of the extrapolations for each fit type as function of α_s^{lat} as colored bands in Fig. 14. While the agreement between the median of Cornell fits and the median of weak-coupling fits is not satisfactory for fits to $T = 0$ data, the different fit results to $T > 0$ data agree very well. Moreover, for the correction factor $\langle K(r/a < 2, \beta)\rangle$, where the data are sensitive to the second-order term in α_s^{lat}, the extrapolation of the results obtained with $T > 0$ data also extends to the $T = 0$ results at larger values of the

coupling, i.e., on coarse lattices. Therefore, in order to obtain our final results, we take the mean of the Cornell fit and the weak-coupling fit to the $N_\tau = 12$ data, and estimate the uncertainty by the largest among the errors of the two results, or half of the spread between the results. Since this uncertainty still underestimates the spread between $\langle K(r/a = 1, \beta)\rangle$ for different fit types and among different β, we add a systematic error estimate of 0.003 in quadrature to the errors of this procedure. We summarize the coefficients corresponding to both fit types in Table XII and show the averaged result as a gray band in Fig. 14. While the corrections are significant and constrained quite well up to $r/a = (2, 1, 1)$, the corrections beyond tree level are consistent with 0 for larger distances.

APPENDIX C: COEFFICIENTS OF THE FORCE AT THREE LOOPS

In this appendix, we summarize the constants appearing in the weak-coupling result of the force at three-loop order in Eqs. (2) and (3). The color factors read

$$C_F = \frac{N_c^2 - 1}{2N_c}, \qquad C_A = N_c, \qquad T_F = \frac{1}{2}, \quad \text{(C1)}$$

where N_c is the number of colors. The QCD beta function is

$$\frac{d\alpha_s(\nu)}{d\ln\nu} = \alpha_s\beta(\alpha_s) = -\frac{\alpha_s^2}{2\pi}\sum_{n=0}^{\infty}\left(\frac{\alpha_s}{4\pi}\right)^n\beta_n$$

$$= -2\alpha_s\left[\beta_0\frac{\alpha_s}{4\pi} + \beta_1\left(\frac{\alpha_s}{4\pi}\right)^2 + \cdots\right], \quad \text{(C2)}$$

where the first three coefficients read [38,39]

$$\beta_0 = \frac{11}{3}C_A - \frac{4}{3}T_F N_f, \quad \text{(C3)}$$

$$\beta_1 = \frac{34}{3}C_A^2 - \frac{20}{3}C_A N_f T_F - 4C_F N_f T_F, \quad \text{(C4)}$$

$$\beta_2 = \frac{2857}{54}C_A^3 + \left(-\frac{1415}{27}C_A^2 - \frac{205}{9}C_A C_F + 2C_F^2\right)N_f T_F$$

$$+ \left(\frac{158}{27}C_A + \frac{44}{9}C_F\right)N_f^2 T_F^2. \quad \text{(C5)}$$

N_f is the number of active massless quarks. Higher-order coefficients β_3, and β_4 are also known [38,39], but are not reproduced for brevity's sake. We reproduce the coefficients $\tilde{a}_i$ as given in Ref. [13],

TABLE XII. Coefficients of the nonperturbative correction procedure up to order $(\alpha_s^{\text{lat}})^2$.

	Tree level	Cornell		Three loop	
r/a	$\kappa_0 = \frac{r_I - r}{r_I}$	κ_1	κ_2	κ_1	κ_2
(1,0,0)	−0.041777	−0.166(49)	0.60(33)	−0.364(56)	1.73(39)
(1,1,0)	+0.013536	−0.142(55)	0.90(39)	−0.183(50)	1.12(37)
(1,1,1)	+0.031486	−0.099(34)	0.92(24)	−0.102(51)	0.87(38)
(2,0,0)	−0.029918	−0.020(04)	0.0	−0.026(04)	0.0
(2,1,0)	−0.005012	+0.001(23)	0.0	−0.004(05)	0.0
(2,1,1)	+0.006374	+0.015(05)	0.0	+0.010(05)	0.0
(2,2,0)	−0.000432	+0.009(47)	0.0	−0.004(05)	0.0
(2,2,1)	+0.004098	+0.026(13)	0.0	+0.004(04)	0.0
(2,2,2)	+0.003694	+0.016(50)	0.0	−0.171(08)	1.47(58)
(3,0,0)	−0.007952	+0.033(27)	0.0	+0.002(50)	0.0
(3,1,0)	−0.003564	+0.020(08)	0.0	−0.004(05)	0.0
(3,1,1)	−0.000562	+0.034(19)	0.0	−0.001(05)	0.0
(3,2,0)	−0.000657	+0.010(53)	0.0	+0.010(06)	0.0
(3,2,1)	+0.000883	+0.002(49)	0.0	−0.177(64)	1.56(48)
(4,0,0)	−0.001854	−0.022(45)	0.0	−0.191(80)	1.72(63)

