Journal of Statistical Physics, Vol. 43, Nos. 5/6, 1986

QCD from Chippewa Falls

Philippe de Forcrand[1]

Recent and new results on $SU(3)$ lattice gauge simulations are presented. They cover the study of finite-size effects for the mass gap and the string tension, the measurement of the central and spin-dependent potentials on large lattices, and the inclusion of dynamical fermions in the simulation.

I have been working at Cray Research now for over 6 months, and I would like to present a summary of my research there. At this conference, Steve Chen gave a general presentation of the Cray machines, and of their performance on a wide range of applications.[1] So I can focus here on QCD applications. The goal of my research at Cray, as I see it, is two-fold: I would like those of you who are still computing on "small" machines to become convinced of the benefits and quasi-necessity to gain access to a supercomputer—that's an easy task. I also wish to make the lucky few who already use a Cray aware of the true computing power of such machines, by showing practical examples. Therefore I will successively review numerical results obtained recently in Chippewa Falls on finite-size effects in pure gauge QCD, on the central quark–antiquark potential, on spin-dependent potentials, and on the simulation of dynamical fermions.

1. FINITE-SIZE EFFECTS IN PURE GAUGE QCD

Suppose one wants to measure the mass gap m_g of pure gauge QCD. One widely used method consists in forming an operator $\mathcal{O}$, which has some overlap with the physical state under investigation (called glueball here), and to measure the connected correlation of this observable as a function of euclidean time t. For large t, the higher excited states generated

[1] Cray Research, Inc., Chippewa Falls, Wisconsin 54729.

0022-4715/86/0600-1077$05.00/0

by the application of $\mathcal{O}$ on the QCD vacuum will have exponentially died off, and one will have

$$C(t) \equiv \langle \mathcal{O}(t_0)\, \mathcal{O}(t_0+t) \rangle - \langle \mathcal{O}(t_0) \rangle \langle \mathcal{O}(t_0+t) \rangle \underset{t \to \infty}{\sim} e^{-m_g t} \tag{1}$$

where $C(t)$ is evaluated by Monte Carlo on a lattice of, say, $N_S^3 \times N_t$ sites and m_g is the energy of a glueball at rest. $\mathcal{O}$ should therefore be a $\vec{0}$-momentum, translation-invariant operator of the form

$$\mathcal{O}(t) = \frac{1}{N_S^3} \sum_{\vec{x}} \omega(\vec{x}, t) \tag{2}$$

where the normalization factor ensures that $\langle \mathcal{O} \rangle$ is independent of the spatial extent N_S of the lattice. Assume N_S can be chosen arbitrarily. The fluctuations of a single measurement of $\mathcal{O}(t)$ around $\langle \mathcal{O} \rangle$ vary like $N_S^{-3/2}$. Thus a measurement of $\langle \mathcal{O}(t_0)\, \mathcal{O}(t_0+t) \rangle$ will yield a connected part of order $N_S^{-3/2}$ on top of the disconnected part $\langle \mathcal{O} \rangle^2$ that must be subtracted. However, $\langle \mathcal{O} \rangle$ is itself known with an error of order $N_S^{-3/2}$, so that after a given number of measurements of $\mathcal{O}(t)$, the signal-to-noise ratio for $C(t)$ is *independent of* N_S. Since the computer time to take measurements grows like N_S^3 it is clearly advantageous to work on long, spatially thin lattices. This strategy, adopted in the early efforts at calculating the glueball spectrum,[2] can be marred by finite-size effects.

An alternative approach consists in perturbing the vacuum at time t by putting a source having an overlap with the desired glueball. Without the source, one would have

$$H|\Omega_0\rangle = E_0|\Omega_0\rangle \tag{3}$$

where H is the QCD Hamiltonian and E_0 the ground-state energy. With the source

$$H|\Omega'\rangle = \sum_{i=0}^{\infty} \alpha_i E_i |\Omega_i\rangle \tag{4}$$

The expectation value of observable $\mathcal{O}$ a distance t from the source is then

$$\begin{aligned} \langle \mathcal{O}(t) \rangle &\equiv \frac{\langle \Omega_0 |\, e^{-Ht} \mathcal{O} e^{-Ht}\, | \Omega_0 \rangle}{\langle \Omega_0 |\, e^{-2Ht}\, | \Omega_0 \rangle} = \frac{\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \alpha_i^* \alpha_j e^{-(E_i+E_j)t} \langle \Omega_i |\, \mathcal{O}\, | \Omega_j \rangle}{\sum_{i=0}^{\infty} |\alpha_i|^2\, e^{-2E_i t}} \\ &= \langle \Omega_0 |\, \mathcal{O}\, | \Omega_0 \rangle + 2Re \left(\frac{\alpha_i}{\alpha_0} \langle \Omega_0 |\, \mathcal{O}\, | \Omega_i \rangle \right) e^{-(E_i - E_0)t} + \cdots \end{aligned} \tag{5}$$

The leading term in (5) is the unperturbed vacuum expectation value $\langle \mathcal{O} \rangle$. The relaxation of $\langle \mathcal{O}(t) \rangle$ toward $\langle \mathcal{O} \rangle$ is exponential, and the exponent $(E_1 - E_0)$ is the desired mass gap.

