Muon Collider Forum Report

Forum Conveners: K.M. Black\textsuperscript{1}, S. Jindariani\textsuperscript{2}, D. Li\textsuperscript{3}, F. Maltoni\textsuperscript{4,5}, P. Mead\textsuperscript{6}, D. Stratakis\textsuperscript{2}, and the full author list is being assembled.

\textsuperscript{1}University of Wisconsin-Madison, United States
\textsuperscript{2}Fermi National Accelerator Laboratory, Batavia, Illinois, United States
\textsuperscript{3}Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA, United States
\textsuperscript{4}Center for Cosmology, Particle Physics and Phenomenology, Université catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium
\textsuperscript{5}Dipartimento di Fisica e Astronomia, Università di Bologna, Bologna, Italy
\textsuperscript{6}YITP, Stony Brook, United States

June 23, 2022

Cross-Frontier Muon Collider Forum Report Submitted to the US Community Study on the Future of Particle Physics (Snowmass 2021)

Contents

1 Executive Summary 3

2 Introduction 5
2.1 Muon Collider Forum and Report ................................. 5
2.2 Why Collide Muons? ................................................. 6
2.3 Muon Collider History ............................................. 7
2.4 International Muon Collider Collaboration .................... 7

3 Physics Case 8
3.1 General Introduction .................................................. 8
3.2 Physics cases at different energies ............................... 11
3.3 Higgs Boson ........................................................... 12
3.4 Dark Matter ............................................................ 16
3.5 Naturalness ............................................................. 19
3.6 Complementary Probes ............................................ 21
3.6.1 Flavor and CP violation ......................................... 22
3.6.2 g-2 Anomaly ....................................................... 23
3.7 Future Theory Development for Simulation .................. 24

4 Accelerator 25
4.1 General Introduction .................................................. 25
4.2 Feasibility statement ................................................. 27
4.2.1 Proton Source ..................................................... 27
1 Executive Summary

“One ring to rule them all, one ring to find them...” J.R. R. Tolkien

A multi-TeV muon collider offers a spectacular opportunity in the direct exploration of the energy frontier. Offering a combination of unprecedented energy collisions in a comparatively clean leptonic environment, a high energy muon collider has the unique potential to provide both precision measurements and the highest energy reach in one machine that cannot be paralleled by any currently available technology.

The LHC has fundamentally altered the high energy particle physics landscape in the last ten years. The discovery of the Higgs boson has laid out a clear path to study numerous foundational questions about our universe. However, also crucial is the glaring absence of definitive signs of new physics up to ~ the TeV scale. We now have a definite picture from data which did not exist a decade ago, where there is a Higgs like particle at 125 GeV and then a (possibly large) gap to the scale of physics beyond the Standard Model. While this reinforces the need to study the Higgs boson properties in greater detail, it also sharpens the need to significantly extend the reach of the energy frontier to its utmost. Motivated by theoretical studies, experimental results, and the desire to explore the energy scale well beyond the LHC, a natural goal for the next Energy Frontier collider is the 10+ TeV scale. Such a scale can be directly probed at lepton colliders running at the corresponding energy, while for hadron colliders one would need much higher energies, in the range of 100 TeV. However, the only indisputable target set by the LHC thus far is the Higgs boson, which resides at the weak scale. Maximizing the physics potential at the energy frontier for precision Higgs studies can profit enormously from a cleaner environment than that offered in hadron collisions. Therefore, given the LHC results, a 10+ TeV lepton collider going beyond the classic precision versus energy dichotomy is an ideal machine. Unfortunately with conventional technology electron based colliders cannot reach this scale. Consequently, to probe this scale with leptons we must investigate completely new options, namely muon colliders or wakefield acceleration of electrons.

A high energy muon collider represents a paradigm change for the community. Its attractiveness is at least twofold. First, it offers a unique range of physics opportunities opened by an accelerator design where luminosity goes in par with the energy, providing the ideal combination of precision and direct reach needed for the exploration of the 10 TeV scale. Second, it is timely, not only from the physics needs, but also from the technological developments. The accelerator challenges can now be overcome using the technological advances achieved over the past decade. Significant progress has been made in the development of high power targets and of high-field HTS magnets, in the demonstration of operation of normal conducting RF cavities in magnetic fields, and of the self-consistent lattice designs of the various subsystems. No fundamental show-stoppers have been identified. Nevertheless, engineering challenges exist in many aspects of the design and targeted R&D is necessary in order to make further engineering and design progress. Consideration of issues related to sustainability makes a high-energy muon collider particularly attractive due to its relatively small footprint and excellent power efficiency.

There is an established plan for supporting muon collider related R&D activities in Europe and it is imperative for Snowmass/P5 to reestablish R&D efforts in the US and to enable participation of US physicists in the International Muon Collider Collaboration (IMCC). Successful execution of the IMCC program relies heavily on US participation, in particular in the areas of unique technology expertise (e.g. ionization cooling, targets, high-field magnets, etc) acquired during the Muon Accelerator Program in the U.S. Furthermore, renewed US contributions in muon collider efforts provides for an ideal opportunity to make designs for a potential US siting. The existing accelerator facilities at Fermilab (e.g. PIP-II) and the proposed booster upgrade or extension of the PIP-II linac could be developed into proton drivers for a Muon Collider. Furthermore, synergies with the neutrino and intensity frontiers programs in the US, as well as overlaps with nuclear science and industry applications provide additional benefits. This makes Fermilab a particularly attractive siting option. A set of Muon Collider design options, with potential siting at FNAL, would allow for deliberations with the IMCC and the international committees to facilitate a global consensus on the selection of siting.

Despite the compelling physics case and the feasibility of the accelerators, a series of challenges arise from the unstable nature of muon beams that have a wide range of implications. Firstly, the decay of muons in the collider ring produces an intense flux of neutrinos that exit from the ground at a significant distance from the collider. Secondary particle production induced by these
high-energy neutrinos has been of concern in the past. Novel ideas to reduce the density of this flux have been developed, including placing the beam line components on movers and tilting the ring periodically in small steps such that the muon beam does not point to a specific location for an extended period of time. It is expected that these will reduce the flux to a negligible level. It should be noted though, that high-energy neutrinos while posing an experimental challenge provide also an opportunity of synergies with other physics projects, as outlined in this report. Secondly, in the context of experiments/detectors, muon decays induce a possibly overwhelming background, i.e., a cloud of low momentum particles that end up in the detector volume that can impact physics performance and has to be curbed. Feasibility of suppressing this background to the levels necessary for precision physics was often questioned in the past. However, major advancements in collider detector technologies have been achieved in the last decade, driven in large by needs of the High Luminosity LHC experiments. These advancements allow one to incorporate precision timing and particle-flow into the tracking and calorimetry systems, which in turn translates into a dramatic reduction of the BIB. Full simulation studies demonstrate that necessary performance can be achieved with a reasonable extrapolation of the known technology. Further improvements are certainly conceivable and can be achieved by further optimization studies, including deployment of Machine Learning techniques.

