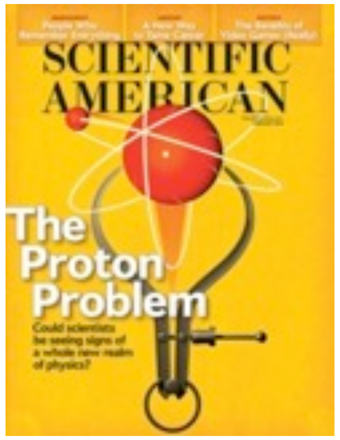
Nucleon Structure on a Lattice :
Present and Future Computing Requirements

Sergey Syritsyn
RIKEN / BNL Research Center
Brookhaven National Lab, Upton, NY

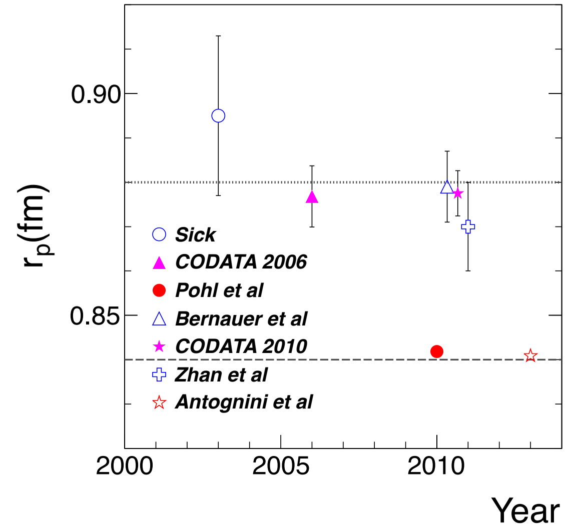
Large Scale Computing and Storage Requirements
for Nuclear Physics : Target 2017
Bethesda, Maryland, April 29, 2014

Scientific Goals (1)

- Proton Charge Radius
  current lattice QCD uncertainty is comparable to the experimental discrepancy



proton radius as seen by muon and electron differs by 7σ

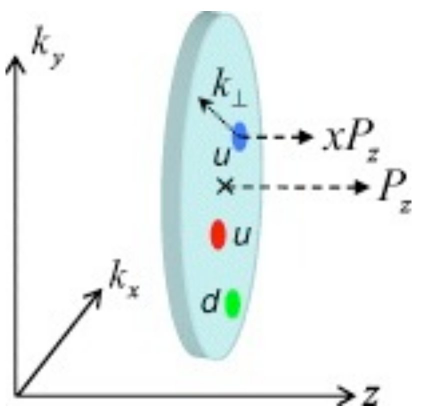


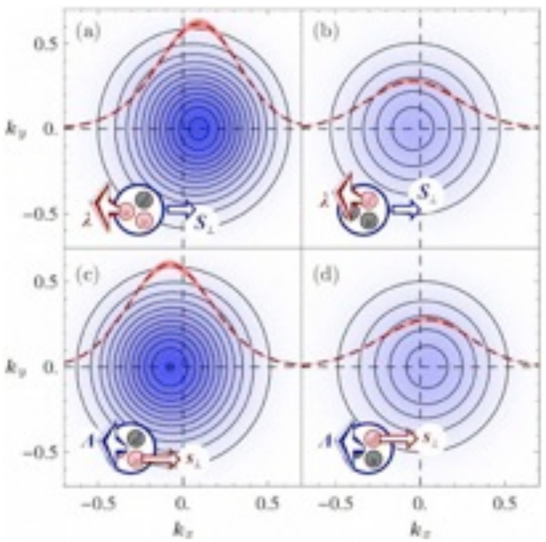
[arXiv:1303.2160]

- Proton and Neutron Charge and Magnetization Distributions (Form Factors)
  fundamental theory counterpart to extensive programs at JLab, Mainz; OLYMPUS, MUSE (planned)

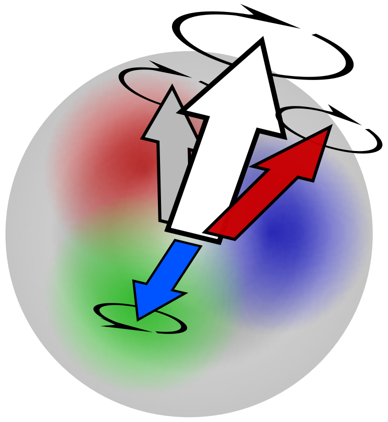
- Quark Density Distributions in the Proton
  current experiments at JLab, planned Electron-Ion Collider

