18. Lattice Quantum Chromodynamics

Updated September 2015 by S. Hashimoto (KEK), J. Laiho (Syracuse University) and S.R. Sharpe (University of Washington).

Many physical processes considered in the Review of Particle Properties (RPP) involve hadrons. The properties of hadrons—which are composed of quarks and gluons—are governed primarily by Quantum Chromodynamics (QCD) (with small corrections from Quantum Electrodynamics [QED]). Theoretical calculations of these properties require non-perturbative methods, and Lattice Quantum Chromodynamics (LQCD) is a tool to carry out such calculations. It has been successfully applied to many properties of hadrons. Most important for the RPP are the calculation of electroweak form factors, which are needed to extract Cabbibo-Kobayashi-Maskawa (CKM) matrix elements when combined with the corresponding experimental measurements. LQCD has also been used to determine other fundamental parameters of the standard model, in particular the strong coupling constant and quark masses.

This review describes the theoretical foundations of LQCD and sketches the methods used to calculate the quantities relevant for the RPP. It also describes the various sources of error that must be controlled in a LQCD calculation. Results for hadronic quantities are given in the corresponding dedicated reviews.

18.1. Lattice regularization of QCD

Gauge theories form the building blocks of the Standard Model. While the SU(2) and U(1) parts have weak couplings and can be studied accurately with perturbative methods, the SU(3) component—QCD—is only amenable to a perturbative treatment at high energies. The growth of the coupling constant in the infrared—the flip-side of asymptotic freedom—requires the use of non-perturbative methods to determine the low energy properties of QCD. Lattice gauge theory, proposed by K. Wilson in 1974 [1], provides such a method, for it gives a non-perturbative definition of vector-like gauge field theories like QCD. In lattice regularized QCD—commonly called lattice QCD or LQCD—Euclidean space-time is discretized, usually on a hypercubic lattice with lattice spacing $a$, with quark fields placed on sites and gauge fields on the links between sites. The lattice spacing plays the role of the ultraviolet regulator, rendering the quantum field theory finite. The continuum theory is recovered by taking the limit of vanishing lattice spacing, which can be reached by tuning the bare coupling constant to zero according to the renormalization group.

Unlike dimensional regularization, which is commonly used in continuum QCD calculations, the definition of LQCD does not rely on the perturbative expansion. Indeed, LQCD allows non-perturbative calculations by numerical evaluation of the path integral that defines the theory.

Practical LQCD calculations are limited by the availability of computational resources and the efficiency of algorithms. Because of this, LQCD results come with both statistical and systematic errors, the former arising from the use of Monte-Carlo integration, the latter, for example, from the use of non-zero values of $a$. There are also different ways
Lattice QCD

in which the QCD action can be discretized, and all must give consistent results in the continuum limit, \( a \to 0 \). It is the purpose of this review to provide an outline of the methods of LQCD, with particular focus on applications to particle physics, and an overview of the various sources of error. This should allow the reader to better understand the LQCD results that are presented in other reviews, primarily those on “Quark Masses”, “Quantum Chromodynamics”, “CKM quark-mixing matrix”, “\( V_{ud} \), \( V_{us} \), Cabibbo angle and CKM Unitarity” and “Semileptonic B-meson decays and the determination of \( V_{cb} \) and \( V_{ub} \)”. For more extensive explanations the reader should consult the available textbooks or lecture notes, the most up-to-date of which are Refs. 2–4.

18.1.1. Gauge invariance, gluon fields and the gluon action:

A key feature of the lattice formulation of QCD is that it preserves gauge invariance. This is in contrast to perturbative calculations, where gauge fixing is an essential step. The preservation of gauge invariance leads to considerable simplifications, e.g. restricting the form of operators that can mix under renormalization.

The gauge transformations of lattice quark fields are just as in the continuum: \( q(x) \rightarrow V(x)q(x) \) and \( \bar{q}(x) \rightarrow \bar{q}(x)V^\dagger(x) \), with \( V(x) \) an arbitrary element of SU(3). The only difference is that the Euclidean space-time positions \( x \) are restricted to lie on the sites of the lattice, i.e. \( x = a(n_1,n_2,n_3,n_4) \) for a hypercubic lattice, with the \( n_j \) being integers. Quark bilinears involving different lattice points can be made gauge invariant by introducing the gluon field \( U_\mu(x) \). For example, for adjacent points the bilinear is \( \bar{q}(x)U_\mu(x)q(x+a\hat{\mu}) \), with \( \hat{\mu} \) the unit vector in the \( \mu \)’th direction. (This form is used in the construction of the lattice covariant derivative.) This is illustrated in Fig. 18.1. The gluon field (or “gauge link”) is an element of the group, SU(3), in contrast to the continuum field \( A_\mu \) which takes values in the Lie algebra. The bilinear is invariant if \( U_\mu \) transforms as \( U_\mu(x) \rightarrow V(x)U_\mu(x)V^\dagger(x+a\hat{\mu}) \). The lattice gluon field is naturally associated with the link joining \( x \) and \( x+a\hat{\mu} \), and corresponds in the continuum to a Wilson line connecting these two points, \( P \exp(i \int_x^{x+a\hat{\mu}} dx_\mu A^\text{cont}_\mu(x)) \) (where \( P \) indicates a path-ordered integral, and the superscript on \( A_\mu \) indicates that it is a continuum field). The trace of a product of the \( U_\mu(x) \) around any closed loop is easily seen to be gauge invariant and is the lattice version of a Wilson loop.

The simplest possible gauge action, usually called the Wilson gauge action, is given by the product of gauge links around elementary plaquettes:

\[
S_g = \beta \sum_{x,\mu,\nu} [1 - \frac{1}{3} \text{Re} \text{Tr}[U_\mu(x)U_\nu(x+a\hat{\mu})U_\mu^\dagger(x+a\hat{\nu})U_\nu^\dagger(x)]] .
\]  

(18.1)

This is illustrated in Fig. 18.1. For small \( a \), assuming that the fields are slowly varying, one can expand the action in powers of \( a \) using \( U_\mu(x) = \exp(iaA_\mu(x)) \). Keeping only the leading non-vanishing term, and replacing the sum with an integral, one finds the continuum form,

\[
S_g \rightarrow \int d^4x \frac{1}{4g_{\text{lat}}^2} \text{Tr}[F^2_{\mu\nu}(x)] , \quad (F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + i[A_\mu, A_\nu])
\]  

(18.2)
as long as one chooses $\beta = 6/g_{\text{lat}}^2$ for the lattice coupling. In this expression, $g_{\text{lat}}$ is the bare coupling constant in the lattice scheme, which can be related (by combining continuum and lattice perturbation theory) to a more conventional coupling constant such as that in the $\overline{\text{MS}}$ scheme (see Sec. 18.3.4 below).

In practice, the lattice spacing $a$ is non-zero, leading to discretization errors. In particular, the lattice breaks Euclidean rotational invariance (which is the Euclidean version of Lorentz invariance) down to a discrete hypercubic subgroup. One wants to reduce discretization errors as much as possible. A very useful tool for understanding and then reducing discretization errors is the Symanzik effective action: the interactions of quarks and gluons with momenta low compared to the lattice cutoff ($|p| \ll 1/a$) are described by a continuum action consisting of the standard continuum terms (e.g. the gauge action given in Eq. (18.2)) augmented by higher dimensional operators suppressed by powers of $a$ [5]. For the Wilson lattice gauge action, the leading corrections come in at $O(a^2)$. They take the form $\sum_j a^2 c_j O_6^{(j)}$, with the sum running over all dimension-six operators $O_6^{(j)}$ allowed by the lattice symmetries, and $c_j$ unknown coefficients. Some of these operators violate Euclidean rotational invariance, and all of them lead to discretization errors of the form $a^2 \Lambda^2$, where $\Lambda$ is a typical momentum scale for the quantity being calculated. These errors can, however, be reduced by adding corresponding operators to the lattice action and tuning their coefficients to eliminate the
dimension-six operators in the effective action to a given order in perturbation theory or even non-perturbatively. This is the idea of the Symanzik improvement program [5]. In the case of the gauge action, one adds Wilson loops involving six gauge links (as opposed to the four links needed for the original plaquette action, Eq. (18.1)) to define the $O(a^2)$ improved (or “Symanzik”) action [6]. In practical implementations, the improvement is either at tree-level (so that residual errors are proportional to $\alpha_s a^2$, where the coupling is evaluated at a scale $\sim 1/a$), or at one loop order (errors proportional to $\alpha_s^2 a^2$). Another popular choice is motivated by studies of renormalization group (RG) flow. It has the same terms as the $O(a^2)$ improved action but with different coefficients, and is called the RG-improved or “Iwasaki” action [7].

