

High-precision c and b masses, and QCD coupling from current-current correlators in lattice and continuum QCD

C. McNeile,^{1,*} C. T. H. Davies,¹ E. Follana,² K. Hornbostel,³ and G. P. Lepage^{4,†}

(HPQCD Collaboration)

¹*Department of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, United Kingdom*²*Departamento de Física Teórica, Universidad de Zaragoza, E-50009 Zaragoza, Spain*³*Southern Methodist University, Dallas, Texas 75275, USA*⁴*Laboratory for Elementary-Particle Physics, Cornell University, Ithaca, New York 14853, USA*

(Received 24 April 2010; published 26 August 2010)

We extend our earlier lattice-QCD analysis of heavy-quark correlators to smaller lattice spacings and larger masses to obtain new values for the c mass and QCD coupling, and, for the first time, values for the b mass: $m_c(3 \text{ GeV}, n_f = 4) = 0.986(6) \text{ GeV}$, $\alpha_{\overline{\text{MS}}}(M_Z, n_f = 5) = 0.1183(7)$, and $m_b(10 \text{ GeV}, n_f = 5) = 3.617(25) \text{ GeV}$. These are among the most accurate determinations by any method. We check our results using a nonperturbative determination of the mass ratio $m_b(\mu, n_f)/m_c(\mu, n_f)$; the two methods agree to within our 1% errors and taken together imply $m_b/m_c = 4.51(4)$. We also update our previous analysis of $\alpha_{\overline{\text{MS}}}$ from Wilson loops to account for revised values for r_1 and r_1/a , finding a new value $\alpha_{\overline{\text{MS}}}(M_Z, n_f = 5) = 0.1184(6)$; and we update our recent values for light-quark masses from the ratio m_c/m_s . Finally, in the Appendix, we derive a procedure for simplifying and accelerating complicated least-squares fits.

DOI: 10.1103/PhysRevD.82.034512

PACS numbers: 11.15.Ha, 12.38.Aw, 12.38.Gc

I. INTRODUCTION

Precise values for the QCD coupling $\alpha_{\overline{\text{MS}}}$ and the quark masses are important for high-precision tests of the standard model of particle physics. In a recent paper we showed how to use realistic lattice-QCD simulations to extract both the coupling and the charm quark's mass m_c from zero-momentum moments of correlators built from the c quark's (UV cutoff-independent) pseudoscalar-density operator $m_c \bar{\psi}_c \gamma_5 \psi_c$ [1]. In this paper we refine our previous analysis and extend it to include other quark masses, up to and including the b -quark mass. As a result our coupling constant and mass determinations from these correlators are among the most accurate by any method.

Low moments of heavy-quark correlators are perturbative and several are now known through $\mathcal{O}(\alpha_s^3)$ in perturbation theory (that is, four-loop order) [2–6]. Moments of correlators built from the electromagnetic currents can be estimated nonperturbatively, using dispersion relations, from experimental data for the electron-positron annihilation cross section, $\sigma(e^+e^- \rightarrow \gamma^* \rightarrow X)$. Accurate values for both the c and b masses can be obtained by comparing these perturbative and nonperturbative determinations of the moments (for a recent discussion see [7]).

In our earlier paper we showed that heavy-quark correlator moments are easily and accurately computed non-

perturbatively using lattice-QCD simulations, in place of experimental data, provided (1) the electromagnetic current is replaced by the pseudoscalar density multiplied by the bare quark mass; (2) the discretization of the quark action has a partially conserved axial vector current (PCAC); and (3) the discretization remains accurate when applied to heavy quarks. In our simulations we use the highly improved staggered quark (HISQ) discretization of the quark action, which is a highly corrected version of the standard staggered-quark action [8]. It has a chiral symmetry (PCAC) and has been used in a wide variety of accurate simulations involving c quarks [8–12].

Here we show that the HISQ action can be pushed to still higher masses—indeed, very close to the b mass—on new lattices, from the MILC Collaboration [13], with the smallest lattice spacings available today (0.044 fm). Currently most high-precision lattice work on b physics relies upon nonrelativistic effective field theories, like NRQCD [10,12,14,15]. In this paper we show how to obtain accurate b physics using the fully relativistic HISQ action on these new lattices.

In what follows, we first review how the QCD coupling and quark masses are extracted from heavy-quark correlators, in Sec. II. Then in Sec. III we describe our lattice-QCD simulations and discuss in detail the chief systematic errors in our simulation results. In Sec. IV we describe our fitting procedure and the results of our analysis of the heavy-quark correlators. We check our calculation using a different, nonperturbative method to determine m_b/m_c in Sec. V. We then, in Sec. VI, update our previous calculation

*Current address: Department of Theoretical Physics, University of Wuppertal, Wuppertal 42199, Germany.

†g.p.lepage@cornell.edu

of the QCD coupling from Wilson loops to compare with our new result from the correlators. We summarize our findings in Sec. VII and compare our results with work by others. There we also update our recent calculations of the light-quark masses from the c mass. In the Appendix we present a powerful simplification for complicated least-squares fits that can greatly reduce the computing required for fits. We use this technique in dealing with finite- a errors in our analysis.

II. HEAVY-QUARK CORRELATOR MOMENTS

Following our earlier paper [1], we focus on correlators formed from the pseudoscalar density of a heavy quark, $j_5 = \bar{\psi}_h \gamma_5 \psi_h$:

$$G(t) = a^6 \sum_{\mathbf{x}} (am_{0h})^2 \langle 0 | j_5(\mathbf{x}, t) j_5(0, 0) | 0 \rangle, \quad (1)$$

where m_{0h} is the heavy quark's bare mass (from the lattice-QCD Lagrangian), t is Euclidean and periodic with period T , heavy-quark annihilation into gluons is omitted (because it is negligible [16]), and the sum over spatial positions \mathbf{x} sets the total three momentum to zero. In our earlier paper we examined only c quarks; here we will consider a range of masses between the c and b masses. While we have written this formula for use with the lattice regulator, it is important to note that the correlator is UV finite because we include the factors of am_{0h} . Consequently lattice and continuum $G(t)$'s are equal when $t \neq 0$ up to $\mathcal{O}((am_h)^m)$ corrections, which vanish in the continuum limit [17].

The moments of $G(t)$ are particularly simple to analyze:

$$G_n \equiv \sum_t (t/a)^n G(t), \quad (2)$$

where, on our periodic lattice [1],

$$t/a \in \{0, 1, 2, \dots, T/2a - 1, 0, -T/2a + 1, \dots, -2, -1\}. \quad (3)$$

Low moments emphasize small t 's and so are perturbative; and moments with $n \geq 4$ are UV cutoff independent [17]. Therefore

$$G_n = \frac{g_n(\alpha_{\overline{\text{MS}}}(\mu), \mu/m_h)}{(am_h(\mu))^{n-4}} + \mathcal{O}((am_h)^m) \quad (4)$$

for small $n \geq 4$, where $m_h(\mu)$ is the heavy quark's $\overline{\text{MS}}$ mass at scale μ , and the dimensionless factor g_n can be computed using continuum perturbation theory.

Again following our previous paper, we introduce reduced moments to suppress both lattice artifacts and tuning errors in the heavy quark's mass [18]:

$$R_n \equiv \begin{cases} G_4/G_4^{(0)} & \text{for } n = 4, \\ \frac{am_{\eta_h}}{2am_{0h}} (G_n/G_n^{(0)})^{1/(n-4)} & \text{for } n \geq 6, \end{cases} \quad (5)$$

where $G_n^{(0)}$ is the moment in lowest-order, weak-coupling

perturbation theory, using the lattice regulator, and m_{η_h} is the (nonperturbative) mass of the pseudo-Goldstone $h\bar{h}$ boson. The reduced moments can again be written in terms of continuum quantities:

$$R_n \equiv \begin{cases} r_4(\alpha_{\overline{\text{MS}}}, \mu/m_h) & \text{for } n = 4, \\ z(\mu/m_h, m_{\eta_h}) r_n(\alpha_{\overline{\text{MS}}}, \mu/m_h) & \text{for } n \geq 6, \end{cases} \quad (6)$$

up to $\mathcal{O}((am_h)^m \alpha_s)$ corrections, where

$$z(\mu/m_h, m_{\eta_h}) \equiv \frac{m_{\eta_h}}{2m_h(\mu)}, \quad (7)$$

and r_n is obtained from g_n [Eq. (4)] and its value, $g_n^{(0)}$, in lowest-order continuum perturbation theory:

$$r_n = \begin{cases} g_4/g_4^{(0)} & \text{for } n = 4, \\ (g_n/g_n^{(0)})^{1/(n-4)} & \text{for } n \geq 6. \end{cases} \quad (8)$$

Our strategy for extracting quark masses and the QCD coupling relies upon lattice simulations to determine non-perturbative values for the R_n , using simulation results for am_{η_h}/am_{0h} . We then compare this simulation ‘‘data’’ to the continuum perturbation theory formulas [Eq. (6)]. That is, we find values for $\alpha_{\overline{\text{MS}}}(\mu)$ and $z(\mu/m_h, m_{\eta_h})$ that make lattice and continuum results agree for small $n \geq 4$. The function $z(\mu/m_h, m_{\eta_h})$ can then be combined with experimental results for m_{η_c} and m_{η_b} to obtain masses for the c and b quarks:

$$m_c(\mu) = \frac{m_{\eta_c}^{\text{exp}}}{2z(\mu/m_c, m_{\eta_c}^{\text{exp}})}, \quad (9)$$

$$m_b(\mu) = \frac{m_{\eta_b}^{\text{exp}}}{2z(\mu/m_b, m_{\eta_b}^{\text{exp}})}.$$

Parameter μ sets the scale for $\alpha_{\overline{\text{MS}}}$ in the perturbative expansions of the r_n . An obvious choice for this parameter is $\mu = m_h$ since the quark mass, together with n , sets the momentum scale in our correlators. As noted in our previous paper, however, perturbation theory is somewhat more convergent if we use larger μ 's in the c -quark case. Consequently here we take $\mu/m_h = 3$, which is approximately what we did in our previous paper.

The mass and coupling determinations were done separately in our previous paper. Here we extract them simultaneously, to guarantee consistency between results. Also in our previous paper we considered only heavy-quark masses near the c mass. Here we explore a variety of masses ranging from just below the c mass to just below the b mass. This allows us to obtain a value for b -quark's mass.

III. LATTICE-QCD SIMULATIONS

A. Simulation results

The gluon-configuration sets we use were created by the MILC Collaboration. The relevant simulation parameters are listed in Table I.

Given a lattice spacing, the QCD action is specified completely by the values of the bare coupling constant and the bare quark masses. In our analyses we set the u and d quark masses equal; this approximation results in negligible errors ($\ll 1\%$) for the quantities studied in this paper. It is too costly to simulate QCD at the correct value for the u/d mass; we typically use masses that are 2–5 times too large and extrapolate to values that give the correct mass for the π^0 meson. We tune the strange quark mass to give the correct mass for the (fictitious) η_s meson [10]. The c and b masses are tuned to give correct masses for the η_c and η_b mesons, respectively.

It is convenient in QCD simulations to specify a value for the bare coupling constant and then extract the value of the lattice spacing from the simulation. We set the lattice spacing using MILC results for r_1/a , computed from the heavy-quark potential, and

$$r_1 = 0.3133(23) \text{ fm} \quad (10)$$

whose value has been determined from multiple physical quantities [10].