The most fruitful path forward towards the development of a conceptual design of a Muon Collider would be the engagement of the U.S. community in the IMCC. The U.S. Muon Collider community is well positioned to provide crucial contributions to physics studies, further advance the accelerator technology and detector instrumentation, and explore options for domestic siting of a muon collider. An Integrated National Collider R&D Initiative being discussed in Snowmass can provide a much needed platform for R&D funding for such accelerator and detector development. From the exploratory studies performed so far, it is already evident that the muon collider has the potential to illuminate various open questions in fundamental interactions, from the origin of electroweak symmetry breaking to the nature of dark matter, to naturalness. It is also capable of covering a broad swathe of well-motivated physics beyond the Standard Model and probing explanations for potential signals in experiments across the many frontiers of particle physics. The accelerator and detector technologies have advanced in the last decade, making the collider feasible on the timescale of approximately 20 years. Engaging in the development of a muon collider will reinvigorate the U.S. HEP community and provide benefit to the field across multiple frontiers.

“We choose to go to the moon. We choose to go to the moon in this decade and do the other things (accomplishments and aspirations), not because they are easy, but because they are hard, because that goal will serve to organize and measure the best of our energies and skills, because that challenge is one that we are willing to accept, one we are unwilling to postpone, and one which we intend to win.” John F. Kennedy
2 Introduction

2.1 Muon Collider Forum and Report

There has been a recent explosion of interest in muon colliders, as evident from a ten-fold increase in the number of related publications submitted to arXiv in the last couple of years. The topic generated a lot of excitement in Snowmass meetings and continues to attract a large number of supporters, including many from the early career community. In light of this very strong interest within the US particle physics community, Snowmass Energy, Theory and Accelerator Frontiers created a cross-frontier Muon Collider Forum in November of 2020. The Forum [1] has since been meeting on a monthly basis and has invited many experts to give their perspective and to educate a broader community about physics potential and technical feasibility of muon colliders. It facilitated strong bond and exchange of new ideas between the particle physics community and accelerator experts. Synergies with the neutrino and intensity frontiers, as well as overlaps with nuclear science and industry applications have also been extensively discussed. Finally, the Forum served as an interface between the US community and the International Muon Collider Collaboration (IMCC) [2] hosted by CERN. The Forum’s mailing list has over 160 active subscribers, while between 50 and 100 physicists participated in each of the regular Forum meetings. In addition to the regular meetings, the Forum organized few topical workshops dedicated to physics, accelerator technology, and detector R&D. There were over 400 registered participants in the Muon Collider Agora event organized jointly by Snowmass and Fermilab Future Colliders group. Activities of the Forum culminated in writing of this report.

The purpose of the report is to summarize findings of the Forum and to provide input to the Snowmass/P5 process. In writing of this report, emphasis was made on a significant shift in physics motivation and technology readiness since the previous Snowmass/P5 process. While the previous Snowmass focused mainly on the 125 GeV Higgs factory concept, more recently the primary interest of theoretical and experimental communities has shifted to a higher energy machine, with \( \sim 10 \text{ TeV} \) collision energy as the target. This is due to the fact that some of the \( e^+e^- \) Higgs factories are technologically more mature than \( \mu^+\mu^- \), while physics potential of a high-energy muon collider is truly unmatched. Staging at 125 GeV is still motivated and provides a complementary to \( e^+e^- \) program with model-independent measurements of the Higgs boson mass and width, and a highly precise muon Yukawa coupling. Additional staging options are also being studied, including (a) operation at \( \sim 1 \text{ TeV} \) with maximum luminosity to enable physics above \( tt \) threshold and to probe certain beyond Standard Model (SM) scenarios; and (b) running at 3 TeV inspired by CLIC physics case.

In the report we also demonstrate that over the past decade we have seen the beginnings of several transformative new developments in accelerator and detector technologies. These developments address many of past concerns about muon collider feasibility. We present a technically limited R&D timeline and argue that investments into both accelerator and detector development are necessary to make muon colliders a reality on the timescale of approximately two decades. We believe that U.S. HEP community possesses critical expertise and is uniquely positioned to make leading contributions to the global muon collider efforts. Activities within the Snowmass Muon Collider Forum already identified key areas of interest and expertise, assuming that P5 will support a revival of the Muon Collider R&D program.

The Muon Collider R&D program would fit well within the timeline of U.S. HEP funding. In order to see this, it is useful to look at the landscape of currently ongoing large-scale facilities. A possible Muon Collider timeline is sketched in Fig. 1. For approximately the next decade, most of U.S. and Europe construction funds will be devoted to LBNF/DUNE facilities and HL-LHC upgrades. Successful completion of these mega-projects and execution of their physics program are essential for the future of particle physics. Without a significant increase in the HEP budget, it is difficult to envision a large scale investment into future colliders before completion of DUNE and HL-LHC upgrades. However, a modest amount of funding for the Muon Collider R&D would enable further technological progress and a preliminary Conceptual Design Report around the time of the next Snowmass process. Following the completion of DUNE in the early 2030s, more funds can become available for a larger investment towards the construction of muon collider demonstration facilities with the goal of producing a Technical Design Report and declaring the project ready for construction in early 2040s. It should be noted that decisions about International Linear Collider (ILC) in Japan and feasibility of Future Circular Collider (FCC) tunnel are expected within the next five years. While these decisions can alter the collider landscape, in all scenarios a high-energy
muon collider will remain a highly attractive future collider option with unique physics capabilities.

![Timeline diagram showing HL-LHC, LBNF/DUNE, and MuColl stages with decision trees](image)

Figure 1: A sketch of what the Muon Collider timeline could look like, superimposed with approximate HL-LHC and LBNF/DUNE schedules. Future collider decision tree adopted from Ref. [3] is also shown. The decision tree is "optimistic" in the sense that the timeline is driven by physics goals and technology readiness rather than financial considerations. We also assume that globally more than one future collider can be pursued at the same time.

2.2 Why Collide Muons?

Despite the incredible success of the SM at predicting various particle physics phenomena, many questions remain open. Discovery of the Higgs boson in 2012 helped to shed light on the origin of mass but does not explain why electroweak symmetry breaking occurs and what sets its scale. Other unanswered questions have to do with the origin of Dark Matter, the origin of flavor, and the nature of the neutrino sector. We also do not know if there is a fundamental reason for the gauge symmetry and what kind of unification of the known forces may exist at the higher energy scales. Conventionally, answers to these questions are pursued by probing small distances with either precision (indirectly) or energy (directly). The Muon Collider has the potential to provide both, leveraging full energy of the accelerator with a relatively clean environment.

