

**Hindered M1 radiative decay of  $\Upsilon(2S)$  from lattice NRQCD**

Ciaran Hughes,<sup>1,\*</sup> Rachel J. Dowdall,<sup>1</sup> Christine T. H. Davies,<sup>2</sup> Ronald R. Horgan,<sup>1</sup>  
Georg von Hippel,<sup>3</sup> and Matthew Wingate<sup>1</sup>  
(HPQCD Collaboration)<sup>†</sup>

<sup>1</sup>*Department of Applied Mathematics and Theoretical Physics, University of Cambridge,  
Cambridge CB3 0WA, United Kingdom*

<sup>2</sup>*SUPA, School of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, United Kingdom*

<sup>3</sup>*Institut für Kernphysik, University of Mainz, Becherweg 45, 55099 Mainz, Germany*

(Received 12 August 2015; published 3 November 2015)

We present a calculation of the hindered M1  $\Upsilon(2S) \rightarrow \eta_b(1S)\gamma$  decay rate using lattice nonrelativistic quantum chromodynamics. The calculation includes spin-dependent relativistic corrections to the NRQCD action through  $\mathcal{O}(v^6)$  in the quark's relative velocity, relativistic corrections to the leading order current which mediates the transition through the quark's magnetic moment, radiative corrections to the leading spin-magnetic coupling and for the first time a full error budget. We also use gluon field ensembles at multiple lattice spacing values, all of which include  $u$ ,  $d$ ,  $s$  and  $c$  quark vacuum polarization. Our result for the branching fraction is  $\mathcal{B}(\Upsilon(2S) \rightarrow \eta_b(1S)\gamma) = 5.4(1.8) \times 10^{-4}$ , which agrees with the current experimental value.

DOI: [10.1103/PhysRevD.92.094501](https://doi.org/10.1103/PhysRevD.92.094501)

PACS numbers: 12.38.Gc, 13.20.Gd, 13.40.Hq, 14.40.Pq

**I. INTRODUCTION**

Quantum chromodynamics (QCD) has been accepted as the theory describing the strong force of nature ever since the discovery of the  $J/\psi$ . Since then, there has been a long history of using the spectrum and decays of heavy quarkonia in order to understand QCD, heavy quarkonia being the ideal theoretical testing grounds when using potential models, and more recently, lattice QCD. Heavy quarkonium states below threshold are very narrow, and electromagnetic transition rates are therefore significant. Comparing the theoretical and experimental rates for these decays then provides a very clear test of our understanding of the internal structure of heavy quarkonia.

A certain class of electromagnetic transitions between quarkonium states, known as hindered M1 transitions, requires a spin-flip between different radial excitations and is particularly sensitive to small relativistic effects [1] which can illuminate the dynamics of the initial and final state systems. These hindered M1 transitions still remain a challenge from both the experimental and theoretical perspective. Within the bottomonium sector, such decays include the  $\Upsilon(2S) \rightarrow \eta_b(1S)\gamma$  radiative transition, where *BABAR* measured  $\mathcal{B}(\Upsilon(2S) \rightarrow \eta_b(1S)\gamma) = 3.9(1.5) \times 10^{-4}$  [2] in 2009.

On the theory side, hindered M1 decays have been notoriously difficult to pin down from within a potential model framework [1], where systematic errors are hard to quantify and branching fractions ranging from  $0.05 \times 10^{-4}$

to  $15 \times 10^{-4}$  are found. The reasons for the difficulty in accurately predicting these decays from within a potential model will be discussed in Sec. VI. The continuum effective field theory approach called potential non-relativistic QCD (pNRQCD) has been used to predict radiative bottomonium decays, including M1 transitions. While these calculations have become quite precise for the allowed  $1S \rightarrow 1S$  M1 transitions, the results for hindered M1 transitions are dominated by theoretical uncertainties and presently can only give an order-of-magnitude estimate [3,4].

Lattice NRQCD is a first principles tool that has been systematically improved by the HPQCD collaboration and can aid in reliably pinning down this difficult to predict decay. Using this formalism, one can accurately overcome each of the issues arising from within a potential model framework. Previous exploratory work on this decay in a lattice NRQCD framework was done in [5,6]. We make a number of improvements to those studies so that an accurate calculation can be done, complete with a full error budget. Some of these improvements include using one-loop radiative corrections in the NRQCD action and we show in Sec. V that these decays are very sensitive to a subset of these radiative corrections.

This paper is organized as follows. In Sec. II we set up notation and formulas relevant to this decay, and in Sec. III we give details of the computational setup including a discussion of states in NRQCD at nonzero momentum. In Sec. IV the different currents mediating this transition in NRQCD are shown and the perturbative calculation of the matching coefficient from the leading order current to full QCD is performed. Finally, analysis of

\*ch558@cam.ac.uk

†<http://www.physics.gla.ac.uk/HPQCD>

the  $\Upsilon(2S) \rightarrow \eta_b(1S)\gamma$  decay rate with a full error budget is given in Sec. V. We conclude with a discussion in Sec. VI.

## II. DECAY RATES FOR RADIATIVE TRANSITIONS

*BABAR* has measured the branching fraction of the  $\Upsilon(2S) \rightarrow \eta_b(1S)\gamma$  decay as  $3.9(1.5) \times 10^{-4}$  [2], which when combined with the  $\Upsilon(2S)$  total width  $31.98 \pm 2.63$  keV [7], gives the decay rate  $1.25(49) \times 10^{-2}$  keV. The large errors on the branching fraction are due to the difficulty in isolating the small  $\eta_b(1S)$  signal from other nearby photon lines ( $\chi_{bJ}(2P, 1P) \rightarrow \Upsilon(1S)\gamma$ ,  $\Upsilon(3S, 2S) \rightarrow \Upsilon(1S)\gamma$ ) and from the large background in the energy spectrum of inclusive decays [8].

We want to perform an accurate and reliable theoretical calculation to compare to this experimental result. Computation of the theoretical decay rate requires the matrix element of the appropriate operator between the  $\Upsilon(2S)$  and  $\eta_b(1S)$  states as input. In a Lorentz-invariant theory, using the fact that the matrix element transforms as a vector under parity (and parity invariance of our theory), the only possible decomposition of the matrix element is

$$\langle \eta_{b(mS)}(k) | j^\mu(0) | \Upsilon_{(nS)}(p, \epsilon(p, \lambda)) \rangle = \frac{2\mathcal{V}_{nm}^{\Upsilon\eta_b}(q^2)}{m_{\Upsilon(nS)} + m_{\eta_b(mS)}} \epsilon^{\mu\nu\rho\sigma} p_\nu k_\rho \epsilon(p, \lambda)_\sigma, \quad (1)$$

where  $q$  is the photon momentum,  $\epsilon(p, \lambda)_\sigma$  is the polarization vector of the  $\Upsilon_{(nS)}$  and  $p = k + q$  by momentum conservation. Using time reversal invariance, one can show that  $\mathcal{V}_{nm}^{\Upsilon\eta_b}(q^2)$  is real [9]. As the  $\Upsilon(2S)$  is a  $\bar{b}b$  bound state, this M1 (spin-flip) transition can occur by flipping the spin on either the quark or the antiquark. Since this is a symmetric process, the form factor resulting from coupling the current to the quark or to the antiquark is then identical. In our lattice calculation we only couple the current to the quark (c.f. Sec. IV) and actually compute  $\mathcal{V}_{nm}^{\Upsilon\eta_b}(q^2)|_{\text{lat}} = \mathcal{V}_{nm}^{\Upsilon\eta_b}(q^2)/2$ .

The decay rate can now be written as

$$\Gamma(\Upsilon(2S) \rightarrow \eta_b(1S)\gamma) = \frac{16\alpha_{\text{QED}}e_q^2}{3} \frac{|\mathbf{q}|^3}{(m_{\Upsilon(2S)} + m_{\eta_b(1S)})^2} |V_{21}^{\Upsilon\eta_b}(0)|_{\text{lat}}|^2, \quad (2)$$

where  $\alpha_{\text{QED}}$  is the fine structure constant,  $e_q$  is the quark charge in units of  $e$  (i.e.,  $-1/3$  for  $b$  quarks) and  $|\mathbf{q}| = (m_{\Upsilon(2S)}^2 - m_{\eta_b(1S)}^2)/2m_{\Upsilon(2S)}$  by energy conservation, ensuring that the photon is on shell with  $q^2 = 0$ . Thus, from the theoretical perspective, the most challenging part of calculating the decay rate from first principles is computing the single unknown dimensionless hadronic form factor  $V_{21}^{\Upsilon\eta_b}(q^2 = 0)$ , which encodes the nonperturbative effects

of QCD. This quantity can be calculated in lattice QCD, and this study will focus on the computation of  $V_{21}^{\Upsilon\eta_b}(q^2 = 0)|_{\text{lat}}$ .

Using the experimental value of the decay rate mentioned above, as well as  $|\mathbf{q}| = 609(5)$  MeV measured from experiment [2] and  $\alpha_{\text{QED}} = 1/137$ , we infer

$$V_{21}^{\Upsilon\eta_b}(q^2 = 0)|_{\text{exp}} = 0.069(14). \quad (3)$$

This form factor can be directly compared to  $V_{21}^{\Upsilon\eta_b}(q^2 = 0)|_{\text{lat}}$ . From now on, we drop the  $|_{\text{lat}}$  subscript to avoid superfluous notation.

## III. COMPUTATIONAL DETAILS

### A. Second generation $N_f = 2 + 1 + 1$ gluon ensembles

Our calculation uses gauge field configurations generated by the MILC collaboration [10]. For the gauge fields, they used the tadpole-improved Lüscher-Weisz gauge action, fully improved to  $\mathcal{O}(\alpha_s a^2)$ . This is possible as the gluon action has coefficients corrected perturbatively through  $\mathcal{O}(\alpha_s)$ , including pieces proportional to the number of quark flavors in the sea [11]. These ensembles are said to have  $2 + 1 + 1$  flavors in the sea, the up and down quarks (treated as two degenerate light quarks with mass  $m_l$ ), the strange quark, and the charm quark. The sea quarks are included using the HISQ formulation of fermions [12], fully improved to  $\mathcal{O}(\alpha_s a^2)$ , removing one-loop taste-changing processes and possessing smaller discretization errors compared to the previous staggered actions.

Five ensembles were chosen, spanning three lattice spacing and three values of  $m_l/m_s$ , so that any dependence on the lattice spacing and sea quark mass could be fit and extrapolated to the physical limit. Details are given in Table I. Due to the computational expense, most of the

TABLE I. Details of the gauge ensembles used in this study.  $\beta$  is the gauge coupling.  $a_\Upsilon$  is the lattice spacing determined from the  $\Upsilon(2S - 1S)$  splitting [13], where the error combines statistics, experiment and the dominant NRQCD systematic error.  $am_q$  are the sea quark masses,  $N_s \times N_T$  gives the spatial and temporal extent of the lattices in lattice units and  $n_{\text{cfg}}$  is the number of configurations in each ensemble. We use 16 time sources on each configuration to increase statistics. Ensemble 1 is referred to as “very coarse,” 2, 3, and 4 as “coarse,” and 5 as “fine.”

Set	$\beta$	$a_\Upsilon$ (fm)	$am_l$	$am_s$	$am_c$	$N_s \times N_T$	$n_{\text{cfg}}$
1	5.8	0.1474(15)	0.013	0.065	0.838	$16 \times 48$	1020
2	6.0	0.1219(9)	0.0102	0.0509	0.635	$24 \times 64$	1052
3	6.0	0.1195(10)	0.00507	0.0507	0.628	$32 \times 64$	1000
4	6.0	0.1189(9)	0.00184	0.0507	0.628	$48 \times 64$	1000
5	6.3	0.0884(6)	0.0074	0.037	0.440	$32 \times 96$	1008

ensembles use heavier  $m_l$  than in the real world; however one of the ensembles used in this study (set 4 in Table I) has physical  $am_l/am_s$ , enabling our calculations to be performed at the physical point and reducing uncertainties associated with unphysically heavy sea quark masses.

Successive configurations generated within each ensemble are expected to be correlated. These autocorrelations in meson correlators were studied in [13] for the ensembles in Table I. There we find that the autocorrelations for bottomonium correlators are not appreciable and that the configurations can be treated as statistically independent. The ensembles have been fixed to Coulomb gauge to allow nongauge invariant smearings to be used, helping extract precise results for the excited states in our calculation (c.f. Sec. III D).

### B. $b$ quarks using NRQCD

This study focuses purely on bottomonium processes, and information on these processes can be computed on the lattice using combinations of  $b$ -quark propagators, calculated on the gluon ensembles listed in Table I. As the  $b$  quark has a Compton wavelength of about 0.04 fm, these lattices cannot resolve relativistic  $b$ -quark formulations, owing to  $a > 0.08$  fm. However, it is well known that  $b$  quarks are very nonrelativistic inside their bound states ( $v^2 \approx 0.1$ ), and thus, using a nonrelativistic effective field theory (NRQCD) for bottomonium states is very appropriate. Within NRQCD, with expansion parameter  $v$  (the velocity of the quark inside the bound state), one writes down a tower of operators to a certain order in  $v$  allowing for a systematic inclusion of ever-decreasing relativistic corrections. This effective field theory is then discretized as lattice NRQCD [14]. There are a number of systematic improvements which need to be made in order to produce highly accurate results. These will be addressed shortly.

We use a lattice NRQCD action correct through  $\mathcal{O}(v^4)$ , with additional spin-dependent  $\mathcal{O}(v^6)$  terms<sup>1</sup> and include discretization corrections. This lattice formalism has already been used successfully to study bottomonium  $S$ ,  $P$  and  $D$  wave mass splittings [13,15], precise hyperfine splittings [16,17],  $B$  meson decay constants [18],  $\Upsilon$  and  $\Upsilon'$  leptonic widths [19] and  $B$ ,  $D$  meson mass splittings [17]. The Hamiltonian evolution equations can be written as

$$\begin{aligned} G(\mathbf{x}, t+1) &= e^{-aH} G(\mathbf{x}, t) \\ G(\mathbf{x}, t_{\text{src}}) &= \phi(\mathbf{x}) \end{aligned} \quad (4)$$

with

<sup>1</sup>The quantities relevant to this study are insensitive to the spin-independent  $\mathcal{O}(v^6)$  terms within our precision.

$$\begin{aligned} e^{-aH} &= \left(1 - \frac{a\delta H|_{t+1}}{2}\right) \left(1 - \frac{aH_0|_{t+1}}{2n}\right)^n U_t^\dagger(x) \\ &\times \left(1 - \frac{aH_0|_t}{2n}\right)^n \left(1 - \frac{a\delta H|_t}{2}\right); \quad aH_0 = -\frac{\Delta^{(2)}}{2am_b}, \\ a\delta H &= a\delta H_{v^4} + a\delta H_{v^6}; \end{aligned} \quad (5)$$

$$\begin{aligned} a\delta H_{v^4} &= -c_1 \frac{(\Delta^{(2)})^2}{8(am_b)^3} + c_2 \frac{i}{8(am_b)^2} (\nabla \cdot \tilde{\mathbf{E}} - \tilde{\mathbf{E}} \cdot \nabla) \\ &- c_3 \frac{1}{8(am_b)^2} \boldsymbol{\sigma} \cdot (\tilde{\nabla} \times \tilde{\mathbf{E}} - \tilde{\mathbf{E}} \times \tilde{\nabla}) \\ &- c_4 \frac{1}{2am_b} \boldsymbol{\sigma} \cdot \tilde{\mathbf{B}} + c_5 \frac{\Delta^{(4)}}{24am_b} - c_6 \frac{(\Delta^{(2)})^2}{16n(am_b)^2} \\ a\delta H_{v^6} &= -c_7 \frac{1}{8(am_b)^3} \{\Delta^{(2)}, \boldsymbol{\sigma} \cdot \tilde{\mathbf{B}}\} \\ &- c_8 \frac{3}{64(am_b)^4} \{\Delta^{(2)}, \boldsymbol{\sigma} \cdot (\tilde{\nabla} \times \tilde{\mathbf{E}} - \tilde{\mathbf{E}} \times \tilde{\nabla})\} \\ &- c_9 \frac{i}{8(am_b)^3} \boldsymbol{\sigma} \cdot \tilde{\mathbf{E}} \times \tilde{\mathbf{E}}. \end{aligned} \quad (6)$$

The parameter  $n$  is used to prevent instabilities at large momentum due to the kinetic energy operator. A value of  $n = 4$  is chosen for all  $am_b$  values. A smearing function  $\phi(\mathbf{x})$  is used to improve projection onto a particular state in the lattice data. Using an array of smearing functions to improve the overlap with the ground state and the first excited state will prove crucial to obtaining accurate results for the  $\Upsilon(2S) \rightarrow \eta_b(1S)\gamma$  decay. To evaluate the propagator, we use random wall sources that are implemented stochastically with  $U(1)$  white noise, significantly improving the precision of the S-wave states [13].

Here,  $am_b$  is the bare  $b$  quark mass,  $\nabla$  is the symmetric lattice derivative, with  $\tilde{\nabla}$  being the improved version, and  $\Delta^{(2)}$ ,  $\Delta^{(4)}$  are the lattice discretizations of  $\Sigma_i D_i^2$ ,  $\Sigma_i D_i^4$  respectively.  $\tilde{\mathbf{E}}$ ,  $\tilde{\mathbf{B}}$  are the improved chromoelectric and chromomagnetic fields, details of which can be found in [13]. Each of these fields, as well as the covariant derivatives, must be tadpole improved using the same improvement procedure as in the perturbative calculation of the matching coefficients [13,20] (thus removing unphysical tadpole diagrams from using the Lie group element rather than the Lie algebra element in the construction of the lattice field theory). We take the mean trace of the gluon field in Landau gauge,  $u_{0L} = \langle \frac{1}{3} \text{Tr} U_\mu(x) \rangle$ , as the tadpole parameter, calculated in [13,18].

The matching coefficients  $c_i$  in the above Hamiltonian take into account the high-energy UV modes from QCD processes that are not present in NRQCD. Each  $c_i$  can be expanded perturbatively as  $c_i = 1 + c_i^{(1)}\alpha_s + \mathcal{O}(\alpha_s^2)$  and, after tadpole improvement, we expect the radiative corrections  $c_i^{(1)}$  to be  $\mathcal{O}(1)$ . Each  $c_i^{(1)}$  can then be fixed by

matching a particular lattice NRQCD formalism<sup>2</sup> to full continuum QCD. These corrections have previously been computed [13,20]. Alternatively, particular  $c_i$ s can be tuned nonperturbatively, which we discuss in Sec. VB 9.

A high-precision calculation with a reliable error budget will require knowledge of at least the  $\mathcal{O}(\alpha_s)$  corrections to the matching coefficients. For example, when tuning the quark mass  $am_b$  fully nonperturbatively in NRQCD, one computes the kinetic mass of a hadron<sup>3</sup> [13]. This kinetic mass depends on the internal kinematics of the hadron, and hence on the terms  $c_1$ ,  $c_5$ , and  $c_6$  in the Hamiltonian. Using the one-loop corrected coefficients to these terms has a small but visible effect on the kinetic masses and hence on the value of the tuned  $am_b$  [13].

