Computational Challenges in Lattice QCD

Vera Gülpers

School of Physics and Astronomy
University of Edinburgh

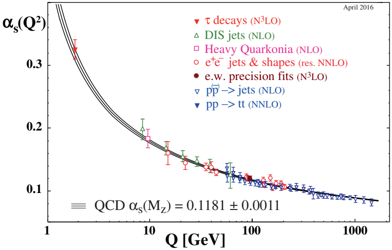
February 18, 2020

European Research Council
Established by the European Commission

THE UNIVERSITY of EDINBURGH

QCD and confinement

- Quantum Chromo Dynamics (QCD) theory of the strong interaction
- strong coupling α_s

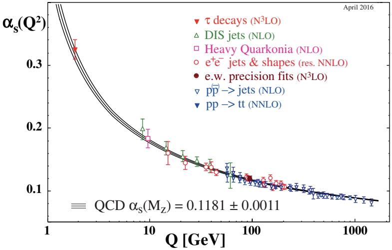


[Particle Data Group (PDG), Phys. Rev. D 98, 030001 (2018)]

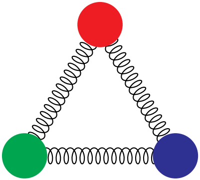
- quarks and gluons confined to bound states (hadrons)

QCD and confinement

- Quantum Chromo Dynamics (QCD) theory of the strong interaction
- strong coupling α_s



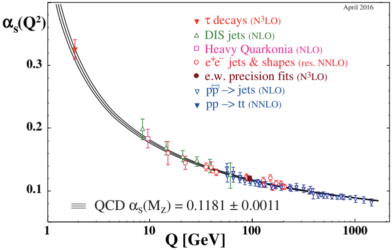
[Particle Data Group (PDG), Phys. Rev. D 98, 030001 (2018)]



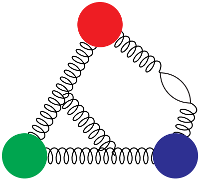
- quarks and gluons confined to bound states (hadrons)

QCD and confinement

- Quantum Chromo Dynamics (QCD) theory of the strong interaction
- strong coupling α_s



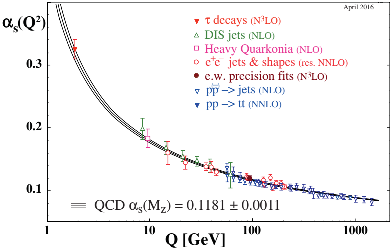
[Particle Data Group (PDG), Phys. Rev. D 98, 030001 (2018)]



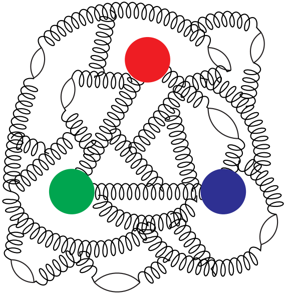
- quarks and gluons confined to bound states (hadrons)

QCD and confinement

- Quantum Chromo Dynamics (QCD) theory of the strong interaction
- strong coupling α_s



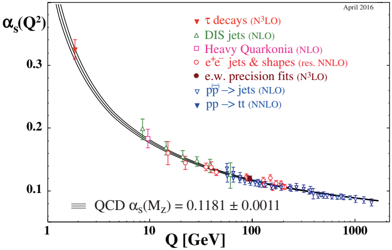
[Particle Data Group (PDG), Phys. Rev. D 98, 030001 (2018)]



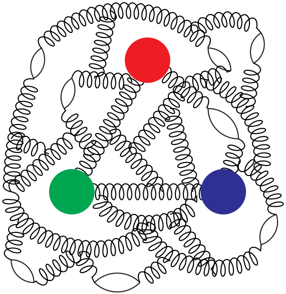
- quarks and gluons confined to bound states (hadrons)

QCD and confinement

- Quantum Chromo Dynamics (QCD) theory of the strong interaction
- strong coupling α_s