The MILC configurations include vacuum polarization contributions from only the lightest three quark flavors, using the ASQTAD discretization. Vacuum polarization effects from the heavier c and b quarks are easily incorporated into our final results for quark masses and the QCD coupling using perturbation theory.

TABLE I. Parameter sets used to generate the gluon configurations analyzed in this paper. The lattice spacing is specified in terms of the static-quark potential parameter $r_1 = 0.3133(23)$ fm; values for r_1/a are from [13]. The bare quark masses are for the ASQTAD formalism and u_0 is the fourth root of the plaquette. The spatial (L) and temporal (T) lengths of the lattices are also listed, as are the number of gluon configurations (N_{cf}) and the number of time sources (N_{ts}) per configuration used in each case. Sets with similar lattice spacings are grouped.

Set	r_1/a	$au_0m_{0u/d}$	au_0m_{0s}	u_0	L/a	T/a	$N_{\text{cf}} \times N_{\text{ts}}$
1	2.152(5)	0.0097	0.0484	0.860	16	48	631×2
2	2.138(4)	0.0194	0.0484	0.861	16	48	631×2
3	2.647(3)	0.005	0.05	0.868	24	64	678×2
4	2.618(3)	0.01	0.05	0.868	20	64	595×2
5	2.618(3)	0.01	0.05	0.868	28	64	269×2
6	3.699(3)	0.0062	0.031	0.878	28	96	566×4
7	3.712(4)	0.0124	0.031	0.879	28	96	265×4
8	5.296(7)	0.0036	0.018	0.888	48	144	201×2
9	7.115(20)	0.0028	0.014	0.895	64	192	208×2

We computed heavy-quark correlators [Eq. (1)] using the HISQ discretization [8] for a variety of bare heavy-quark masses am_{0h} on the MILC gluon configurations. Our results for the reduced moments R_n with $n = 4$ –18 are given in Table II.

In Table II, we also give masses am_{η_h} from the simulations for the pseudo-Goldstone meson made from two heavy quarks. These were computed using single-exponential fits to $G(t)$ for the middle 30% of t 's on the lattice for all configurations except the two smallest lattice spacings where we used only 8% of the t 's. We have less statistics for the two finest lattice spacings and consequently the fits did not work as well for these. We increased the statistical errors on our results for am_{η_h} by factors of 1.4 and 2 for the next-to-finest and finest lattice spacings (sets 8 and 9), respectively, to ensure that the χ^2 per degree of freedom was 1 or less. The statistical errors here are very small and have only a small impact on our final results. We also verified our results with multiexponential fits in every case.

B. Systematic errors

As discussed above, our goal is to find values for $\alpha_{\overline{\text{MS}}}(\mu)$ and $z(\mu/m_h, m_{\eta_h})$ [Eq. (7)] that make the theoretical results from perturbative QCD agree, to within statistical and systematic errors, with Monte Carlo simulation data for the reduced moments. We simultaneously analyze results for all of our lattice spacings and most of our masses, and for moments with $4 \leq n \leq 10$. We focus on these particular moments for our final results since their perturbation theory is known to third order.

Systematic errors are larger here than statistical errors, which contribute less than 0.3%. We discuss the most important sources of systematic error in this section.

1. m_h extrapolations

We need the m_{η_h} dependence of the mass-ratio function $z(\mu/m_h = 3, m_{\eta_h})$ in order to extract c and b masses from our simulation [using Eq. (9)]. We parametrize this dependence as follows:

$$z(\mu/m_h, m_{\eta_h}) = \sum_{j=0}^{N_z} z_j(\mu/m_h) \left(\frac{2\Lambda}{m_{\eta_h}} \right)^j, \quad (11)$$

where the z_j 's are determined in our fit. This is an expansion in the QCD scale, which we take to be

$$\Lambda = 0.5 \text{ GeV}, \quad (12)$$

divided by $m_{\eta_h}/2$, which we use as a proxy for the quark mass. The expansion is adequate for the range of quark masses used in our analysis, where $(2\Lambda/m_{\eta_h})^2$ ranges approximately between $1/m_{\eta_b}^2 = 0.01$ and $(1/m_{\eta_c})^2 = 0.1$; the singular point $m_h = 0$ is infinitely far away in this parametrization. In our fits we keep terms only through

TABLE II. Results for the reduced moments R_n and pseudoscalar-meson mass am_{η_h} obtained from ($n_f = 3$) simulations using different bare heavy-quark (HISQ) masses am_{0h} and gluon-configuration sets (see Table I). The errors listed here are statistical errors from the Monte Carlo simulation. Results where $am_{\eta_h} > 1.95$ are omitted from our final analysis, as are R_n 's with $n > 10$.

Set	am_{0h}	am_{η_h}	R_4	R_6	R_8	R_{10}	R_{12}	R_{14}	R_{16}	R_{18}	
1	0.660	1.9202(1)	1.2132(3)	1.5364(3)	1.4151(2)	1.3476(1)	1.3001(1)	1.2649(1)	1.2378(1)	1.2164(1)	
	0.810	2.1938(1)	1.1643(2)	1.4427(2)	1.3619(1)	1.3148(1)	1.2780(1)	1.2481(1)	1.2238(1)	1.2039(1)	
	0.825	2.2202(1)	1.1604(2)	1.4339(2)	1.3563(1)	1.3111(1)	1.2754(1)	1.2462(1)	1.2222(1)	1.2025(1)	
2	0.825	2.2196(1)	1.1591(2)	1.4327(2)	1.3556(1)	1.3106(1)	1.2751(1)	1.2459(1)	1.2221(1)	1.2024(1)	
	3	0.650	1.8458(1)	1.1809(2)	1.4805(2)	1.3755(1)	1.3160(1)	1.2740(1)	1.2429(1)	1.2190(1)	
4	0.440	1.4241(1)	1.2752(4)	1.6144(4)	1.4397(2)	1.3561(2)	1.3041(1)	1.2678(1)	1.2408(1)	1.2200(1)	
	0.630	1.8085(1)	1.1881(3)	1.4935(2)	1.3826(1)	1.3205(1)	1.2773(1)	1.2456(1)	1.2214(1)	1.2021(1)	
	0.660	1.8667(1)	1.1782(2)	1.4764(2)	1.3738(1)	1.3152(1)	1.2736(1)	1.2426(1)	1.2187(1)	1.1997(1)	
	0.720	1.9811(1)	1.1605(2)	1.4435(2)	1.3559(1)	1.3044(1)	1.2662(1)	1.2367(1)	1.2136(1)	1.1950(1)	
	0.850	2.2194(1)	1.1301(2)	1.3763(1)	1.3145(1)	1.2774(1)	1.2473(1)	1.2221(1)	1.2012(1)	1.1839(1)	
5	0.630	1.8086(1)	1.1882(1)	1.4936(1)	1.3826(1)	1.3205(1)	1.2774(1)	1.2457(1)	1.2214(0)	1.2022(0)	
	6	0.300	1.0314(1)	1.2930(3)	1.6061(3)	1.4249(2)	1.3444(1)	1.2953(1)	1.2610(1)	1.2353(1)	
6	0.413	1.2806(1)	1.2224(2)	1.5216(2)	1.3796(1)	1.3115(1)	1.2689(1)	1.2390(1)	1.2164(1)	1.1985(1)	
	0.430	1.3169(1)	1.2145(2)	1.5113(2)	1.3743(1)	1.3076(1)	1.2658(1)	1.2363(1)	1.2141(1)	1.1964(1)	
	0.440	1.3382(1)	1.2100(2)	1.5054(2)	1.3712(1)	1.3054(1)	1.2640(1)	1.2348(1)	1.2127(1)	1.1952(1)	
	0.450	1.3593(1)	1.2057(2)	1.4996(2)	1.3683(1)	1.3033(1)	1.2623(1)	1.2333(1)	1.2114(1)	1.1941(1)	
	0.700	1.8654(1)	1.1301(1)	1.3782(1)	1.3053(1)	1.2616(1)	1.2294(1)	1.2048(0)	1.1857(0)	1.1705(0)	
	0.850	2.1498(1)	1.1026(1)	1.3163(1)	1.2671(1)	1.2366(0)	1.2114(0)	1.1903(0)	1.1729(0)	1.1584(0)	
	7	0.427	1.3074(1)	1.2131(3)	1.5091(3)	1.3729(2)	1.3066(1)	1.2651(1)	1.2358(1)	1.2137(1)	1.1961(1)
	8	0.273	0.8993(3)	1.2454(8)	1.5234(9)	1.3739(7)	1.3069(6)	1.2657(6)	1.2366(6)	1.2145(6)	1.1969(5)
0.280		0.9154(2)	1.2403(5)	1.5175(5)	1.3706(3)	1.3045(3)	1.2638(2)	1.2350(2)	1.2132(2)	1.1958(2)	
0.564		1.5254(1)	1.1324(2)	1.3674(2)	1.2857(2)	1.2405(1)	1.2102(1)	1.1885(1)	1.1719(1)	1.1587(1)	
0.705		1.8084(1)	1.1043(2)	1.3156(2)	1.2574(1)	1.2217(1)	1.1952(1)	1.1750(1)	1.1593(1)	1.1467(1)	
0.760		1.9157(1)	1.0955(2)	1.2965(2)	1.2460(1)	1.2142(1)	1.1895(1)	1.1701(1)	1.1547(1)	1.1423(1)	
0.850		2.0875(1)	1.0831(2)	1.2666(1)	1.2266(1)	1.2010(1)	1.1799(1)	1.1621(1)	1.1474(1)	1.1353(1)	
9		0.195	0.6710(2)	1.2583(5)	1.5243(5)	1.3733(4)	1.3066(3)	1.2655(3)	1.2364(2)	1.2144(2)	1.1968(2)
	0.400	1.1325(2)	1.1532(3)	1.3800(3)	1.2838(2)	1.2370(2)	1.2077(2)	1.1869(2)	1.1710(2)	1.1583(2)	
	0.500	1.3446(2)	1.1267(2)	1.3410(2)	1.2616(1)	1.2198(1)	1.1927(1)	1.1734(1)	1.1588(1)	1.1471(1)	
	0.700	1.7518(1)	1.0900(1)	1.2765(1)	1.2261(1)	1.1949(1)	1.1718(1)	1.1542(1)	1.1407(1)	1.1299(1)	
	0.850	2.0428(1)	1.0712(1)	1.2327(1)	1.1983(1)	1.1760(1)	1.1574(1)	1.1418(1)	1.1290(1)	1.1185(1)	

order $N_z = 4$, but, as we discuss later, our results are unchanged by additional terms. On dimensional grounds, we assume *a priori* that the coefficients are

$$z_j(3) = 0 \pm 1. \quad (13)$$

2. Finite-lattice spacing errors

Discretization errors are of order $(am_h)^{2i}\alpha_s$ for $i \geq 1$. We model these by

$$R_n^{\text{latt}} = R_n(\mu, m_{\eta_h}, a, N_{am}), \quad (14)$$

where fit function $R_n(\mu, m_{\eta_h}, a, N_{am})$ has the double expansion

$$R_n(\mu, m_{\eta_h}, a, N_{am}) \equiv R_n^{\text{cont}} / \left(1 + \sum_{i=1}^{N_{am}} \sum_{j=0}^{N_z} c_{ij}^{(n)} \left(\frac{am_{\eta_h}}{2} \right)^{2i} \times \left(\frac{2\Lambda}{m_{\eta_h}} \right)^j \right), \quad (15)$$

the $c_{ij}^{(n)}$'s are determined in our fit, R_n^{cont} is given by Eq. (6),

$$i + j \leq \max(N_{am}, N_z), \quad (16)$$

and again we use $m_{\eta_h}/2$ in place of the quark mass. This expansion allows for finite- a corrections involving $(am_{\eta_h}/2)^2$, $(a\Lambda)^2$, and cross terms, with m_{η_h} -dependent coefficients. We assume *a priori* that

$$c_{ij}^{(n)} = 0 \pm 2/n, \quad (17)$$

which implies smaller a dependence for larger n 's. This is expected (and obvious in our simulation data) since the reduced moments become more infrared as n increases. The exact functional form of the n dependence has little effect on our results, as we show later.

