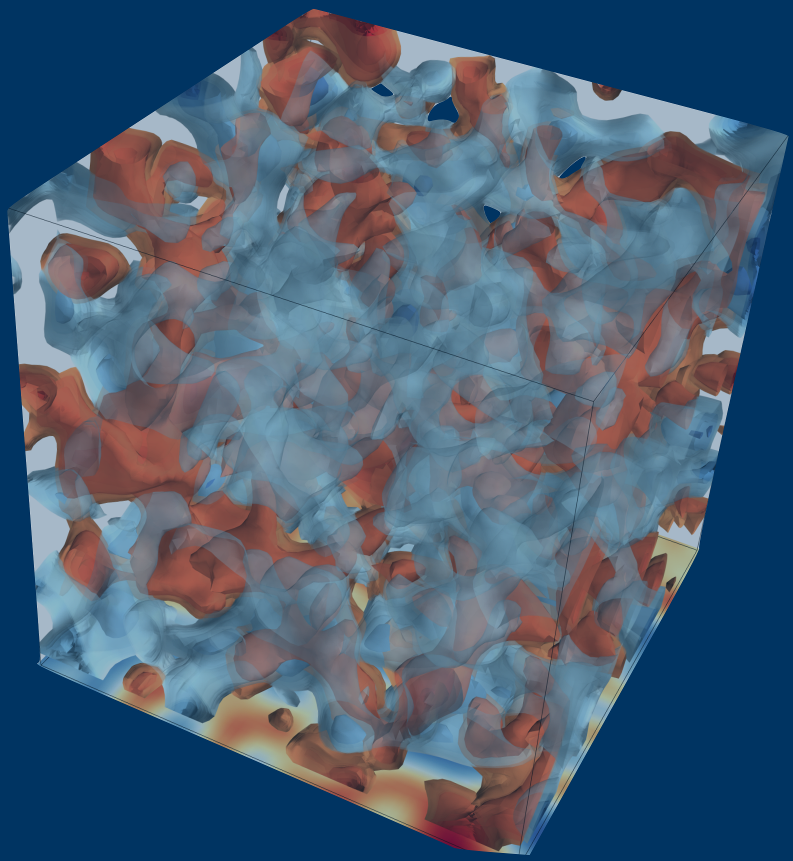


Lattice Quantum Chromodynamics

Software Infrastructure and Computing

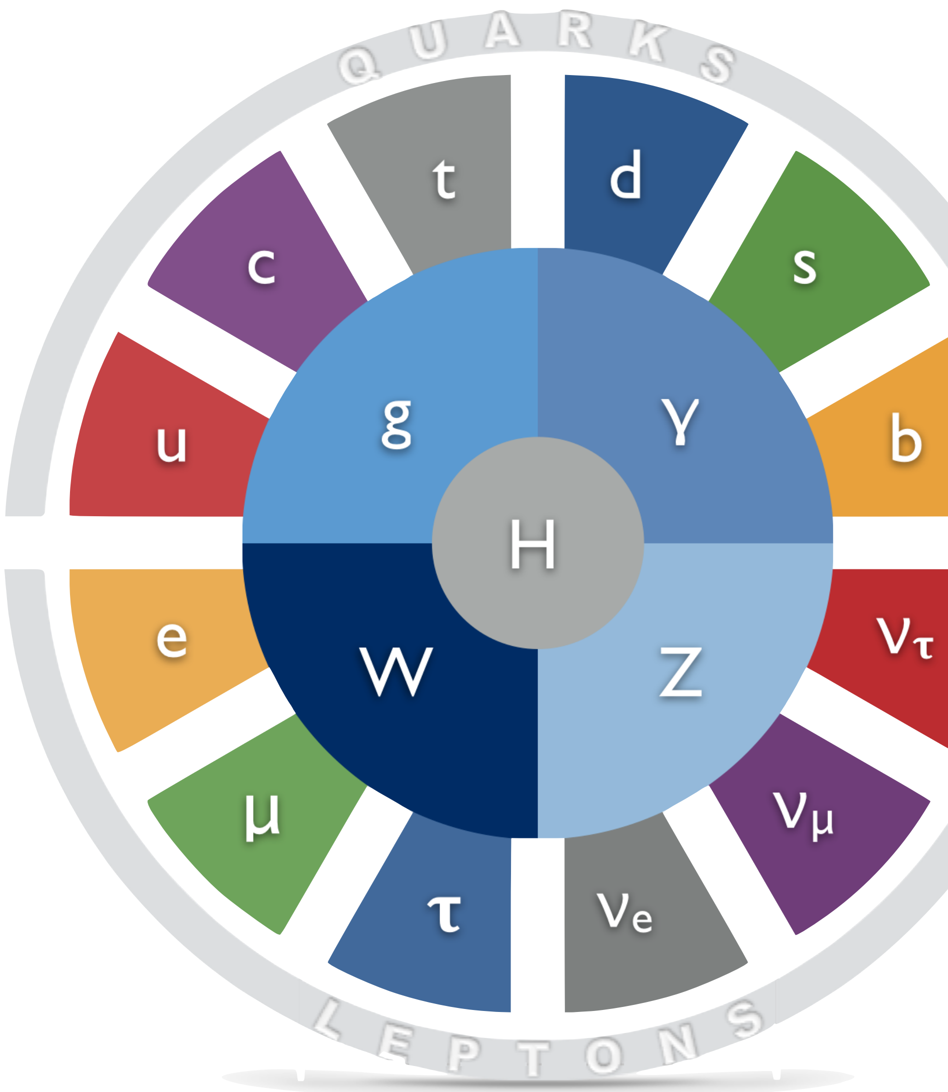
WORKSHOP ON
SOFTWARE INFRASTRUCTURE FOR ADVANCED NUCLEAR PHYSICS COMPUTING JUNE 20-22, 2024

William Detmold, MIT

The femtoscale universe

The Standard Model

- Understand the world around us from first principles in the Standard Model (SM)
    - Quarks, leptons & bosons (17 fundamental fields)
- What is the structure of matter?
    - All the way from the proton to neutron stars and supernovae
- How does it depend on SM parameters?
- How do we constrain the SM in experiments?
- How do we search for new physics beyond the SM?
    - Dark matter, dark energy, origin of neutrino masses,…



Quantum Chromodynamics (QCD)

The strong interaction

- Standard Model = electromagnetism + weak + strong nuclear forces
- QCD is the strong force that binds protons together and into nuclei
- Interaction strength depends on energy
- At high energy: analytic small-coupling expansion
- At low energies/long distances: no analytic control
    - The regime of protons and nuclei
    - Numerical calculations through Lattice QCD

