Precision calculations for high-energy collider processes

Contact persons: Charalampos Anastasiou, Stefan Dittmaier, Thomas Gehrmann, Nigel Glover, Massimiliano Grazzini, Michelangelo Mangano, Stefano Pozzorini, Gavin Salam, Giulia Zanderighi

December 2018

Abstract

For the field of particle physics to benefit from the immense body of data from present and possible future colliders, it is essential to be able to accurately relate the experimental observations to the underlying Lagrangian. This is the case whether searching for physics beyond the Standard Model, exploring the many facets of the Higgs sector, or pinning down fundamental parameters such as coupling constants and particles masses, which have implications for foundational questions such as the stability of the universe. This document discusses the current status and future objectives of precision theory for collider processes, and outlines the needs required to accomplish them.

1 Introduction

Theoretical predictions based on the Standard Model (SM) are one of the most fundamental ingredients for the interpretation of collider data. The vast majority of experimental analyses make use of theoretical predictions in the form perturbative calculations at parton level or in combination with parton showers, embedded into the well-established factorisation picture for hadronic collisions.

Theoretical predictions are relevant for testing the SM at the level of the measured fiducial cross sections and distributions, and for the determination of SM parameters and parton distribution functions (PDFs). Moreover they are used as ingredients of the experimental measurements, both for the description of acceptance efficiencies and for the modelling of backgrounds in SM measurements and beyond the SM (BSM) analyses.

As a result of the continuously growing precision of LHC data, an increasing number of experimental analyses is limited by theoretical uncertainties. The full deployment of state-of-the art theory precision and its further improvement are thus going to be crucial for the full exploitation of the physics potential of the LHC, the HL-LHC and future colliders.

2 Progress since last European Strategy exercise

Since the start of the last European Strategy exercise, there have been several major axes of progress in the field of precision calculations for high-energy colliders.

1. **Automated next-to-leading-order (NLO) tools:** These were made possible by the emergence of new concepts and techniques during the last decade, and have opened the

door to NLO precision for a vast range of collider processes, including complex multi-particle processes that are copiously observed at the LHC. Automated NLO methods and tools have recently been extended from quantum chromodynamics (QCD) to electroweak (EW), Yukawa and scalar interactions in the SM. Full NLO SM precision is going to be especially important in the context of high-precision tests as well as for searches at the TeV scale, where higher-order EW effects are strongly enhanced.

Automated NLO tools start to be used routinely to compute the one-loop parts of $2 \to 2$ next-to-next-to-leading order (NNLO) calculations, and the technical one-loop improvements required for future higher multiplicity NNLO calculations are nontrivial but within reach.

2. **Monte Carlo generators:** The full chain of operations that is needed to generate NLO hadron-level predictions, including also the matching to parton showers, is already supported by various automated Monte Carlo (MC) frameworks. Thanks to their flexibility, such tools make it possible to deploy NLO precision in the framework of realistic experimental analyses, and to implement new general MC techniques in a way that renders them rapidly applicable on a large scale.

   The recent advent of NLO merging techniques has opened the route to a precise description of multi-jet final states. However, the full exploitation of NLO precision for high-multiplicity processes at the LHC will require very significant efficiency improvements in various building blocks of MC generators.

3. **NNLO calculations:** Until a few years ago, NNLO accuracy was available only for $2 \to 1$ processes. The last few years have seen a major step forward, with the development of new computational techniques and mathematical methods that allowed us to reach NNLO precision in the strong interaction for essentially all important $2 \to 2$ LHC processes. This has enabled theoretical predictions at a few-percent accuracy for many such processes.

   It should be stressed that these NNLO predictions are the result of a global effort of our community, and often come by combining results of different project stages, sometimes spanning a period of several years. NNLO QCD calculations require the combination of tree-level, one-loop and two-loop contributions, and a treatment of associated infrared singularities. Major innovations in the methods to carry out that combination have been crucial to the progress made, as has the development of flexible partonic event generators, so as to be able to take into account realistic experimental cuts.