In our fits we take $N_z = 4$. While low orders suffice for the $2\Lambda/m_{\eta_h}$ expansion, expansion parameter $am_{\eta_h}/2$ ranges between 0.3 and 1.1, and higher orders are necessary, especially given our tiny statistical errors. We find that our fit results do not converge well unless N_{am} is larger

than 10–20. Also we have difficulty getting good fits if we include data with $am_{\eta_h} > 1.95$ from Table II. The $am_{\eta_h}/2$ expansion may not converge for these last cases and therefore we exclude such data from our final analysis.

The fit function has many more fit parameters $c_{ij}^{(n)}$ than we have simulation data points when N_{am} is so large. This does not cause problems in (Bayesian) constrained fits since the parameters' priors [Eq. (17)] are included in the fit as extra data [19]. Each parameter has a *prior* and therefore we always have more data than parameters.

It is, however, very time consuming to fit a function with so many fit parameters. Although it is not essential for our analysis, there is a trick that greatly accelerates this kind of fit. The idea is to fit a modified moment \bar{R}_n^{latt} in place of R_n^{latt} , where

$$\bar{R}_n^{\text{latt}} \equiv R_n^{\text{latt}} + R_n^{\text{latt}} \sum_{i=\bar{N}_{am}+1}^{N_{am}} \sum_{j=0}^{N_z} c_{ij}^{(n)} \left(\frac{am_{\eta_h}}{2} \right)^{2i} \left(\frac{2\Lambda}{m_{\eta_h}} \right)^j \quad (18)$$

and $\bar{N}_{am} \ll N_{am}$. The modified moment is fit with the much simpler formula (simpler since $\bar{N}_{am} \ll N_{am}$)

$$\bar{R}_n^{\text{latt}} = R_n(\mu, m_{\eta_h}, a, \bar{N}_{am}), \quad (19)$$

where $R_n(\cdot \cdot \cdot)$ is again given by Eq. (15). To evaluate \bar{R}_n^{latt} from Eq. (18), we treat the coefficients $c_{ij}^{(n)}$ with $i > \bar{N}_{am}$ as new data with means and standard deviations specified by the prior, Eq. (17). Uncertainties coming from the $c_{ij}^{(n)}$'s are combined in quadrature with the statistical error in R_n^{latt} to obtain a new error estimate for \bar{R}_n^{latt} (but leaving the central value unchanged). In effect we are increasing the error in the reduced moment to account for high-order $(am_{\eta_h}/2)^{2i}$ terms omitted from the fit formula Eq. (19). By choosing $\bar{N}_{am} \ll N_{am}$, most of the $am_{\eta_h}/2$ terms are incorporated into \bar{R}_n^{latt} [Eq. (18)], where they are inexpensive, and relatively few end up in the fit function $\bar{R}_n(\cdot \cdot \cdot)$ [Eq. (15)], where they add parameters to the fit and increase its cost. Note that the new errors introduce correlations between \bar{R}_n^{latt} 's computed with different lattice spacings or quark masses, since the same $c_{ij}^{(n)}$'s are used for all a 's and m_{η_h} 's. These correlations are important and need to be preserved in the fit.

Our procedure, whereby terms are moved out of the fitting function and incorporated into new (correlated) errors in the Monte Carlo fit data, is generally useful. Somewhat remarkably, final fit results are completely (or almost completely) independent of the number of terms that are transferred when fits are linear (or almost linear) in the associated parameters. (The general theorem from which this result follows is proven in the Appendix.) Consequently, in our analysis here, we can take N_{am} very large—say, $N_{am} = 80$ —and still have very fast fits by keeping \bar{N}_{am} very small. With $N_{am} = 80$ we find, for example, that setting $\bar{N}_{am} = 0$ in $\bar{R}_n(\cdot \cdot \cdot)$ (no terms) gives essentially identical results for our quark masses and cou-

pling as setting $\bar{N}_{am} = 30$ (140 terms), even though the latter fit requires 22 times more computing. We used this procedure, with $\bar{N}_{am} = 0$, for most of our testing and development in this project.

3. Truncated perturbation theory

The perturbative part,

$$r_n(\alpha_{\overline{\text{MS}}}, \mu/m_h) = 1 + \sum_{j=1}^{N_{\text{pth}}} r_{nj}(\mu/m_h) \alpha_{\overline{\text{MS}}}^j(\mu), \quad (20)$$

of the reduced moments is known at best through third order. We present coefficients r_{nj} through $j = 3$ in Table III [2–6]; the values for $n = 4$ –10 are exact, while r_{n3} is estimated for the others. In our fits we include higher-order terms by treating the coefficients of these terms as fit parameters with prior

$$r_{nj}(1) = 0 \pm 0.5 \quad (21)$$

for any coefficient that has not been computed in perturbation theory. We set $N_{\text{pth}} = 6$ since then contributions from still higher orders should be less than 0.1% (and setting $N_{\text{pth}} = 8$ does not change our results).

The perturbative coefficients for $\mu/m_h = 1$ (Table III) are small and relatively uncorrelated from order to order. This is less true for $\mu/m_h = 3$, which is where we wish to work (see Sec. II), because of $\log(\mu/m_h)^m$ terms. In order to capture these effects, we use renormalization-group equations to express the $r_{nj}(3)$ coefficients (for all $j \leq N_{\text{pth}}$) in terms of the $r_{nj}(1)$ coefficients and $\log(\mu/m_h)$, and substitute the results from Table III for $j \leq 3$ and from the prior [Eq. (21)] for $j > 3$. This procedure generates (correlated) priors for the unknown coefficients at $\mu/m_h = 3$ that properly account for renormalization-group logarithms.

TABLE III. Perturbation theory coefficients ($n_f = 3$) for r_n [2–6]. Coefficients are defined by $r_n = 1 + \sum_{j=1}^{N_{\text{pth}}} r_{nj} \alpha_{\overline{\text{MS}}}^j(\mu)$ for $\mu = m_h(\mu)$. The third-order coefficients are exact for $4 \leq n \leq 10$. The other coefficients are based upon estimates; we assign conservative errors to these.

n	r_{n1}	r_{n2}	r_{n3}
4	0.7427	−0.0577	0.0591
6	0.6160	0.4767	−0.0527
8	0.3164	0.3446	0.0634
10	0.1861	0.2696	0.1238
12	0.1081	0.2130	0.1(3)
14	0.0544	0.1674	0.1(3)
16	0.0146	0.1293	0.1(3)
18	−0.0165	0.0965	0.1(3)

4. $\alpha_{\overline{\text{MS}}}$ evolution

As discussed above, we fix the ratio of $\mu/m_h(\mu)$ in our analysis. This means that the renormalization scale μ varies over a wide range of values for the different m_h 's we use. The coupling constant $\alpha_{\overline{\text{MS}}}(\mu)$ used in the perturbative expansions for the r_n 's is specified at $\mu = 5$ GeV by fit parameter α_0 , with prior

$$\alpha_0 = 0.20 \pm 0.01, \quad (22)$$

where

$$\alpha_0 \equiv \alpha_{\overline{\text{MS}}}(5 \text{ GeV}, n_f = 3). \quad (23)$$

The prior corresponds to $\alpha_{\overline{\text{MS}}}(M_Z) = 0.118(3)$ —a very broad range, which means that the prior has little impact on our final fit results. The coupling value at any scale $\mu \neq 5$ GeV is obtained by integrating (numerically) the QCD evolution equation for $\alpha_{\overline{\text{MS}}}(\mu)$ starting with value α_0 at scale 5 GeV. We use the $\overline{\text{MS}}$ beta function through sixth order in $\alpha_{\overline{\text{MS}}}$,

$$\begin{aligned} \mu^2 \frac{d\alpha_{\overline{\text{MS}}}(\mu)}{d\mu^2} = & -\beta_0 \alpha_{\overline{\text{MS}}}^2 - \beta_1 \alpha_{\overline{\text{MS}}}^3 - \beta_2 \alpha_{\overline{\text{MS}}}^4 \\ & - \beta_3 \alpha_{\overline{\text{MS}}}^5 - \beta_4 \alpha_{\overline{\text{MS}}}^6, \end{aligned} \quad (24)$$

where $\beta_0 \cdots \beta_3$ are known from perturbation theory and β_4 is taken as a fit parameter with prior

$$\beta_4 = 0 \pm \sigma_\beta, \quad (25)$$

where σ_β is the root-mean-square average of $\beta_0 \cdots \beta_3$ [20,21]. We include this last term to estimate the uncertainties in our final results caused by unknown terms in the beta function.

Our simulations include vacuum polarization effects from only the three lightest quarks. We use perturbation theory, together with the c and b masses that come out of our analysis, to incorporate vacuum polarization effects from the heavier quarks into our final results for the masses and QCD coupling (using formulas from [22–24] to add the c and b quarks at scales $\mu = m_c$ and m_b , respectively).

5. Nonperturbative condensates

As discussed in our previous paper, nonperturbative effects dominate the reduced moments when n is large. The dominant nonperturbative contribution, which is from the gluon condensate, is quite small, however, for the range of n 's and quark masses we use here. We correct for it by replacing

$$R_n^{\text{latt}} \rightarrow R_n^{\text{latt}} \left(1 + d_n \frac{\langle \alpha_s G^2 / \pi \rangle}{(2m_h)^4} \right), \quad (26)$$

where d_n is computed to leading order in perturbation theory [25] with $m_h = m_h(m_h)$, which we approximate by $m_{\eta_h}/2.27$. We take

$$\langle \alpha_s G^2 / \pi \rangle = 0 \pm 0.012 \text{ GeV}^4, \quad (27)$$

which covers the range of most current estimates [26]. The correction factor in Eq. (26) adds (slightly) to the error in R_n^{latt} (and introduces new correlations between different moments, since the same $\langle \alpha_s G^2 / \pi \rangle$ is assumed for every moment, lattice spacing, and quark mass).