A facility colliding high-energy muon beams has a number of advantages when compared to its electron-positron and proton-proton counterparts [4]. First, since the muon is a lepton, all of the beam energy is available in the collision. Second, since the muon is roughly 200 times heavier than the electron and thus emits around $10^9$ times less synchrotron radiation at the same energy, it is possible to produce multi-TeV collisions in a reasonably compact circular collider. Finally, a high-energy muon collider is the most efficient machine in terms of power per luminosity [5], a very important consideration in light of the global push for a more energy efficient and sustainable future.

In principle, muon colliders can reach very high energies in excess of 100 TeV. In order for this to happen, the size of the accelerator ring will have to be sufficiently large (e.g. 100 km ring would enable a 40-60 TeV collider). Further considerations such as cost, power consumption, and construction time may also impose practical limitations for what energy and luminosity are achievable, e.g. for energies much greater than 10 TeV synchrotron radiation must be taken into account similar to electron colliders at much lower energy. However, there are no fundamental physics reasons that would prevent it from going well beyond what is achievable by any other currently proposed technology.

While the above arguments are highly appealing, there are several challenges with muons. First, muons are obtained from decay of pions made by higher energy protons impinging on a target. The
proton source must have a high intensity, and very efficient capture of pions is required. Second, muons have very large emittance and must be cooled quickly before their decay. Given their short time, ionization cooling [6] is the only viable option. Moreover, conventional synchrotron accelerators are too slow and recirculating accelerators and/or pulsed synchrotrons must be considered. Because they decay while stored in the collider, muons irradiate the ring and detector with decay electrons. Shielding is essential and backgrounds need to be strongly suppressed.

2.3 Muon Collider History

The concept of a muon collider is not new. Muon storage rings were mentioned in the literature in 1965 [7] and concepts for a muon collider and for the required muon cooling were developed in the 1970s and 1980s. A muon collider collaboration was formed in the U.S. in the 1990s which delivered a design study in 1999 [8]. In 2000 the Neutrino Factory and Muon Collider Collaboration (NFMCC) was formed [9] which set out to perform a multi-year R&D program aimed at validating the critical design concepts for the Neutrino Factory (NF) and the Muon Collider (MC). The Muon Accelerator Program (MAP) [10] was a follow-on (approved in 2011) program to the NFMCC and was tasked to assess the feasibility of the technologies required for the construction of the NF and the MC. At the conclusion of MAP the program had produced a number of significant milestones summarized in Section 4.1.

Although MAP was terminated in 2016, work continued on documenting the program’s results and has provided a “jumping-off” point for the recently formed International Muon Collider Collaboration (section 2.4).

2.4 International Muon Collider Collaboration

The 2019 update of the European Strategy for Particle Physics (ESPPU) identified muon colliders as a highly promising path to reaching very high center-of-mass energies in leptonic collisions. In response to these findings, the European Laboratory Directors Group (LDG) formed a muon beam panel and charged it with delivering input to the European Accelerator R&D Roadmap covering the development and evaluation of a muon collider option. In parallel, CERN initiated formation of a new IMCC to assess feasibility of building a high energy muon collider, identify critical challenges, and develop an R&D program aimed to address them. The effort includes development of the machine-detector interface (MDI), detector concepts, and an evaluation of the physics potential.

Figure 2: Schematic layout of 10 TeV-class muon collider complex being studied within the International Muon Collider Collaboration. From https://muoncollider.web.cern.ch/

The collaboration is hosted by CERN. The near-term goal is to establish whether an investment into a full Conceptual Design Report and a demonstration program are scientifically justified for the next European Strategy for Particle Physics Update. Depending on the outcome of this study and the decisions made at the next ESPPU, the design can be further optimised and a demonstration
program can be executed in the following years. The latter contains one or more test facilities as well as the development and testing of individual components and potentially dedicated beam tests. The resulting conceptual design will demonstrate the possibility to technically commit to the collider. In this case a technical design phase will follow to prepare the approval and ultimate implementation of the collider.

The design strategy taken by IMCC relies heavily on the concepts developed by the MAP collaboration. In the baseline design, muons are produced in decays on pions produced by colliding a multi-megawatt proton beam onto a target. The muons are then cooled to the emittances necessary to achieve target luminosities, rapidly accelerated to the desired energies in order to minimize the number of muon decays, and injected into a collider ring with two interaction points. IMCC envisions a staged approach with the first stage collider operating at the center-of-mass energy of 3 TeV and the second stage at 10+ TeV (Fig. 2). Staging allows for demonstration of performance at the lower energy and also facilitates stretching out the construction time, while executing a vibrant physics program. The front end and most of the cooling chain in the accelerator complex are common to all stages. An alternative approach (LEMMA), which uses positrons to produce muon pairs at threshold, was also considered but had difficulties with achieving a high muon beam current and hence the necessary luminosity.

Despite strong interest and expertise, U.S. participation in IMCC has been limited to the work done in the context of Snowmass. The European muon beam panel included two representatives (including the co-chair) from the U.S., and a large number of scientists helped to organize the IMCC working group activities. U.S. scientists made key contributions to most areas of the IMCC design development and planning, including magnets, RF cavities, muon production and cooling, muon acceleration, beam dynamics, machine-detector interface, and the high-energy complex. Besides the accelerator design, the Energy and Theory Frontier communities in the U.S. provided strong contributions in the areas of physics studies and detector design.

3 Physics Case

3.1 General Introduction

The most fundamental physics case for the energy frontier is not collider specific. The investigation of the unknown is an ubiquitous motivation in any area of science. Therefore exploring the shortest distances, the highest energies, and the earliest times in our universe will always be the prime motivation for the energy frontier. Any collider that can push the furthest in these directions will always be compelling regardless of the particles collided. However, what differentiates a muon collider at high energies is how it reaches the energy frontier and the unique range of physics that comes along with it. Colliders are often grouped into two distinct categories of precision (electron based) and discovery (proton based) machines. Muon colliders, however, do not really fit this naïve classification and have to be considered as a fundamentally different option for our field. Muon colliders being circular and compact provide a unique combination of energy, precision, and high luminosity. Thus they are a distinctly attractive option.

At its core what enables the remarkable physics potential of a high-energy muon collider is that it accelerates fundamental rather than composite particles. This has two key advantages which are normally competing in usual electron and proton based colliders, i.e., i) equivalent high energy collisions reached in a more compact setting and ii) a cleaner (non QCD dominated) environment to undertake physics studies in.