In addition to this, for an  $\mathcal{O}(v^4)$  NRQCD action with  $c_4 = 1$ , the kinetic mass for the  $\eta_b$  is actually found to be larger than that of the  $\Upsilon$  [13], opposite to what is seen at zero momentum and, more importantly, in experiment. The explanation is that the  $\sigma \cdot B$  term gives rise to the hyperfine splitting, and the splitting from this term is correctly included in the static mass (the mass at zero energy, offset due to removing the mass term from the Lagrangian). However, relativistic corrections to  $\sigma \cdot B$  (the term proportional to  $c_7$  in the Hamiltonian above) are needed to correctly feed this splitting into the kinetic mass. On a fine lattice, a value of  $c_4 = 1.18$  and  $c_7 = 1.25$  was needed to yield a hyperfine splitting using kinetic masses which agreed with experiment within errors [16]. In order to remove the sensitivity to the  $\sigma \cdot B$  term when tuning  $am_b$ , one does not use the kinetic mass of a single state, but the spin-averaged kinetic mass of the  $\Upsilon$  and  $\eta_b$  [13,21]. Including  $a\delta H_{v^6}$  terms in the evolution equations makes the  $\eta_b$  kinetic mass lower than that of the  $\Upsilon$ , as they include relativistic corrections to the  $\sigma \cdot B$  term. The spin-averaged kinetic mass gets smaller and the bare quark mass gets larger [16].

The parameters used in this study are summarized in Table II. There,  $c_1$ ,  $c_5$  and  $c_6$  are the correct values for a  $v^4$  NRQCD action [13], but the small changes to these coefficients in going to a  $v^6$  NRQCD action have a negligible effect on the quantities studied here, as shown in Fig. 7. While the  $am_b$  values from ensembles 1, 2 and 5 listed in Table II have all been tuned against the spin-averaged kinetic mass using the Hamiltonian above [16], the  $am_b$  values from ensembles 3 and 4 were previously tuned without the  $a\delta H_{v^6}$  terms [18]. Ensembles 2, 3 and 4 are all coarse lattices and only differ by having different light quark masses in the sea. Ensemble 2 has a correctly

TABLE II. Parameters used for the valence quarks.  $am_b$  is the bare  $b$ -quark mass in lattice units,  $u_{0L}$  is the tadpole parameter. The  $c_i$  are coefficients of terms in the NRQCD Hamiltonian [see Eq. (6)]. Details of their calculation can be found in [13,20].  $c_3$ ,  $c_7$ ,  $c_8$  and  $c_9$  are included at tree level. We also list the values of  $\alpha_s$  used to determine the one-loop corrections in the perturbative matching in Sec. IVA and for the error budget in Sec. VD.

Set	$am_b$	$u_{0L}$	$c_1, c_6$	$c_2$	$c_4$	$c_5$	$\alpha_s(\pi/a)$
1	3.31	0.8195	1.36	1.29	1.23	1.21	0.275
2	2.73	0.8346	1.31	1.02	1.19	1.16	0.255
3	2.68	0.8349	1.31	1.02	1.19	1.16	0.255
4	2.66	0.8350	1.31	1.02	1.19	1.16	0.255
5	1.95	0.8525	1.21	0.68	1.18	1.12	0.225

tuned  $am_b = 2.73$  for the Hamiltonian we use, corresponding to  $m_b = 4.418$  GeV. It is appropriate to tune the  $am_b$  values on the other coarse lattices to match this physical value. Using the lattice spacings listed in Table I, we find the  $am_b$  values on ensemble 3 and 4 listed in Table II. All these ensembles have essentially the same value of the lattice spacing, so the running of the bare mass is a negligible effect. This was observed with a  $\mathcal{O}(v^4)$  Hamiltonian [13].

Within NRQCD, the Dirac field  $\Psi$  can be written in terms of the quark  $\psi$  and antiquark  $\chi$  as  $\Psi = (\psi, \chi)^T$ . The propagator is then found to be

$$S(x|y) = \begin{pmatrix} G_\psi(x|y) & 0 \\ 0 & -G_\chi(x|y) \end{pmatrix},$$

where  $G_\psi(x|y)$  is the two-spinor component quark propagator and  $G_\chi(x|y)$  is the two-spinor component antiquark propagator.  $\gamma^5$  hermiticity becomes  $G_\psi(x|y) = -G_\chi^\dagger(y|x)$ . As such, we write our interpolating operators as in Table III and then use the above decomposition, with suitable boundary conditions, to write the correlator in terms of  $G_\psi(x|y)$ .

TABLE III. The local bilinear operators used in this study. Note the  $i\gamma^5$  is needed to make the overlaps real [9]. The second column gives the  $J^{PC}$  states that these operators create at rest in an infinite-volume continuum. The third column gives the helicity eigenvalues  $\lambda$  that these operators create at nonzero momentum in an infinite-volume continuum which is only rotationally invariant, while the  $J$  in brackets are the states which contribute to that helicity (c.f. Sec. III E).

$\mathcal{O}^\Gamma(x)$	$J^{PC}$	$\lambda(\leftarrow J^P)$
$\bar{\psi}i\gamma^5\psi$	$0^{-+}$	$0^-(\leftarrow J^P = 0^-, 1^+, 2^-, \dots)$
$\bar{\psi}\gamma^i\psi$	$1^{--}$	$0^+(\leftarrow J^P = 0^+, 1^-, 2^+, \dots)$ $ 1 (\leftarrow J = 1, 2, 3, \dots)$

<sup>2</sup>Changing the NRQCD action can modify the Feynman rules used in the computation of  $c_i^{(1)}$  in perturbation theory, in general changing its value.

<sup>3</sup>The static mass (the energy corresponding to zero-spatial momentum) in lattice NRQCD [13] is shifted due to the removal of the mass term from the Hamiltonian and so one can only tune static mass differences fully nonperturbatively.

### C. Noninteger momentum on the lattice

Using periodic boundary conditions (PBC) for the quark fields forces the momentum components to be  $p_i = 2\pi n_i/L$ , where  $n_i$  is an integer. The issue with this is that processes which occur at a specific momentum, such as those needed for an on-shell photon in the form factor  $V_{21}^{\eta_b}(q^2 = 0)$ , cannot be reached at an integer-valued momentum. Here, we use “twisted boundary conditions” ( $\theta$ BC) [22,23] in order to find the matrix element at the physical  $q^2 = 0$  point. There are some subtleties with using  $\theta$ BC in our calculation that, to our knowledge, are not found in the literature, and we give an explicit example of the construction of our twisted correlators in Appendix B. As seen there, and confirmed by numerical data, the twisted and untwisted correlator data should agree (if the same momentum is used) on a configuration level if everything is done correctly.

In our calculations, we choose  $\mathbf{p}_i = \mathbf{p}_f = \mathbf{q} = \mathbf{0}$  and only twist a single propagator so that  $\mathbf{p}_f^\theta = -\mathbf{q}^\theta = \boldsymbol{\theta}$ . The choice of isotropic twist momentum  $\boldsymbol{\theta} = \chi_0(1, 1, 1) \times 2\pi/L$  that gives  $q^2 = 0$  depends on the specific process under study and for the  $\Upsilon(2S) \rightarrow \eta_b(1S)\gamma$  decay  $\chi_0$  is found from (2) as

$$\chi_0 = \frac{L}{2\sqrt{3}\pi} \frac{m_{\Upsilon(2S)}^2 - m_{\eta_b(1S)}^2}{2m_{\Upsilon(2S)}} \quad (7)$$

yielding  $|\mathbf{q}^\theta|^2 = |\boldsymbol{\theta}|^2$ . We choose an isotropic momentum as it has been shown to reduce discretization errors from rotational symmetry breaking [13]. Since static masses obtained from correlators at rest are shifted by an arbitrary value in NRQCD, tuning  $\chi_0$  from lattice data would require a more lengthy computation of the kinetic masses. Instead, we use the experimental values of these masses [7] to tune  $\chi_0$  and check that  $q^2 = 0$  from the results.

### D. Energies and amplitudes from lattice QCD

Extracting matrix elements on the lattice requires knowledge of the lattice amplitudes and energies corresponding to the states being studied. The lattice quantity which most naturally encodes information on these is the two-point correlator

$$C_{2\text{pt}}(n_{\text{src}}, n_{\text{snk}}, \mathbf{p}^\theta; t) = \sum_{\mathbf{x}} e^{-i\mathbf{x}\cdot\mathbf{p}^\theta} \langle \mathcal{O}(n_{\text{snk}}; \mathbf{x}, t + t_0) \mathcal{O}^\dagger(n_{\text{src}}; \mathbf{0}, t_0) \rangle. \quad (8)$$

Here,  $t_0$  is the source time,  $n_{\text{src}}, n_{\text{snk}}$  are the smearing type (discussed below) and  $\mathbf{p}^\theta$  is the twisted momentum. After performing the Wick contractions with the bilinear operators listed in Table III, the connected<sup>4</sup> correlator has the form

<sup>4</sup>Disconnected diagrams for heavy quarkonia are expected to be negligible as they are suppressed by the heavy quark mass [9].

$$C_{2\text{pt}}(n_{\text{src}}, n_{\text{snk}}, \mathbf{p}^\theta; t) = \sum_{\mathbf{x}} e^{-i\mathbf{x}\cdot\mathbf{p}^\theta} \text{Tr}[\Gamma_{\text{src}} S(0|\mathbf{x}; n_{\text{src}}; n_{\text{snk}}) \Gamma_{\text{snk}} \tilde{S}^\theta(\mathbf{x}|0)],$$

where  $\tilde{S}^\theta$  is the twisted propagator (c.f. Appendix B). We use smearing functions  $\phi^{\text{src}}(r)$ ,  $\phi^{\text{snk}}(r)$  on the antiquark field at the source and sink respectively. We employ hydrogenlike wave functions which have been successful in previous studies of  $b$ -physics:  $\phi(r) = \delta_{r,0}$ ,  $\exp(-r/r_0)$ ,  $(2r_0 - r) \exp(-r/2r_0)$ .  $r_0$  is the smearing radius, and we point the reader to [13] for further details on the smearings.<sup>5</sup> The different smearing combinations used in this study give a  $3 \times 3$  matrix of correlators. We do not smear the quark fields due to complications on using twisted-smear fields as outlined in Appendix B.

The two-point correlator in (8) can be spectrally decomposed as

$$C_{2\text{pt}}(n_{\text{src}}, n_{\text{snk}}, \mathbf{p}^\theta; t) = \sum_{k=1}^{n_{\text{exp}}} a(n_{\text{snk}}, k) a^*(n_{\text{src}}, k) e^{-E_k t}, \quad (9)$$

where  $E_k$  is the  $(k-1)^{\text{th}}$  energy excitation of the interpolating operator  $\mathcal{O}(x)$  used in the construction of the correlator and  $a(n_{\text{src}}/n_{\text{snk}}, k)$  are the corresponding amplitudes, labeled by the smearing used at the source or sink. We are only interested in the first few excited states, so we do not need to worry about multiparticle states or the open  $b$ -threshold. Our two-point correlators are propagated for a maximum of  $t/a = 15$  time slices, as after this the locally smeared correlator on a fine lattice is largely saturated by the ground state. In addition, correlators were calculated with 16 different time sources on each configuration in order to increase statistics. To avoid complications due to correlations between these time sources, correlators were then averaged over all sources on the same configuration.

We fit the  $3 \times 3$  matrix of correlators from  $t/a = 1 - 15$  using a simultaneous multiexponential Bayesian fit [24,25] to the spectral decomposition in (9). Different smearings give rise to different amplitudes and so we take priors on them to be 0.1(1.5). The priors on the ground state energies are estimated from previous results and given a suitably wide width [13]. For the zero-momentum case, prior information tells us that the energy splittings  $E_{n+1} - E_n$  are of the order 500(250) MeV, while for the nonzero-momentum case, priors of 480(250) MeV are used (due to the inclusion of additional states in the correlator; see Sec. III E). Logarithms of the energy splittings are taken in the fit to ensure that the ordering of states is preserved, helping the stability of the fit [25].

<sup>5</sup>We use the smearing types  $l, g, e$  as described in that reference.

### E. Energy eigenstates in lattice NRQCD

Theoretically, particle states living in the Hilbert space are classified in terms of invariant quantities within irreducible representations (irreps) of the symmetry group of a theory. For our calculation, two groups need to be considered: the Lorentz group and the continuous rotational group in three dimensions. Appendix A reviews the construction of the irreps of both these groups at zero and nonzero momentum.

As is well known, the irreps of the Lorentz group at rest are described by  $|p^2 = m^2; J^{PC}, M\rangle$ , where  $J$ ,  $M$  are the total and third component of angular momentum respectively.  $P$  is the parity quantum number and for quarkonia  $C$  is the charge conjugation. The quantum numbers  $J^{PC}$  classify all particles seen in experiments to date [7].

However, the symmetry group of NRQCD is only the rotational group. At zero momentum, the states within such a theory are also described by  $|\mathbf{p} = \mathbf{0}; J^{PC}, M\rangle$ . At nonzero momentum, the situation is significantly different, and the irreps are described by  $|\mathbf{p} \neq \mathbf{0}; \lambda\rangle$ , where  $\lambda$  is an eigenvalue of the helicity operator  $\hat{\lambda} = \hat{\mathbf{p}} \cdot \hat{\mathbf{J}}/E$ . This has important consequences for the energy spectrum extracted from our lattice calculation (compare the zero- and nonzero-momentum lattice spectrum seen in Figs. 1 and 2) and therefore needs to be fully understood in order to have a reliable computation.

At rest the bilinear operators that we use in our calculation, listed in Table III with  $\Gamma = i\gamma^5, \gamma^i$ , overlap onto definite  $J^{PC} = 0^{-+}, 1^{--}$  energy eigenstates respectively in the infinite-volume continuum version of our theory (which is rotationally invariant) [26]. In Appendix A (as in [26]) it is shown that at nonzero momentum,  $\mathcal{O}^5(\mathbf{p})$  is a helicity operator which creates a definite  $\lambda = 0^-$  energy eigenstate, but  $\mathcal{O}^i(\mathbf{p})$  creates an admixture of  $\lambda = 0^+, \pm 1$  eigenstates, where these  $\lambda$  get contributions from  $J^P$  values as listed in the third column of Table III. The  $\pm$  superscript on the  $\lambda = 0$  represents the eigenvalue  $\tilde{\eta} \equiv P(-1)^J$  from the  $\hat{\Pi}$  symmetry (a parity transformation followed by a rotation to bring the momentum direction back to the original direction) [26].

In the correlator data from using  $\mathcal{O}^5(\mathbf{p} \neq \mathbf{0})$ , guided by the experimental masses and this analysis, the lowest states in the spectrum should be  $\eta_b(1S)(= 0^{-+})$ ,  $\chi_{b1}(1P)(= 1^{++})$ ,  $\eta_b(2S)(= 0^{-+})$ , etc. whereas from using  $\mathcal{O}^i(\mathbf{p})$  the lowest states in the spectrum should be  $\Upsilon(1S)(= 1^{--})$ ,  $h_b(1P)(= 1^{+-})$ ,  $\Upsilon(2S)(= 1^{--})$ , etc. These are the  $J^P$  states which we see in our lattice spectrum at nonzero momentum.

The first three states extracted from the spectrum with the operator  $\mathcal{O}^5$ ,  $\mathcal{O}^i$  are shown in Figs. 1 and 2 respectively. On the same plot, the solid lines represent the energy of the states according to a nonrelativistic, rotational dispersion relation reconstructed using the experimental

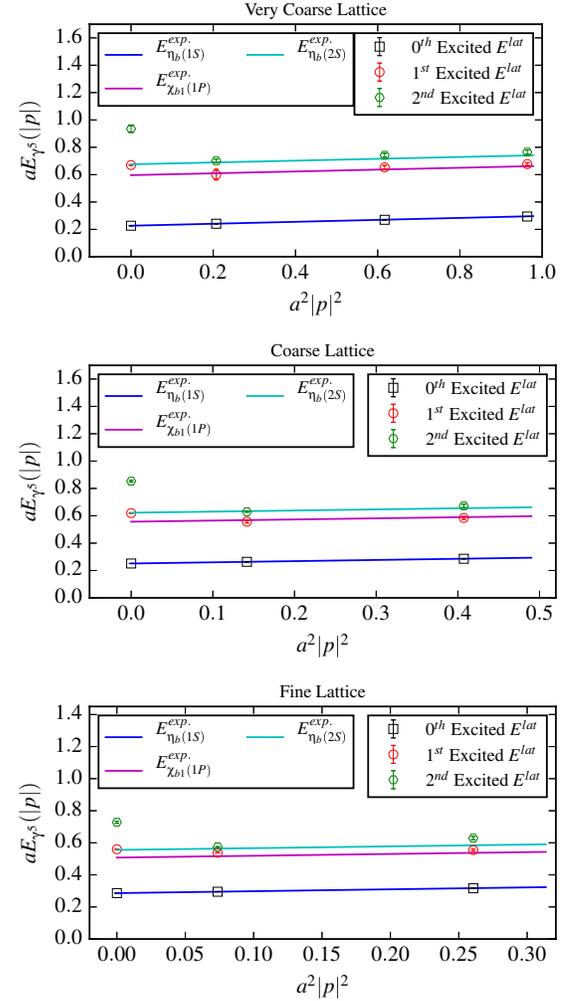


FIG. 1 (color online). The first three energies extracted from the lattice NRQCD correlator data with the operator  $\mathcal{O}^5$  across multiple momenta. Statistical errors only. At nonzero momentum, the energy of the first excited state is lower than the energy of the first excited state at zero momentum. This is a consequence of new states being present in the correlator data at nonzero momentum, as described in Sec. III E. Thus, care must be taken not to misidentify states.  $aE^{exp}$  represents the energy of the states according to a nonrelativistic, rotational dispersion relation reconstructed using the experimental masses, details of which can be found in the text.

masses, e.g.,  $aE(|\mathbf{p}|) = am^{sim} + |\mathbf{p}|^2/2am^{kin}$ , where  $m^{kin}$  is the kinetic mass which we set equal to the experimental mass, and  $m^{sim}$  is the static mass offset due to neglecting the mass term in the NRQCD Hamiltonian. We find  $am^{sim}$  in the correlator data from the  $\mathcal{O}^5$  operator by taking the ground state lattice energy at zero momentum and finding the shift in the static mass as the difference  $a\Delta = am_{\eta_b(1S)}^{exp} - am_{\eta_b(1S)}^{lat}$ . We then use this value of the shift to find  $am_{J^{PC}}^{exp,sim} = am_{J^{PC}}^{exp} - a\Delta$ , to be used in the above dispersion relation. We found the shift in the  $\mathcal{O}^i$  correlator data in the same way.

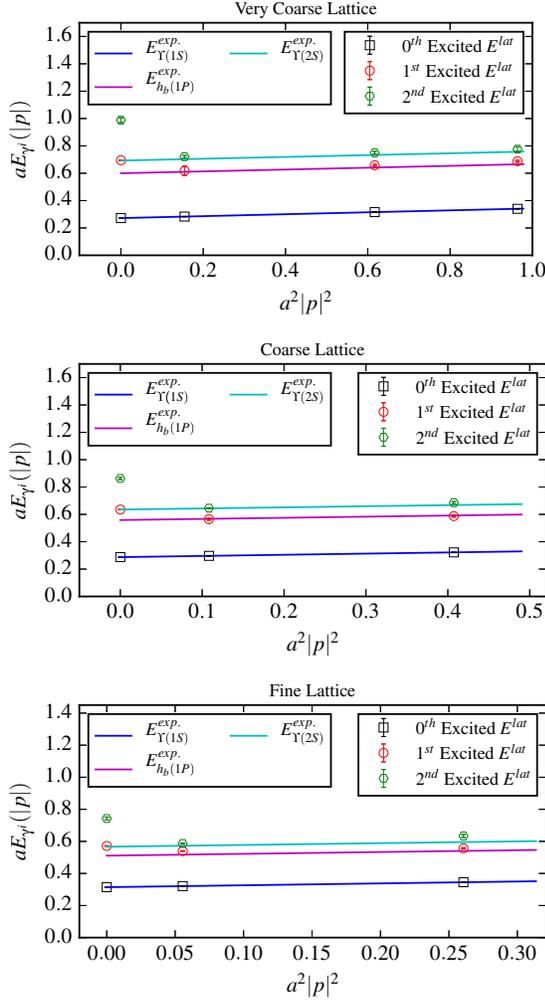


FIG. 2 (color online). As in Fig. 1 but with the operator  $\mathcal{O}^j$ .