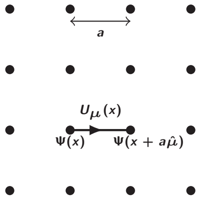
[Particle Data Group (PDG), Phys. Rev. D 98, 030001 (2018)]



- quarks and gluons confined to bound states (hadrons)
- each additional gluon line or quark-antiquark pair comes with α_s ($\alpha_s \sim \mathcal{O}(1)$ at small energies)

  $\rightarrow$ Monte Carlo sampling

QCD on the lattice



- Wick rotation $(t \to -ix_0)$ to Euclidean space-time
- Discretize space-time by a hypercubic lattice Λ
- Quantize QCD using Euclidean path integrals

$$\langle A \rangle = \frac{1}{Z} \int \mathcal{D}[\Psi, \overline{\Psi}] \mathcal{D}[U] \, e^{-S_E[\Psi, \overline{\Psi}, U]} \, A(U, \Psi, \bar{\Psi})$$

$\longrightarrow$ can be split into fermionic and gluonic part

- Calculate gluonic expectation values using Monte Carlo techniques:

$$\langle \langle A \rangle_F \rangle_G = \int \mathcal{D}[U] \, \langle A \rangle_F \, P(U) \approx \frac{1}{N_{\text{cfg}}} \sum_{n=1}^{N_{\text{cfg}}} \langle A \rangle_F$$

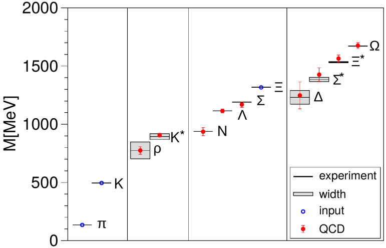
average over gluonic gauge configurations U distributed according to

$$P(U) = \frac{1}{Z} (\det D)^{N_f} \, e^{-S_G[U]}$$

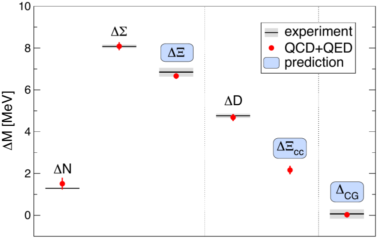
- extrapolate to the continuum $(a \to 0)$ and infinite volume $(V \to \infty)$

QCD on the lattice

- successfully used to calculate hadronic observables, e.g. hadron spectrum



[S. Dürr *et al*, Science 322 (2008) 1224-1227]



[Sz. Borsanyi *et al*, Science 347 (2015) 1452-1455]

- other quantities, e.g.
  hadronic form factors, decay amplitudes, hadronic contributions to $g-2$, PDFs, ...

Computational Challenges

- two energy scales in the problem, box size $\boldsymbol{L}$, lattice spacing $\boldsymbol{a}$

$$\mathcal{O}(1/L) \ll E \ll \mathcal{O}(1/a)$$

- typical size of a lattice

$$N = L^3 \times T = 64^3 \times 128 \sim \mathcal{O}(10^7 - 10^8)$$

- Dirac-operator $\boldsymbol{D}$: matrix of size $(12 \cdot N) \times (12 \cdot N)$

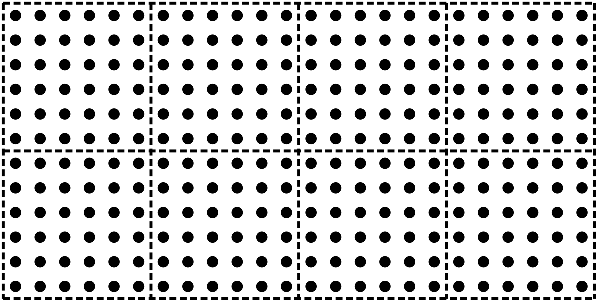
- anatomy of a lattice calclation
  - generate gauge configurations $\rightarrow$ need determinant **det** $\boldsymbol{D}$
  - "measurements": calculate quark propagators $\rightarrow$ need the inverse $\boldsymbol{D}^{-1}$
    $\rightarrow$ solve the Dirac equation using appropriate sources $\boldsymbol{\eta}$