18.1.2. Lattice fermions:

Discretizing the fermion action turns out to involve subtle issues, and the range of actions being used is more extensive than for gauge fields. Recall that the continuum fermion action is $S_f = \int d^4x \bar{q}iD\gamma q + m_q q$, where $D\gamma = \partial \gamma + iA\gamma$ is the gauge-covariant derivative. The simplest discretization replaces the derivative with a symmetric difference:

$$D\gamma q(x) \longrightarrow \frac{1}{2a} [U\gamma(x)q(x + a\hat{\mu}) - U\gamma(x - a\hat{\mu})^q(x - a\hat{\mu})]. \quad (18.3)$$

The factors of $U\gamma$ ensure that $D\gamma q(x)$ transforms under gauge transformations in the same way as $q(x)$, so that the discretized version of $\bar{q}(x)D\gamma q(x)$ is gauge invariant. The choice in Eq. (18.3) leads to the so-called naive fermion action. This, however, suffers from the fermion doubling problem—in $d$ dimensions it describes $2^d$ equivalent fermion fields in the continuum limit. The appearance of the extra “doubler” fermions is related to the deeper theoretical problem of formulating chirally symmetric fermions on the lattice. This is encapsulated by the Nielsen-Ninomiya theorem [8]: one cannot define lattice fermions having exact, continuum-like chiral symmetry without producing doublers. Naive lattice fermions do have chiral symmetry but at the cost of introducing 15 unwanted doublers (for $d = 4$).

There are a number of different strategies for dealing with the doubling problem, each with their own theoretical and computational advantages and disadvantages. Wilson fermions [1] add a term proportional to $a\bar{q}\Delta q$ to the fermion action (the “Wilson term”—in which $\Delta$ is a covariant lattice Laplacian). This gives a mass of $O(1/a)$ to the doublers, so that they decouple in the continuum limit. The Wilson term, however, violates chiral symmetry, and also introduces discretization errors linear in $a$. A commonly used variant that eliminates the $O(a)$ discretization error is the $O(a)$-improved Wilson (or “clover”) fermion [9]. In this application of Symanzik improvement, methods have been developed to remove $O(a)$ terms non-perturbatively using auxiliary simulations to tune parameters [10]. Such “non-perturbative improvement” is of great practical importance as it brings the discretization error from the fermion action down to the same level as that from the gauge action. It is used by essentially all simulations using clover fermions.

The advantages of Wilson fermions are their theoretical simplicity and relatively low computational cost. Their main disadvantage is the lack of chiral symmetry, which makes them difficult to use in cases where mixing with wrong chirality operators can occur,
particularly if this involves divergences proportional to powers of $1/a$. A related problem is the presence of potential numerical instabilities due to spurious near-zero modes of the lattice Dirac operator. Ongoing work has, however, been successful at ameliorating these problems and increasing the range of quantities for which Wilson fermions can be used (see, e.g., Refs. 11–13).

Twisted-mass fermions [14] are a variant of Wilson fermions in which two flavors are treated together with an isospin-breaking mass term (the “twisted mass” term). The main advantage of this approach is that all errors linear in $a$ are automatically removed (without the need for tuning of parameters) by a clever choice of twisted mass and operators [15]. A disadvantage is the presence of isospin breaking effects (such as a splitting between charged and neutral pion masses even when up and down quarks are degenerate), which, however, vanish as $a^2 \Lambda^2$ in the continuum limit. Strange and charm quarks can be added as a second pair, with a term added to split their masses.

Staggered fermions are a reduced version of naive fermions in which there is only a single fermion Dirac component on each lattice site, with the full Dirac structure built up from neighboring sites [16]. They have the advantages of being somewhat faster to simulate than Wilson-like fermions, of preserving some chiral symmetry, and of having discretization errors of $O(a^2)$. Their disadvantage is that they retain some of the doublers (3 for $d = 4$). The action thus describes four degenerate fermions in the continuum limit. These are usually called “tastes”, to distinguish them from physical flavors, and the corresponding SU(4) symmetry is referred to as the “taste symmetry”. The preserved chiral symmetry in this formulation has non-singlet taste. Practical applications usually introduce one staggered fermion for each physical flavor, and remove contributions from the unwanted tastes by taking the fourth-root of the fermion determinant appearing in the path integral. The validity of this “rooting” procedure is not obvious because taste symmetry is violated for non-zero lattice spacing. Theoretical arguments, supported by numerical evidence, suggest that the procedure is valid as long as one takes the continuum limit before approaching the light quark mass region [17]. Additional issues arise for the valence quarks (those appearing in quark propagators, as described in Sec. 18.2 below), where rooting is not possible, and one must remove the extra tastes by hand [18].

Just as for Wilson fermions, the staggered action can be improved, so as to reduce discretization errors. The Asqtad ($a$-squared tadpole improved) action [19] was used until recently in many large scale simulations [20]. Most recent calculations use the HISQ (highly improved staggered quark) action, introduced in Ref. 21. This removes tree-level $O(a^2)$ errors, and leads to a substantial reduction in the breaking of taste symmetry, as well as a general reduction in the size of other discretization errors. It is tuned to reduce discretization errors for both light and heavier quarks, and is being used to directly simulate charm quarks.

There is an important class of lattice fermions, “Ginsparg-Wilsons fermions”, that possess a continuum-like chiral symmetry without introducing unwanted doublers. The lattice Dirac operator $D$ for these fermions satisfies the Ginsparg-Wilson relation $D \gamma_5 + \gamma_5 D = aD \gamma_5 D$ [22]. In the continuum, the right-hand-side vanishes, leading to chiral symmetry. On the lattice, it is non-vanishing, but with a particular form (with two factors of $D$) that restricts the violations of chiral symmetry in Ward-Takahashi identities.
to short-distance terms that do not contribute to physical matrix elements [23]. In fact, one can define a modified chiral transformation on the lattice (by including dependence on the gauge fields) such that Ginsparg-Wilson fermions have an exact chiral symmetry for on-shell quantities [24]. The net result is that such fermions essentially have the same properties under chiral transformations as do continuum fermions, including the index theorem [23]. Their leading discretization errors are of $O(a^2)$.

Two types of Ginsparg-Wilson fermions are currently being used in large-scale numerical simulations. The first is Domain-wall fermions (DWF). These are defined on a five-dimensional space, in which the fifth dimension is fictitious [25]. The action is chosen so that the low-lying modes are chiral, with left- and right-handed modes localized on opposite four-dimensional surfaces. For an infinite fifth dimension, these fermions satisfy the Ginsparg-Wilson relation. In practice, the fifth dimension is kept finite, and there remains a small, controllable violation of chiral symmetry. The second type is Overlap fermions. These appeared from a completely different context and have an explicit form that exactly satisfies the Ginsparg-Wilson relation [26]. Their numerical implementation requires an approximation of the matrix sign function of a Wilson-like fermion operator, and various approaches are being used. In fact, it is possible to rewrite these approximations in terms of a five-dimensional formulation, showing that the DWF and Overlap approaches are essentially equivalent [27]. Numerically, the five-dimensional approach appears to be more computationally efficient.

The various lattice fermion formulations are often combined with the technique of link smearing. Here one couples the fermions to a smoother gauge link, defined by averaging with adjacent links in a gauge invariant manner. Several closely related implementations are being used. All reduce the coupling of fermions to the short-distance fluctuations in the gauge field, leading to an improvement in the numerical stability and speed of algorithms. One cannot perform this smearing too aggressively, however, since the smearing may distort short distance physics and enhance discretization errors.

As noted above, each fermion formulation has its own advantages and disadvantages. For instance, domain-wall and overlap fermions are theoretically preferred as they have chiral symmetry without doublers, but their computational cost is greater than for other choices. If the physics application of interest and the target precision do not require near-exact chiral symmetry, there is no strong motivation to use these expensive formulations. On the other hand, there is a class of applications (including the calculation of the $\Delta I = 1/2$ amplitude for $K \rightarrow \pi\pi$ decays and the S-parameter [28]) where chiral symmetry plays an essential role and for which the use of Ginsparg-Wilson fermions is strongly favored.

### 18.1.3. Heavy quarks on the lattice

The fermion formulations described in the previous subsection can be used straightforwardly only for quarks whose masses are small compared to the lattice cutoff, $m_q \lesssim 1/a$. This is because there are discretization errors proportional to powers of $am_q$, and if $am_q \gtrsim 1$ these errors are large and uncontrolled. Present LQCD simulations typically have cutoffs in the range of $1/a = 2 - 4 \text{ GeV}$ (corresponding to $a \approx 0.1 - 0.05 \text{ fm}$). Thus, while for the up, down and strange quarks one has $am_q \ll 1$, for heavy quarks one has $am_q \gtrsim 1$.
for bottom quarks (with $m_b \approx 4.5$ GeV) one must use alternative approaches. Charm quarks ($m_c \approx 1.5$ GeV) are an intermediate case, allowing simulations using both direct and alternative approaches.