The important point to observe in these figures is that at nonzero momentum the energy of the first excited state is actually lower than the energy of the second excited state at zero momentum, opposite to what one would expect from a dispersion relation. The reason is clear: at nonzero momentum energy eigenstates have definite helicity, not definite  $J^P$ . Therefore our correlator data get contributions from the  $J^P$  states listed in Table III.

We conclude that, as Figs. 1 and 2 show, one has to be careful in equating the states found in NRQCD at nonzero momentum with continuum  $J^{PC}$  quantum numbers and also in extracting matrix elements involving a state in flight. However, here we only extract excited states at zero momentum in order to avoid unnecessary complications and to obtain high-precision results, which can be muddled when extracting excited states in flight due to the addition of extra states in the spectrum and their small overlap factors as described in Appendix A. After our analysis, we can then be sure that we have extracted the correct matrix element for the  $\Upsilon(2S) \rightarrow \eta_b(1S)\gamma$  decay.

## F. Matrix elements from lattice QCD

The simplest quantity which encodes information on a meson-to-meson decay matrix element from within lattice QCD is the three-point correlator

$$C_{3\text{pt}}^{mn}(n_{\text{src}}, n_{\text{snk}}, \mathbf{p}_f^\theta = -\mathbf{q}^\theta; t, T) = \sum_{\mathbf{x}, \mathbf{y}} e^{-i\mathbf{x}\cdot\mathbf{p}^\theta} \langle \mathcal{O}_f(n_{\text{snk}}; \mathbf{x}, T) J^n(\mathbf{q}^\theta; \mathbf{y}, t) \mathcal{O}_i^{m\dagger}(n_{\text{src}}; \mathbf{0}, 0) \rangle, \quad (10)$$

where  $\mathcal{O}_i^m, \mathcal{O}_f$  are interpolating operators which create the initial state with polarization  $m$  and final state respectively,  $J^n(\mathbf{q}^\theta; \mathbf{y}, t) = \psi^\dagger \Gamma^n(\mathbf{q}^\theta; \mathbf{y}, t) \psi$  is the current which induces the transition with  $n$  labeling the polarization of the photon, and the twisted momenta are described in Sec. III C. The three-point correlator is visualized as in Fig. 3 where the three points in lattice units correspond to the source point of the initial particle at time  $t_0$  [equal to zero in (10)]; the position and time of the current causing the transition at  $(\mathbf{y}, t)$ ; and the position and time of the final state at  $(\mathbf{x}, T)$ . After performing Wick contractions on the three-point correlator the connected contribution, written in terms of NRQCD propagators as discussed in Sec. III B, is

$$C_{3\text{pt}}^{mn}(n_{\text{src}}, n_{\text{snk}}, \mathbf{p}_f^\theta = -\mathbf{q}^\theta; t, T) = - \sum_{\mathbf{x}, \mathbf{y}} e^{-i\mathbf{x}\cdot\mathbf{p}^\theta} \text{Tr}[\Gamma_i^m G_\chi(0|x) \Gamma_f \tilde{G}_\psi^\theta(x|y) \Gamma^n(\mathbf{q}^\theta; y) G_\psi(y|0)], \quad (11)$$

where the twisted propagator  $\tilde{G}^\theta(x|y)$  is defined in Appendix B. Direct computation of the propagator  $G(x|y)$  is unnecessarily expensive as we can use the sequential source technique (SST) [9,27] to yield the desired propagator, which only requires one further evolution. There are two ways to package the  $G(x|y)$  propagator in the three-point correlator when using the SST. The first is called the fixed current method, which requires the insertion time  $t$  to be fixed and for propagator 2 in Fig. 3 to be used as a source for propagator  $\theta$ . However, this method

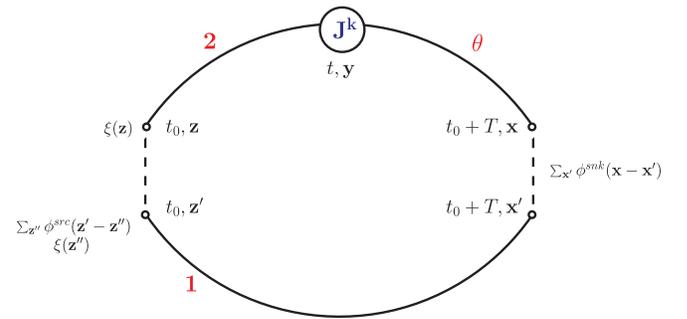


FIG. 3 (color online). Setup for the three-point correlator calculation as described in Sec. III F. Propagator 1 is the antiquark and  $\xi(x)$  is the random noise source as described in the text.

does not scale well and is undesirably expensive for relativistic quark formalisms.

The second approach is called the fixed sink method. In this approach, one fixes the sink time  $T$  and factorizes (11) as

$$C_{3\text{pt}}^{mn}(n_{\text{src}}, n_{\text{snk}}, \mathbf{p}_f^\theta = -\mathbf{q}^\theta; t, T) = -\sum_y e^{-iy\cdot\theta} \text{Tr}[\Gamma_i^m H^{\theta\dagger}(y|0) \Gamma^n(\mathbf{q}^\theta; y) G_\psi(y|0)] \quad (12)$$

with

$$H^\theta(y|0) = \sum_x e^{ix\cdot\mathbf{p}} G_\psi^\theta(y|x) \Gamma_f^\dagger G_\psi(x|0),$$

where we have written  $H(y|0)$  in terms of the twisted propagator that satisfies periodic boundary conditions and used the fact that  $\Gamma_f$  commutes with the exponential as described in Appendix B. We have also used the NRQCD  $\gamma^5$ -hermicity conditions from Sec. III B, and used  $G_\psi^\dagger(x|y) = -G_\psi(y|x)$  because  $G(x|y) = \langle \psi(x) \psi^\dagger(y) \rangle$ . We can obtain  $H^\theta(y|0)$  by using the twisted evolution equations with the source  $e^{ix\cdot\mathbf{p}} \Gamma_f^\dagger G_\psi(x|0)$ .

Clearly, the two methods should give the same correlator data as they only differ in how  $G(x|y)$  is packaged. We have checked this numerically and found it to be true on any given configuration up to machine precision. As the fixed sink method is more cost effective, this method was used for the calculation. Our program structure can be visualized in Fig. 3. Propagator 1 is generated with a smeared, random wall source at time  $t_0$  and propagated to time  $T$  where the sink smearing is applied.  $H^\theta(y|0)$  is found by using the source  $e^{ix\cdot\mathbf{p}} \Gamma_f^\dagger G_\psi^1(x|0)$  and evolving backwards in time using the twisted configurations to a time  $0 \leq t \leq T$ . Propagator 2 is made from the same random wall as 1. We then combine propagator 2,  $H^\theta(y|0)$  and the current as in (12) to obtain the three-point correlator. We use the same 16 time sources as in the two-point correlator and prior to fitting, all data are translated to a common  $t_0 = 0$ .

The three-point correlator (10) can be related to matrix elements of the current by inserting a complete set of states [9]. By doing so, and using the rotational parametrization of the overlaps as described in Appendix A,  $C_{3\text{pt}}^{mn}$  is seen to be antisymmetric. We average over the six nonzero contributions using an isotropic momentum as

$$C_{3\text{pt}}^V = \frac{1}{6} \sum_{l=1}^3 \epsilon_{lmn} C_{3\text{pt}}^{nm}. \quad (13)$$

In addition, inserting the complete set of states also leads to the functional form of the fitting function

$$C_{3\text{pt}}^V(n_{\text{src}}, n_{\text{snk}}, \theta; t, T) = \sum_{i,f} a(n_{\text{snk}}, i) V_{ij}^{\text{fit}} b^*(n_{\text{src}}, f) e^{-E_i t} e^{-E_f(T-t)}, \quad (14)$$

where  $a(n_{\text{snk}}, i)$  and  $b(n_{\text{src}}, f)$  are amplitudes from the two-point fitting function in (9). The two-point and three-point correlators can be simultaneously fit to (9) and (14) respectively using multiexponential chained [28], marginalized [29] Bayesian fitting. Chained, marginalized fitting has been shown to significantly decrease the fitting time and produce reliable, precise and accurate results if the data are in the limit of high statistics (Gaussianly distributed) [30]. We check that results are compatible from both with and without chained, marginalized fits on a subset of the data. We use a prior of 0.1(0.2) for all  $V_{ij}^{\text{fit}}$  and the same priors for the amplitudes and energies as in the two-point fits described in Sec. III D. For each current, we obtain data for fixed  $T = 9, 12, 15$  and the same  $3 \times 3$  matrix of smearings as in the two-point correlators. This allows accurate extractions of the matrix element as it includes excited state contributions.

The use of a singular value decomposition stabilizes the fit and is standard practice in the literature [28]. In our Bayesian fit, this is performed by setting a tolerance and replacing all eigenvalues of the correlation matrix smaller than this tolerance times the maximum eigenvalue to this value [28]. By doing so, this leads to larger errors in the fit results and so is a conservative step. We use a tolerance of  $10^{-4}$ .

The matrix element for the  $\Upsilon(2S) \rightarrow \eta_b(1S)\gamma$  decay will be proportional to  $V_{21}^{\Upsilon\eta_b}$ . By equating the fitting functions to their continuum correlator counterparts with conventional relativistic normalization, parametrizing our overlaps using rotational invariance with the initial particle at rest, we find

$$V_{21}^{\Upsilon\eta_b}(q^2) = \frac{m_{\Upsilon(2S)} + m_{\eta_b(1S)}}{m_{\Upsilon(2S)}\theta_i} \sqrt{m_{\Upsilon(2S)} E_{\eta_b}} V_{21}^{\text{fit}}, \quad (15)$$

where  $\theta$  is the twisted momentum described in Sec. III C. Since the static masses obtained from an NRQCD calculation are shifted, as explained previously, we extract  $V_{21}^{\Upsilon\eta_b}(q^2)$  from  $V_{21}^{\text{fit}}$  using the same experimental masses as in Sec. II. A nonrelativistic dispersion relation was used to find  $E_{\eta_b(1S)}$ , which is appropriate as shown in Fig. 1.

#### IV. M1 RADIATIVE DECAY CURRENTS

In order to compute the form factor  $V_{nm}^{\Upsilon\eta_b}(q^2)$ , we need to choose currents which will induce a hindered M1 radiative decay. Within a nonrelativistic framework, it is a standard result in the literature [31–33] that the leading order contribution to the matrix element is suppressed due to the orthogonality of the radial wave functions and relativistic corrections are necessary. This suppression introduces a sensitivity to a range of effects that we must test and quantify in order to perform an accurate calculation. The first of these effects is the fact that next-to-leading order current contributions are appreciable and we need to include them.

As we are using NRQCD to simulate the  $b$  quark, choosing the currents from a NRQCD and non-relativistic quantum electrodynamics (NRQED) effective field theory is most appropriate. This effective field theory can be found straightforwardly by extending the  $SU(3)$  Lie algebra of NRQCD to a  $SU(3) \times U(1)$  Lie algebra to produce NRQCD + NRQED [3]. Then, in principle, one could discretize the  $SU(3) \times U(1)$  theory and choose appropriate currents from the resulting operators. However, this introduces complications, e.g. the  $U(1)$  magnetic field only decouples from the  $SU(3)$  chromomagnetic field to leading order in the lattice spacing, resulting in lattice artefact currents which are not present in the continuum. Calculating such currents would require more computational resources and make the computation of the matching coefficients more difficult.

Instead, we are free to choose the currents from the continuum NRQCD + NRQED theory and renormalize these. It is important to understand the power counting in the NRQCD + NRQED effective field theory in order to choose our currents appropriately. Given that NRQCD + NRQED is a  $SU(3) \times U(1)$  effective field theory, it has two expansion parameters. For NRQCD, we have the standard expansion parameter  $v$ , where  $v^2 \sim 0.1$  for bottomonium. The only scale available for the on-shell emitted photon is the photon's energy  $|\vec{q}_\gamma| \sim 0.6$  GeV. Since the photon's energy is the difference between the masses of two heavy S-wave quarkonia, it is of the order  $|\vec{q}_\gamma| \sim mv^2 \sim 0.4$  GeV. Thus we can expand our effective field theory in terms of  $v$  only.

We summarize the power counting as

- (i)  $A_{\text{QED}} \sim |\vec{q}_\gamma|$ .
- (ii)  $B_{\text{QED}}, E_{\text{QED}} \sim |\vec{q}_\gamma|^2$ .
- (iii) The standard QCD power-counting rules for the QCD fields.
- (iv) The knowledge that when a derivative acts on the photon field, it gives a factor of  $|\vec{q}_\gamma|$  and when acting on the quark field a factor of  $p_q \sim mv$  (as the valence quark knows nothing of the photon momentum in the initial quarkonium rest frame).

Ordering the operators that induce a M1 (spin-flip) transition from NRQCD + NRQED, we find (to next-to-leading order for our decay and borrowing notation from [34])

$$\begin{aligned}
O_F &= \omega_F \frac{ee_b}{2m_b} \psi^\dagger \boldsymbol{\sigma} \cdot \mathbf{B}_{\text{QED}} \psi \\
O_{W1} &= \omega_{W1} \frac{ee_b}{8m_b^3} \psi^\dagger \{ \mathbf{D}^2, \boldsymbol{\sigma} \cdot \mathbf{B}_{\text{QED}} \} \psi \\
O_S &= \omega_S \frac{iee_b}{8m_b^2} \psi^\dagger \boldsymbol{\sigma} \cdot (\mathbf{D} \times \mathbf{E}_{\text{QED}} - \mathbf{E}_{\text{QED}} \times \mathbf{D}) \psi \\
O_{S2} &= \omega_{S2} \frac{i3ee_b}{64m_b^4} \psi^\dagger \{ \mathbf{D}^2, \boldsymbol{\sigma} \cdot (\mathbf{D} \times \mathbf{E}_{\text{QED}} - \mathbf{E}_{\text{QED}} \times \mathbf{D}) \} \psi \\
O_{\text{tot}} &= O_F + O_{W1} + O_S + O_{S2}. \tag{16}
\end{aligned}$$

Here,  $i\vec{D} = i\vec{\nabla} + g\vec{A}_{\text{QCD}}^a T^a$  are all pure QCD covariant derivatives, fields marked QED (QCD) are the QED (QCD) fields and  $\omega_i$  are the matching coefficients needed to reproduce full QCD + QED from our effective theory. Using the power-counting rules above, we find  $O_F \sim v^4$ ,  $O_{W1} \sim v^6$ ,  $O_S \sim v^5$  and  $O_{S2} \sim v^7$ . We can then factor out the photon and electric charge in order to derive the currents  $J_k(\mathbf{q}^\theta; \mathbf{y}, t)$  which give the decomposition of the matrix element in (1). For example, the operator  $O_F$  gives rise to the current

$$J_F^k = -\omega_F \frac{1}{2m_b} \psi^\dagger (\boldsymbol{\sigma} \times i\mathbf{q})^k e^{-i\mathbf{q} \cdot \mathbf{x}} \psi.$$

We then write all currents as  $J^k(\mathbf{q}^\theta; \mathbf{y}, t) = \psi^\dagger \Gamma_k(\mathbf{q}^\theta; \mathbf{y}, t) \psi$  so that  $\Gamma_k(\mathbf{q}^\theta; \mathbf{y}, t)$  will be what enters the three-point correlator as in (12). We use the terminology that the form factor coming from the current  $J_F$  is called  $V_{21}^{Y\eta_b}|_F = \omega_F \tilde{V}_{21}^{Y\eta_b}|_F$ , where the tilde implies we have factored off the matching coefficient from the form factor in the numerical calculation and this should be applied later in the analysis. Similar notation is used for the other currents and we refer to  $\tilde{V}_{21}^{Y\eta_b}|_i$  as unrenormalized form factors. The final form factor is  $V_{21}^{Y\eta_b} = \sum_i \omega_i \tilde{V}_{21}^{Y\eta_b}|_i$ .

It should be noted that there are other currents (suppressed by  $v$  or  $\alpha_s$ ) that contribute to this decay and which might be of interest, notably, the QCD analogues of the  $O_{W1}$ ,  $O_S$ , operators arising from choosing the electric (magnetic) fields in (16) to be gluon fields and the photon coming from the full  $SU(3) \times U(1)$  covariant derivative. Other operators are those which only occur at loop level in the full QCD + QED theory. These can be written as

$$\begin{aligned}
O_{W1\text{QCD}} &= -\omega_{W1\text{QCD}} \frac{iee_b}{8m_b^3} \psi^\dagger \\
&\quad \times \{ \mathbf{A}_{\text{QED}} \cdot \mathbf{D} + \mathbf{D} \cdot \mathbf{A}_{\text{QED}}, \boldsymbol{\sigma} \cdot \mathbf{gB}_{\text{QCD}} \} \psi \\
O_{\text{SQCD}} &= \omega_{\text{SQCD}} \frac{ee_b}{8m_b^2} \psi^\dagger \\
&\quad \times \boldsymbol{\sigma} \cdot (\mathbf{A}_{\text{QED}} \times \mathbf{gE}_{\text{QCD}} - \mathbf{gE}_{\text{QCD}} \times \mathbf{A}_{\text{QED}}) \psi \\
O_{W2} &= \omega_{W2} \frac{ee_b}{4m_b^3} \psi^\dagger D^i \boldsymbol{\sigma} \cdot \mathbf{B}_{\text{QED}} D^i \psi \\
O_{p'p} &= \omega_{p'p} \frac{ee_b}{8m_b^3} \psi^\dagger \boldsymbol{\sigma} \cdot \mathbf{DB}_{\text{QED}} \cdot \mathbf{D} + \mathbf{D} \cdot \mathbf{B}_{\text{QED}} \boldsymbol{\sigma} \cdot \mathbf{D} \psi. \tag{17}
\end{aligned}$$

When attempting power counting on the QCD operators above, it is helpful to draw the Feynman diagram that such an operator would produce. Essentially, we need to contract the gluon field with another, producing another factor of  $gv^3$  at least [14]. Consequently these operators are expected to be of order  $\alpha_s v^8$  at most. We confirm numerically that the form factors from these QCD operators are suppressed as

expected and they are negligible within the errors of our final results. Since  $\omega_{W2}, \omega_{p'p}$  occur only at loop level they are suppressed by  $\mathcal{O}(\alpha_s)$  relative to  $O_{W1}$ . We introduce a systematic error for neglected currents in the final analysis.

### A. Matching coefficients for the currents

The matching coefficients,  $\omega_i$ , appearing in the operators in (16) are needed to take into account the high-energy UV modes from processes in the full theory but not present in our effective field theory. They have the expansion  $1 + \omega_i^{(1)}\alpha_s + \mathcal{O}(\alpha_s^2)$ . Here we compute the one-loop correction to the coefficient  $\omega_F$  from the leading order current. Following this, we estimate the errors from neglecting corrections that we do not calculate.

