Lattice QCD on the BGQ: Achieving 1 PFlops Production Jobs

Sequoia at LLNL

Through advances in computer hardware and software, production lattice QCD jobs run by the RBC and UKQCD Collaborations on BGQ installations are sustaining 1 PFlops. The largest computational cost is in the solution of the Dirac equation (a linear equation) in the presence of a fixed gauge background and we use Peter Boyle's (Univ. of Edinburgh) BAGEL based solver for this important calculation. Extensive multithreading via OpenMP in the rest of our code base (Chulwoo Jung, BNL and Hantao Yin, Columbia) has also been vital to achieving good performance. In addition, we employ a number of new deflation and variance reduction strategies in our calculations which give a further large speed-up (~ 10×) beyond the software improvements. With all of these techniques, we are able to simulate lattice QCD with physical pion masses in large volumes, of size (5.5 fm)^3. This has led to markedly reduced statistical and systematic errors for our results. With these advances, as well as new theoretical ideas, we are now begining production calculations for kaon decay involving disconnected quark diagrams, where the signal to noise ratio is much worse. These new calculations improve constraints on standard model physics.

Mira at ANL

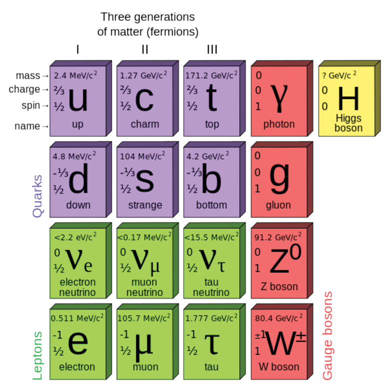
BGQ at BNL

Robert Mawhinney
Columbia University

The calculations reported here have been run on Mira at the ALCF, Sequoia at LLNL, the BGQ computers of the University of Edinburgh and the BGQ computers of BNL and the RBRC.

Quantum Chromodynamics

Known Elementary Particles

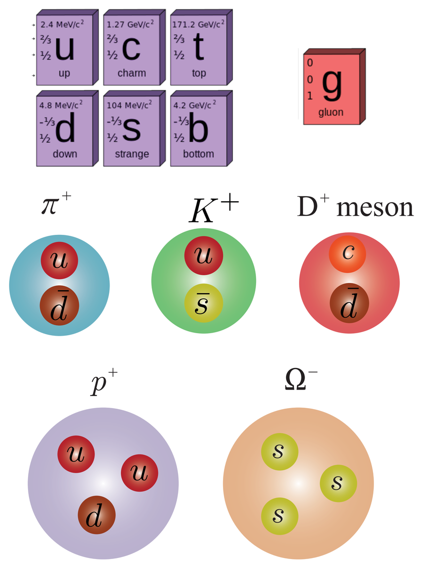


$m_u = 2.19 \pm 0.15\,\mathrm{MeV}$
$m_d = 4.67 \pm 0.20\,\mathrm{MeV}$
$m_s = 94 \pm 3\,\mathrm{MeV}$
$m_c = 1.275 \pm 0.025\,\mathrm{GeV}$
$m_b = 4.18 \pm 0.03\,\mathrm{GeV}$
$m_t = 173.5 \pm 0.6 \pm 0.8\,\mathrm{GeV}$

QCD

Theory of interactions of quarks

Interactions mediated by gluons

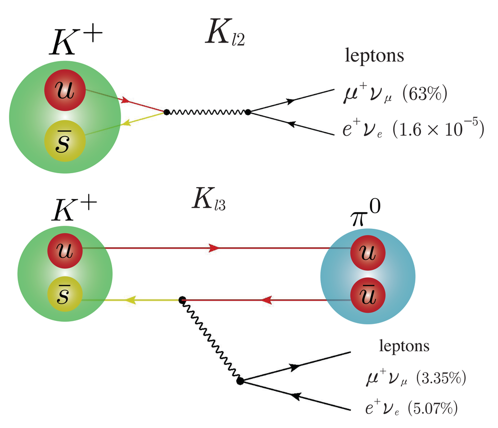


QCD + Electroweak Interactions Produce Particle Decays

- Decays of quarks via weak interactions predicted by Standard Model.
- Experiments measure decays of hadrons
- Standard Model quark decays involve elements of a 3 by 3 unitary matrix, the CKM matrix, described by 4 parameters