The direct reach of a muon collider at energy $E_{CM} = \sqrt{s}$ for a heavy particle of mass $M$, can be easily estimated considering the $\mu^+\mu^- s$-channel annihilation into a pair of heavy particles, whose kinematical threshold is at $M = \sqrt{s}/2$. Heavier states, such as $Z', W'$, or heavy Higgses [11] could be produced singly in association with a soft/collinear vector boson ($\gamma, Z, \pm \pm$), extending the mass reach beyond $M = \sqrt{s}/2$ and possible almost to $\sqrt{s}$.

A useful way to estimate the reach of a muon collider is to compare it to a $pp$ machine by means of an effective parton luminosity. At a hadron collider high-$Q^2$ events are produced through the collision of elementary constituents of the protons, the partons. By employing standard collinear factorization, the inclusive cross section for producing a particle final state $F$ can be written as

$$\sigma(pp \rightarrow F + X) = \int_{t_n}^{1} d\tau \sum_{ij} \frac{dL_{ij}}{d\tau} \hat{\sigma}(ij \rightarrow F),$$

(1)
Figure 3: The COM energy for a proton collider $\sqrt{s_p}$ and a muon collider $\sqrt{s_\mu}$ such that the $2 \rightarrow 2$ cross sections are the same based on different assumptions for the partonic cross sections characterized by $\beta$. Separate curves for gluon and quark annihilation channels are shown, with the bands given by two choices of PDFs, i.e. NNPDF3.0LO and CT18NNLO [12].

where $\hat{\sigma}$ represents the partonic cross section of partons (quarks or gluons) $i, j$ to produce the final state $F$, and $L_{ij}$ represents the parton luminosity constructed from the parton distribution functions (PDFs) as a function of $\tau = \hat{s}/s$. Given that a muon collider is colliding fundamental particles, in a $2 \rightarrow 2$ process it can in principle produce states of mass $M$ up to $\sqrt{\hat{s}}/2$, while at a proton collider there is an extreme suppression from the falloff of PDFs with $\sqrt{s}$. To estimate when a proton collider and muon collider would have the same inclusive cross section

$$\sigma_p = \sigma_\mu$$

(2)

to produce a pair of particles of mass $M$ in a $2 \rightarrow 2$ process, one needs the partonic cross section. Production processes could be different for a proton collider compared to a muon collider, yet we can characterize it by the ratio

$$\beta \equiv \frac{\hat{\sigma}_p}{\hat{\sigma}_\mu},$$

(3)

where the annihilation cross section $\sigma_\mu$ is calculated close to its maximum, i.e., slightly above threshold $\sqrt{\hat{s}} = M/2$. For annihilation processes involving EW processes $\beta \sim 1$, whereas for producing colored states one could imagine a scaling due to the coupling of order $(\alpha_S/\alpha)^2$.

In Figure 3, we show the resulting COM energies which yield equivalent cross sections. For example, a 100 TeV pp is equivalent to a muon collider of $\sqrt{s_\mu} \sim 20$ TeV. This is only illustrative since it depends on the details of the partonic cross section. For example, for a $2 \rightarrow 2$, the equivalent muon collider scale could be in the 5 TeV range [12]. However, as shown in numerous studies [12], the details depend on the specific physics cases, some of which we will be reviewed in this section.

Based on what we have learned from the LHC thus far since the last Snowmass/P5 process, the ability to reach higher energies in a compact collider is crucial. As of today, ATLAS and CMS have discovered a Higgs-like state at 125 GeV and nothing else [13]. This determines two priorities for future colliders. First, it is crucial to make as large of step possible in energy, e.g., to the $\gtrsim 10$ TeV scale. From what we know now, there may be a significant energy gap between the TeV scale that the LHC/HL-LHC is probing to the scale of new physics. The largest step possible in energy is motivated not just for the sake of exploration, but also by specific targets related to the origin of EWSB, Naturalness, Dark Matter, complementarity with other frontiers, and even...
existing experimental anomalies as we will discuss. Second, it is mandatory to achieve the most precise determination of the Higgs boson properties and its interactions. The Higgs boson is the only unambiguous discovery we have made at the LHC so far. While this is often thought as the domain of "Higgs Factories", a high energy muon collider also offers unique abilities that can take us beyond the first generation of Higgs factories. This is because not only a muon collider provides a compact energy frontier machine, but also because at high energies emission of W and Z bosons from the initial state muon is enhanced and vector boson fusion becomes dominant. This is the reason why a muon collider is often also dubbed a "vector boson collider" as well [14, 15, 12].

The vector boson fusion channel provides the dominant production mechanism not only for single but also for multi-Higgs final states. Cross sections are significant, increase with the COM energy and do not suffer from the copious QCD backgrounds that come along with an energy frontier proton collider, e.g., the FCC-hh. The dual nature of a muon collider, i.e., "high energy with high precision", allows for an unparalleled number of Higgs bosons that can be produced in the cleanest environment. This very fact has further ramifications. The Higgs precision measurement program is normally thought as staged in two main phases: the first phase at a Higgs factory, where deviations of Higgs properties predicted by the SM are detected, followed by a second phase at a different high energy machine to directly search for the cause of the deviations, e.g., at CERN the FCC-ee followed by the FCC-hh. This is because, when viewed through the lens of an EFT, the expected deviations in the Higgs couplings enter as $O(v^2/M^2)$ where $M$ is the new physics scale, which typically resides well above the energy probed at Higgs factories. On the other hand, the muon collider opens up a completely novel possibility for Higgs physics. A high energy muon collider will provide precision and energy at the same time, achieving the highest precision possible for many Higgs observables and directly test the physics which causes it. Thus a high energy muon collider compared to a standard multi-staged approach based on different colliders in the same facility, such as LEP and then the LHC, can be thought as a "two-colliders-in-one" or "one-stage" option. Putting this together with the opportunity of increasing the COM energy as the accelerator technology will advance, clearly makes the muon collider option a very attractive one.

The novel nature of a high energy muon collider is clear from these simple arguments. It allows the highest energy reach with the most compact design in space (actual size of the facility) and in time (high-luminosity can be obtained in a handful of years of running), providing a new type of collider that bridges the usual precision versus energy dichotomy. This unique features call for a systematic assessment of the BSM opportunities that a specific muon collider at 10 TeV could open up, to establish the reach in specific scenarios with physics studies that go beyond the simple parton luminosity scaling argument presented above. The 2021 Snowmass process has catalysed an explosion of interest in the high energy theoretical community, leading to a wealth of phenomenological studies, whose current number and breadth make it already impossible to cover all results in detail. Several references exist that try to summarize a great deal of the physics case [12, 16, 15]. For this forum report we emphasize a few specific areas that will make easy for the reader to grasp the impressive physics reach of a 10 TeV muon collider. In section 3.3, we review the current status of Higgs precision physics with muon colliders. In particular, we outline some of the novel probes only available at high energy, and give examples of the aforementioned complementarity between searching for deviations through precision and searching for their causes directly with the same collider. In section 3.4, we explore one of the other pressing mysteries of particle physics, i.e., the nature and origin of dark matter. By briefly reviewing the current status of the field, we show how muon colliders are able to probe minimal models featuring WIMP dark matter that are out of reach of present and future experiments except for the FCC-hh.