In this approach only expectation values are measured, with a signal-to-noise ratio proportional to $N_S^{-3/2}$. The unperturbed vacuum expectation value must be subtracted, but it is kept as a free parameter with no statistical noise. Therefore this method should be more efficient than the previous one by a factor $C \times N_S^3$, where C appears to be not much smaller than 1.[3,4]

The source at time t_0 can be arbitrarily large. An early attempt[5] considered an additive source term proportional to the plaquette. We simply set all spatial links at time t_0 to the identity. All spatial loops at t_0 are therefore set to 1, and we can study the relaxation of a wide choice of observables. All moderate-size loops turn out to probe the O^{++} glueball equally well. A Polyakov loop winds around the periodic lattice in one spatial direction. It is topologically different. Correlations $C(N_S, t)$ between two parallel Polyakov loops of length $(N_S \times a)$ at separation $(t \times a)$ probe the force between static quarks

$$\lim_{t \to \infty} \lim_{N_S \to \infty} \frac{-\ln C(N_S, t)}{N_S \cdot t} = \sigma a^2 \tag{6}$$

where σ is the string tension and a is the lattice spacing. The vacuum expectation value of a Polyakov loop is 0 because of the Z_3 symmetry of the pure gauge QCD theory. Therefore, in the presence of our source, the average value P will relax from 1 to 0 according to

$$\langle P(N_S, t) \rangle \sim e^{-tN_S \sigma a^2} \qquad \text{for} \quad N_S \gg t \gg 1 \tag{7}$$

Spectacular results have been obtained using this method for the string tension[3,6,7] and the mass gap.[3,8] Indeed, the results become accurate enough to study their dependence on N_S, i.e., finite-size effects.

The mass gap must vary with N_S according to Ref. 9

$$m_g(N_S) = m_g(\infty) \left\{ 1 - G \frac{e^{-(\sqrt{3}/2) m_g(\infty) a N_S}}{m_g(\infty)\, a N_S} \left[1 + \mathcal{O}\left(\frac{1}{N_S} \right) \right] \right\} \tag{8}$$

A strong coupling expansion indicates that G could be quite large, $O(100\text{–}300)$.[10] For the Polyakov loop, a classical string model was considered,[7] yielding a fist-order correction

$$\langle P(N_S, t) \rangle \sim e^{-t[N_S \sigma a^2 - (\lambda / N_S)]} \qquad \text{for} \qquad t \gg N_S \gg 1 \tag{9}$$

with $\lambda = \pi/3$. A first attempt was made[7,8] at checking the effective values of G and λ. Two simulations were run at $\beta = 5.7$, on a $6^3 \times 16$ and an $8^3 \times 16$ lattice. The data for the 2×2 loop and the Polyakov loop are plotted in Fig. 1. They yield

$$G = 155 \pm 45 \qquad \lambda = 1.09 \pm .39 \tag{10}$$

I recently repeated the same simulations at $\beta = 5.9$ on lattices of sizes $8^3 \times 20$, $10^3 \times 20$, and $12^3 \times 20$, with 45000, 35000, 25000 sweeps, respectively. A preliminary analysis of the results indicates[11]

$$G \sim 55 \qquad \lambda \sim 1.5 \tag{11}$$

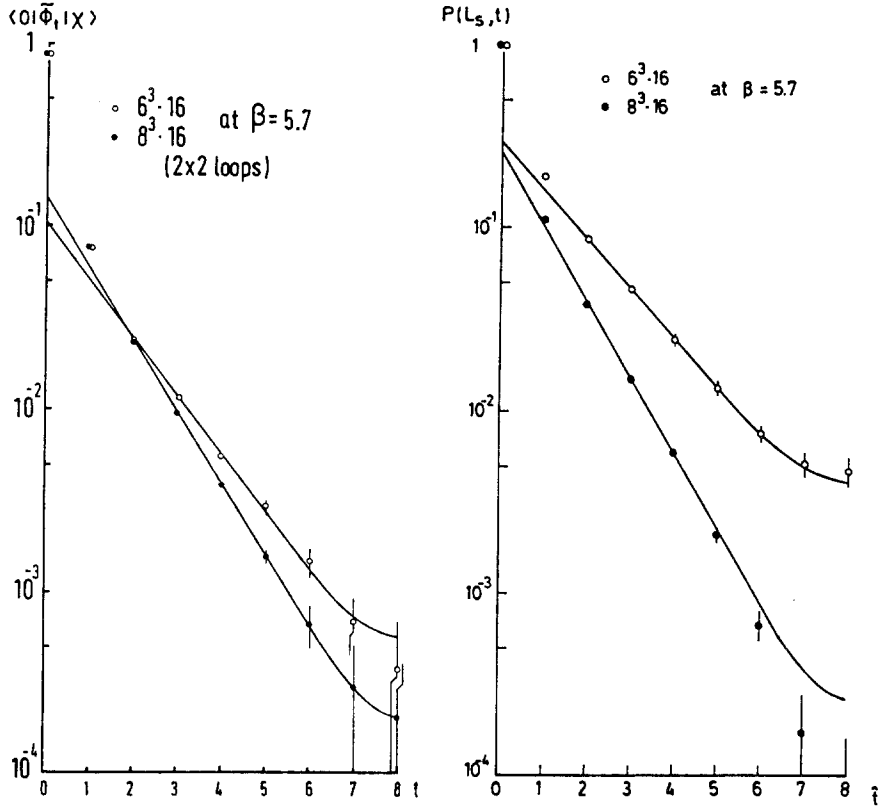

Fig. 1. (a) Relaxation of the 2×2 loop and (b) Polyakov loop to their unperturbed vacuum expectation value as a function of their distance from the source. Two lattices of different spatial sizes, 6 and 8, are compared. A ratio of relaxation exponents different from 1 or 6/8 (b) provides evidence of finite-size effects; taken from Refs. 7 and 8.

A careful error analysis still needs to be performed. But there is a clear reduction of finite-size effects for the glueball: as β increases, G decreases from the strong coupling prediction. The value for λ is still roughly consistent with the string picture. The three sets of results are not accurate enough to allow a check of the functional forms (8) and (9). Taking these first-order formulas for granted, an extrapolation to $N_S \to \infty$ yields the infinite volume values

$$m_g a \sim .685 \qquad \sigma a^2 \sim .065 \qquad \text{at} \quad \beta = 5.9 \tag{12}$$

These values are very close to those of Ref. 7 and 8. It would be nice to express m_g and $\sigma^{1/2}$ in units of the cutoff Λ_L. However, the perturbative two-loop formula

$$a(\beta)\,\Lambda_L = e^{-(4\pi^2/33)\beta} \left(\frac{8\pi^2}{33}\beta\right)^{51/121} \tag{13}$$

does not hold until β gets much larger than 5.9, as we shall see in Section 2. A less stringent assumption is that of scaling, which implies

$$\left.\frac{m_g}{\sigma^{1/2}}\right|_{\beta=5.9} = \left.\frac{m_g}{\sigma^{1/2}}\right|_{\beta=\infty} \sim 2.7 \tag{14}$$

In a world of gluons and infinitely heavy quarks, where the string tension would still be $\sigma^{1/2} \sim 420$ MeV, the O^{++} glueball might be expected to lie around 1.1 GeV.