   The two-loop amplitudes themselves have in some cases been available for almost twenty years, e.g. for massless $2 \to 2$ processes, relevant for inclusive jet production, while for massive final states, relevant for example for vector-boson pair production, they have been computed only recently.

   The dramatic impact of NNLO calculations in ensuring agreement between theory and a wide range of data is illustrated in Fig. 1. That progress has also fed into full NLO calculations for loop-induced processes such as double-Higgs and high-p_T single-Higgs production.

4. **N3LO Higgs calculations:** At the time of the last ESPP many would have been hard-pressed to believe that the N3LO calculation of Higgs production through gluon fusion would be complete by now. Without this important progress, theoretical uncertainties

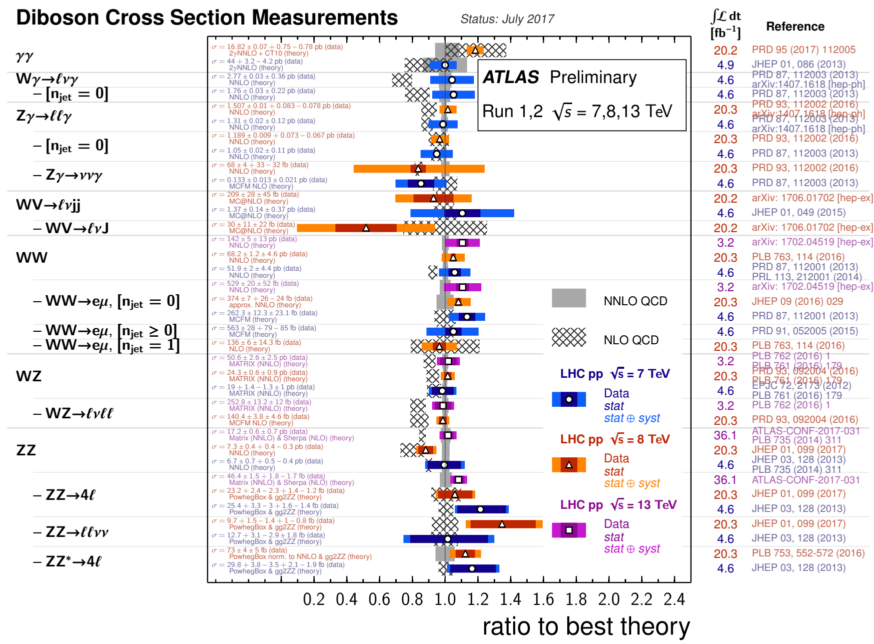

Figure 1: Current status of data/theory comparison for diboson cross sections.

would now start to dominate in a number of comparisons between Higgs production measurements and SM expectations and in the corresponding extraction of Higgs boson properties, as illustrated in Figure 2.

5. **Resummation:** Our ability to compute all-order logarithmic terms in physical cross sections, either using standard coherent branching methods or using soft-collinear effective theory (SCET), has advanced significantly. As a result we have next-to-next-to-leading logarithmic (NNLL) accurate predictions for a wide range of multi-scale processes. One example is in Higgs production, where the combination of NNLL and NNLO results has made it possible to have confidence in the ability to extract quantitatively reliable information from Higgs studies even in the presence of jet vetoes that help remove backgrounds. Another example is N3LL predictions for the Higgs and DY transverse momentum distributions, the former being of interest because it is sensitive to the coupling of the Higgs to relatively light quarks, the latter being of particular relevance for example for precise hadron-collider W-mass extractions.

   The theoretical developments in resummation have equally contributed to the formulation of efficient methods to handle and cancel infrared singularities at NNLO and beyond.