$$\tilde{a}_1 = a_1 + 2\gamma_E\beta_0, \tag{C6}$$

$$\tilde{a}_2 = a_2 + \left(\frac{\pi^2}{3} + 4\gamma_E^2\right)\beta_0^2 + \gamma_E(4a_1\beta_0 + 2\beta_1), \tag{C7}$$

$$\tilde{a}_3 = a_3 + (8\gamma_E^3 + 2\gamma_E\pi^2 + 16\zeta(3))\beta_0^3 + 2\gamma_E\beta_2$$
$$+ [(12\gamma_E^2 + \pi^2)\beta_0^2 + 4\gamma_E\beta_1]a_1$$
$$+ \left[6a_2\gamma_E + \frac{5}{2}\left(4\gamma_E^2 + \frac{\pi^2}{3}\right)\beta_1\right]\beta_0, \tag{C8}$$

where the coefficients a_1 and a_2 read [40–42]

$$a_1 = \frac{31}{9}C_A - \frac{20}{9}T_FN_f, \tag{C9}$$

$$a_2 = \left(\frac{4343}{162} + 4\pi^2 - \frac{\pi^4}{4} + \frac{22}{3}\zeta(3)\right)C_A^2$$
$$- \left(\frac{1798}{81} + \frac{56}{3}\zeta(3)\right)C_AT_FN_f$$
$$- \left(\frac{55}{3} - 16\zeta(3)\right)C_FT_FN_f + \left(\frac{20}{9}T_FN_f\right)^2. \tag{C10}$$

The coefficient a_3 reads [43,44]

$$a_3 = a_3^{(3)}N_f^3 + a_3^{(2)}N_f^2 + a_3^{(1)}N_f + a_3^{(0)}, \tag{C11}$$

$$a_3^{(3)} = -\left(\frac{20}{9}\right)^3 T_F^3$$

$$a_3^{(2)} = \left(\frac{12541}{243} + \frac{368}{3}\zeta(3) + \frac{64\pi^4}{135}\right)C_AT_F^2$$
$$+ \left(\frac{14002}{81} - \frac{416}{3}\zeta(3)\right)C_FT_F^2 \tag{C12}$$

$$a_3^{(1)} = (-709.717)C_A^2T_F$$
$$+ \left(-\frac{71281}{162} + 264\zeta(3) + 80\zeta(5)\right)C_AC_FT_F$$
$$+ \left(\frac{286}{9} + \frac{296}{3}\zeta(3) - 160\zeta(5)\right)C_F^2T_F$$
$$+ (-56.83(1))\frac{18 - 6N_c^2 + N_c^4}{96N_c^2} \tag{C13}$$

$$a_3^{(0)} = 502.24(1)C_A^3 - 136.39(12)\frac{N_c^3 + 6N_c}{48}$$
$$+ \frac{8}{3}\pi^2C_A^3\left(-\frac{5}{3} + 2\gamma_E + 2\log 2\right). \tag{C14}$$

The three numerical values given above are nowadays known analytically [45,46]. Lastly, we reproduce the coefficient a_3^L as given in Ref. [15],

$$a_3^L = \frac{16\pi^2}{3}C_A^3. \tag{C15}$$

[1] M. Tanabashi *et al.* (Particle Data Group), Phys. Rev. D **98**, 030001 (2018).

[2] A. Bazavov, P. Petreczky, and J. H. Weber, Phys. Rev. D **97**, 014510 (2018).

[3] A. Bazavov, N. Brambilla, P. Petreczky, A. Vairo, and J. H. Weber (TUMQCD Collaboration), Phys. Rev. D **98**, 054511 (2018).

[4] M. Berwein, N. Brambilla, P. Petreczky, and A. Vairo, Phys. Rev. D **96**, 014025 (2017).

[5] A. Bazavov, N. Brambilla, X. Garcia i Tormo, P. Petreczky, J. Soto, and A. Vairo, Phys. Rev. D **90**, 074038 (2014).

[6] E. Follana, Q. Mason, C. Davies, K. Hornbostel, G. P. Lepage, J. Shigemitsu, H. Trottier, and K. Wong (HPQCD, UKQCD Collaborations), Phys. Rev. D **75**, 054502 (2007).

[7] A. Bazavov *et al.* (HotQCD Collaboration), Phys. Rev. D **90**, 094503 (2014).

[8] J. H. Weber, A. Bazavov, and P. Petreczky, Proc. Sci Confinement2018 (**2018**) 166.

[9] A. Hasenfratz and F. Knechtli, Phys. Rev. D **64**, 034504 (2001).

[10] A. Bazavov, N. Brambilla, X. Garcia i Tormo, P. Petreczky, J. Soto, and A. Vairo, Phys. Rev. D **86**, 114031 (2012).

[11] H. Takaura, T. Kaneko, Y. Kiyo, and Y. Sumino, Phys. Lett. B **789**, 598 (2019).

