Lattice Gauge Theories at the Exascale Frontier: Part 1

Biagio Lucini

SWIFT-HEP and ExaTEPP joint event, Warwick (UK), 11th-12th November 2024

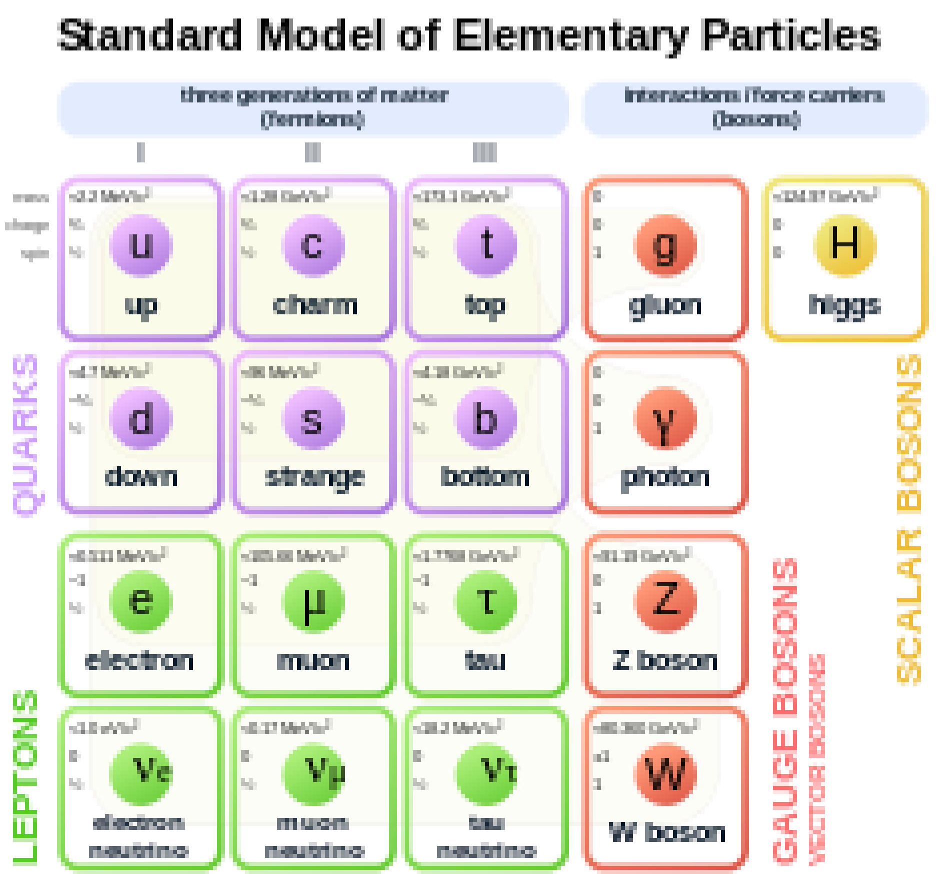
Overview

- A very quick introduction to the Lattice
- Numerical strategies for the Lattice
- Current physics calculations using the Lattice

The science programme

The Standard Model of Particle Physics

Prifysgol Abertawe
Swansea University



- The SM describes the strong and electroweak interactions (including the Higgs sector) very successfully
- The electroweak interactions are amenable to analytical predictions
- The strong interactions require non-perturbative ab initio methods