6. **Parton distribution functions:** The recent progress on NNLO corrections for $2 \to 2$ processes has had an immediate effect on the determination of parton distributions from collider data. They enabled many more observables to be consistently included in fits at this order, thereby sharpening our understanding of the parton structure of the proton. However, several open issues remain. Specifically a full characterisation of the statistical meaning of PDF uncertainties is still in the making, and the inclusion of theoretical uncer-

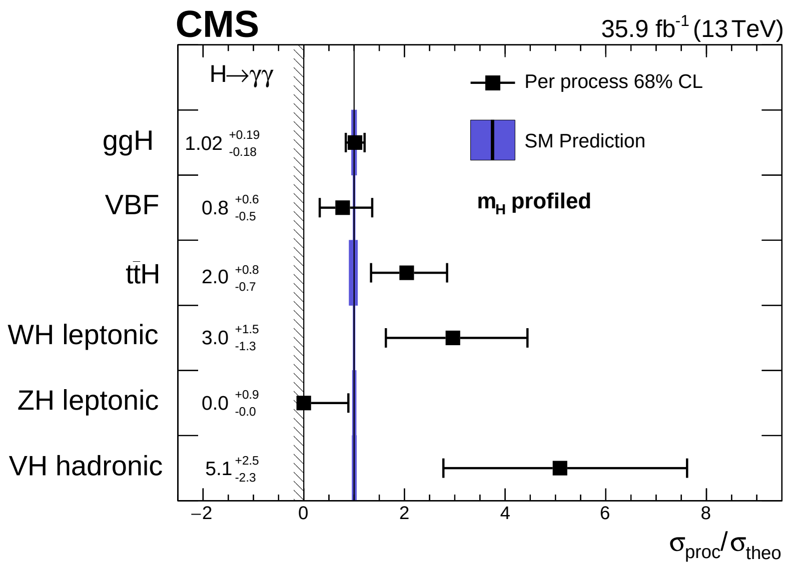

Figure 2: Current status of data/theory comparison for Higgs production cross sections.

tainties is missing as yet. The validity of certain legacy datasets is controversial, thereby highlighting the need of complementary information from more collider processes.

3 Roadmap for precision theory in the context of HL-LHC

The LHC experiments have reached a very high level of sophistication in the reconstruction of collision events, thereby enabling to make precise measurements despite the complex environment including substantial pileup. As an example, measurements of Drell-Yan production have reached half-percent precision (aside from an overall luminosity uncertainty) for a range of kinematic distributions. On this basis, and in the light of detailed projections from the experiments, one can expect that substantial further progress will be needed in theoretical calculations if these are not to become the limiting factor in interpreting a wide range of HL-LHC data. Again there are several broad directions:

1. **Core processes at high accuracy:** The experimental precision for many core $2 \to 1$ and $2 \to 2$ processes is likely to approach 1% precision, over a substantial range of phase space. Even where today's measurements are already limited by systematics, for example $t\bar{t}$ cross section measurements, it is not unreasonable to expect (a) that experimental systematics will come under better control and (b) that the field will devise analysis techniques that eliminate certain classes of experimental systematic errors. Many current NNLO predictions do not normally reach 1% precision and therefore there is a strong case for seeking to achieve N3LO accuracy for a range of $2 \to 1$ and $2 \to 2$ processes, and also in PDF evolution.

   Theory-data comparisons of these simple processes will feed back into our knowledge of PDFs, with a knock-on effect across a whole range of LHC physics. They also have the potential to increase sensitivity to small deviations from BSM effects that cannot otherwise be directly observed. For example, in the case of BSM operators that mix with SM ones, the new physics mass range Λ that can be probed scales as $1/\sqrt{\epsilon}$ for relative precision ϵ on the related physics observables.