Our calculation of the one-loop coefficient  $\omega_F^{(1)}$  is very similar to the computation of the one-loop correction of  $c_4$  in [20]. Following that analysis, by matching the current from NRQCD + NRQED to continuum QCD + QED, we find

$$\omega_F^{(1)} = b_{\sigma, \text{QED}}^{(1)} - Z_m^{\text{NR},(1)} - Z_m^{\text{tad},(1)} - Z_2^{\text{NR},(1)} - Z_{\sigma, \text{QED}}^{\text{NR},(1)}, \quad (18)$$

where  $b_{\sigma, \text{QED}}^{(1)} = C_F/2\pi$  is the coefficient of the first order correction to the quark's magnetic moment, computed analytically in continuum QCD following standard techniques. As  $b_{\sigma, \text{QED}}$  is UV finite, this allows us to directly equate results obtained on the lattice to those obtained in the continuum, since the difference between the schemes for UV regulation is then irrelevant. In the general matching procedure the continuum and lattice IR divergences cancel in the computation of the radiative correction; here, because of the standard Ward identity, the continuum and lattice contributions to  $\omega_F^{(1)}$  are separately finite.

$Z_m^{\text{NR}}, Z_2^{\text{NR}}, Z_{\sigma, \text{QED}}^{\text{NR}}$  are the renormalization factors of the bare quark mass, the wave function and the current  $J_F$  from (16). These are calculated in lattice NRQCD. We automatically generate the Feynman rules for a specific NRQCD action (along with the Symanzik-improved gluonic action) using the HiPPY package, then compute the numerical evaluation of these diagrams using the HPsrc package [35,36]. We use the full  $v^4$  NRQCD Hamiltonian with spin-dependent  $v^6$  pieces as defined in (6). Computation of  $Z_m^{\text{NR},(1)}$  and  $Z_2^{\text{NR},(1)}$  is identical to [20].  $Z_m^{\text{NR},(1)}$  will get contributions from mean-field corrections which we denote as  $Z_m^{\text{tad},(1)}$ . We use the Landau mean link  $u_0^{(2)} = 0.750$  [37]. For the action that we use, the tadpole correction is [20]

$$Z_m^{\text{tad}} = -\left(\frac{2}{3} + \frac{3}{(am_b)^2}\right)u_0^{(2)}. \quad (19)$$

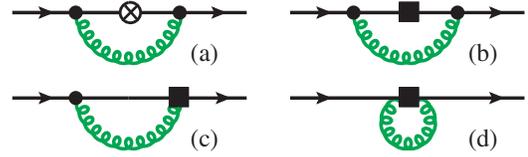


FIG. 4 (color online). Classes of one-loop diagrams which contribute to  $Z_{\sigma, \text{QED}}^{(1)}$  as described in the text. The cross inside a circle represents the  $J_F$  current obtained from (16), while the solid box represents the higher order currents from (16) and the exchange of a gluon is denoted by a curly line.

The NRQCD diagrams contributing to  $Z_{\sigma, \text{QED}}^{\text{NR},(1)}$  are shown in Fig. 4. Since we do not actually include the QED field in our calculation, there are no tadpole factors from this term. Note that Fig. 4(a) is generated by the current coming from  $\psi^\dagger \boldsymbol{\sigma} \cdot \mathbf{B}_{\text{QED}} \psi / 2am_b$  being inserted at the vertex, and Figs. 4(b)–4(d) arise from mixing effects from the higher order currents (that we include in the calculation of the decay rate) from (16). Computation of the Feynman diagrams shown in Figs. 4(b)–4(d) is more involved than that of Fig. 4(a), so they are not included in this calculation, but we plan on computing them in future work. For now, we introduce a systematic error from neglecting these contributions.

A breakdown of the numerical values of the various terms that enter  $\omega_F^{(1)}$  for the masses that we use in this calculation is shown in Table IV.  $\omega_F^{(1)}$  was computed for a range of masses (neglecting the mixing down) and we give these values in Table V.

We show the values of  $\omega_F^{(1)}$  with a smooth interpolating curve in Fig. 5. This interpolating curve was chosen to be a polynomial in  $1/am_b$  in order to reproduce the static limit as  $m_b \rightarrow \infty$ . To fit these values easily we increased the errors on the points returned by HPsrc to 1%. We use a

TABLE IV. Breakdown of the different terms that go into  $\omega_F^{(1)}$ . The  $\alpha_s(q^* = \pi/a)$  values are taken from Table II.

$am_b$	1.90	2.70	3.30
$Z_m^{(1)} + Z_2^{(1)} + Z_{\sigma, \text{QED}}^{(1)}$	1.2961(5)	0.9061(4)	0.7585(6)
$Z_m^{\text{tad}}$	-1.1233	-0.8086	-0.7066
$\omega_F^{(1)}$	0.0394(5)	0.1148(4)	0.1603(6)
$\alpha_s(\pi/a)\omega_F^{(1)}$	0.0089(1)	0.0293(1)	0.0441(2)

TABLE V. Values of  $\omega_F^{(1)}$  at various  $am_b$  values.

$am_b$	1.1	1.5	2.1
$\omega_F^{(1)}$	-0.211(2)	-0.030(1)	0.0626(9)
$am_b$	2.4	4.0	4.6
$\omega_F^{(1)}$	0.0918(7)	0.2039(4)	0.2372(5)

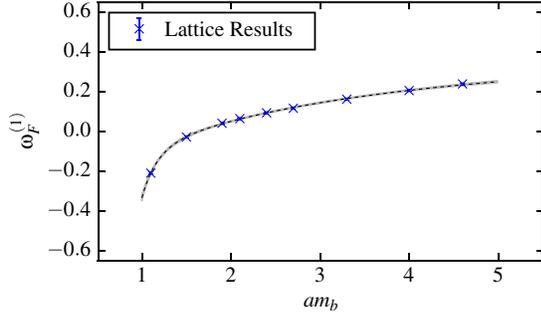


FIG. 5 (color online). The values  $\omega_F^{(1)}$  with a smooth interpolating curve as described in the text.

Bayesian fit to all points in Fig. 5 against a polynomial in  $1/am_b$ . We found the smallest  $\chi^2/\text{dof}(\text{dof}) = 0.7(9)$  and largest Gaussian Bayes factor [24] when including all terms in the polynomial up to and including the quartic term. We used a prior for the constant piece as the polynomial of  $0.4(2)$  and priors for the coefficients of the  $1/(am_b)^n$  pieces of  $0(1)$ .

### B. Systematic error from current matching coefficients

We need to include a systematic error from not knowing the matching coefficients in the currents to infinite precision. There are two distinct types of errors in this case: the first is from neglecting the  $\mathcal{O}(\alpha_s^2)$  corrections in  $\omega_F$  and the  $\mathcal{O}(\alpha_s)$  corrections to the matching coefficients of the other currents; the second is from neglecting the mixing down effects on the values of  $\omega_F^{(1)}$  used in the calculation. We estimate each of these in turn.

To estimate the effect of neglecting the higher order corrections that we have not calculated, it is helpful to compare our result to the pure NRQED calculation of [34]. There, the authors find that the continuum QED contribution to their  $\omega_F^{(1)}$  is the anomalous magnetic moment of the electron  $\alpha/2\pi$ , while for us it is the anomalous magnetic moment of the quark  $C_F\alpha_s/2\pi$ . For the NRQED contribution, they find neither an IR log nor a constant piece and in their continuum approach the UV power law divergences may be omitted. Although we find no IR log in our data, we cannot neglect the UV power law divergence associated with the momentum cutoff. This shows up as a polynomial in  $1/am_b$  as mentioned above. We observe that this lattice artefact contribution gives a negative contribution to the continuum value, as shown in Table IV, and for the  $am_b$  range that we are interested in,  $|\alpha_s\omega_F^{(1)}| < C_F\alpha_s/2\pi$ . As we are observing similar behavior over this mass range as the pure NRQED calculation, we can use that calculation to estimate the error conservatively.

As shown in [34] and confirmed by the small values of our numerical data, the matching coefficients can actually be expanded in  $\alpha_s/\pi$ . In principle, the second order

coefficient of  $\omega_F$  could be  $\mathcal{O}(1)$ , and then this contribution could be  $\mathcal{O}(\alpha_s^2/\pi^2)$ . As such, we allow for an additive systematic error (assumed to be correlated across all ensembles) of  $1 \pm \alpha_s^2/\pi^2$  from not knowing higher order contributions to  $\omega_F$ .

We have not included the  $\mathcal{O}(\alpha_s/\pi)$  contributions to the other matching coefficients in (16), namely  $\omega_S$ ,  $\omega_{W1}$  and  $\omega_{S2}$ . A difficult calculation would be necessary to determine them. Again, we use the equivalent parameters from the pure NRQED calculation [34] to estimate the systematic error. The pure NRQED equivalent of  $\omega_{W1}$  has log contributions in its first order coefficient and so we allow for an additive correlated systematic error of  $1 \pm \alpha_s/\pi$  to the tree level value. We allow the same error on  $\omega_{S2}$ .

The one-loop correction of the pure NRQED equivalent of  $\omega_S$  was found to be  $2\omega_F^{(1)} = \alpha/\pi$ . As such, we allow for an additive correlated systematic error on our  $\omega_S$  of  $1 \pm C_F\alpha_s/\pi$ , to compensate for using the tree level value in the calculation of the decay rate. This is a conservative estimate as we see above that the lattice artefacts actually subtract away some of this contribution over the mass range we are interested in.

The mixing down effects from diagrams (b)–(d) in Fig. 4 are difficult to estimate since each graph by itself can be IR divergent but  $\omega_F^{(1)}$  is IR finite. We allow an uncertainty of 30% in the one-loop coefficient (correlated across all lattice spacings) from neglecting the mixing down. There is no substitute for the actual calculation though, and we intend to do this in the future.

## V. RESULTS FOR THE $\Upsilon(2S) \rightarrow \eta_b(1S)\gamma$ DECAY

The unrenormalized form factors,  $\tilde{V}_{21}^{\Upsilon\eta}(q^2=0)|_i$ , for each of the currents obtained from (16) are computed for each of the ensembles listed in Table I and their values are given in Tables VI–X. A visual representation of  $\tilde{V}_{21}^{\Upsilon\eta}(q^2=0)|_i$  is shown in Fig. 6. From this, we can see that the form factor from the current  $J_F$  is leading order, and the other currents give a negative contribution to  $J_F$  of approximately 30%, 20%, 3% for  $J_{W1}$ ,  $J_S$ ,  $J_{S1}$  respectively across all ensembles. Note that these values do not appear

TABLE VI. Values of the unrenormalized form factors  $\tilde{V}_{21}^{\Upsilon\eta_b}|_i$ , as described in Sec. V, from the lattice NRQCD data on the ensemble labeled set 1 in Table I. We also give elements of the correlation matrix. A value of  $a^2q^2 = 0.0034(21)$  was found from the data.

$p$	Value	$C(p, \tilde{V} _F)$	$C(p, \tilde{V} _{W1})$	$C(p, \tilde{V} _S)$
$\tilde{V}_{21}^{\Upsilon\eta_b} _F$	0.1818(42)			
$\tilde{V}_{21}^{\Upsilon\eta_b} _{W1}$	-0.0594(12)	-0.4010		
$\tilde{V}_{21}^{\Upsilon\eta_b} _S$	-0.0339(17)	-0.2932	0.1261	
$\tilde{V}_{21}^{\Upsilon\eta_b} _{S1}$	-0.0037(3)	-0.0624	0.3488	-0.2678

to obey the power counting for the currents given in Sec. IV; however, we understand (and explain below) that the leading order contribution is suppressed for these hindered transitions. Similar behavior was seen in previous lattice NRQCD studies of this decay [5,6].

TABLE VII. Values and correlation matrix elements of the  $\tilde{V}_{21}^{\chi_{nb}}|_i$  from the ensemble labeled set 2 in Table I. A value of  $a^2q^2 = 0.00338(92)$  was found from the data.

$p$	Value	$C(p, \tilde{V} _F)$	$C(p, \tilde{V} _{W1})$	$C(p, \tilde{V} _S)$
$\tilde{V}_{21}^{\chi_{nb}} _F$	0.1765(22)			
$\tilde{V}_{21}^{\chi_{nb}} _{W1}$	-0.0593(7)	-0.5298		
$\tilde{V}_{21}^{\chi_{nb}} _S$	-0.0293(8)	-0.3803	0.3065	
$\tilde{V}_{21}^{\chi_{nb}} _{S1}$	-0.0045(2)	-0.0134	0.3962	-0.2264

TABLE VIII. Values and correlation matrix elements of the  $\tilde{V}_{21}^{\chi_{nb}}|_i$  from set 3 in Table I. A value of  $a^2q^2 = 0.0007(12)$  was found from the data.

$p$	Value	$C(p, \tilde{V} _F)$	$C(p, \tilde{V} _{W1})$	$C(p, \tilde{V} _S)$
$\tilde{V}_{21}^{\chi_{nb}} _F$	0.1720(36)			
$\tilde{V}_{21}^{\chi_{nb}} _{W1}$	-0.0577(10)	-0.2634		
$\tilde{V}_{21}^{\chi_{nb}} _S$	-0.0309(12)	-0.1887	0.2733	
$\tilde{V}_{21}^{\chi_{nb}} _{S1}$	-0.0032(3)	0.0213	0.1346	-0.1634

TABLE IX. Values and correlation matrix elements of the  $\tilde{V}_{21}^{\chi_{nb}}|_i$ , from set 4 in Table I. A value of  $a^2q^2 = 0.00066(70)$  was found from the data.

$p$	Value	$C(p, \tilde{V} _F)$	$C(p, \tilde{V} _{W1})$	$C(p, \tilde{V} _S)$
$\tilde{V}_{21}^{\chi_{nb}} _F$	0.1710(27)			
$\tilde{V}_{21}^{\chi_{nb}} _{W1}$	-0.0596(7)	-0.4441		
$\tilde{V}_{21}^{\chi_{nb}} _S$	-0.0289(10)	-0.3281	0.2708	
$\tilde{V}_{21}^{\chi_{nb}} _{S1}$	-0.0038(2)	0.0206	0.1493	-0.3195

TABLE X. Values and correlation matrix elements of the  $\tilde{V}_{21}^{\chi_{nb}}|_i$ , from set 5 in Table I. A value of  $a^2q^2 = -0.0021(6)$  was found from the data.

$p$	Value	$C(p, \tilde{V} _F)$	$C(p, \tilde{V} _{W1})$	$C(p, \tilde{V} _S)$
$\tilde{V}_{21}^{\chi_{nb}} _F$	0.1785(31)			
$\tilde{V}_{21}^{\chi_{nb}} _{W1}$	-0.0618(15)	-0.0703		
$\tilde{V}_{21}^{\chi_{nb}} _S$	-0.0276(10)	-0.0925	0.1526	
$\tilde{V}_{21}^{\chi_{nb}} _{S1}$	-0.0060(5)	0.0457	0.3260	-0.0266

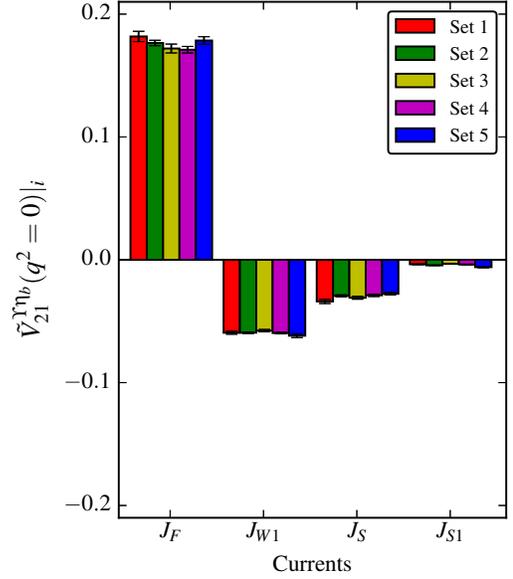


FIG. 6 (color online). The value of the unrenormalized form factor, as described in the text, arising from each current across the different ensembles listed in Table I. Statistical error only ( $\approx 2\% - 3\%$  for each current).

We also need to determine the sensitivity of our form factors to the different parameters used in our calculation and use this analysis to give a reliable error budget. This is easily done in lattice NRQCD, as we can simply change the value of a single parameter and rerun the whole calculation. The results are shown in Fig. 7, where we denote  $p$  as a parameter to vary (either  $c_i$  or  $m_b$ ) and use  $\Delta = p^{\text{test}} - p^{\mathcal{O}(\alpha_s)}$  to signify an upwards/downwards shift from the  $\mathcal{O}(\alpha_s)$  correct value  $p^{\mathcal{O}(\alpha_s)}$  [ $am_b$  is tuned fully nonperturbatively but we use  $am_b = p^{\mathcal{O}(\alpha_s)}$  to avoid additional superfluous notation]. The values of the changed parameters are given in Table XI.

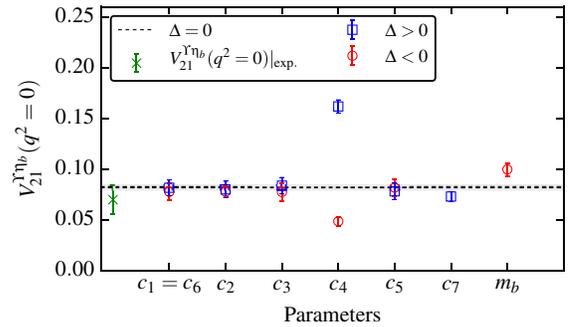


FIG. 7 (color online). The variation of the form factor with the parameters used in this study.  $\Delta > 0$  ( $\Delta < 0$ ) denotes an upward (downward) shift in the parameter as described in the text, and the values of the varied parameters can be found in Table XI. The data for  $\Delta > 0$  ( $\Delta < 0$ ) were generated on a subset of 400 configurations of the coarse lattice denoted set 2 in Table I. Error is statistical only.

TABLE XI. Values of the varied parameters used to obtain Fig. 7.  $\Delta > 0$  ( $\Delta < 0$ ) denotes an upward (downward) shift in the parameter as described in the text.  $p^{\mathcal{O}(\alpha_s)}$  for  $\Delta = 0$  values are taken from Table II and reproduced here for convenience.

Parameter	$p^{\text{test}}$ for $\Delta < 0$	$p^{\mathcal{O}(\alpha_s)}$ for $\Delta = 0$	$p^{\text{test}}$ for $\Delta > 0$
$c_1 = c_6$	1.00	1.31	1.50
$c_2$	0.75	1.02	1.25
$c_3$	0.75	1.00	1.25
$c_4$	1.00	1.19	1.50
$c_5$	1.00	1.16	1.50
$c_7$	...	1.00	1.50
$m_b$	2.5935	2.73	...

From Fig. 7 we can see that the form factor is most sensitive to the value of  $c_4$ , while  $c_7$  and  $m_b$  are also important. We need to describe this sensitivity in order to give a reliable estimate on the error from not knowing each of these parameters to infinite precision. Interestingly, it is useful to note that the sensitivity to these parameters comes from the  $J_F$  current, as shown in Fig. 8. We use a simple potential model analysis to understand the deficiencies in the naive power counting, where these sensitivities arise from, and to gain insight into this hindered M1 decay.