- Proton Spin Puzzle

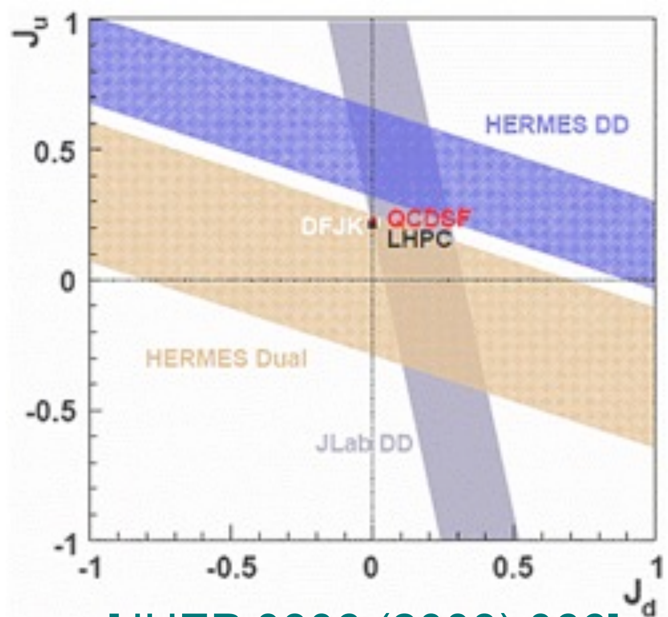




[arXiv:0908.1283]



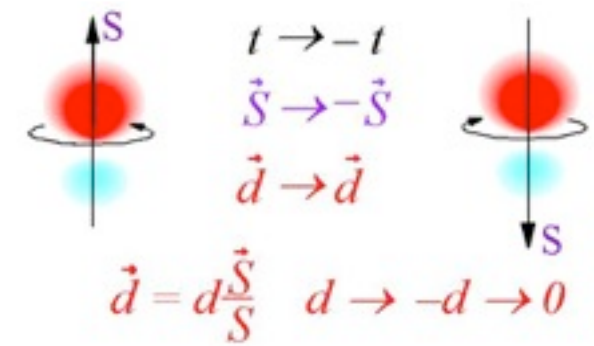
only 30% of the proton spin can be explained by quark spins



[JHEP 0806 (2008) 066]

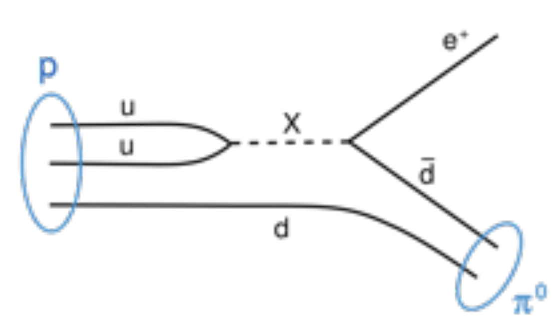
Scientific Goals (2)

Searches for New Physics and explanations of Matter-Antimatter imbalance in the Universe



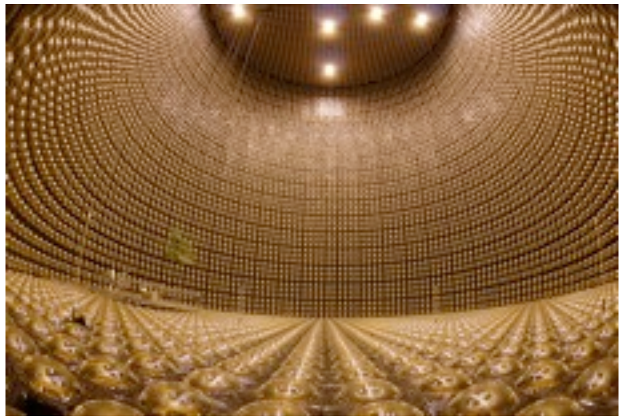
- Sensitivity to Dark Matter

- Neutron Electric Dipole Moment
  experiments to improve bounds on nEDM proposed at Spallation Neutron Source (ORNL), TRIUMF, PSI

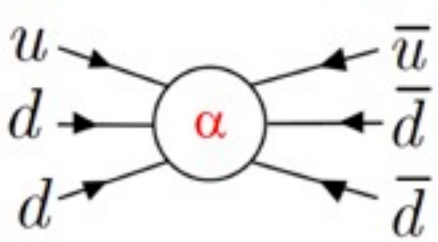


- CP violation
  sensitivity of ordinary matter to CPV effects from Beyond the Standard Model (BSM) physics

- Proton decay
  on-going and planned experiments (e.g. Hyper-Kamiokande)