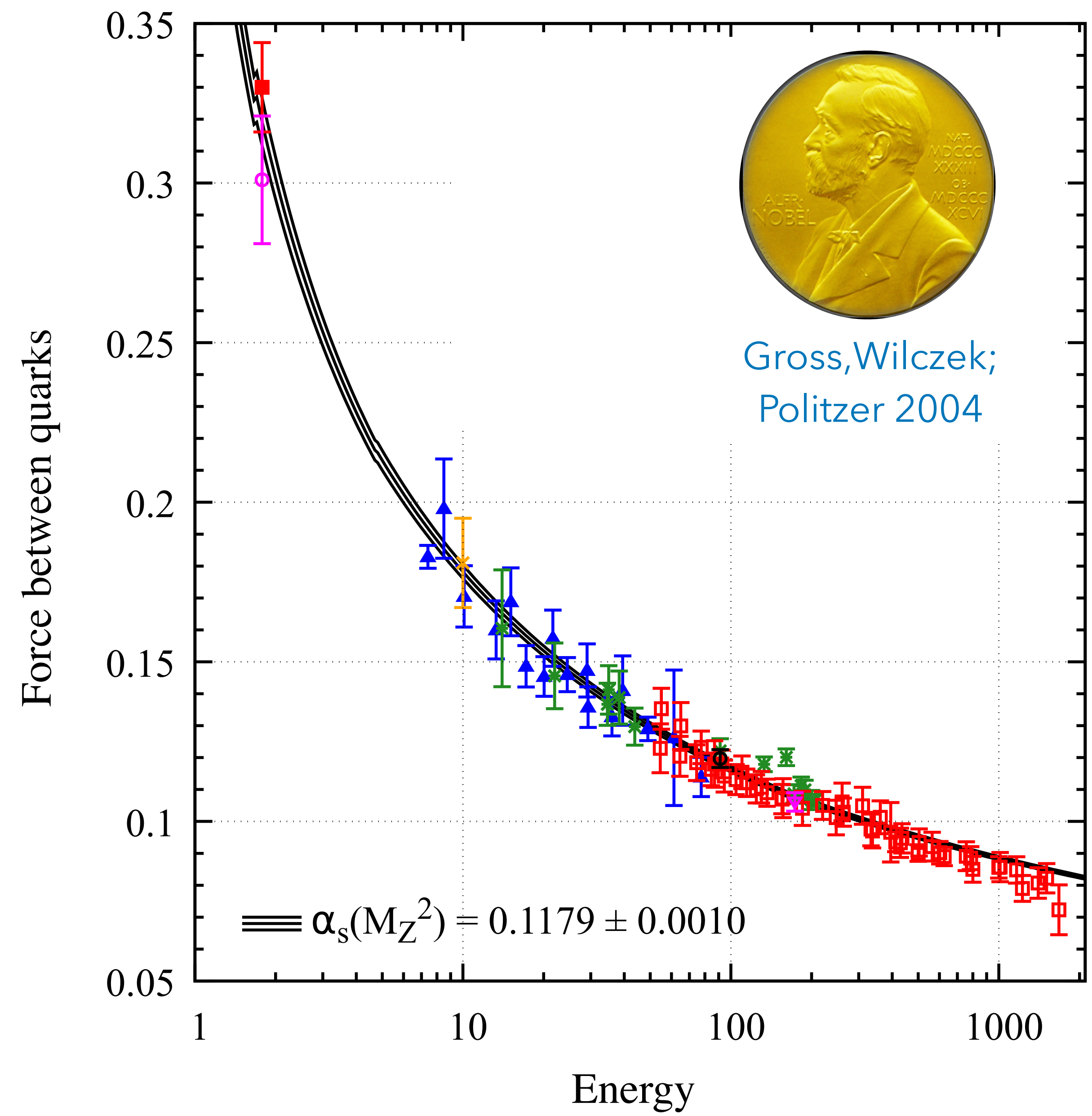


Figure credit: Particle Data Group, 2019

Anatomy of lattice QCD

In theory

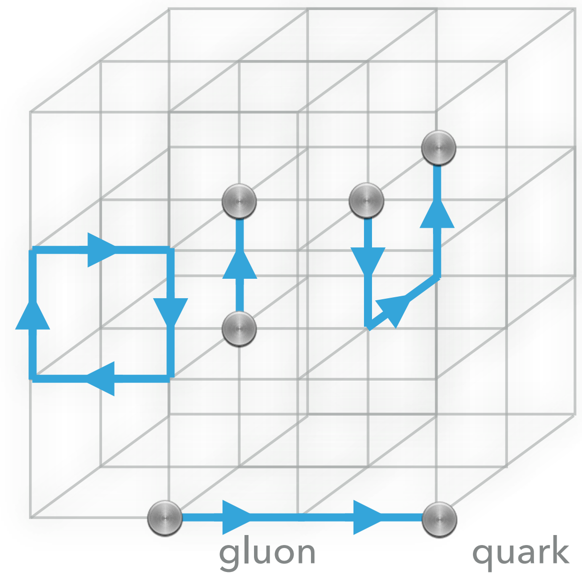
- Lattice QCD: tool to deal with quarks and gluons (also a rigorous definition of QCD)
- Formulate problem as functional integral over quark $(q, \overline{q})$ and gluon (A) d.o.f. in 3+1D spacetime

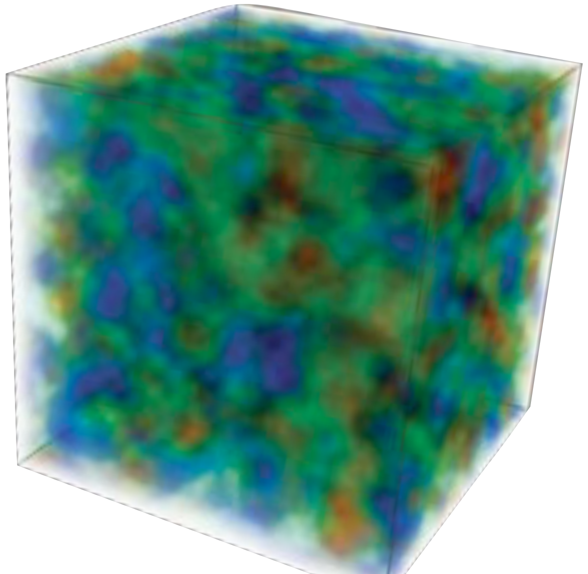
$$\langle \mathcal{O} \rangle = \int \mathcal{D}A \mathcal{D}q \mathcal{D}\overline{q} e^{-S_{\mathrm{QCD}}[A,q,\overline{q}]} \mathcal{O}[A, q, \overline{q}]$$

- Discretise and compactify system
    - Finite but large number of d.o.f (~10^{8}–10^{12})
- Integrate via importance sampling according to probability measure

$$\langle \mathcal{O} \rangle = \int \mathcal{D}P(A) \mathcal{O}[A] \simeq \frac{1}{N_{\mathrm{samp}}} \sum_{i}^{N_{\mathrm{samp}}} \mathcal{O}[A_{(i)}]$$

- Undo the harm done in previous steps





Anatomy of lattice QCD

In practice

1. Generate representative samples of gluon fields in a MC chain ($N_{\text{samp}} \sim 1000$)

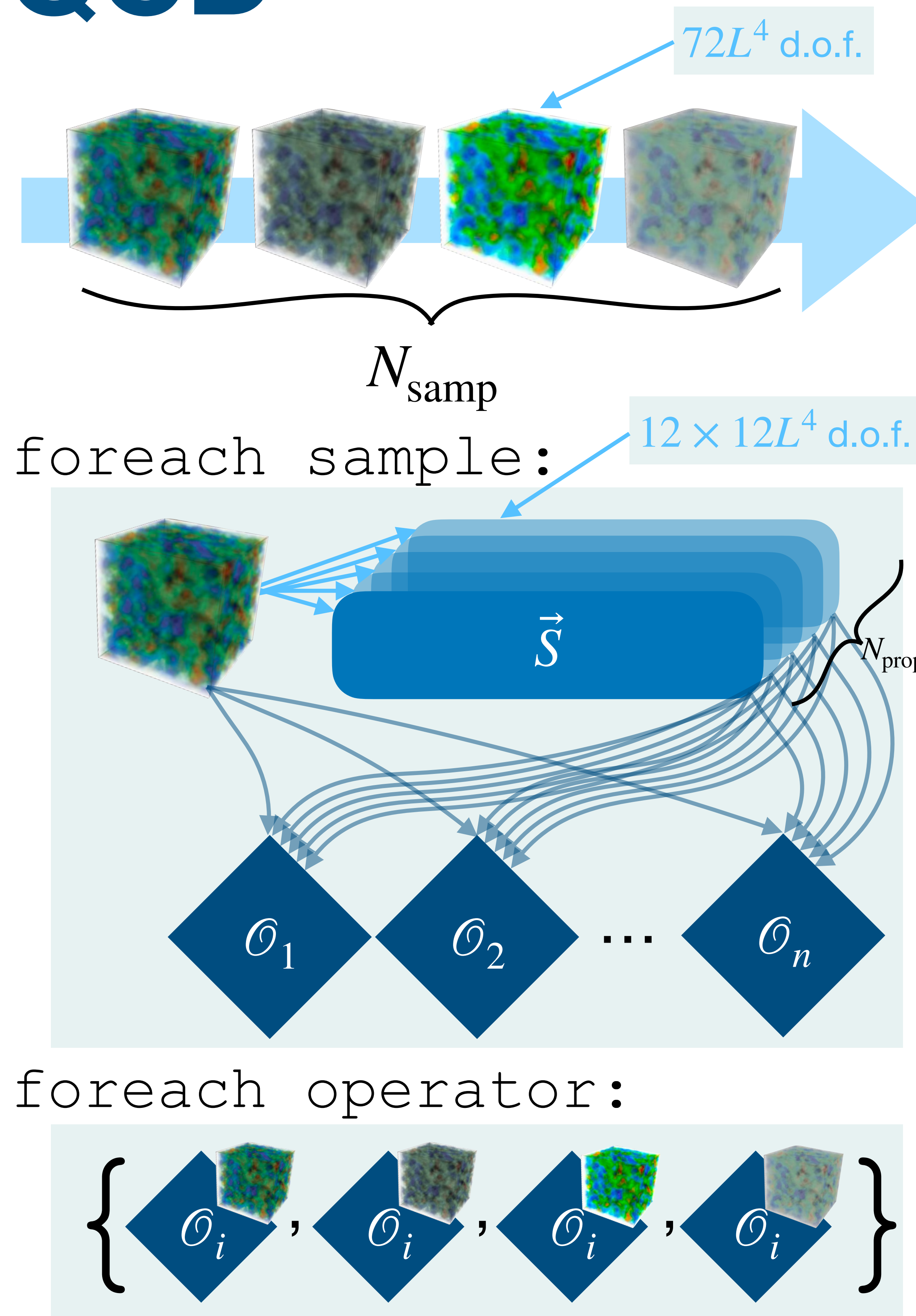
2. On each gluon field calculate "quark propagators" $\vec{S}$ by solving linear systems

$$\not{D}_{\text{Dirac}}[A_{(i)}]\vec{S} = \vec{\phi}$$

   Dirac operator depends on the gluon field; typically $N_{\text{prop}} \sim O(100s)$ RHSs used

3. Construct operators $\mathcal{O}_i$ for a particular physics goal from the gluon field and associated quark propagators
   - Tensor contractions