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}$$

- Relating standard model parameters to measured results generically requires knowing the value of a quark process inside a hadron (a matrix element) such as as f+(0)



For K_{l3} we have: $\Gamma_{K\to\pi l\nu} = C_K^2 \frac{G_F^2 m_K^5}{192\pi^3} I\, S_{EW}[1+2\Delta_{SU(2)}+2\Delta_{EM}]|V_{us}|^2|f_+(0)|^2$

- QCD Simulations are done in a four (or five) dimensional box, with O(50-100) grid points in each dimension
- Sample the phase space of the system via Monte Carlo, following the Euclidean space Feynamn path integral. The continuum, Minkowski space path integral is

$$Z = \int [dA] \prod_{i=1}^{3} \det\left[D(A,g_0,m_0^i)\right] \exp\left\{-\frac{i}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu + g_0 f^{abc}A_\mu^b A_\nu^c)^2\right\}$$
$$D(A,g_0,m_0^i) \equiv i\gamma^\mu(\partial_\mu - igA_\mu^a t^a/2) - m_0^i$$

- Only parameters are an input coupling and three (or four) light quark masses.
- Gluon self-interactions yields a very non-linear system.

Numerical Path Integral Including the Fermionic Effects of Quarks

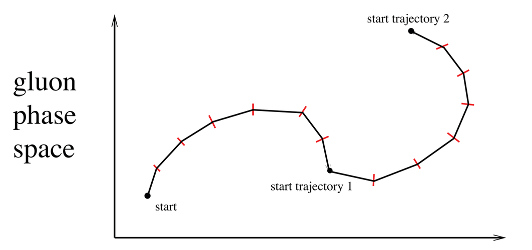
- Require positive definite measure for Monte Carlo

$$Z = \int [dA] \prod_{i=u,d,s} \left(\frac{\det[D^\dagger(m_i)D(m_i)]}{\det[D^\dagger(1)D(1)]}\right)^{1/2} \exp\{-S_g(A)\}$$
$$= \int [dA][d\Phi_i^\dagger][d\Phi_i] \exp\left\{-S_g(A) - \sum_{i=u,d,s} \Phi_i^\dagger [D(1)\,D^\dagger(m_i)^{-1}\,D(m_i)^{-1}\,D^\dagger(1)]^{1/2}\Phi_i\right\}$$

- From $(A, \Phi, \Pi)_{ini}$ use molecular dynamics to move to $(A, \Phi, \Pi)_{fin}$
- Then do Monte Carlo accept/reject
- Rational Hybrid Monte Carlo of Clark and Kennedy approximates

$$\mathcal{M}^{1/2}\Phi = \left(\alpha_0 + \sum_{k=1}^{n} \frac{\alpha_k}{\mathcal{M}+\beta_k}\right)\Phi$$