For the charm quark, the straightforward approach is to simultaneously reduce the lattice spacing and to improve the fermion action so as to reduce the size of errors proportional to powers of $am_c$. This approach has, for example, been followed successfully using the HISQ and twisted-mass actions [21,29,30]. It is important to note, however, that reducing $a$ increases the computational cost because an increased number of lattice points are needed for the same physical volume. One cannot reduce the spatial size below $2 - 3$ fm without introducing finite volume errors. Present lattices have typical sizes of $\sim 64^3 \times 128$ (with the long direction being Euclidean time), and thus allow a lattice cutoff up to $1/a \sim 4$ GeV.

Alternative approaches for discretizing heavy quarks are motivated by effective field theories. For a bottom quark in heavy-light hadrons, one can use Heavy Quark Effective Theory (HQET) to expand about the infinite quark-mass limit. In this limit, the bottom quark is a static color source, and one can straightforwardly write the corresponding lattice action [31]. Corrections, proportional to powers of $1/m_b$, can be introduced as operator insertions, with coefficients that can be determined non-perturbatively using existing techniques [32]. This method allows the continuum limit to be taken controlling all $1/m_b$ corrections.

Another way of introducing the $1/m_b$ corrections is to include the relevant terms in the effective action. This leads to a non-relativistic QCD (NRQCD) action, in which the heavy quark is described by a two-component spinor [33]. This approach has the advantage over HQET that it can also be used for heavy-heavy systems, such as the Upsilon states. A disadvantage is that some of the parameters in this effective theory are determined perturbatively (originally at tree-level, but more recently at one-loop), which limits the precision of the final results. Although discretization effects can be controlled with good numerical precision for a range of lattice spacings, at fine enough lattice spacing the NRQCD effective theory no longer applies since power divergent terms become important, and taking the continuum limit would require fine-tuning a large number of couplings non-perturbatively.

This problem can be avoided if one uses HQET power counting to analyze and reduce discretization effects for heavy quarks while using conventional fermion actions [34]. For instance, one can tune the parameters of an improved Wilson quark action so that the leading HQET corrections to the static quark limit are correctly accounted for. As the lattice spacing becomes finer, the action smoothly goes over to that of a light Wilson quark action, where the continuum limit can be taken as usual. In principle, one can improve the action in the heavy quark regime up to arbitrarily high orders using HQET, but so far large-scale simulations have typically used clover improved Wilson quarks, where tuning the parameters of the action corresponds to including all corrections through next-to-leading order in HQET. Three different methods for tuning the parameters of the clover action are being used: the Fermilab [34], Tsukuba [35] and Columbia [36] approaches. An advantage of this HQET approach is that the $c$ and $b$ quarks can be treated on the same footing. Parameter tuning has typically been done perturbatively,
18. Lattice QCD

as in NRQCD, but recent work using the Columbia approach has used non-perturbative tuning of some of the parameters \[37,38\].

Another approach is the “ratio method” introduced in \[39\]. Here one uses quarks with masses lying at, or slightly above, the charm mass \(m_c\), which can be simulated with a relativistic action, and extrapolates to \(m_b\) incorporating the behavior predicted by HQET. The particular implementation relies on the use of ratios. As an example, consider the B meson decay constant \(f_B\). According to HQET, this scales as \(1/\sqrt{m_B}\) for \(m_B \gg \Lambda_{\text{QCD}}\), up to a logarithmic dependence that is calculable in perturbative QCD (but will be suppressed in the following). Here \(m_B\) is the B meson mass, which differs from \(m_b\) by \(\sim \Lambda_{\text{QCD}}\). One considers the ratio \(y(\lambda, m_b') = f_B'/\sqrt{m_B'} / f_B / \sqrt{m_B}\) for fictitious B mesons containing b quarks with unphysical masses \(m_b' = \lambda m_c\), all well below the physical \(m_b\), and for each the continuum limit is taken. The form of the ratio for larger values of \(m_b'\) is obtained by fitting, incorporating the constraints implied by HQET. The result for \(f_B / \sqrt{m_B}\) is then obtained as a product of \(y\)'s with \(f_D / \sqrt{m_D}\).

18.1.4. QED on the lattice:

Quarks in nature are electrically charged, and the resultant coupling to photons leads to shifts in the properties of hadrons that are generically of \(O(\alpha_{\text{EM}})\). Thus, for example, the proton mass is increased by \(\sim 1\) MeV relative to that of the neutron due to its overall charge although this effect is more than compensated for by the \(\sim 2.5\) MeV relative decrease due to the up quark being lighter than the down quark \[40\]. This example shows that once pure QCD, isospin-symmetric lattice calculations reach percent level accuracy, further improvement requires the inclusion of effects due to both electromagnetism and the up-down mass difference. This level of accuracy has in fact been obtained for various quantities, e.g. light hadron masses and decay constants (see Ref. 41), and simulations including QED in addition to QCD are now beginning.

The extension of lattice methods to include QED is straightforward, although some new subtleties arise. The essential change is that the quark must now propagate through a background field containing both gluons and photons. The gauge field \(U_\mu\) that appears in covariant derivative Eq. (18.3) is extended from an SU(3) matrix to one living in U(3):

\[ U_\mu \to U_\mu e^{iaq_e A^\text{EM}_\mu}. \]

Here \(A^\text{EM}_\mu\) is the photon field, \(e\) the electromagnetic coupling, and \(q\) the charge of the quark, e.g. \(q = 2/3\) for up and \(-1/3\) for down and strange quarks. The lattice action for the photon that is typically used is a discretized version of the continuum action Eq. (18.2), rather than the form used for the gluons, Eq. (18.1). This “non-compact” action has the advantage that it is quadratic in \(A^\text{EM}_\mu\), which simplifies the QED part of the generation of configurations.

One subtlety that arises is that Gauss’ law forbids a charged particle in a box with periodic boundary conditions. This finite volume effect can be overcome by including a uniform background charge, and this can be shown to be equivalent to removing the zero-momentum mode from the photon field. This is an example of the enhanced finite-volume effects that arise in the presence of the massless photon.

Simulations including QED have progressed over the last few years, and now a full
inclusion of QED has been achieved with almost physical quark masses [40]. Alternative approaches have also been considered: reweighting the QCD fields \textit{a posteriori} [42,43], and keeping only the linear term in an expansion in \( \alpha_{\text{EM}} \) about the QCD only case [44]. In addition, some calculations have included QED effects for the valence quarks but not the sea quarks (the “electroquenched approximation”).

18.1.5. Basic inputs for lattice calculations:

Since LQCD is nothing but a regularization of QCD, the renormalizability of QCD implies that the number of input parameters in LQCD is the same as for continuum QCD—the strong coupling constant \( \alpha_s = g^2/(4\pi) \), the quark masses for each flavor, and the CP violating phase \( \theta \). The \( \theta \) parameter is usually assumed to be zero, while the other parameters must be determined using experimental inputs.

18.1.5.1. Lattice spacing: In QCD, the coupling constant is a function of scale. With lattice regularization, this scale is the inverse lattice spacing \( 1/a \), and choosing the bare coupling constant is equivalent to fixing the lattice spacing.

In principle, \( a \) can be determined using any dimensionful quantity measured by experiments. For example, using the mass of hadron \( H \) one has \( a = (am_H)^{\text{lat}}/m_H^{\text{exp}} \). (Of course, one must first tune the quark masses to their physical values, as discussed below.) In practice, one chooses quantities that can be calculated accurately on the lattice, and that are only weakly dependent on the light quark masses. The latter property minimizes errors from extrapolating or interpolating to the physical light quark masses or from mistuning of these masses. Commonly used choices are the spin-averaged 1S-1P or 1S-2S splittings in the Upsilon system, the mass of the \( \Omega^- \) baryon, and the pion decay constant \( f_\pi \). Ultimately, all choices must give consistent results for \( a \), and that this is the case provides a highly non-trivial check of both the calculational method and of QCD.

18.1.5.2. Light quark masses:

In LQCD simulations, the up, down and strange quarks are usually referred to as the light quarks, in the sense that \( m_q < \Lambda_{\text{QCD}} \). (The standard definition of \( \Lambda_{\text{QCD}} \) is given in the “Quantum Chromodynamics” review; in this review we are using it only to indicate the approximate non-perturbative scale of QCD.) This condition is stronger than that used above to distinguish quarks with small discretization errors, \( m_q < 1/a \). Loop effects from light quarks must be included in the simulations to accurately represent QCD. At present, most simulations are done in the isospin symmetric limit \( m_u = m_d \equiv m_\ell < m_s \), and are often referred to as \( N_f = 2 + 1 \) simulations. Increasingly, simulations also include loops of charm quarks (denoted \( N_f = 2 + 1 + 1 \) simulations), although the effect of charmed sea quarks on low-energy physics is generically expected to be at the sub-percent level [45]. Precision is now reaching the point where isospin breaking effects must be included. To do so without approximation requires simulating with nondegenerate up and down quarks (leading to \( N_f = 1 + 1 + 1 \) or \( 1 + 1 + 1 + 1 \) simulations) as well as including electromagnetism (as described above). This has been done in Ref. 40. Alternatively, one can use a perturbative approach, expanding about the isospin symmetric theory and working to linear order in \( \alpha_{\text{EM}} \) and \( m_u - m_d \) [44].
18. Lattice QCD

We now describe the tuning of $m_\ell$, $m_s$ and $m_c$ to their physical values. (For brevity, we ignore isospin violation in the following discussion.) The most commonly used quantities for these tunings are, respectively, $m_\pi$, $m_K$ and $m_D$. If the scale is being set by $m_\Omega$, then one adjusts the lattice quark masses until the ratios $m_\pi/m_\Omega$, $m_K/m_\Omega$ and $m_D/m_\Omega$ take their physical values. In the past, most calculations needed to extrapolate to the physical value of $m_\ell$ (typically using forms based on chiral perturbation theory [ChPT]), while simulating directly at or near to the physical values of $m_s$ and $m_c$. Present calculations are increasingly done with physical or near physical values of $m_\ell$, requiring at most only a short extrapolation.