4. Analyse statistics of the resulting quantities to extract physics



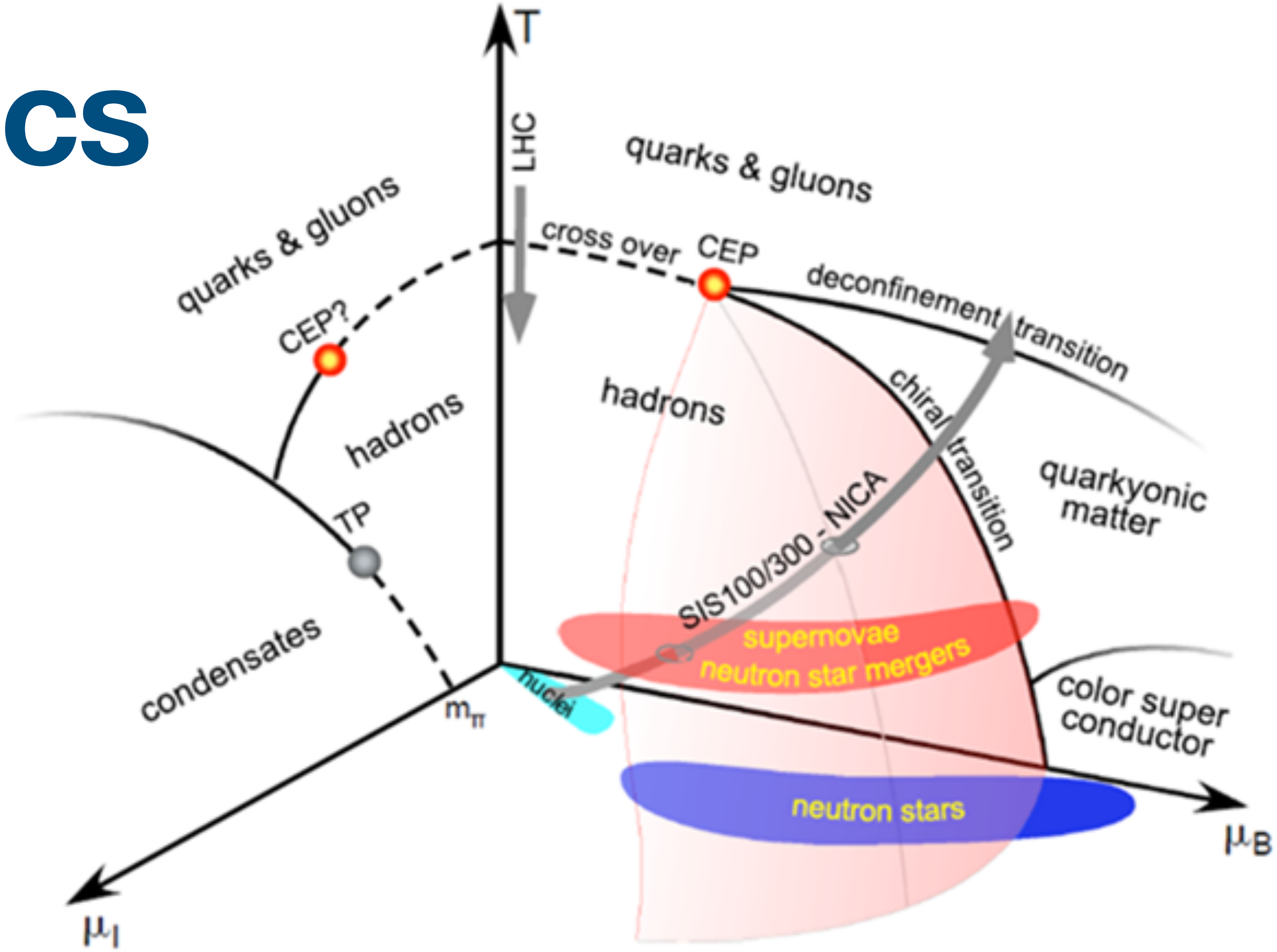
LQCD Science

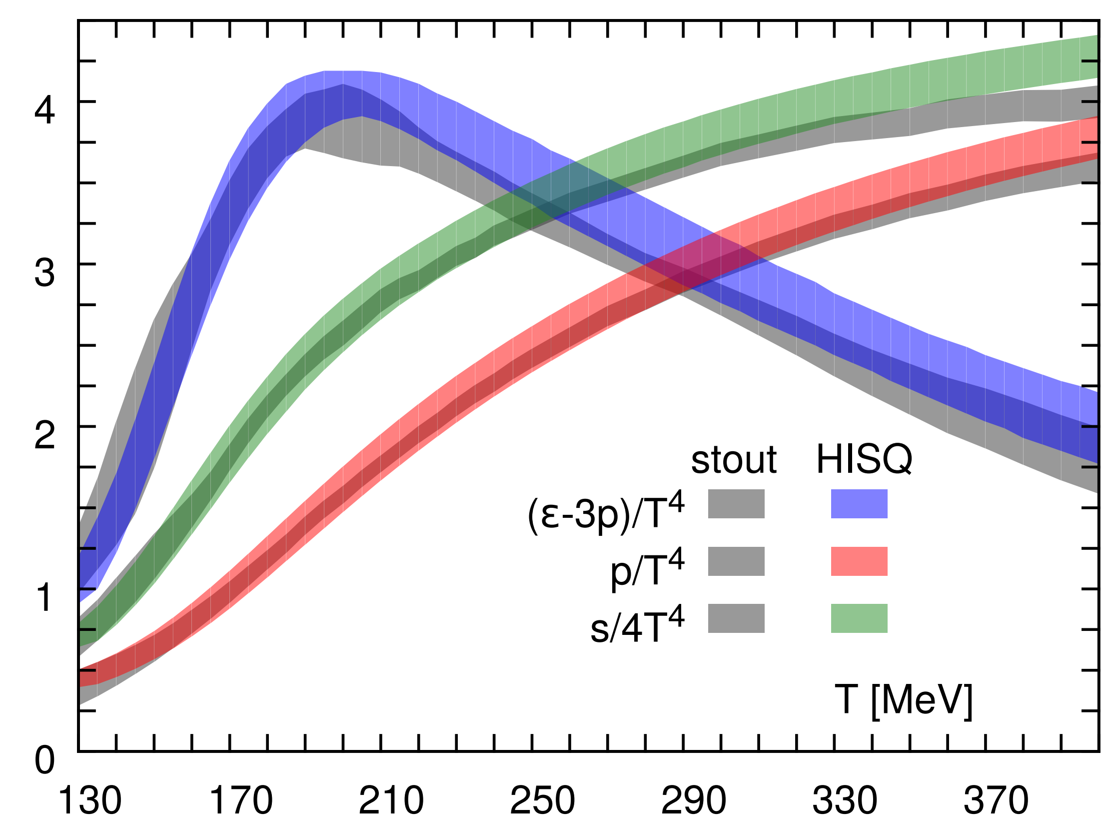
Achievements and Goals in Nuclear Physics

Thermodynamics

QCD Phase structure

- Understand the phases of strongly interacting matter
- Deconfinement transition at $\mu_B = \mu_I = 0,\ T \neq 0$
  - Crossover transition at $T_c = 155$ MeV
- Nonzero isospin density
  - Superconductivity
- Baryon density
  - Hints for critical end point
  - “Sign problem” prevents direct Monte-Carlo calculations
- Many open problems: neutron star equation of state, transport in heavy-ion collisions,…



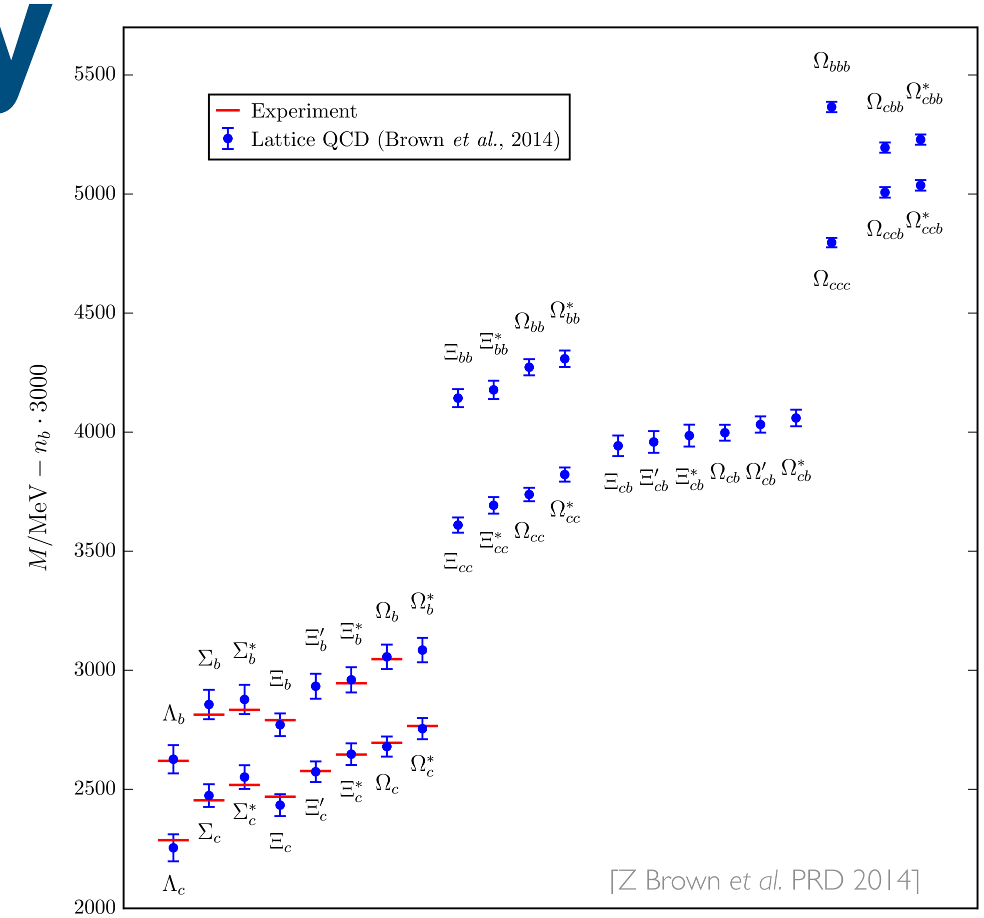


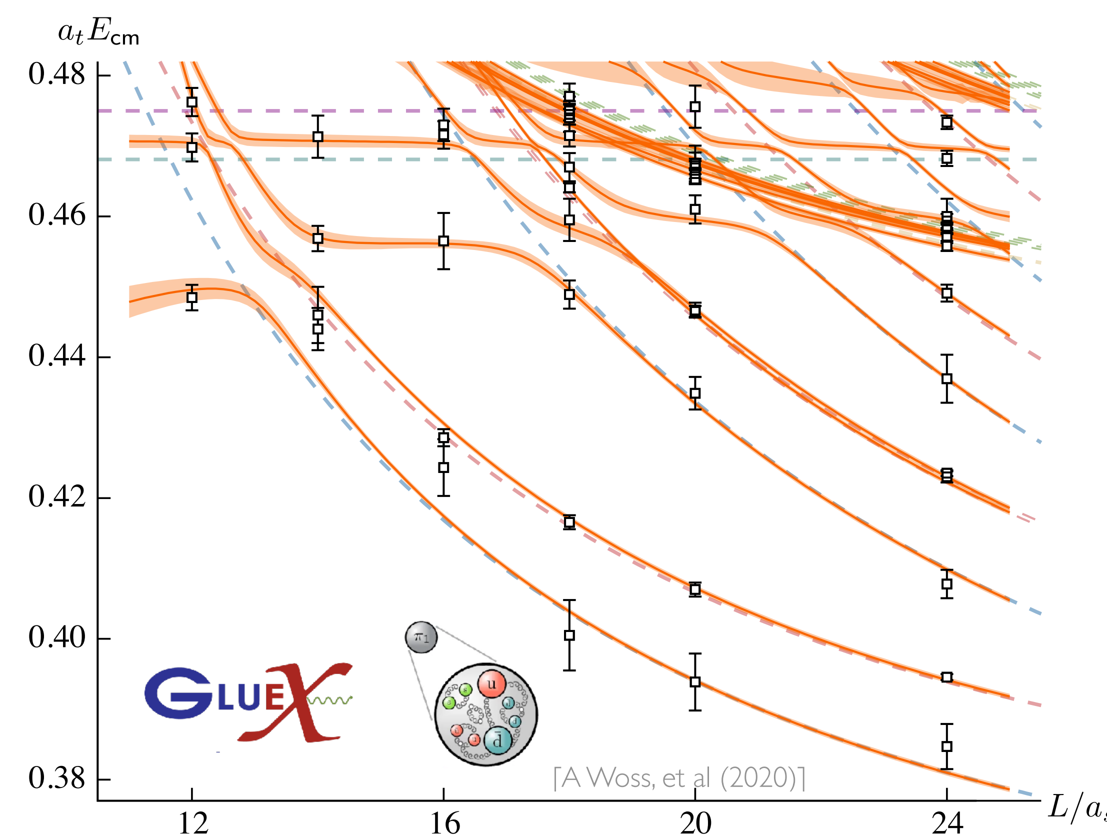
[Bazavov et al [HotQCD *Phys.Rev.D* 90 (2014) 094503]