The Lattice as a computational paradigm

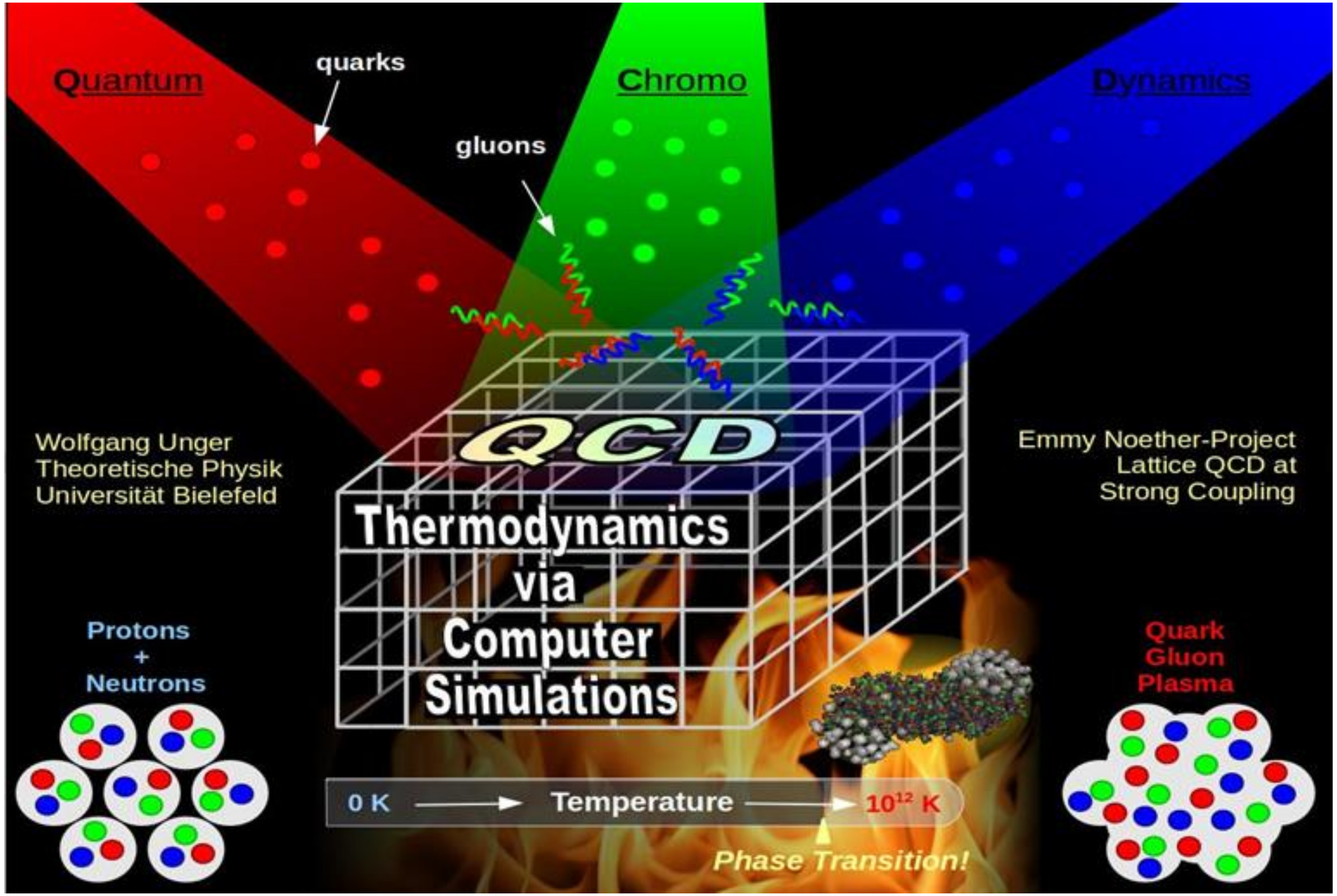


Image credits: Wolfgang Unger

- The theory gives rise to a spectrum of hadrons (baryons, mesons, glueballs and exotica) at zero temperature
- Deconfinement happens at a critical/crossover temperature
- Exotic phases at non-zero density
- All these are non-perturbative phenomena

Enters the Lattice…

Open problems in the Standard Model

- Gravity is not accounted for
- Asymmetry matter-antimatter
- Absence of CP violation in QCD
- Dark matter/dark energy
- Fundamental mechanism for electroweak symmetry breaking
- …

Investigation paradigms

1. Precision frontier

   Observables are computed theoretically and determined experimentally to very high accuracy, with deviations providing evidence for new physics and agreement setting stringent bounds for the latter

   The Lattice provides a robust calculation tool for precision physics

2. Energy frontier

   Theoretically motivated interactions beyond the standard model are studied, with their observables providing input to phenomenology

   Non-perturbative calculation on a lattice often crucial

From the continuum to the lattice (and back)

1. Start from the Euclidean Path Integral formulation of the theory

$$\langle O \rangle = \frac{\int (\mathcal{D}\phi)\, O[\phi] e^{-S}}{\int (\mathcal{D}\phi)\, e^{-S}}$$

