Accelerating lattice QCD calculations with G-parity Boundary Conditions

2024 SciDAC-5 Principal Investigator (PI) Meeting

Multiscale acceleration: Powering future discoveries in High Energy Physics

PI: Peter Boyle

Christopher Kelly

07/16/24

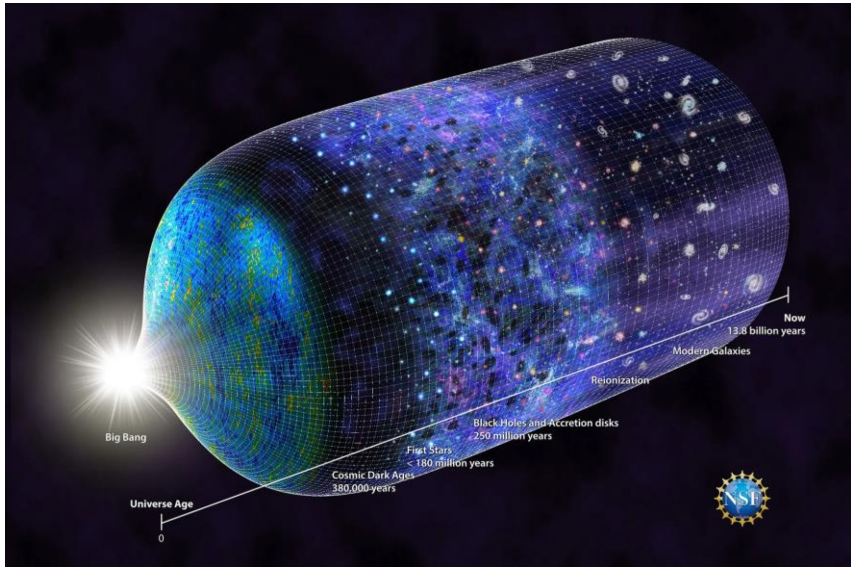
Brookhaven National Laboratory

U.S. DEPARTMENT OF ENERGY

@BrookhavenLab

Motivation

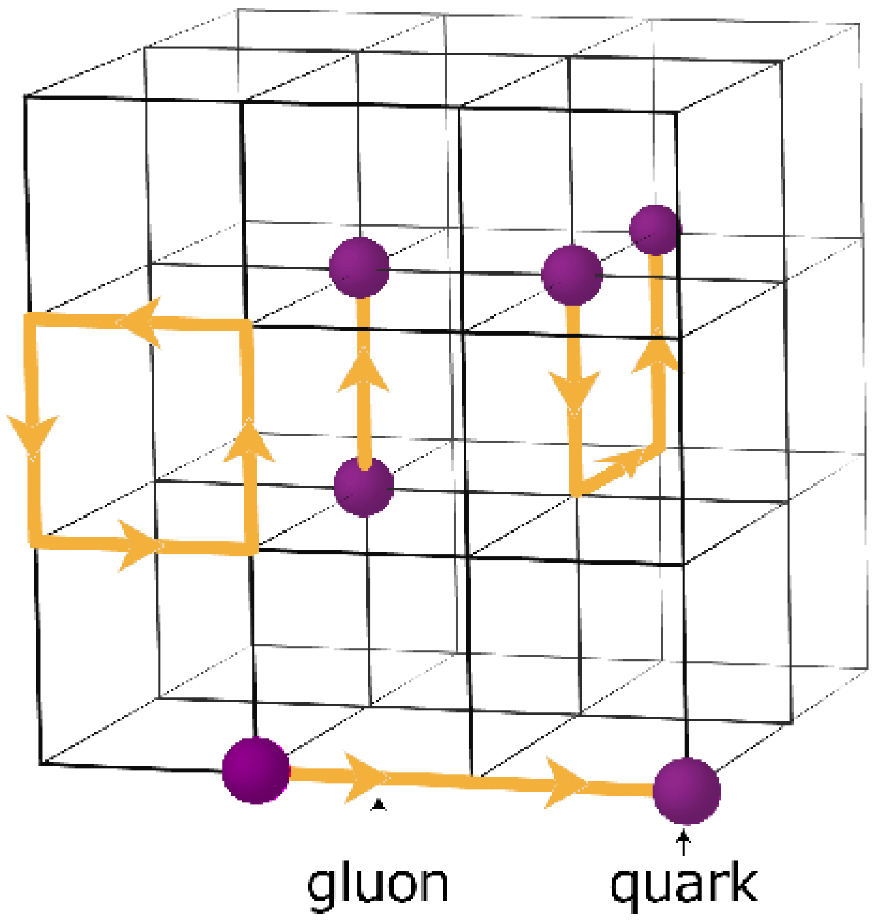
- The RBC & UKQCD collaborations have a long-running series of lattice QCD calculations of **Direct CP-violation in $K \to \pi\pi$ decays**.
- Comparing theory to experiment allows us to probe for new physics.
- *Such new sources of CPV are needed to explain the dominance of matter over antimatter in the Universe.*