### A. Phenomenological insight: potential model analysis

In a potential model framework one would consider periodic harmonic time-dependent perturbations and find the matrix element as the overlap between the spatial part of

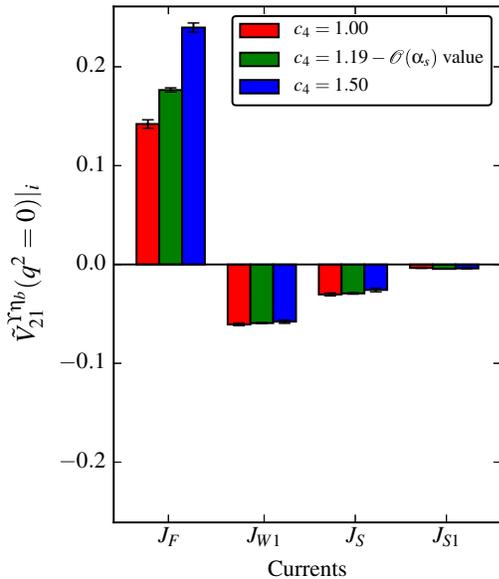


FIG. 8 (color online). How each of the unrenormalized form factors from the different currents vary with  $c_4$ . As can be seen, the sensitivity comes from the  $J_F$  current. The reason for this is described in Sec. VB 1.

the potential and the initial and final states under study. For an M1 decay, mediated by either of the constituent quarks' magnetic moment  $\boldsymbol{\sigma} \cdot \mathbf{B}$ , one can find the matrix element as [38] (labeling the spatial part of the potential as  $J_F$ , similar to the current we use in Sec. IV to highlight comparisons)

$$\begin{aligned} & \langle \eta_b(mS) | J_F | \Upsilon(nS) \rangle \\ &= S_{fi} \int_0^\infty dr r^2 R_{m,\eta_b}^*(r) j_0\left(\frac{|q|r}{2}\right) R_{n,\Upsilon}(r) \end{aligned}$$

with the integral expanded as

$$\begin{aligned} & \int_0^\infty dr r^2 R_{m,\eta_b}^*(r) j_0\left(\frac{|q|r}{2}\right) R_{n,\Upsilon}(r) \\ &= \delta_{nm} + a_2 |q_\gamma|^2 r_0^2 + a_4 |q_\gamma|^4 r_0^4 + \dots \end{aligned} \quad (20)$$

Here, we have factored the spin piece  $S_{fi}$  in the matrix element from the radial integral (appropriate in the non-relativistic limit) and used the Taylor expansion of  $j_0(x) = \sin(x)/x = \sum_n (-1)^n x^{2n}/(2n+1)!$  to see that it is a polynomial in  $|q_\gamma|^2$ . Additionally, the only scale in the wave functions capable of being combined with  $|q_\gamma|^2$  to make it dimensionless is some combination of the Bohr radii of each state, which we call  $r_0$ . The  $a_{2l}$  are coefficients which could be calculated if wave functions were supplied. The leading Kronecker  $\delta$ -function in (20) comes from noting the orthogonality condition in the extreme non-relativistic limit,  $|q_\gamma|^2 \rightarrow 0$ .

As can be seen, for a  $nS \rightarrow nS$  transition, the leading order term in (20) is one. However, for transitions between different radial excitations, the  $\delta_{nm}$  vanishes and we are left with a leading order  $|q_\gamma|^2 r_0^2$  term. The radii of the bottomonium states under study are of the order of the reciprocal of the typical momentum, e.g.  $r_0 \sim 1/mv$ . Thus, as  $|q_\gamma|^2 r_0^2 \sim m^2 v^4 / (m^2 v^2) \sim v^2$ , the leading order matrix element from  $J_F$  in a radially excited decay is suppressed by a factor of  $v^2$  more than naively expected from using power-counting rules on the currents alone. This suppression leads to an array of sensitivities that make this decay particularly difficult to pin down theoretically from within a potential model [1], as we expand upon in Sec. VI.

Due to the derivatives in the other currents listed in (16), the matrix elements of these currents give rise to wave function overlaps that are not orthogonal in the extreme nonrelativistic limit, and as such are not more suppressed for radially excited transitions. The derivatives act on the initial bottomonium state and give a leading order  $p \sim \mathcal{O}(mv)$  effect, which does not depend on the photon momentum, as can be seen by taking the  $|q_\gamma| \rightarrow 0$  limit. This results in the relativistic corrections to the leading order  $J_F$  current, which we have included in our calculation, having appreciable effects (see Fig. 6), namely  $J_{W1}$ ,  $J_S$ . The orthogonality of the radial wave function muddles

up the power counting of the first few currents, but additional derivatives in relativistic corrections to these currents will suppress them further. By including the current  $J_{S2}$ , we check that added derivatives do suppress the contribution of the current further as expected.

By examining (20), we found that the leading order matrix element for the radially excited radiative transition can be suppressed more than we would naively expect from just power counting the current alone. Relativistic corrections to the  $J_F$  current are then appreciable, explaining the behavior seen in Fig. 6. Even if we included the relativistic corrections to the current in a potential model, we still would not get the correct value for this decay, as we also need to consider all relativistic corrections to the wave functions arising from perturbative potentials in the Hamiltonian. This gives rise to the sensitivities to the different parameters as seen in Fig. 8, which we explain below. To do so, it is sufficient to consider first order time-independent perturbation theory.

### B. Sensitivity and errors from terms in the NRQCD action

We want to consider potentials arising from relativistic corrections in the NRQCD action causing perturbations of the wave function. To first order in  $\alpha_s$  we have

$$\begin{aligned} |\eta_b(1S)\rangle^{(1)} &= |\eta_b(1S)\rangle^{(0)} - \sum_{m \neq 1} |\eta_b(mS)\rangle^{(0)} \frac{V_{m1}^{\eta_b}}{E_{m1}^{\eta_b}} \\ |\Upsilon(2S)\rangle^{(1)} &= |\Upsilon(2S)\rangle^{(0)} - \sum_{n \neq 2} |\Upsilon(nS)\rangle^{(0)} \frac{V_{n2}^{\Upsilon}}{E_{n2}^{\Upsilon}}. \end{aligned} \quad (21)$$

The state  $|n\rangle^{(1)}$  ( $|n\rangle^{(0)}$ ) is the first order perturbed state (the unperturbed state),  $V_{nm} = {}^{(0)}\langle n|V|m\rangle^{(0)}$  with  $V$  being the potential representing the perturbation and  $E_{nm} = E_n^{(0)} - E_m^{(0)}$ . Now, we take currents of interest between these states to yield

$$\begin{aligned} &{}^{(1)}\langle \eta_b(1S)|J_i|\Upsilon(2S)\rangle^{(1)} \\ &= {}^{(0)}\langle \eta_b(1S)|J_i|\Upsilon(2S)\rangle^{(0)} \\ &\quad - \sum_{m \neq 1} \frac{V_{m1}^{\eta_b *}}{E_{m1}^{\eta_b}} {}^{(0)}\langle \eta_b(mS)|J_i|\Upsilon(2S)\rangle^{(0)} \\ &\quad - \sum_{n \neq 2} \frac{V_{n2}^{\Upsilon}}{E_{n2}^{\Upsilon}} {}^{(0)}\langle \eta_b(1S)|J_i|\Upsilon(nS)\rangle^{(0)}. \end{aligned} \quad (22)$$

As mentioned above, for the current  $J_F$ , due to the fact that  ${}^{(0)}\langle \eta_b(1S)|J_F|\Upsilon(2S)\rangle^{(0)}$  is suppressed for radially excited decays, the  ${}^{(0)}\langle \eta_b(nS)|J_F|\Upsilon(nS)\rangle^{(0)}$  pieces in the second term in (22) become appreciable. The matrix elements arising from currents with derivatives are already suppressed, and the first order corrections to these matrix elements are not appreciable, as seen in Fig. 8.

#### 1. Sensitivity and error from $c_4$

Including a potential from the exchange of a single gluon between two vertices involving the chromomagnetic operator as shown in Appendix C, we find

$$\begin{aligned} &{}^{(1)}\langle \eta_b(1S)|J_F|\Upsilon(2S)\rangle^{(1)} = {}^{(0)}\langle \eta_b(1S)|J_F|\Upsilon(2S)\rangle^{(0)} + \frac{c_4^2 g^2}{9m_b^2 E_{21}} \psi_1^*(0) \psi_2(0) \\ &\quad \times (6 {}^{(0)}\langle \eta_b(2S)|J_F|\Upsilon(2S)\rangle^{(0)} + 2 {}^{(0)}\langle \eta_b(1S)|J_F|\Upsilon(1S)\rangle^{(0)} + \mathcal{O}(v^2)) \\ &= {}^{(0)}\langle \eta_b(1S)|J_F|\Upsilon(2S)\rangle^{(0)} + \frac{8c_4^2 g^2}{9m_b^2 E_{21}} \mathcal{S}_{fi} \psi_1^*(0) \psi_2(0) + \mathcal{O}(v^2). \end{aligned} \quad (23)$$

The reason for the sensitivity to  $c_4$  is clear. The matrix element  ${}^{(0)}\langle \eta_b(1S)|J_F|\Upsilon(2S)\rangle^{(0)}$  is suppressed due to the orthogonality of the radial wave functions in (20), while  ${}^{(0)}\langle \eta_b(nS)|J_F|\Upsilon(nS)\rangle^{(0)}$  is not. This results in the second term in (23) being sizeable compared to the first.

Since we have values of the form factor at three values of  $c_4$  on a coarse lattice as shown in Fig. 8, and an understanding that the functional dependence of the form factor on  $c_4$  should be  $\tilde{V}_F^{\Upsilon} = a_{c_4} + c_4^2 b_{c_4}$  from (23), we should check that this is consistent. We use the  $c_4 = 1.00$  and  $c_4 = 1.19$  values from our lattice NRQCD calculation listed in Table XII to find the values of  $a_{c_4}$  and  $b_{c_4}$  in Table XIII.

We can also relate the second term from the leading order approximation in (23) to quantities that are measured in experiment and check the consistency of the value of  $b_{c_4}$  given in Table XIII. By comparing the decay rate formulas from a potential model calculation [9] with the one given in (2), we find

$$\begin{aligned} V_{21}^{\Upsilon \eta_b} &= \left( \frac{m_{\Upsilon(2S)} + m_{\eta_b(1S)}}{2m_b} \right) \\ &\quad \times \int_0^\infty dr r^2 R_{m, \eta_b}^*(r) j_0\left(\frac{|q|r}{2}\right) R_{n, \Upsilon}(r) \end{aligned}$$

TABLE XII. Values of the form factor  $\tilde{V}_{21}^{Y\eta_b}|_F$  with a variation of certain parameters from the lattice NRQCD data on a coarse lattice (set 2 in Table I). Error is statistical only.

$p$	Value	$C(p, \tilde{V}_{21}^{Y\eta_b} _{F,c_4=1.00})$	$C(p, \tilde{V}_{21}^{Y\eta_b} _{F,c_4=1.19})$	$C(p, \tilde{V}_{21}^{Y\eta_b} _{F,c_7=1.50})$	$C(p, \tilde{V}_{21}^{Y\eta_b} _{F,c_2=1.25})$
$\tilde{V}_{21}^{Y\eta_b} _{F,c_4=1.00}$	0.1426(47)				
$\tilde{V}_{21}^{Y\eta_b} _{F,c_4=1.19}$	0.1772(44)	0.3040			
$\tilde{V}_{21}^{Y\eta_b} _{F,c_7=1.50}$	0.1687(67)	0.0342	0.0472		
$\tilde{V}_{21}^{Y\eta_b} _{F,c_2=1.25}$	0.1769(46)	0.2979	0.3352	0.0467	
$\tilde{V}_{21}^{Y\eta_b} _{F,m_b=2.59}$	0.1939(48)	0.3070	0.3479	0.0508	0.3411

TABLE XIII. Values of the functional dependency of  $\tilde{V}_{21}^{Y\eta_b}|_F$  with parameters from the action using data from Table XII. See the text for details. Error is statistical only.

$p$	Value	$C(p, a_{c_4})$	$C(p, b_{c_4})$	$C(p, a_{c_7})$	$C(p, b_{c_7})$	$C(p, a_{c_2})$	$C(p, b_{c_2})$	$C(p, a_{m_b})$	$C(p, b_{m_b})$
$a_{c_4}$	0.060(16)								
$b_{c_4}$	0.083(13)	-0.974							
$a_{c_7}$	0.194(18)	-0.257	0.395						
$b_{c_7}$	-0.017(16)	0.202	-0.307	-0.979					
$a_{c_2}$	0.179(24)	-0.389	0.510	0.486	-0.378				
$b_{c_2}$	-0.001(21)	0.366	-0.459	-0.403	0.316	-0.988			
$a_{m_b}$	0.077(34)	-0.382	0.487	0.442	-0.346	0.562	-0.506		
$b_{m_b}$	2.04(65)	0.363	-0.448	-0.381	0.300	-0.512	0.469	-0.994	

and then using this in (23) yields

$$c_4^2 b_{c_4} = \left( \frac{m_{\Upsilon(2S)} + m_{\eta_b(1S)}}{2m_b} \right) \frac{\sqrt{\Delta(2S)\Delta(1S)}}{E_{21}} + \mathcal{O}(v^2), \quad (24)$$

where  $\Delta(iS)$  is the hyperfine splitting between  $i$ 'th radial excitations. Using the values of  $c_4$ ,  $a$  and  $am_b$  from set 2 in Table II, along with the Particle Data Group (PDG) average [7] values for  $\Delta(iS)$  and the spin-averaged  $E_{21}$ , we find  $b_{c_4} = 0.105(14)$ . This is consistent with the value of  $b_{c_4}$  from Table XIII.

In Fig. 9, we show the strong  $c_4$  dependence of  $\tilde{V}_{21}^{Y\eta_b}|_F = a_{c_4} + c_4^2 b_{c_4}$ , along with the the lattice values of  $\tilde{V}_{21}^{Y\eta_b}|_F$  shown in Fig. 8. This illustrates both the need for at least the  $\mathcal{O}(\alpha_s)$ -correct value of  $c_4$  and the consistency of  $a_{c_4}$  and  $b_{c_4}$  with all our lattice data.

Since we only know  $c_4$  to one loop in perturbation theory, there will be a systematic error associated with not knowing it to higher orders. With the above functional dependence of  $\tilde{V}_{21}^{Y\eta_b}|_F = a_{c_4} + c_4^2 b_{c_4}$ , an error of  $2\alpha_s^2 b_{c_4}$  should be introduced from not knowing  $c_4$  to second order. As there is little lattice spacing dependence in the unrenormalized form factors as shown in Fig. 6, we use the value of  $b_{c_4}$  from Table XII across all ensembles and introduce an additive systematic error (correlated across lattice spacings) of  $2\alpha_s^2 b_{c_4}$  from not knowing  $c_4$  to more than one loop. We

also allow for the statistical error in the determination of  $c_4^{(1)}$  coming from the Vegas integration [20] by adding an error of  $2\alpha_s \delta c_4^{(1)} b_{c_4}$ .

With the other currents that have derivatives, the situation is significantly different. Due to the derivatives, the second term in (22) is always suppressed and relativistic corrections are not an appreciable effect, as seen in Fig. 8. Variations of these currents with  $c_4$  are not appreciable within the other errors.

## 2. Sensitivity and error from $c_7$

The  $c_7$  operator is a  $D^2$  correction to the  $c_4$  term and is expected to be a  $\mathcal{O}(v^2)$  effect. We can proceed as before,

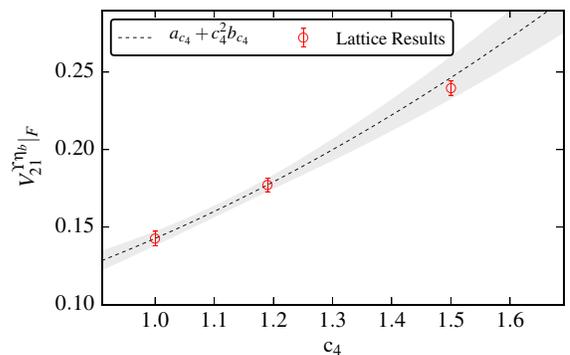


FIG. 9 (color online). The  $c_4$  dependence of  $\tilde{V}_{21}^{Y\eta_b}|_F$  as described in the text, along with the lattice values of  $\tilde{V}_{21}^{Y\eta_b}|_F$ .

assuming a linear functional dependence on  $c_7$  as  $\tilde{V}_{21}^{Y\eta_b}|_F = a_{c_7} + c_7 b_{c_7}$ , coming from the exchange of a single gluon from a  $c_4$  vertex and a  $c_7$  vertex. Using our data points in Table XII, we find  $a_{c_7}, b_{c_7}$  listed in Table XIII.

It is seen that  $b_{c_7}$  gives a negative contribution as a consequence of the  $D^2$  and the ratio  $b_{c_7}/b_{c_4} = -0.20(18)$  should be a  $\mathcal{O}(v^2)$  effect. This is roughly consistent. We assume a dependence on  $c_7$  as  $b_{c_7} \approx 2v^2 b_{c_4} = 0.2b_{c_4}$ , and similarly to the  $c_4$  error above, introduce an additive systematic error (correlated across lattice spacings) of  $2\alpha_s v^2 b_{c_4}$  from not knowing  $c_7$  past tree level. Just as with variations of  $c_4$ , the currents with derivatives are insensitive to variations of  $c_7$  and are all consistent within our small statistical errors.

### 3. Sensitivity and error from $m_b$

Using the fact that radial splittings are expected to be  $E_{21} \sim m_b v^2$ , by examining (23) we observe that the form factor should have a functional dependence on  $m_b$  as  $\tilde{V}_{21}^{Y\eta_b}|_F = a_{m_b} + b_{m_b}/m_b^3$ . Using our data points in Table XII, we find  $a_{m_b}, b_{m_b}$  listed in Table XIII.

Again, we can check consistency within this first order approximation. Comparing the assumed functional forms against the equation from which they came (23), we find  $b_{m_b} = c_4^2 m_b^3 b_{c_4}$ . Thus, using the values of  $b_{c_4}, b_{m_b}$  we obtain from the lattice data, we find the ratio  $b_{m_b}/c_4^2 m_b^3 b_{c_4} = 0.85(35)$ , consistent with 1.0.

We allow for a systematic error from the (small) uncertainty in mistuning the  $b$ -quark mass estimated from [13]. By using the above inverse cubic functional dependence on  $m_b$ , we find an error of  $3b_{m_b}\delta_{m_b}/m_b^4$ . Using the estimate of  $b_{m_b}$  in terms of  $b_{c_4}$ , we find the error as  $3c_4^2 b_{c_4}\delta_{m_b}/m_b$ .

### 4. Sensitivity and error from $c_2$

From our numerical data, it appears as if the form factor is not sensitive to a variation in  $c_2$ . We can understand this and use it in our analysis of the errors. In Appendix C we show how the the leading spin-independent perturbative potential from the exchange of a single gluon involving the Darwin term at one of the vertices [20] gives rise to a correction to the leading order matrix element that is  $\mathcal{O}(\alpha_s v^2)$ . Using the data in Table XII for how  $\tilde{V}_{21}^{Y\eta_b}|_F$  varies with  $c_2$ , and using the functional form  $\tilde{V}_{21}^{Y\eta_b}|_F = a_{c_2} + c_2 b_{c_2}$ , we find the values listed in Table XIII.