6. Finite-volume errors

We expect small errors due to the fact that our simulation lattices are only about 2.5 fm across. We allow for the possibility of finite-volume errors by replacing

$$R_n^{\text{latt}} \rightarrow R_n^{\text{latt}} \left(1 + f_n \frac{\Delta R_n^{\text{pth}}}{R_n^{\text{pth}}} \right), \quad (28)$$

where ΔR_n^{pth} is the finite-volume error in leading-order perturbation theory and

$$f_n = 0 \pm 0.5. \quad (29)$$

The true finite-volume errors are expected to be smaller, because of quark confinement, than the perturbative errors that we use to model them here. We verified this by running two sets of simulations that were identical except for the spatial volume (gluon-configuration sets 4 and 5 in Table I). The differences between the two simulations are smaller than our statistical errors, but the statistical errors are much smaller than our estimate above. Our error estimate here is very conservative, but has negligible impact on our final results.

7. Sea-quark masses

The sea-quark masses used in our simulations are not exactly correct. To correct for this we replace

$$R_n^{\text{latt}} \rightarrow R_n^{\text{latt}} \left(1 + g_n \frac{2\delta m_l + \delta m_s}{m_s} \right), \quad (30)$$

where δm_l and δm_s are the errors in the u/d and s masses (see [10] for more details), respectively, and

$$g_n = 0 \pm 0.01. \quad (31)$$

This correction introduces (correlated) errors into the R_n^{latt} 's that are of order 0.5%–1%. Direct comparison of results from configuration sets 6 and 7 (or 1–2 and 3–4) in Table I suggests that sea-quark mass effects are no larger than 0.1%, so our error estimate is conservative.

We have only included the leading dependence on the sea-quark mass, which comes from nonperturbative (chiral) effects. Quadratic terms from perturbation theory and other nonperturbative sources are negligible.

IV. ANALYSIS AND RESULTS

We have computed reduced moments for 30 different sets of lattice spacing, lattice volume, and quark masses (Table II). To extract quark masses and the QCD coupling,

we fit moments with $4 \leq n \leq 10$ from 22 of these parameter sets (the ones with $am_{\eta_h} \leq 1.95$)—88 pieces of simulation data in all. In this section we first describe the fitting method used to extract the masses and coupling, and then we review our results.

A. Constrained fits

We analyze all four R_n 's for all 22 parameter sets simultaneously using a constrained fitting procedure based upon Bayesian ideas [19]. In this procedure we minimize an augmented χ^2 function of the form

$$\chi^2 = \sum_{in,jm} \Delta R_{ni} (\sigma_R^{-2})_{in,jm} \Delta R_{mj} + \sum_{\xi} \delta \chi_{\xi}^2, \quad (32)$$

where

$$\Delta R_{ni} \equiv R_{ni}^{\text{latt}} - R_n(\mu_i, m_{\eta_h}, a_i, N_{am}); \quad (33)$$

the R_n^{latt} come from Table II with corrections from Eqs. (26), (28), and (30); fit function $R_n(\dots)$ is defined by Eq. (15); and σ_R^2 is the error covariance matrix for the R_n^{latt} . The sums i, j are over the 22 sets of lattice spacings and quark masses; the sums n, m range over the moments 4, 6, 8, and 10.

Function $R_n(\mu_i, m_{\eta_h}, a_i, N_{am})$ depends upon a large number of parameters, all of which are varied in the fit to minimize χ^2 . Priors $\delta \chi_{\xi}^2$ are included for each of these:

- (i) parameters z_j , with prior Eq. (13), from the $1/m_{\eta_h}$ expansion of $z(\mu/m_h, m_{\eta_h})$;
- (ii) parameters $c_{ij}^{(n)}$, with prior Eq. (17), from the finite-lattice spacing corrections;
- (iii) unknown perturbative coefficients r_{nj} , with prior Eq. (21) (evolved to $\mu/m_h = 3$);
- (iv) coupling parameter $\log(\alpha_0)$, with prior Eq. (22);
- (v) β_4 in the QCD β function, with prior Eq. (25);
- (vi) lattice spacings a_i for each gluon-configuration set, with priors specified by simulation results for r_1/a (Table I) and the current value for r_1 [Eq. (10)];
- (vii) values for am_{η_h} , with priors specified by our simulation results (Table II).

The renormalization scales μ_i are obtained from the ratio $\mu/m_h = 3$, simulation results for m_{η_h} , and Eq. (7). We take $N_{am} = 30$ for our final results.

B. Results

We fit our simulation data for the reduced moments R_n^{latt} (Table II) using fit function $R_n(\dots)$ [Eq. (15)] with $N_{am} = 30$, as discussed in the previous section. The best-fit values for parameters z_j give us the mass-ratio function $z(\mu/m_h = 3, m_{\eta_h})$ [Eq. (7)], which we plot in Fig. 1. We also show our simulation results there for R_n^{latt}/r_n , together with the best-fit lines for each lattice spacing. Results are shown for the three moments that depend upon z , 5 different lattice spacings, and quark masses ranging from below the c mass almost to the b mass. The simulation data were

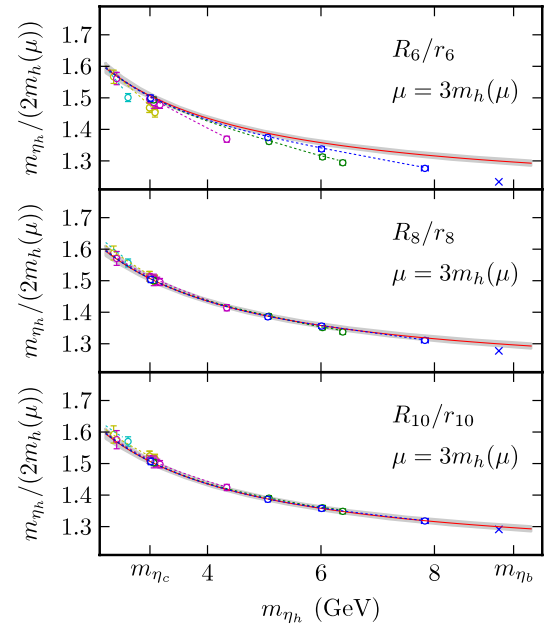


FIG. 1 (color online). Function $z(\mu/m_h = 3, m_{\eta_h}) \equiv m_{\eta_h}/(2m_h)$ as a function of m_{η_h} . The solid line, plus gray error envelope, shows the $a = 0$ extrapolation obtained from our fit. This is compared with simulation results for R_n/r_n for $n = 6, 8, 10$ from our 5 different lattice spacings, together with the best fits (dashed lines) corresponding to those lattice spacings. Dashed lines for smaller lattice spacings extend farther to the right. The points marked by an \times are for the largest mass we tried (last line in Table II); these are not included in the fit because am_{η_h} is too large. Finite- a errors become very small for the larger- n moments, causing points from different lattice spacings to overlap.

all fit simultaneously, using the same functions $z(3, m_{\eta_h})$ and $\alpha_{\overline{\text{MS}}}(\mu)$ [with $\mu = 3m_{\eta_h}/(2z)$] for all moments. The fits are excellent, with $\chi^2/88 = 0.19$ for the 88 pieces of simulation data we fit.

Evaluated at $m_{\eta_c} = 2.985(3)$ GeV [27], the mass-ratio function is $z(3, m_{\eta_c}) = 1.507(7)$. Combining this with Eq. (9) and perturbation theory, we can obtain the following results for the $\overline{\text{MS}}$ c -quark mass at different scales:

$$\begin{aligned} m_c(3m_c, n_f = 3) &= 0.991(5) \text{ GeV}, \\ m_c(3 \text{ GeV}, n_f = 4) &= 0.986(6) \text{ GeV}, \\ m_c(m_c, n_f = 4) &= 1.273(6) \text{ GeV}. \end{aligned} \quad (34)$$

Similarly at $m_{\eta_b} = 9.395(5)$ GeV [28], the mass-ratio function is $z(3, m_{\eta_b}) = 1.296(8)$, and we obtain the following results for the $\overline{\text{MS}}$ b -quark mass at different scales:

$$\begin{aligned} m_b(3m_b, n_f = 3) &= 3.622(22) \text{ GeV}, \\ m_b(10 \text{ GeV}, n_f = 5) &= 3.617(25) \text{ GeV}, \\ m_b(m_b, n_f = 5) &= 4.164(23) \text{ GeV}. \end{aligned} \quad (35)$$

Note that the ratio $m_b(\mu, n_f)/m_c(\mu, n_f)$ is independent of μ and n_f . We obtain the following result for this mass ratio:

$$m_b/m_c = 4.53(4). \quad (36)$$

The other important output from our fit is a value for the parameter

$$\alpha_0 \equiv \alpha_{\overline{\text{MS}}}(5 \text{ GeV}, n_f = 3) = 0.2034(21). \quad (37)$$

To compare with other determinations of the coupling, we add vacuum polarization corrections from the c and b quarks, using the masses above, and evolve to the Z -meson mass [20–23]:

$$\alpha_{\overline{\text{MS}}}(M_Z, n_f = 5) = 0.1183(7). \quad (38)$$

Figure 2 shows how consistent our simulation results are with the theoretical curve for $\alpha_{\overline{\text{MS}}}(\mu, n_f = 3)$ corresponding to our value for α_0 . For this figure we extracted values for $\alpha_{\overline{\text{MS}}}$ from each R_n separately by dividing out the a^2

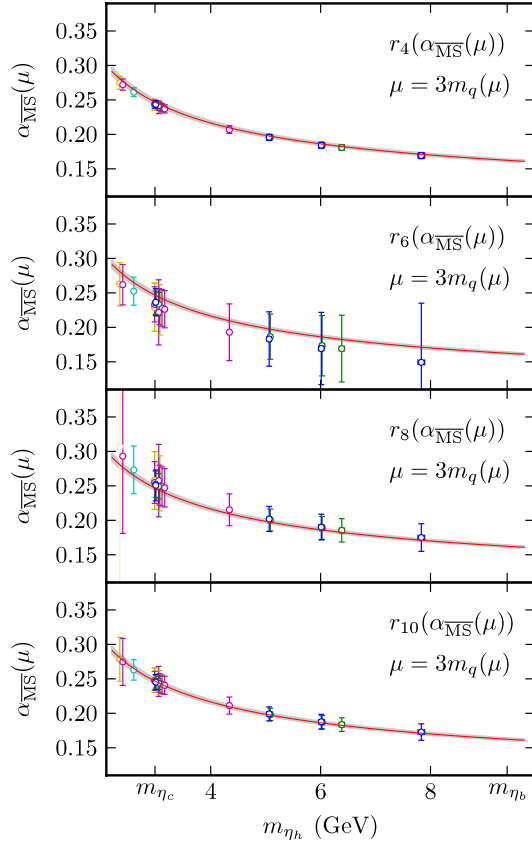


FIG. 2 (color online). QCD coupling $\alpha_{\overline{\text{MS}}}(\mu, n_f = 3)$ as a function of m_{η_h} where $\mu = 3m_h$. The solid line, plus gray error envelope, shows the best-fit coupling from our fit when perturbative evolution is assumed. The data points are values of $\alpha_{\overline{\text{MS}}}$ extracted from individual simulation results for R_n after extrapolating to $a = 0$ and dividing out $z(3, m_{\eta_h})$ ($n > 4$). Results are given for moments $n = 4$ – 10 and all 5 lattice spacings. Several points from different lattice spacings overlap in these plots.

dependence and $z(3, m_{\eta_h})$ using our best-fit parameters, and then solving for $\alpha_{\overline{\text{MS}}}$ by matching with perturbation theory for r_n . (In our fit, of course, we fit all R_n 's simultaneously to obtain a single $\alpha_{\overline{\text{MS}}}$ for all of them.)