2. QUARK–ANTIQUARK CENTRAL POTENTIAL

The method to obtain the static $q\bar{q}$ potential $V(R)$ via Wilson loops is well-known. If the average value of a rectangular loop of sizes R and T is called $W(R, T)$, then

$$\lim_{T\to\infty} \frac{-\ln W(R, T)}{T} = V(R) \tag{15}$$

So the recipe goes. Start with as much computer time as possible. Choose the biggest lattice that your budget and the memory of your machine will accommodate. Choose the smallest lattice spacing that will still allow a proton to sit "comfortably" in your lattice. Generate by Monte Carlo "enough" configurations to protect yourself against thermalization effects and critical slowing down of the algorithm (there is no accepted definition for the terms "comfortably" and "enough"). Measure elongated loops (planar or nonplanar) on your configurations. It is a waste of time to measure loops too often, since they take some Monte Carlo steps to

decorrelate. A grandmother's rule of thumb says: Spend at most half of your time on loop measurements.

Along these guidelines, my effort in Chippewa has been the following. The lattice size is $24^3 \times 48$ at $\beta = 6.3$. Thermalization was achieved by 3500 sweeps (starting from a thermalized $24^3 \times 6$ lattice replicated eight times). Then measurements were taken for 10,000 sweeps. Planar loops of size up to 24×12 in one spatial plane were measured every other sweep (the measurement plane rotated every sweep). Simulations of a 16^4 lattice were also performed at $\beta = 6.0$ and 6.3, each with 20,000 sweeps for measurement. The lower β value was chosen to check the program against earlier work[12,13] and to allow a study of the scaling behavior of the potential. The higher β value was chosen to check for finite-size effects between the 16^4 and $24^3 \times 48$ lattices. It turns out that all Wilson loops $W(R, T)$ measured on the smaller lattice are consistent with those on the larger one, except for $R \geqslant 15$. The effects due to the periodicity of the lattice are not visible until the loop very nearly wraps around. This result supports our use of loops up to $T = 22$ for the extraction of the potential on the larger lattice.

The $24^3 \times 48$ lattice represents over 21 million degrees of freedom and over 48 million words of data [an 18 word $SU(3)$ matrix only has 8 degrees of freedom. These simulations are therefore something of a technological feat, and I should emphasize the following technical aspects (a nutshell presentation of the project is given in Ref. 14). The core memory used is organized as a circular buffer containing four time-slices of data: one time-slice, say i, is updated, using information from its two neighbors $(i-1)$ and $(i+1)$; meanwhile, time-slice $(i-2)$ is written to disk and $(i+2)$ read from disk in its place. The update and the loop measurement on time-slice i are multitasked, using the four processors of Cray-XMP/48 at a sustained 490 MFlops. The speed-up, defined as the ratio (wall-clock time on 1 CPU)/(wall-clock time on 4 CPUs), is 3.77. To keep pace with the computation, I/O is handled by a 128 Megaword SSD (solid-state storage device) rather than by conventional disks. This reduces I/O wait time from several days to zero. The update routine uses the pseudo-heatbath algorithm.[15] The program can update a link in less than 6 μs, using 2 $SU(2)$ subgroups. In production mode, a slower update with three subgroups (or more), is preferred because it decorrelates more efficiently; the program then executes the update *and* the loop measurement in less than 12 μs/link.

I would like to mention a few algorithmic features which give the program its high performance. The Wilson loops are measured each sweep in parallel 24^2 planes, of orientation say (x, y). First y links are replaced by their meanfield value, as suggested in Ref. 16. This replacement brings

down the statistical noise by a large factor, $O(10)$. The mean-field averages are not computed by Monte Carlo as in Ref. 16, but by fast numerical integration, as developed in Ref. 6. At this stage all x links, except those in one band, are rotated to the identity by local gauge transformations in the (x, y) plane. Then products of adjacent y links are computed, and assembled by pairs to yield Wilson loops (whose x sides are all 1). The gauge transformations save almost an order of magnitude in the number of operations to perform. Of course, loops containing the x links not gauged to 1 must be left out (or computed separately), but we do not have time to measure all loops in all planes anyway. The update program was developed from Ref. 17, although the pseudo-heatbath routine was modified following Ref. 18.

It might be useful here to present a variation on the method of Ref. 18. The crux of the pseudo-heatbath method consists in generating a random variable between -1 and $+1$ with probability density proportional to $P(x) = \sqrt{1-x^2}\, e^{kx}$. The function

$$f(y) = \int_{-1}^{y} P(x)\, dx$$

is not easily invertible, so that the distribution $P(x)$ is generated in two steps. $P(x)$ is factored into $P_1(x) \cdot P_2(x)$, where distribution P_1 can be generated directly. Then the random sample distributed according to P_1 is passed through a filter of shape P_2, which should reject as small a fraction as possible. The decomposition suggested in Refs. 18, 19, and 20, which is appropriate for large values of k, is

$$\begin{aligned} P(x)\, dx &= \sqrt{1-x^2}\, e^{kx}\, dx \qquad (16) \\ &= 2e^k \underbrace{\sqrt{2-u^2}}_{P_2} \underbrace{e^{-ku^2} u^2}_{P_1}\, du, \qquad \text{where} \quad x \in [-1, +1],\ u \in [0, \sqrt{2}] \end{aligned}$$

P_1 can be generated from Gaussian distributions. But it can also be tabulated, noticing that

$$e^{-ku^2} u^2\, du = k^{-3/2} e^{-v^2} v^2\, dv \qquad \text{for} \quad u = vk^{-1/2} \qquad (17)$$

A single table can thus be constructed and used for all values of k: a uniform random entry into the table yields the random variable v, then u, by rescaling. The restriction $v \leqslant (2k)^{1/2}$ is always satisfied for the k range of interest. Tests using a 10,000-element table allowed a shorter update time, less than 20 μs/link on one processor ($\approx 5\mu$s/link on four CPUs). No bias could be detected in the measured plaquette after 30,000 sweeps on a 4^4 lattice. But the table was not implemented in production runs: we did not

want any bias toward disorder due to the discreteness of v, however small, to affect our high-statistics results.