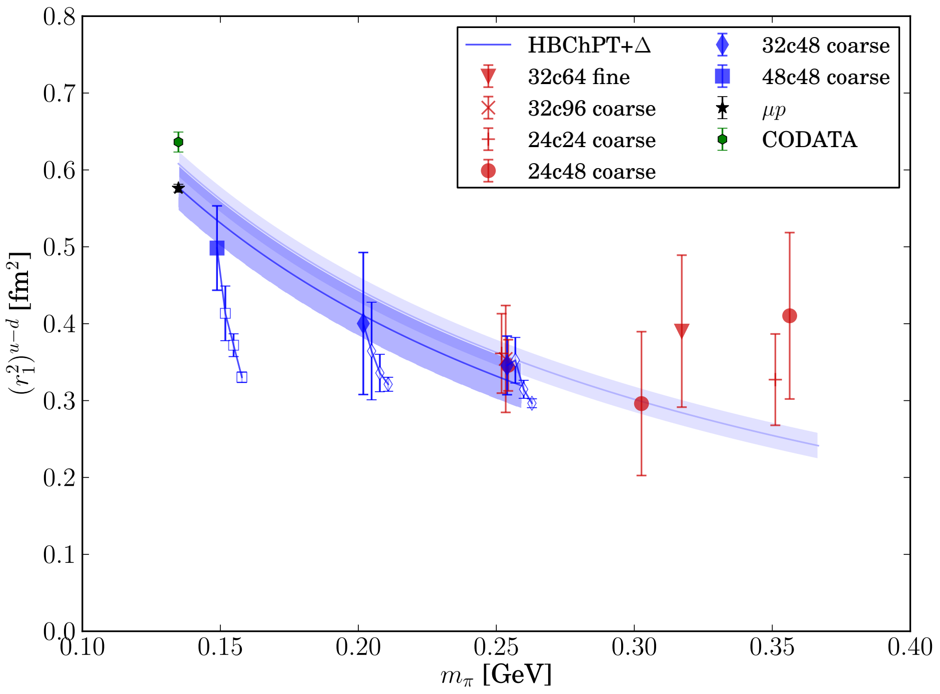
- Neutron-antineutron oscillation
  Proposed Project X at Fermilab

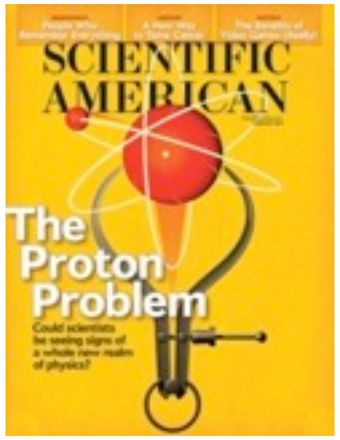


Lattice QCD calculations are essential for gauging sensitivity of current and next generation BSM searches

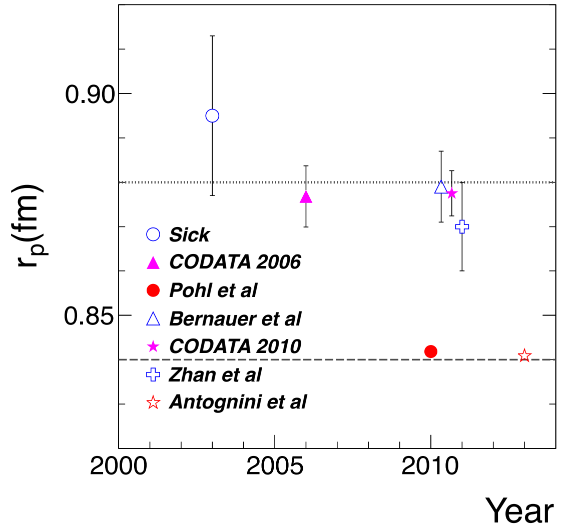
Nucleon Structure Highlights

Dirac Radius





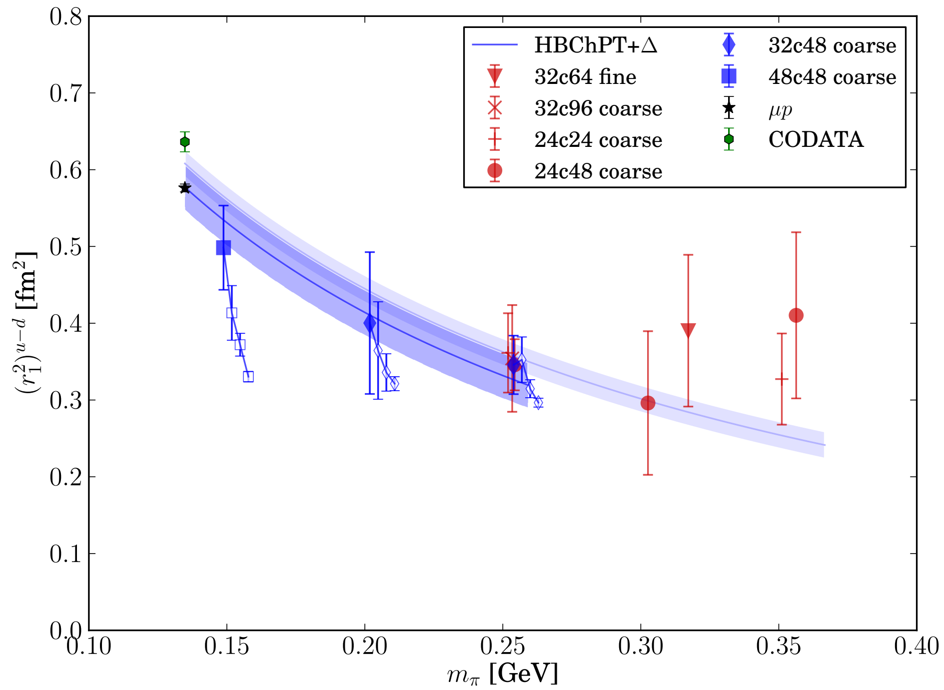
proton radius as seen by muon and electron differs by 7σ