Brookhaven National Laboratory

G-parity Boundary Conditions

- Allowed finite-volume momenta are dictated by boundary conditions (BCs)
- To ensure dominance of the physical, energy-conserving decay amplitude we pioneered the use of **G-parity BCs**.
- *Pion states become antiperiodic*, hence have a ground-state energy tunable by varying the box size.
- This advantage comes with the downside of *greatly increasing numerical cost*.
- <u>Under SciDAC-5 we have developed a series of new algorithms that significantly reduce this cost to a level comparable with a standard lattice calculation*</u>.



*assuming the cost of ensemble generation is likewise amortized

Brookhaven National Laboratory

Two quarks for Muster Mark!

- Under the G-parity operation, pions simply pick up a sign

$$\hat{G}\pi^{\pm,0}\hat{G}^{-1} = -\pi^{\pm,0}$$

- <u>But we cannot directly control the BCs of composite particles</u>, only those of their constituent *quarks*.
- *G-parity on quarks is more complicated!*

$$\hat{G}\begin{pmatrix} u \\ d \end{pmatrix}\hat{G}^{-1} = \begin{pmatrix} -C\bar{d}^T \\ C\bar{u}^T \end{pmatrix}$$

up quark

down quark

4x4 complex (spin) matrix

antiquark!

- Convenient to reformulate:

$$\psi \equiv \begin{pmatrix} d \\ C\bar{u}^T \end{pmatrix}$$

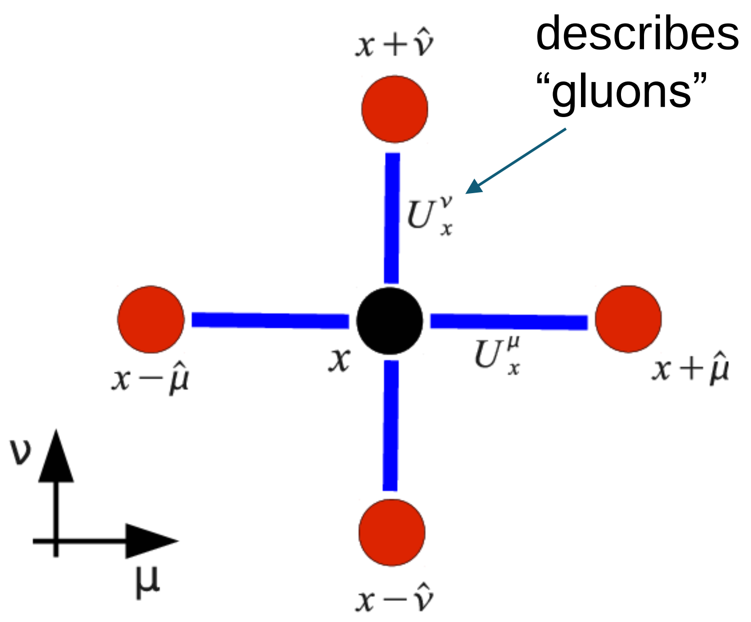
$$\hat{G}\psi\hat{G}^{-1} = \sigma_2\psi$$