2. **Complex processes at few-percent accuracy:** There are a number of crucially important signal and background processes that involve a $2 \to 3$ structure, which is beyond today's state-of-the-art for NNLO calculations. These include processes involving vector bosons, the Higgs boson, jets and heavy quarks (e.g. 3-jet production, $t\bar{t}H$, $t\bar{t}V$, $H+2$ jets, ...) which would benefit from the extension to NNLO. As an example $t\bar{t}H$ production, whose cross section is measured with roughly 15% statistical precision, might be expected to have $\sim 2\%$ statistical precision at the end of the HL-LHC. Without NNLO QCD and NLO EW calculations for the signal and the backgrounds this experimental precision can not be matched on the theory side, thereby preventing full exploitation of the results for physics studies.

   Work to break the $2 \to 3$ barrier for two-loop amplitudes is ongoing, and, at the same time, an effort is ongoing to improve available methods to handle and cancel infrared singularities. Given the complexity of these calculations, cross validation through the use of independent results and tools will be mandatory.

3. **Accuracy at high p_T:** Current measurements have only explored a limited portion of the available phase space. NNLO for fully differential cross sections paves the way for more detailed data/theory comparisons in remote phase-space regions where new physics effects could be concealed. Two important examples are high-p_T Higgs production and dark-matter searches. In the case of the former, recent studies of the ATLAS and CMS collaborations anticipate a $\mathcal{O}(10\%)$ precision in the Higgs rate for $p_T \geq 350$ GeV at the end of the HL phase of the LHC.

   Recent breakthroughs in the numerical evaluation of $2 \to 2$ amplitudes mediated by massive quarks, combined with NNLO calculations in the heavy-top limit provide a comparable precision in the SM prediction, and will therefore allow us to disentangle possible new physics effects. As concerns dark-matter searches, precision control of SM processes that generate missing E_T and corresponding visible normalisation channels will be important to exploit the full experimental reach. In this context, the understanding of logarithmically enhanced EW effects at high p_T will be crucial.

4. **Technical requirements for NLO multi-particle precision:** The full deployment of NLO precision through automated MC frameworks and its consistent exploitation in the wide landscape of HL-LHC analyses raises major technical challenges. Establishing the predictivity of MC tools at precision levels of order 10%, as well as their correct usage within the experiments, will require quantitatively and qualitatively unprecedented validation work. Given the high complexity of NLO MC frameworks, their continuous evolution, and the huge number of processes and kinematic regions that need to be scrutinized, this is a challenge that can only be addressed by a community of well-trained experts. In this context, the availability of various redundant state-of-the-art NLO tools will be crucial.

   Already now, the generation of NLO accurate event samples for $2 \to 4$ processes is limited by dramatic efficiency bottlenecks related to the convergence of phase space integration and various other technical aspects. The MC requirements for the HL–LHC era call for order-of-magnitude efficiency improvements that can only be addressed through significant algorithmic optimizations and entirely new techniques.

5. **Multi-variate analyses and background uncertainties:** The possibility of estimating theory uncertainties in a realistic way is among the main advantages of (N)NLO accuracy. This will be an increasingly important input for a large variety of experimental analyses that start being dominated by theoretical systematics. However, estimating theory uncertainties in the presence of nontrivial experimental cuts or in the context of sophisticated multi-variate analyses is a challenging and poorly studied problem.

   A typical example is provided by $t\bar{t}H$ searches in the $H \to b\bar{b}$ channel, which are presently limited by theory uncertainties in the $t\bar{t}b\bar{b}$ QCD background. In this kind of analysis, MC predictions for the large QCD background are constrained by data through a profile likelihood fit of several kinematic distributions in different event categories. In this context, theoretical predictions for the correlations across different categories and kinematic regions play a key role. All related uncertainties, e.g. at the level of NLO matrix elements, parton showers and NLO matching, need to be identified and modelled in a realistic way. This task is further complicated by the presence of multiple scales, both in the kinematic observables and in the process itself.

   This type of problem is characteristic for a broad range of LHC analyses, and its systematic solution will require a close collaboration between experts in the multiple theoretical and experimental facets of the problem. In particular, a much closer collaboration and communication between theory and experiments will be mandatory.

