High-Precision $c$ and $b$ Masses, and QCD Coupling from Current-Current Correlators in Lattice and Continuum QCD

C. McNeile, C. T. H. Davies, E. Follana, K. Hornbostel, and G. P. Lepage

(HPQCD Collaboration)

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

(Dated: March 26, 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: $m_c(3\text{GeV},n_f=4)=0.986(6)\text{ GeV}$, $\alpha_{\text{QCD}}(M_Z,n_f=5)=0.1183(7)$, 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 $m_b(m_c,m_f)/m_c(m_f)$; 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 $\alpha_{\text{QCD}}$ from Wilson loops to account for revised values for $r_1$ and $r_2/a$, finding a new value $\alpha_{\text{QCD}}(M_Z,n_f=5)=0.1184(6)$; 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.

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

I. INTRODUCTION

Precise values for the QCD coupling $\alpha_{\text{QCD}}$ 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 $O(\alpha_s^3)$ 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 nonperturbatively 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 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 Section II. Then in Section III we describe our lattice QCD simulations and discuss in detail the chief systematic errors in our simulation results. In Section 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 Section V. We then, in Section VI, update our previous calculation of the QCD coupling from Wilson loops to compare with our new result from the correlators. We summarize our findings in Section 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-

---

*Current address: Dept. of Theoretical Physics, Univ. of Wuppertal, Wuppertal 42199, Germany
†Electronic address: g.p.lepage@cornell.edu
squares fits that can greatly reduce the computing re-
quired 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,
j5 = ψhγ5ψh:

\[ G(t) = a^6 \sum_x (amat)^2 0)j5(x,t)j5(0,0)0 \]  (1)

where mat is the heavy quark’s bare mass (from the lat-
tice QCD lagrangian), t is euclidean and periodic with
period T, and the sum over spatial positions 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 amat. Consequently
lattice and continuum G(t)s are equal when t ≠ 0 up to
O((amat)m) corrections, which vanish in the continuum
limit.

The moments of G(t) are particularly simple to ana-
lyze:

\[ G_n(t) = \sum_t (t/a)^n G(t) \]  (2)

where, on our periodic lattice,

\[ t/a \in \{0,1,2...T/2a-1,0,-T/2a+1...2,-1\}. \]  (3)

Low moments emphasize small ts and so are perturbative;
and moments with n ≥ 4 are UV-cutoff independent.
Therefore

\[ G_n = \frac{g_n(\alpha_{MS}(\mu),\mu/m_h)}{(amat(\mu))^{n-4}} + O((amat)^m) \]  (4)

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

Again following our previous paper, we introduce re-
duced moments to suppress both lattice artifacts and
tuning errors in the heavy quark’s mass [10]:

\[ R_n = \begin{cases} \frac{G_4/G_4(0)}{am_h} & \text{for } n = 4, \\ \frac{am_h}{2am_\text{MS}} (G_n/G_n(0))^{1/(n-4)} & \text{for } n \geq 6 \end{cases} \]  (5)

where \( G_n(0) \) is the moment in lowest-order, weak-coupling
perturbation theory, using the lattice regulator, and amh
is the (nonperturbative) mass of the pseudo-Goldstone
boson. The reduced moments can again be written in
terms of continuum quantities:

\[ R_n = \begin{cases} \frac{r_4(\alpha_{MS},\mu/m_h)}{2am_\text{MS}} & \text{for } n = 4, \\ z(\mu/m_h,m_{\text{MS}}) r_n(\alpha_{MS},\mu/m_h) & \text{for } n \geq 6 \end{cases} \]  (6)

up to \( O((amat)^m \alpha_s) \) corrections, where

\[ z(\mu/m_h,m_{\text{MS}}) = \frac{m_h}{2m_h(\mu)} \]  (7)

and \( r_n \) is obtained from \( g_n \) (Eq. (3)) and its value, \( g_n(0) \),
in lowest-order continuum perturbation theory:

\[ r_n = \begin{cases} \frac{g_4(0)}{g_4(0)} & \text{for } n = 4, \\ \frac{g_n(0)}{g_4(0)}^{1/(n-4)} & \text{for } n \geq 6 \end{cases} \]  (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 amh/amat. We then compare this simulation “data”
to the continuum perturbation theory formulas (Eq. (5)).
That is, we find values for \( \alpha_{\overline{\text{MS}}}(\mu) \) and \( z(\mu/m_h,m_{\text{MS}}) \)
that make lattice and continuum results agree for small \( n \geq 4 \).
The function \( z(\mu/m_h,m_{\text{MS}}) \) can then be combined with
experimental results for \( m_c \) and \( m_b \) to obtain masses
for the c and b quarks:

\[ m_c(\mu) = \frac{m_c^{\text{exp}}}{2z(\mu/m_c,m_c^{\text{exp}})} \]

\[ m_b(\mu) = \frac{m_b^{\text{exp}}}{2z(\mu/m_b,m_b^{\text{exp}})} \]  (9)

Parameter \( \mu \) sets the scale for \( \alpha_{\overline{\text{MS}}} \) in the perturbative
expansions of the \( r_n \). An obvious choice for this param-
eter 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 \( \mu \) 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 sep-
arátely in our previous paper. Here we extract them si-
multaneously, 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 param-
eters are listed in Table [1].