I am deeply indebted to John Stack for the analysis of the results. He also provided the figures accompanying this section and the next. Results, which are still preliminary, have changed slightly since the Berkeley conference last May,[21] because more statistics have been accumulating. Figure 2 shows the quality of the fit ln $W(R, T)$ versus T for R up to 8, and Fig. 3 shows $V(R)$ versus R, both obtained from the $24^3 \times 48$ lattice at $\beta = 6.3$. Similar results have been obtained on the 16^4 lattice at $\beta = 6.0$. The

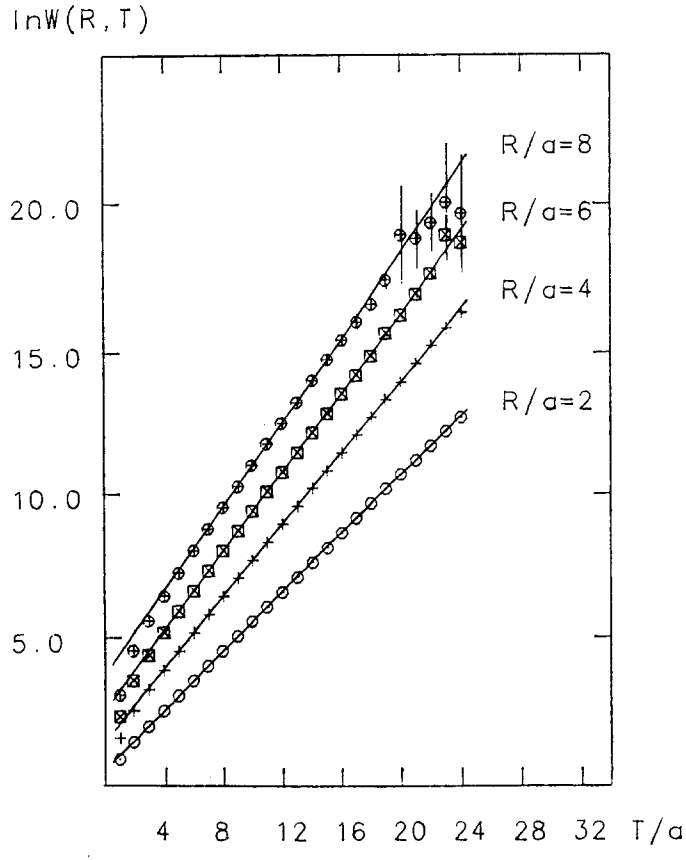

Fig. 2. Logarithms of the measured Wilson loops $W(R, T)$ versus T, on a $24^3 \times 48$ lattice. The slope for large T measures the potential $V(R)$.

favored coefficients of the Coulomb (α) and linear (σa^2) terms in fits to $V(R)$ seem to be

$$\begin{aligned} &\beta = 6.0 && \alpha \sim -.335 && \sigma a^2 \sim .046 && \text{ie.} && \frac{\sigma^{1/2}}{\Lambda_L} \sim 92 \\ &(R = 2 \text{ to } 6) \\ &\beta = 6.3 && \alpha \sim -.34 && \sigma a^2 \sim .0173 && \text{ie.} && \frac{\sigma^{1/2}}{\Lambda_L} \sim 79 \\ &(R = 3 \text{ to } 8) \end{aligned} \tag{18}$$

The Coulomb terms at both β values are nicely consistent with each other, but rather far from the value $-\pi/12 \approx -.26$ expected when probing the

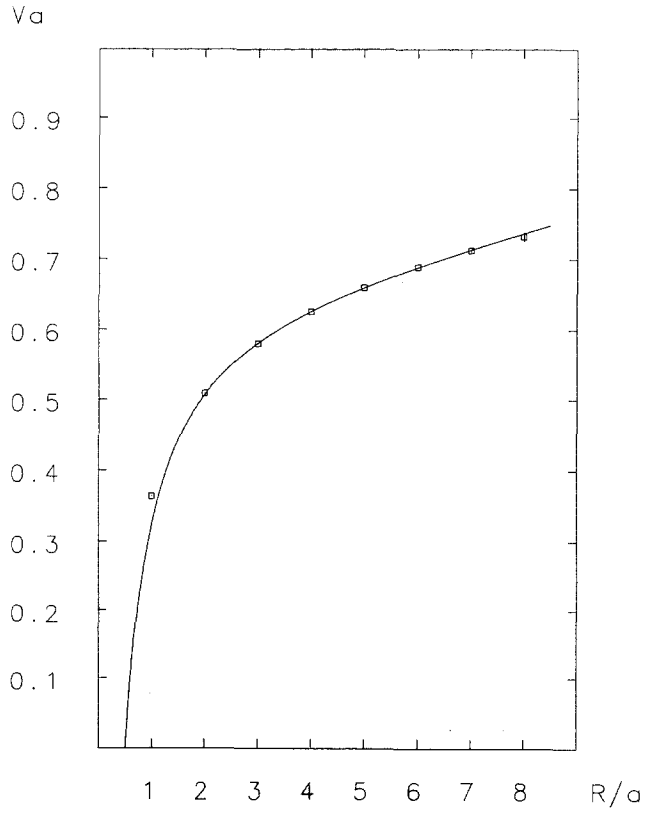

Fig. 3. Potential $V(R)$ extracted from the fitting procedure of Fig. 2. The solid line is a (Coulomb + linear) fit to the points $R = 3$ to 8.

large R region.[22] The values for the string tension hit a new low in the Hubble-constant-like history of that quantity. It is also noteworthy that they show a violation of asymptotic scaling between $\beta = 6.0$ and 6.3. This violation is consistent with early tests conducted on loop ratios as in Ref. 23: the scaling factor between the two values of β is greater than 1.5, whereas asymptotic scaling (13) predicts ≈ 1.4. Such evidence would support the project of simulating a larger lattice at a slightly larger β—with the next Cray machine.