$$\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$ Pauli “flavor” matrix

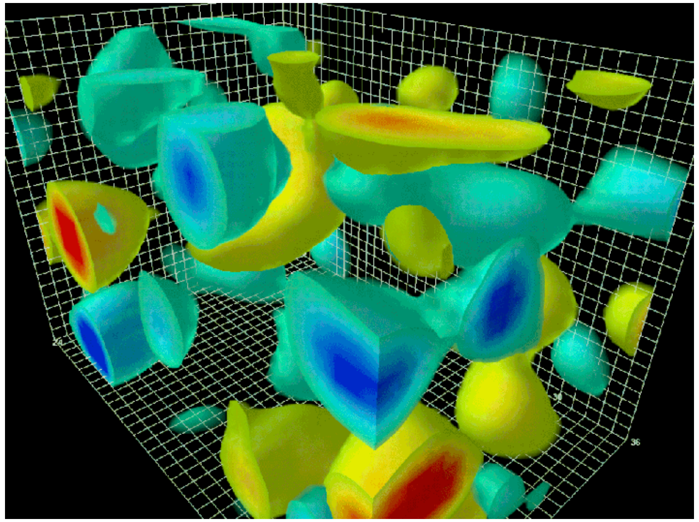
Brookhaven National Laboratory

Two for the price of two

- Quark propagation described by the inverse of the "**Dirac operator**", $\mathcal{M}$: a matrix in spin/color/space-time:
  - Dense in spin,color
  - Sparse (stencil operator) in space-time
- In a modern sim this matrix is ~10^9x 10^9 !
- Inversion of this matrix (via iterative methods) is the most computationally expensive aspect of a lattice calc.
- With G-parity BCs, we must now boost it to a dense matrix in "flavor", **doubling the cost!**



An ensemble of woes



- We use *Markov chain Monte Carlo* to **sample the "gauge fields" U_x**
- **G-parity BCs need custom ensembles**
- Probability weight for n "quarks" ➡ $\sim \exp\left(-\phi^\dagger (\mathcal{M}\mathcal{M}^\dagger)^{-n/2}\phi\right)$
- For 1 quark, computing matrix square-root is expensive, especially for light quarks (condition #).
- *Normally, however we use n=2 for the light quarks.*

Dirac op, contains U_x

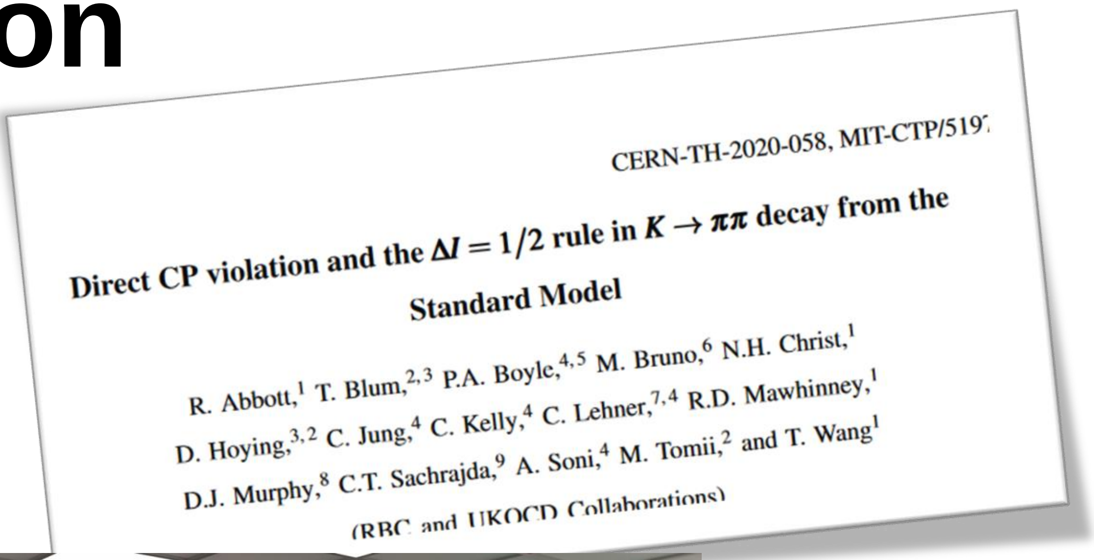
complex vector "pseudofermion" (also randomly sampled in MCMC)

- With G-parity BCs, $\mathcal{M}$ is already two-flavor, must use n=1 for a physical calc.
- **Makes sim even more expensive!**