Hadron Spectroscopy

Hadrons and resonances

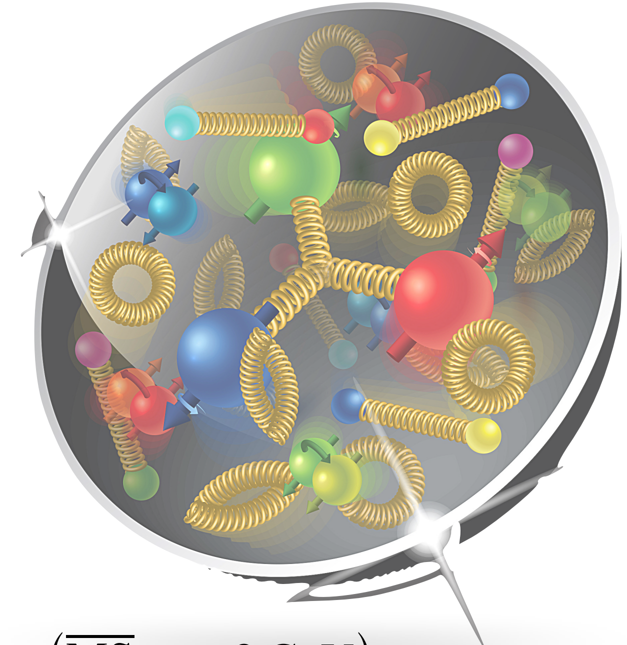
- LQCD reproduces masses of stable hadrons
    - Demonstrates QCD describes nature at low energies
    - Go beyond experiment: separate QED & isospin contributions to $M_n - M_p$
- Heavy hadrons
    - Predictions subsequently seen in experiment (eg Ω_b, Ξ_{cc})
- Predictions for exotic hadrons
    - Part of the motivation for the GlueX experiment
- Future:
    - More complicated decay channels
    - Branching fractions and structure of exotics



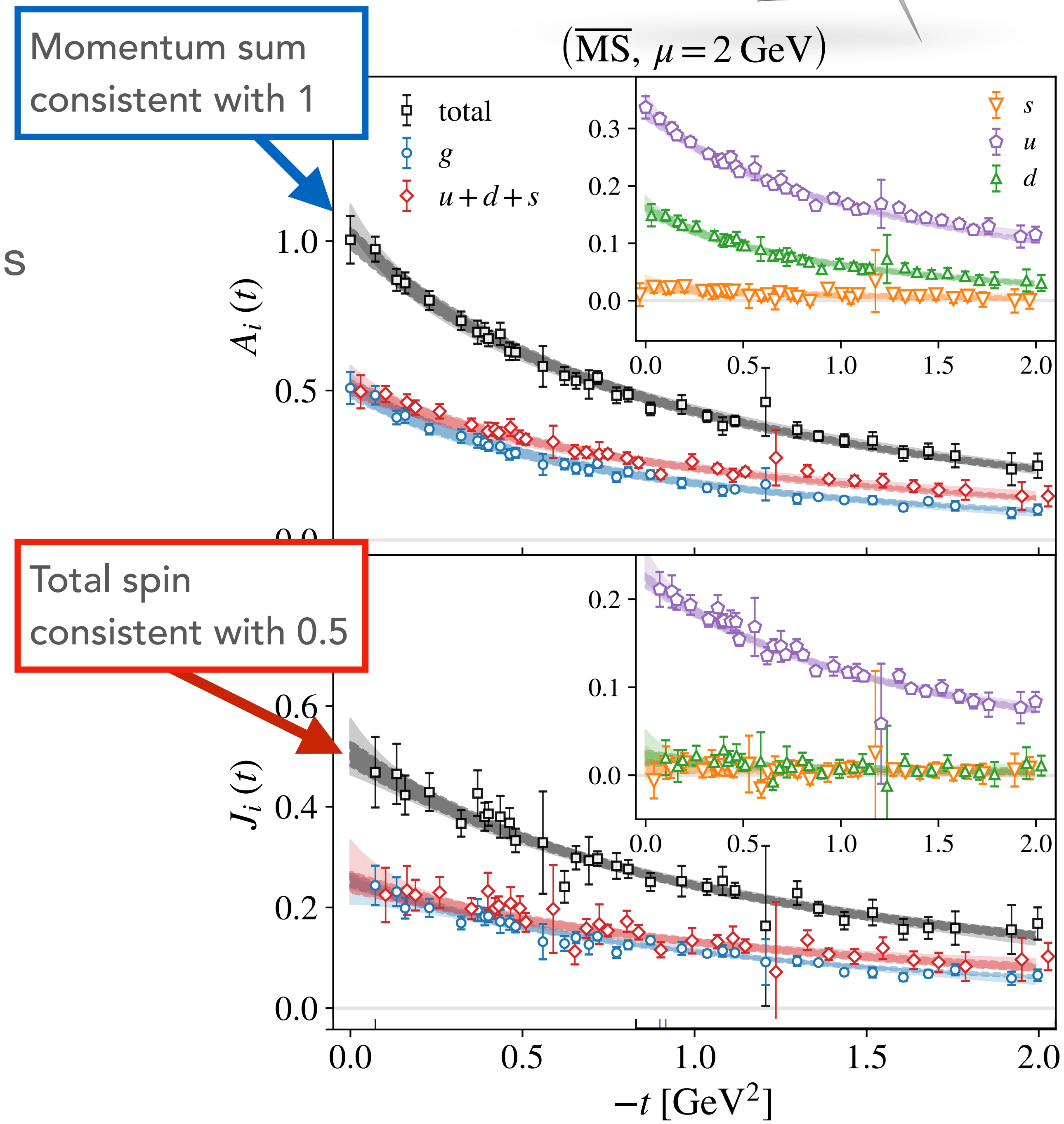


Hadron Structure

The internal dynamics of the proton



- QCD studies of EM form factors
- Partonic structure: PDFs, GPDs, TMDs
    - Key inputs for EIC projections and predictions
- Gravitational form factors & the energy-momentum tensor
    - Origin of mass, spin and mechanical properties
    - Complementary to JLab DVCS experiments
- Future:
    - Increased precision as needed to match experiment for simple quantities
    - Predictions of harder-to-measure quantities