$$D\phi = \eta$$

- size of lattices restricted by available computer power

Parallelisation of the Problem

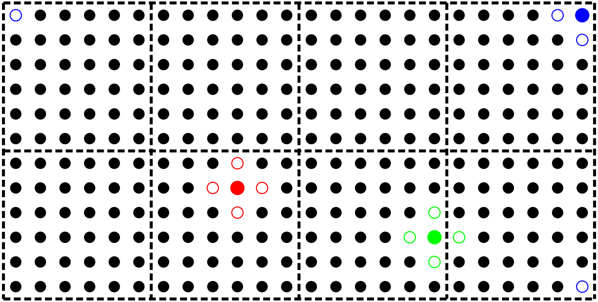
- split lattice into smaller local lattices



- need good scalabilty up to large number of compute nodes
  $\rightarrow$ balance between
  - time for computation on local lattice
  - time for communication
  - time for loading data from memory to cache
  - time for using data from cache

Parallelisation of the Problem

- split lattice into smaller local lattices

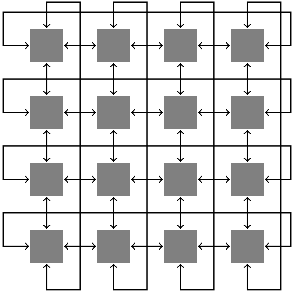


- need good scalabilty up to large number of compute nodes

- Dirac operator is a sparse matrix
- mostly nearest neighbour communication between processes

Supercomputers for Lattice QCD

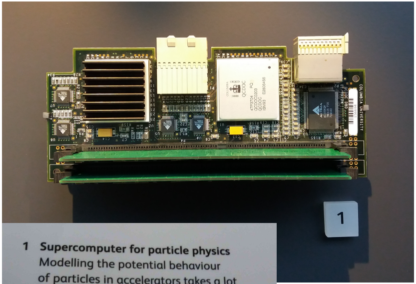
- torus interconnect

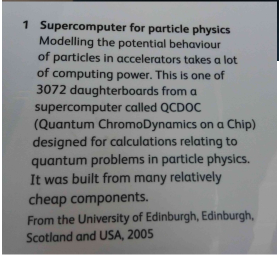


- "tradition" of custom designed supercomputers tailored to lattice calculations, e.g.
  - QCDSP ("QCD on digital Signal Processors"), designed in the 90's
    $\rightarrow$ 4-dimensional torus
  - QCDOC ("QCD On a Chip")
    $\rightarrow$ 6-dimensional torus

QCDOC ("QCD On a Chip")

- devloped by IBM, University of Edinburgh, Columbia University and RIKEN-BNL see, e.g., [D. Chen *et al*, Nucl.Phys.Proc.Suppl. 94 (2001) 825-832], [P. Boyle *et al*, Nucl.Phys.Proc.Suppl. 129 (2004) 838-843]
- application-specific integrated circuit
- custom-designed communications hardware
- $\approx$ **1** GFlop/s per node
- sustained performance of $\approx$ **50%**
- QCDOC in Edinburgh
  - installed in 2004
  - running until 2011
  - **12288** processing nodes
    $\rightarrow \approx$ **10** TFLOPS/s peak





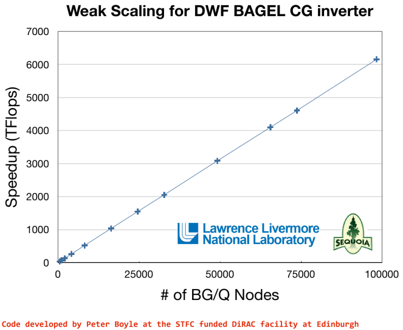
IBM BlueGene/Q

- three generations of supercomputers: BG/L, BG/P, BG/Q
- Top 500 June 2012 [https://www.top500.org/lists/2012/06/]: Nr 1,3,7,8 BG/Q