18.1.5.3. Heavy quark masses:

The $b$ quark is usually treated only as a valence quark, with no loop effects included. The errors introduced by this approximation can be estimated to be $\sim \alpha_s(m_b)\Lambda_{\text{QCD}}^2/m_b^2$ and are likely to be very small. In the past, the same approximation has been made for the $c$ quark, leading to errors $\sim \alpha_s(m_c)\Lambda_{\text{QCD}}^2/m_c^2$. (See [45] for a quantitative estimate of the effects of including the charm quark on some low energy physical quantities.) For high precision, however, dynamical charm quarks are necessary, and some of the most recent simulations now include them.

The $b$ quark mass can be tuned by setting heavy-heavy ($\Upsilon$) or heavy-light ($B$) meson masses to their experimental values. Consistency between these two determinations provides an important check that the determination of parameters in the heavy quark lattice formulations is being done correctly (see, e.g. Ref. 46).

18.1.6. Sources of systematic error:

Lattice results have statistical and systematic errors that must be quantified for any calculation in order for the result to be a useful input to phenomenology. The statistical error is due to the use of Monte Carlo importance sampling to evaluate the path integral (a method discussed below). There are, in addition, a number of systematic errors that are always present to some degree in lattice calculations, although the size of any given error depends on the particular quantity under consideration and the parameters of the ensembles being used. The most common lattice errors are reviewed below.

Although not strictly a systematic error, it is important to note that the presence of long autocorrelations in the sequence of lattice configurations generated by the Monte Carlo method can lead to underestimates of statistical errors [47]. It is known that the global topological charge of the gauge fields decorrelates very slowly with certain algorithms [48]. The effect of poorly sampling topological charge is expected to be most significant for the pion mass and related quantities [49,50]. This issue becomes more relevant as the precision of the final results increases.
18. Lattice QCD

18.1.6.1. Continuum limit: Physical results are obtained in the limit that the lattice spacing \( a \) goes to zero. The Symanzik effective theory determines the scaling of lattice artefacts with \( a \). Most lattice calculations use improved actions with leading discretizations errors of \( \mathcal{O}(a^2\Lambda^2) \), \( \mathcal{O}(\alpha_s a^2\Lambda^2) \), or \( \mathcal{O}(\alpha_s a\Lambda) \), where \( \Lambda \) is a typical momentum scale in the system. Knowledge of the scaling of the leading discretization errors allows controlled extrapolation to \( a = 0 \) when multiple lattice spacings are available, as in current state-of-the-art calculations. Residual errors arise from the exclusion of subleading \( a \) dependence from the fits.

For many quantities the typical momentum scale in the system is \( \sim \Lambda_{\text{QCD}} \approx 300 \text{ MeV} \). Discretization errors are expected to be larger for quantities involving larger scales, for example form factors or decays involving particles with momenta larger than \( \Lambda_{\text{QCD}} \).

18.1.6.2. Infinite volume limit: LQCD calculations are necessarily carried out in finite space-time boxes, leading to departures of physical quantities (masses, decay constants, etc.) from their measured, infinite volume values. These finite-volume shifts are an important systematic that must be estimated and minimized.

Typical lattices are asymmetric, with \( N_s \) points in the three spatial directions and \( N_t \) in the (Euclidean) temporal direction. The spatial and temporal sizes in physical units are thus \( L_s = aN_s \) and \( L_t = aN_t \), respectively. (Anisotropic lattice spacings are also sometimes used, as discussed below in Sec. 1.3.1.) Typically, \( L_t \geq 2L_s \), a longer temporal direction being used to allow excited-state contributions to correlators to decay. This means that the dominant impact of using finite volume is from the presence of a finite spatial box.

High-precision LQCD calculations are of quantities involving no more than a single particle in initial and final states (with the exception of the \( K \to \pi\pi \) decay amplitudes). For such quantities, once the volume exceeds about 2 fm (so that the particle is not “squeezed”), the dominant finite-volume effect comes from virtual pions wrapping around the lattice in the spatial directions. This effect is exponentially suppressed as the volume becomes large, roughly as \( \sim \exp(-m_\pi L_s) \), and has been estimated using ChPT [51] or other methods [52]. The estimates suggest that finite volume shifts are sub-percent effects when \( m_\pi L_s \gtrsim 4 \), and most large-scale simulations use lattices satisfying this condition. This becomes challenging as one approaches the physical pion mass, for which \( L_s \gtrsim 5 \text{ fm} \) is required. At present, this can only be achieved by using relatively coarse lattices, \( a \gtrsim 0.07 \text{ fm} \).

Finite volume errors are usually determined by repeating the simulations on two or more different volumes (with other parameters fixed). If different volumes are not available, the ChPT estimate can be used, often inflated to account for the fact that the ChPT calculation is truncated at some order.

In the future, LQCD calculations involving more than a single hadron will become increasingly precise. Examples include the calculation of resonance parameters and the above-mentioned \( K \to \pi\pi \) amplitudes. Finite volume effects are much larger in these cases, with power-law terms (e.g. \( 1/L_s^3 \)) in addition to exponential dependence. Indeed, as will be discussed in Sec. 1.2.4., one can use the volume dependence to indirectly extract infinite-volume quantities such as scattering lengths. Doing so, however, requires a set of
lattice volumes satisfying $m_\pi L_s \gtrsim 4$ and is thus more challenging than for single-particle quantities.

18.1.6.3. Chiral extrapolation:

Until recently, an important source of systematic error in LQCD calculations was the need to extrapolate in $m_u$ and $m_d$ (or, equivalently, in $m_\pi$). This extrapolation was usually done using functional forms based on ChPT, or with analytic functions, with the difference between different fits used as an estimate of the systematic error, which was often substantial. Increasingly, however, calculations work directly at, or very close to, the physical quark masses. This either removes entirely, or greatly reduces, the uncertainties in the extrapolation, such that this error is subdominant.

18.1.6.4. Operator matching:

Many of the quantities that LQCD can precisely calculate involve hadronic matrix elements of operators from the electroweak Hamiltonian. Examples include the pion and kaon decay constants, semileptonic form factors and the kaon mixing parameter $B_K$ (the latter defined in Eq. (18.13)). The operators in the lattice matrix elements are defined in the lattice regularization scheme. To be used in tests of the Standard Model, however, they must be matched to the continuum regularization scheme in which the corresponding Wilson coefficients have been calculated. The only case in which such matching is not needed is if the operator is a conserved or partially conserved current. Similar matching is also needed for the conversion of lattice bare quark masses to those in the continuum MS scheme.

Three methods are used to calculate the matching factors: perturbation theory (usually to one- or two-loop order), non-perturbative renormalization (NPR) using Landau-gauge quark and gluon propagators [53], and NPR using gauge-invariant methods based on the Schrödinger functional [54]. The NPR methods replace truncation errors (which can only be approximately estimated) by statistical and systematic errors which can be determined reliably and systematically reduced.

A common issue that arises in many such calculations (e.g. for quark masses and $B_K$) is that, using NPR, one ends up with operators regularized in a MOM-like (or Schrödinger functional) scheme, rather than the $\overline{\text{MS}}$ scheme mostly used for calculating the Wilson coefficients. To make contact with this scheme requires a purely continuum perturbative matching calculation. The resultant truncation error can, however, be minimized by pushing up the momentum scale at which the matching is done using step-scaling techniques as part of the NPR calculation [55]. It should also be noted that this final step in the conversion to the $\overline{\text{MS}}$ scheme could be avoided if continuum calculations used a MOM-like scheme.
18.2. Methods and status

Once the lattice action is chosen, it is straightforward to define the quantum theory using the path integral formulation. The Euclidean-space partition function is

\[ Z = \int [dU] \prod_f [dq_f][dq_f] e^{-S_g[U] - \sum_f \bar{q}_f (D[U] + m_f) q_f}, \]

where link variables are integrated over the SU(3) manifold, \( q_f \) and \( \bar{q}_f \) are Grassmann (anticommuting) quark and antiquark fields of flavor \( f \), and \( D[U] \) is the chosen lattice Dirac operator with \( m_f \) the quark mass in lattice units. Integrating out the quark and antiquark fields, one arrives at a form suitable for simulation:

\[ Z = \int [dU] e^{-S_g[U]} \prod_f \det(D[U] + m_f). \]

The building blocks for calculations are expectation values of multi-local gauge-invariant operators, also known as “correlation functions”,

\[ \langle O(U, q, \bar{q}) \rangle = (1/Z) \int [dU] \prod_f [dq_f][dq_f] O(U, q, \bar{q}) e^{-S_g[U] - \sum_f \bar{q}_f (D[U] + m_f) q_f}. \]

If the operators depend on the (anti-)quark fields \( q_f \) and \( \bar{q}_f \), then integrating these fields out leads not only to the fermion determinant but also, through Wick’s theorem, to a series of quark “propagators”, \( (D[U] + m_f)^{-1} \), connecting the positions of the fields.