The dominant sources of error for our results are listed in Table IV. The largest uncertainties come from extrapolations to $a = 0$, especially for quantities involving b quarks; unknown higher-order terms in perturbation theory, especially for quantities involving c quarks; statistical fluctuations; extrapolations in the heavy quark mass, especially for quantities involving b quarks; and uncertainties in static-quark parameters r_1/a and r_1 . The pattern of errors is as expected in each case. The nonperturbative contribution from the gluon condensate is negligible except for m_c , again as expected; and errors due to mistuned sea-quark masses, finite-volume errors, and uncertainties in $\overline{\text{MS}}$ coupling and mass evolution are negligible ($< 0.05\%$).

The a^2 extrapolations of our data are not large. This is illustrated for $m_h \approx m_c$ in Fig. 3, which shows the a^2 dependence of the reduced moments. The smallest two lattice spacings are sufficiently close to $a = 0$ so that the extrapolation is almost linear from those points. The $a = 0$ extrapolated values we obtain here for the R_n agree to within (smaller) errors with those in our previous paper: here we get 1.282(4), 1.527(4), 1.373(3), 1.304(2) with $n = 4, 6, 8, 10$, respectively, for the masses used in the figure.

We tested the stability of our analysis in several ways:

- (i) Vary perturbation theory: We chose $\mu = 3m_h$ in order to keep scales large and $\alpha_{\overline{\text{MS}}}(\mu)$ small. Our results are quite insensitive to μ , however. Choosing $\mu = m_h$, for example, shifts none of our results by more than 0.2σ , and leaves all errors unchanged except for $m_c(3)$, where the error increases by a third. Taking $\mu = 9m_h$ shifts results by less than 0.4σ , and reduces the m_c error by one-third, leaving others only slightly reduced. Adding more terms to the perturbative expansions ($N_{\text{pth}} = 6 \rightarrow 8$) also has essentially no effect on the results. The prior for the

TABLE IV. Sources of uncertainty for the QCD coupling and mass determinations in this paper. In each case the uncertainty is given as a percentage of the final value.

	$m_c(3)$	$m_b(10)$	m_b/m_c	$\alpha_{\overline{\text{MS}}}(M_Z)$
a^2 extrapolation	0.2%	0.6%	0.5%	0.2%
Perturbation theory	0.5	0.1	0.5	0.4
Statistical errors	0.1	0.3	0.3	0.2
m_h extrapolation	0.1	0.1	0.2	0.0
Errors in r_1	0.2	0.1	0.1	0.1
Errors in r_1/a	0.1	0.3	0.2	0.1
Errors in m_{η_c}, m_{η_b}	0.2	0.1	0.2	0.0
α_0 prior	0.1	0.1	0.1	0.1
Gluon condensate	0.0	0.0	0.0	0.2
Total	0.6%	0.7%	0.8%	0.6%

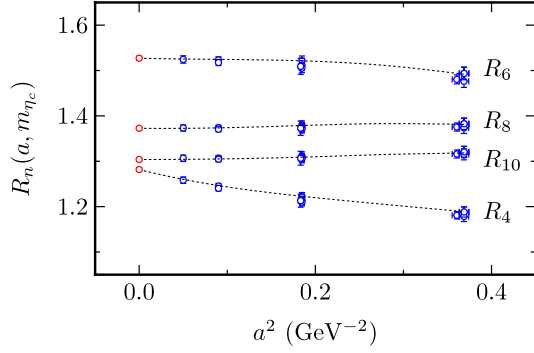


FIG. 3 (color online). Lattice-spacing dependence of R_n for masses m_{η_h} within 5% of m_{η_c} and moments $n = 4, 6, 8,$ and 10 . The dashed lines show our fit for the average of these masses, and the points at $a = 0$ are the continuum extrapolations of our data.

unknown perturbative coefficients [Eq. (21)] is twice as wide as suggested by our simulation results (using the empirical Bayes criterion [19]); we choose the larger width to be conservative.

- (ii) Include more/fewer finite- a corrections: We set $N_{am} = 30$ for our results above. Using $N_{am} = 15$ gives results that differ by less than 0.5σ for m_b and much less for the other quantities. Much larger N_{am} 's can be tested easily using the trick described in Sec. III B 2. For example, replacing R_n^{latt} by $\tilde{R}_n^{\text{latt}}$ [Eq. (18)] with $N_{am} = 80$ and $\tilde{N}_{am} = 30$ gives results that are essentially identical to those above. As discussed above, taking $\tilde{N}_{am} = 0$ with the same N_{am} also gives the same results and is 22 times faster (see the Appendix for further discussion).
- (iii) Change n dependence of finite- a corrections: Replacing the n -dependent prior for the expansion coefficients [Eq. (17)] by the n -independent prior 0 ± 0.5 causes changes that are less than 0.3σ . The width of the original prior is optimal according to the empirical Bayes criterion—that is, it is the width suggested by the size of finite- a deviations observed in our simulation data.
- (iv) Add more/fewer $\Lambda/m\eta_h$ terms in z : Increasing the number of terms in the expansion for z from $N_z = 4$ to 6 changes nothing by more than 0.1σ . Decreasing to $N_z = 3$ also has no effect. Again the width of the prior is optimal according to the empirical Bayes criterion.
- (v) Include more/fewer moments: Keeping all moments $4 \leq n \leq 18$ changes nothing by more than 0.5σ and reduces errors slightly for everything other than m_b , where the errors are cut almost in half: $m_b(10) = 3.623(15)$ GeV or $m_b(m_b) = 4.170(13)$ GeV, both for $n_f = 5$. We continue to restrict ourselves to moments with $n \leq 10$ because these are the only moments for which we have exact third-order perturbation theory. Keeping just $n = 4, 6$ gives al-

most identical results for m_c and $\alpha_{\overline{\text{MS}}}$, with almost the same errors, but doubles the error on m_b .

- (vi) Omit simulation data: The coarsest two lattice spacings (configuration sets 1–5) affect our results only weakly. Leaving these out shifts no result by more than 0.5σ and leaves errors almost unchanged. Leaving out the smallest lattice spacing, however, increases errors significantly (almost double for $\alpha_{\overline{\text{MS}}}$), while still shifting central values by less than 0.5σ .
- (vii) Add large masses: Including cases with $am_{\eta_h} > 1.95$ from Table II leads to poor fits. The excluded data, however, do not deviate far from the best-fit lines. For example, the points marked with an \times in Fig. 1 are for the largest mass we studied, corresponding to $m_{\eta_h} = 9.15$ GeV (last line in Table II). Although am_{η_h} is too large for this case to be included in our fit, the values of R_n/r_n are only slightly below the fit results.

V. NONPERTURBATIVE m_b/m_c

It is possible to extract the ratio of quark masses m_b/m_c directly, without using the moments and without using perturbation theory. This provides an excellent nonperturbative check on our results from the moments.

Ratios of quark masses are UV cutoff independent and therefore the ratio of $\overline{\text{MS}}$ masses

$$\frac{m_b(\mu, n_f)}{m_c(\mu, n_f)} = \frac{m_{0b}}{m_{0c}} + \mathcal{O}(\alpha_s a^2 m_b^2) \quad (39)$$

for any μ and n_f , where m_{0b} and m_{0c} are the bare quark masses in the lattice quark action that give correct masses for the η_c and η_b , respectively. We obtain accurate mass ratios from this relationship by extrapolating to $a = 0$. We used such a method recently to determine m_c/m_s [11].

Here we have to modify our earlier method slightly because we cannot reach the b -quark mass directly, but rather must simultaneously extrapolate to the b mass and the continuum limit. This is most simply done by determining the functional dependence of the ratio

$$w(m_{\eta_h}, a) \equiv \frac{2m_{0h}}{m_{\eta_h}} \quad (40)$$

on the η_h mass and the lattice spacing. The ratio of $\overline{\text{MS}}$ masses is then given by the experimental masses of the η_c and η_b and the equation:

$$\frac{m_b(\mu, n_f)}{m_c(\mu, n_f)} = \frac{m_{\eta_b}^{\text{exp}} w(m_{\eta_b}^{\text{exp}}, 0)}{m_{\eta_c}^{\text{exp}} w(m_{\eta_c}^{\text{exp}}, 0)}. \quad (41)$$

It might seem simpler to fit m_{0h} directly, rather than the ratio w ; but using w significantly reduces the m_{η_h} dependence (and therefore our extrapolation errors), and also

makes our results quite insensitive to uncertainties in our values for the lattice spacing.

We parametrize function w with an expansion modeled after the one we used to fit the moments:

$$w(m_{\eta_h}, a) = Z_m(a) \left(1 + \sum_{n=1}^{N_w} w_n \left(\frac{2\Lambda}{m_{\eta_h}} \right)^n \right) / \left(1 + \sum_{i=1}^{N_{am}} \sum_{j=0}^{N_w} c_{ij} \left(\frac{am_{\eta_h}}{2} \right)^{2i} \left(\frac{2\Lambda}{m_{\eta_h}} \right)^j \right), \quad (42)$$

where, as for the moments,

$$i + j \leq \max(N_{am}, N_w). \quad (43)$$

Coefficients c_{ij} and w_n are determined by fitting function $w(m_{\eta_h}, a)$ to the values of $2am_{0h}/(am_{\eta_h})$ from Table II. The fit also determines the parameters $Z_m(a)$, one for each lattice spacing, which account for the running of the bare quark masses between different lattice spacings.

The finite- a dependence is smaller here than for the moments, because the η_h is nonrelativistic (finite- a errors are suppressed by additional powers of v/c [8]), and the variation with m_{η_h} stronger [twice that of $z(3, m_{\eta_h})$]. So here we use priors

$$c_{ij} = 0 \pm 0.05, \quad w_n = 0 \pm 4, \quad Z_m(a) = 1 \pm 0.5, \quad (44)$$

with $N_w = 8$. We again take $N_{am} = 30$, although identical results are obtained with $N_{am} = 15$.

Our fit results are illustrated by Fig. 4 which plots the ratio m_{0h}/m_{η_h} divided by m_{0c}/m_{η_c} for a range of η_h masses. Our data for different lattice spacings are com-

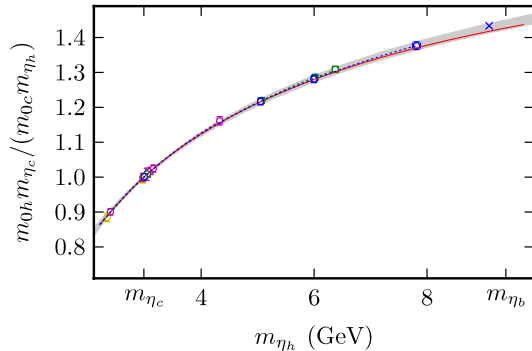


FIG. 4 (color online). Ratio m_{0h}/m_{η_h} divided by m_{0c}/m_{η_c} [which we approximate by $w(m_{\eta_c}, a)/2$ from our fit] as a function of m_{η_h} . The solid line shows the $a = 0$ extrapolation obtained from our fit. This is compared with simulation results for our 4 smallest lattice spacings, together with the best fits (dashed lines) corresponding to those lattice spacings. The gray band shows the $a = 0$ extrapolation for this quantity obtained from our best fit to the moments in the perturbative analysis of the previous section. The point marked by an \times is for the largest mass we tried (last line in Table II); this was not included in the fit because am_{η_h} is too large.

pared with our fit, and with the $a = 0$ limit of our fit (solid line). The fit is excellent, with $\chi^2/22 = 0.42$ for the 22 pieces of data we fit (we again exclude cases with $am_{\eta_h} > 1.95$). We also plot the range of values expected (at $a = 0$) for each mass from our perturbative analysis of the moments (gray band). The two methods agree to within errors for all quark masses.