- Use multimass Krylov space solver (CG) to do all poles at once



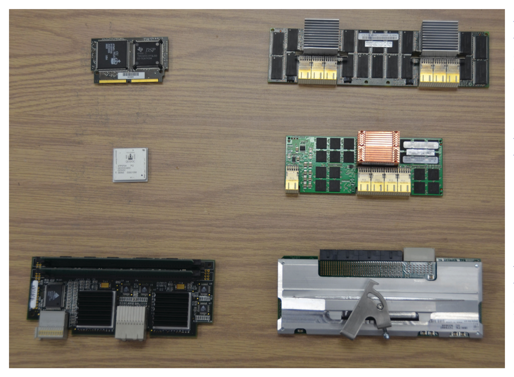
Computers, Algorithms and Software

- Physicists at Columbia built the 8,000 and 12,000 node QCDSP computers for QCD in 1997
- Columbia, RBRC and UK physicists built 3, 12,000 node QCDOC computers for QCD in 2004, working with IBM on the custom ASIC. These machines were built while IBM was producing the BG/L computer.
- Columbia (Christ, Kim) and Edinburgh (Boyle) physicists worked on the design of the BG/Q computer.
- Boyle has an extensively optimzed linear solver for the lattice Dirac equation that makes extensive use of the hardware features of the BG/Q. This solver was used in the chip design and testing stages to help validate the design.

Computers

Columbia/RBRC QCDSP 1998-2005 0.050 GFlops/node

Columbia/RBRC/ UKQCD QCDOC 2005-2011 0.8 GFlops/node



IBM BGL 2005-2013 2.8 GFlops/node

IBM BGP 2007- 13.6 GFlops/node

IBM BGQ 2012- 200 GFlops/node

RBC/UKQCD have production jobs on the Argonne ALCF BGQ that sustain 1 PFlops on 32 racks = 32k nodes = 0.5 M cores.

This performance comes from very carefully tuned assembly code on BGQ, produced by Peter Boyle, using his BAGEL code generator

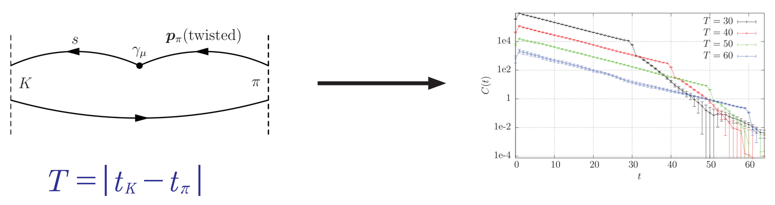
Algorithms for Gauge Field Production

- Producing gauge fields:
  - Use classical molecular dynamics to move through gauge field space
  - Quark loops give back reaction on gauge fields by solving Dirac equation
  - Hasenbusch mass preconditioning allows tuning back reaction

$$\det[D(m)] = \underbrace{\frac{\det[D(m)]}{\det[D(m_1)]}}_{\text{For } m \approx m_1 \text{ gives small force but expensive to calculate}} \times \underbrace{\frac{\det[D(m_1)]}{\det[D(m_2)]}}_{\text{Control force size from } m_1 \text{ and } m_2\text{, less expensive to calculate}} \cdots \det[D(m_s)]$$

  - RBC/UKQCD uses 7 levels of intermediate masses
  - Integrate different d.o.f on different time scales (Sexton-Weingarten integrators)
  - Use higher order integrators, currently RBC/UKQCD use force gradient, O(dt^4)
- These are giving 10-100× speed-up over a decade ago.
  - Hard to be completely quantitative here, since without speeds-ups, we could not even try current simulations

