TWG8.2: Nuclear physics tools
Machine Learning, Artificial Intelligence, and Quantum Computing

Valerio Bertone

IRFU, CEA, Université Paris-Saclay

On behalf of TWG8.2: ML, AI, QC

April 16, 2024, NuPECC LRP 2024 Town Meeting, Bucharest

List of contributors

- Conveners:
  - V. Bertone (Paris-Saclay), J. N. Günther (Wuppertal)
- Liaisons:
  - H. Moutarde (Paris-Saclay), E. Nappi (Bari)
- **16 contributors** (in alphabetical order):
  - Joerg Aichelin (Nantes)
  - Mohammad Al-Turany (GSI)
  - Sara Collins (Regensburg)
  - Tommaso Dorigo (Padova)
  - Christian S. Fischer (Gießen)
  - Philipp Hauke (Trento)
  - Andreas Ipp (Vienna)
  - Marc Labiche (STFC)
  - Denis Lacroix (Orsay)
  - Johan Messchendorp (GSI)
  - Stefano Piano (Trieste)
  - Arnau Rios (Barcelona)
  - Alessandro Roggero (Trento)
  - Olga Soloveva (Frankfurt)
  - Vittorio Somà (Paris-Saclay)
  - Lingxiao Wang (Frankfurt)
- **15 pages** (limit almost saturated) and **2 figures**.
- **1 highlight box**: one removed due to overlap with WG1.

Disclaimer

- For the **first time,** a chapter *totally* devoted to computing tools was included in the NuPECC LRP report.
- This is particularly **timely** in view of the rise of new computing tools that have come to the forefront.
- However, this subject is **transversal** to all branches of nuclear and particle physics.
- This has required a **selection** of the subjects treated in this chapter
    - fit in 15 pages with a limited number of contributors.
- Yet, we tried to issue **general recommendations** that apply as much as possible to the whole field.

Introduction

- Tremendous progress in nuclear and (high-energy) particle physics has led to the need of **appropriate numerical tools**:
  - encompass the topics of the Joint ECFA-NuPECC-APPEC activities (**JENAA**).
- Algorithms based on Machine-Learning (**ML**) and Artificial-Intelligence (**AI**) and the progress Quantum Computing (**QC**) have often come to rescue.
- So far, the exploration of possible applications of ML, AI, and QC to nuclear and particle physics in Europe has mostly proceeded in an incoherent fashion with **local or at most national initiatives** devoted to it.
- Provide an as-comprehensive-as-possible **overview** of the current status of numerical techniques in nuclear and particle physics, with the purpose of **coordinating** this effort at a European level.
- The final purpose is to identify the main **guidelines** to deliver a set of general **recommendations**.

Structure of TWG8.2

- To this end, TWG8.2 is organised in **four subgroups**:

  1. Machine learning and artificial intelligence
  2. Quantum computing
  3. Tools and techniques
  4. Resources and infrastructure

- This organisation allowed us to focus on the specific aspects of these branches to eventually issue a set of **general recommendations** that have the purpose of **optimising the European resources**.

ML and AI

- ML is suitable for the rich data and sophisticated explorations, with unprecedented opportunities of understanding nuclear and particle physics in today's **AI-driven world**.
- ML applications have been employed/explored in:
  - nuclear experiments, nuclear astrophysics, processing large datasets, particle identification, event reconstruction, experiment design and control, analyse signals, processing data from noisy environments, determining properties of dense matter, hadron structure, astrophysical simulations, Lattice QCD (LQCD), etc.
  - Neural networks (NNs) are used to design accurate ansätze for the wave-function of strongly correlated many-body systems, study the real time dynamics of nuclear many-body systems, realise real-time event reconstruction or online physics analysis at, *e.g.*, STAR and LHCb.