Using the η_c and η_b masses from Sec. IV B, and Eq. (41) with the best-fit values for the parameters, we obtain finally

$$\frac{m_{0b}}{m_{0c}} \rightarrow 4.49(4) \quad \text{as } a \rightarrow 0 = \frac{m_b(\mu, n_f)}{m_c(\mu, n_f)}. \quad (45)$$

This agrees well with our result from the moments [Eq. (36)].

VI. $\alpha_{\overline{\text{MS}}}$ FROM WILSON LOOPS

In a recent paper [29], we presented a very accurate determination of the QCD coupling from simulation results for Wilson loops. Here we want to compare those results to the value we obtain from heavy-quark correlators. First,

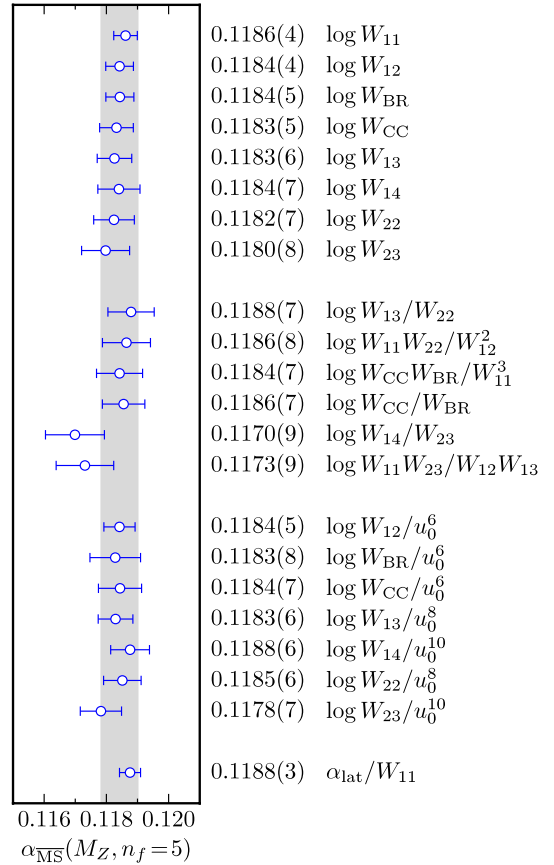


FIG. 5 (color online). Updated values for the 5-flavor $\alpha_{\overline{\text{MS}}}$ at the Z-meson mass from each of 22 different short-distance quantities built from Wilson loops. The gray band indicates a composite average, 0.1184(6). χ^2 per data point is 0.3.

however, we must update our earlier analysis to take account of the new value for r_1 [10] given in Eq. (10) and improved values for r_1/a [13] given in Table I. [The Wilson-loop paper uses some additional configuration sets: from Table II in that paper, sets 1, 6, 9, and 11 whose new r_1/a 's are 1.813(8), 2.644(3), 5.281(8), and 5.283(8), respectively.] We have rerun our earlier analysis, updating r_1 , r_1/a , and the c and b masses. The results are shown in Fig. 5. Combining results as in the earlier paper we obtain a final value from the Wilson-loop quantities of

$$\alpha_{\overline{\text{MS}}}(M_Z, n_f = 5) = 0.1184(6), \quad (46)$$

with $\chi^2/22 = 0.3$ for the 22 quantities in the figure. This agrees very well with the result in the earlier paper, $\alpha_{\overline{\text{MS}}}(M_Z) = 0.1183(8)$, but has a slightly smaller error, as expected given the smaller error in r_1 . This new value also agrees well with our very different determination from heavy-quark correlators [Eq. (38)]. A breakdown of the error into its different sources can be found in Table IV of [29] (reduce the r_1 and r_1/a errors in that table by one-half to account for the improved values used here).

VII. CONCLUSIONS

In this paper, we improve significantly on our previous determinations of the QCD coupling and c -quark mass from heavy-quark correlators. This is principally due to the inclusion of a new, smaller lattice spacing in our analysis. We also generated results for a variety of quark masses near m_c , allowing us to interpolate more accurately to the physical value of m_c . New third-order perturbation theory makes R_{10} as useful now as R_4 , R_6 , and R_8 were in the earlier paper. Finally, in this paper, we fit multiple moments simultaneously, determining consistent values simultaneously for both the QCD coupling and the quark masses for all moments. Previously we examined each moment or ratio of moments independently, extracting m_c 's or $\alpha_{\overline{\text{MS}}}$'s independently of each other. Our new results,

$$\begin{aligned} m_c(3 \text{ GeV}, n_f = 4) &= 0.986(6) \text{ GeV}, \\ \alpha_{\overline{\text{MS}}}(M_Z, n_f = 5) &= 0.1183(7), \end{aligned} \quad (47)$$

agree well with our older results of 0.986(10) GeV and 0.1174(12), respectively [1].

The much heavier b quark is usually analyzed using effective field theories like NRQCD or the static-quark approximation. By using very small lattice spacings and the very highly improved HISQ discretization for the heavy quarks, we are able to extend our analysis almost to the b -quark mass, using the same relativistic discretization that we use for c and lighter quarks. A 1.5% extrapolation of $z(3, m_h)$, from the largest m_{η_h} used in our fits to m_{η_b} , gives us a new, accurate determination of the b -quark mass,

$$m_b(10 \text{ GeV}, n_f = 5) = 3.617(25) \text{ GeV}. \quad (48)$$

This calculation demonstrates the utility of the HISQ formalism for studying b quarks on lattices that are computationally accessible today. This represents a breakthrough for b physics on the lattice since far greater precision becomes possible when all quarks are treated using the same formalism, and that formalism is relativistic and has a chiral symmetry. Even better would be to work right at the b mass, as opposed to extrapolating from nearby; this would require a lattice spacing of order 0.03 fm.

Both of our new c and b masses agree well with non-lattice determinations from vector-current correlators and experimental e^+e^- collisions. A recent analysis of the continuum data gives [7]

$$\begin{aligned} m_c(3 \text{ GeV}, n_f = 4) &= 0.986(13) \text{ GeV}, \\ m_b(m_b, n_f = 5) &= 4.163(16) \text{ GeV}, \end{aligned} \quad (49)$$

which compare well with our values of 0.986(6) and 4.164(23) GeV, respectively. This provides strong evidence that the different systematic errors in each calculation are understood.

Function $z(\mu/m_h, m_{\eta_h})$ is a by-product of our analysis. It relates the $\overline{\text{MS}}$ quark mass $m_h(\mu)$ to the η_h mass [Eq. (7)]. We show our result again in Fig. 6 for $\mu = 3m_h$, as well as for $\mu = m_h$ and $\mu = m_h/2$, which we obtain by evolving perturbatively from $\mu = 3m_h$. The latter two curves are relatively flat, and the last surprisingly close to 1 for most masses.

Questions have been raised about the way perturbation theory is used in analyzing the perturbative parts of the moments [30]. Like [7] we favor using larger scales than m_c for c -quark correlators, but, as we have shown, our results are quite insensitive to μ over a broad range. Furthermore, the fact that our results, from pseudoscalar-density correlators, agree so well with the continuum results, from vector-current correlators, is also compelling

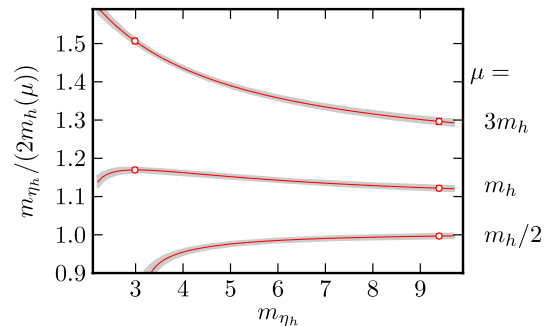


FIG. 6 (color online). $z(\mu/m_h, m_{\eta_h})$ versus m_{η_h} (in GeV) for three different values of μ/m_h . The curve for $\mu = 3m_h$ comes from the best fit to the moments. The other curves are obtained by evolving perturbatively from $\mu = 3m_h$.

evidence that perturbation theory is being handled correctly. We also find consistent results from several different moments, which is only possible if perturbation theory is working well. Compare, for example, Fig. 7 for the moments, as a function of m_{η_h} , with the plots of R_n/r_n in Fig. 1. Figure 7 shows very different m_{η_h} behavior, at the 10%–20% level, for different moments of R_n ; Fig. 1, where the perturbative part r_n is divided out, shows behavior that is almost moment independent.

An additional check on our use of perturbation theory comes from the close agreement between our perturbative result for the ratio m_b/m_c of $\overline{\text{MS}}$ masses [Eq. (36)] and our nonperturbative result for the ratio of HISQ masses [Eq. (45)]. These should be and are equal to within our 1% errors. Taken together they suggest a composite result of

$$\frac{m_b(\mu, n_f)}{m_c(\mu, n_f)} = 4.51(4) \quad (\text{composite}). \quad (50)$$

The validity of our perturbative analyses is further supported by the close agreement between the QCD coupling we get from the heavy-quark correlators, $\alpha_{\overline{\text{MS}}}(M_Z) = 0.1183(7)$, and that obtained from Wilson loops, 0.1184(6). These are radically different methods for determining the coupling. The first relies upon a continuum quantity, extrapolated to $a = 0$, and continuum perturbation theory. The second relies upon quantities that are highly sensitive to the UV cutoff (π/a) but are analyzed to all orders in the cutoff using lattice perturbation theory. Systematic errors are almost completely different in the two cases. The fact that they agree to within our 0.6% uncertainties is highly nontrivial evidence that perturbative and other potential errors are understood.

Our coupling values also agree well with determinations from nonlattice methods. Figure 8 summarizes recent re-

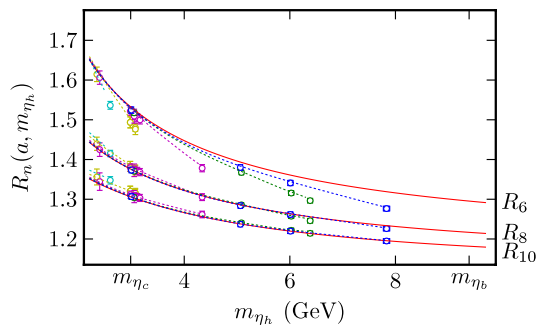


FIG. 7 (color online). Simulation results for reduced moments R_n with $n = 6, 8,$ and 10 as functions of m_{η_h} for 5 different lattice spacings. The dashed lines show the corresponding behavior of our fit function, with the best-fit parameters. The curves for smaller lattice spacings extend farther to the right. The solid lines show the $a = 0$ limit of our best fit.