Algorithms for Measurements

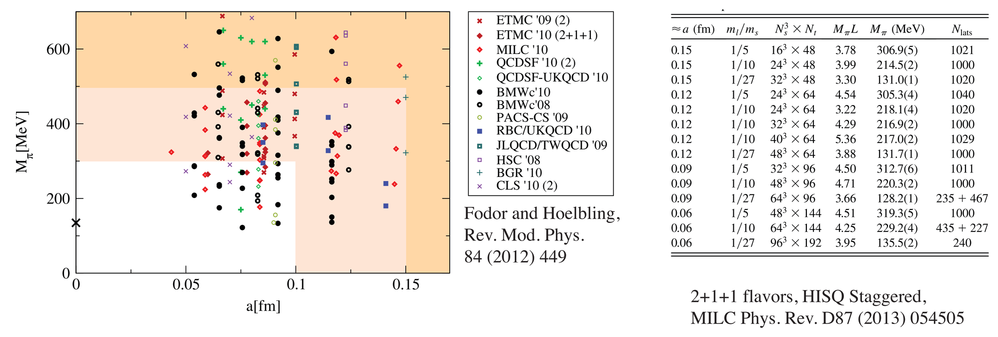


- Time translated the n-point function, on a fixed background gauge field, are sufficiently decorrelated (independent enough) to make them worth calculating
- This means many solutions of the Dirac equation D[U$_\mu$] Ψ = s for fixed U$_\mu$
- Calculating eigenvectors of D[U$_\mu$] with small eigenvalues (low-modes) speeds up subsequent solves. Can be done with EigCG or Lanczos algorithms
- Alternatives for Wilson fermions are domain decomposition and multigrid, giving similar speed-up with smaller memory requirements.
- Further improvement from all-mode-averaging of Blum, Izubuchi and Shintani
  - Separates measurements into expensive parts, with small statistical errors after a few measurements, and inexpensive parts, where many measurements are needed.
- These improvements make measurements ~ 10× faster than a year ago.
  - This particular method takes substantial computer memory O(100 TBytes)
  - No checkpointing is done, so computer must work reliabilty for duration of job
  - Our smaller volume simulations takes 6 days on 1 rack of BGQ
  - Our larger volume simulations take 6 hours on 32 BGQ racks at 1 PFlops.
  - This speed-up was accomplished through the work of Hantao Yin at Columbia.
- Boyle has developed a Heirarchically Deflated Conjugate Gradient (HDCG) to markedly decrease the memory footprint, since eigenvectors are not needed, while also needing fewer iterations to converge. Production level testing is currently underway by Chulwoo Jung of BNL

Physics Results and Prospects

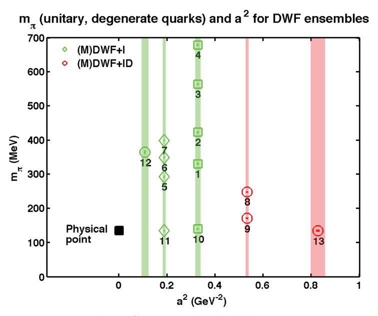
Major Development: Ensembles with Physical Quark Masses

- Given advances in hardware, software and algorithms, lattice QCD physicists are now regularly doing simulations with physically light pions, a major goal since the beginning of numerical QCD in the early 1980's.



Fodor and Hoelbling, Rev. Mod. Phys. 84 (2012) 449

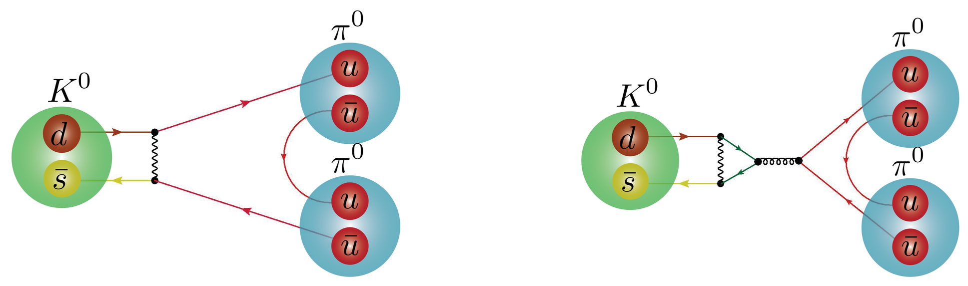
2+1+1 flavors, HISQ Staggered, MILC Phys. Rev. D87 (2013) 054505



2+1 flavors, (M)DWF RBC and UKQCD Collaborations

Large Vol. Ensembles with Physical Quark Masses
- BMW Hex-smeared clover fermions
- MILC HISQ Staggered fermions
- RBC/UKQCD DWF fermions

- The RBC and UKQCD collaborations, using the numerically expensive domain wall formulation, now have two large simulations at different lattice spacings with physically light quarks.
  - Domain wall fermions preserve all the continuum symmetries of QCD at finite lattice spacing
  - This is vital for measurements of many observables
- For the first time, physical results have been produced for a process with 2 particles in the final state, the amplitude for a kaon to decay to 2 pions (K -> ππ) with the pions in an isospin 2 final state. Physical Review Letters, 108 (2012) 141601 and Physical Review D 86 (2012) 074513.