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 $\pi^0$-meson. We tune the strange quark mass to give the correct mass for the (fictitious) $\eta_s$ meson $^{10}$). The $c$ and $b$ masses are tuned to give correct masses for the $\eta_c$ and $\eta_b$ 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 $^{10}$

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

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.

We computed heavy-quark correlators (Eq. (11)) using the HISQ discretization $^8$ for a variety of bare heavy-quark masses $am_{\eta_b}$ on the MILC gluon configurations. Our results for the reduced moments $R_n$ with $n = 4-18$ are given in Table I.

In Table I we also give masses $am_{\eta_b}$ 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 $am_{\eta_b}$ by factors of 1.4 and 2 for the next-to-finest and finest lattice spacings (sets 8 and 9), respectively, to account for this. The statistical errors here are very small and have only a small impact on our final results. We also verified our results with multi-exponential fits in every case.

### 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 $r_1 = 0.3133(23) \text{ fm}$; values for $r_1/a$ are from $^{12}$. The bare quark masses are for the ASQRTAD 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 ($N_{gl}$) and the number of time sources ($N_c \times N_t$) per configuration used in each case. Sets with similar lattice spacings are grouped.

| Set | $r_1/a$ | $u_0$ | $L/a$ | $T/a$ | $N_{gl}$ | $N_c \times N_t$ |
|-----|---------|-------|-------|-------|----------|-----------------|
| 1   | 2.152(5)| 0.0097| 0.0484| 0.860 | 16       | 48 \( \times \) 631 \( \times \) 2 |  
| 2   | 2.138(4)| 0.0194| 0.0484| 0.861 | 16       | 48 \( \times \) 631 \( \times \) 2 |  
| 3   | 2.647(3)| 0.005 | 0.05  | 0.868 | 24       | 64 \( \times \) 678 \( \times \) 2 |  
| 4   | 2.618(3)| 0.01  | 0.05  | 0.868 | 20       | 64 \( \times \) 595 \( \times \) 2 |  
| 5   | 2.618(3)| 0.01  | 0.05  | 0.868 | 28       | 64 \( \times \) 269 \( \times \) 2 |  
| 6   | 3.679(3)| 0.0062| 0.031 | 0.878 | 28       | 64 \( \times \) 566 \( \times \) 2 |  
| 7   | 3.712(4)| 0.0124| 0.031 | 0.878 | 28       | 96 \( \times \) 265 \( \times \) 2 |  
| 8   | 5.296(7)| 0.0036| 0.018 | 0.888 | 48       | 144 \( \times \) 201 \( \times \) 2 |  
| 9   | 7.115(20)| 0.0028| 0.014 | 0.895 | 64       | 192 \( \times \) 208 \( \times \) 2 |  

### B. Systematic Errors

As discussed above, our goal is to find values for $\alpha_s(M)$ and $z(\mu/m_h, m_{\eta_b})$ (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 $4 \leq n \leq 10$.

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

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

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 $m_h/2$, 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 $(2\Lambda/m_{\eta_b})^2$ ranges approximately between 1/16 and 1/6; the singular point $m_h = 0$ is infinitely far away in this parameterization. In our fits we keep terms only through 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 $(am_h)^2 \alpha_s$ for $i \geq 1$. We model these by

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

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

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

where $c_i(n)$ are obtained from Eq. (11) with $\mu/m_h = 3$.
the $c_{ij}^{(n)}$s are determined in our fit, $R_n^{\text{cont}}$ is given by Eq. (9),

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

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

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

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 $2\Lambda/m_{qs}$ expansion, expansion parameter $am_{qs}/2$ ranges between 0.3 and 1.1, and higher orders are necessary, especially given our tiny statistical errors. We find that our fit results don’t converge well unless $N_{am}$ is larger than 10-20. Also we have difficulty getting good fits if we include data with $am_{qs} > 1.95$ from Table II. The $am_{qs}/2$ 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. 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 $R_n^{\text{latt}}$ in place of $R_n^{\text{latt}}$ where

$$R_n^{\text{latt}} = R_n^{\text{latt}} + \sum_{i=N_{am}+1}^{N_z} \sum_{n=0}^{N_z} c_{ij}^{(n)} \left( \frac{am_{qs}}{2} \right)^{2i} \left( \frac{2\Lambda}{m_{qs}} \right)^j$$

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

$$R_n^{\text{latt}} = R_n(\mu, m_h, a, \bar{N}_{am}).$$ \hspace{1cm} (19)

where $R_n(\ldots)$ is again given by Eq. (15). To evaluate $R_n^{\text{latt}}$ 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 $R_n^{\text{latt}}$ to obtain a new error estimate for $R_n^{\text{latt}}$ (but leaving the central value unchanged). In effect we are increasing the error in the reduced moment to account for high-order ($am_h/2$) terms omitted from the fit formula Eq. (19). By choosing $\bar{N}_{am} \ll N_{am}$, most of the $am_h/2$ terms are transferred when fits are linear (or almost linear) and the fit parameters with prior $N_{am}$, since the same $c_{ij}^{(n)}$’s are used for all as and $m_h$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 $\bar{N}_{am}$ very large—say, $\bar{N}_{am} = 80$—and still have very fast fits by keeping $\bar{N}_{am}$ very small. With $\bar{N}_{am} = 80$ we find, for example, that setting $\bar{N}_{am} = 0$ in $R_n(\ldots)$ (no terms) gives essentially identical results for our quark masses and coupling 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,

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

of the reduced moments is known at best through third order. We present coefficients $r_{nj}$ through $j = 3$ in Table III; 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

$$r_{n3}(1) = 0 \pm 0.5$$ \hspace{1cm} (21)

for any coefficient that hasn’t been computed in perturbation theory. We set $N_{\text{nth}} = 6$ since then contributions from still higher orders should be less than 0.1% (and setting $N_{\text{nth}} = 8$ doesn’t 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 Section 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 $j \leq N_{\text{nth}}$) in terms of the $r_{nj}(1)$ coefficients and $\log(\mu/m_h)$, and substitute the results from Table III for $j < 3$ 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.