3. $SU(3)$ SPIN-DEPENDENT POTENTIALS

The work reviewed in this section results from a continuing collaboration with John Stack. Other groups have been studying spin-dependent effects in the $SU(2)$[24] and $U(1)$[25] lattice gauge theories. In all cases the approach is based on the $1/m$ expansion of Eichten and Feinberg.[26] I will only try to make plausible the formulas we used. Reference 27 gives a pedagogical introduction to this subject.

We want to investigate spin–spin and spin–orbit potentials. The spin of a static quark at position $(\vec{x}, t)$ will orient the chromomagnetic field $\vec{B}$ at that location. Similarly, an elementary displacement dx_i of this quark can be decomposed into a series of kinks of its world-line, probing the field strength tensor F_{0i} (i.e., the chromoelectric field E_i) at successive times. Perhaps then it will come as no surprise that we must measure elongated loops with insertions on the long (timelike) sides. Two magnetic-type insertions (probing $\vec{B}$) will contribute to spin–spin potentials. An electric and a magnetic insertion will contribute to spin–orbit potentials. In the accepted terminology, the potentials V_1 to V_4 can be measured by

$$\begin{aligned}
\hat{R}_k \frac{dV_1}{dR} &= \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} \int_{-T/2}^{T/2} dt_1\, dt_2\, \frac{t_1 - t_2}{2}\, \varepsilon_{ijk}\, \frac{\langle E_i(\vec{0}, t_1)\, B_j(\vec{0}, t_2)\rangle}{\langle W(R, T)\rangle} \\
\hat{R}_k \frac{dV_2}{dR} &= \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} \int_{-T/2}^{T/2} dt_1\, dt_2\, \frac{t_1 - t_2}{2}\, \varepsilon_{ijk}\, \frac{\langle E_i(\vec{0}, t_i)\, B_j(\vec{R}, t_2)\rangle}{\langle W(R, T)\rangle} \\
&\left(\hat{R}_i \hat{R}_j - \frac{1}{3}\delta_{ij}\right) V_3 + \frac{1}{3}\delta_{ij} V_4 \\
&\quad = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} \int_{-T/2}^{T/2} dt_1\, dt_2\, \frac{\langle B_i(\vec{0}, t_1)\, B_j(\vec{R}, t_2)\rangle}{\langle W(R, T)\rangle}
\end{aligned} \tag{19}$$

In (19), $W(R, T)$ is the usual rectangular $R \times T$ loop. $\langle E_i(\vec{0}, t_1)\, B_j(\vec{R}, t_2)\rangle$ means the average value of the loop of size $R \times T$, having an insertion probing E_i at location $(\vec{0}, t_1)$ and one probing B_j at location $(\vec{R}, t_2)$. There

is a lot of arbitrariness in the insertion chosen to probe $\vec{E}$ or $\vec{B}$ on the lattice. Various choices, all equivalent in the continuum, introduce important changes in the normalization of $V_1 - V_4$ and even distort the potentials for small or moderate R. Finally, renormalization of these insertions is not under control, so that the overall scale of $V_1 - V_4$ cannot be reliably determined—as yet—from lattice studies. In our case we chose to estimate the field strength by

$$F_{\alpha\beta}(\vec{R}, t) \sim \frac{1}{2ia^2}(M_{\alpha\beta} - M_{\alpha\beta}^{+}) \tag{20}$$

where $M_{\alpha\beta}$ is the product of four links around a plaquette in plane (α, β) attached by its corner to the point $(\vec{R}, t)$.

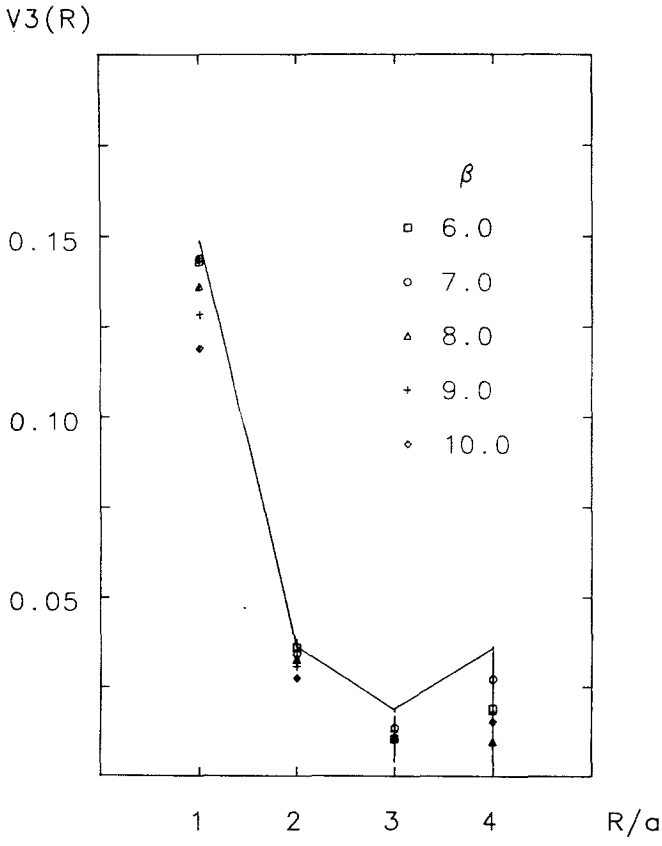

Fig. 4. Tensor (spin–spin) potential measured on a $6^3 \times 12$ lattice at $\beta = 6$, 7, 8, 9, 10. The solid line represents perturbation theory at $\beta = 10$. Taken from Ref. 28.