2. Approximate the integral on a grid of spacing a and of size $V = N_t \times N_s^3$

3. Compute the integral with Monte Carlo methods

4. Extrapolate to $V \rightarrow \infty$ and $a \rightarrow 0$

Fields on the lattice

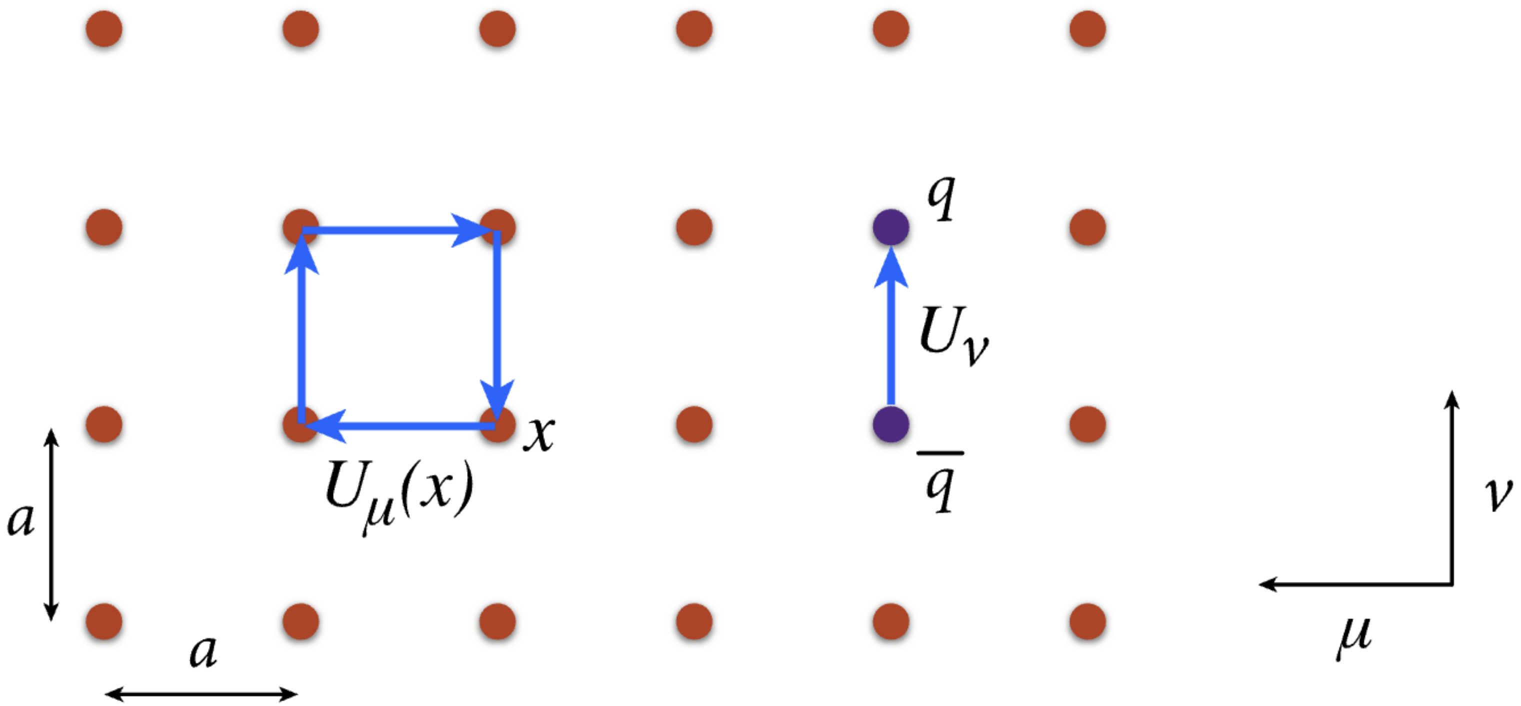


Figure from Particle Data Group

Example: mass determination

$$\langle \phi^\dagger(0)\phi(x)\rangle = \sum_n |c_n|^2 e^{-m_n x}$$

$$\overset{x\to\infty}{=} |c_0^2| e^{-m_0 x}$$

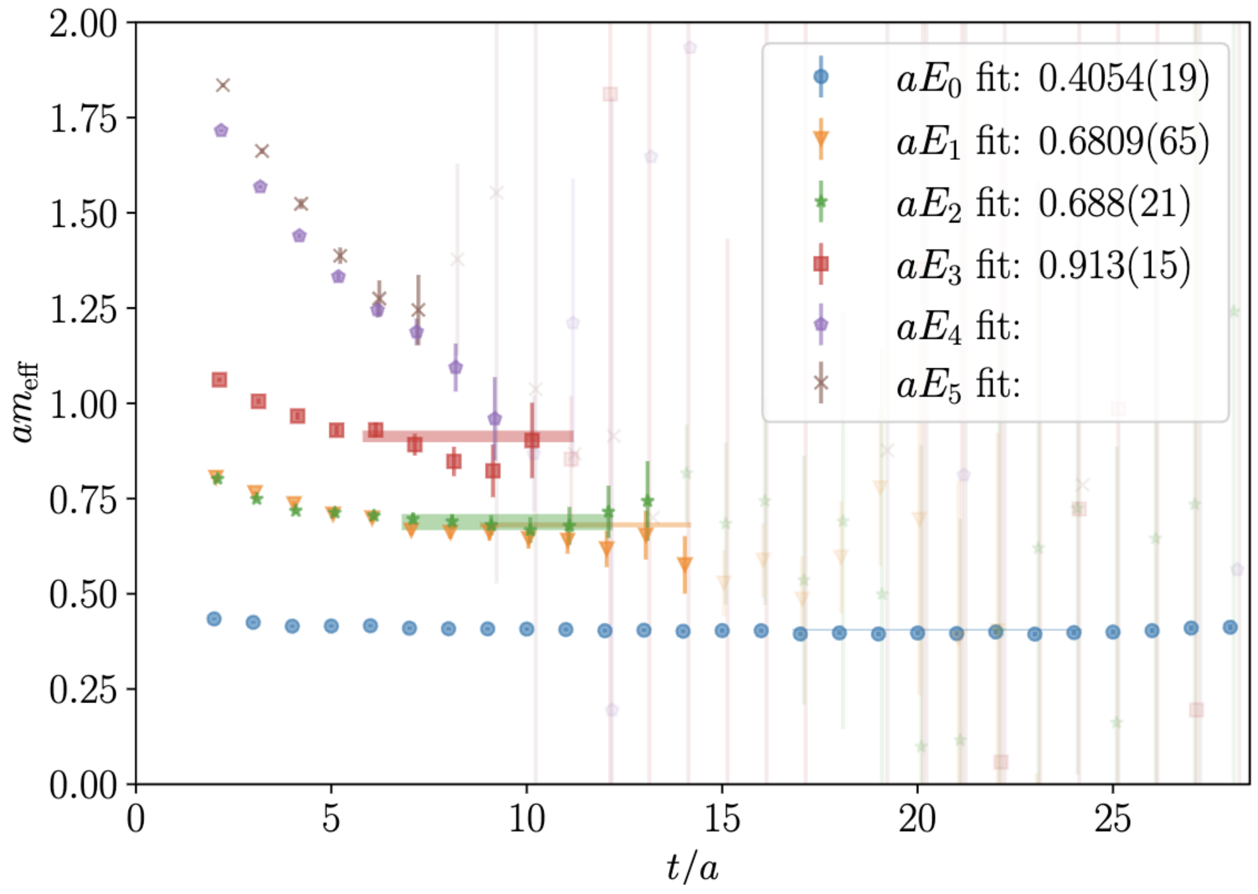
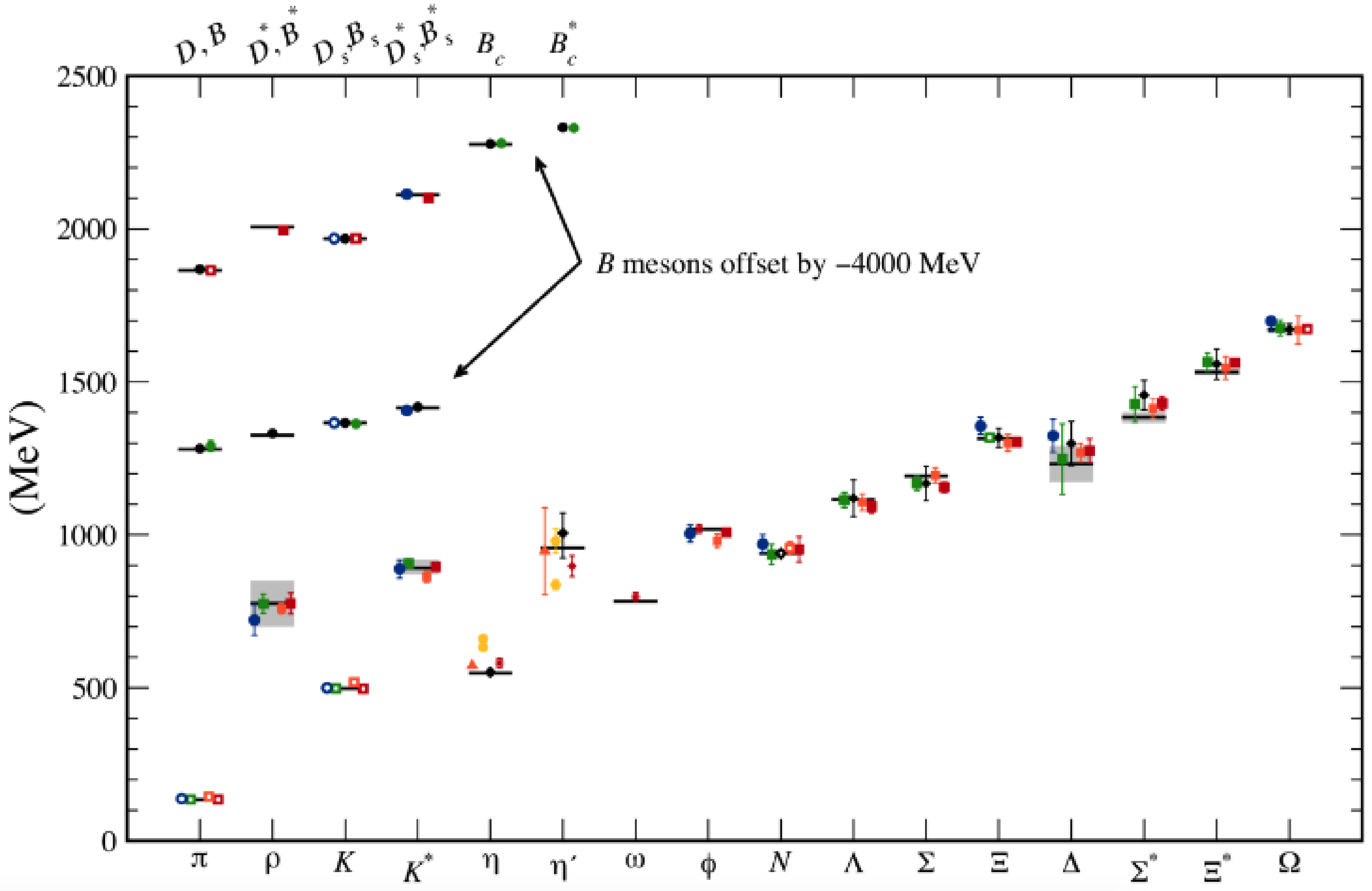