This set-up allows one to choose, by hand, the masses of the quarks in the determinant (the sea quarks) differently from those in the propagators (valence quarks). This is called “partial quenching”, and is used by some calculations as a way of obtaining more data points from which to extrapolate both sea and valence quarks to their physical values.

18.2.1. Monte-Carlo method:

Since the number of integration variables \( U \) is huge \((N_s^3 \times N_t \times 4 \times 9)\), direct numerical integration is impractical and one has to use Monte-Carlo techniques. In this method, one generates a Markov chain of gauge configurations (a “configuration” being the set of \( U \)’s on all links) distributed according to the probability measure \([dU] e^{-S_g[U]} \prod_f \det(D[U] + m_f)\). Once the configurations are generated, expectation values \( \langle O(U, q, \bar{q}) \rangle \) are calculated by averaging over those configurations. In this way the configurations can be used repeatedly for many different calculations, and there are several large collections of ensembles of configurations (with a range of values of \( a \), lattice sizes and quark masses) that are publicly available through the International Lattice Data Grid (ILDG). As the number of the configurations, \( N \), is increased, the error decreases as \( 1/\sqrt{N} \).

The most challenging part of the generation of gauge configurations is the need to include the fermion determinant. Direct evaluation of the determinant is not feasible,
18. Lattice QCD

as it requires $O((N_s^3 \times N_t)^3)$ computations. Instead, one rewrites it in terms of “pseudofermion” fields $\phi$ (auxiliary fermion fields with bosonic statistics). For example, for two degenerate quarks one has

$$\det(D[U] + m_f)^2 = \int [d\phi] e^{-\phi^\dagger (D[U] + m_f)^{-1} \phi}.$$  \hspace{1cm} (18.7)

By treating the pseudofermions as additional integration variables in the path integral, one obtains a totally bosonic representation. The price one pays is that the pseudofermion effective action is highly non-local since it includes the inverse Dirac operator $(D[U] + m_f)^{-1}$. Thus, the large sparse matrix $(D[U] + m)$ has to be inverted every time one needs an evaluation of the effective action.

Present simulations generate gauge configurations using the Hybrid Monte Carlo (HMC) algorithm [56], or variants thereof. This algorithm combines molecular dynamics (MD) evolution in a fictitious time (which is also discretized) with a Metropolis “accept-reject” step. It makes a global update of the configuration, and is made exact by the Metropolis step. In its original form it can be used only for two degenerate flavors, but extensions (particularly the rational HMC [57]) are available for single flavors. Considerable speed-up of the algorithms has been achieved over the last two decades using a variety of techniques.

All these algorithms spend the bulk of their computational time on the repeated inversion of $(D[U] + m)$ acting on a source (which is required at every step of the MD evolution). Inversions are done using a variety of iterative algorithms, e.g. the conjugate gradient algorithm. In this class of algorithms, computational cost is proportional to the condition number of the matrix, which is the ratio of maximum and minimum eigenvalues. For $(D[U] + m)$ the smallest eigenvalue is $\approx m$, so the condition number and cost are inversely proportional to the quark mass. This is a major reason why simulations at the physical quark mass are challenging. Recent algorithmic studies are making progress in significantly reducing this problem.

A practical concern is the inevitable presence of correlations between configurations in the Markov chain. These are characterized by an autocorrelation length in the fictitious MD time. One aims to use configurations separated in MD time by greater than this autocorrelation length. In practice, it is difficult to measure this length accurately, and this leads to some uncertainty in the resulting statistical errors, as well as the possibility of insufficient equilibration.

For most of the applications of LQCD discussed in this review, the cost of generating gauge configurations is larger than or similar to that of performing the “measurements” on those configurations. The computational cost of gauge generation grows with the lattice volume, $V_{\text{lat}} = N_s^3 N_t$, as $V_{\text{lat}}^{1+\delta}$. Here $\delta = 1/4$ for the HMC algorithm [58] and can be reduced slightly using modern variants. Such growth with $V_{\text{lat}}$ provides a (time-dependent) limit on the largest lattice volumes that can be simulated. At present, the largest lattices being used have $N_s = 144$ and $N_t = 288$. Typically one aims to create an ensemble of $\sim 10^3$ statistically independent configurations at each choice of parameters ($a$, $m_q$ and $V_{\text{lat}}$). For most physical quantities of interest, this is sufficient to make the resulting statistical errors smaller than or comparable to the systematic errors.
18.2.2. Two-point functions:

One can extract properties of stable hadrons using two-point correlation functions, \( \langle O_X(x)O_Y^\dagger(0) \rangle \). Here \( O_{X,Y}(x) \) are operators that have non-zero overlaps with the hadronic state of interest \( |H\rangle \), i.e. \( \langle 0|O_{X,Y}(x)|H\rangle \neq 0 \). One usually Fourier-transforms in the spatial directions and considers correlators as a function of Euclidean time:

\[
C_{XY}(t; \vec{p}) = \sum_{\vec{x}} \langle O_X(t, \vec{x})O_Y^\dagger(0) \rangle e^{-i\vec{p}\cdot\vec{x}}.
\] (18.8)

(Here and throughout this section all quantities are expressed in dimensionless lattice units, so that, for example, \( \vec{p} = a\vec{p}_{\text{phys}} \)). By inserting a complete set of states having spatial momentum \( \vec{p} \), the two-point function can be written as

\[
C_{XY}(t; \vec{p}) = \sum_{i=0}^{\infty} \frac{1}{2E_i(\vec{p})} \langle 0|O_X(0)|H_i(\vec{p})\rangle \langle H_i(\vec{p})|O_Y^\dagger(0)|0\rangle e^{-E_i(\vec{p})t},
\] (18.9)

where the energy of the \( i \)-th state \( E_i(\vec{p}) \) appears as an eigenvalue of the time evolution operator \( e^{-Ht} \) in the Euclidean time direction. The factor of \( 1/[2E_i(\vec{p})] \) is due to the relativistic normalization used for the states. For large enough \( t \), the dominant contribution is that of the lowest energy state \( |H_0(\vec{p})\rangle \):

\[
C_{XY}(t) \xrightarrow{t \to \infty} \frac{1}{2E_0(\vec{p})} \langle 0|O_X(0)|H_0(\vec{p})\rangle \langle H_0(\vec{p})|O_Y^\dagger(0)|0\rangle e^{-E_0(\vec{p})t}.
\] (18.10)

One can thus obtain the energy \( E_0(\vec{p}) \), which equals the hadron mass \( m_H \) when \( \vec{p} = 0 \), and the product of matrix elements \( \langle 0|O_X(0)|H_i(\vec{p})\rangle \langle H_i(\vec{p})|O_Y^\dagger(0)|0\rangle \).

This method can be used to determine the masses of all the stable mesons and baryons by making appropriate choices of operators. For example, if one uses the axial current, \( O_X = O_Y = A_\mu = \bar{d}\gamma_\mu\gamma_5 u \), then one can determine \( m_{a_\mu} \) from the rate of exponential fall-off, and in addition the decay constant \( f_\pi \) from the coefficient of the exponential. A complication arises for states with high spins (\( j \geq 4 \) for bosons) because the spatial rotation group on the lattice is a discrete subgroup of the continuum group \( SO(3) \). This implies that lattice operators, even when chosen to lie in irreducible representations of the lattice rotation group, have overlap with states that have a number of values of \( j \) in the continuum limit [59]. For example \( j = 0 \) operators can also create mesons with \( j = 4 \). A method to overcome this problem has recently been introduced [60,61].