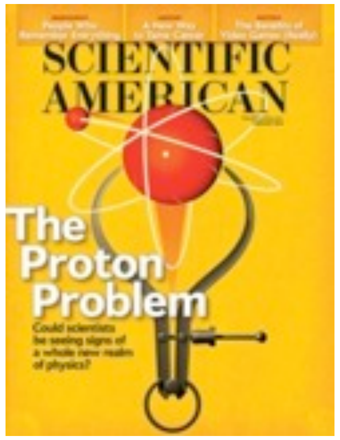


[arXiv:1303.2160]

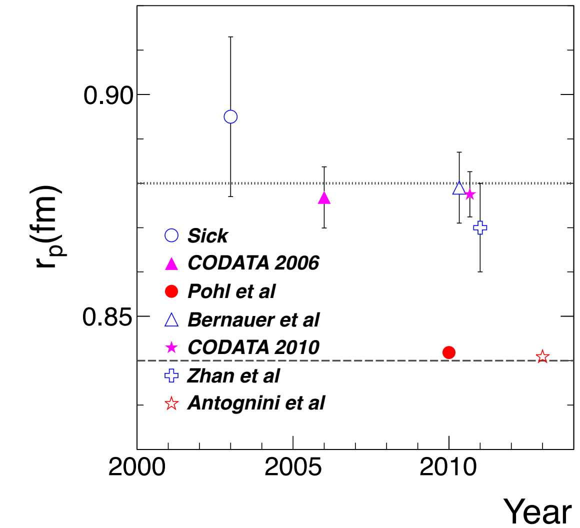
Nucleon Structure Highlights

Dirac Radius



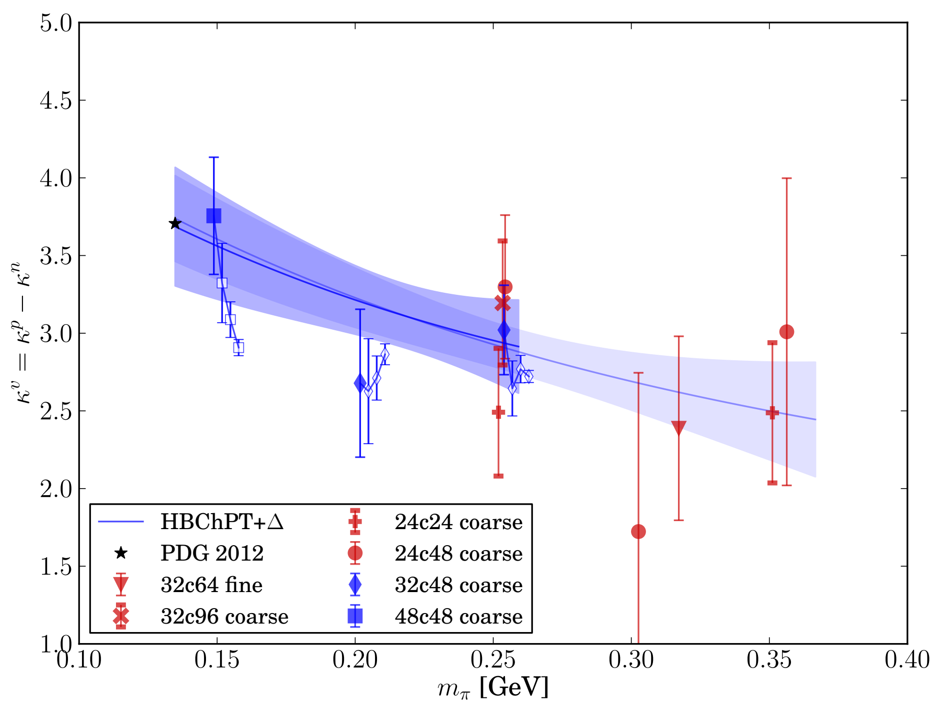


proton radius as seen by muon and electron differs by 7σ

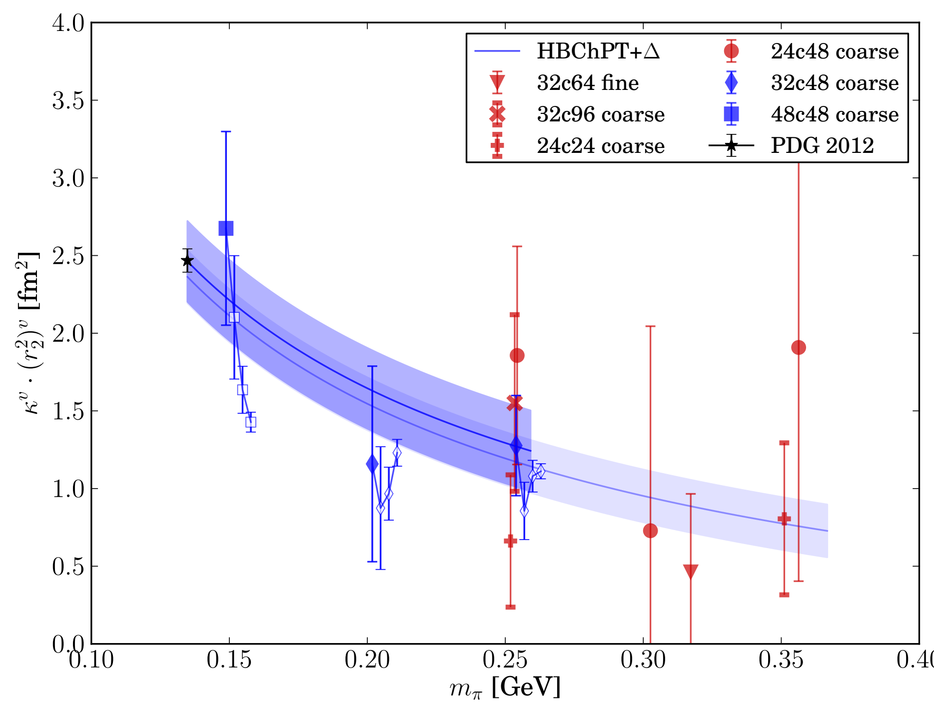


[arXiv:1303.2160]

Magnetic Moment



Pauli Radius



Lattice Objectives 2014-2017 : Hadron Structure

Perform lattice QCD calculations of proton and neutron structure with

- *physical chirally-symmetric quarks,*
- *at least two different lattice discretization spacings, and*
- *sufficiently large lattice volume for F.V. systematic << stochastic errors*

Examples of Specific Objectives for 2017

- Compute Isovector Nucleon Charge Radius to 3% (exp.discrepancy ~10%)
- Compute Nucleon Axial Charge to 2%; (experimental uncertainty ~0.2%)