In section 3.5, we then discuss one of the questions that inexorably comes along with the LHC discovery of the Higgs, i.e., naturalness. From what we have inferred so far from the data collected at the LHC, and in particular, the values of the mass of the Higgs at 125 GeV and the lack of deviations from the SM predictions or of signs of new states, naturalness has to be understood in ways that were not prominent during the last Snowmass process. Studies performed so far have indicated that the muon collider has a significant reach in constraining scenarios motivated by the naturalness guiding principle, to an extent similar to the FCC-hh in some cases and clearly surpassing it in others. In Section 3.6, we demonstrate how a muon collider is naturally suited as a complementary probe for many other types of particle physics experiments that are already planned or will be undertaken in the coming future. As example of this complementarity we highlight how a muon collider is uniquely suited to probe existing anomalies in particle physics such as the Fermilab $g-2$ experiment. Our point here is not motivating a muon collider by
leveraging on the current (yet possibly to be solved) anomalies (such as \( g - 2 \), \( B \)-anomalies, \( W \) mass), but to provide evidence that a muon collider will play a role in addressing a wide range of possible hints of new physics emerging at scales probed at current and planned experiments. In Section 3.7, we outline some of the additional lines of research that should be undertaken to further improve the projections for muon colliders. Finally, although it is manifest that the full physics potential of a muon collider is unlocked by realizing it at the highest energies possible \( \gtrsim 10 \text{ TeV} \), in Section 3.2 we also consider the physics reach and targets that could be attained by staging it at lower energies. In particular, we discuss some complementary observables that could be measured at lower energies, which could be used to leverage the physics results of the high energy muon collider operation. We conclude by stressing in Section 6 further synergistic aspects, beyond the specific physics case, such as the impact that a muon collider could have on the neutrino physics program more so than other colliders and vice versa with thanks also to the shared infrastructure.

3.2 Physics cases at different energies

The ultimate physics reach of a muon collider depends on the highest energy that it can achieve. In this report we have focused primarily on the 10 TeV muon collider, because it is realizable with current technology and does not require any “miracles” to occur. However, there is not a hard upper limit at 10 TeV, in fact some physics studies have been performed up to 100 TeV [others, 12]. However, in the context of a 10 TeV muon collider, it can be useful to comment about the staging possibilities in reaching this energy. From the accelerator point of view, staging is a logical and prudent approach. For example, the IMCC investigated the option of a 3 TeV stage before achieving the 10 TeV energy [17]. It is therefore important to consider whether or not a complementary physics program exists with the earlier stage, and whether a multi-staged muon collider could actually be instrumental to the physics program. With the largest possible BSM direct reach in mind, a staged approach is obviously not particularly interesting. However, for accurate SM measurements, or for any process that benefits from lower energies, such as production of new light and weakly interacting states in annihilation, a staged plan could be beneficial. For example, as shown in Figure 4, for single Higgs cross sections as a function of \( \sqrt{s} \) there are qualitatively different aspects of physics that can be tested depending on the choice of center-of-mass energy.

![Figure 4: Figure reproduced from [18] showing various Higgs process as a function of COM energy. The dashed curves correspond to s-channel annihilation processes, while the solid curves all are from Vector Boson Fusion.](image)

In particular, we see that for Higgs physics, 3 TeV, is not a particularly attractive staging option in the absolute sense, as the s-channel processes have all fallen significantly from their maximum, and the VBF processes would only have larger cross sections at 10 TeV (including di-Higgs). Therefore in principle any Higgs measurements done at 3 TeV would only be superseded by 10 TeV measurements without offering new production channels that would be complementary to 10 TeV. This can be contrasted to the complementarity present potentially for \( e^+ e^- \) Higgs...
factories at low energy, where the processes are dominated by associated production rather than VBF. Although as a first stage 3 TeV would offer the highest precision compared to the HL-LHC and $e^+e^-$ Higgs factories in for example the di-Higgs channel until superseded by higher energy. Given this general chain of logic, it is therefore useful to consider various staging options based on the physics outcomes that could be achieved.

The canonical example of a low energy stage of a muon collider is a machine running at 125 GeV COM designed to produce the Higgs through an s-channel resonance. The idea of producing a Higgs through the s-channel muon-antimuon annihilation goes back many decades, see for example [19]. However, since the discovery of the Higgs in 2012, more detailed investigations have become available, pointing to the unique opportunity of directly measuring the Higgs width by a lineshape mapping process. Such a collider would also provide complementary Higgs coupling measurements, including a measurement of the muon Yukawa coupling at the subpercent level. The scenario where one or more $e^+e^-$ Higgs factories are constructed will not render the 125 GeV muon collider an uninteresting staging option. On the contrary, as shown in Ref. [20], there is a strong synergy between a 240 GeV $e^+e^-$ and a 125 GeV muon collider due to their different production channels. A combination of the two provides significantly better results on the coupling determination than individual ones. This will be discussed further in Section 3.3, where also the complementarity with measurements at 10 TeV is explicitly discussed. A possible disadvantage to this particular staging option is that it requires beam conditions that are more challenging than those needed at a higher energy muon collider.

There are of course other staging possibilities than 125 GeV or 3 TeV, although these have been worked out in the most detail thus far. A balance has to be struck between what makes a high energy muon collider so attractive in terms of luminosity and power consumption and the fact that lower energy stages would necessarily come with lower luminosity. At this moment, further studies are needed to optimize a staging plan to achieve the largest physics reach of a combined program within a muon collider and in conjunction with other colliders. Nevertheless, there are some obvious interesting options, some of which have received at least preliminary attention. For example, it was recently proposed that a muon collider running at the $2m_t$ with foreseeable luminosities could provide a sufficiently precise top mass measurement to answer the question of whether our universe is stable or metastable [21]. Additional possibilities for measuring the W mass and top mass more precisely are given in Barger:1997yk. At 10 TeV, a high energy muon collider also does not provide a strong measurement of the top Yukawa coupling compared the the LHC in the standard $tth$ analysis as can be understood from Figure 4. Although the ultimate sensitivity of a high energy muon collider still needs investigated, as new methods for measuring couplings can open up with energy, for example the $W^+W^- \rightarrow t\bar{t}$ process is sensitive to deviations in the top Yukawa [22, 15, 12]. Nevertheless, similar to the ILC, a muon collider stage in the 500 GeV to 1 TeV range, could provide complementary information to a high energy muon collider. Finally, if there is new physics at low mass with muon specific couplings there also could be benefits to a sub TeV staging option. An example of this is models built to account for the current muon g-2 anomaly and discussed further in Section 3.6.