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 $\mu$ 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 $\alpha_0$, with prior

$$\alpha_0 = 0.20 \pm 0.01,$$ \hspace{1cm} (22)

where

$$\alpha_0 \equiv \alpha_{\overline{\text{MS}}}(5 \text{ GeV}, n_f = 3).$$ \hspace{1cm} (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 $\alpha_0$ at scale 5 GeV. We use the $\overline{\text{MS}}$ beta function.

### Table III: Perturbation theory coefficients ($n_f = 3$)

| $n$ | $r_{n1}$ | $r_{n2}$ | $r_{n3}$ |
|-----|----------|----------|----------|
| 4   | 0.7427   | -0.0827  | 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   | 0.1133   |
| 14  | 0.0544   | 0.1674   | 0.1133   |
| 16  | 0.0146   | 0.1293   | 0.1133   |
| 18  | -0.0165  | 0.0965   | 0.1133   |
through sixth order in $\alpha_{\text{MS}}$

\[ \mu^2 \frac{d \alpha_{\text{MS}}(\mu)}{d \mu^2} = - \beta_0 \alpha_{\text{MS}}^2 - \beta_1 \alpha_{\text{MS}}^3 - \beta_2 \alpha_{\text{MS}}^4 \] (24)

\[ - \beta_3 \alpha_{\text{MS}}^5 - \beta_4 \alpha_{\text{MS}}^6, \]

where $\beta_0 \ldots \beta_3$ are known from perturbation theory and $\beta_4$ is taken as a fit parameter with prior

\[ \beta_4 = 0 \pm \sigma_\beta \] (25)

where $\sigma_\beta$ is the root-mean-square average of $\beta_0 \ldots \beta_3$ \[18, 19\]. 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 \[20, 21\] 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 $n$ and quark masses we use here. We correct for it by replacing

\[ R_{n}^{\text{latt}} \rightarrow R_{n}^{\text{latt}} \left( 1 + d_n \frac{\langle \alpha_s G^2 / \pi \rangle}{(2m_h)^4} \right) \] (26)

where $d_n$ is computed to leading order in perturbation theory \[22\] with $m_h = m_b(m_h)$, which we approximate by $m_{h/2}$. We take

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

which covers the range of most current estimates \[23\]. The correction factor in Eq. (26) adds (slightly) to the error in $R_{n}^{\text{latt}}$ (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

\[ R_{n}^{\text{latt}} \rightarrow R_{n}^{\text{latt}} \left( 1 + f_n \frac{\Delta R_{n}^{\text{th}}}{R_{n}^{\text{th}}} \right) \] (28)

where $\Delta R_{n}^{\text{th}}$ is the finite volume error in leading-order perturbation theory and

\[ f_n = 0 \pm 0.5. \] (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 II). 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

\[ R_{n}^{\text{latt}} \rightarrow R_{n}^{\text{latt}} \left( 1 + g_n \frac{2\delta m_t + \delta m_s}{m_s} \right) \] (30)

where $\delta m_t$ and $\delta m_s$ are the errors in the u/d and s masses (see \[10\] for more details), respectively, and

\[ g_n = 0 \pm 0.01. \] (31)

This correction introduces (correlated) errors into the $R_{n}^{\text{latt}}$’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 II 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 \(n = 4\) in Table I suggests that sea-quark mass effects are no larger than 0.1%, so our error estimate is conservative.

\[ \text{IV. ANALYSIS AND RESULTS} \]

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 \(n = 4\) in Table I. Our error estimate here is very conservative, but our error estimate is conservative.

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 \cite{17}. In this procedure we minimize an augmented \( \chi^2 \) function of the form

\[
\chi^2 = \sum_{in,jm} \Delta R_{ni} \left( \sigma^{-2}_{R} \right)_{in,jm} \Delta R_{mj} + \sum_i \delta \chi_i^2
\]  

(32)

where:

\[
\Delta R_{ni} \equiv R_{n}^{\text{platt}} - R_{n}(\mu_i, m_{\eta}, a_i; N_{am});
\]

(33)

the \( R_{n}^{\text{platt}} \) come from Table 11 with corrections from Eqs. (20), (28) and (30); fit function \( R_{n}(\ldots) \) is defined by Eq. (15); and \( \sigma^{-2}_{R} \) is the error covariance matrix for the \( R_{n}^{\text{platt}} \). The sums \( i, j \) are over the 22 sets of lattice spacings and quark masses; the sums \( n, m \) range over of the moments 4, 6, 8, 10.