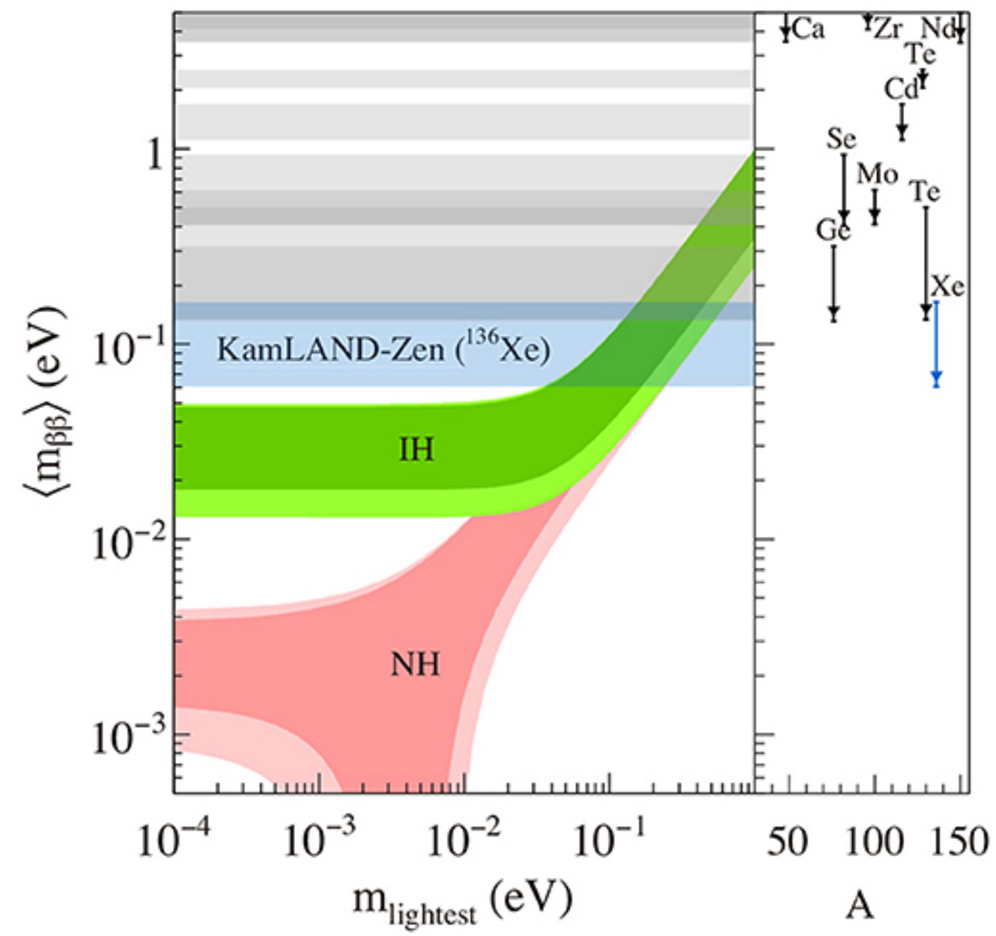


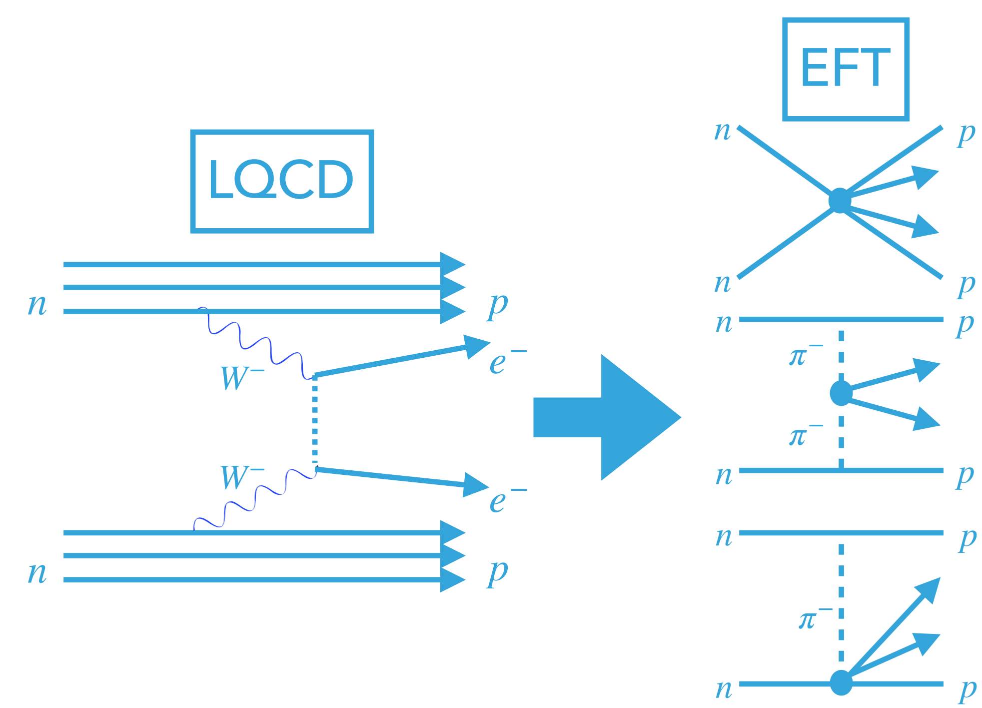
[Hackett, Pefkou, Shanahan PRL (2024)]

Fundamental Symmetries

Neutrinoless $\beta\beta$ decay

- Search for lepton number (L) violation/nature of neutrino [high priority in 2023 NSAC Long Range Plan]
- Uncontrolled nuclear model uncertainties cloud interpretation of experimental searches
- Ground nuclear physics in QCD: LQCD +EFT→ Many-body
- Calculations of subprocess $\pi^- \to \pi^+ e^- e^-$
  - Light Majorana neutrino
  - Short distance L-violation from BSM
- First QCD calculations of $nn \to ppe^-e^-$ for light Majorana neutrino
- Future: controlled inputs for future ton-scale $0\nu\beta\beta$ experiments

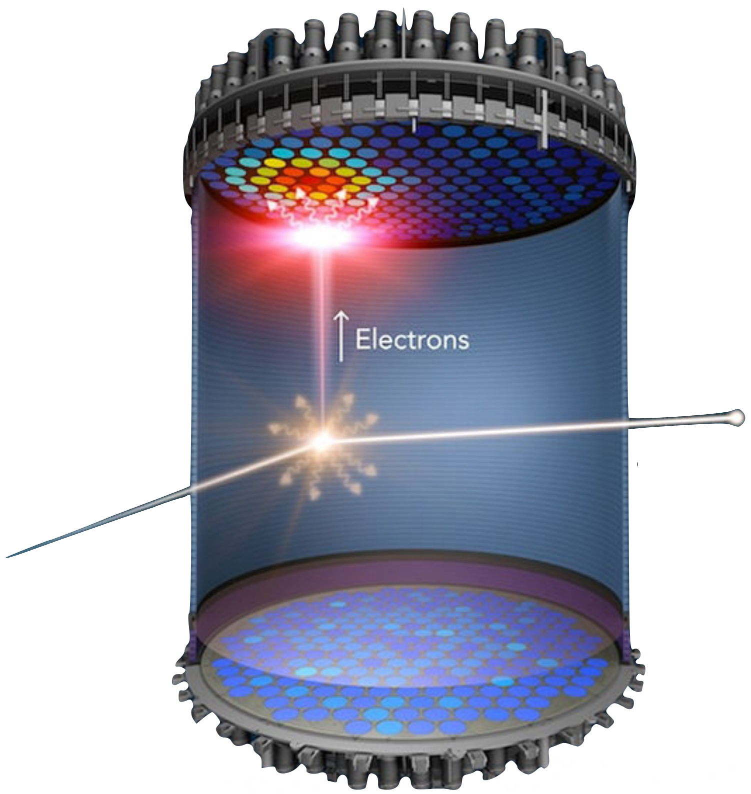


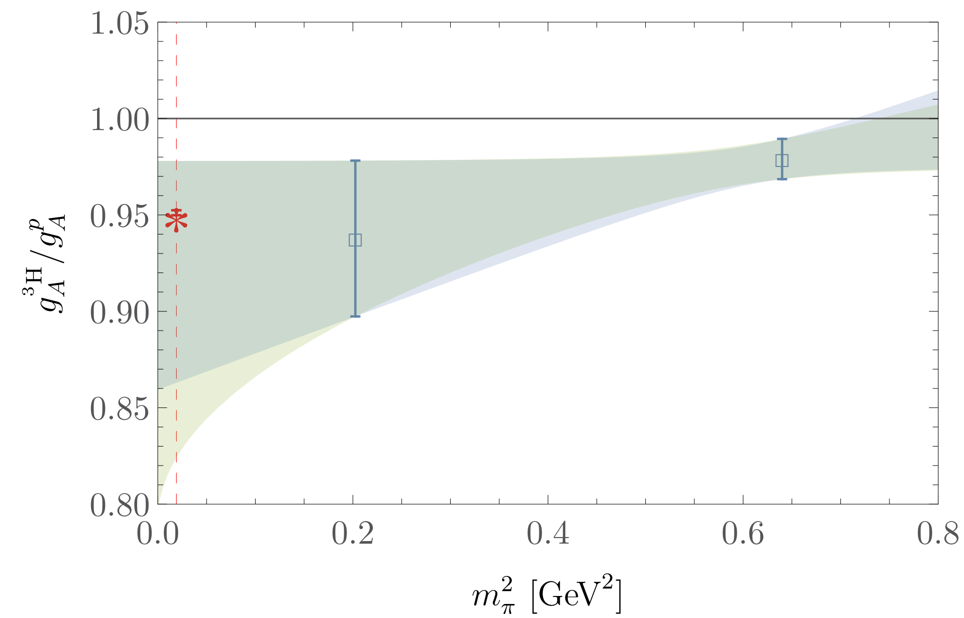


Beyond Standard Model physics

QCD @ the intensity frontier

- Many Intensity Frontier (HEP) experiments use nuclear targets
    - Neutrino physics: DUNE, HyperK,…
    - Dark matter: LZ, XENON-nT, PANDAX-4T,…
    - Lepton flavour violation: $\mu 2e$
- Must control QCD for rigorous BSM searches
- Current highlights:
    - Weak current nucleon form factors for DUNE
    - First calculations of nuclear effects
- Future: matrix elements for design and analysis of different detectors





[Parreño et al. [NPLQCD] *Phys.Rev.D* 103 (2021) 7, 074511]

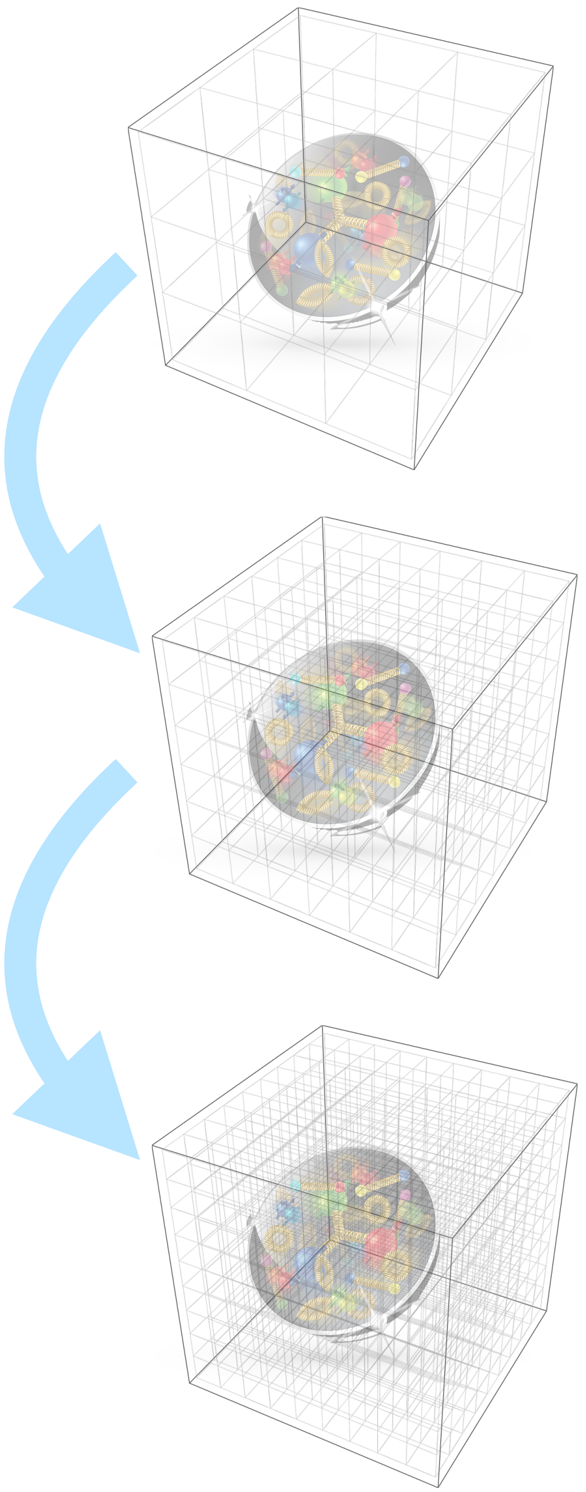
LQCD Science

Challenges and Developments

Precision physics

Continuum limit, QED, and isospin-breaking

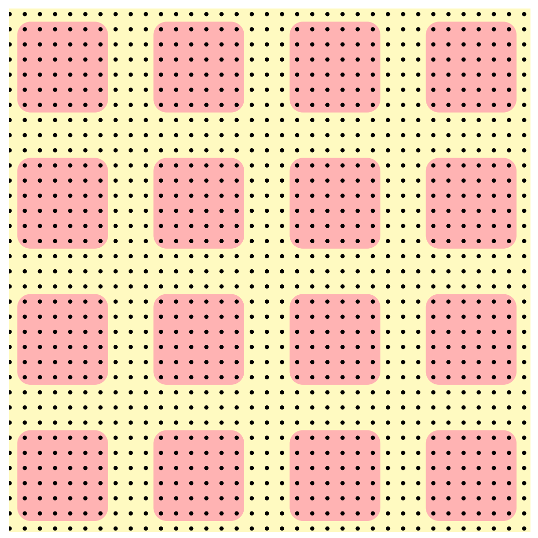
- Turning LQCD into physics requires continuum limit
  - Sampling becomes harder as update algorithms are quasi-local
  - New machine-learned normalising flow gauge generation algorithms seem promising
- <1% precision requires QED and effects of isospin breaking of quark masses ($m_u \neq m_d$)
  - Eg axial charge g_A precision limited by these effects
- New algorithmic requirements
  - Sampling with QED: complications of infinite volume limit
  - Sampling for $m_u \neq m_d$: Single fermion flavours