- Compute separate contributions to the proton spin and energy-momentum from u and d quark momenta, spins and orbital angular momenta
- Compute moments of nucleon generalized parton distributions (GPDs) to complement experimental programs at the future EIC

[NSAC Long Range Plan] HP9 milestone: “Perform lattice calculations in full QCD of nucleon form factors, low moments of nucleon structure functions and low moments of generalized parton distributions, including flavor and spin dependence”

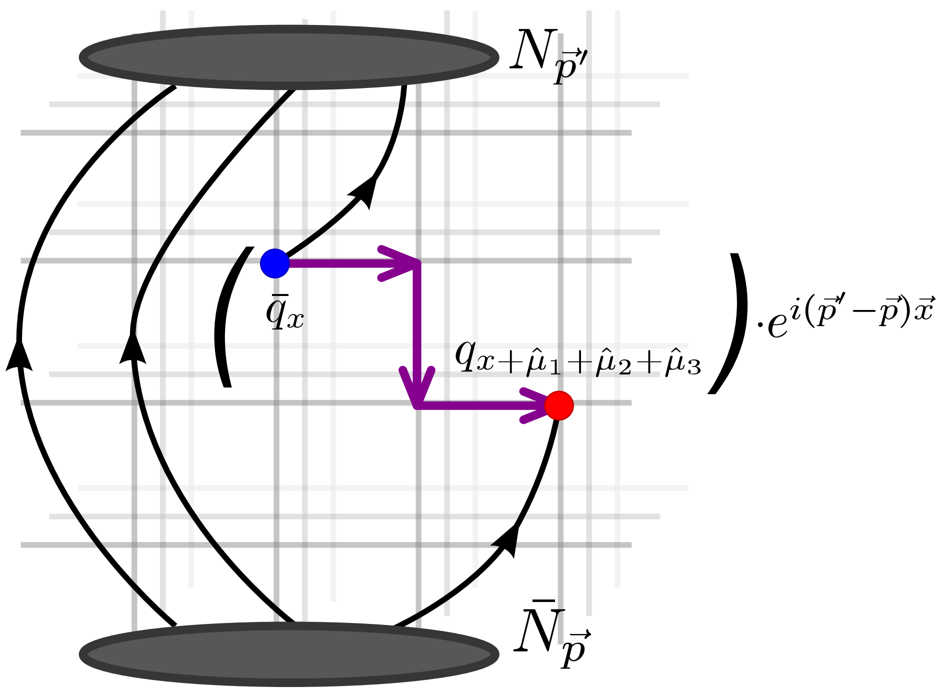
Nucleon Structure on a Lattice Project

- Gauge field configurations are generated elsewhere (normally BG/Qs), gauge generation cost << nucleon structure cost
- Capacity computing requirement: collect independent Monte Carlo samples, trivially parallel
- Variance reduction techniques are essential and put strong requirements on total memory per job