Function \( R_{n}(\mu_i, m_{\eta}, a_i; N_{am}) \) depends upon a large number of parameters, all of which are varied in the fit to minimize \( \chi^2 \). Priors \( \delta \chi_i^2 \) are included for each of these:

- parameters \( z_j \), with prior Eq. (13), from the \( 1/m_{\eta} \) expansion of \( z(\mu/m_h, m_{\eta}) \);
- parameters \( c_i^{(m)} \), with prior Eq. (17), from the finite-lattice spacing corrections;
- unknown perturbative coefficients \( r_{nj} \), with prior Eq. (21) (evolved to \( \mu/m_h = 3 \));
- coupling parameter \( \log(\alpha_0) \), with prior Eq. (22);
- \( \beta_4 \) in the QCD \( \beta \)-function, with prior Eq. (25);
- lattice spacings \( a_n \) for each gluon configuration set, with priors specified by simulation results for \( r_2/a \) (Table 1) and the current value for \( r_1 \) (Eq. (10));
- values for \( am_{\eta} \), with priors specified by our simulation results (Table 11).

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

**B. Results**

We fit our simulation data for the reduced moments \( R_{n}^{\text{platt}} \) (Table 11) using fit function \( R_{n}(\ldots) \) (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 \( z(\mu/m_h = 3, m_{\eta}) \) (Eq. (7)), which we plot in Figure 1. We also show our simulation results there for \( R_{n}^{\text{platt}}/r_n \), 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 all fit simultaneously, using the same functions \( z(3, m_{\eta}) \) and \( \alpha_{MS}(\mu) \) (with \( \mu = 3m_{\eta}/(2z) \)) for all moments. The fits are excellent, with \( \chi^2/88 = 0.19 \) for the 88 pieces of simulation data we fit.

**FIG. 1:** Function \( z(\mu/m_h = 3, m_{\eta}) \equiv m_{\eta}/(2m_h) \) as a function of \( m_{\eta} \). 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 further to the right. The points marked by an “x” are for the largest mass we tried (last line in Table 11); these are not included in the fit because \( am_{\eta} \) is too large. Finite-a errors become very small for the larger-\( \eta \) moments, causing points from different lattice spacings to overlap.

Evaluated at \( m_{\eta c} = 2.985(3) \) GeV \cite{24}, the mass-ratio function is \( z(3, m_{\eta}) = 1.507(7) \). Combining this with Eq. (9) and perturbation theory, we can obtain the following results for the \( MS \) c-quark mass at different scales:

\[
m_c(3m_c, n_f = 3) = 0.991(5) \text{ GeV},
\]

(34)

\[
m_c(3 \text{ GeV}, n_f = 4) = 0.986(6) \text{ GeV},
\]

\[
m_c(3 \text{ GeV}, n_f = 4) = 1.273(6) \text{ GeV}.
\]

Similarly at \( m_{\eta_b} = 9.395(5) \) GeV \cite{23}, the mass-ratio function is \( z(3, m_{\eta_b}) = 1.296(8) \), and we obtain the following results for the \( MS \) b-quark mass at different scales:

\[
m_b(3m_b, n_f = 3) = 3.622(22) \text{ GeV},
\]

(35)

\[
m_b(10 \text{ GeV}, n_f = 5) = 3.617(25) \text{ GeV},
\]

\[
m_b(10 \text{ GeV}, n_f = 5) = 4.164(23) \text{ GeV}.
\]

Note that the ratio \( m_b(\mu, n_f)/m_c(\mu, n_f) \) is independent of \( \mu \) and \( n_f \). We obtain the following result for this
we add vacuum polarization corrections from the quarks, using the masses above, and evolve to the b meson mass [18–21]:

\[ \alpha \approx \sqrt{\frac{2}{3} \alpha_0} \]

The data points are values of the best-fit coupling from our fit when perturbative evolution is assumed. The points from different lattice spacings overlap in these plots. Several points from different lattice spacings overlap in these plots.

\[ m_b/m_c = 4.53(4) \]

(36)

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

\[ \alpha_0 = \alpha_{\overline{\text{MS}}}(5 \, \text{GeV}, n_f = 3) = 0.2034(21). \]

(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 [18–21]:

\[ \alpha_{\overline{\text{MS}}}(M_Z, n_f = 5) = 0.1183(7). \]

(38)

Figure 2 shows how consistent our simulation results are with the theoretical curve for \( \alpha_{\overline{\text{MS}}}^p(\mu, n_f = 3) \) corresponding to our value for \( \alpha_0 \). For this figure we extracted values for \( \alpha_{\overline{\text{MS}}} \) from each \( R_n \) separately by dividing out the \( a^2 \) dependence and \( z(3, m_{\eta_b}) \) 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 non-perturbative 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 \( m_b \approx m_c \) in Figure 3 which shows the \( a^2 \) dependence of the reduced moments. The smallest two lattice spacings are sufficiently close to \( a = 0 \) 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:

- **Vary perturbation theory:** We chose \( \mu = 3m_b \) in order to keep scales large and \( \alpha_{\overline{\text{MS}}}^p(\mu) \) small. Our results are quite insensitive to \( \mu \), however. Choosing \( \mu = m_b \), for example, shifts none of our results by more than 0.2\( \sigma \), and leaves all errors unchanged except for \( m_c(3) \), where the error increases by a third. Taking \( \mu = 9m_b \) shifts results by less than

| Source of Uncertainty | \( \alpha_{\overline{\text{MS}}}^p(\mu, n_f = 3) \) | Total |
|-----------------------|-----------------------------|-------|
| \( \sigma \) extrapolation | 0.2% | 0.6% |
| perturbation theory | 0.5 | 0.1 |
| statistical errors | 0.1 | 0.3 |
| \( m_b \) extrapolation | 0.1 | 0.2 |
| errors in \( r_1 \) | 0.1 | 0.1 |
| errors in \( r_1/a \) | 0.1 | 0.2 |
| errors in \( m_{\eta_c}, m_{\eta_b} \) | 0.1 | 0.1 |
| \( \alpha_0 \) prior | 0.1 | 0.1 |
| gluon condensate | 0.0 | 0.0 |
| Total | 0.6% | 0.7% |

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.
0.4σ, and reduces the me error by a third, leaving others only slightly reduced. Adding more terms to the perturbative expansions (N_{nth}=6 \rightarrow 8) also has essentially no effect on the results. The prior for the unknown perturbative coefficients (Eq. (21)) is twice as wide as suggested by our simulation results (using the empirical Bayes criterion [17]); we choose the larger width to be conservative.

- **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 mb and much less for the other quantities. Much larger N_{ams} can be tested easily using the trick described in Section II.B.2. For example, replacing R_n by R^{latt}_n (Eq. (18)) with N_{am}=80 and N_{am}=30 gives results that are essentially identical to those above. As discussed above, taking N_{am}=0 with the same N_{am} also gives the same results and is 22 times faster (see the appendix for further discussion).

- **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.

- **Add more/fewer \Lambda/m_{th} terms in z:** Increasing the number of terms in the expansion for z from \bar{N}_z=4 to 6 changes nothing by more than 0.1σ. Decreasing to \bar{N}_z=3 also has no effect. Again the width of the prior is optimal according to the empirical Bayes criterion.

- **Include more/fewer moments:** Keeping all moments 4 ≤ n ≤ 18 changes nothing by more than 0.5σ and reduces errors slightly for everything other than mb, 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 ≤ 10 because these are the only moments for which we have exact third-order perturbation theory. Keeping just n=4,6 gives almost identical results for m_e and \alpha_{\overline{MS}} with almost the same errors, but doubles the error on mb.

- **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{MS}}), while still shifting central values by less than 0.5σ.

- **Add large masses:** Including cases with am_{th} > 1.95 from Table III leads to poor fits. The excluded data, however, do not deviate far from the best-fit lines. For example, the points marked with an “x” in Figure 1 are for the largest mass we studied, corresponding to m_{th}=9.15 GeV (last line in Table III). Although am_{th} 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 mb/mb

It is possible to extract the ratio of quark masses mb/me 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{MS} masses

\[
\frac{m_b(\mu, n_f)}{m_c(\mu, n_f)} = \frac{m_{0b}}{m_{0c}} + \mathcal{O}(\alpha_s a^2 m_0^2)
\]  

(39)

for any \mu 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 \eta_c and \eta_b, respectively. We obtain accurate mass ratios from this relationship by extrapolating to a = 0. We used such a method recently to determine me/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

\[
w(m_{th}, a) \equiv \frac{2m_{0b}}{m_{0b}}
\]  

(40)

FIG. 3: Lattice spacing dependence of R_n for masses m_{th} within 5% of m_{0c}, and moments n = 4, 6, 8, 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.
on the $\eta_h$ mass and the lattice spacing. The ratio of MS masses is then given by the experimental masses of the $\eta_c$ and $\eta_h$ and the equation:

$$\frac{m_h(\mu, n_f)}{m_c(\mu, n_f)} = \frac{m_{\exp h}(\mu_{\exp}, 0)}{m_{\exp c}(\mu_{\exp}, 0)},$$  

(41)

It might seem simpler to fit $m_{oh}$ directly, rather than the ratio $w$; but using $w$ significantly reduces the $m_{\eta_h}$ dependence (and therefore our extrapolation errors), and also makes our results quite insensitive to uncertainties in our values for the lattice spacing.

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

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

(42)

where, as for the moments,

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

(43)

Coefficients $c_{ij}$ and $w_n$ are determined by fitting function $w(m_{\eta_h}, a)$ to the values of $2am_{oh} / am_{\eta_h}$ 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 $\eta_h$ is nonrelativistic (finite-$a$ errors are suppressed by additional powers of $v/c$ [8]), and the variation with $m_{\eta_h}$ stronger (twice that of $z(3, m_{\eta_h})$). So here we use priors

$$c_{ij} = 0 \pm 0.05$$  

(44)

$$w_n = 0 \pm 4$$  

$$Z_m(a) = 1 \pm 0.5$$

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 Figure 4 which plots the ratio $m_{oh}/m_{\eta_h}$ divided by $m_{oh}/m_{\eta_h}$ over a range of $\eta_h$ masses. Our data for different lattice spacings is compared 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 $am_{\eta_h} > 1.95$). Using the $\eta_c$ and $\eta_h$ masses from Section [V.B] and Eq. (41) with the best-fit values for the parameters, we obtain finally

$$\frac{m_{oh}}{m_{0c}} \rightarrow 4.49(4) \quad \text{as } a \rightarrow 0$$  

(45)

$$\frac{m_h(\mu, n_f)}{m_c(\mu, n_f)} = \frac{m_{\exp h}(\mu_{\exp}, 0)}{m_{\exp c}(\mu_{\exp}, 0)},$$

which agrees well with our result from the moments (Eq. 39).