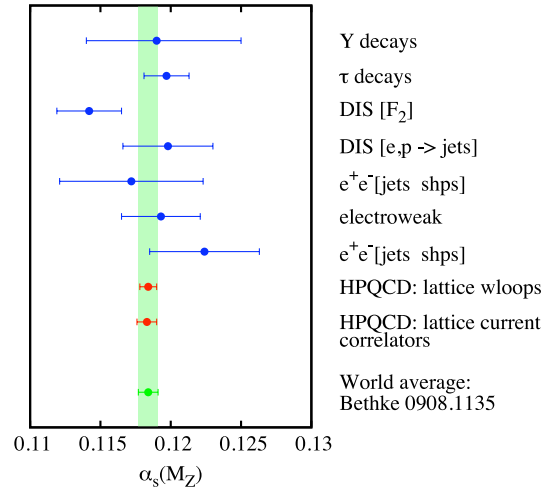


FIG. 8 (color online). The 5-flavor QCD coupling $\alpha_{\overline{\text{MS}}}$ at the Z mass as determined by a variety of different methods. The nonlattice numbers used here are from the review in [31].

sults that were included in a world average by Bethke [31]. The world average result, 0.1184(7), was dominated by our previous determination from the Wilson-loop analysis. The average excluding our result was 0.1186(11), which also agrees well. Including our new results into a new error-weighted world average gives $\alpha_{\overline{\text{MS}}}(M_Z) = 0.1184(4)$.

Our new c mass is the most accurate currently available. With it we can improve slightly on our recent determination of light-quark masses using an accurate value for m_c/m_s , 11.85(16), derived completely nonperturbatively from lattice calculations [11]. Our new c mass, which becomes 1.093(6) GeV when converted to $n_f = 3$ at 2 GeV, implies

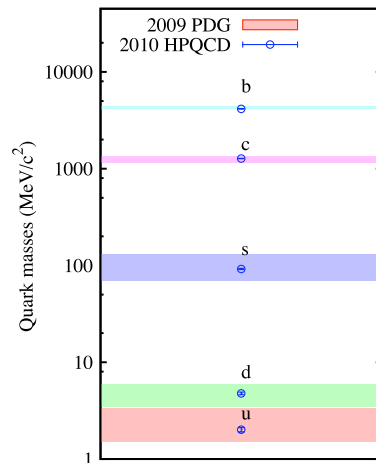


FIG. 9 (color online). $\overline{\text{MS}}$ masses, for the 5 lightest quarks, from this paper compared with the Particle Data Group’s current estimates [41]. Each mass is quoted at its conventional scale: 2 GeV for u, d, s ($n_f = 3$); m_c for c ($n_f = 4$); and m_b for b ($n_f = 5$).

$$\begin{aligned}
m_s(2 \text{ GeV}, n_f = 3) &= 92.2(1.3) \text{ MeV}, \\
m_d(2 \text{ GeV}, n_f = 3) &= 4.77(15) \text{ MeV}, \\
m_u(2 \text{ GeV}, n_f = 3) &= 2.01(14) \text{ MeV}, \\
(m_u + m_d)/2 &= 3.39(6) \text{ MeV},
\end{aligned} \tag{51}$$

where we have allowed for correlations in lattice-spacing errors between the m_c and m_c/m_s analyses. Our results for all 5 quark masses are compared with the Particle Data Group's 2009 values in Fig. 9. Agreement is excellent, but our uncertainties are much smaller in every case, and by an order of magnitude for the strange and light quarks.

Finally we note that the consistency between quark masses from lattice and nonlattice analyses, and between couplings from heavy-quark correlators and Wilson loops provides further evidence that taste-changing interactions in the HISQ and ASQTAD quark formalisms are understood and vanish as $a \rightarrow 0$. While early concerns about the validity of these formalisms have been largely addressed both by formal arguments [13,32–35,13] and by extensive empirical studies [8–11,29,36–40], it remains important to test the simulation technology of lattice QCD with increasing precision, given the growing importance of lattice results for phenomenology.

ACKNOWLEDGMENTS

We are grateful to MILC for configurations and thank Hans Kühn and Jack Laiho for useful discussions. Computing was done at the Ohio Supercomputer Center, USQCDs Fermilab cluster, and at the Argonne Leadership Computing Facility supported by DOE-AC02-06CH11357. We used chroma for some analyses [42]. We acknowledge support by the STFC, SUPA, MICINN, NSF, and DoE.

APPENDIX: ACCELERATED FITTING

In Sec. III B 2 we used a trick to simplify our fits by, in effect, transferring fit terms from the fit function into the errors of the fit data. This trick can greatly speed up complicated fits. Here we present a formal derivation of this procedure for three increasingly complicated situations.

1. Linear least squares—exact data

Assuming we know D values y_i for a quantity y which can be expressed as a power series in x ,

$$y = \sum_n c_n x^n, \tag{A1}$$

we wish to obtain a best fit for the first F unknown coefficients c_n . The c_n are then our random variables. If we are able to make reasonable estimates for their means and standard deviations σ_n , in the absence of additional information, maximizing entropy suggests a Gaussian

prior of

$$P(c) \propto e^{-\sum_n c_n^2 / 2\sigma_n^2}. \tag{A2}$$

For simplicity, we assume throughout that the c_n are uncorrelated and have a prior mean of zero; extending to more general cases is straightforward.

If we knew all coefficient values, then the data y_i would be completely determined, with

$$P(y|c) \propto \prod_{i=0}^{D-1} \delta\left(y_i - \sum_n c_n x_i^n\right). \tag{A3}$$

Bayes' theorem

$$P(c|y) \propto P(y|c)P(c) \tag{A4}$$

allows us to convert this into a distribution for c given the data y .

If we are only interested in fitting a subset of coefficients $c_{n<}$ with $n < F$, we integrate over the remaining $c_{n>}$, giving

$$\begin{aligned}
P(c_{<}|y) &\propto e^{-\sum_{n<} c_{n<}^2 / 2\sigma_n^2} \\
&\times \left[\int dc_{>} \delta^D\left(y - \sum_n c_n x^n\right) e^{-\sum_{n>} c_{n>}^2 / 2\sigma_n^2} \right]. \tag{A5}
\end{aligned}$$

We replace the delta function by its Fourier representation, integrate over first the $c_{n>}$, then the Fourier variables, to obtain

$$P(c_{<}|y) \propto e^{-\sum_{n<} c_{n<}^2 / 2\sigma_n^2} (\det \sigma_\Delta^2)^{-1/2} e^{-\Delta y \cdot (2\sigma_\Delta^2)^{-1} \cdot \Delta y}. \tag{A6}$$

Here

$$\Delta y_i \equiv y_i - \sum_{n<} c_{n<} x_i^n \tag{A7}$$

is the discrepancy between the measured y_i and the portion of the series to be kept in the fit, the dot product sums over the D data points, and

$$\sigma_{\Delta ij}^2 \equiv \sum_{n>} x_i^n \sigma_n^2 x_j^n. \tag{A8}$$

The correlation matrix σ_Δ^2 is independent of $c_{<}$ (so the determinant is constant), and is the same as one would compute directly by

$$\langle \Delta y_i \Delta y_j \rangle_{c_{>}} = \left\langle \sum_{m>} c_{m>} x_i^m \sum_{n>} c_{n>} x_j^n \right\rangle_{c_{>}} \tag{A9}$$

using

$$\langle c_{m>} c_{n>} \rangle_{c_{>}} = \sigma_n^2 \delta_{mn}. \tag{A10}$$

Finally, we fit $c_{<}$ by minimizing χ^2 , which includes these correlations and is augmented by the remaining $c_{<}$ priors.

Because the distribution is Gaussian, the $c_<$ at their minima are equal to their average values.

The correlation matrix σ_Δ^2 properly accounts for correlations in the discrepancy, due to the neglected terms, between y and the portion of the series retained. If F terms are kept in the series, σ_Δ is $\mathcal{O}(x^F)$, enforcing agreement between y and the finite series to this order, as appropriate. It also suggests an alternative but equivalent approach. We may define new random (rather than exact) versions of y , whose correlation matrix is σ_Δ^2 , by moving the $c_>$ terms to the left side of Eq. (A1). Using the truncated series as a model for these random data, straightforward application of Bayes' theorem [19] again implies the distribution in Eq. (A6).

One useful consequence is that, as long as we include the correlations for the $c_>$, we may arbitrarily reduce the number of coefficients $c_<$ retained, even to as few as 1, and still obtain the same minimization values. To see this, note that to compute a particular $\langle c_{n_<} \rangle$, we could start with the full distribution and integrate over all c 's. The integral over $c_>$ produces $P(c_<|y)$, which we then use in the integral over $c_<$; the result will be the same regardless of where the dividing line is set, as long as it does not include $c_{n_<}$. (We could even include in σ_Δ^2 terms of order less than n .) Because averaging and minimization give the same result, the minimization value for $c_{n_<}$ will also remain unchanged. This is also true of the $c_{n_<}$ error. While the result is the same, reducing the number of terms in the series to fit can significantly improve the fitting time.

2. Fits to nonlinear functions—exact data

We now consider fitting to the data y_i a general function $g_i(c_n)$ not necessarily linear in the parameters c , and where we assume $y_i = g_i(c_n)$ exactly for properly chosen c_n . Now

$$P(y|c) \propto \prod_{i=0}^{D-1} \delta\left(y_i - \sum_n g_i(c_n)\right). \quad (\text{A11})$$

Combining with the prior $P(c)$ and integrating over the $c_>$ gives $P(c_<|y)$.

If our estimate of prior means is good, expanding g around $c_> = 0$ should give a reasonable approximation; an expansion to first order gives a Gaussian. More specifically, defining

$$g_i(c_<) \equiv g_i(c_<, c_> = 0) \quad (\text{A12})$$

and

$$\Delta y_i \equiv y_i - g_i(c_<), \quad (\text{A13})$$

and integrating over $c_>$ in this Gaussian approximation gives as before

$$P(c_<|y) \propto e^{-\sum_{n_<} c_{n_<}^2/2\sigma_n^2} (\det \sigma_\Delta^2(c_<))^{-1/2} e^{-\Delta y \cdot (2\sigma_\Delta^2(c_<))^{-1} \cdot \Delta y}, \quad (\text{A14})$$

but with

$$\sigma_\Delta^2(c_<)_{ij} \equiv \sum_{n_>} \partial_n g_i(c_<) \sigma_n^2 \partial_n g_j(c_<). \quad (\text{A15})$$

This is again the correlation one would compute directly for $\langle \Delta y_i \Delta y_j \rangle_{c_>}$ after expanding g to first order in $c_>$.

We have not expanded in $c_<$, so σ_Δ^2 depends on $c_<$, the determinant in front is not constant, and the dependence of Δy on $c_<$ is not in general linear. In practice, however, we will often further approximate the distribution by setting the $c_<$ to their prior means in $\sigma_\Delta^2(c_<)$ before minimization.

Because $g(c_<)$ is nonlinear, $c_<$ from minimization can differ slightly from $\langle c_< \rangle$, and due to approximations made, can vary somewhat with the number of terms retained.

3. Fits to data with intrinsic statistical errors

Finally we consider the most general case, in which the data y contribute intrinsic statistical uncertainties in addition to those associated with the truncated series. If we measure a range of values for y with an average $\langle y \rangle$ and correlation matrix σ_y^2 , then for sufficiently large samples we expect a Gaussian distribution

$$P(\langle y \rangle | c) \propto e^{-\langle (y-g(c)) \cdot (2\sigma_y^2)^{-1} \cdot (y-g(c)) \rangle} \quad (\text{A16})$$

rather than the delta function above. Combining with the prior $P(c)$ gives $P(c|\langle y \rangle)$.