Very preliminary investigations show that leveraging on runs at different energies is very promising and certainly deserves more studies. Staging options could exist that optimize several complementary open physics questions to the high energy muon collider. Ultimately, a detailed plan might depend on the physics landscape at the end of the HL-LHC and also on other accelerator projects foreseen. Nevertheless, the very possibility for a muon collider to operate at different energies from the sub-TeV to the multi-TeV range could significantly enhance its physics reach. Such studies will therefore be carried out in parallel leading up to the next Snowmass and ESG studies.

### 3.3 Higgs Boson

The Higgs boson, the only new fundamental particle discovered so far at the LHC, is the central figure of the Standard Model. As shown in Figure 5, taken from the corresponding Energy Frontier topical report for the Higgs, it connects to many of the deepest questions about our universe. Accurately measuring its properties to assess its nature and its role in phenomena that currently escape our understanding, is one of the top priorities of the high-energy community.

Measurements of the Higgs properties are also a powerful probe of new physics, and they play a central role in the physics programs of all foreseen future colliders. Muon colliders, in
particular, provide a unique setting to probe Higgs properties for two main reasons. First, multi-
TeV energies allow for the production of a larger sample of Higgs bosons than what is attainable
at Higgs factories, with a similarly clean environment to study them. In addition, they allow
multi-Higgs production and therefore an unmatched probe of the Higgs potential. Second, high
energy muon colliders offer the unique ability to simultaneously access Higgs properties with very
high precision/accuracy, and in case of deviations, directly probe their origin, as we discuss below.

To demonstrate the first point, we consider the precision on the Higgs couplings that can be
achieved at muon colliders. Drawing on the Higgs exclusive channel inputs of Refs. [18, 20] one can
perform a global fit analysis. There are two main approaches that are followed for doing the global
fits. The first is by assuming the same type of couplings as in the SM, but associating to each
of them a rescaling factor $\kappa_i$. This approach has been dubbed “kappa framework” and enjoys the
simplicity of a direct translation between different channels and the Higgs property precision. A
second approach employs a full-fledged effective field theory, the SMEFT, which provides a consis-
tent deformation of the SM which allows to perform accurate predictions and combine information
across different scales and experiments as long as new physics exists only at a parameterically
larger scale than probed. For consistency with the electroweak precision fit group at Snowmass,
we use a modified SMEFT framework, where the Higgs width can be considered as an additional
free parameter, yet not only Higgs measurements, but also electroweak precision observables and
possibly other low-energy measurements are included to achieve a consistent projection of the
overall precision. \footnote{We thank EF04 electroweak fitting group for various communications in developing the results.}

We show the SMEFT projection results in Figure 6. Here we only report the Higgs couplings
part in the Higgs basis, marginalizing on other parameters. The corresponding precision for the
electroweak sector and trilinear gauge couplings can be found in the Snowmass report [23]. In this
plot, all muon collider projections are combined with the HL-LHC. The muon collider scenarios
considered include a 3 TeV muon collider with 1 ab$^{-1}$ of luminosity, a 10 TeV muon collider
with 10 ab$^{-1}$ and also its combination with a 125 GeV resonant muon collider Higgs factory with
0.02 ab$^{-1}$ integrated luminosity. The semi-opaque and opaque bars represent the results with
and without the Higgs width $\Gamma_H$ left as a free parameter. As one can anticipate, considering
$\Gamma_H$ as a calculable parameter in the SMEFT allows to attain a better precision. On the other
hand, considering it a free parameter, introduces a "flat" direction in the fit, that needs very
specific measurements (such as the direct $\Gamma_H$ measurement at the resonance peak $\sqrt{s} = m_H$) to
be resolved. At high energies this can also be investigated by using indirect methods such as the
"ofshell" methods employed at LHC, and should have roughly the same precision as the direct
lineshape measurement but with added theory assumptions. We would like to emphasize that

![Figure 5: Figure from Energy Frontier Higgs topical report illustrating the centrality of the Higgs and the connections to numerous fundamental questions.](image)
Figure 6: The one-sigma precision reach on the effective Higgs couplings from a global fit of the Higgs and electroweak measurements in the SMEFT framework. The first set of (red) columns represents the HL-LHC S2 scenario with electroweak measurements at LEP and SLD. The second (blue) and third (yellow) sets of columns represent the 3 TeV muon collider and 10 TeV muon collider projections, respectively. The fourth (green) sets of columns represent the 10 TeV muon collider combined with a 125 GeV muon collider Higgs factory. The measurements are combined with the HL-LHC S2 and LEP/SLD measurements for all the muon collider scenarios. The semi-opaque bars represent the results with the Higgs width being a free parameter, e.g., allowing for exotic decays that are hard to constrain through direct searches. The solid bars are for the results without exotic Higgs decays.
these different frameworks and/or basis choices can be also associate to different UV hypotheses and are therefore useful also develop an idea of different new physics effects. It is important to keep in mind that there is no best approach that can be single out and on which one solely rely on in establishing the Higgs physics potential of various machines. In a rather general way, one can see that the precision on the Higgs couplings increases by an order of magnitude or more at a muon collider compared to the HL-LHC. In simple words, a muon collider could probe with an unprecedented resolution the inner structure of the Higgs boson. Furthermore, it is also evident that a 10 TeV muon collider can generally achieve much better precision compared to the 3 TeV muon collider. This gain in Higgs precision comes from the inherently higher luminosity at higher energies and the logarithmically enhanced weak boson fusion rates for Higgs boson production.

The precision Higgs program at the muon collider naturally goes beyond the study Higgs couplings to lighter SM states. One example is the Higgs top Yukawa which can be probed in very non-trivial way through the VBF $t\bar{t}$ process [15, 12]. In fact, a wide variety of differential Higgs measurements could be explored at a high energy muon collider. For example, as shown in [17], bounds on composite Higgs models go well beyond a 100 TeV proton collider by exploiting the differential measurements available. Furthermore a high energy muon collider allows for unprecedented measurements of multi-Higgs production and thereby the Higgs trilinear and even quartic couplings. These could be determined at better than 10% [24, 28] and O(1) [26] precision, respectively. One can further connect various Higgs-related processes to new physics hints from low energy precision measurements [12, 27], as well as directly probing hidden sector physics through Higgs exotic decays [28].