Figure from E. Bennett et al., Phys. Rev. D 110 (2024) 7, 074504 (arXiv:2405.05765)

Status of Lattice QCD



Amsler, Creder and DeGrand Particle Data Group, 2023

Computational strategies for the Lattice

Algorithms

- Goal: sample the Euclidean path integral on a grid and reconstruct continuum physics from it
- Work flow can be broken in two main (distinct) components: generation of gauge configurations via MCMC and measurements of observables
- Generally the computational cost generation:measurements is between 80:20 and 90:10
- For the generation, the bulk of the time is spent on the inversion of a large sparse matrix
- Details are different, as there are a few community codes taking a different angle in order to optimally target different Physics applications
- Broadly speaking, the main algorithm is conjugated gradient (and improvements), which is accelerated with preconditioners

Computational challenges

- $O(10^9)$ SU(3)/SU(N)/Sp(2N) matrices at each step
- Need to invert sparse matrix of ~ this linear size
- Hierarchic programming: low-level vs accelerator vs shmem vs distmem
- Need to interlace communication and computation
- Equivalent different formulations of the physical problem implemented in community codes

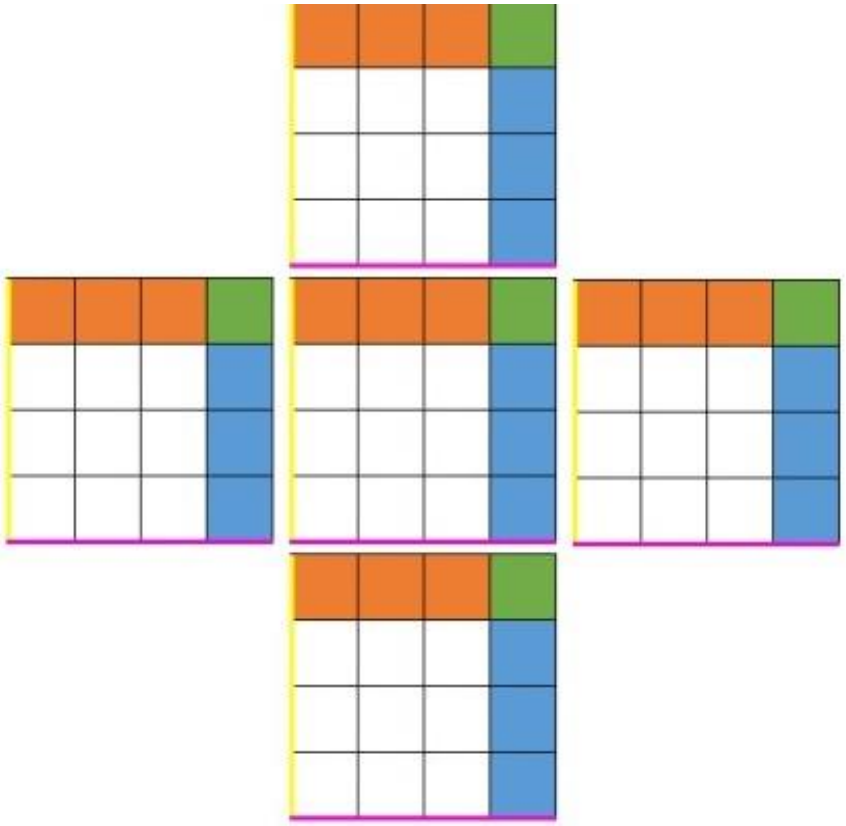


Image credits: A. Portelli

Numerical challenges

Prifysgol Abertawe
Swansea University

- Taking the infinite volume limit at constant lattice spacing generates a polynomial growth of the computational time
- Taking the chiral limit results in (nearly) ill-conditioned inversions of large sparse matrices
- Taking the continuum limit at fixed physical size causes an exponential growth of the required operations
- At high temperatures large lattice artefacts need to be tamed
- Large cancellations happen in the Monte Carlo at finite density
- Additional cost for varying gauge content and fermion representations