To test the consistency of this description, by comparing the value  $b_{c_4}$  associated with the second term in (23) and the second term in (C5) we see  $b_{c_2} \approx 3v^2 b_{c_4}/8$ . Using the values in Table XIII gives  $3v^2 b_{c_4}/8 = 0.00311(49)$ , consistent with  $b_{c_2} = 0.001(21)$ . Due to the smallness of this dependency, we can safely neglect the systematic error from not knowing  $c_2$  to two-loop order.

### 5. Sensitivity and error from $c_3$

Since the bottomonium states under study have no orbital angular momentum, there is no sensitivity to  $c_3$  arising from a spin-orbit perturbing potential. This is confirmed by the numerical data in Fig. 7. We introduce no error from  $c_3$ .

### 6. Sensitivity and error from four-quark operators

The four-quark operators in NRQCD [13] are contact terms between the quark and antiquark fields arising from  $\alpha_s^2$  processes in relativistic QCD. These can have a noticeable effect on the hyperfine splitting [16]. Since the matrix element in (23) is sensitive to parameters in much the same way as the hyperfine splitting, we would expect contributions from the four-quark operators. In Appendix C, we show the effect of the four-quark potential on the matrix element to first order.

We introduce a systematic error (correlated across lattice sites) for neglecting these leading order four-quark operators in our calculation. We estimate this by comparing the second term in (23) with the second term in (C7) to find an error  $27b_{c_4}(d_1\alpha_s - d_2\alpha_s)/16\pi$  and then use the values of  $d_1\alpha_s - d_2\alpha_s$  from [20] (as corrected per [39]).

### 7. Error from missing higher order operators in the NRQCD action

The terms in the action that have not been considered are the  $\mathcal{O}(v^2)$  corrections to the  $c_2$  and  $c_7$  terms. Since the coefficient  $b_{c_2}$  is already quite small, the  $v^2$  correction to this will be negligible within our numerical precision and can be neglected. The error from  $v^2$  corrections to  $c_7$  is estimated as  $v^2 b_{c_7} = 2v^4 b_{c_4}$ .

### 8. Total error on $\tilde{V}_{21}^{Y\eta_b}|_F$ from terms in the NRQCD action

After performing the final continuum and chiral extrapolation as shown in Sec. VD, we can obtain a breakdown of how each of the uncertainties arising from the NRQCD action effects the error in  $\tilde{V}_{21}^{Y\eta_b}|_F$  as a percentage of the error on the total form factor given in Table XIV. We find that the errors from the NRQCD action contribute to a 10.4% systematic error in  $\tilde{V}_{21}^{Y\eta_b}|_F$  as a percentage of the total error on the total form factor. In order of dominance, the most sizable of these errors is a 7.9% error from neglecting the  $\mathcal{O}(\alpha_s^2)$  correction in  $c_4$ , then a 4.4% error from the statistical error in  $c_4^{(1)}$  while 3.9% comes from neglecting the one-loop correction to  $c_7$ . These numbers should be added in quadrature and each is a percentage of the total error on the total form factor.

Note that due to the destructive interference between the leading order form factor,  $\tilde{V}_{21}^{Y\eta_b}|_F$ , and the other currents as shown in Sec. V, the error coming from  $\tilde{V}_{21}^{Y\eta_b}|_F$  as a

TABLE XIV. Full error budget for the total form factor  $V_{21}^{\Upsilon\eta_b}$  relevant for the  $\Upsilon(2S) \rightarrow \eta_b(1S)\gamma$  decay from Fig. 10. A discussion of the uncertainties in  $\tilde{V}_{21}^{\Upsilon\eta_b}|_F$  is given in Sec. VB 8. The form factor inferred from experimental data in Sec. II is  $V_{21}^{\Upsilon\eta_b}|_{\text{exp}} = 0.069(14)$  and has a relative error of 19.74%.

Error %	$V_{21}^{\Upsilon\eta_b}$
Systematic $\tilde{V}_{21}^{\Upsilon\eta_b} _F$	10.36
Stats in $V_{21}^{\Upsilon\eta_b}$	5.48
Radiative $\alpha_s^2$ in $\omega_F$	0.83
Radiative $\alpha_s$ in $\omega_{W1}$	4.71
Radiative $\alpha_s$ in $\omega_S$	2.36
Radiative $\alpha_s$ in $\omega_{S1}$	0.51
Mixing down in $\omega_F^{(1)}$	3.92
Missing currents	7.08
$a_{fm}$ scale	1.07
Experimental masses	0.03
Priors	4.18
Total	15.81

percentage of the total error on  $V_{21}^{\Upsilon\eta_b}$  is larger than the errors on  $\tilde{V}_{21}^{\Upsilon\eta_b}|_F$  alone. As a result, improvement of the errors coming from the NRQCD action has an appreciable effect.

### 9. Test of uncertainties from the NRQCD action

To ensure that we have performed a reasonable estimation of the errors arising from the NRQCD action, we have also tuned  $c_4$  against the  $\Upsilon(1S) - \eta_b(1S)$  hyperfine splitting on the coarse lattice denoted set 2 in Table I. In a perturbative framework as described above, the hyperfine splitting can be pictured as a result of perturbative potentials shifting the unperturbed energies. The most sizable of these is the leading order  $c_4^2$  potential, as described in Sec. VB 1, and then the four-quark potential, as described in Sec. VB 6. In a numerical calculation with no four-fermion operators, tuning the numerical hyperfine splitting against the experimental one would have the effect of absorbing the above four-fermion term (among others) into the tuned  $c_4$ . Stated more concretely,

$$(c_4^{\text{lat}})^2 \rightarrow (c_4^{\text{tuned}})^2 = c_4^2 - \frac{27}{16\pi}(d_1 - d_2)\alpha_s. \quad (25)$$

Then, putting  $(c_4^{\text{tuned}})^2$  into (23) gives exactly the four-fermion term which we need in (C7). As such, using  $c_4^{\text{tuned}}$  numerically would include the effect of the four-fermion operator for this decay automatically. For the nonperturbative tuned  $c_4^{\text{tuned}}$  error budget, there are no  $c_7$ , leading order four-quark, or missing  $v^8$  operator errors as these will be absorbed into the value of  $c_4^{\text{tuned}}$  and fed back into the matrix element calculation automatically. However, from (C7) we see there is still an additive systematic error of

$3v^2(27/16\pi)\alpha_s b_{c_4}$  from only knowing the difference  $(d_1 - d_2)$ , and not  $d_1$  and  $d_2$  individually.

The Particle Data Group average for the hyperfine splitting is  $\Delta^{\text{exp}} = 62.3(3.2)$  MeV [7], while our lattice calculation with  $c_4 = 1.23$  gives  $\Delta^{\text{lat}} = 62.54(46)$  MeV (statistical and scale setting error only). We get a value of  $c_4^{\text{tune}} = 1.230(5)(31)$  from tuning  $c_4$  against the experimental hyperfine splitting, where the first error is from the lattice, and the second is from experiment. The change from the one-loop perturbative value 1.19 to the non-perturbatively tuned 1.230(5)(31) is well accounted for in the error budget (see Sec. VB 8) from the statistical error on  $\delta c_4^{(1)}$  alone, and including the higher order corrections to  $c_4$  and the four-quark error is significantly overcompensating for this change.

Rerunning the computation of the form factor with  $c_4 = 1.23$  gives a value of  $V_{21}^{\Upsilon\eta_b} = 0.097(14)$ . This includes all errors, and the only difference from the above error budget is that the error in  $\tilde{V}_{21}^{\Upsilon\eta_b}|_F$  now comes from  $c_4^{\text{tune}}$  and the error from knowing only the difference  $d_2 - d_1$ . This value is to be compared with the form factor from a perturbatively tuned  $c_4$  shown in Sec. VD, i.e.,  $V_{21}^{\Upsilon\eta_b} = 0.089(22)$ . These are entirely consistent, giving evidence that our error budget is a reliable estimation of the errors.

The four-quark operators appear to increase the value of the form factor, in a similar way as they do for the hyperfine splitting. However, it was found that including the four-quark operators in the calculation of the hyperfine splitting largely changed the slope of the continuum extrapolation but did not shift the final result away from the value computed without the four-fermion operators included [16].

Based on our analysis, we estimate that by tuning  $c_4$  against the hyperfine splitting on all ensembles and redoing the full calculation, one could reduce the error on  $\tilde{V}_{21}^{\Upsilon\eta_b}|_F$  to  $\sim 4\%$ . Also, we estimate that such a calculation would give an error on the final form factor of  $\sim 11\%$  (compared against the value given in Table XIV), where now the uncertainties in order of dominance are from the neglected currents, neglecting the mixing down in  $\omega_F^{(1)}$ , and neglecting the one-loop correction to  $\omega_{W1}$ .

### C. Errors from missing higher order currents

Since we are using an effective field theory to study this transition, there will be higher order currents which we have not included in this study but that contribute to the final form factor. The most sizable current which we have not included is the  $D^2$  correction to  $J_{W1}$ . Therefore, we include a systematic uncertainty (correlated across all lattice sites) of  $v^2\tilde{V}_{21}^{\Upsilon\eta_b}|_{W1}$ .

### D. Full error budget

After the analysis performed in the previous sections, we are now in a position to give a full error budget for the

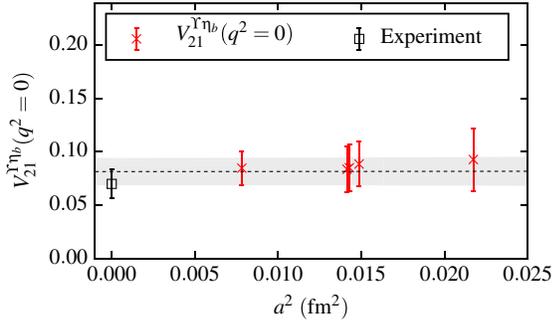


FIG. 10 (color online). Fit results for the form factor relevant to the  $\Upsilon(2S) \rightarrow \eta_b(1S)\gamma$  decay, all errors included. The error budget is shown in Table XIV.

form factor  $V_{21}^{\Upsilon\eta_b}$ . To compare to experiment, we perform a simultaneous lattice spacing and sea quark mass extrapolation. We fit results from all ensembles to the form [13,40]

$$V(a^2, am_b) = V_{\text{phys}} \times \left[ 1 + \sum_{j=1,2} (a\Lambda)^{2j} k_j (1 + k_{j1} \delta x_m + k_{j2} (\delta x_m)^2) + 2l_1 \delta m (1 + l_2 (a\Lambda)^2) \right]. \quad (26)$$

The lattice spacing dependence is set by a scale  $\Lambda = 500$  MeV,  $\delta x_m = (am_b - 2.7)/1.5$  allows for a mild dependence on the effective theory cutoff  $am_b$ , and  $\delta x_l = (am_l/am_s) - (am_l/am_s)_{\text{phys}}$  for each ensemble with  $(m_l/m_s)_{\text{phys}} = 27.4(1)$  is taken from lattice QCD [41]. We take a Gaussian prior on the leading order  $a^2$  term to be 0.0(3), as the HISQ action is correct through  $\mathcal{O}(\alpha_s a^2)$ ; a prior of 0.0(1.0) on the higher order  $a$  terms; a prior of 0.00(3) on  $l_1$  allowing for a 3% shift if the light quarks were as heavy as the strange; and a prior of 0.10(5) on  $V_{\text{phys}}$ .<sup>6</sup> The extrapolation with all errors is shown in Fig. 10 and a full error budget is shown in Table XIV.

By studying the error budget we see that the main sources of error are from the systematics in  $\tilde{V}_{21}^{\Upsilon\eta_b}|_F$ . Here, as discussed in Sec. VB 8, the main sources of uncertainty come from the statistical error in  $c_4^{(1)}$  and from not knowing the coefficient of  $\alpha_s^2$  in the expansion of  $c_4$ . While the statistical error on  $c_4^{(1)}$  could potentially be reduced from 7%–10% to 2%–3% [20], computation of the two-loop coefficient of  $\alpha_s^2$  would be difficult and lengthy, and unlikely to be done in the near future. Alternatively, one could tune  $c_4$  against the hyperfine splitting on all ensembles, as shown in Sec. VB 9, and the error on  $V_{21}^{\Upsilon\eta_b}$  could be reduced to  $\sim 11\%$ .

<sup>6</sup>The width on this prior is chosen so as to ensure that the fitted result is insensitive to the central value.

After this, the main uncertainty comes from the missing currents. These could be included with more computational time if necessary. While the statistical error on each current alone is around 3%, these statistical errors do not allow the correlations between the data points in the fit to constrain the final result as much as we would like, and the final error from statistics in the error budget is 5% as a result. Reducing the error from statistics is unlikely to have a sizable effect.

Based on our analysis, we estimate that by including the next order of relativistic corrections to the current, the mixing down in  $\omega_F^{(1)}$ , and tuning  $c_4$  against the hyperfine splitting on all ensembles, an error on  $V_{21}^{\Upsilon\eta_b}$  of 8% could be possible (compared against an error of 19% on the value inferred from experiment), where the uncertainties in order of dominance would be from the one-loop corrections to  $\omega_{W1}$  and  $\omega_S$  and the systematic error coming from  $\tilde{V}_{21}^{\Upsilon\eta_b}|_F$ .

Our final answer for the form factor is

$$V_{21}^{\Upsilon\eta_b}(q^2 = 0) = 0.081(13). \quad (27)$$

Final values for the decay rate and branching fraction are given in Sec. VI.

## VI. DISCUSSION AND CONCLUSIONS

In this paper we have computed the hindered M1  $\Upsilon(2S) \rightarrow \eta_b(1S)\gamma$  decay rate using a lattice NRQCD formalism for the  $b$  quark. We include several improvements on earlier exploratory work [5,6] which are fundamental to obtaining an accurate value for this decay rate. The key improvements are the following:

- (i) Previous work only had one lattice spacing. We use five ensembles with a fully  $\mathcal{O}(\alpha_s a^2)$  tadpole-improved Lüscher-Weisz gluon action with HISQ  $u$ ,  $d$ ,  $s$  and  $c$  quarks in the sea, provided by the MILC collaboration. These ensembles each have  $\sim 1000$  configurations and one has physical light quark masses.
- (ii) We use three relativistic corrections to the leading order current as described in Sec. IV and we also test the sensitivities of the form factors from all these currents to the parameters in our action as shown in Fig. 7.
- (iii) We use  $\mathcal{O}(\alpha_s)$  correct values for the matching coefficients in the NRQCD action. We also take into account issues in tuning the  $b$ -quark mass as described in Sec. III B. As shown in Fig. 7, this decay is very sensitive to a subset of these parameters.
- (iv) We calculate the  $\mathcal{O}(\alpha_s)$  contribution to the matching coefficient of the leading order  $\psi^\dagger \sigma \cdot \mathbf{B}_{\text{QED}} \psi / 2m_b$  current which mediates this decay, as described in Sec. IV A.
- (v) While previous work extracted the matrix element by extrapolating/interpolating to the  $|\mathbf{q}|_{\text{phys}}$  point,

which only gives the photon on-shell contribution  $q^2 = 0$  if the hyperfine splitting is correct, we use twisted boundary conditions to extract the form factor relevant to this decay at the physical  $q^2 = 0$  point.

In Sec. III E we performed an analysis of the energy eigenstates of NRQCD at nonzero momentum. This is necessary as the energy eigenstates of a rotationally invariant theory, like NRQCD, in an infinite-volume continuum at nonzero momentum are classified by helicity, unlike in a Lorentz-invariant theory where they are described by the standard angular momentum  $J$ . This has important consequences for a lattice NRQCD calculation as additional states appear in the spectrum at nonzero momentum (see Fig. 1) and one has to be careful to ensure that the correct matrix elements are extracted from the correlator data.

In Sec. V, we show results for the four form factors from the currents listed in Sec. IV which when renormalized, summed and extrapolated to the continuum limit can be compared to the form factor inferred from experimental data. We found that relativistic corrections to the leading order current gave a negative contribution causing destructive interference, the power counting of the currents deviated from what one would naively expect in NRQCD, and a range of sensitivities needed to be explained.

In Sec. V B, using a simple potential model, we explained that the matrix element of the leading order current was suppressed due to the orthogonality of the radial wave functions, and this causes the matrix element to become sensitive to a multitude of effects such as relativistic corrections to the leading order current, and certain parameters in the NRQCD action that give rise to perturbing potentials causing relativistic corrections to the wave functions, particularly those which effect the hyperfine splitting.

It has been suggested [5,6] that the large changes experienced in going from an unimproved calculation to an improved calculation may mean that it would be beneficial to avoid nonrelativistic approximations. We come to a different conclusion and illustrate that although such a calculation is intrinsically difficult, NRQCD does indeed show that a systematic approach works while also giving insight into the process under study.

After performing the continuum and sea quark mass extrapolation, we obtain the form factor  $V_{21}^{\Upsilon\eta_b}(0)|_{\text{lat}} = 0.081(13)$ , with a full error budget in Table XIV. This form factor can be combined with the experimental masses used in Sec. II to produce the decay rate:

$$\Gamma^{\text{lat}}(\Upsilon(2S) \rightarrow \eta_b(1S)\gamma) = 1.72(55) \times 10^{-2} \text{ keV} \quad (28)$$

which can be compared against the experimental decay rate  $\Gamma^{\text{exp}}(\Upsilon(2S) \rightarrow \eta_b(1S)\gamma) = 1.25(49) \times 10^{-2} \text{ keV}$  [2,7]. Using the experimental total width from the PDG average

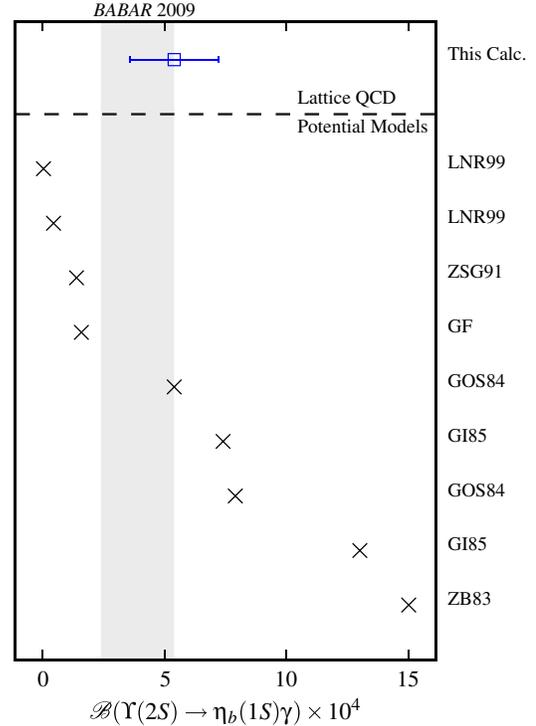


FIG. 11 (color online). Comparison of our result for the branching fraction (square) with experiment (vertical gray band) and potential model estimates from [1] (crosses). The y axis labels the different references [42–46] and more information about these can be found in [1]. Using the pNRQCD decay rate [4], combined with the experimental total width from the PDG average given in Sec. II, gives a branching fraction of  $1.9_{-1.9}^{+8.1} \times 10^{-4}$ .

given in Sec. II with our decay rate gives a branching fraction of  $\mathcal{B}(\Upsilon(2S) \rightarrow \eta_b(1S)\gamma) = 5.4(1.8) \times 10^{-4}$  which can be compared against the *BABAR* result of  $3.9(1.5) \times 10^{-4}$  [2]. A comparison of our calculation with potential model results including relativistic corrections [1] is shown in Fig. 11.