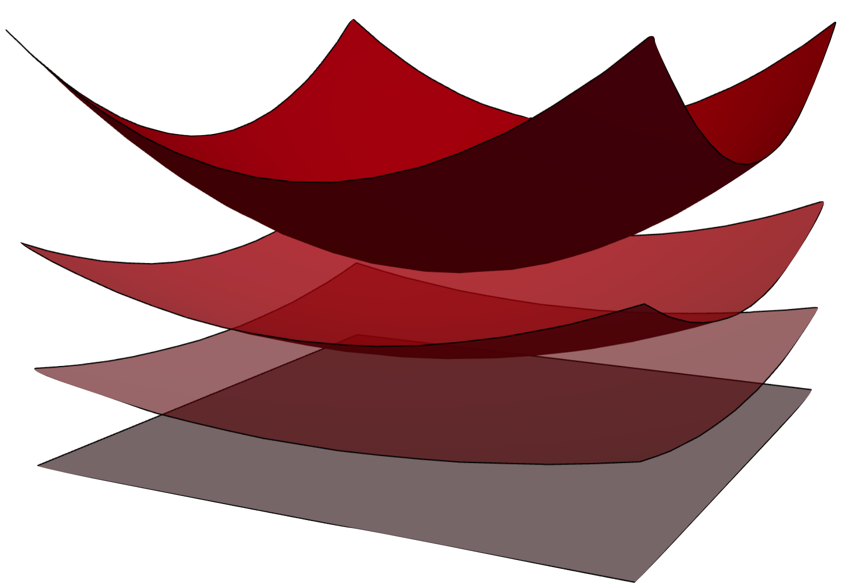


Statistical limitations

New sampling procedures

- LQCD is a Monte Carlo sampling method
    - High momentum structure of hadrons from LaMET/short distance factorisation requires very boosted states: statistical noise grows
    - Multi-hadron systems (eg nuclei): exponential growth of noise
- Better statistical sampling approaches
    - Hierarchical integration
    - Machine learned changes of integration variables

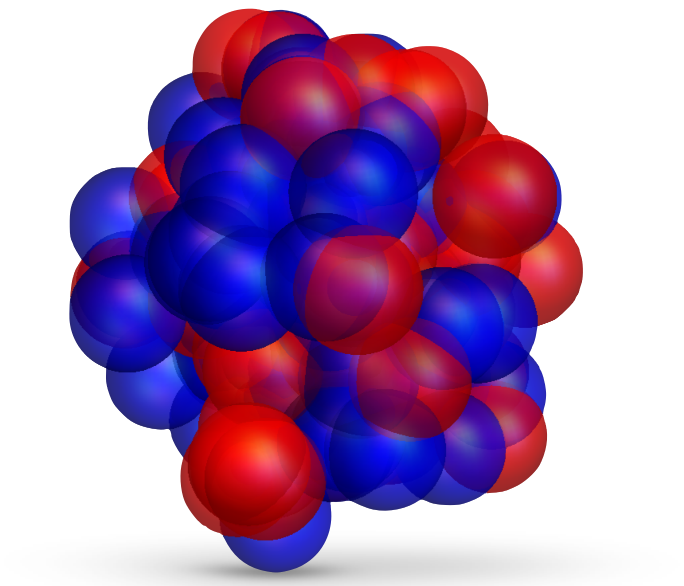


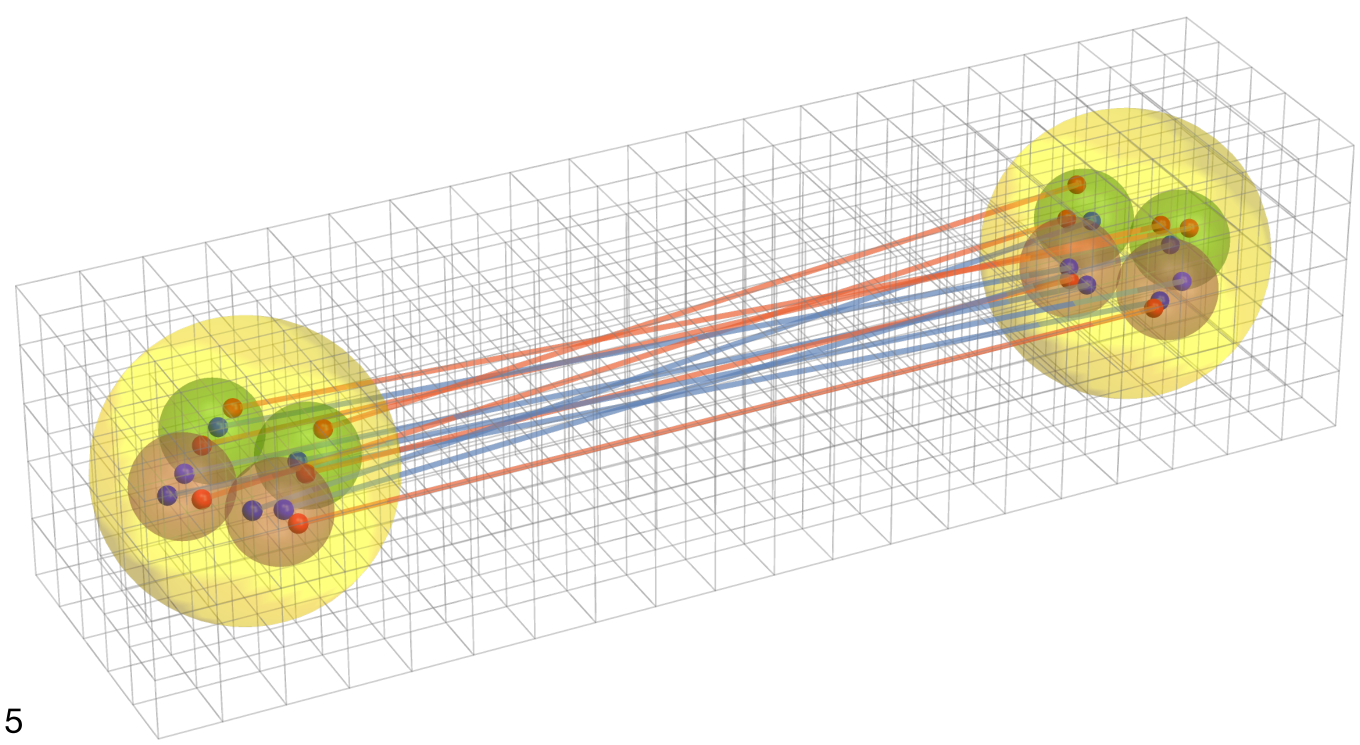


Complexity

Nuclear Contractions

- Physics of nuclei requires calculation of correlation functions with factorial number of Wick contractions: $N_u!N_d!N_s!$
  - Wick contactions = tensor contractions
  - Costs grow dramatically with system size
  - Graph-theoretic methods minimise redundant work
- Efficient calculation requires many tricks
  - Common subexpression elimination
  - Tiling of partial sums
  - Code generators
- Current state-of-the-art: $A = 5$ nuclei, but $A \sim 250$ in nature (a long way to go!)





LQCD Ecosystem

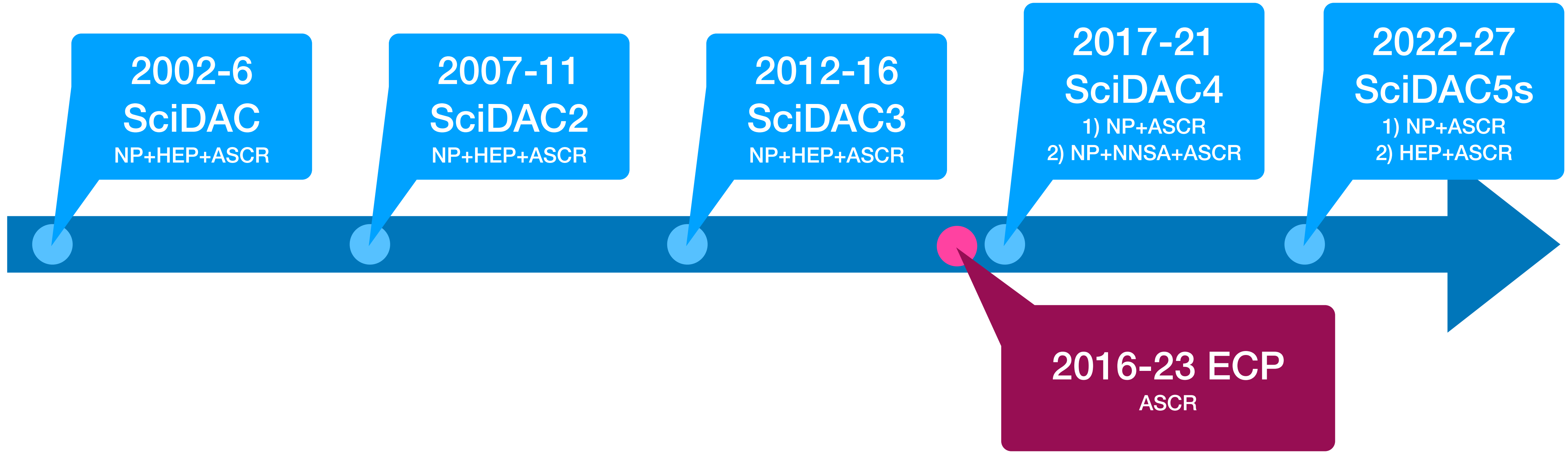
USQCD collaboration

- Federation of scientific collaborations
  - 90+% of LQCD (NP+HEP) researchers in the US
  - 180+ members from grad students to permanent staff
- Collaborate on software development
  - Train students and postdocs in LQCD
- Seek funding for, deploy, and allocate mid-scale computing resources at JLab, BNL and FNAL
  - Allocations support development of new projects and research directions
  - Build community and support junior scientists