The expression given above for the correlator \( C_{XY}(t; \vec{p}) \) shows how, in principle, one can determine the energies of the excited hadron states having the same quantum numbers as the operators \( O_{X,Y} \), by fitting the correlation function to a sum of exponentials. In practice, this usually requires using a large basis of operators and adopting the variational approach such as that of Ref. 62. One can also use an anisotropic lattice in which \( a_t \), the lattice spacing in the time direction, is smaller than its spatial counterpart \( a_s \). This allows better separation of the different exponentials. Using a combination of these and other technical improvements extensive excited-state spectra have recently been obtained [61,63,64].
18.2.3. Three-point functions:

Hadronic matrix elements needed to calculate semileptonic form factors and neutral meson mixing amplitudes can be computed from three-point correlation functions. We discuss here, as a representative example, the \( D \to K \) amplitude. As in the case of two-point correlation functions one constructs operators \( O_D \) and \( O_K \) having overlap, respectively, with the \( D \) and \( K \) mesons. We are interested in calculating the matrix element \( \langle K | V_\mu | D \rangle \), with \( V_\mu = \bar{c} \gamma_\mu s \) the vector current. To obtain this, we use the three-point correlator

\[
C_{KV_\mu D}(t_\chi, t_\eta; \vec{p}) = \sum_{\vec{x}, \vec{y}} \langle O_K(t_\chi, \vec{x}) V_\mu(0) O_D^\dagger(t_\eta, \vec{y}) \rangle e^{-i\vec{p} \cdot \vec{x}},
\]

and focus on the limit \( t_\chi \to \infty, t_\eta \to -\infty \). In this example we set the \( D \)-meson at rest while the kaon carries three-momentum \( \vec{p} \). Momentum conservation then implies that the weak operator \( V_\mu \) inserts three-momentum \( -\vec{p} \). Inserting a pair of complete sets of states between each pair of operators, we find

\[
C_{KV_\mu D}(t_\chi, t_\eta; \vec{p}) = \sum_{i,j} \frac{1}{2m_D 2E_K} e^{-m_D t_\chi - E_K|t_\eta|} \times
\]

\[
\times \langle 0 | O_K(t_\chi, \vec{x}) | K_i(\vec{p}) \rangle \langle K_i(\vec{p}) | V_\mu(0) | D_j(\vec{0}) \rangle \langle D_j(\vec{0}) | O_D^\dagger(0) | 0 \rangle.
\]

The matrix element \( \langle K_i(\vec{p}) | V_\mu(0) | D_j(\vec{0}) \rangle \) can then be extracted, since all other quantities in this expression can be obtained from two-point correlation functions. Typically one is interested in the weak matrix elements of ground states, such as the lightest pseudoscalar mesons. In the limit of large separation between the three operators in Euclidean time, the three-point correlation function yields the weak matrix element of the transition between ground states.

18.2.4. Scattering amplitudes and resonances:

The methods described thus far yield matrix elements involving single, stable particles (where by stable here we mean absolutely stable to strong interaction decays). Most of the particles listed in the Review of Particle Properties are, however, unstable—they are resonances decaying into final states consisting of multiple strongly interacting particles. LQCD simulations cannot directly calculate resonance properties, but methods have been developed to do so indirectly for resonances coupled to two-particle final states in the elastic regime [65].

The difficulty faced by LQCD calculations is that, to obtain resonance properties, or, more generally, scattering phase-shifts, one must calculate multiparticle scattering amplitudes in momentum space and put the external particles on their mass-shells. This requires analytically continuing from Euclidean to Minkowski momenta. Although it is straightforward in LQCD to generalize the methods described above to calculate four- and higher-point correlation functions, one necessarily obtains them at a discrete and finite set of Euclidean momenta. Analytic continuation to \( p_E^2 = -m^2 \) is then an ill-posed
and numerically unstable problem. The same problem arises for single-particle states, but can be largely overcome by picking out the exponential fall-off of the Euclidean correlator, as described above. With a multi-particle state, however, there is no corresponding trick, except for two particles at threshold [66].

What LQCD can calculate are the energies of the eigenstates of the QCD Hamiltonian in a finite box. The energies of states containing two stable particles, e.g. two pions, clearly depend on the interactions between the particles. It is possible to invert this dependence and, with plausible assumptions, determine the scattering phase-shifts at a discrete set of momenta from a calculation of the two-particle energy levels for a variety of spatial volumes [65]. This is a challenging calculation, but it has recently been carried through in several channels with quark masses approaching physical values. Channels studied include \( \pi \pi \) (for \( I = 2, 1 \) and \( 0 \)), \( K \pi \), \( K D \) and \( DD^* \). For recent reviews see Ref. 67. Extensions to nucleon interactions are also being actively studied [68]. The generalization of the formalism to the case of three particles is under active consideration [69].

It is also possible to extend the methodology to calculate electroweak decay amplitudes to two particles below the inelastic threshold, e.g. \( \Gamma(K \rightarrow \pi \pi) \) [70]. Results for both the \( \Delta I = 3/2 \) and \( 1/2 \) amplitudes with physical quark masses have been obtained [71], the former now including a controlled continuum limit [72]. Partial extensions of the formalism above the elastic threshold have been worked out, in particular for the case of multiple two-particle channels [73]. An extension to decays with many multiparticle channels, e.g. hadronic \( B \) decays, has, however, yet to be formulated.

### 18.2.5. Recent advances

In some physics applications, one is interested in the two-point correlation function \( \langle O(x)O^\dagger(0) \rangle \) for all values of the separation \( x \), not just its asymptotic form for large separations (which is used to determine the hadron spectrum as sketched above). A topical example is the hadronic vacuum polarization function \( \Pi_{\mu\nu}(x) = \langle V_{\mu}(x)V_{\nu}(0) \rangle \) and its Fourier transform \( \Pi_{\mu\nu}(q^2) \). Since the lattice is in Euclidean space-time, only space-like momenta, \( q^2 = -Q^2 < 0 \), are accessible. Nevertheless, this quantity is of significant interest. It is related by a dispersion relation to the cross section for \( e^+ + e^- \rightarrow \) hadrons, and is needed for a first-principles calculation of the “hadronic vacuum polarization” contribution to the muon anomalous magnetic moment \( a_\mu \). This is the contribution with the largest theoretical uncertainty at present. There are a number of lattice calculations of this contribution (see, e.g., [74–77] following the pioneering work of Ref. 78).

Exploratory calculations of the light-by-light scattering contribution to \( a_\mu \) are also underway. These involve the evaluation of a four-point correlation function (see, e.g., Ref. 79). Another process under consideration is the long-distance contribution to the neutral kaon mass splitting, \( \Delta M_K \). This also requires the evaluation of a four-point function, constructed from the two-point functions described above by the insertion of two electroweak Hamiltonians [80].
18.2.6. Status of LQCD simulations:

Until the 1990s, most large-scale lattice simulations were limited to the “quenched” approximation, wherein the fermion determinant is omitted from the path integral. While much of the basic methodology was developed in this era, the results obtained had uncontrolled systematic errors and were not suitable for use in placing precision constraints on the Standard Model. During the 1990s, more extensive simulations including the fermion determinant (also known as simulations with “dynamical” fermions) were begun, but with unphysically heavy quark masses \(m_\ell \sim 50 - 100\) MeV, such that the extrapolation to the physical light quark masses was a source of large systematic errors \[81\]. During the 2000s, advances in both algorithms and computers allowed simulations to reach much smaller quark masses \(m_\ell \sim 10 - 20\) MeV such that LQCD calculations of selected quantities with all sources of error controlled and small became available. Their results played an important role in constraints on the CKM matrix and other phenomenological analyses. In the last few years, simulations directly at the physical isospin-symmetric light quark masses have become standard, removing the need for a chiral extrapolation and thus significantly reducing the overall error. The present frontier, as noted above, is the inclusion of isospin breaking. This will be needed to push the accuracy of calculations below the percent level.

On a more qualitative level, analytic and numerical results from LQCD have demonstrated that QCD confines color and spontaneously breaks chiral symmetry. Confinement can be seen as a linearly rising potential between heavy quark and anti-quark in the absence of quark loops. Analytically, this can be shown in the strong coupling limit \(g_{\text{lat}} \rightarrow \infty\) \[1\]. At weaker couplings there are precise numerical calculations of the potential that clearly show that this behavior persists in the continuum limit \[82,83,84\].

Chiral symmetry breaking was also demonstrated in the strong coupling limit on the lattice \[16,85\], and there have been a number of numerical studies showing that this holds also in the continuum limit. The accumulation of low-lying modes of the Dirac operator, which is the analog of Cooper pair condensation in superconductors, has been observed, yielding a determination of the chiral condensate \[86-90\]. Many relations among physical quantities that can be derived under the assumption of broken chiral symmetry have been confirmed by a number of lattice groups \[41\].

18.3. Physics applications

In this section we describe the main applications of LQCD that are both computationally mature and relevant for the determination of particle properties.