Potential model predictions of hindered M1 decay rates are known to be particularly difficult to pin down [38] and can mischaracterize the experimental data by an order of magnitude without relativistic corrections [8]. According to the Quarkonium Working Group reviews [8,38], sources of uncertainty that contribute to making such decays complicated to calculate include the form of the long range potential chosen, and the results depending explicitly on the quark mass and the perturbative potential chosen. Without relativistic corrections, the branching fraction of the  $\Upsilon(2S) \rightarrow \eta(1S)\gamma$  decay from potential model predictions ranges from  $(0.67 - 11.0) \times 10^{-4}$  [1]. Due to the suppression mentioned above, the value of the decay rate is very dependent on good knowledge of the relativistic corrections [1]. Including relativistic corrections, potential model predictions for the same branching fraction have a wider range  $(0.05 - 15.0) \times 10^{-4}$ , showing indeed that the

decay rates may be sensitive to small details of the potential [1].

The  $\Upsilon(2S) \rightarrow \eta_b(1S)\gamma$  decay is sensitive to many of the same effects as the hyperfine splitting and an accurate calculation of this decay relies on having the correct hyperfine splitting. Given the large range of estimates of the hyperfine splitting from potential model predictions (46–87 MeV [38]), we should not be surprised that the potential model estimates for this decay rate also have a large range.

Additionally, radiative transitions between bottomonium states provide a search for new-physics effects, separate from the weak-sector searches common in the literature [47]. For example, the hyperfine splitting between the  $\Upsilon(1S)$  and  $\eta_b(1S)$  has been an important quantity in bottomonium physics, being historically difficult for both experimentalists and theorists to predict reliably. Using hindered M1 decays, the *BABAR* [2,48] and *CLEO* [49] experiments inferred this hyperfine splitting to be  $\Delta_{M1}^{\text{exp}} = 69.3 \pm 2.8$  MeV [50]. However, in 2012, *BELLE* measured the  $h_b(2P, 1P) \rightarrow \eta_b(1S)\gamma$  branching fractions (called E1 decays in the literature), removing the dependence on hindered M1 decays, and used a significantly larger sample of events to yield a hyperfine splitting of  $\Delta_{E1}^{\text{exp}} = 57.9 \pm 2.3$  MeV [51], where  $\Delta_{M1}^{\text{exp}} - \Delta_{E1}^{\text{exp}}$  has a  $3.2\sigma$  tension with being zero.

It has been suggested that the tension of  $\Delta_{E1}^{\text{exp}}$  and theory [16] with  $\Delta_{M1}^{\text{exp}}$  could, if it persists, indicate a hint at new physics [52,53]. For example, in a multiple-Higgs extension to the standard model, one would speculate that the  $\eta_b^{\text{exp}}$  seen in experiments is actually an admixture of the true  $\eta_b$  and a charge parity (*CP*)-odd Higgs boson with mass ranging from 9.4–10.5 GeV. A relatively light *CP*-odd Higgs scalar can appear in nonminimal supersymmetric extensions of the standard model, such as the next-to-minimal supersymmetric standard model [53]. In such cases, the measured decay rate for  $\Upsilon(2S) \rightarrow \eta_b(1S)\gamma$  would likely differ from the standard model prediction. As stated above, this decay is sensitive in much the same way as the hyperfine splitting. To observe a similar tension between theory and experiment here as that existing between  $\Delta_{E1}^{\text{exp}}$  and  $\Delta_{M1}^{\text{exp}}$  would require a 5% uncertainty on the form factor ( $\sim 10\%$  on the decay rate). The error on the lattice form factor could be reduced to  $\sim 8\%$  (as discussed in Sec. V D) if more precise experimental results became available. Any hint of new physics arising from a deviation between the experimental  $\Upsilon(2S) \rightarrow \eta_b(1S)\gamma$  decay rate and theory could then be explored more concretely. Additionally, the  $\eta_b(2S) \rightarrow \Upsilon(1S)\gamma$  decay is an alternative approach to studying such effects and a study of this decay rate is already underway.

E1 radiative decays are more easily computed than hindered M1 decays, and so the E1 decay rates  $h_b(1P) \rightarrow \eta_b(1S)\gamma$  and  $h_b(2P) \rightarrow \eta_b(1S)\gamma$  could be calculated within this NRQCD framework. Additionally, E1 currents can be

readily renormalized nonperturbatively. Combined with the experimental branching fraction of these decays [51], this could give a prediction of the total width of the  $h_b(1P)$  and  $h_b(2P)$ .

## ACKNOWLEDGMENTS

We thank Christopher Thomas for the many insightful discussions on the finite-volume lattice in flight spectrum. We are also grateful to the MILC collaboration for the use of their gauge configurations. This work was funded by The Science and Technology Facilities Council. The results described here were obtained using the Darwin Supercomputer of the University of Cambridge High Performance Computing Service as part of the DiRAC facility jointly funded by STFC, the Large Facilities Capital Fund of The Department for Business, Innovation & Skills and the Universities of Cambridge and Glasgow.

## APPENDIX A: CLASSIFICATION OF PARTICLE STATES

Theoretically, particle states living in the Hilbert space are classified by unitary irreps of the symmetry group of a theory. We need to consider two symmetry groups here: the Lorentz group and the continuous rotational group in three dimensions (the symmetry group of NRQCD). The standard procedure to build infinite dimensional unitary irreps of these groups is via the method of induced representations, where one considers finite dimensional unitary irreps of the little group and then uses these to build unitary irreps of the full group.

The Poincaré group is the symmetry group of a relativistic quantum field theory, and is given by the semidirect product of the Lorentz group and four translations. For massive irreps of the Poincaré group, the little group is  $SO(3)$ <sup>7</sup> [54]. Thus in a Lorentz-invariant theory, massive irreps are defined as  $|p^2; J, M\rangle$ . Note that for quarkonia these states are eigenvectors of the charge-conjugation operator and parity is also a conserved quantum number,<sup>8</sup> giving the standard  $|p^2; J^{PC}, M\rangle$  decomposition. This description classifies experimental states seen to date [7].

In a continuum theory that is only rotationally invariant, the analogue of the Poincaré group is the semidirect product of the rotational group  $SO(3)$  with the three translations. For a rotationally invariant theory with zero momentum, the little group is also  $SO(3)$  and the states are classified as  $|\mathbf{p}^2; J, M\rangle$ . Thus states in a rotationally invariant theory at rest overlap with those in a Lorentz-invariant theory at rest, where again, parity and charge

<sup>7</sup>At nonzero momentum we can perform a Lorentz boost back to the rest frame, ensuring the little group is the same for zero and nonzero three-momentum.

<sup>8</sup>At nonzero momentum, these states are not eigenvectors of the parity operator, but are eigenstates of the  $\hat{\Pi}$  operator defined in the text, which conserves parity.

conjugation are good quantum numbers in similar situations. Given that at nonzero momentum in a rotationally invariant theory we cannot perform a Lorentzian boost to the rest frame, the little group at nonzero momentum is now different to the zero-momentum little group. The little group is now  $SO(2)$ <sup>9</sup> [54]. In this case, the unitary irreps are classified by  $|\mathbf{p}^2; \lambda\rangle$ , where  $\lambda$  is an eigenvalue of the helicity operator  $\hat{\lambda} = \hat{\mathbf{p}} \cdot \hat{\mathbf{J}}/E$ . The helicity  $\lambda = \lambda_0$  will get contributions from all  $J$  with  $\lambda_0 \leq J$ . This can have important consequences for the extracted energy spectrum in NRQCD, c.f., Figs. 1 and 2, and is fundamentally different from the Lorentzian theory.

At zero momentum, the operators  $i\gamma^5$  and  $\gamma^i$  that we use in this calculation overlap onto  $0^{++}$  and  $1^{--}$  states in a rotationally invariant continuum theory [26]. We now need to find which helicity eigenstates these operators overlap with at nonzero momentum. The authors of [26] illustrate how to construct helicity operators via

$$\mathbb{O}^{J,\lambda}(\mathbf{p}) = \sum_M \mathcal{D}_{M\lambda}^{J*}(R) \mathcal{O}^{J,M}(\mathbf{p}), \quad (\text{A1})$$

where  $\mathcal{D}_{M\lambda}^J(R)$  is a Wigner- $\mathcal{D}$  matrix,  $R$  is the active transformation which rotates  $(0, 0, |\mathbf{p}|)$  to  $\mathbf{p}$ ,  $\mathbb{O}^{J,\lambda}(\mathbf{p})$  is a helicity operator with helicity  $\lambda$  in an infinite-volume continuum, e.g.,

$$\langle 0 | \mathbb{O}^{J,\lambda}(\mathbf{p}) | \mathbf{p}; J', \lambda' \rangle = Z^{[J, J', \lambda]} \delta_{\lambda\lambda'} \quad (\text{A2})$$

and we refer the reader to Ref. [26] for further details. For quarkonium, the possible values of  $\lambda = \{0^+, 0^-, |1|, |2|, \dots\}$ , where the  $+/-$  on the  $\lambda = 0$  represent the  $\hat{\Pi}$  symmetry with eigenvalue  $\tilde{\eta} \equiv P(-1)^J$  [26]. Using the fact that the Wigner- $\mathcal{D}$  matrices with  $J = 0$  are  $\delta_{\lambda M}$ , the  $\mathcal{O}^{\gamma^5}$ ,  $\mathcal{O}^{\gamma^i}$  bilinear operators which we use in this calculation give rise to the helicity operators at nonzero momentum

$$\begin{aligned} \mathbb{O}^{J=0,\lambda=0^-}(\mathbf{p}) &= \mathcal{O}^{\gamma^5}(\mathbf{p}) \\ \mathbb{O}^{J=1,\lambda=0^+}(\mathbf{p}) &= \sum_M \mathcal{D}_{M\lambda=0}^{J=1*}(R) \mathcal{O}^{\gamma^M}(\mathbf{p}) \\ \mathbb{O}^{J=1,\lambda=|1|}(\mathbf{p}) &= \sum_M \mathcal{D}_{M\lambda=|1|}^{J=1*}(R) \mathcal{O}^{\gamma^M}(\mathbf{p}). \end{aligned} \quad (\text{A3})$$

As can be seen,  $\mathcal{O}^{\gamma^5}(\mathbf{p})$  is a helicity operator which creates a  $\lambda = 0^-$  state, but  $\mathcal{O}^{\gamma^i}(\mathbf{p})$  creates an admixture of  $\lambda = 0^+, |1|$  states.

The question now is the following: how do we identify which  $J^{PC}$  contributes to each  $\lambda$ , and how do we parametrize the amplitudes? By noticing that the helicity  $\hat{\lambda} = J_z$  when the momentum is projected onto the  $z$  axis, all

states with  $J \geq \lambda$  will have a  $J_z$  large enough to give a contribution to this helicity state (see Table III).

We also want to know how to quantify the amplitudes. In a rotationally invariant theory, the invariant quantities are  $\delta_{ij}$  and  $\varepsilon_{ijk}$ . For a  $J^P$  state, we also have the momentum  $p_j^i$  and the symmetric polarization tensor  $\varepsilon^{i_1 \dots i_J}$ . We can use these to parametrize the amplitudes relevant for a rotationally invariant theory. For the operator  $\mathcal{O}^{\gamma^i}$ , Table XI in [26] has the possible decompositions and we reproduce the parametrizations for the  $\mathcal{O}^{\gamma^5}$  operator which are important for our calculation

$$\begin{aligned} \langle 0 | \mathcal{O}^{\gamma^5}(\mathbf{p}) | n 0^{++}(p) \rangle &= Z_n \\ \langle 0 | \mathcal{O}^{\gamma^5}(\mathbf{p}) | n 1^{++}(\epsilon, p) \rangle &= Z_n' \epsilon_i p_i / m_{n1^{++}} \\ \langle 0 | \mathcal{O}^{\gamma^5}(\mathbf{p}) | n 2^{++}(\epsilon, p) \rangle &= Z_n' \epsilon_{ii} + Z_n'' \epsilon_{ij} p_i p_j / m_{n2^{++}}^2, \end{aligned} \quad (\text{A4})$$

where  $n$  is the radial label. Using the overlap for the  $1^{++}$  from (A4) to parametrize the continuum two-point correlator with nonzero momentum, one finds that the amplitudes from our fit with local smearing should be suppressed by  $|\mathbf{p}|/m_{1^{++}}$  relative to states which overlap with the operator at zero momentum. For the momentum that we use in our calculation, this factor is around 7%, and we observe that in our correlator data, the amplitudes for the states which do not overlap at zero momentum (and for which we get a signal) such as the  $1^{++}$ , are suppressed by this factor while the other amplitudes are  $\mathcal{O}(1)$ . We observe that as the momentum increases, so does the value of the amplitude at fixed lattice spacing.

Additionally, the symmetry group giving rise to the invariants which classify states, e.g., the little group, is broken by a finite-volume lattice to a reduced symmetry group [55]. At zero momentum with a cubic lattice, this reduced symmetry group for quarkonia is the octahedral group,  $O_h$ . States are now classified in terms of irreps of  $O_h$ , denoted  $\Lambda^{PC}$ , where [56] shows how to subduce operators with continuum spin  $J^{PC}$  to operators with definite  $\Lambda^{PC}$  on the lattice. As mentioned above, in an infinite-volume continuum theory, the  $\mathcal{O}^{\gamma^5}$  ( $\mathcal{O}^{\gamma^i}$ ) operator overlaps only with  $J^{PC} = 0^{++}(1^{--})$  at rest, but this operator falls into the  $A_1^{++}(T_1^{--})$  irrep of  $O_h$  on the lattice, where mixing with the  $J^{PC} = 4^{++}(3^{--})$  state (and higher spins) is possible. However we do not see this mixing: rotational symmetry breaking is found to be weakly broken with a fine lattice and with a rotationally invariant smearing for a particular lattice setup [56], where the spectrum and overlaps were compatible with an effective restoration of rotational symmetry. For this reason, we choose to use a rotationally invariant smearing and an isotropic lattice, and to have discretization improvements in our action. Secondly, the masses of the additional states are too large to be seen in the first few energy levels which we are interested in. As such, they will only potentially contribute

<sup>9</sup>The construction of the irreps for a rotationally invariant theory at nonzero momentum is similar to a massless representation in a Lorentz-invariant theory.

as additional discretization effects in the lowest energy modes. Indeed, studies of the spectrum from NRQCD by the HPQCD collaboration indicate this to be the case (see Appendix C of [13]).

For the nonzero momentum case, the reduced little group actually depends on the type of momenta. This is due to the fact that a general integer-valued momentum on the lattice cannot be rotated into the  $z$  axis like in an infinite-volume continuum,<sup>10</sup> e.g. there is no octahedral transformation which rotates (0,1,1) to the  $z$  axis. We use an isotropic momentum (rather than an on-axis momentum) as it has been shown to break rotational invariance less and lead to smaller discretization effects [13]. For our isotropic momentum, the reduced little group is  $\text{Dic}_3$  [26]. The operator  $\mathcal{O}^s$  ( $\mathcal{O}^i$ ) falls into the  $A_2$  ( $A_1$  and  $E_2$ ) irrep of  $\text{Dic}_3$ , where mixing with  $\lambda = 3$  (3 and 2) states is possible. For  $\mathcal{O}^s$ , this gives rise to potential mixing from  $3^{\pm+}$ ,  $4^{\pm+}$  states (and higher spin). As in the zero-momentum case, this mixing due to the lattice was found to be negligible with a fine lattice and a rotationally invariant smearing for a particular setup [26]. These states should be of higher energy than the first few states in our spectrum, and we see no evidence of them in our low-lying spectrum. For the  $\mathcal{O}^i$  operator, there can be mixing with  $\lambda = 2$  ( $2 \leq J$  with  $J_z = 2$  states) which is not important for our analysis.

There is an important distinction to be understood from using a rotationally invariant formalism for the quark versus a Lorentz-invariant one. If each of these formalisms is discretized, then at fixed nonzero momentum, the discretized version of the Lorentz-invariant theory might be broken to a rotationally invariant theory, e.g., by using an anisotropic lattice spacing in the time direction. As such, as the infinite-volume continuum limit is taken, any overlap onto  $J^{PC}$  as a result of helicity eigenstates (such as the  $1^{++}$  from the  $\mathcal{O}^s$  operator) would disappear [57]. However, in a rotationally invariant theory like NRQCD, as the lattice spacing is taken to zero, these overlaps are still present as they are an infinite-volume continuum effect. This is why we find a similar signal across all lattice spacings for these states in NRQCD.

## APPENDIX B: TWISTED CORRELATORS WITH DERIVATIVE OPERATORS

For clarity, we describe the construction of the twisted correlators with derivative operators in this section. To gain access to arbitrary momenta on the lattice, one can define a quark field [22,23] that satisfies  $\theta\text{BC}$  via  $\tilde{\psi}^\theta(x + e_i L) = e^{i2\pi\chi_i} \tilde{\psi}^\theta(x)$ , where  $\theta_i = 2\pi\chi_i/L$ . Now the

<sup>10</sup>With twisted boundary conditions, the momenta are still discretized but just shifted by an arbitrary value. As such, the little group of momentum with a twist is the same as the little group of momentum without a twist.

available momentum space is  $\tilde{\Lambda} = \{\mathbf{k} = \mathbf{p} + \boldsymbol{\theta} | k_i = 2\pi(n_i + \chi_i)/L, \text{ where } n_i \in \mathbb{Z}\}$ . Notice that the available momentum space has an arbitrary shifted value  $\boldsymbol{\theta}$  that we can choose to give the physical point  $q^2 = 0$ . One now builds interpolating operators from these  $\theta\text{BC}$  fields as  $O(x; \boldsymbol{\theta}_2 \boldsymbol{\theta}_1) = \tilde{\psi}^{\boldsymbol{\theta}_2}(x) \Gamma \tilde{\psi}^{\boldsymbol{\theta}_1}(x)$ , which gives rise to the two-point correlator

$$\begin{aligned} C_{2\text{pt}}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2 + \mathbf{p}, t) &= \sum_{\mathbf{x}} e^{-i(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2 + \mathbf{p}) \cdot \mathbf{x}} \\ &\times \text{Tr}[(\Gamma_i \tilde{S}^{\boldsymbol{\theta}_2}(\mathbf{0}, 0 | \mathbf{x}, t)) (\Gamma_f \tilde{S}^{\boldsymbol{\theta}_1}(\mathbf{x}, t | \mathbf{0}, 0))], \end{aligned} \quad (\text{B1})$$

where  $\tilde{S}^\theta(\mathbf{0}, 0 | \mathbf{x}, t)$  is a quark propagator found by inverting the Dirac matrix,  $\tilde{D}^\theta(x, y)$ , defined via  $S[\tilde{\psi}^\theta] = \sum_{x,y} \tilde{\psi}^\theta(x) \tilde{D}^\theta(x, y) \tilde{\psi}^\theta(y)$ . As a consequence of  $\tilde{\psi}^\theta$  satisfying  $\theta\text{BC}$ , the Dirac matrix  $\tilde{D}^\theta(x, y)$  also satisfies the same boundary conditions. This is an inconvenience as typical inverters are built with PBC. However, it is possible to use a trick in order to use the PBC invertors yet still get access to the  $\theta\text{BC}$  correlator data in (B1).