ML and AI

- **Deep learning** (DL) has proven to be successful in representing **complex processes**:
  - **data-driven nature** well suited for nuclear physics, especially given the extensive data from experiments, observations, and simulations.
- **Generative models** are designed to **generate data samples** resembling the input data and found applications across various domains of nuclear and particle physics:
  - Improve the efficiency of **Monte-Carlo** algorithms in lattice QCD.
  - Incorporation of gauge symmetry by means of convolutional NNs.
  - Data augmentation, noise reduction, and event simulation, such as optimisation of particle-accelerator operations.
  - **Generative Pre-trained Transformer** (GPT) models are set to play pivotal roles in both experimental and theoretical advancements.

ML and AI

- ML offers benefits for data interpretability by means of **physics-informed** and **physics-driven** methods
  - **Automatic differentiation** (AD) is being employed to refine experimental designs, ensuring that the functional designs align with physical principles.
  - AD is also used for the determination of the hadronic structure.
  - Tackling **inverse problems** when one wants to extract spectral functions from LQCD.
  - Build dense nuclear matter equation of states (EoSs) from neutron star observations.
  - Obtain QCD matter properties from heavy-ion collision experiments.

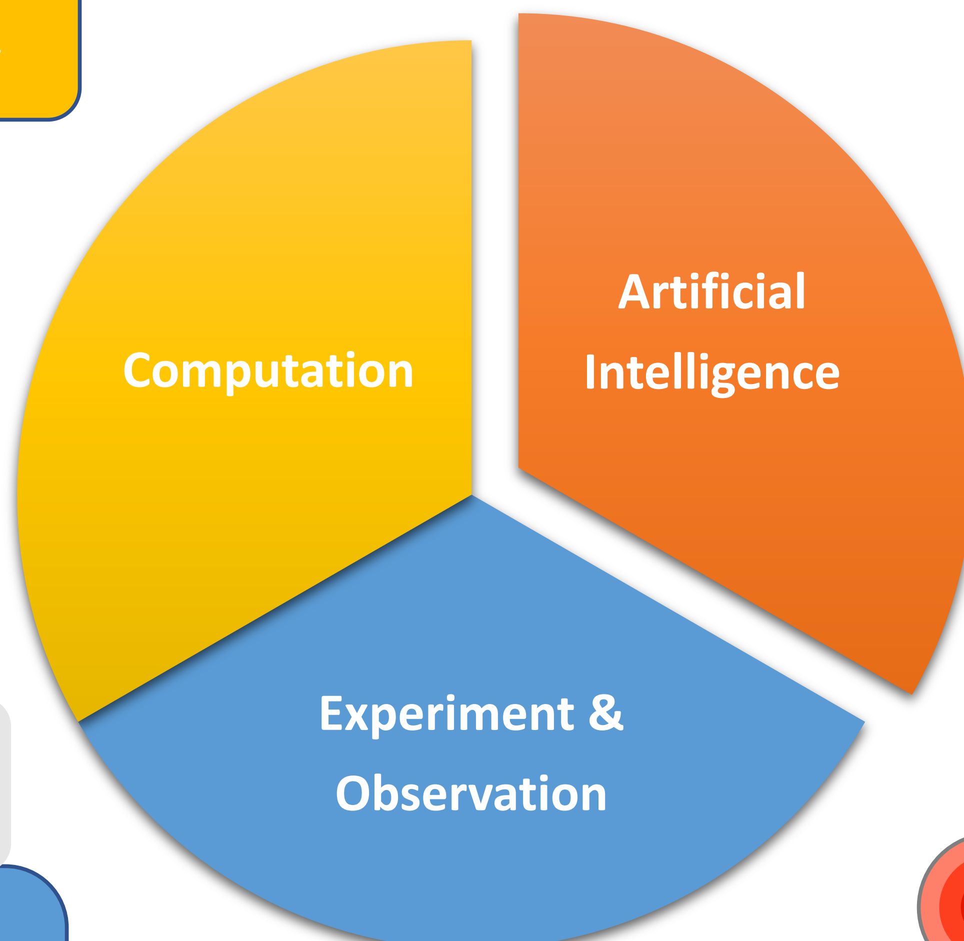
ML and AI

Lattice QCD

Property Extraction,
Computational Efficiency

Machine Learning

Data Analysis,
Neural Network Emulator,
Automatic Differentiation,
Generative Model