A general feature to keep in mind is that, since there are many different choices for lattice actions, all of which lead to the same continuum theory, a crucial test is that results for any given quantity are consistent. In many cases, different lattice calculations are completely independent and often have very different systematic errors. Thus final agreement, if found, is a highly non-trivial check, just as it is for different experimental measurements.
The number, variety and precision of the calculations has progressed to the point that an international “Flavour Lattice Averaging Group” (FLAG) has been formed. The main aims of FLAG include collecting all lattice results of relevance for a variety of phenomenologically interesting quantities and providing averages of those results which pass appropriate quality criteria. The averages attempt to account for possible correlations between results (which can arise, for example, if they use common gauge configurations). The quantities considered are those we discuss in this section, with the exception of the hadron spectrum. The most recent FLAG review is from 2013 [41]( with an update due in 2015-16). The interested reader can consult the FLAG review for very extensive discussions of the details of the calculations and of the sources of systematic errors.

We stress that the results we quote below are those obtained using the physical complement of light quarks (i.e. $N_f = 2 + 1$ or $2 + 1 + 1$ simulations).

18.3.1. Spectrum:

The most basic prediction of LQCD is of the hadron spectrum. Once the input parameters are fixed as described in Sec. 18.1.5, the masses or resonance parameters of all other states can be predicted. This includes hadrons composed of light ($u$, $d$ and $s$) quarks, as well as heavy-light and heavy-heavy hadrons. It also includes quark-model exotics (e.g. $J^{PC} = 1^{--}$ mesons) and glueballs. Thus, in principle, LQCD calculations should be able to reproduce many of the experimental results compiled in the Review of Particle Properties. Doing so would test both that the error budgets of LQCD calculations are accurate and that QCD indeed describes the strong interactions in the low-energy domain. The importance of the latter test can hardly be overstated.

What is the status of this fundamental test? As discussed in Sec. 1.2, LQCD calculations are most straightforward for stable, low-lying hadrons. Calculations of the properties of resonances which can decay into only two particles are more challenging, though substantial progress has been made. First theoretical work on decays to more than two particles has begun, but the methodology is not yet practical. It is also more technically challenging to calculate masses of flavor singlet states (which can annihilate into purely gluonic intermediate states) than those of flavor non-singlets, although again algorithmic and computational advances have begun to make such calculations accessible, although not yet for physical quark masses. The present status for light hadrons is that fully controlled results are available for the masses of the octet light baryons, while results with less than complete control are available for the decuplet baryon resonances, the vector meson resonances and the $\eta$ and $\eta'$. In addition, it has been possible to calculate the isospin splitting in light mesons and baryons (due to the up-down mass difference and the incorporation of QED) [40]. There are also extensive results for heavy-light ($D$ and $B$ systems) and heavy-heavy ($J/\psi$ and $\Upsilon$ systems). All present results, which are discussed in the “Quark Model” review, are consistent with experimental values, and several predictions have been made. For a recent extensive review of lattice results see also Ref. 91.
18.3.2. Decay constants and bag parameters:

The pseudoscalar decay constants can be determined from two-point correlation functions involving the axial-vector current, as discussed in Sec. 18.2.2. The decay constant $f_P$ of a meson $P$ is extracted from the weak matrix element involving the axial-vector current using the relation

$$\langle 0 | A_\mu(x)| P(\vec{p}) \rangle = f_P p_\mu \exp(-ip \cdot x),$$

where $p_\mu$ is the momentum of $P$ and $A_\mu(x)$ is the axial-vector current. Since they are among the simplest quantities to calculate, decay constants provide good benchmarks for lattice methods, in addition to being important inputs for flavor physics phenomenology in their own right. Results from many lattice groups for the pion and kaon decay constants now have errors at the percent level or better. The decay constants in the charm and bottom sectors, $f_D$, $f_{D_s}$, $f_B$, and $f_{B_s}$, have also been calculated to high precision. Lattice results for all of these decay constants are discussed in detail in the review “Leptonic Decays of Charged Pseudoscalar Mesons.”

Another important lattice quantity is the kaon bag parameter $B_K$, which is needed to turn the precise measurement of CP-violation in kaon mixing into a constraint on the Standard Model. It is defined by

$$\frac{8}{3} m_K^2 f_K^2 B_K(\mu) = (\overline{K}_0 | Q_{\Delta S=2}(\mu) | K^0),$$

where $m_K$ is the kaon mass, $f_K$ is the kaon decay constant, $Q_{\Delta S=2} = \bar{s} \gamma_\mu (1-\gamma_5) d \bar{s} \gamma_\mu (1-\gamma_5) d$ is the four-quark operator of the effective electroweak Hamiltonian and $\mu$ is the renormalization scale. The short distance contribution to the electroweak Hamiltonian can be calculated perturbatively, but the hadronic matrix element parameterized by $B_K$ must be computed using non-perturbative methods. In order to be of use to phenomenology, the renormalization factor of the four-quark operator must be matched to a continuum renormalization scheme, e.g. to $\overline{MS}$, as described in Sec. 18.1.6.4. Determinations with percent-level precision using different fermion actions and $N_f = 2 + 1$ light sea quarks are now available using DWF [92], staggered fermions [93], DWF valence on staggered sea quarks [94], and Wilson fermions [12]. The results are all consistent. Based on results available in 2013, FLAG quoted an average of $\tilde{B}_K = 0.766(10)$ [41]. The updates of Refs. 92 and 93 are consistent with this result.

The bag parameters for $B$ and $B_s$ meson mixing are defined analogously to that for kaon mixing. The $B$ and $B_s$ mesons contain a valence $b$-quark so that calculations of these quantities must use one of the methods for heavy quarks described above. Calculations with $N_f = 2 + 1$ light fermions have been done using NRQCD [95], the Fermilab formalism [96], and static heavy quarks [97]. All results are consistent. The FLAG averages from 2013 for the quantities relevant for $B_s$ and $B$ mixing are $f_{B_s} \sqrt{B_{B_s}} = 266(18)$ MeV and $f_B \sqrt{B_B} = 216(15)$ MeV, with their ratio (which is somewhat better determined) being $\xi = 1.268(63)$ [41]. Note that the errors for quantities involving $b$ quarks are larger than those for quantities involving only light quarks.

The results for mixing matrix elements are used in the reviews “The CKM Quark-Mixing Matrix,” and “$B_0 - \bar{B}_0$ Mixing.”

October 1, 2016 19:59
18.3.3. Form factors ($K \to \pi \ell \nu$, $D \to K \ell \nu$, $B \to \pi \ell \nu$, $B \to D^{(*)} \ell \nu$):

Semileptonic decay rates can be used to extract CKM matrix elements once the semileptonic form factors are known from lattice calculations. For example, the matrix element of a pseudoscalar meson $P$ undergoing semileptonic decay to another pseudoscalar meson $D$ is mediated by the vector current, and can be written in terms of form factors as

$$
\langle D(p_D)|V_\mu|P(p_P)\rangle = f_+(q^2)(p_D + p_P - \Delta)_\mu + f_0(q^2)\Delta_\mu,
$$

where $q = p_D - p_P$, $\Delta_\mu = (m_D^2 - m_P^2)q_\mu/q^2$ and $V_\mu$ is the quark vector current. The shape of the form factor is typically well determined by experiment, and the value of $f_+(q^2)$ at some reference value of $q^2$ is needed from the lattice in order to extract CKM matrix elements. Typically $f_+(q^2)$ dominates the decay rate, since the contribution from $f_0(q^2)$ is suppressed when the final state lepton is light.

The form factor $f_+(0)$ for $K \to \pi \ell \nu$ decays is highly constrained by the Ademollo-Gatto theorem [98] and chiral symmetry. Old estimates using chiral perturbation theory combined with quark models quote sub-percent precision [99], though they suffer from some model dependence. Utilizing the constraint from the vector current conservation that $f_+(0)$ is normalized to unity in the limit of degenerate up and strange quark masses, the lattice calculation can be made very precise and has now matched the precision of the phenomenological estimates [100–106]. The FLAG average from its 2013 edition is $f_+(0) = 0.967(4)$ [41].

Charm meson semileptonic decays have been calculated by different groups using methods similar to those used for charm decay constants, and results are steadily improving in precision [107,108]. For semileptonic decays involving a bottom quark, one uses HQET or NRQCD to control the discretization errors of the bottom quark. The form factors for the semileptonic decay $B \to \pi \ell \nu$ have been calculated in unquenched lattice QCD by a number of groups [109–111]. These $B$ semileptonic form factors are difficult to calculate at low $q^2$, i.e. when the mass of the $B$-meson must be balanced by a large pion momentum, in order to transfer a small momentum to the lepton pair. The low $q^2$ region has large discretization errors and very large statistical errors, while the high $q^2$ region is much more accessible to the lattice. For experiment, the opposite is true. To combine lattice and experimental results it has proved helpful to use the $z$-parameter expansion [112]. This provides a theoretically constrained parameterization of the entire $q^2$ range, and allows one to obtain $|V_{ub}|$ without model dependence [113,114].