6. **Non-perturbative effects:** While the perturbative calculations follow a systematic first-principles approach, the understanding of non-perturbative effects is still far more rudimentary. This lack of predictivity limits the precision that can be attained for final states containing hadronic objects, thereby potentially influencing important parameter extractions such as the top quark mass.

7. **Resummation:** For key observables depending on disparate scales, advances in the all-order resummation of large logarithmic corrections will be crucial. These require to increase the logarithmic order in the standard coherent branching and SCET frameworks, as well as extensions to multiply-differential resummation, sub-eikonal effects, high-energy resummation, as well as the understanding of sub-leading and super-leading structures.

8. **Accurate predictions for BSM effects:** The great success of the SM in describing all phenomena observed at the LHC up to now seems to tell us that the key to a potential discovery of new physics is precision, in particular in the light of the planned HL-LHC. The ability to straightforwardly make predictions with reasonable accuracy for BSM effects is important both in terms of timely interpretation of any emerging BSM signals and for limits on BSM effects to be reliably compared across different types of measurements.

   In a model-independent (bottom-up) approach one generically describes deviations from the SM within the so-called SM effective field theory (SMEFT) that is based on the SM particle content, supplemented by higher-dimensional operators. In SMEFT it will be important to provide a setup within which high-precision data/theory comparisons can be performed for global fits to all the available data, with a consistent inclusion of QCD and EW radiative corrections.

A model-dependent (top-down) approach confronts data with predictions from specific SM extensions, which may be based on comprehensive models (e.g. supersymmetry) or rather minimalistic variants (e.g. simple extensions of the SM Higgs sector) carrying generic features of more comprehensive theories. Evaluating such models at least at the NLO level is necessary to to establish constrains on their parameters to the 10% level or better.

To enable this leap in precision and multiplicity for a large number of processes, calculations will require a high degree of automation, comparable to what is presently feasible for NLO QCD and NLO EW calculations. Current technical methods for calculating multi-loop corrections to scattering amplitudes and for handling multiple real-radiation corrections in scattering processes are usually restricted to low-multiplicity processes. More automated calculations will first require the development of better-suited technical approaches, for which a much deeper understanding of the high-order structure of SM perturbation theory will likely be a crucial prerequisite.

4 Precision theory beyond HL-LHC

4.1 Lepton colliders

The FCC-ee project as well as other electron–positron colliders at comparable or higher energy probe particle physics dynamics at the weak scale. Their physics exploitation will demand advances in precision theory in various directions. Main areas include:

1. **Electroweak precision observables:** Where overlapping in energy with LEP, these colliders will provide a data sample that is larger by three orders of magnitude, translating into at least one order of magnitude improvement in precision in physics observables. To arrive at a comparable level of precision in theory will require most of the LEP-era calculations at the Z-pole and the WW and ZZ threshold to be revisited, often calling for one or more orders in perturbation theory both in QCD and the EW theory.

2. **Higgs-boson and top-quark physics:** Acting as Higgs boson factories, novel electron-positron colliders will produce a unique sample of Higgs boson decays, enabling a detailed study of the Higgs boson properties, including invisible Higgs boson decays and therefore providing the total Higgs boson width, as well as searches for exotic decay modes. Depending on the collider energy, top quark pair production or multiple-Higgs-boson production may equally be accessible. Some theoretical work has already been accomplished for these processes in the context of earlier linear collider studies. Progress will be needed for precision predictions including particle decays, for realistic descriptions of backgrounds and for the interplay of higher-order QCD and EW effects on these observables.

3. **Photon radiation and two-photon physics:** With larger collider energy, initial-state photon radiation will become increasingly important. It induces significant corrections to the dynamics of electron-positron collisions, and provides a source for photon-photon interactions. With the interplay of both types of processes, final-state event dynamics will start to resemble the case of hadron colliders, thereby calling for a new framework for its theory description.