For all the above reasons I think that only the qualitative features of the results obtained so far are relevant. A comparison with one-gluon exchange (modulo a poorly controlled constant factor) is particularly instructive. Figures 4 and 5, taken from Ref. 28, show our measurements of V_3 and dV_2/dR on a $6^3 \times 12$ lattice over a wide range of β. The solid line corresponds to one-gluon exchange. The most significant hint of nonperturbative effects occurs at $\beta = 6.0$ for dV_2/dR. Even so the statistical errors are quite large enough to accommodate perturbation theory. A more accurate check is provided by the integrand itself

$$f_k(t_1 - t_2) \equiv \varepsilon_{ijk} \langle E_i(\vec{0}, t_1)\, B_j(\vec{R}, t_2) \rangle \tag{21}$$

Figure 6 was obtained by measurements on previously stored $24^3 \times 48$ configurations. The agreement with perturbation theory for this spin–orbit

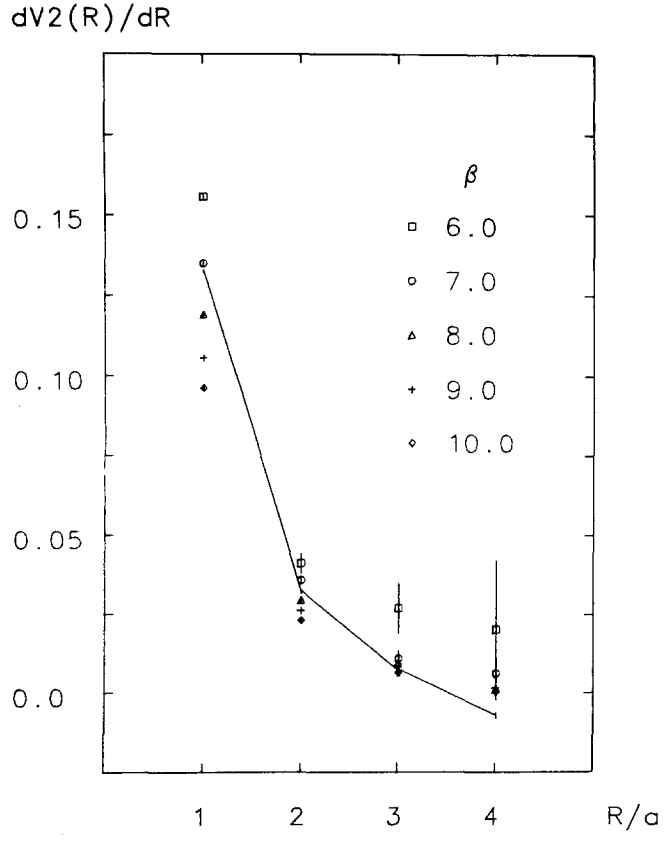

Fig. 5. Spin-orbit potential measured on a $6^3 \times 12$ lattice at $\beta = 6$, 7, 8, 9, 10. The solid line represents perturbation theory at $\beta = 10$. Taken from Ref. 28.

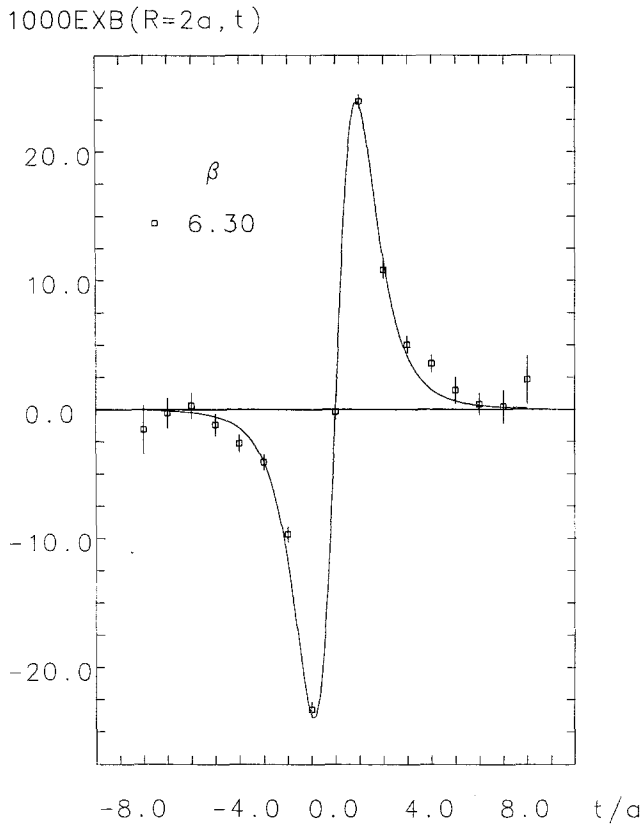

Fig. 6. $E \times B$ versus t at $R = 2a$, measured on a $24^3 \times 48$ lattice at $\beta = 6.3$. The solid line represents perturbation theory.

term is rather striking. A good knowledge of the integrand should also help control the error introduced by the finite (small) T bound to the integrals in (19). More loop measurements are in progress, using techniques explained in Section 2, like mean-field averaging of timelike links and even field insertions. Much more work is needed to go beyond Ref. 28 and start probing R regions where the choice of insertion becomes irrelevant.

4. DYNAMICAL FERMIONS

I will now review results from an ongoing collaboration with I. O. Stamatescu. There is a well-known problem with simulating the complete QCD theory. The anticommuting property of the quark fields translates, after proper integration, into an effective action

$$S_{\text{eff}} = S_G - N_f \ln(\det W) \tag{22}$$

where S_G is the local Yang–Mills action, N_f the number of dynamical (light) quark flavors, and W a discretized version of the Dirac operator $(\not{D}+m)$. The presence of the determinant makes (22) nonlocal. For each Monte Carlo update of one link variable, the evaluation of the determinant will require an amount of computation at least proportional to the volume V of the lattice (if not to V^2), so that the algorithm for dynamical fermions is slower than the pure gauge algorithm by a factor $O(V)$ at least. Or so it seems.