- 18 cores per chip

- L1 (**16 + 16** KB/core) and L2 (**32** MB/node) cache
- L1p prefetch engine designed by Edinburgh and Columbia

- interconnect: high-bandwidth 5d torus network



[P. Boyle, PoS LATTICE2012 (2012) 020]

- fixed local volume $\mathbf{8^4} \times \mathbf{16}$

GFlops/s	L1 GB/s	L2 GB/s	DDR GB/s	Torus GB/s
204.8	**820**	**536**	**42.7**	**40**

[P. Boyle, PoS LATTICE2012 (2012) 020]

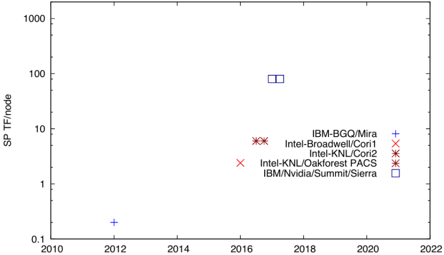
IBM BlueGene/Q in Edinburgh

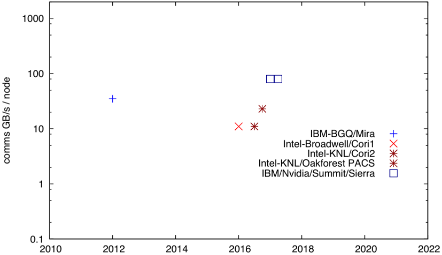
- BlueGene/Q in Edinburgh
- part of the DiRAC supercomputing facility
- hosted by Edinburgh Parallel Computing Centre (epcc)
- 6144-node (6 racks) BlueGene/Q with 1.26Pflop/s peak performance (Nr. 20 in Top500 List, June 2012)
- running from 2012 until January 2018

[P. Boyle, PoS LATTICE2012 (2012) 020]

Current Supercomputers

- performance and communication bandwidth for recent supercomputers





[P. Boyle, PoS LATTICE2016 (2016) 013]

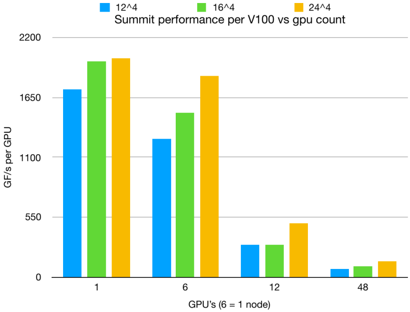
- computational intensity for lattice algorithms typically $\sim \mathbf{1}$ Byte/Flop
- single node performance increasing much faster than interconnect bandwidth $\rightarrow$ affects scaling of lattice calculations

GPU

- typically many processing units per GPU
  (picture of NVIDIA Volta V100 die on the right)

[https://images.anandtech.com/doci/11367/voltablockdiagram.png

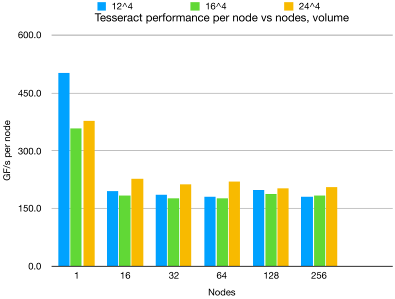
- GPU in supercomputers, e.g.
- Summit (Oak Ridge National Laboratory):
  - Nr 1 in Top500 (**200** PFlops)
  - 4608 compute nodes, each with
    - two IBM POWER9 CPUs
    - six Nvidia Volta GV100 GPUs
    - connected by NVLink
  - Dual-rail Mellanox EDR Infiniband



[talk by P. Boyle, USQCD All-Hands Collaboration Meeting 2019]

Tesseract

- Extreme scaling service of DiRAC, hosted by EPCC in Edinburgh
- **1468** compute nodes, each with two Intel Xeon Silver 4116 processors
- Omnipath interconnect