4.2 Hadron colliders

The HE-LHC and FCC-hh, as well as similar projects outside Europe, would take particle physics to energy scales well beyond the Fermi scale which is currently probed at the LHC. They offer opportunities for the direct discovery of new particles in an unprecedented mass range, and will equally enable a unique range of precision measurements:

1. **Higgs mechanism:** The dynamics of EW symmetry breaking can be probed through precision studies of the Higgs boson, then covering all production modes and most decay channels, as well as measuring the Higgs self-interaction through Higgs-boson pair production. Complementary to direct Higgs production, the mechanism of EW symmetry breaking is probed via off-shell Higgs bosons which play a crucial role in the unitarization of the SM at high energies. Electroweak vector-boson pair production and vector-boson scattering scrutinise the high-energy dynamics of the EW interaction in a novel energy regime, approaching symmetry restoration.

2. **Established benchmark processes:** Final states such as the production of Z bosons can be measured to sub-per-cent accuracy, thereby enabling their usage as luminosity calibration, provided their theory description attains a comparable level of precision. Jet observables will test the running of the strong coupling constant almost up to scales comparable to the collider energy, potentially yielding indirect signatures of new coloured states that are hard to discover in direct searches.

3. **Proton structure:** The parton structure of the proton will be probed over a much-enlarged range in resolution and momentum fraction, potentially uncovering signatures of saturation and non-linear evolution.

4. **Top-quark studies:** Top-quark production will increase considerably in yield, allowing to discover rare top-quark decays, and kinematical range, enabling to study direct and indirect signatures of new physics coupled to the third generation.

The physics exploitation of these measurements will require a close interplay with theory, posing entirely novel challenges in terms of particle multiplicity, complexity and dynamics. At scales considerably above the LHC energy, the EW interaction is no longer suppressed compared to the strong interaction, such that NLO and NNLO EW corrections will play an equally important role as their QCD counterparts. Many EW final-state objects can be identified with full kinematical information, thus requiring novel approaches to higher-order calculations. At the highest energies, multiple emission of EW gauge bosons will become increasingly significant, requiring it to be accounted for in parton-shower simulations and resummation. A coherent and accurate theoretical description of final states at future hadron colliders will require the incorporation of these advances into a single framework, combining QCD and EW corrections with multi-final-state matching, ideally interfaced to parton showers.

5 Methods and interdiscipinarity

In many cases, the cutting-edge calculations are not the result of a brute-force application of well-established methods, rather they are the outcome of the development of new and innovative concepts and calculational methods. Modern NLO approaches such as generalized unitarity connect loop integrals to tree-level amplitudes, making it possible to compute one-loop results by exploiting the powerful technology available at tree-level, such as highly efficient recursion relations. Novel algebraic methods have been developed that allow one to obtain coefficients of basis integrals appearing in one-loop amplitudes by solving a system of equations. Reverse unitarity is a concept that allows one to transform phase-space integrals in loop-integrals, which in turns makes it possible to take advantage of the most advanced results in multi-loop virtual amplitudes. Other ground-breaking ideas allow to extract complicated poly-logarithmic two-loop integrals as a solution of a system of differential equations. In all examples above, one could reformulate an unexplored, complicated problem as a simpler one, where one could apply and extend established computational methods. This ability to make connections between loosely related topics has accelerated enormously the pace in perturbative calculations in recent years.

Today, many technically challenging questions are still open, and their solution is required to address outstanding phenomenological questions. Referring to the examples above, it is not clear how to formulate generalized unitarity systematically at two-loop level and beyond, or whether it is possible to develop algebraic integrand reduction methods at higher loops. It is not clear how to extend reverse unitarity when phase-space cuts are present, making the method useful only for the calculation of fully inclusive cross sections. It is not known how to systematically use differential equations when elliptic integrals are involved. These functions typically appear when massive internal and external particles are considered, such as Higgs production at large transverse momentum. Finding solutions to these and other problems will be crucial to make progress in the field. Often steps towards progress in this mathematical directions are made in collaboration of particle theorists and mathematicians, as seen for example in recent years in the field of applied computational algebraic geometry.