Brookhaven National Laboratory

Motivation for optimization

- RBC & UKQCD are working to <u>repeat our 2020 result on a larger, finer lattice</u>.
- Aim to address the **“discretization” sys. err.** from simulating on just 1 lattice spacing.
- Large resource grants on *Polaris* (~135k node-hours ALCC) + comparable on *Perlmutter*.
- Aim for ~750 measurements, expect multiple **years** of running to accumulate.
- **Strong motivation for algorithmic development.**


CERN-TH-2020-058, MIT-CTP/519

Direct CP violation and the $\Delta I = 1/2$ rule in $K \to \pi\pi$ decay from the Standard Model

R. Abbott,[1] T. Blum,[2,3] P.A. Boyle,[4,5] M. Bruno,[6] N.H. Christ,[1]
D. Hoying,[3,2] C. Jung,[4] C. Kelly,[4] C. Lehner,[7,4] R.D. Mawhinney,[1]
D.J. Murphy,[8] C.T. Sachrajda,[9] A. Soni,[4] M. Tomii,[2] and T. Wang[1]
(RBC and UKQCD Collaborations)

A curious complexity

- Inspired by work on DDHMC, we looked for a way to *flavor-decompose the Dirac operator* $\mathcal{M}_{GP}$
- $\mathcal{M}_{GP}$ has a curious property that *complex conjugation is equivalent to a unitary transformation*
- We discovered that a related unitary transformation makes $\mathcal{M}_{GP}$ **pure-real**

4x4 complex spin matrix

$$\mathcal{M}_{GP}^{*} = -X\sigma_2\mathcal{M}_{GP}\sigma_2 X$$

$$\mathcal{M}_{\mathrm{re}} \equiv R^{\dagger}\mathcal{M}_{GP}R$$

$$R = \frac{1}{\sqrt{2}}\begin{pmatrix} -X & i \\ -1 & iX \end{pmatrix}$$

Two for the price of one!

- Can then separate contributions to probability weight from real and imaginary parts of the pseudofermion.
- 4-quark weight decomposes into 2x 2-quark weights with real pseudofermions.
- <u>If we just evaluate with one set of real pseudofermions we get the 2-quark weight without needing a square root!</u>
- *Furthermore, the cost of real-matrix times real-vector is <2x complex equivalent!*

$$\exp\left(-\phi^\dagger(\mathcal{M}_{GP}\mathcal{M}_{GP}^\dagger)^{-1}\phi\right)$$

$R^\dagger\phi$

$$= \exp\left(-\phi'^\dagger(\mathcal{M}_{\rm re}\mathcal{M}_{\rm re}^\dagger)^{-1}\phi'\right)$$

$(\phi' = \phi'_r + i\phi'_i)$

$$= \exp\left(-\phi_r'^\dagger(\mathcal{M}_{\rm re}\mathcal{M}_{\rm re}^\dagger)^{-1}\phi'_r\right)$$

$$\times \exp\left(-\phi_i'^\dagger(\mathcal{M}_{\rm re}\mathcal{M}_{\rm re}^\dagger)^{-1}\phi'_i\right)$$

Brookhaven National Laboratory

The X-factor

- Implementing the real Dirac matrix is non-trivial in a library (Grid).
- Fortunately we can avoid this with another trick:

$$\mathcal{M}_{\mathrm{re}}\phi_r = R^\dagger \mathcal{M}_{GP}\phi_X$$

$-X\phi_{r,1} + i\phi_{r,2}$

$$\phi_X = R\begin{pmatrix}\phi_{r,1}\\ \phi_{r,2}\end{pmatrix} = \begin{pmatrix}\chi\\ -X\chi^*\end{pmatrix}$$

An “X-conjugate” vector

$$\mathcal{M}_{GP}\begin{pmatrix}\chi\\ -X\chi^*\end{pmatrix} = \begin{pmatrix}\rho\\ -X\rho^*\end{pmatrix}$$