The two most popular methods to keep the computations tractable are the microcanonical[29] and pseudo-fermion[30] methods. The approach taken in Chippewa is a different one, conceived as an improvement over the pseudo-fermion method. It is also characterized by the choice of the Wilson formulation for the quark degrees of freedom, as opposed to the more popular staggered fermion formulation. These two formulations are equivalent in the continuum, but lead to different operators W in (22). We chose Wilson fermions because they are easier to understand and because they may reflect continuum physics at a coarser lattice spacing than staggered fermions. The algorithm itself is of the Metropolis type, as in the pseudo-fermion method. Each link update requires the comparison of $e^{-[S(u')-S(u)]}$ with a random number. The crucial step is the evaluation of the determinant ratio

$$\rho \equiv \frac{\det W(u')}{\det W(u)} \tag{23}$$

In either case this is done by an inner Monte Carlo on bosonic degrees of freedom (the "pseudo-fermions"), since

$$\rho^{-2} = \frac{\det W^+ W(u)}{\det W^+ W(u')} = \frac{\int d\varphi_i^*\, d\varphi_j\, e^{-\sum_{i,j} \varphi_i^* W^+ W(u')\varphi_j}}{\int d\varphi_i^*\, d\varphi_j\, e^{-\sum_{i,j} \varphi_i^* W^+ W(u)\varphi_j}} = \langle e^{-\Delta S_\varphi} \rangle_u \tag{24}$$

where

$$\Delta S_\varphi \equiv \sum_{i,j} \varphi_i^* [W^+ W(u') - W^+ W(u)]\, \varphi_j$$

and the observable $e^{-\Delta S_\varphi}$ is evaluated in the old background gauge configuration.

The two main differences between our algorithm, described in Ref. 31, and the standard pseudo-fermionic one, are the following. The standard algorithm linearizes the exponential and takes

$$\rho^{-2} \approx e^{-\langle \Delta S_\varphi \rangle_u} \tag{25}$$

whereas because of the convexity of the exponential (Janssen's inequality)

$$e^{-\langle x \rangle} \leqslant \langle e^{-x} \rangle \tag{26}$$

approximation (25) leads to a *systematic underestimate* of the fermionic effects. We keep the original observable (24) to avoid this bias. The other main ingredient is due to Kuti and Kennedy.[32] The Monte Carlo integration in (24) leaves us with a very noisy estimate $\tilde{\rho}^{-2}$ for ρ^{-2} unless the number of pseudo-fermionic Monte Carlo sweeps is of $O(10^4)$ or more. Therefore it is judicious to modify the Metropolis algorithm so that it allows for noise provided the estimator is unbiased. This way the number of pseudo-fermionic sweeps can be reduced to $O(10)$, as we shall see. Final touches in our algorithm include mean-field averaging on some pseudo-fermions and sparse grouping of the links for a single pseudo-fermionic Monte Carlo integration. Reference 31 provides complete details. Our emphasis is on keeping the remaining approximations under control by monitoring the leading error terms. For instance, the error due to grouping links for pseudo-fermion integration is monitored and always remains in the noise ($\lesssim 10\%$) in the following simulations.

Checks were first conducted by comparing our method with an exact calculation on a 2^4 lattice. Then simulations on a $4^3 \times 2$ lattice with three light flavors revealed a very smooth croosover as a function of β.[33] There was no indication of a phase transition as observed with other methods on larger lattices. We want of course to simulate larger lattices too. As a first step we ran more tests on a 4^4 lattice with three light flavors. We evolve in a four-dimensional parameter space: the bare coupling constant β, the hopping parameter k, related to the bare quark mass m_b by

$$k = \frac{1}{2(dr + m_b)} \tag{27}$$

the Wilson parameter r, and the number N_{pf} of pseudo-fermionic sweeps per link update (actually N_{pf1} sweeps to equilibrate the pseudo-fermion fields, and N_{pf2} sweeps to take measurements). N_{pf} should be an irrelevant parameter; r also, but only in the continuum limit.

To check the dependence on N_{pf} we ran simulations at $\beta = 5.0$, $k = .12$, $r = 1$, with the results

N_{pf} ($N_{\text{pf1}} + N_{\text{pf2}}$)	$\square$	$\bar{\psi}\psi$	Re (Polyakov)
$2 + 10$	.4074	2.849	.010
	(9)	(1)	(2)
$5 + 10$	.4083	2.849	.008
	(10)	(1)	(3)
$10 + 20$	.4069	2.847	.006
	(11)	(1)	(4)
$20 + 20$	.4081	2.849	.010
	(9)	(1)	(3)

As we changed N_{pf} we changed the average acceptance rate α of the Kuti–Metropolis scheme to keep the probability for the estimator $\tilde{\rho}^{(-N_f/2)}$ of $\rho^{(-N_f/2)}$ to exceed $1/\alpha$ less than $O(3\times 10^{-3})$. No trend can be detected for this set (β, k, r).

The β dependence was investigated by the comparing $\beta = 5.0$, 5.4, and 5.6 at $k = .12$, $r = 1$, $N_{pf} = (5 + 10)$ sweeps.

β	$\square_{YM}$	$\square$	$\bar{\psi}\psi$	Re (Polyakov)
5.0	.4005	.408	2.849	.008
	(10)	(1)	(1)	(3)
5.4	.473	.483	2.839	.032
	(1)	(2)	(1)	(8)
5.6	.525	.557	2.818	.101
	(1)	(1)	(2)	(34)

Although we stay below the pure gauge phase transition ($\beta \sim 5.7$), the gradual ordering effect of the fermions is clearly visible on the Polyakov loop.