Nuclear Physics Experiments

Signal Analysis,
Event Selection,
Fast Simulation,
Experiment Design and Control

Nuclear Astrophysics

Denoising,
Signal Analysis,
EoS Reconstruction

ML and AI

- Recommendations:
  - **Transform ML prototypes into applications for production.** Formulate a strategic approach to advance from current short-term, proof-of-concept ML projects towards practical applications usable in production. […]
  - **Fostering data sharing in nuclear physics.** Promote data-sharing initiatives in nuclear physics by establishing a user-friendly hub for open databases, akin to existing platforms in related scientific fields. […]
  - **Strengthen computational resources.** Allocate funding for enhanced GPU clusters within established HPC centres across Europe. […]
  - **Training of scientific foundation models.** We recommend dedicated investments in training and fine-tuning models tailored for scientific purposes. […]

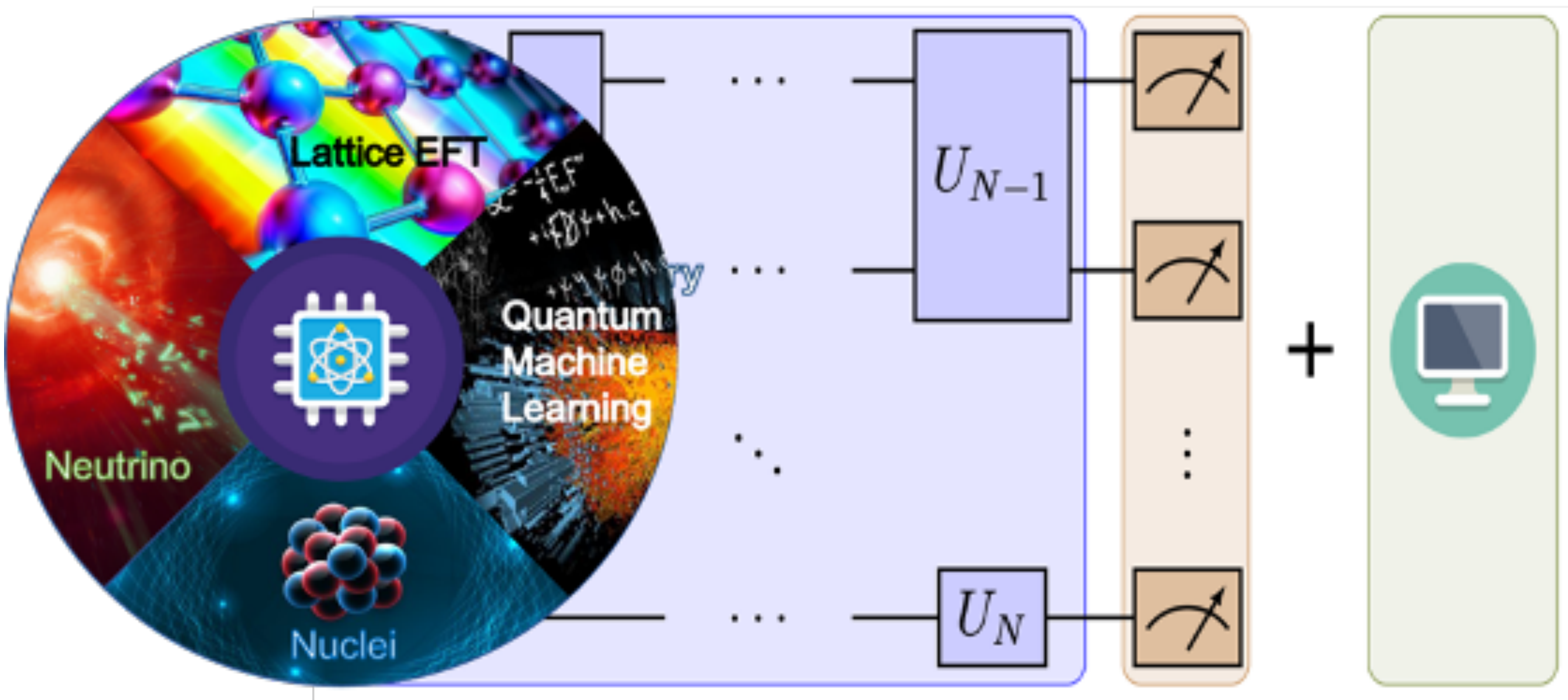
Quantum computing

- The field of quantum QC has seen a **remarkable growth** worldwide over the past few years.
- Interest motivated by the exciting **potential** of these technologies to simulate strongly correlated **many-body systems**.
- Data mining through the development of advanced **Quantum ML** (QML) techniques blends QC and ML.
    - possibility of defining quantum and classical algorithms to run on classic or quantum data.
    - QML is an excellent candidate to deal with the incoming computational tools outbreak caused by increasingly complex data.

Quantum computing

- Small-scale quantum computers are now widely available, often as computational devices accessible as cloud services.
- Complex quantum computers with **hundreds of qubits** are being commissioned or built by a wide variety of academic institutions and industrial (private) partners worldwide.
- We are currently in a **transition period** where the size and the fidelity of quantum resources are not sufficient for fault tolerance, and decoherence is still an issue.
  - The technological challenges associated with these issues are being addressed with substantial progress every year.
  - Despite these difficulties, quantum computers are considered a **disruptive technology** that opens new horizons.

Quantum computing