LQCD Ecosystem

Software Development under SciDAC and ECP

- Strong DoE support for software development through SciDAC and ECP



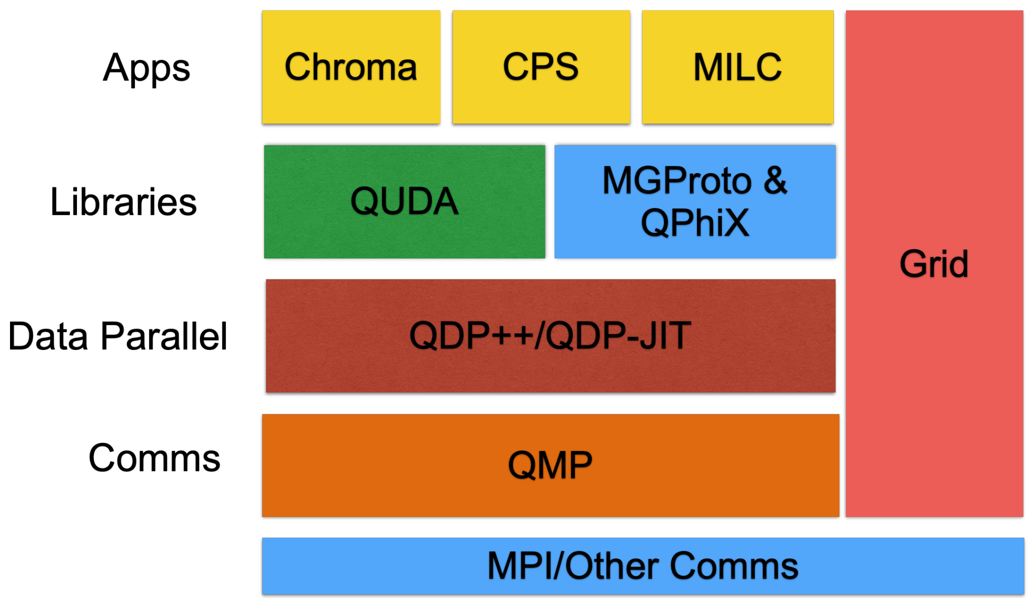
- Physics + Applied Math + Computer Science (SciDAC5 includes FastMath & RAPIDS SciDAC Institute collaborators)

- $50M+ investment in LQCD software and algorithms

LQCD Ecosystem

Codes and Libraries

- Layered software infrastructure
  - QMP: QCD message passing (abstracts different MPI frameworks)
  - QDP/C, QDP++, QDP-JIT: QCD data parallel
  - Linear system solver libraries: QUDA, MDWF, QPhiX,…
  - General application frameworks
    - Chroma, QLua, CPS, Grid, Grid Python Toolkit, MILC, Redstar, …
    - Most physicists work at this level
- Codes continue to evolve to address new physics problems

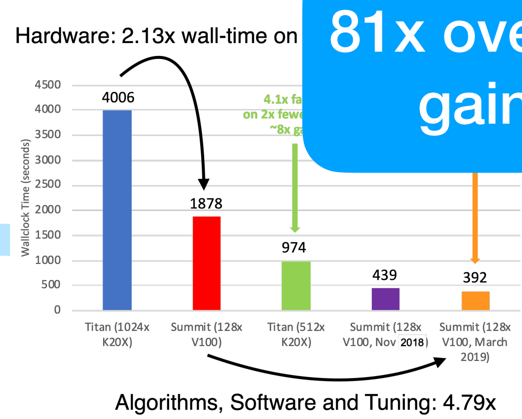


The USQCD software layers

LQCD Ecosystem

Codes and Libraries

- Example: porting Chroma from Titan to Summit to Frontier
  - Major changes to low level layers: QDP-JIT, QUDA
  - Smaller tweaks to application level code: Chroma



81x overall gain

113x now, in progress

LQCD Ecosystem

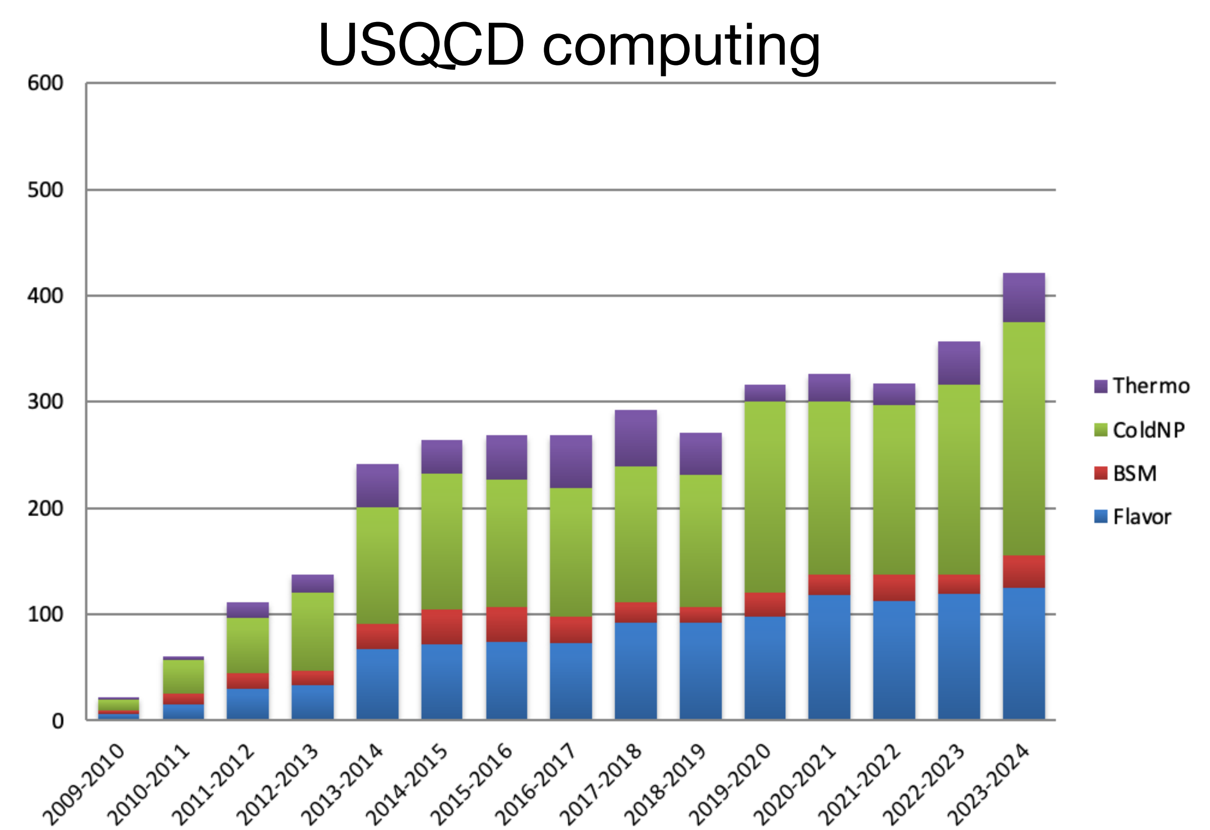
Codes and Libraries

- Connections to industry
- QUDA GPU library development based at nVIDIA and critical for LQCD
  - Ported to AMD and Intel GPUs with significant vendor assistance
- Code generators for new hardware (LLVM, Halide etc) build upon industry efforts
- Flow of talent from academia to industry helps maintain this
  - SciDAC supports development of the necessary skills

LQCD computing

Sources of computing

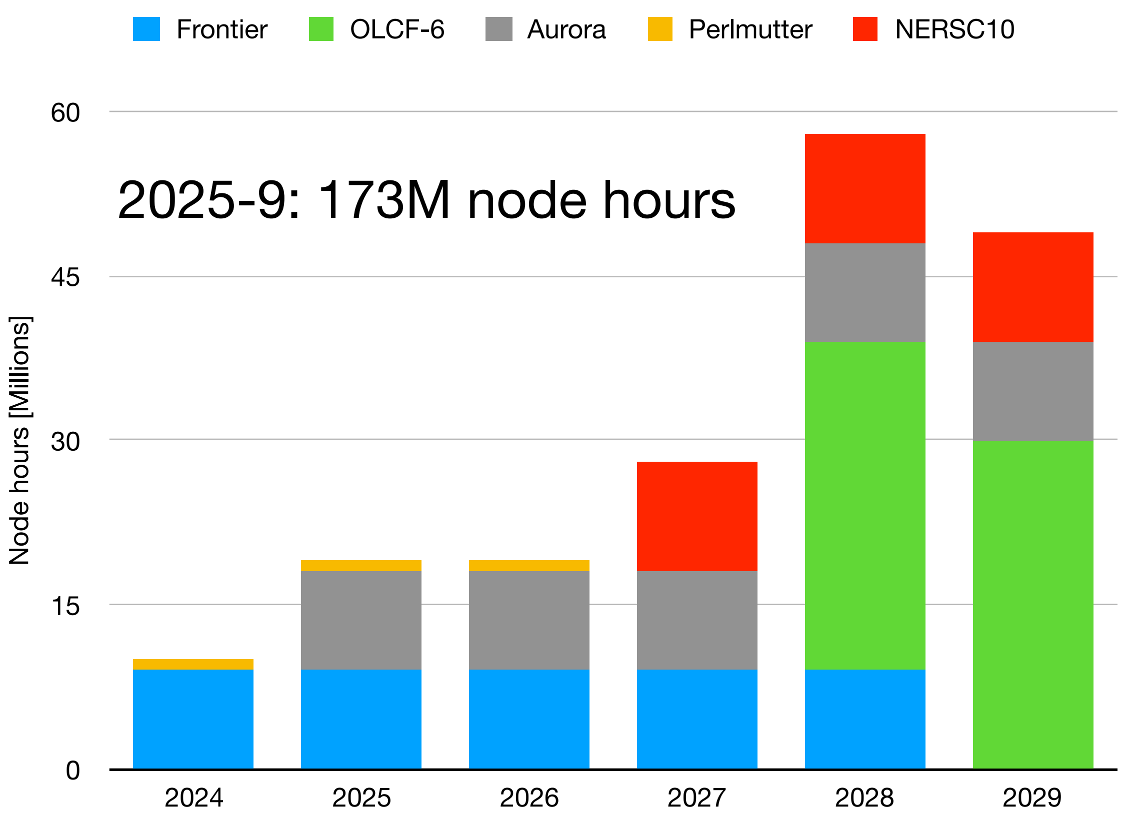
- INCITE 2020-4: 21 awards: 14.9M node hours
- ALCC 2019-23: 26 awards,14.1M node hours
- NERSC (2023: 21 PIs, 1.4M/1.0M CPU/GPU node hours)
- NSF: O(5) Frontera LRAC awards per cycle
- LQCD(HEP) & NPLCC(NP) hardware projects
  - 1/10th of INCITE in 2024
  - 25-30 projects per year