Approach

- Design, implementation and development of highly specialized algorithms
- Parallelism exploited at all possible levels
- Intense use of cutting/bleeding edge computational resources
- Development of suites of open-source community codes (Grid, HiRep, MILC, OpenQCD, QUDA…)
- Continuous dialogue with hardware and technology providers

Grid

https://github.com/paboyle/Grid

GRID

- Purpose: Lattice QFT (QCD, QCD+QED, BSM, ...)
- Parallelism:
  - CPU: SIMD, Multi-thread, Multi-processing (MPI).
  - GPU: SIMT (Cuda, HIP, Sycl), Multi-processing (MPI).
  - Expression template engine abstracts site wise operations (automatically parallel).
  - High level cshift and stencil interfaces.
- Multiplatform: vectorisation for many instruction sets (SSE,AVX,AVX2,...)
- Implements popular lattice QCD fermion actions (Wilson, DWF, Staggered,...)
- Variety of solver algorithms already implemented (CG, Multi-grid, Lanczos,...)
- Full HMC/RHMC interface included.
- Workflow management: Hadrons [https://github.com/aportelli/Hadrons]

HiRep
https://github.com/claudiopica/HiRep

- Purpose of the code: explore novel strong interactions
- Main physics motivations: fundamental mechanism of electroweak symmetry breaking in the standard model and dark matter
- Lack of clear experimental guidance suggests to use safe methods
- Different theories are implemented at compilation time, through a PERL preprocessor
- Uses arrays of structures, enrolling and inlining of mathematical operations
- The original code is high-level, developed in C, and uses MPI with latence masking
- More recent implementations using OpenMP and CUDA available

A (biased!) selection of current calculations

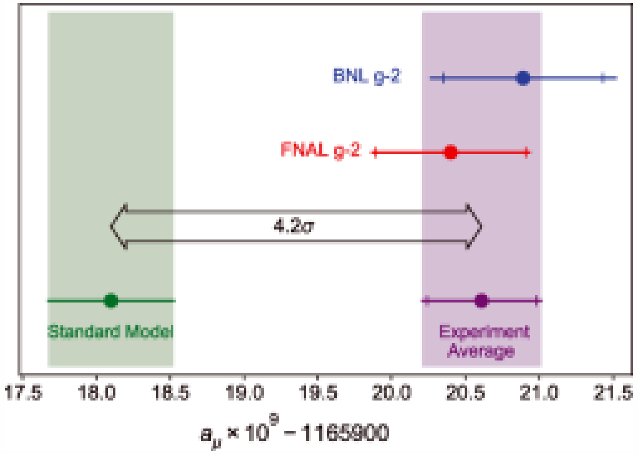
The muon g-2 (in 2021)