Differences from Hadron Spectrum

- Specific requirements to lattice discretization of quarks, more expensive than discretization than for Hadron Spectrum
- Already at the physical mass
- Fewer MC samples, but each sample is more expensive

Computing Strategy



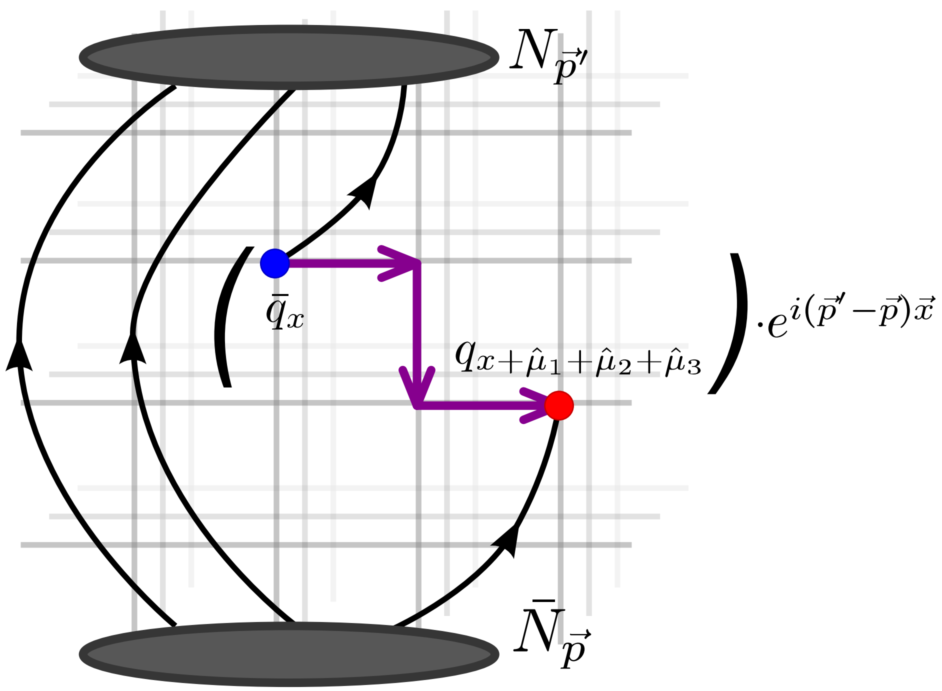
- 1 MC sample = 15-20 quark propagators
- sample over positions on the same gauge cfg
- 1 propagator = 12 Dslash inversions
- CG with even-odd preconditioning works best on CPU clusters and BG/Q
- Deflation speedup x10 with 500 e.vectors on a 48c96x24 lattice
- Cost of deflating eigenvectors is amortized by many samples per gauge cfg

$$\langle N(z)\mathcal{O}(x)\bar{N}(y)\rangle = \frac{1}{N_{\text{samp}}}\sum_{N_{\text{samp}}}[\not{D}^{-1}_{z,y}\otimes\not{D}^{-1}_{z,y}\otimes\not{D}^{-1}_{z,x}\Gamma_{\mathcal{O}}\not{D}^{-1}_{x,y}]$$

Quark propagators

$$\not{D}_{x,y} = (\text{sparse in coordinates } x,y)\otimes(4\times 4\text{ spin}\otimes 3\times 3\text{ color local matrices})$$

Computing Strategy : Variance Reduction



"All-mode averaging":

$$\not{D}^{-1} \cdot u \approx V_k \frac{1}{\Lambda} V_k^\dagger \cdot u + \not{D}^{-1}_{\text{truncated CG}} \cdot \left(1 - V_k V_k^\dagger\right) \cdot u$$

$$V_k = \text{low-mode eigenvectors}$$

$$y_{\text{corrected}} = \langle y_{\text{approx}} \rangle + \langle \Delta y \rangle \,, \quad \Delta y = y_{\text{exact}} - y_{\text{approx}}$$

$$\text{Var}\{y\} = \text{Var}\{\Delta y\} + \frac{1}{N_{\text{approx}}} \text{Var}\{y_{\text{approx}}\}$$