VI. $\alpha_{MF}$ FROM WILSON LOOPS

In a recent paper [26], 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, 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 $r_1/a$ [12] 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 $r_1/\alpha s$ are 1.813(8), 2.644(3), 5.281(8) and 5.283(8), respectively.) We have rerun our earlier analysis, updating $r_1/a$ and the $c$ and $b$ masses. The results are shown in Figure 5. Combining results as in the earlier paper we obtain a final value from the Wilson-loop quantities of

$$\alpha_{MF}(M_Z, n_f = 5) = 0.1184(6),$$  

(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, $\alpha_{MF}(M_Z) = 0.1183(8)$, 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 [26] (reduce the $r_1$ and $r_1/a$ errors in that table by half to account for the improved values used here).
FIG. 5: Updated values for the 5-flavor $\alpha_{\text{MS}}$ 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). $\chi^2$ per data point is 0.3.

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 $\alpha_{\text{MS}}$ independently of each other. Our new results,

$$m_c(3 \text{ GeV}, n_f = 4) = 0.986(6) \text{ GeV}$$

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

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 $z(3/m_b)$, from the largest $m_b$ used in our fits to $m_b$, gives us a new, accurate determination of the $b$-quark mass,

$$m_b(10 \text{ GeV}, n_f = 5) = 3.617(25) \text{ GeV}. $$
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 \( \mathbf{7} \)

\[
m_c(3\text{GeV}, n_f = 4) = 0.986(13)\text{GeV} \\
m_b(m_b, n_f = 5) = 4.163(16)\text{GeV}
\]  

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

Function \( z(\mu/m_h, m_{y_h}) \) is a by-product of our analysis. It relates the \( \overline{\text{MS}} \) quark mass \( m_{y_h}(\mu) \) to the \( y_h \) mass (Eq. (5)). We show our result again in Figure 6 for \( \mu = 3m_h \), as well as for \( \mu = m_h \) and \( \mu = m_b/2 \), 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 \( \mathbf{27} \). Like \( \mathbf{7} \) we favor using larger scales than \( m_c \) for \( c \)-quark correlators, but, as we have shown, our results are quite insensitive to \( \mu \) 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 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, Figure 6 for the moments, as a function of \( m_{y_h} \), with the plots of \( R_n/n_\alpha \) in Figure 1. Figure 7 shows very different \( m_{y_h} \) behavior, at the 10–20\% level, for different moments \( R_n \); Figure 11 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. (5)) and our nonperturbative result for the ratio of HISQ masses (Eq. (15)). These should be and are equal to within our 1\% errors. Taken together they suggest a composite result of:

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

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 (\( \pi/a \)) 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 non-lattice methods. Figure 6 summarises recent results that were included in a world average by Bethke \( \mathbf{28} \). 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 \( \mathbf{11} \). Our new \( c \) mass, which becomes 1.093(6)GeV when converted to \( n_f = 3 \) at 2GeV, implies:

\[
m_u(2\text{GeV}, n_f = 3) = 92.2(1.3)\text{MeV}, \\
m_d(2\text{GeV}, n_f = 3) = 4.77(15)\text{MeV}, \\
m_s(2\text{GeV}, n_f = 3) = 2.01(10)\text{MeV}.
\]  

Our results for all 5 quark masses are compared with the
Particle Data Group’s 2009 values in Figure 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 non-lattice 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 $a \to 0$. While early concerns about the validity of these formalisms have been largely addressed both by formal arguments and by extensive empirical studies, it remains important to test the simulation technology of lattice QCD with increasing precision, given the growing importance of lattice results for phenomenology.

## Acknowledgements

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

### Appendix: Accelerated Fitting

In Section 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 formal derivations.

#### A. 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 $x$,

$$\sum_{\nu} c_{\nu} x_{i}^\nu,$$

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

$$P(c) \propto e^{−\sum_{\nu} c_{\nu}^2/2\sigma_{\nu}^2}.$$  

For simplicity, we assume throughout that the $c_{\nu}$ 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

$$P(y|c) \propto \prod_{i=0}^{D-1} \delta(y_i − \sum_{\nu} c_{\nu} x_{i}^\nu).$$

Bayes’ theorem

$$P(c|y) \propto P(y|c)P(c)$$  

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 $c_{\nu_{<}}$ with $n < F$, we integrate over the remaining $c_{\nu_{>}}$, giving

$$P(c_{\nu_{<}}|y) \propto e^{−\sum_{\nu_{<}} c_{\nu_{<}}^2/2\sigma_{\nu_{<}}^2} \times$$

$$\left[ ∫ dc_{\nu_{>}} \delta^{D}(y − \sum_{\nu} c_{\nu} x_{i}^\nu) e^{−\sum_{\nu_{>}} c_{\nu_{>}}^2/2\sigma_{\nu_{>}}^2} \right].$$

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

$$P(c_{\nu_{<}}|y) \propto e^{−\sum_{\nu_{<}} c_{\nu_{<}}^2/2\sigma_{\nu_{<}}^2} \times$$

$$(\det \Delta_{\nu})^{-1/2} e^{−\Delta y (2\sigma_{\nu_{<}}^2)^{-1} \cdot \Delta y}.$$
Here
\[ \Delta y_i \equiv y_i - \sum_{n<} c_{n<} x_i^n \]  
(58)
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
\[ \sigma_{\Delta}^2_{ij} \equiv \sum_{n>} x_i^n \sigma_n^2 x_j^n. \]  
(59)