Experiment vs Standard Model prediction

Exp: $\boldsymbol{a_\mu = 0.00116592061(41)}$

SM: $\boldsymbol{a_\mu = 0.00116591810(43)}$

Muon g—2 Coll., Phys. Rev. Lett. 126, 141801

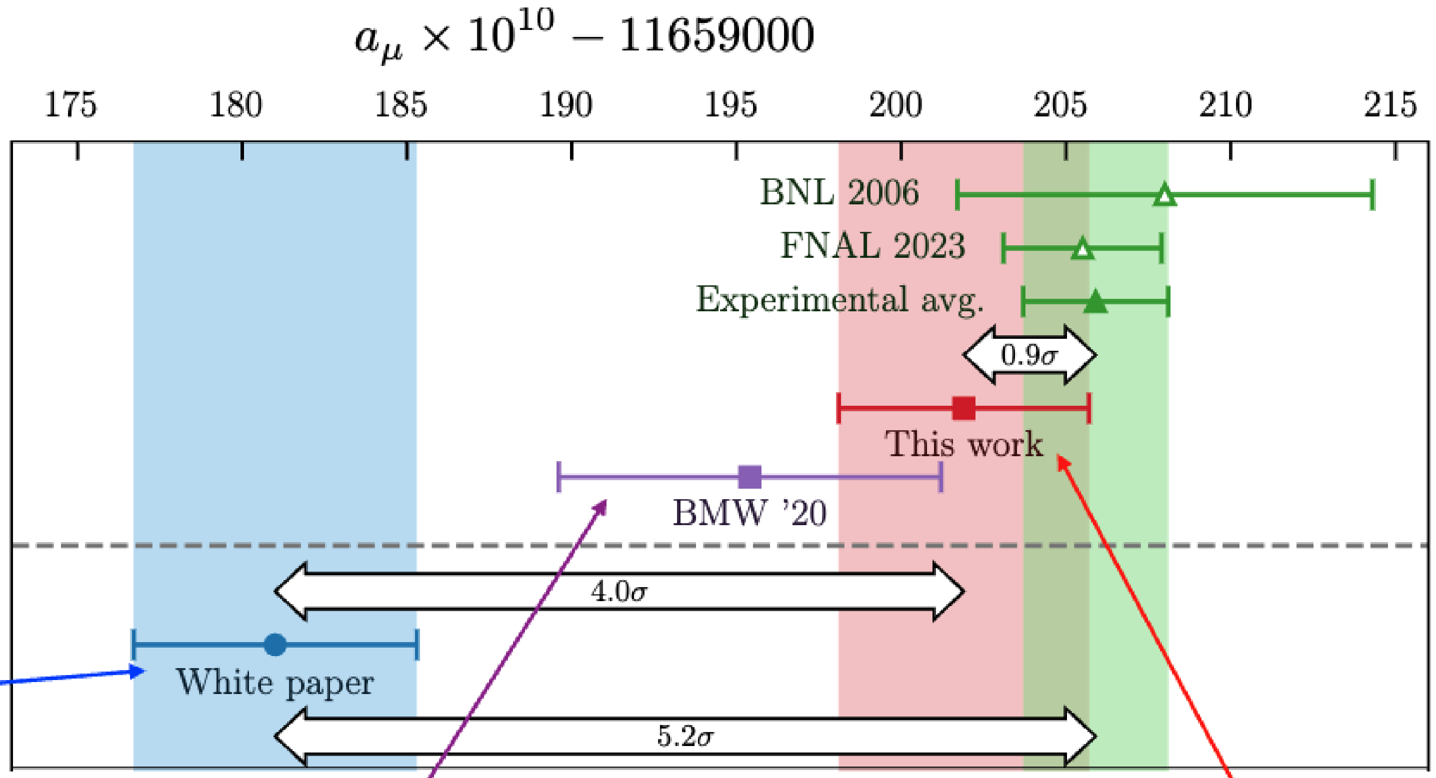


Credits: V. Guelpers

The muon g-2
2024 update

BMW/DMZ24, 2407.10913

adds 0.048fm ensemble, reduces finite L/T error. Uses data-driven for large-t tail. Blinded analysis.



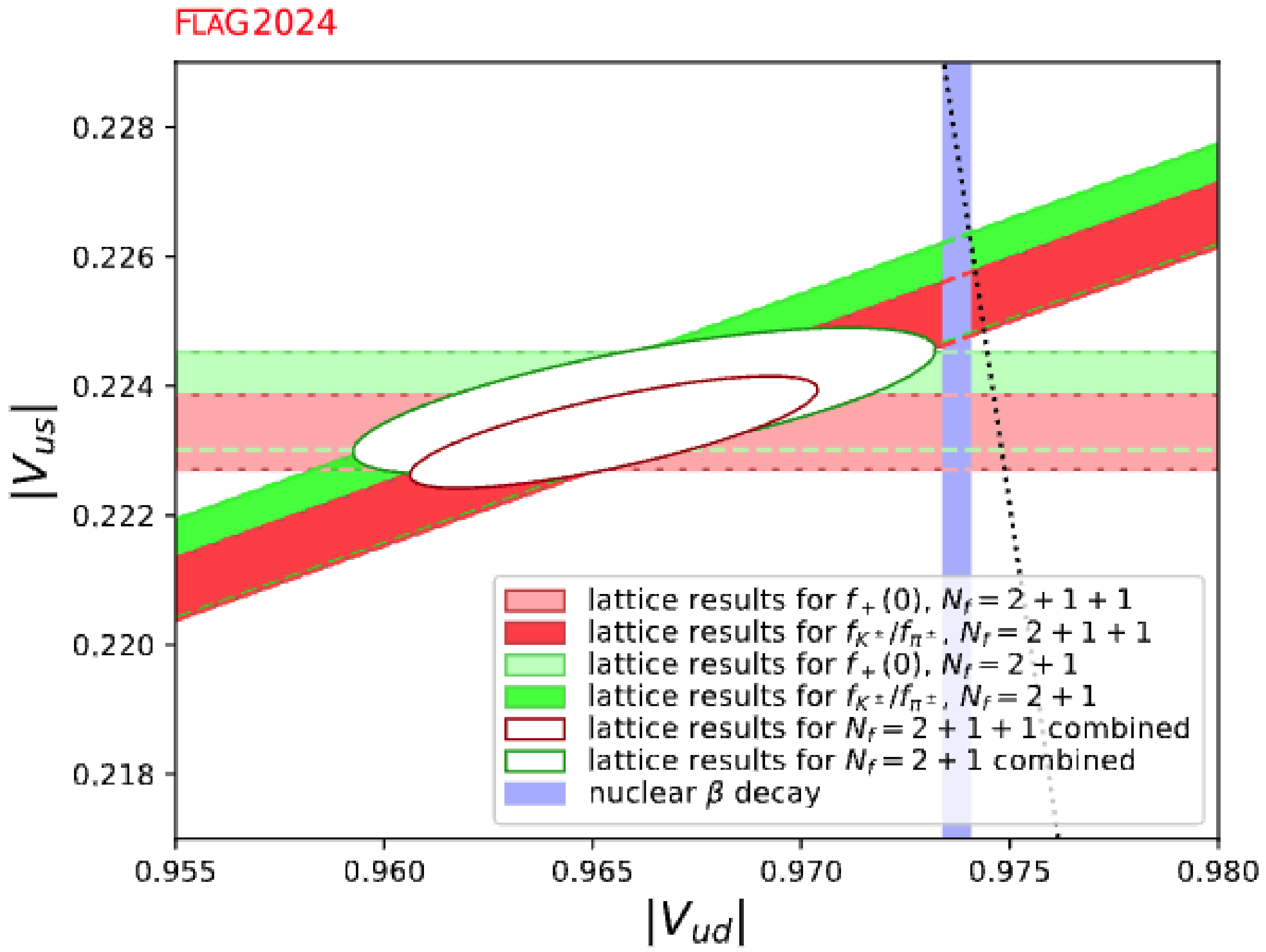
WP20 data-driven: 693.1(4.0)