Schematic illustration of some of the fields where quantum computing is now being explored (left). Examples include physics of quarks, neutrinos and nuclei, as well as a pictorial view of quantum machine learning. All these problems are being now considered as pilot applications that could be treated on digital quantum computers using quantum circuits (right).

Quantum computing

- Effort to advance QC towards realistic applications.
- Recommendations:
  - **Establish a transnational European network on quantum activities.** Create a collaborative network at the European level, focused on QC […]. Bridge QC theory and machine learning. […]
  - **Organise workshops, schools and training programs.** Facilitate the transition towards quantum computing by partnering with academic institutions to provide comprehensive training opportunities in QC.
  - **Facilitate access to quantum platforms.** Ensure access to state-of-the-art platforms by bridging the gap between academic and private institutions. […]
  - **Develop Strategies for Quantum-Classical Interface.** Formulate a clear strategy for interfacing quantum and classical machines. […]

Tools and techniques

- An exhaustive treatment of the numerical tools employed in nuclear and high-energy particle physics is an enormous task impossible to fit in this report.
- Discuss some **prominent cases** that will eventually help us formulate a set of general recommendations concerning the development of new numerical technologies.
- To be as representative as possible, we distinguish between:
  - **theory** viewpoint: heavy-ion collisions and hadronic structure, computations on lattice, low-energy nuclear structure.
  - **experimental** viewpoint: Monte Carlo generators for the development of detector systems.

Heavy-ions and nucl. structure

- Experiments, such as those at RHIC and the LHC, have moved on from an **exploratory phase to a high-precision phase**.
- Theoretical approaches evolved from simple models to more complex frameworks aiming at a **global** description of data.
- These developments have led to advancements of great **complexity**:
    - developing them further requires **software specialists** and new computational approaches.
- Underlying theories, like lattice or perturbative QCD, do not provide predictions in the whole kinematic range:
    - **forceful extrapolations** are often necessary ⇒ an assessment of the different approaches necessary.
- In the US, different approaches have been **centralised** and manpower made available to make the programs comparable and benchmark the results (*e.g.* JETSCAPE, MUSES, HEFTY).