The correlation matrix \( \sigma_{\Delta}^2 \) is independent of \( c_< \) (so the determinant is constant), and is the same as one would compute directly by
\[ \langle \Delta y_i \Delta y_j \rangle_{c>} = \left( \sum_{m>} \langle x_{m>} \sum_{n>} x_{n>} \right)_{c>} \]  
(60)
using
\[ \langle c_{m>} c_{n>} \rangle_{c>} = \sigma_n^2 \delta_{mn}. \]  
(61)

Finally, we fit \( c_< \) by minimizing \( \chi^2 \), 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 \( \sigma_{\Delta}^2 \) 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, \( \sigma_{\Delta} \) 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 \( \sigma_{\Delta}^2 \), by moving the \( c_> \) terms to the left side of Eq. (52). Using the truncated series as a model for these random data, straightforward application of Bayes’ theorem again implies the distribution in Eq. (57).

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 one, 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 \( c_> \). 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 \( \sigma_{\Delta}^2 \) 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.

**B. 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
\[ P(y|c) \propto \prod_{i=0}^{D-1} \delta(y_i - \sum_n g_i(c_n)) \]  
(62)

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
\[ g_i(c_<) \equiv g_i(c_<, c_> = 0) \]  
(63)
and
\[ \Delta y_i \equiv y_i - g_i(c_<), \]  
(64)
and integrating over \( c_> \) in this Gaussian approximation gives as before
\[ P(c_<|y) \propto e^{-\sum_n c_n^2/2\sigma_n^2} \times \]  
(65)
\[ (\det \sigma_{\Delta}^2(c_<))^{-1/2} e^{-\Delta y(2\sigma_{\Delta}^2(c_<))^{-1} \Delta y} \]
but with
\[ \sigma_{\Delta}^2_{ij} \equiv \sum_{n>} \partial_n g_i(c_<) \sigma_n^2 \partial_n g_j(c_<). \]  
(66)
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 \( \sigma_{\Delta}^2 \) depends on \( c_< \), the determinant in front is not constant, and the dependence of \( \Delta y \) 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.

**C. 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 \( \sigma^2_y \), then for sufficiently large samples we expect a Gaussian distribution
\[ P(\langle y \rangle|c) \propto e^{-\langle (y-g(c))(2\sigma^2_y)^{-1}(y-g(c)) \rangle} \]  
(67)
rather than the delta function above. Combining with the prior \( P(c) \) gives \( P(c|y) \).
Expanding \( g_i(c_<,c_>) \) to first order around \( c_> = 0 \), defining
\[
\Delta y_i \equiv \langle y_i \rangle - g_i(c_<),
\]
and integrating \( P(c|\langle y \rangle) \) over \( c_> \) gives
\[
P(c|\langle y \rangle) \propto e^{-\sum_n c_n^2/2\sigma_n^2} \times \left( \det \sigma_\Delta^2(c_<) \right)^{-1/2} e^{-\Delta y^2(2\sigma_\Delta^2(c_<))^{-1} \cdot \Delta y}.
\]
The resulting correlation matrix is a combination of true statistical and neglected series contributions, with
\[
\sigma_\Delta^2(c_<) \equiv \sigma_y^2 + \sigma_\Delta^2(c_<),
\]
as one would obtain by including both sources of uncertainty in computing \( \langle \Delta y, \Delta y \rangle \) directly. With no statistical fluctuations in \( y, \sigma_y^2 = 0 \), and it reduces to the previous result. When \( \sigma_y^2 \) is nonzero but small, \( \sigma_\Delta^2 \) still makes an important contribution.

D. 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 \((am_{m_4}/2)^2\) fit. As described in Section 11.1.3 we keep corrections through order \( N_{am} = 80 \) but moved all but \( \tilde{N}_{am} \ll N_{am} \) out of the fit function and into the errors for the reduced moments. If we set \( \tilde{N}_{am} = 3 \), for example, our fit to the \( R_6 \) simulation data changes from Figure 7 to Figure 10. The small \( N_{am} \) means that each point in Figure 10 has much larger error bars, coming from \((am_{m_4}/2)^2\) terms moved into the \( R_6 \)s. The final fit results, however, are almost identical in both cases (to within less than 0.1\%) with the same errors. Note that the \( R_6 \) errors in Figure 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 \( N_{am} \).

![Graph](image-url)  
**FIG. 10:** Same as Figure 7 but with \( N_{am} = 80 \) and \( \tilde{N}_{am} = 3 \), instead of \( N_{am} = \tilde{N}_{am} = 30 \). The error bars are almost entirely due to systematic errors caused by \( am_{m_4}/2 \) corrections omitted from the fit function.