The semileptonic decays $B \to D \ell \nu$ and $B \to D^{*} \ell \nu$ can be used to extract $|V_{cb}|$ once the corresponding form factors are known. At present only one unquenched calculation exists for the $B \to D^{*} \ell \nu$ form factor, where the Fermilab formulation of the heavy quark was adopted [115,116]. This calculation is done at zero-recoil because that is where the lattice systematic errors are smallest. Calculations at non-zero recoil in unquenched lattice QCD have recently been done for the first time for the form factors needed to extract $|V_{cb}|$ from $B \to D \ell \nu$ decays [117].

The results discussed in this section are used in the reviews “The CKM Quark-Mixing Matrix,” “$V_{ud}$, $V_{us}$, the Cabibbo Angle and CKM Unitarity,” and “$V_{cb}$ and $V_{ub}$ CKM Matrix Elements.”
18.3.4. Strong coupling constant:

As explained in Sec. 18.1.5.1, for a given lattice action, the choice of bare lattice coupling constant, $g_{\text{lat}}$, determines the lattice spacing $a$. If one then calculates $a$ as described in Sec. 18.1.5.1, one knows the strong coupling constant in the bare lattice scheme at the scale $1/a$, $\alpha_{\text{lat}} = g_{\text{lat}}^2 / (4\pi)$. This is not, however, useful for comparing to results for $\alpha_s$ obtained from other inputs, such as deep inelastic scattering or jet shape variables. This is because the latter results give $\alpha_s$ in the $\overline{\text{MS}}$ scheme, which is commonly used in such analyses, and the conversion factor between these two schemes is known to converge extremely poorly in perturbation theory. Instead one must use a method which directly determines $\alpha_s$ on the lattice in a scheme closer to $\overline{\text{MS}}$.

Several such methods have been used, all following a similar strategy. One calculates a short-distance quantity $K$ both perturbatively ($K^{\text{PT}}$) and non-perturbatively ($K^{\text{NP}}$) on the lattice, and requires equality: $K^{\text{NP}} = K^{\text{PT}} = \sum_{i=0}^{n} c_i \alpha_s^i$. Solving this equation one obtains $\alpha_s$ at a scale related to the quantity being used. Often, $\alpha_s$ thus obtained is not defined in the conventional $\overline{\text{MS}}$ scheme, and one has to convert among the different schemes using perturbation theory. Unlike for the bare lattice scheme, the required conversion factors are reasonably convergent. As a final step, one uses the renormalization group to run the resulting coupling to a canonical scale (such as $M_Z$).

In the work of the HPQCD collaboration [118], the short-distance quantities are Wilson loops of several sizes and their ratios. These quantities are perturbatively calculated to $O(\alpha_s^3)$ using the $V$-scheme defined through the heavy quark potential. The coefficients of even higher orders are estimated using the data at various values of $a$.

Another choice of short-distance quantities is to use current-current correlators. Appropriate moments of these correlators are ultraviolet finite, and by matching lattice results to the continuum perturbative predictions, one can directly extract the $\overline{\text{MS}}$ coupling. The JLQCD collaboration [119] uses this approach with light overlap fermions, while the HPQCD collaboration uses charm-quark correlators and HISQ fermions [120–122]. Yet another choice of short-distance quantity is the static-quark potential, where the lattice result for the potential is compared to perturbative calculations; this method was used to compute $\alpha_s$ within 2+1 flavor QCD [123]. The ETM Collaboration obtains $\alpha_s$ by a comparison of lattice data for the ghost-gluon coupling with that of perturbation theory [124], providing the first determination of $\alpha_s$ with 2+1+1 flavors of dynamical quarks.

With a definition of $\alpha_s$ given using the Schrödinger functional, one can non-perturbatively control the evolution of $\alpha_s$ to high-energy scales, such as 100 GeV, where the perturbative expansion converges very well. This method developed by the ALPHA collaboration [55] has been applied to 2+1-flavor QCD in Refs. 125,126.

The various lattice methods for calculating $\alpha_s$ have significantly different sources of systematic error. Thus the good agreement between the approaches (which can be seen in the “Quantum Chromodynamics” review) provides a strong check on the final result.
18.3.5. Quark masses:

Once the quark mass parameters are tuned in the lattice action, the remaining task is to convert them to those of the conventional definition. Since the quarks do not appear as asymptotic states due to confinement, the pole mass of the quark propagator is not a physical quantity. Instead, one defines the quark mass after subtracting the ultra-violet divergences in some particular way. The conventional choice is again the $\overline{\text{MS}}$ scheme at a canonical scale such as 2 or 3 GeV. Ratios such as $m_c/m_s$ and $m_b/m_c$ are also useful as they are free from multiplicative renormalization (in a mass-independent scheme).

As discussed in Sec. 18.1.6.4, one must convert the lattice bare quark mass to that in the $\overline{\text{MS}}$ scheme. Older calculations did so directly using perturbation theory; most recent calculations use an intermediate NPR method (e.g. RI/MOM or RI/SMOM) which is then converted to the $\overline{\text{MS}}$ scheme using perturbation theory.

Alternatively, one can use a definition based on the Schrödinger functional, which allows one to evolve the quark mass to a high scale non-perturbatively [127]. In practice, one can reach scales as high as $\sim 100$ GeV, at which matching to the $\overline{\text{MS}}$ scheme can be reliably calculated in perturbation theory.

Another approach available for heavy quarks is to match current-current correlators at short distances calculated on the lattice to those obtained in continuum perturbation theory in the $\overline{\text{MS}}$ scheme [120,121]. This has allowed an accurate determination of $m_c$ and is also beginning to be used for $m_b$ [121,122].

The ratio method for heavy quarks (discussed earlier) can also be used to determine $m_b$ [128].

Results are summarized in the review of “Quark Masses”.

18.3.6. Other applications:

In this review we have concentrated on applications of LQCD that are relevant to the quantities discussed in the Review of Particle Properties. We have not discussed at all several other applications which are being actively pursued by simulations. Here we list the major such applications. The reader can consult the texts [2–4] for further details, as well as the proceedings of recent lattice conferences [129].

LQCD can be used, in principle, to simulate QCD at non-zero temperature and density, and in particular to study how confinement and chiral-symmetry breaking are lost as $T$ and $\mu$ (the chemical potential) are increased. This is of relevance to heavy-ion collisions, the early Universe and neutron-star structure. In practice, finite temperature simulations are computationally tractable and relatively mature, while simulations at finite $\mu$ suffer from a “sign problem” and are at a rudimentary stage.

Another topic under active investigation is nucleon structure and inter-nucleon interactions. The simplest nucleon matrix elements are calculable with reasonable precision. Of particular interest are those of the axial current (leading to $g_A$) and of the scalar density (with $\langle N|\bar{s}s|N\rangle$ needed for dark matter searches).

Finally, we note that there is much recent interest in studying QCD-like theories with more fermions, possibly in other representations of the gauge group. The main interest...
Lattice QCD is to find nearly conformal theories which might be candidates for “walking technicolor” models.

18.4. Outlook

While LQCD calculations have made major strides in the last decade, and are now playing an important role in constraining the Standard Model, there are many calculations that could be done in principle but are not yet mature due to limitations in computational resources. As we move to exascale resources (e.g. $10^{18}$ floating point operations per second), the list of mature calculations will grow. Examples that we expect to mature in the next few years are results for excited hadrons, including quark-model exotics, at close to physical light-quark masses; results for moments of structure functions; $K \rightarrow \pi\pi$ amplitudes (allowing a prediction of $\epsilon'/\epsilon$ from the Standard Model); $\bar{K} \leftrightarrow K$ and $\bar{B} \leftrightarrow B$ mixing amplitudes from operators arising in models of new physics (allowing one to constrain these models in a manner complementary to the direct searches at the LHC); hadronic vacuum polarization contributions to muon $g-2$, the running of $\alpha_{\text{EM}}$ and $\alpha_s$; $\pi \rightarrow \gamma\gamma$ and related amplitudes; long-distance contribution to $\overline{K} \leftrightarrow K$ mixing and the light-by-light contribution to muon $g-2$. There will also be steady improvement in the precision attained for the mature quantities discussed above. As already noted, this will ultimately require simulations with $m_u \neq m_d$ and including electromagnetic effects.

References:
7. Y. Iwasaki, UT-HEP-118.
20. A. Bazavov et al. [MILC Collab.], Rev. Mod. Phys. 82, 1349 (2010).
27. A. Borici, hep-lat/9912040; A.D. Kennedy, hep-lat/060703.
30. R. Baron et al. [ETM Collab.], JHEP 1006, 111 (2010).
32. J. Heitger & R. Sommer [ALPHA Collab.], JHEP 0402, 022 (2004); B. Blossier et al. [ALPHA Collab.], JHEP 1012, 039 (2010).
42. S. Aoki et al. [PACS-CS Collab.], PTEP 2012, 01A102 (2012).
18. Lattice QCD

88. K. Cichy et al., JHEP 1310, 175 (2013).
28 18. Lattice QCD


125. S. Aoki et al. [PACS-CS Collab.], JHEP 0910, 053 (2009).


128. A. Bussone et al., arXiv:1411.0484.