[talk by P. Boyle, USQCD All-Hands Collaboration Meeting 2019]

picture from https://www.epcc.ed.ac.uk/facilities/dirac

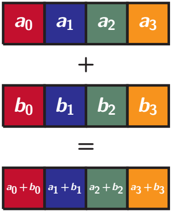
- new: **8** GPU compute nodes:
  two Intel Xeon and four NVidia V100 (Volta) GPU accelerators

Software for Lattice QCD

- exploit all forms of parallelism
  - multi processing
    - communication between different nodes
  - multi threading
    - execute several instructions simultaneously
    - parallelisation within a node
  - single instruction multiple data (SIMD)
    - simultaneously perform the same operation on a vector of data
    - parallelisation within a core
- Software that is aware of the hardware, e.g.
  - network geometry
  - vector instructions
  - memory hierarchy, cache reuse
- reduce message size in communication, e.g.
  - communication avoiding algorithms
  - reduced data precision

Single Instruction Multiple Data (SIMD)

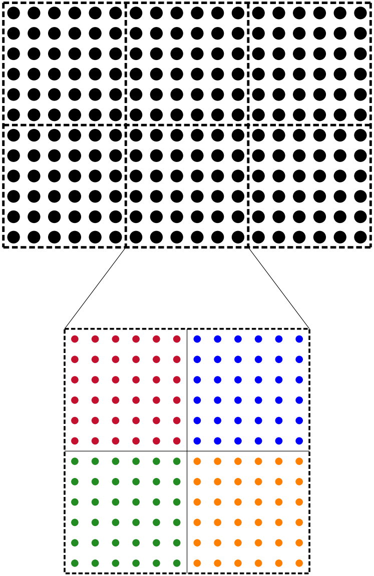
- SIMD as vector instructions



- various instruction sets, e.g.

Intel	SSE (128), SSE2 (128)
Intel	AVX (256), AVX2 (256), AVX512 (512)
IBM PowerPC	QPX (256)
. . .	

- SIMD for Lattice
- data layout: SIMD vector



GRID Software package

- Free (GPLv2) data parallel C++11 QCD library
- P. Boyle *et al*, `https://github.com/paboyle/Grid`
  [P. Boyle *et al*, PoS LATTICE2015 (2016) 023]

GRID

- high level data parallel approach: MPI, OpenMP and
- various vector instructions, SSE, AVX, AVX2, AVX512, QPX, NEONv8
- adjust data layout automatically to vector length of given architecture
- high portability between many architectures
- aims to performance on upcoming/planed exascale machines

- New: GPU-Port

- various different solvers and lattice actions

Getting ready for Exascale

- Top500 List, November 2019 https://top500.org/lists/2019/11/

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	**Summit** - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM DOE/SC/Oak Ridge National Laboratory United States	2,414,592	148,600.0	200,794.9	10,096
2	**Sierra** - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	1,572,480	94,640.0	125,712.0	7,438
3	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway , NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371

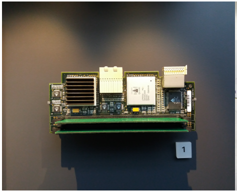
- Exascale will come soon
  - Aurora at Argonne National Laboratory, US
    [https://www.intel.co.uk/content/www/uk/en/highperformancecomputing/supercomputing/exascalecomputing.html
  - Frontier at Oak Ridge National Laboratory, US
    [https://www.olcf.ornl.gov/2018/02/13/frontier-olcfs-exascale-future/
  - Tianhe-3, China?
    https://www.top500.org/news/china-reveals-third-exascale-prototype/

Summary

- Lattice QCD is a successful tool

- computationally very expensive
  - supercomputers tailored for lattice calculations
  - need good scalabilty
  - need highly optimised software

- Exascale will come soon

Thank you

VG has received funding from the European Research Council (ERC) under the
European Union's Horizon 2020 research and innovation programme under grant agreement No 757646.