[12] H. Takaura, T. Kaneko, Y. Kiyo, and Y. Sumino, J. High Energy Phys. 04 (2019) 155.

[13] X. Garcia i Tormo, Mod. Phys. Lett. A **28**, 1330028 (2013).

[14] N. Brambilla, A. Pineda, J. Soto, and A. Vairo, Nucl. Phys. **B566**, 275 (2000).

[15] N. Brambilla, A. Pineda, J. Soto, and A. Vairo, Phys. Rev. D **60**, 091502 (1999).

[16] T. Appelquist, M. Dine, and I. J. Muzinich, Phys. Lett. **69B**, 231 (1977).

[17] A. Pineda and J. Soto, Phys. Lett. B **495**, 323 (2000).

[18] N. Brambilla, A. Pineda, J. Soto, and A. Vairo, Rev. Mod. Phys. **77**, 1423 (2005).

[19] N. Brambilla, A. Vairo, X. Garcia i Tormo, and J. Soto, Phys. Rev. D **80**, 034016 (2009).

[20] N. Husung, M. Koren, P. Krah, and R. Sommer, EPJ Web Conf. **175**, 14024 (2018).

[21] S. Aoki *et al.* (Flavour Lattice Averaging Group), arXiv:1902.08191.

[22] M. Bruno, M. Dalla Brida, P. Fritzsch, T. Korzec, A. Ramos, S. Schaefer, H. Simma, S. Sint, and R. Sommer (ALPHA Collaboration), Phys. Rev. Lett. **119**, 102001 (2017).

[23] C. McNeile, C. T. H. Davies, E. Follana, K. Hornbostel, and G. P. Lepage, Phys. Rev. D **82**, 034512 (2010).

[24] B. Chakraborty, C. T. H. Davies, B. Galloway, P. Knecht, J. Koponen, G. C. Donald, R. J. Dowdall, G. P. Lepage, and C. McNeile, Phys. Rev. D **91**, 054508 (2015).

[25] P. Petreczky and J. H. Weber, Phys. Rev. D **100**, 034519 (2019).

[26] Y. Maezawa and P. Petreczky, Phys. Rev. D **94**, 034507 (2016).

[27] K. Nakayama, B. Fahy, and S. Hashimoto, Phys. Rev. D **94**, 054507 (2016).

[28] R. J. Hudspith, R. Lewis, K. Maltman, and E. Shintani, arXiv:1804.10286.

[29] S. Zafeiropoulos, P. Boucaud, F. De Soto, J. Rodríguez-Quintero, and J. Segovia, Phys. Rev. Lett. **122**, 162002 (2019).

[30] V. Mateu and P. G. Ortega, J. High Energy Phys. 01 (2018) 122.

[31] C. Peset, A. Pineda, and J. Segovia, J. High Energy Phys. 09 (2018) 167.

[32] N. Brambilla, X. Garcia i Tormo, J. Soto, and A. Vairo, Phys. Rev. D **75**, 074014 (2007).

[33] V. Mateu, P. G. Ortega, D. R. Entem, and F. Fernndez, Eur. Phys. J. C **79**, 323 (2019).

[34] R Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria, 2018).

[35] J. Pinheiro, D. Bates, S. DebRoy, and D. Sarkar (R Core Team), nlme: Linear and Nonlinear Mixed Effects Models (2018), r package version 3.1-137, https://CRAN.R-project.org/package=nlme.

[36] A. Bazavov *et al.*, Phys. Rev. D **85**, 054503 (2012).

[37] A. Bazavov, H. T. Ding, P. Hegde, F. Karsch, C. Miao, S. Mukherjee, P. Petreczky, C. Schmidt, and A. Velytsky, Phys. Rev. D **88**, 094021 (2013).

[38] F. Herzog, B. Ruijl, T. Ueda, J. A. M. Vermaseren, and A. Vogt, J. High Energy Phys. 02 (2017) 090.

[39] T. Luthe, A. Maier, P. Marquard, and Y. Schröder, J. High Energy Phys. 10 (2017) 166.

[40] W. Fischler, Nucl. Phys. **B129**, 157 (1977).

[41] A. Billoire, Phys. Lett. **92B**, 343 (1980).

[42] Y. Schröder, Phys. Lett. B **447**, 321 (1999).

[43] C. Anzai, Y. Kiyo, and Y. Sumino, Phys. Rev. Lett. **104**, 112003 (2010).

[44] A. V. Smirnov, V. A. Smirnov, and M. Steinhauser, Phys. Rev. Lett. **104**, 112002 (2010).

[45] R. N. Lee, A. V. Smirnov, V. A. Smirnov, and M. Steinhauser, Phys. Rev. D **94**, 054029 (2016).

[46] R. N. Lee and V. A. Smirnov, J. High Energy Phys. 10 (2016) 089.