Heavy-ions and nucl. structure

- she study of the hadronic structure in the context of high-energy collisions has also witnessed an impressive development catalysed by **experimental**, **theoretical**, and **technological** advances.
- Recent experimental results require an unprecedentedly detailed understanding of the internal structure of hadrons.
- More accurate theoretical and numerical calculations have been performed to streamline the interpretation of experimental results:
  - ML techniques have also come to rescue by enhancing our ability to manage complex data structures.
- In Europe, STRONG2020 has invested significant resources on open-source numerical frameworks for precision phenomenology at high-energy colliders.
  - In the US, the Centre for Nuclear Femtography has similar purposes.
- The Electron-Ion Collider to be built in the US is a further stepping stone towards a detailed understanding of the hadronic and nuclear structure.

Lattice QCD (LQCD)

- Since the LRP 2017, there has been tremendous progress in LQCD.
- New theoretical approaches, simulation algorithms, and solvers as well as the possibility to access large **supercomputing resources** enabled by PRACE and EuroHPC Joint Undertaking (JU).
- Huge computational resources required largely organised into European and worldwide collaborations, each with **their own codes**:
  - *e.g.* Chroma, CPS, MILC, USQCD, GRID, OpenQCD, tmLQCD (TWG1).
- LQCD necessitates high-performance computing capabilities and continuous access to large computational resources to maintain their **competitiveness** and achieve their scientific objectives
  - resources on GPUs and CPUs (O(10^9) core hours on CPUs).
  - Double precision floating point performance must be matched with adequate memory and internode communication bandwidth.
  - Develop highly optimised algorithms and codes.

Lattice QCD (LQCD)

- LQCD groups are early adopters of cutting-edge technologies and are expected to become users of **exascale** infrastructure.
- The need for **coordinated data handling** has led to the formation of the International Lattice Data Grid (ILDG), with several regional sub-grids.
- Emerging techniques based of LQCD have recently broken through in the field of hadronic structure, which allow us to extract information on **partonic distributions defined on the light cone**:
    - this challenge has now been taken on by many groups in Europe and all around the world leading to significant advances.

Low-energy nuclear structure

- Nuclear structure theory aims at describing the properties of nuclei and nuclear matter by solving the many-nucleon **Schrödinger equations**.
- *Ab initio* calculations have lately witnessed considerable progress extending beyond light nuclei to more complex systems.
  - heavy and/or deformed isotopes, accurate nuclear interactions and many-body techniques, bridging nuclear structure and reactions.
- The modelling of nuclear forces requires sophisticated fitting procedures in **multi-dimensional parameter spaces**.
- Many-body techniques adopt large-scale **diagonalisations** or iterative solutions of tensor networks, where the dimensions increase with the required accuracy.
  - Applications to nuclear reactions further increase size and complexity of the objects to be manipulated.
- This workflow entails CPU- and memory-greedy algorithms that require **state-of-the-art high-performance clusters**.
- Existing codes will be ported to **exascale** machines. The possibilities offered by ML and QC will be increasingly explored by the community.

Tools and techniques (Theory)

- Recommendations for heavy-ion collisions and hadronic structure:
  - **Creation of a new set of open-source numerical tools** aimed at improving the physical interpretation of present and future high-energy data. […]
  - **Provide a solid basis for systematic comparisons** of the experimental results to theoretical predictions (*e.g.* RIVET). […]
- Recommendations for Lattice QCD (LQCD):
  - **International collaborations and freely available software** packages for LQCD should be maintained and extended.
  - **Access to substantial computational resources** and funding for the development of optimised algorithms and codes.
  - **Education and training for early-career researchers in LQCD**. […]
  - **Long-term career perspectives for software developers**. […]
  - **Long-term storage solutions for gauge ensembles.** […]

Tools and techniques (Theory)

- Recommendations for low-energy structure:
  - **Promote cross-fertilisation from/to other fields** utilising many-body methods and applied mathematics.
  - **Enhance and systematise collaborations with computer scientists** to exploit new computer architectures.
  - **Support virtual-access facilities** such as Theo4Exp of EURO-LABS.
  - **Encourage the sharing and publication of numerical codes**.