The second unique aspect of a muon collider is its ability to probe the causes of possible Higgs property deviations. The discovery of a deviation in the measured Higgs couplings would, at a Higgs factory, generally point to new physics outside the direct discovery reach of that collider. By contrast, a muon collider offers a unique opportunity: a single collider could both carry out precision measurements illustrating indirect effects of new physics on Higgs properties and directly discover the particles responsible. This is a powerful argument in favor of a high-energy muon collider. Other precision colliders generally aim to make a case for the next energy-frontier collider. The muon collider would already be such an energy-frontier discovery machine. Here we will give some examples that illustrate this capacity. As a first example, let us consider modifications to the Higgs gluon coupling due the existence of heavy colored partners of the quarks in supersymmetry. The fractional deviation in the Higgs coupling to gluons due to a loop of stops is estimated by [29, 30]

$$\delta\kappa_g = \frac{1}{4} \left( \frac{m_t^2}{m_{\tilde{t}_1}^2} + \frac{m_t^2}{m_{\tilde{t}_2}^2} - \frac{m_t^2 X_t^2}{m_{\tilde{t}_1} m_{\tilde{t}_2}} \right),$$

(4)

where $X_t$ is a trilinear mixing term. Using the precision for a 10 TeV muon collider quoted in [18] as a benchmark, and taking the two stops to be degenerate with $X_t = 0$, a muon collider Higgs precision measurement of $|\kappa_g - \kappa_g^{\text{max}}|$ would translate into a constraint of

$$m_t \gtrsim 1.5 \text{ TeV} \sqrt{\frac{0.67\%}{\delta\kappa_g^{\text{max}}}}.$$

(5)

On the other hand, the direct discovery reach for each stop squark at a high-energy muon collider extends up to very nearly $m_t = \sqrt{5}/2$, or about 5 TeV for a 10 TeV center-of-mass muon collider [12]. Thus, the collider will discover the same physics responsible for the measured Higgs coupling deviation. On the other hand, if the stops are sufficiently light, the measurement of $\delta\kappa_g$ could play a role in elucidating the detailed structure of stop mixing by helping to pin down $X_t$.

In composite models, the Higgs couplings to $W$ and $Z$ bosons receive corrections of order $v^2/f^2$, a result that follows from universal model-independent considerations when the Higgs is a pseudo-Nambu-Goldstone boson [31]. In the minimal composite Higgs model [32], for instance, one finds $\kappa_{W,Z} = \sqrt{1 - v^2/f^2}$. In this case, in the absence of a positive signal, we would obtain a bound on the scale $f$:

$$f \gtrsim 4.8 \text{ TeV} \sqrt{\frac{0.13\%}{\delta\kappa_{W,Z}^{\text{max}}}}.$$

(6)

The decay constant $f$ does not directly determine the mass scale of all composite states; generic composite mesons lying at the naive dimensional analysis scale $m_\rho \sim 4\pi f/\sqrt{N}$ could remain out of
reach of a 10 TeV muon collider if $f \sim 5$ TeV. However, composite Higgs models also contain other particles, like top partners, lying around the scale $f$ itself, and these particles play a major role in the naturalness of the theory (see, e.g., [33]). Thus, much of the parameter space where a precision deviation in $\kappa_W$ is observable could also lead to a direct discovery of new particles associated with compositeness. Because this scenario involves tree-level states with strong coupling, it is one of the cases where precision is expected to be farthest ahead of direct reach, so this is an encouraging conclusion.

Similar remarks hold for the Higgs oblique operator $\partial_\mu(H^\dagger H)\partial^\mu(H^\dagger H)$, which can affect the precision measurement of the di-Higgs production rate at a high energy muon collider [25]. One possible origin for such an operator is a singlet scalar mixing with the Higgs, as in the Twin Higgs scenario [34]. A 10 TeV muon collider could probe $f \sim 10$ TeV in direct searches for such a scalar $\phi \rightarrow hh \rightarrow (bb)(bb)$ [12]. In this case, the precision constraint and the direct search are extremely similar, with the former being a non-resonant search for the di-Higgs process and the latter a resonant search.

3.4 Dark Matter

Weakly Interacting Massive Particles (WIMPs) are natural cold dark matter (DM) candidates [35, 36, 37]. A representative case among the WIMP scenarios is the dark matter particle being the lightest member of an electroweak (EW) multiplet. The electroweak mass splitting among the members of the same multiplet is small compared to the overall mass scale. The high mass scale and near degeneracy render the DM searches at colliders extremely challenging. The model-independent mono-$X$ signals ($X = g, \gamma, W/Z, h...$) are not expected to reach a mass beyond two to three hundreds GeV at the high luminosity upgrade of the LHC (HL-LHC) [38, 39]. On the other hand, the disappearing track-based searches can extend the coverage up to 900 GeV for a triplet (Wino) [40]. This signature class relies on the mass gap between the members of the EW multiplet, which can introduce additional dependencies, in particular for the case of the Higgsino and the scalar multiplets.

A high energy muon collider can make powerful statements about the electroweak WIMP Dark Matter for a fermionic DM particle in connection with its thermal relic abundance. We adopt the benchmark choices of the collider energies and the corresponding integrated luminosities:

$$\sqrt{s} = 3, 6, 10, 14, 30 \text{ and } 100 \text{ TeV}, \quad \text{with } L = 1, 4, 10, 20, 90, \text{ and } 1000 \text{ ab}^{-1},$$

respectively. In the universal and inclusive signals, the particles in an EW multiplet are produced in association with at least one energetic SM particle. The soft particles or disappearing tracks are treated as invisible. This signal class is more inclusive and model-independent since it is not sensitive to the mass splittings within the EW multiplet. The most obvious channel is the pair production of the EW multiplet associated with a photon, which dominates the sensitivity to higher-dimensional EW multiplets. Additionally, vector boson fusion (VBF) channels unique to a high-energy muon collider [15] can also contribute. In particular, the mono-muon channel shows the most promise. After considering the inclusive signatures, we also perform a phenomenological estimate of the size of the disappearing track signal. Our findings show that the high-energy muon collider could substantially improve the constraints on thermal dark matter, serving as one of the main physics drivers for a high-energy muon collider.

We present the sensitivity results for the WIMP DM of Dirac fermion (DF), Majorana fermion (MF) in Fig. 7, obtained from Refs. [41, 42]. The reaches for 2$\sigma$ exclusion are shown for individual search channels, and various muon collider running scenarios are indicated by the color codes:

- The mono-muon channel, a unique signal for muon collider, shows much potential and it is especially promising for lower-dimensional EW multiplets, i.e. with $n \leq 3$.
- The traditional mono-photon channel at lepton colliders is suitable for higher-dimensional EW multiplets due to the coupling enhancement for high EW $n$-plets and the high multiplicity of the final state. In principle, one can consider radiation of other EW gauge bosons such as $W$ and $Z$ to improve the sensitivity [43, 44].
- The disappearing track signature will play an indispensable role in searching for EW multiplets. The mono-photon channel with one disappearing track will have the most significant signal rate and can extend the reach significantly for all odd-dimensional cases. Requiring
Figure 7: 2σ exclusion of fermion DM masses with horizontal bars for individual search channels and muon collider energies by the different colors. The vertical bars indicate the thermal mass targets [41, 42].

disappearing-track pairs will reduce the reach. However, providing a cleaner signal could turn out instrumental if the single disappearing track signature does now allow for enough background suppression.