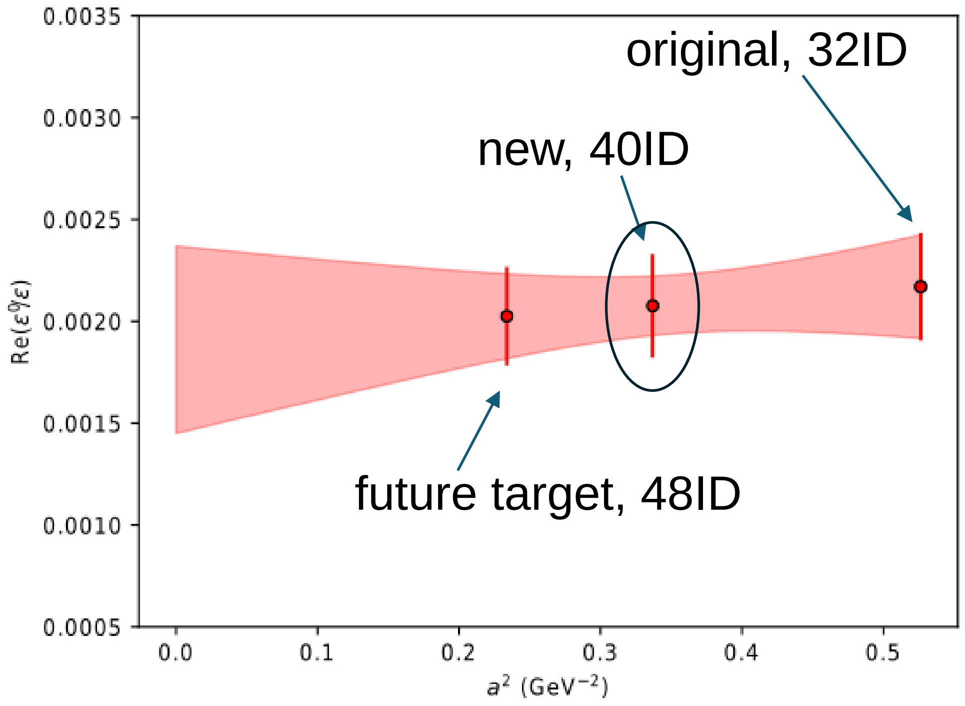
$$\rho = \mathcal{M}_{GP,11}\chi - \mathcal{M}_{GP,12}X\chi^* \equiv \mathcal{M}_X\chi$$

Can be evaluated using a standard (non-G-parity) Dirac op but with “X-conjugate” boundary conditions

$$\chi(x+L) = -X\chi^*(x)$$

40ID ensemble

- $40^3 \times 64$ DWF+Iwasaki-DSDR ensemble
  - $a^{-1} = 1.73$ GeV vs 1.38 GeV previous
  - Same physical volume, physical masses
- Evolving on Perlmutter GPU
- Switched to X-conjugate action and retuned evolution:

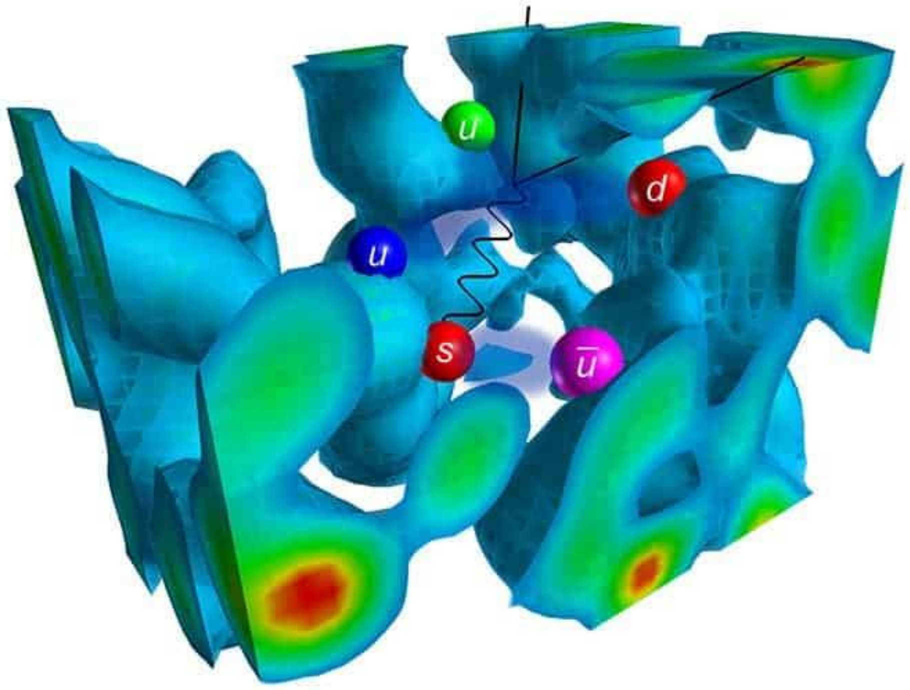


  - Original: 4.36hrs (32 nodes) – 139.5 node-hrs
  - New : 1.12hrs (32 nodes) – 35.8 node-hrs
    : 1.61hrs (16 nodes) – 25.76 node-hrs

5.4x (or 3.9x) reduction in cost, 2.7x (or 3.9x) speedup

Measurements

- $K \to \pi\pi$ measurements much more expensive (per config) than ensemble gen:
  ~7.2 hrs on 80 nodes Polaris
  (576 node-hours)
- ~60% cost on inversions:
  - 2000 light eigenvectors, single-prec block Lanczos: 1.1 hrs
  - Deflated mixed-precision (split-Grid) G-parity light quark inversions: 2.25 hrs
  - Mixed-precision (split-Grid) G-parity strange quark inversions: 0.8 hrs



Eigenvectors of Hermitian Dirac Op

$$\mathcal{M}^\dagger\mathcal{M}\psi = \lambda\psi \quad \Rightarrow \quad R^\dagger\mathcal{M}^\dagger\mathcal{M}R(R^\dagger\psi) = \lambda(R^\dagger\psi)$$

$$\Rightarrow \quad \mathcal{M}_{\rm re}^T\mathcal{M}_{\rm re}(R^\dagger\psi) = \lambda\underbrace{(R^\dagger\psi)}_{v_r}$$

- $\mathcal{M}_{\rm re}^T\mathcal{M}_{\rm re}$ is real, symmetric:
  - evecs v_r can be chosen to be real vectors

Evecs of $\mathcal{M}_{GP}$ $\Rightarrow$ $\psi \equiv Rv_r = \frac{1}{\sqrt{2}}\begin{pmatrix} -Xv_{r,1} + iv_{r,2} \\ -v_{r,1} + iv_{r,2} \end{pmatrix} = \begin{pmatrix} \psi_1 \\ -X\psi_1^* \end{pmatrix}$

- Evecs ψ can be expressed as X-conjugate vectors!
- <u>Compute using Lanczos method with X-conjugate Dirac op</u>

2x cost reduction in generating evecs!
2x reduction in memory and disk footprint for storing!

Quark propagators

- We use "all-to-all" propagators

A2A:

$$\psi^i = \mathcal{M}_{GP}^{-1}\eta^i$$

$$\eta: \quad \frac{1}{N}\sum_{i=1}^{N}\eta^i(x)\eta^{i\,\dagger}(y) \approx \delta_{x,y}$$

$$\mathcal{M}^{-1}(x,y) \approx \frac{1}{N}\sum_{i=1}^{N}\psi^i(x)\eta^{i\,\dagger}(y)$$

- We identified a variety of sources satisfying this with X-conjugate columns.
- Inverse can be performed with X-conjugate Dirac op: 2x cost reduction
- Identified optimal candidate via numerical expt. on small lattice.
- *Evidence of improvement in statistical error on measurement result!?*

"X-conjugate A2A source"

$$\eta^i(x) = \begin{pmatrix} a(x) & b(x) \\ -Xa^*(x) & -Xb^*(x) \end{pmatrix}$$

a, b random, $\eta^i(x)$ unitary

Brookhaven National Laboratory

Conclusions

- Under SciDAC-5 we have upended the status quo regarding calculations with G-parity BCs.
  - 4x speed-up in ensemble generation
  - 2x speed-up / 2x memory usage reduction in eigenvector generation
  - 2x speed-up in inversions (for suitable sources)
- G-parity simulations are now comparable in cost to conventional calculations!
- *This has dramatically accelerated RBC & UKQCD's $K \to \pi\pi$ program.*

Brookhaven National Laboratory