BMW20: $10^{10} a_\mu^{\mathrm{LOHVP}} = 707.5(5.5)$

BMW/DMZ24: $10^{10} a_\mu^{\mathrm{LOHVP}} = 714.1(3.3)$

Credits: C. Davies

Radiative corrections to weak decays



Y. Aoki *et al.,* arXiv:2411.04268

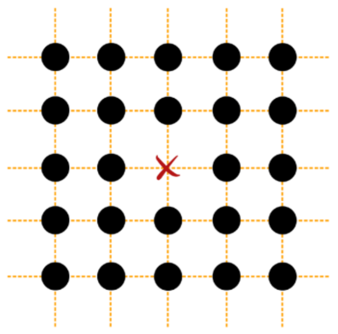
QCD + QED

Gauss law: only zero net charge is allowed in a finite volume with periodic boundary conditions

$$Q = \int d^3x \, j_0(t, x) = \int d^3x \, \nabla \cdot E(t, x) \stackrel{!}{=} 0$$

Possible solutions:

QED$_L$

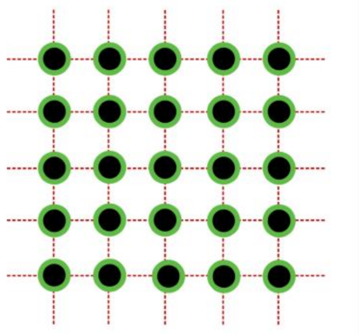


$k_3 = 2\pi \mathbb{Z}^3/L$

remove spatial zero-mode of the photon field

M.Hayakawa & S.Uno, PTP 120 (2008)

QED$_m$

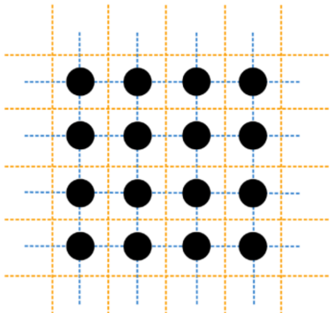


$k_4 = 2\pi\{\mathbb{Z}^3/L, \mathbb{Z}/T\}$

use massive photon m_γ

M.G.Endres et al., [1507.08916]

QED$_{C^*}$



$k_3 = 2\pi \mathbb{Z}^3/L$ $\quad k'_3 = (2\mathbb{Z}^3 + \bar{n})\pi/L$

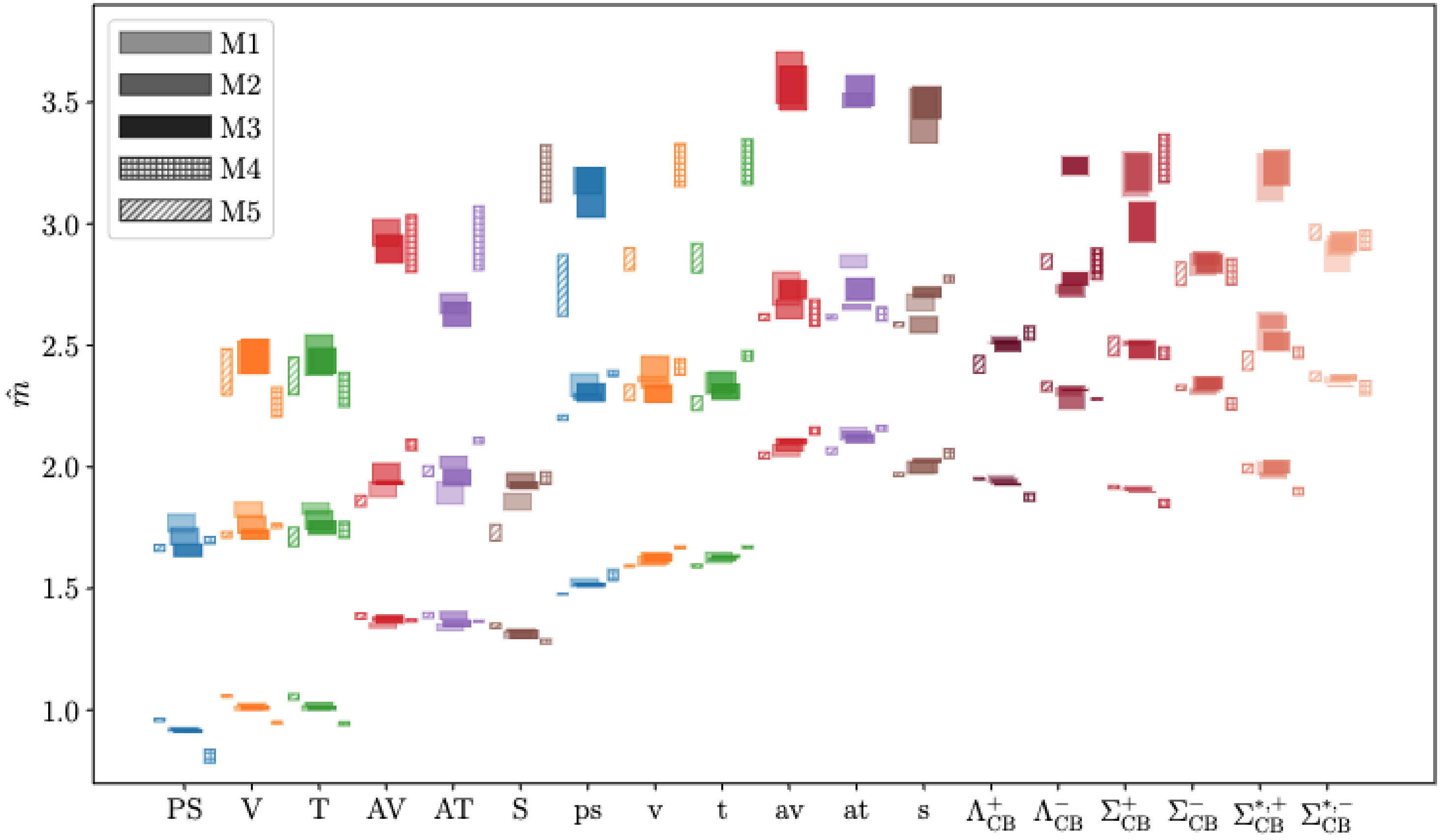
employ C* boundary conditions

A.S.Kronfeld & U.-J.Wiese, NPB 357 (1991)
B.Lucini et al., JHEP 1602 (2016)