Monte Carlo generators

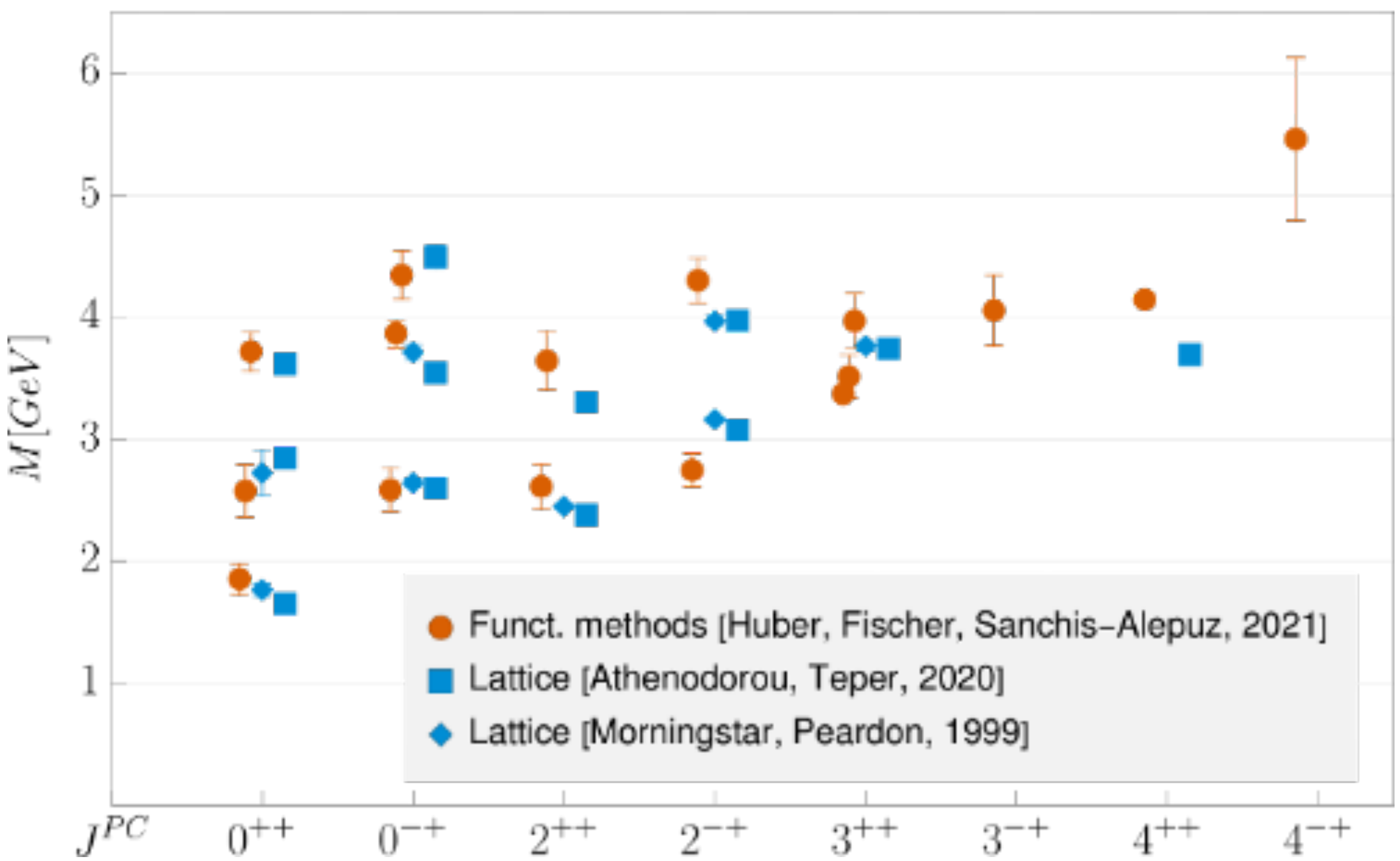
- Monte-Carlo (MC) event generators are extremely valuable numerical tools in experimental nuclear and high-energy physics.
- MC simulations are used for the development and optimisation of detector systems.
- MC generators rely on the generation of pseudo-random numbers which is typically a **CPU-bound task**.
- Since MC simulations are **parallelisable**, multi-core systems can be exploited.
- Large data storage and backup systems would allow users to log in remotely and execute data analysis codes directly on the server with obvious benefits:
  - such systems have already been used with great benefits in high-energy physics and could be equally beneficial in nuclear physics.
- A **centralised data storage** would secure large data sets collected from laboratories all around the world and allow to make them freely available.
  - Such a system would need a very large array of redundant hard disks for storage but would only have basic requirements for CPU and computing power.