[1] I. Allison et al [HPQCD Collaboration], K. G. Chetyrkin, J. H. Kühn, M. Steinhauser and C. Sturm, Phys. Rev. D78:054513 (2008) [arXiv:0805.2999].
[2] K. G. Chetyrkin, J. H. Kuhn and C. Sturm, Eur. Phys. J. C 48, 107 (2006) [arXiv:hep-ph/0604234].
[3] R. Boughezal, M. Czakon and T. Schultze,meier, Phys. Rev. D 74, 074006 (2006) [arXiv:hep-ph/0605023].
[4] A. Maier, P. Maierhofer and P. Marquard, Phys. Lett. B 669, 88 (2008) [arXiv:0805.3405 [hep-ph]].
[5] A. Maier, P. Maierhofer, P. Marquard and A. V. Smirnov, Nucl. Phys. B 824, 1 (2010) [arXiv:0907.2117 [hep-ph]].
[6] Y. Kiyo, A. Maier, P. Maierhofer and P. Marquard, Nucl. Phys. B 823, 269 (2009) [arXiv:0907.2120 [hep-ph]].
[7] K. G. Chetyrkin, J. H. Kuhn, A. Maier, P. Maierhofer, P. Marquard, M. Steinhauser and C. Sturm, Phys. Rev. D 80, 074010 (2009) [arXiv:0907.2110 [hep-ph]].
[8] E. Follana et al [HPQCD Collaboration], Phys. Rev. D 75, 054502 (2007) [arXiv:hep-lat/0610092].
[9] E. Follana, C. T. H. Davies, G. P. Lepage and J. Shigemitsu [HPQCD Collaboration], Phys. Rev. Lett. 100, 062002 (2008) [arXiv:0706.1720 [hep-lat]].
[10] C. T. H. Davies, E. Follana, I. D. Kendall, G. P. Lepage and C. McNeile [HPQCD Collaboration], Phys. Rev. D 81, 034506 (2010) [arXiv:0910.1229 [hep-lat]].
[11] C. T. H. Davies et al [HPQCD Collaboration], Phys. Rev. Lett. 104, 132003 (2010) [arXiv:0910.3102 [hep-ph]].
[12] E. B. Gregory et al [HPQCD Collaboration], Phys. Rev. Lett. 104, 022001 (2010) [arXiv:0909.4462 [hep-lat]].
[13] A. Bazavov et al, arXiv:0903.3598.
[14] A. Gray, I. Allison, C. T. H. Davies, E. Dalgic, G. P. Lepage, J. Shigemitsu and M. Wingate [HPQCD Collaboration], Phys. Rev. D 72, 094507 (2005) [arXiv:hep-lat/0507013].
[15] E. Gamiz, C. T. H. Davies, G. P. Lepage, J. Shigemitsu and M. Wingate [HPQCD Collaboration], Phys. Rev. D 80, 014503 (2009) [arXiv:0902.1815 [hep-lat]].
[16] We have simplified the definition of \( R_6 \) 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 \( O((am_{m_4})^2) \), which vanishes in the continuum limit.
[17] G. P. Lepage, B. Clark, C. T. H. Davies, K. Hornbostel, P. B. Mackenzie, C. Morningstar and H. Trottier, Nucl. Phys. Proc. Suppl. 106, 12 (2002) [arXiv:hep-lat/0110175].
[18] T. van Ritbergen, J. A. M. Vermaseren and S. A. Larin, Phys. Lett. B 400, 379 (1997) [arXiv:hep-ph/9701390].
[19] M. Czakon, Nucl. Phys. B 710, 485 (2005) [arXiv:hep-ph/0411261].
[20] K. G. Chetyrkin, Phys. Lett. B 404, 161 (1997).
[arXiv:hep-ph/9703278].

[21] J. A. M. Vermaseren, S. A. Larin and T. van Ritbergen, Phys. Lett. B 405, 327 (1997) [arXiv:hep-ph/9703284].

[22] D. J. Broadhurst, P. A. Baikov, V. A. Ilyin, J. Fleischer, O. V. Tarasov and V. A. Smirnov, Phys. Lett. B 329, 103 (1994) [arXiv:hep-ph/9403274].

[23] For a recent analysis of condensate values see B. L. Ioffe, Prog. Part. Nucl. Phys. 56, 232 (2006) [arXiv:hep-ph/0502148].

[24] We use the Particle Data Group value for $m_{\eta_c}$: C. Am-
sler et al. (Particle Data Group), PL B 667, 1 (2008)
and 2009 partial update for the 2010 edition (URL: http://pdg.lbl.gov). We correct the $\eta_c$ mass for quark
annihilation (2.4(1.2) MeV [8]), and electromagnetic cor-
rections (2.6(1.2) MeV), which are dominated by the
Coulombic interaction between the heavy quark and an-
tiquark. Neither effect is in our simulation. See [10] for
more details.

[25] We average over recent experiments to obtain a value
of 9.931(3) GeV for $m_{\eta_b}$, and then include corrections
for quark annihilation and electromagnetic corrections
(see [24]) 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)
[Erratum-ibid. 102, 029901 (2009)] [arXiv:0807.1086
[hep-ex]], G. Bonvicini et al. [CLEO Collaboration],
Phys. Rev. D 81, 031104 (2010) [arXiv:0909.5474
[hep-ex]], B. Aubert et al. [BABAR Collaboration], Phys. Rev.
Lett. 103, 161801 (2009) [arXiv:0903.1124 [hep-ex]].

[26] C. T. H. Davies, K. Hornbostel, I. D. Kendall, G. P. Lepage, C. McNeile, J. Shigemitsu and H. Trot-
tier [HPQCD Collaboration], Phys. Rev. D 78, 114507
(2008) [arXiv:0807.1687 [hep-lat]].

[27] For a related discussion see: J. H. Kuhn, arXiv:1001.5173
[hep-ph].