Precision calculations rely heavily on the usage of computer algebra. They have often been driving the conception and continuous development of algebraic manipulation languages and packages, that have subsequently made their way to a much larger user base. Turning the results of precision calculations into numerical predictions for collider observables often requires high-dimensional integration, usually accomplished using Monte Carlo techniques. The development of analytical and numerical computational tools is taking place at the interface of particle physics, mathematics and computer science, requiring close interaction between researchers from different backgrounds. It will be absolutely crucial to ensure technical progress and conceptual leaps on these computational tools in order to meet the future challenges of precision calculations.

Machine learning (ML) and deep learning tools are playing an increasingly important role in particle physics and could very well become a breakthrough in addressing a large number of numerical and algebraic problems. ML is already applied in particle physics experiments to data taking, event identification, and event reconstruction. ML could also be used to detect departures of data from the SM in a model-independent way, to devise much more efficient adaptive MC integration algorithms, or to optimize the internal workflow of computer algebra operations.

While mathematical and computational skills are very valuable, so is the phenomenological intuition and the ability to formulate physics analyses that allow to exploit collected data in the most effective way. This includes designing observables that allow one to best measure fundamental SM parameters or have the highest sensitivity to new physics effects. As an example one can consider the photon parton distribution function which was first extracted with an order one uncertainty using LHC Drell-Yan data. It was later realized that one could use accurate data on the proton structure functions measured in deep-inelastic scattering to extract the photon PDF with an uncertainty of just a few percent. It is rare to obtain such a dramatic reduction in uncertainties, but it is rather common to substantially improve in sensitivity with optimized analyses and better observables. The above example also stresses the complementarity of pp and ep colliders.

Openness of our field and fruitful exchanges with neighbouring research areas such as formal quantum field theory, mathematics, computer science and very close contact to the experimental analyses will thus be crucial for the precision calculations that enable the full exploitation of future collider data.

6 Considerations on resources

The novel challenges of precision calculations for future colliders can only be addressed by a committed community of researchers that combines skills from particle theory, collider phenomenology, applied mathematics and computer science.

1. **Long-term project support:** Most of the research objectives in specific calculations, automation and integration into a coherent framework can only be attained in larger collaborations on substantial timescales. Assuring long-term support for projects that last often far longer than typical funding cycles is thus essential.

2. **Career development of researchers:** To encourage young researchers to engage in this challenging field, it will be of foremost importance to ensure the appropriate recognition of individual contributions, including technical developments and implementation. The need for an increasing specialisation of individuals in these projects will need to be properly accounted for when evaluating researchers for fellowships, grants, academic appointments and promotions. It will be crucial to train future research leaders who are able to go across disciplinary boundaries, identifying novel and unconventional problem-solving approaches, and maintaining the diversity of the research portfolio.

3. **Support for computing codes:** Tools for algebraic and numerical computations (such as computer algebra languages or Monte Carlo codes) are most widely used in precision calculations. Ensuring their development and long-term support will be crucial for progress in the field. At present, many of the relevant codes are developed and maintained by small research groups, often funded through a single institution or funding agency. These crucial activities call for a broader community-wide support model ensuring their long-term sustainability.

4. **Computing resources:** Today's state-of-the art calculations can involve the extended use of computer farms with several thousand cores and hundreds of high-end GPU units. The future availability of comparable or even increased computing resources will be of the utmost importance for continued progress in the field. The required resources vary considerably in type and scope in the course of a project, ranging from smaller scale requirements for algorithm and code development and testing of novel hardware architectures to large-scale structures for production. In particular for smaller university-based groups, it becomes increasingly challenging to ensure the timely availibilty and operation of these resources, thereby calling for new national and international models of resource sharing.