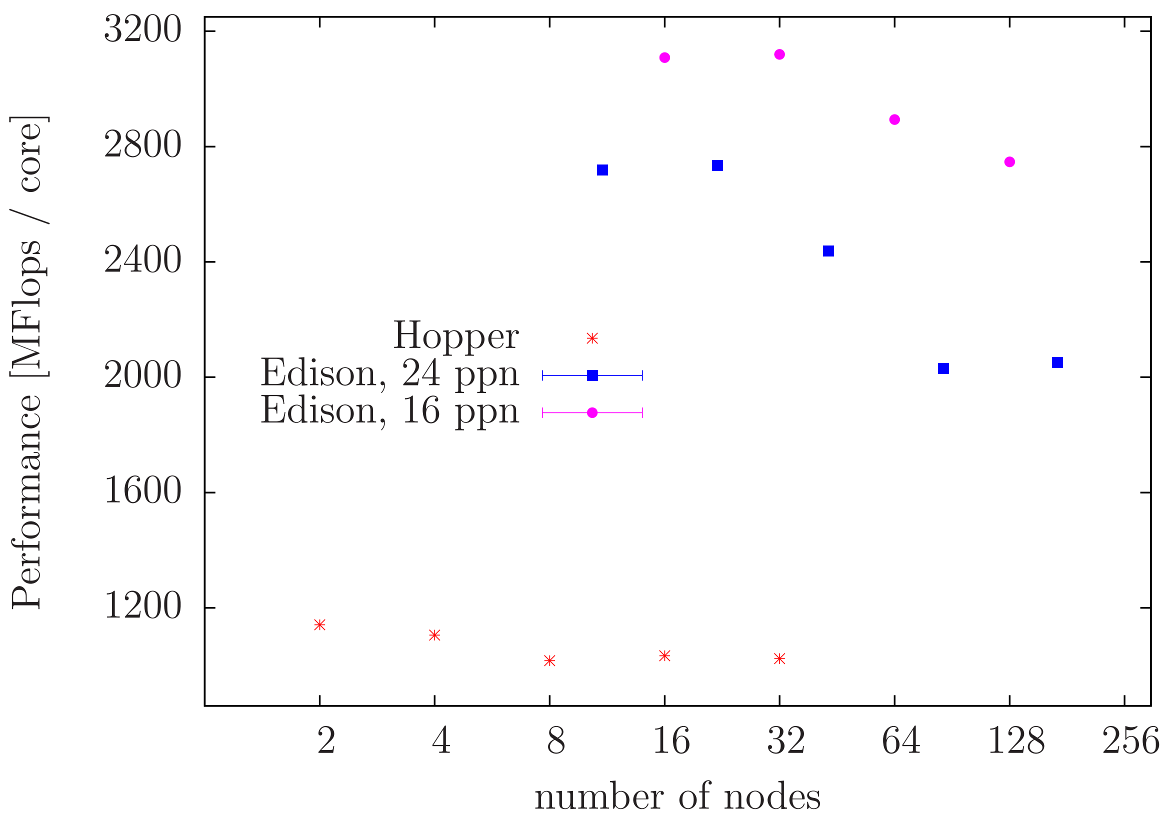
$$\text{Cost}\{y_{\text{approx}}\} \ll \text{Cost}\{y_{\text{exact}}\}$$

- Deflation of eigenmodes between light and strange quark masses
- Truncated CG solver : 1/4 ... 1/16 cost
- Bias is computed on a much smaller statistics (1/32 .. 1/64)
- Deflation is **essential** for small bias correction
- Eigenvectors are computed with Impl.Restarted Arnoldi (ARPACK), with *n=200...400* Chebyshev polynomial acceleration

Software Summary

Complete production cycle is implemented in Qlua
[https://usqcd.lns.mit.edu/redmine/projects/qlua_code]

- Interface to MDWF (A.Pochinsky) [https://usqcd.lns.mit.edu/redmine/projects/mdwf]
  optimized for BG/Q and CPU clusters
  efficient kernel for CG and Arnoldi algorithms
  mixed-precision CG with EV deflation
- PARPACK for computing eigenpairs
- USQCD QIO library:
  disk read & write rate ~75 GB/sec (Edison)
- Support HDF5 for future intensive I/O



performance on 32^3x64
[2014 request]

Current HPC Usage

- Fermilab USQCD cluster ("Ds"): 20M "JPsi" core*hours
  2GHz AMD 6128 Opteron 64GB, 32 cores and 64 GB mem per node, QDR (40Gbit/sec) Infiniband

  48^3x96x24
  use only 500+100 (~1/4) of required e.vec's due to lack of memory & I/O
  run on 128nodes * 32cores, 64GB
  job duration ~35 hours
  I/O (reading e.vec's, Lustre) ~1.5 hour

- 2014 NERSC Award (mp133): 55M MPP

  48^3x96x24 with 2000 e.vec's will require 512 nodes of Edison
  with peak SCRATCH3 I/O throughput, ~2.5min
  with sustained in a real test ~5min