To do this, one notices that a second quark field, defined via the scaling  $\psi^\theta(x) = e^{-2\pi i \boldsymbol{\theta} \cdot \mathbf{x} / L} \tilde{\psi}^\theta(x)$ , satisfies PBC yet still includes information on the twist. Since

$$\tilde{S}^\theta(x|y) = e^{i\boldsymbol{\theta} \cdot (\mathbf{x} - \mathbf{y})} S^\theta(x|y) \quad (\text{B2})$$

$S^\theta(x|y)$  is a quark propagator found by inverting the Dirac matrix,  $D^\theta(x, y)$ , where  $D^\theta(x, y) = e^{-i\boldsymbol{\theta} \cdot \mathbf{x}} \tilde{D}^\theta(x, y) e^{i\boldsymbol{\theta} \cdot \mathbf{y}}$ .  $D^\theta(x|y)$  satisfies PBC by construction and the two exponentials only alter the derivative in the Dirac action and can be implemented by scaling the gluonic fields (before inverting) as  $U_\mu(x) \rightarrow U_\mu^\theta(x) = e^{i2\pi/L\theta_\mu} U_\mu(x)$  with  $\theta_\mu = (0, \boldsymbol{\theta})$  [22].

The final step is to rewrite the twisted correlator in (B1) in terms of the propagator we actually compute using (B2)

$$\begin{aligned} C_{2\text{pt}}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2 + \mathbf{p}, t) &= \sum_{\mathbf{x}} e^{-i(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2 + \mathbf{p}) \cdot \mathbf{x}} \\ &\times \text{Tr}[(\Gamma_i e^{-i\boldsymbol{\theta}_2 \cdot \mathbf{x}} S^{\boldsymbol{\theta}_2}(\mathbf{0}, 0 | \mathbf{x}, t)) ((\Gamma_f e^{i\boldsymbol{\theta}_1 \cdot \mathbf{x}} S^{\boldsymbol{\theta}_1}(\mathbf{x}, t | \mathbf{0}, 0))]. \end{aligned} \quad (\text{B3})$$

If  $\Gamma = \nabla$ , then

$$\begin{aligned} C_{2\text{pt}}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2 + \mathbf{p}, t) &= \sum_{\mathbf{x}} e^{-i\mathbf{p} \cdot \mathbf{x}} \text{Tr}[(e^{i\boldsymbol{\theta}_2 \cdot \mathbf{x}} \nabla_k e^{-i\boldsymbol{\theta}_2 \cdot \mathbf{x}} S^{\boldsymbol{\theta}_2}(\mathbf{0}, 0 | \mathbf{x}, t)) \\ &\times (e^{-i\boldsymbol{\theta}_1 \cdot \mathbf{x}} \nabla_k e^{i\boldsymbol{\theta}_1 \cdot \mathbf{x}} S^{\boldsymbol{\theta}_1}(\mathbf{x}, t | \mathbf{0}, 0))]. \end{aligned} \quad (\text{B4})$$

This can be implemented in the same way as the twist in the Dirac inverter, by using  $U_\mu^\theta(x)$  in the construction

of the covariant derivative operator. This “changing the derivatives” issue does not occur in our two-point correlators, but does occur in the (more complicated) three-point correlators with currents  $J_{W1}$ ,  $J_S$ ,  $J_{S1}$  from (16). To give an explicit example of the three-point correlator using the current  $J_{W1}$ , by keeping the initial state at rest, and twisting only one propagator in the final state with  $\theta_f$ , we have

$$\begin{aligned} C_{3\text{pt}}^{nm}(\mathbf{p}_f^\theta = \mathbf{p}_f + \boldsymbol{\theta}_f, \mathbf{q}^\theta = \mathbf{q} - \boldsymbol{\theta}_f; t, T) \\ = -i \sum_{\mathbf{x}, \mathbf{y}} e^{-i\mathbf{p}_f \cdot \mathbf{x}} \text{Tr} \left[ S^{\theta_f}(\mathbf{x}, T | \mathbf{y}, t) \right. \\ \times \left( \frac{e^{-i\theta_f \cdot \mathbf{y}}}{8m_b^3} \{ \mathbf{D}^2, (\boldsymbol{\sigma} \times \mathbf{q}^{\theta_f})^n e^{-i(\mathbf{q} - \boldsymbol{\theta}_f) \cdot \mathbf{y}} \} S(\mathbf{y}, t | \mathbf{0}, 0) \right) \\ \left. \times \sigma^m S(\mathbf{0}, 0 | \mathbf{x}, T) \right], \end{aligned} \quad (\text{B5})$$

where we can clearly see that  $\mathbf{D}^2$  does not commute with  $e^{-i\theta_f \cdot \mathbf{y}}$ , but not all derivatives are twisted due to the commutation. Since there are no derivatives in the  $J_F$  current, the  $\boldsymbol{\theta}_f$  terms cancel and this issue is avoided. Smearing the twisted fields leads to a similar issue as presented above with the derivative, and so we do not smear the twisted fields. Analogous complications arise when using point-split operators with twisted momentum in staggered quark formalisms [58]. If done correctly, and any smearings are applied appropriately, the correlator data from using  $\theta\text{BC}$  and  $\text{PBC}$  should agree on a configuration basis to machine precision (if the total momentum is identical for all states).

### APPENDIX C: ERROR ANALYSIS USING A SIMPLE POTENTIAL MODEL

First, we want to find the sensitivity of the matrix element to  $c_4$  using a potential from the exchange of a single gluon between two vertices involving the chromomagnetic operator [20]. We find (assuming the wave functions at the origin for the  $\eta_b$  and  $\Upsilon$  are the same)

$$\begin{aligned} V_{nm}^{\eta_b} &= -\frac{6c_4^2 g^2}{9m_b^2} \psi_n^*(0) \psi_m(0) \\ V_{nm}^\Upsilon &= \frac{2c_4^2 g^2}{9m_b^2} \psi_n^*(0) \psi_m(0). \end{aligned} \quad (\text{C1})$$

Putting this back into (22) with the  $J_F$  current yields

$$\begin{aligned} {}^{(1)}\langle \eta_b(1S) | J_F | \Upsilon(2S) \rangle^{(1)} \\ = {}^{(0)}\langle \eta_b(1S) | J_F | \Upsilon(2S) \rangle^{(0)} \\ + \frac{c_4^2 g^2}{9m_b^2} \left( \sum_{m \neq 1} \frac{6\psi_1^*(0) \psi_m(0)}{E_{m1}} {}^{(0)}\langle \eta_b(mS) | J_F | \Upsilon(2S) \rangle^{(0)} \right. \\ \left. - \sum_{n \neq 2} \frac{2\psi_n^*(0) \psi_2(0)}{E_{n2}} {}^{(0)}\langle \eta_b(1S) | J_F | \Upsilon(nS) \rangle^{(0)} \right). \end{aligned} \quad (\text{C2})$$

In getting to (C2) we have used the fact that  $E_{nm}^\Upsilon = E_{nm}^{\eta_b}$  as the unperturbed Hamiltonian has no spin terms. We have neglected the  $\Upsilon(pS) \rightarrow \eta_b(1S)$  transitions for  $p \geq 2$  in the sum due to the fact that the radial overlap, (20), is suppressed by at least  $\mathcal{O}(v^2)$ . In fact, they will be suppressed more due to the radial difference getting larger and the wave function at the origin getting smaller for higher radial excitations. Equation (23) can be found straightforwardly by factoring the spin part of the matrix element from the radial part, i.e., using (20).

If we now consider a potential from the exchange of a single gluon involving the Darwin term at one of the vertices, we find [20]

$$V_{nm}^{\eta_b} = V_{nm}^\Upsilon = -\frac{c_2 g^2}{3m_b^2} \psi_n^*(0) \psi_m(0). \quad (\text{C3})$$

Then substituting this back into (22) we find

$$\begin{aligned} {}^{(1)}\langle \eta_b(1S) | J_F | \Upsilon(2S) \rangle^{(1)} \\ = {}^{(0)}\langle \eta_b(1S) | J_F | \Upsilon(2S) \rangle^{(0)} \\ - \frac{c_2 g^2}{3m_b^2} \left( \sum_{m \neq 1} \frac{\psi_1^*(0) \psi_m(0)}{E_{m1}} {}^{(0)}\langle \eta_b(mS) | J_F | \Upsilon(2S) \rangle^{(0)} \right. \\ \left. + \sum_{n \neq 2} \frac{\psi_n^*(0) \psi_2(0)}{E_{n2}} {}^{(0)}\langle \eta_b(1S) | J_F | \Upsilon(nS) \rangle^{(0)} \right) \quad (\text{C4}) \\ = {}^{(0)}\langle \eta_b(1S) | J_F | \Upsilon(2S) \rangle^{(0)} \\ - \frac{c_2 g^2}{3m_b^2 E_{21}} \left( \psi_1^*(0) \psi_2(0) {}^{(0)}\langle \eta_b(2S) | J_F | \Upsilon(2S) \rangle^{(0)} \right. \\ \left. - {}^{(0)}\langle \eta_b(1S) | J_F | \Upsilon(nS) \rangle^{(0)} + \mathcal{O}(v^2) \right). \end{aligned} \quad (\text{C5})$$

Using (20), we see the leading order terms in the second piece of (C5) cancel and we are left with  $\mathcal{O}(\alpha_s v^2)$  corrections to the unperturbed matrix element.

The four-quark potential is (assuming the wave functions at the origin of the two states are the same) [20]

$$\begin{aligned} V_{nm}^{\eta_b} &= \frac{9d_1 \alpha_s^2}{2} \frac{4}{3m_b^2} \psi_n^*(0) \psi_m(0) \\ V_{nm}^\Upsilon &= \frac{9d_2 \alpha_s^2}{2} \frac{4}{3m_b^2} \psi_n^*(0) \psi_m(0). \end{aligned} \quad (\text{C6})$$

Putting this into (22) and performing an identical analysis as done above gives

$$\begin{aligned}
{}^{(1)}\langle\eta_b(1S)|J_F|\Upsilon(2S)\rangle^{(1)} &= {}^{(0)}\langle\eta_b(1S)|J_F|\Upsilon(2S)\rangle^{(0)} - \frac{9d_1\alpha_s^2}{2} \frac{4}{3m_b^2} \sum_{m\neq 1} \frac{\psi_1^*(0)\psi_m(0)}{E_{m1}} {}^{(0)}\langle\eta_b(mS)|J_F|\Upsilon(2S)\rangle^{(0)} \\
&\quad - \frac{9d_2\alpha_s^2}{2} \frac{4}{3m_b^2} \sum_{n\neq 2} \frac{\psi_n^*(0)\psi_2(0)}{E_{n2}} {}^{(0)}\langle\eta_b(1S)|J_F|\Upsilon(nS)\rangle^{(0)} \\
&= {}^{(0)}\langle\eta_b(1S)|J_F|\Upsilon(2S)\rangle^{(0)} + \frac{9}{2} \frac{4}{3m_b^2} \frac{\psi_1^*(0)\psi_2(0)}{E_{21}} \mathcal{S}_{if}(d_2\alpha_s^2 - d_1\alpha_s^2 + \mathcal{O}((2d_2\alpha_s^2 - d_1\alpha_s^2)v^2)).
\end{aligned} \tag{C7}$$

The error in the last line was introduced by expanding out the radial overlap (20) and noting that the two matrix elements do not have to be identical to first order in  $|q_\gamma|^2$ . Even if we did include the four-fermion operators in the calculation, since only the combination  $d_1 - d_2$  is currently known perturbatively, and not  $d_1$  and  $d_2$  individually, we would still need to introduce the  $\mathcal{O}(v^2)$  error in our calculation.

- 
- [1] S. Godfrey and J. L. Rosner, *Phys. Rev. D* **64**, 074011 (2001).
- [2] B. Aubert *et al.* (BABAR Collaboration), *Phys. Rev. Lett.* **103**, 161801 (2009).
- [3] N. Brambilla, Y. Jia, and A. Vairo, *Phys. Rev. D* **73**, 054005 (2006).
- [4] A. Pineda and J. Segovia, *Phys. Rev. D* **87**, 074024 (2013).
- [5] R. Lewis and R. M. Woloshyn, *Phys. Rev. D* **84**, 094501 (2011).
- [6] R. Lewis and R. M. Woloshyn, *Phys. Rev. D* **86**, 057501 (2012).
- [7] K. Olive *et al.* (Particle Data Group), *Chin. Phys. C* **38**, 090001 (2014).
- [8] N. Brambilla *et al.*, *Eur. Phys. J. C* **71**, 1534 (2011).
- [9] J. J. Dudek, R. G. Edwards, and D. G. Richards, *Phys. Rev. D* **73**, 074507 (2006).
- [10] A. Bazavov *et al.* (MILC Collaboration), *Phys. Rev. D* **82**, 074501 (2010).
- [11] A. Hart, G. M. von Hippel, and R. R. Horgan (HPQCD Collaboration), *Phys. Rev. D* **79**, 074008 (2009).
- [12] E. Follana, Q. Mason, C. Davies, K. Hornbostel, G. P. Lepage, J. Shigemitsu, H. Trotter, and K. Wong (HPQCD and UKQCD Collaborations), *Phys. Rev. D* **75**, 054502 (2007).
- [13] R. J. Dowdall *et al.* (HPQCD Collaboration), *Phys. Rev. D* **85**, 054509 (2012).
- [14] G. P. Lepage, L. Magnea, C. Nakhleh, U. Magnea, and K. Hornbostel, *Phys. Rev. D* **46**, 4052 (1992).
- [15] J. O. Daldrop, C. T. H. Davies, and R. J. Dowdall (HPQCD Collaboration), *Phys. Rev. Lett.* **108**, 102003 (2012).
- [16] R. J. Dowdall, C. T. H. Davies, T. Hammant, and R. R. Horgan (HPQCD Collaboration), *Phys. Rev. D* **92**, 039904 (2015).
- [17] R. J. Dowdall, C. T. H. Davies, T. C. Hammant, and R. R. Horgan (HPQCD Collaboration), *Phys. Rev. D* **86**, 094510 (2012).
- [18] R. J. Dowdall, C. T. H. Davies, R. R. Horgan, C. J. Monahan, and J. Shigemitsu (HPQCD Collaboration), *Phys. Rev. Lett.* **110**, 222003 (2013).
- [19] B. Colquhoun, R. J. Dowdall, C. T. H. Davies, K. Hornbostel, and G. P. Lepage (HPQCD Collaboration), *Phys. Rev. D* **91**, 074514 (2015).
- [20] T. C. Hammant, A. G. Hart, G. M. von Hippel, R. R. Horgan, and C. J. Monahan (HPQCD Collaboration), *Phys. Rev. D* **88**, 014505 (2013).
- [21] S. Meinel, *Phys. Rev. D* **82**, 114502 (2010).
- [22] G. de Divitiis, R. Petronzio, and N. Tantalo, *Phys. Lett. B* **595**, 408 (2004).
- [23] D. Guadagnoli, F. Mescia, and S. Simula, *Phys. Rev. D* **73**, 114504 (2006).
- [24] G. P. Lepage, <https://github.com/gplepage/corrfitter>.
- [25] G. 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).
- [26] C. E. Thomas, R. G. Edwards, and J. J. Dudek, *Phys. Rev. D* **85**, 014507 (2012).
- [27] F. D. R. Bonnet, R. G. Edwards, G. T. Fleming, R. Lewis, and D. G. Richards (Lattice Hadron Physics), *Phys. Rev. D* **72**, 054506 (2005).
- [28] C. M. Bouchard, G. P. Lepage, C. Monahan, H. Na, and J. Shigemitsu (HPQCD Collaboration), *Phys. Rev. D* **90**, 054506 (2014).
- [29] C. McNeile, C. T. H. Davies, E. Follana, K. Hornbostel, and G. P. Lepage (HPQCD Collaboration), *Phys. Rev. D* **82**, 034512 (2010).
- [30] K. Hornbostel, G. P. Lepage, C. T. H. Davies, R. J. Dowdall, H. Na, and J. Shigemitsu (HPQCD Collaboration), *Phys. Rev. D* **85**, 031504 (2012).
- [31] G. Feinberg and J. Sucher, *Phys. Rev. Lett.* **35**, 1740 (1975).
- [32] J. Sucher, *Rep. Prog. Phys.* **41**, 1781 (1978).
- [33] J. S. Kang and J. Sucher, *Phys. Rev. D* **18**, 2698 (1978).
- [34] T. Kinoshita and M. Nio, *Phys. Rev. D* **53**, 4909 (1996).
- [35] A. Hart, G. von Hippel, R. Horgan, and E. Muller, *Comput. Phys. Commun.* **180**, 2698 (2009).
- [36] A. Hart, G. von Hippel, R. Horgan, and L. Storoni, *J. Comput. Phys.* **209**, 340 (2005).
- [37] M. A. Nobes, H. D. Trotter, G. P. Lepage, and Q. Mason, *Nucl. Phys. B, Proc. Suppl.* **106–107**, 838 (2002).

- [38] N. Brambilla *et al.* (Quarkonium Working Group), [arXiv: hep-ph/0412158](#).
- [39] R. R. Horgan (private communication).
- [40] B. Colquhoun, C. T. H. Davies, J. Kettle, J. Koponen, A. T. Lytle, R. J. Dowdall, and G. P. Lepage (HPQCD Collaboration), *Phys. Rev. D* **91**, 114509 (2015).
- [41] A. Bazavov *et al.* (Fermilab Lattice and MILC Collaborations), *Phys. Rev. D* **90**, 074509 (2014).
- [42] V. Zambetakis and N. Byers, *Phys. Rev. D* **28**, 2908 (1983).
- [43] H. Grotch, D. A. Owen, and K. J. Sebastian, *Phys. Rev. D* **30**, 1924 (1984).
- [44] S. Godfrey and N. Isgur, *Phys. Rev. D* **32**, 189 (1985).
- [45] X. Zhang, K. J. Sebastian, and H. Grotch, *Phys. Rev. D* **44**, 1606 (1991).
- [46] T. Lahde, C. Nyfalt, and D. Riska, *Nucl. Phys.* **A645**, 587 (1999).
- [47] J. Beringer *et al.*, *Phys. Rev. D* **86**, 010001 (2012).
- [48] B. Aubert *et al.* (BABAR Collaboration), *Phys. Rev. Lett.* **101**, 071801 (2008).
- [49] G. Bonvicini *et al.* (CLEO Collaboration), *Phys. Rev. D* **81**, 031104 (2010).
- [50] K. Nakamura *et al.* (Particle Data Group), *J. Phys. G* **37**, 075021 (2010).
- [51] R. Mizuk *et al.* (Belle Collaboration), *Phys. Rev. Lett.* **109**, 232002 (2012).
- [52] D. Becirevic and F. Sanfilippo, *J. High Energy Phys.* **01** (2013) 028.
- [53] F. Domingo, U. Ellwanger, and M.-A. Sanchis-Lozano, *Phys. Rev. Lett.* **103**, 111802 (2009).
- [54] S. Weinberg, *The Quantum Theory of Fields* (Cambridge University Press, Cambridge, 1995), Vol. 1.
- [55] D. C. Moore and G. T. Fleming, *Phys. Rev. D* **73**, 014504 (2006).
- [56] J. J. Dudek, R. G. Edwards, M. J. Peardon, D. G. Richards, and C. E. Thomas, *Phys. Rev. D* **82**, 034508 (2010).
- [57] Z. Davoudi and M. J. Savage, *Phys. Rev. D* **86**, 054505 (2012).
- [58] G. C. Donald, C. T. H. Davies, J. Koponen, and G. P. Lepage (HPQCD Collaboration), *Phys. Rev. D* **90**, 074506 (2014).