Large power-law finite size effects arise

Credits: M. Di Carlo

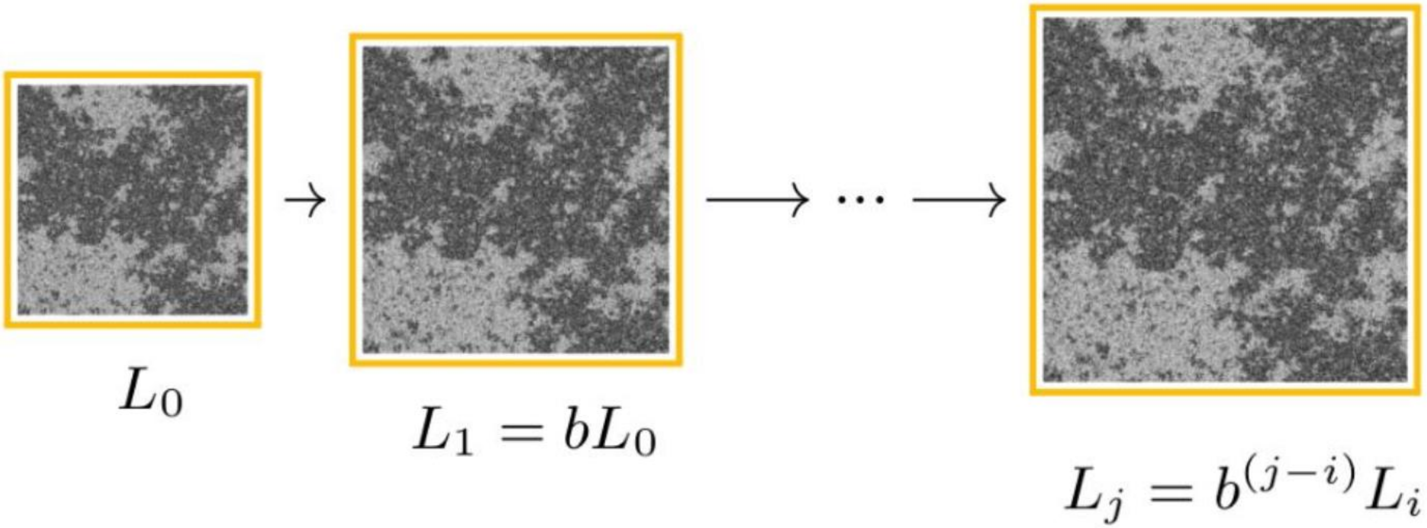
Strongly interacting theories beyond the Standard Model – Example Spectrum

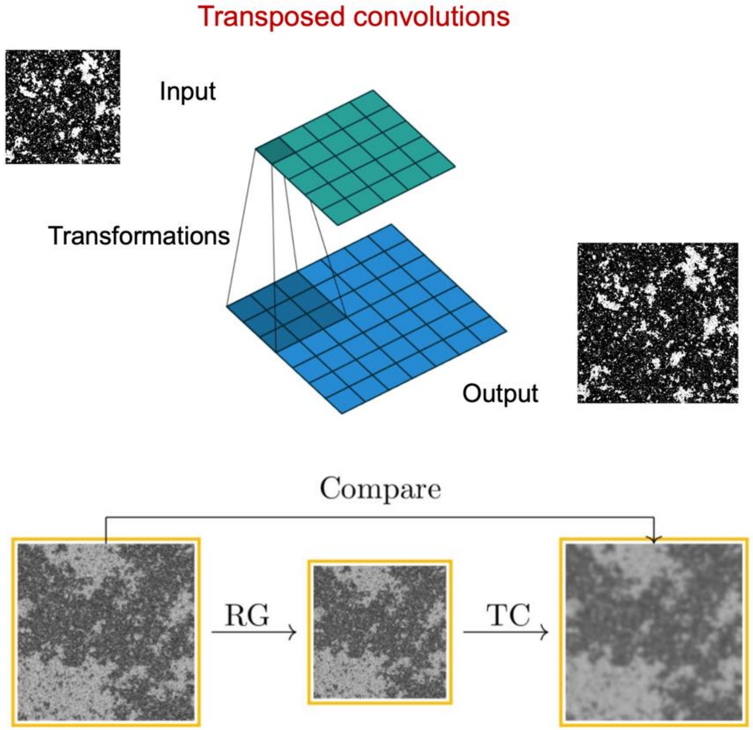


E. Bennett *et al.*, in preparation

Integration of machine learning methods

- ML for acceleration of algorithms
- ML for generation of configurations
- ML for improvement of noise/signal ratio
- ML for discovery of improved/novel observables





[D. Bachtis, G. Aarts, F. Di Renzo, and B. Lucini, Phys. Rev. Lett., 128:081603 (2022)]

Summary

- Lattice QCD (or better, Lattice Gauge Theory) is a mature computational branch of theoretical particle physics
- Computational demands of the science questions keep being a challenge also at the exascale
- On the bright side, LGT can drive exascale development, both in hardware and software, that are transferable across disciplines
- A set of community codes are being ported to future architectures
- Benchmarks that are derived from those codes can provide measures of performance that usefully inform other applications (and the vendors!)
- Integration of machine learning methods offer new opportunities

Acknowledgements

Prifysgol Abertawe
Swansea University

DIRAC HIGH PERFORMANCE COMPUTING FACILITY
Supporting the STFC Theory Community

ExCALIBUR 10

UKRI
Science and Technology Facilities Council