	2014
Lattice Size	48^3x96x24
# Eigenvectors	2000+400
E.vector size	11.4 GB
Memory	27.3 TB
Startup I/O	22.8 TB
Total storage	1.14 PB
Min # nodes (Edison)	512
Min # cores (Edison)	12k
Max # cores (local vol 4x4x4x4)	40.5k
Computing Requirements	200M MPP

Memory & I/O Requirements

Using deflation and variance reduction effectively requires a lot of memory and storage I/O throughput

	2014-2015(*)	2015-2017	2017-??
Lattice Size	48^3x96x24	64^3x128x12	80^2x96x192x16
# Eigenvectors	2000+400	2000+400	2500+500
E.vector size	11.4 GB	18 GB	84.4 GB
Memory	27.3 TB	43.2 TB	253 TB
Startup I/O	22.8 TB	36 TB	211 TB
Total storage	1.14 PB	1.80 PB	10.55 PB
Min # nodes (Edison)	512	810	4800
Min # cores (Edison)	12k	19k	112.5k
Max # cores (local vol 4x4x4x4)	40.5k	128k	450k
Computing Requirements	200M MPP	550M MPP (**)	3000M MPP (**)

(*) on-going project: 55M MPP at NERSC, 20M Jpsi core*hours at Fermilab USQCD clusters

(**) ballpark estimates; exploratory studies are required

Alternatives to Exact Deflation

- HDCG [arXiv:1402.2585]: deflation on a coarsened grid,
  turn a sparse problem into dense, much smaller dimension;
  reduce memory and I/O requirements; a much better fit to GPUs and MICs
- Inexact deflation with approximate eigenvectors :
  recompute approximate eigenspace in each job ;
  reduce IO requirements
- Approximate operator with (1/2)...(1/3) smaller 5-th dimension
  (with proportional decrease to memory and I/O needs)
  reduce memory and IO requirements

All methods must be studied for efficiency at Variance Reduction

Emerging Architectures

- **Does your software have CUDA/OpenCL directives?**
- **Does your software run in production now on Titan using the GPUs?**

We currently do not use GPUs in our calculations. We plan to integrate GPU code into our production if multiple-GPU code (e.g. QUDA) is efficient enough, given our memory requirements (e.g. on Titan with 16 cores, 1 GPU, 32 GB per node, local volume 12x12x12x6 is too small).

- **Does your software have OpenMP directives now?**

We don't use OpenMP directives. Early tests of OpenMP have shown performance inferior to MPI. The approach has not been investigated further.

- **Does your software run in production now on Mira and Sequoia using threading?**

The current version runs on Mira without threading, but using 64 processes on a node instead, achieving 20% efficiency. The next generation software will use threads on nodes.

- **Is porting to, and optimizing for, the Intel MIC architecture underway or planned?**

Intel MIC support is in the plans, following software redesign to use threads.

Emerging Architectures

- **Have there been or are there now other funded groups or researchers engaged to help with these activities?**

QUDA group (M.Clark, R.Babich and others) are working on efficient GPU implementation of the MDWF operator.

- **Explain your strategy for transitioning your software to energy-efficient, manycore architectures**

Transition to GPUs will likely require exploration of alternative variance reduction techniques.

- **What role should NERSC play in the transition to these architectures?**

Providing early access to these architectures for code development is critical for a successful transition.

- **What role should DOE and ASCR play in the transition to these architectures?**

Emplacing small systems both at HPC centers and universities is crucial for early development and attracting young talent to the field.

Special I/O Needs

- **Does your code use checkpoint/restart capability now?**

Saving and loading eigenvectors effectively provides a checkpoint;
associated I/O by far exceeds any other I/O needs

- **Do you foresee that a burst buffer architecture would provide significant benefit to you or users of your code?**

Our code runs most efficiently with large local volumes;
deflation eigenvectors (currently ~35 TB total) occupy all memory on a node;
fast loading of eigenvectors (burst buffer?) will reduce CPU idle time

Summary

- Current project plan will deliver a broad spectrum of results on quantities associated with proton and neutron structure relevant for the current and future experiments in Nuclear and High-Energy Physics.
- **Recommendation on NERSC architecture, system configuration, etc**
  NERSC architecture fits our project very well, including the plan to increase the capacity x2 per year. The major bottlenecks may be memory per node (especially with GPUs) and I/O bandwidth
- **Expected significant scientific results?**
  After decades, the current state of lattice QCD is mature enough to provide precise answers about the structure of hadrons
  - Expected precision for the proton charge radius will exceed the experimental discrepancy.
  - Computing moments of GPDs is complementary to experimental program at the future EIC
- **Expanded HPC resources**
  Planned “Burst buffer” may be advantageous to avoid CPU idle on startup.