Expanding $g_i(c_<, c_>)$ to first order around $c_> = 0$, defining

$$\Delta y_i \equiv \langle y_i \rangle - g_i(c_<), \quad (\text{A17})$$

and integrating $P(c|\langle y \rangle)$ over $c_>$ gives

$$P(c_<|\langle y \rangle) \propto e^{-\sum_{n_<} c_{n_<}^2/2\sigma_n^2} (\det \sigma_{y\Delta}^2(c_<))^{-1/2} e^{-\Delta y \cdot (2\sigma_{y\Delta}^2(c_<))^{-1} \cdot \Delta y}. \quad (\text{A18})$$

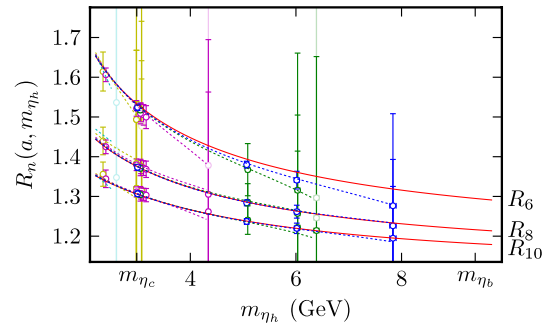


FIG. 10 (color online). Same as Fig. 7 but with $N_{am} = 80$ and $\bar{N}_{am} = 3$, instead of $N_{am} = \bar{N}_{am} = 30$. The error bars are almost entirely due to systematic errors caused by $am_{\eta_b}/2$ corrections omitted from the fit function.

The resulting correlation matrix is a combination of true statistical and neglected series contributions, with

$$\sigma_{y\Delta}^2(c_<) \equiv \sigma_y^2 + \sigma_\Delta^2(c_<), \quad (\text{A19})$$

as one would obtain by including both sources of uncertainty in computing $\langle \Delta y_i \Delta y_j \rangle$ directly. With no statistical fluctuations in y , $\sigma_y^2 = 0$, and it reduces to the previous result. When σ_y^2 is nonzero but small, σ_Δ^2 still makes an important contribution.

4. Application to this paper

We used the technique described here in much of our testing and tuning (but not for our final results) to speed up the $(am_{\eta_h}/2)^2$ fit. As described in Sec. III B 2, we kept

corrections through order $N_{am} = 80$ but moved all but $\bar{N}_{am} \ll N_{am}$ out of the fit function and into the errors for the reduced moments. If we set $\bar{N}_{am} = 3$, for example, our fit to the R_n simulation data changes from Fig. 7 to Fig. 10. The small \bar{N}_{am} means that each point in Fig. 10 has much larger error bars, coming from $(am_{\eta_h}/2)^2$ terms moved into the R_n 's. The final fit results, however, are almost identical in both cases (to within less than 0.1σ), with the same errors. Note that the R_n errors in Fig. 10 are highly correlated, which is why the fit curve passes through the central value for each point. As discussed above these correlations are essential if results are to be independent of the value of \bar{N}_{am} .

-
- [1] I. Allison *et al.* (HPQCD Collaboration), *Phys. Rev. D* **78**, 054513 (2008).
- [2] K. G. Chetyrkin, J. H. Kuhn, and C. Sturm, *Eur. Phys. J. C* **48**, 107 (2006).
- [3] R. Boughezal, M. Czakon, and T. Schutzmeier, *Phys. Rev. D* **74**, 074006 (2006).
- [4] A. Maier, P. Maierhofer, and P. Marquard, *Phys. Lett. B* **669**, 88 (2008).
- [5] A. Maier, P. Maierhofer, P. Marquard, and A. V. Smirnov, *Nucl. Phys.* **B824**, 1 (2010).
- [6] Y. Kiyo, A. Maier, P. Maierhofer, and P. Marquard, *Nucl. Phys.* **B823**, 269 (2009).
- [7] K. G. Chetyrkin, J. H. Kuhn, A. Maier, P. Maierhofer, P. Marquard, M. Steinhauser, and C. Sturm, *Phys. Rev. D* **80**, 074010 (2009).
- [8] E. Follana *et al.* (HPQCD Collaboration), *Phys. Rev. D* **75**, 054502 (2007).
- [9] E. Follana, C. T. H. Davies, G. P. Lepage, and J. Shigemitsu (HPQCD Collaboration), *Phys. Rev. Lett.* **100**, 062002 (2008).
- [10] We use the value from C. T. H. Davies, E. Follana, I. D. Kendall, G. P. Lepage, and C. McNeile (HPQCD Collaboration), *Phys. Rev. D* **81**, 034506 (2010). This paper references a separate determination by the MILC Collaboration, which agrees well with the value used here.
- [11] C. T. H. Davies *et al.* (HPQCD Collaboration), *Phys. Rev. Lett.* **104**, 132003 (2010).
- [12] E. B. Gregory *et al.* (HPQCD Collaboration), *Phys. Rev. Lett.* **104**, 022001 (2010).
- [13] A. Bazavov *et al.*, *Rev. Mod. Phys.* **82**, 1349 (2010).
- [14] A. Gray, I. Allison, C. T. H. Davies, E. Gulez, G. P. Lepage, J. Shigemitsu, and M. Wingate (HPQCD and UKQCD Collaborations), *Phys. Rev. D* **72**, 094507 (2005).
- [15] E. Gamiz, C. T. H. Davies, G. P. Lepage, J. Shigemitsu, and M. Wingate (HPQCD Collaboration), *Phys. Rev. D* **80**, 014503 (2009).
- [16] Annihilation contributions from $h\bar{h} \rightarrow \text{gluons} \rightarrow h\bar{h}$ are omitted from both our perturbation theory and our simulation. They are also omitted from our calculation of the η_h mass, but are negligible there (see Sec. II in [1]).
- [17] Both $G(t)$ and its moments G_n are expressed in units of the lattice spacing (“lattice units”), following standard practice in lattice QCD. This means that extra factors of a or $1/a$ are included in the definitions of $G(t)$ and G_n to make them dimensionless. These factors must be dropped, obviously, to obtain the analogous continuum quantity. When we say that something is UV-cutoff independent, we mean that it is independent after it has been converted into physical units (by dropping the extra factors of a or $1/a$).
- [18] We have simplified the definition of R_n relative to our first paper by replacing the tree-level pole mass of the heavy quark with the bare mass parameter from the lattice. The two masses differ by $\mathcal{O}((am_h)^4)$, which vanishes in the continuum limit.
- [19] G. P. Lepage, B. Clark, C. T. H. Davies, K. Hornbostel, P. B. Mackenzie, C. Morningstar, and H. Trotter, *Nucl. Phys. B, Proc. Suppl.* **106–107**, 12 (2002).
- [20] T. van Ritbergen, J. A. M. Vermaseren, and S. A. Larin, *Phys. Lett. B* **400**, 379 (1997).
- [21] M. Czakon, *Nucl. Phys.* **B710**, 485 (2005).
- [22] K. G. Chetyrkin, *Phys. Lett. B* **404**, 161 (1997).
- [23] J. A. M. Vermaseren, S. A. Larin, and T. van Ritbergen, *Phys. Lett. B* **405**, 327 (1997).
- [24] K. G. Chetyrkin, B. A. Kniehl, and M. Steinhauser, *Nucl. Phys.* **B510**, 61 (1998).
- [25] D. J. Broadhurst, P. A. Baikov, V. A. Ilyin, J. Fleischer, O. V. Tarasov, and V. A. Smirnov, *Phys. Lett. B* **329**, 103 (1994).
- [26] For a recent analysis of condensate values see B. L. Ioffe, *Prog. Part. Nucl. Phys.* **56**, 232 (2006).
- [27] We use the Particle Data Group value for m_{η_c} : C. Amsler *et al.* (Particle Data Group), *Phys. Lett. B* **667**, 1 (2008) and 2009 partial update for the 2010 edition [http://pdg.lbl.gov]. We correct the η_c mass for quark annihilation [2.4(1.2) MeV [8]], and electromagnetic corrections [2.6(1.2) MeV], which are dominated by the Coulombic

- interaction between the heavy quark and antiquark. Neither effect is in our simulation. See [10] for more details.
- [28] We average over recent experiments to obtain a value of 9.391(3) GeV for m_{η_b} , and then include corrections for quark annihilation and electromagnetic corrections (see [27]) that together we estimate at 4(4) MeV. For the experimental results see B. Aubert *et al.* (BABAR Collaboration), *Phys. Rev. Lett.* **101**, 071801 (2008); **102**, 029901(E) (2009); G. Bonvicini *et al.* (CLEO Collaboration), *Phys. Rev. D* **81**, 031104 (2010); B. Aubert *et al.* (BABAR Collaboration), *Phys. Rev. Lett.* **103**, 161801 (2009).
- [29] C. T. H. Davies, K. Hornbostel, I. D. Kendall, G. P. Lepage, C. McNeile, J. Shigemitsu, and H. Trotter (HPQCD Collaboration), *Phys. Rev. D* **78**, 114507 (2008).
- [30] For a related discussion, see J. H. Kuhn, Proc. Sci., RADCOR2009 (2010) 035.
- [31] S. Bethke, *Eur. Phys. J. C* **64**, 689 (2009).
- [32] S. R. Sharpe, Proc. Sci., LAT2006 (2006) 022 [arXiv:hep-lat/0610094].
- [33] C. Bernard, M. Golterman, Y. Shamir, and S. R. Sharpe, *Phys. Rev. D* **77**, 114504 (2008).
- [34] A. S. Kronfeld, Proc. Sci., LAT2007 (2007) 016 [arXiv:0711.0699].
- [35] M. Golterman, Proc. Sci., CONFINEMENT8 (2008) 014 [arXiv:0812.3110].
- [36] C. T. H. Davies *et al.* (HPQCD Collaboration, UKQCD Collaboration, MILC Collaboration, and Fermilab Lattice Collaboration), *Phys. Rev. Lett.* **92**, 022001 (2004).
- [37] E. Follana, A. Hart, and C. T. H. Davies (HPQCD Collaboration and UKQCD Collaboration), *Phys. Rev. Lett.* **93**, 241601 (2004).
- [38] E. Follana, A. Hart, C. T. H. Davies, and Q. Mason (HPQCD Collaboration and UKQCD Collaboration), *Phys. Rev. D* **72**, 054501 (2005).
- [39] E. Follana, A. Hart, and C. T. H. Davies, Proc. Sci., LAT2005 (2006) 298.
- [40] E. Follana, C. T. H. Davies, and A. Hart (UKQCD and HPQCD Collaborations), Proc. Sci., LAT2006 (2006) 051.
- [41] C. Amsler *et al.* (Particle Data Group), *Phys. Lett. B* **667**, 1 (2008) and 2009 partial update for the 2010 edition (<http://pdg.lbl.gov>).
- [42] R. G. Edwards and B. Joo (SciDAC Collaboration, LHPC Collaboration, and UKQCD Collaboration), Nucl. Phys. B, Proc. Suppl. **140**, 832 (2005).