- Previous results

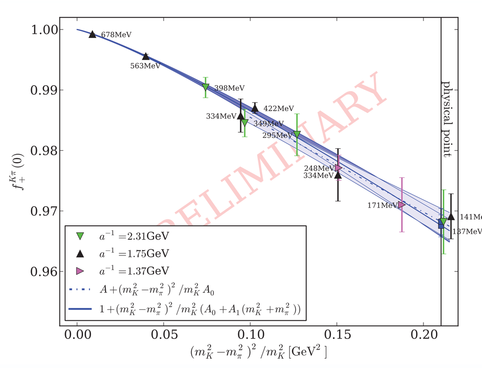
$$\mathrm{Re}\,A_2 = (1.381 \pm 0.046_{\rm stat} \pm 0.135_{{\rm sys\,no}\,a^2} \pm 0.207_{a^2}) \times 10^{-8}\ \mathrm{GeV}$$
$$\mathrm{Im}\,A_2 = -(6.54 \pm 0.46_{\rm stat} \pm 0.72_{{\rm sys\,no}\,a^2} \pm 0.98_{a^2}) \times 10^{-13}\ \mathrm{GeV}$$

- New, preliminary results

$$\mathrm{Re}\,A_2 = (1.424 \pm 0.041_{\rm stat} \pm 0.135_{{\rm sys\,no}\,a^2} \pm 0.0_{a^2}) \times 10^{-8}\ \mathrm{GeV}$$
$$\mathrm{Im}\,A_2 = -(5.88 \pm 0.16_{\rm stat} \pm 0.72_{{\rm sys\,no}\,a^2} \pm 0.0_{a^2}) \times 10^{-13}\ \mathrm{GeV}$$
(Preliminary)

- These new measurements are reducing the errors on measurements of properties of decay of kaons into a pion plus leptons, called Kl3 decays

- $\langle\pi|V|K\rangle \to f_+^{K\pi}(q^2=0)$
- part. twisted boundary conditions
- $N_f = 2+1$ domain wall fermions
- a^2-scaling study (0.09fm-0.14fm) → tiny cut-off effects
- physical point simulation
  m_π: 171–670MeV → arXiv:1305.7217
  137–670MeV → PRELIMINARY
  - polynomial ansatz describes data over entire mass range
  - phys. point data eliminates large systematic due to χ extrapolation



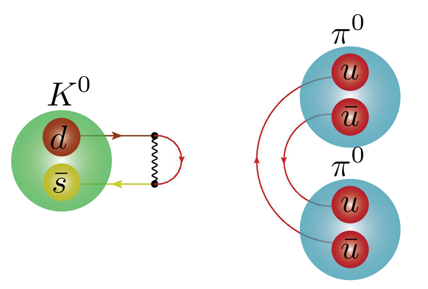
$$f_+^{K\pi}(0) = 0.9670(20)\left({}^{+0}_{-42}\right)_{m_q}(7)_{\rm FSE}(17)_a$$
$$|V_{us}| = 0.2237\,(7)\left({}^{+10}_{-0}\right)_{m_q}(2)_{\rm FSE}(4)_a$$

→≈0 with phys. point data

precision $\lesssim 0.3\%$ feasible!

Challenges and Prospects

- Simulations are just beginning for the more difficult case of K -> ππ decays with the pions in an isospin 0 final state. This process includes disconnected diagrams, which are much noisier



- HDCG or other linear solver improvements could be very helpful for these measurements
- Simulations with a smaller underlying lattice spacing pose problems with the rate of sampling of the path integral phase space. Greg McGlynn at Columbia has made some progress on this important question.