The r dependence was investigated crudely with the two ideas:

(i) keep $k = .12$ constant

(ii) keep the bare mass $m_b = \frac{1}{6}$ constant, as per (27)

The results are, at $\beta = 5.0$, with $(5 + 10)$ sweeps

r	(i)			(ii)			
	$\square$	$\bar{\psi}\psi$	Re (Pol)	k	$\square$	$\bar{\psi}\psi$	Re (Pol)
1.	.408	2.849	.008	.12	.408	2.849	.008
	(1)	(1)	(3)		(1)	(1)	(3)
.7	.405	2.710	.005	.1685	.415	3.569	.017
	(1)	(1)	(2)		(2)	(2)	(3)
.4	.402	2.632	.003	.2830	.448	4.567	.040
	(1)	(1)	(2)		(3)	(3)	(7)
.1	.4035	2.600	.000	.8824	—	—	—
	(10)	(1)	(3)				

It is clear that in (i) a decrease in r increases the bare quark mass, so that the fermionic effects drop to zero. And in (ii) the unwanted 15 fermionic modes go down in mass with r, so that the cumulated fermionic effects rise dramatically (as would be expected from 48 flavors). This exercise emphasizes the main difficulty of Wilson fermions, that is, knowing the

mass of the dynamical quarks during the simulation. Future work including hadron spectroscopy should solve that problem.

Now that we have, I hope, convinced you of the validity and the practicality of our algorithm, we are pursuing work on an $8^3 \times 4$ lattice to study the temperature behavior of the theory and the $q\bar{q}$ potential.

5. CONCLUSION

I hope the work reviewed here will provide more impetus to the field of QCD simulations. A concrete example is given by Steve Sharpe at this conference[34]: he and his collaborators used the $24^3 \times 48$ configurations generated in Chippewa to study the properties of their $3^{1/2}$ blocking transformation. I would be even more pleased if Monte Carlo physicists in other fields would now contemplate problems with 21 million degrees of freedom—or more. In any case, my goal was to try and impress you with all that can be done on a Cray, and is being done at Cray Research itself. I have been fully rewarded with the question: "How does the size of your basic research group compare with that of Bell Labs?"

ACKNOWLEDGMENTS

I would like to thank the organizers, especially Gerry Guralnik, for this exciting interdisciplinary conference. Thanks to John Larson for reading the manuscript, and particularly to Jim Gubernatis for his patience while awaiting this paper.

REFERENCES

1. See Steve Chen, these proceedings.
2. B. Berg and A. Billoire, *Phys. Lett. B* **113**:65 (1982) K. Ishikawa, M. Teper, and G. Schierholz, *Phys. Lett. B* **116**:429 (1982).
3. Ph. de Forcrand, Ec. Polytechnique preprint A615.0784, unpublished.
4. Ph. de Forcrand, G. Schierholz, H. Schneider, and M. Teper, CERN preprint TH-4167/85, to appear in Zeitschrift für Physik.
5. C. Michael and I. Teasdale, *Nucl. Phys. B* **215**:433 (1983).
6. Ph. de Forcrand and C. Roiesnel, *Phys. Lett. B* **151**:77 (1985).
7. Ph. de Forcrand, G. Schierholz, H. Schneider, and M. Teper, *Phys. Lett. B* **160**:137 (1985).
8. Ph. de Forcrand, G. Schierholz, H. Schneider, and M. Teper, *Phys. Lett. B* **152**:107 (1985).
9. M. Luscher, DESY preprint 83-116.
10. G. Munster, DESY preprint 84-088.
11. Ph. de Forcrand, in writing.
12. K. C. Bowler, P. Hasenfratz, U. Heller, F. Karsch, R. D. Kenway, I. Montvay, G. S. Pawley, J. Smit, and D. J. Wallace, Amsterdam preprint ITFA-85-07.
13. D. Barkai, K. J. M. Moriarty, and C. Rebbi, *Phys. Rev. D* **30**:1292 (1984); *Phys. Rev. D* **30**:2201 (1984).

14. Ph. de Forcrand and J. L. Larson, CRAY Channels, winter 1985.
15. N. Cabibbo and E. Marinari, *Phys. Lett. B* **119**:387 (1982).
16. G. Parisi, R. Petronzio, and F. Rapuano, *Phys. Lett. B* **128**:418 (1983).
17. Ph. de Forcrand, D. Lellouch, and C. Roiesnel, *J. Comput. Phys.* **59**:324 (1985).
18. A. D. Kennedy and B. J. Pendleton, Santa Barbara preprint NSF-ITP-85-07.
19. K. Fabricius and O. Haan, *Phys. Lett. B* **143**:459 (1984).
20. W. Celmaster, private communication.
21. J. D. Stack, proceedings of the conference "Quark Confinement and Liberation: Numerical Results and Theory," Berkeley, June 1985.
22. M. Luscher, K. Symanzik, and P. Weisz, *Nucl. Phys. B* **173**:365 (1980).
23. A. Hasenfratz, P. Hasenfratz, U. Heller, and F. Karsch, *Phys. Lett. B* 143:**193** (1984).
24. C. Michael and P. E. L. Rakow, *Nucl. Phys. B* **256**:640 (1985).
25. M. Campostrini, *Nucl. Phys. B* **256**:717 (1985).
26. E. Eichten and F. Feinberg, *Phys. Rev. D* **23**:2724 (1981).
27. M. E. Peskin, proceedings of the 11th SLAC Summer Institute (1983).
28. Ph. de Forcrand and J. D. Stack, *Phys. Rev. Lett.* **55**:1254 (1985).
29. J. Polonyi and H. W. Wylf, *Phys. Rev. Lett.* **51**:2257 (1983).
30. F. Fucito, E. Marinari, G. Parisi, and C. Rebbi, *Nucl. Phys. B* **180**:369 (1981).
31. Ph. de Forcrand and I. O. Stamatescu, *Nucl. Phys. B* **261**:613 (1985).
32. A. D. Kennedy and J. Kuti, Santa Barbara preprint NSF-ITP-85-21.
33. I. O. Stamatescu, proceedings of the conference "Advances in Lattice Gauge Theories," Tallahassee, Florida, April 1985.
34. See Stephen S. Sharpe, these proceedings.