LQCD computing

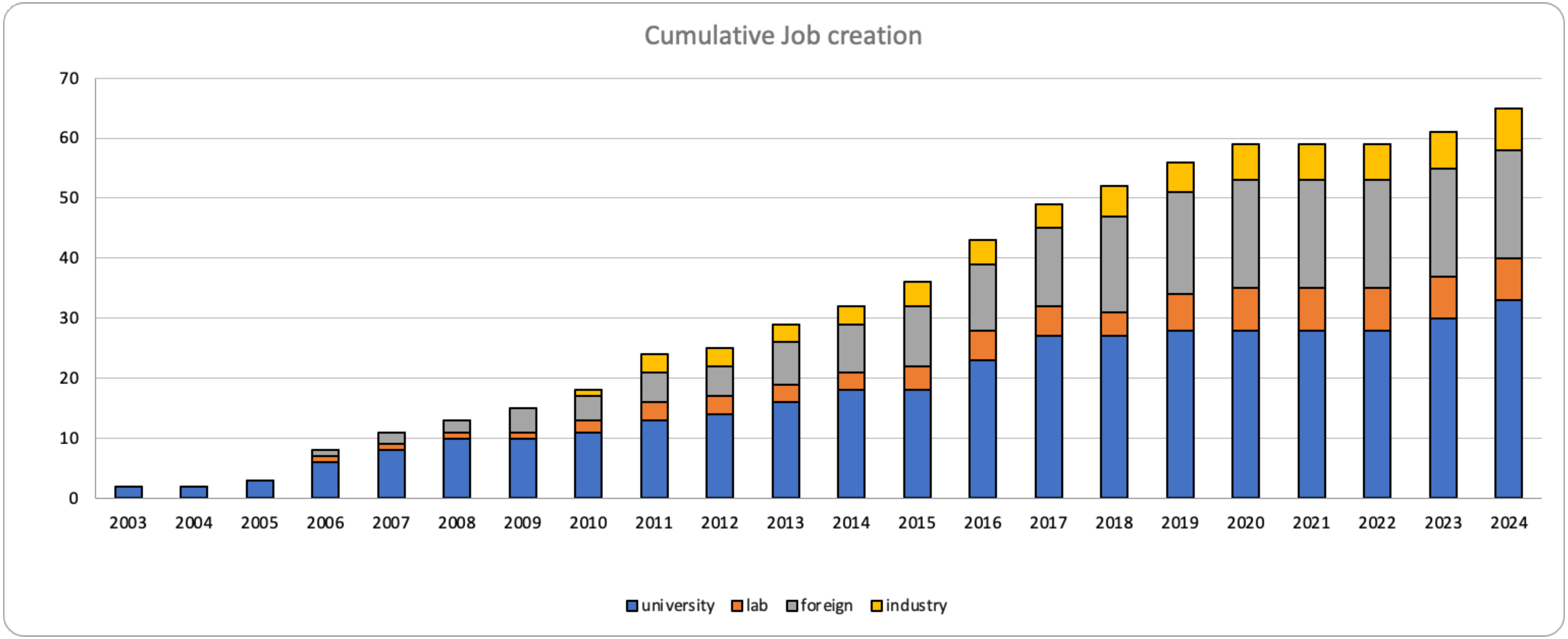
Computing 2025-30

- Projections for INCITE/ALCC/NERSC computing for LQCD
    - 173M Frontier-equivalent node hours over 2025-9
- USQCD maintains detailed timelines for all major projects
    - Project goals align with anticipated growth of computing
    - Assume support for code development on new systems



LQCD workforce
Workforce development

- University/Lab/NPC industry LQCD hiring



LQCD workforce

Workforce development

Early career awards

U.S. National Science Foundation

- Alexandru (GWU)
- Shanahan (MIT)
- Lin (MSU)
- Monahan (W&M)

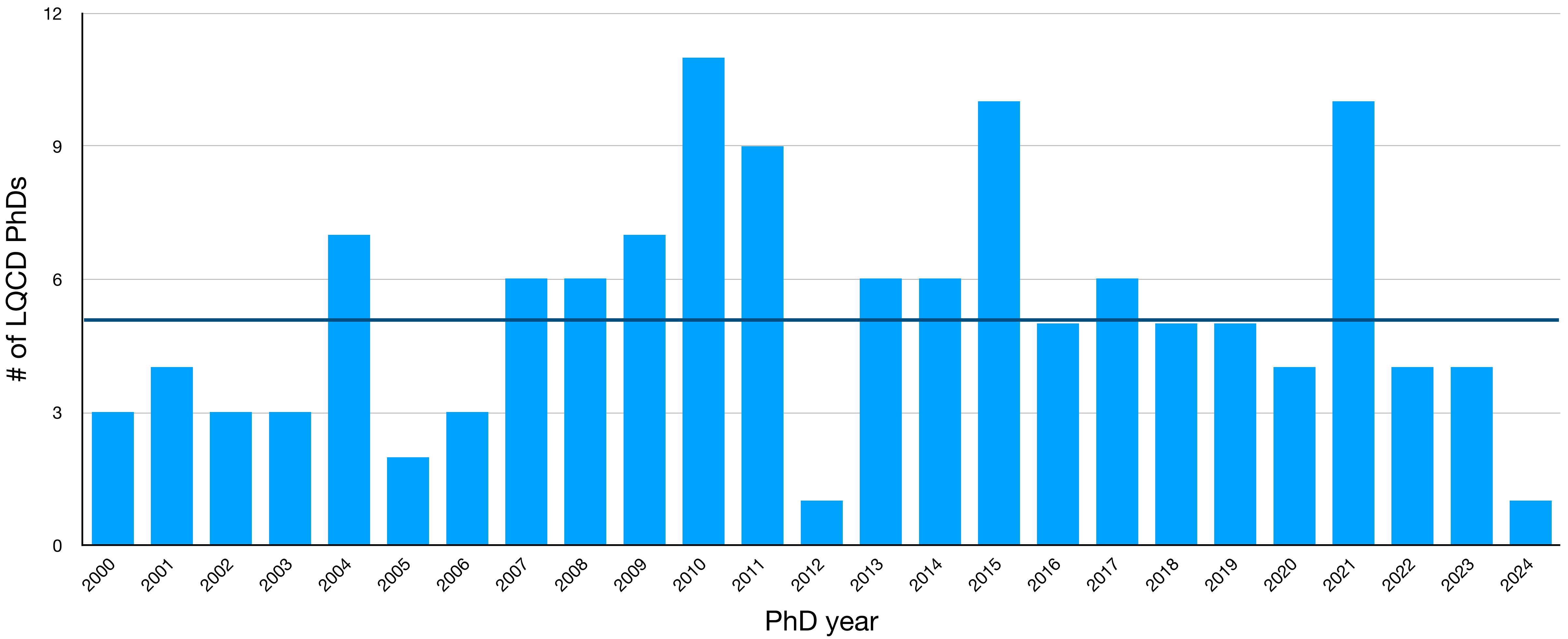
U.S. DEPARTMENT OF ENERGY | Office of Science

- Blum (UConn)
- Orginos (W&M)
- Dudek (W&M)
- Detmold (MIT, 2x)
- Walker-Loud (LBNL)
- Lehner (BNL)
- Davoudi (UMd)
- Jin (UConn)
- Constantinou (Temple)
- Briceño (Berkeley)
- Monahan (W&M)

LQCD workforce

Workforce development

- Training (average of 5.3) PhDs per year
- 35 current students



LQCD outlook

Some thoughts

- DOE SciDAC has been a successful pathway to support LQCD over the last 2+ decades: brings together a community and enables frontier scientific progress
  - Very much support its continuation
  - Broad scope for further innovation and abundant scientific goals
- ECP focus on code performance metrics and industry connections also very helpful
- Current framework (NP+ASCR) might not be necessary for smaller software projects, maybe just NP in some cases
- Storage of (and access to) community resources such as gauge fields is vital but lacking infrastructure
- Machine learning
  - Off the shelf ML tools will rarely solve the problems of the NP domain - interaction with data scientists needs to be two-way and we need to be prepared to develop problem-specific ML
  - Need for new allocation mechanisms to support ML applications where outcomes are not as predictable as with traditional modelling/simulation workloads
- Evaporating talent: flat budgets for many years and recent NT budget cuts do not provide an attractive recruiting environment
  - Students and postdocs leave the field for better-paying industry positions (8 LQCD researchers now @ nVIDIA)
  - Advocate for support for computational scientists at universities: very few mechanisms for scientific computing specialists at universities and technically minded postdocs who focus on software often leave the field