Tools and techniques (Exp.)

- Recommendations:
  - **Deployment of simulation software** such as GEANT4, MNCP, and FLUKA on fast computer systems to enable generation and analysis of large numbers of events and the simulation of complex experiment setups.
  - **Adaptation of physics models** and **development dedicated interfaces** to extend the range of physics applications that event generators can currently deal with […] along with a centralised computing resource with high performance. […]

Highlight: Functional methods

- The functional approach to QCD via **Dyson-Schwinger** equations (DSEs), **functional renormalisation group** equations (FRGs) and **Bethe-Salpeter** equations (BSEs) works at the level of non-perturbative correlation functions that encode all information on the physics content of QCD.



Glueball spectrum of pure Yang-Mills theory determined on the lattice and with **functional methods**.

- Applications to three (baryon, hybrid), four (tetraquarks) and multi-body problems including scattering processes, many current projects in the functional community have reached a level where the bottleneck is **computational power**.

Resources and infrastructure

- Modern CPUs commonly have **homogeneous** multiple cores, enabling parallel processing of tasks.
  - Need for software optimisation to fully exploit the potential of multiple cores. More specialised hardware components (coprocessor) are needed to perform specific tasks.
- The need for efficient and flexible architectures that can handle diverse workloads has become increasingly important.
- **Chiplets** (tiles) break down the CPU into smaller, specialised components.
- Developing programs to exploit the **parallelism** and **heterogeneity** poses several challenges:
  - exploit softwares such as Alpaka, Kokkos, oneAPI and SYCL.
- HPC exascale and QC offer an unprecedented processing capacity. However, exploiting HPC centres comes with several challenges:
  - a cohesive data processing system, the SPECTRUM project, aims to integrate different European computing resources.

Resources and infrastructure

- The high data rates expected from the next generation of particle-physics experiments (*e.g.* at FAIR/GSI and CERN) call for dedicated attention to the design of the computing infrastructure needed **online *and* offline**.

- The traditional DAQ/Trigger designs are not able to handle the amount of data coming from the detection systems:

    - more data processing is needed online which will introduce more software-based components than ever before.

- Handling such systems will require:

    - the design of efficient and scalable algorithms,

    - the development of key software building blocks for ultrafast data processing on large-scale heterogeneous computing infrastructures,

    - synchronisation of multiple data streams, transport services, container orchestration, and efficient binding to storage and network.

Resources and infrastructure

- Recommendations :
  - **Significant investment in software frameworks** needed to support parallelism on multi- or many-core CPU and the execution of algorithms on heterogeneous platforms.
  - Establishment of a program aimed at **facilitating access for nuclear physics researchers to large HPC centres**.
  - **Foster interoperability among diverse architectures**. Create multi-architectural software and measures to ensure seamless operation.
  - **Promote the implementation of a federated computing model**. Ensure accessibility of resources facilitating interoperability among IT activities within the nuclear-physics community and related fields. […]
  - **Develop interfaces located between the system software layers** (i.e. storage, RDMA networks, container runtime, etc.) and the core software developed by the domain scientist encouraging collaboration between IT experts. […]

Thank you