The 2σ reaches for fermionic and scalar DM are summarized in a zoom-in version with fewer energies of $\sqrt{s} = 3, 10$ and 14 TeV in Fig. 8. The thick (darker) bars represent the reach in DM mass (horizontal axis) by combining different inclusive missing-mass signals. The thin (fainter) bars are the estimates of the mono-photon plus one disappearing track search. We have also included the target masses (vertical bars in black) for which the dark matter thermal relic abundance is saturated by the EW multiplets DM under consideration. When combining the inclusive (missing mass) channels, the overall reach is less than the kinematical limit $m_\chi \sim \sqrt{s}/2$, especially for EW multiplets with $n \leq 3$ due to the low signal-to-background ratio. It is possible to cover (with 2σ) the thermal targets of the doublet and Dirac fermion triplet with a 10 TeV muon collider. A 14 TeV muon collider can cover the complex scalar triplet. For the real scalar and Majorana fermion triplet, a 30 TeV option would suffice. The thermal targets of complex scalar and Dirac fermion (real scalar and Majorana fermion) 5-plet would be covered by 30 (100) TeV muon colliders. The 100 TeV option will also cover the thermal target for the complex scalar and Dirac fermion 7-plet. The real scalar and Majorana fermion 7-plet can be probed up to 30 – 40 TeV in mass at a 100 TeV muon collider, with their thermal target still out of reach. We note that in order
to cover the thermal target, the necessary center-of-mass energy and luminosity in many cases can be much lower than the benchmark values we showed in Equation 7. At the same time, the disappearing track signal has excellent potential and could be the leading probe for 5-plet or lower EW multiplet. Based on our study, it could bring the reach very close to the kinematical threshold $m_\chi \sim \sqrt{s}/2$. We note that a 6 TeV muon collider with a disappearing track search can cover the thermal target of the doublet case, motivating further detailed studies in this direction. A 3 TeV muon collider has sufficient energy to access the pure-Higgsino DM through the disappearing track channel kinematically. However, with the current detector layout design [45] and the short lifetime, the signal efficiency would still be too low [41]. The maximal signal efficiency can be estimated as follows. At $E_{CM} = 3$ TeV, the Higgsino would be produced relatively close to the threshold. With a lifetime of 0.02 ns, it would have a lab frame lifetime smaller than 0.56 cm, with a smaller transverse displacement. The single disappearing track reconstruction would have an efficiency at most $2 \times 10^{-4}$ without taking into account any experimental acceptance. The Higgsino production rate without the requirement of the existence of a 25 GeV $p_T$ photon is 10 fb. After requiring such a photon associated with the single track, the cross-section is 1 fb. Higgsinos will be produced with a pseudorapidity distribution, yielding an even smaller number of signal events in the acceptance region. All of these point towards less than one signal event. At the same time, the background would be around 20 from BIB. For a comparative analysis between our study and the full detector simulation study, see discussions in Sec. 5.6.2.

A powerful future constraint on dark matter via indirect detection, i.e., searching for annihilation products, will come from the CTA gamma ray telescope [46]. CTA observations of the Galactic Center will have sensitivity in the range of thermal relic higgsino cross sections. However, substantial astrophysical uncertainties remain, and if the dark matter distribution in the galaxy is cored on length scales of multiple kiloparsecs, higgsino DM could evade even CTA [47]. Similarly, searches for cosmic ray antiprotons from DM annihilation have the potential to constrain higgsino DM, but currently suffer from considerable modeling uncertainties related to cosmic ray propagation in the galaxy. The other class of dark matter detection, direct detection, tends to be less sensitive to the EW multiplet dark matter since their cross sections are loop suppressed and are subject to accidental cancellations between different diagrams [48, 49]. In addition, direct detection also depends crucially on the local dark matter energy density and velocity distributions,
which could vary the sensitivity a lot. In short, the high-energy muon collider could provide the leading and most solid probe of heavy EW dark matter, such as higgsino dark matter, independent of the astrophysical uncertainties.

### 3.5 Naturalness

The naturalness puzzle of the electroweak scale, also known as the hierarchy problem, has been a major driver for searches for new physics above the weak scale accessible at colliders (for a recent review see the Snowmass contribution [50]). It is deeply rooted in our quest for an explanation of electroweak symmetry breaking, or equivalently, a mechanism to generate the familiar Higgs potential, which is put in by hand in the Standard Model. This problem has only come to the forefront since the discovery of the Higgs boson rather than some other mechanism like Technicolor. Fine tuning measures the sensitivity of the Higgs potential to the UV physics in new physics models with a dynamically generated weak scale. While the computation of fine tuning might be taken with a grain of salt due to the lack of a sharp definition, the origin of electroweak symmetry breaking is a clear fundamental question of nature, which calls for continuous efforts at the energy frontier.

There are two categories of scenarios addressing the naturalness puzzle with different strategies: solutions of the “big” hierarchy problem, stabilizing the Higgs potential from scales around or not far above the weak scale all the way to the Planck scale; and solutions of the “little” hierarchy problem, extending the weak scale to some intermediate high energy scale, e.g., the highest energy scale probed experimentally. Currently LHC data has not provided a confirmed statistically significant hint of new physics related to the Higgs boson. This suggests two different possibilities. The first one is that there exists a gap between the weak scale and the new physics scale that solves the (big) hierarchy problem. In other words, the weak scale emerges and is stabilized up to the scale of quantum gravity, but with a large residual fine-tuning. The second possibility contains new particles filling in the gap between the weak scale and the TeV scale which is being probed at the LHC, as suggested in solutions to the little hierarchy problem. But these new particles have exotic collider signatures and are more elusive search targets. A high-energy muon collider demonstrates great potential in covering both possibilities, with either precision measurements (e.g., Higgs couplings) or direct searches. We will sketch some possibilities here. More details could be found in [12].

So far there are only two well-studied solutions to the big hierarchy problem: supersymmetry and compositeness. Let us consider supersymmetry first. In supersymmetry, three types of key superparticles, higgsinos, stops and gluinos, contribute to the Higgs potential at tree level, one loop and two loops respectively and determine the level of electroweak fine tuning. The specific reach at a muon collider requires detailed simulations. But a simple rule of thumb is that for distinctive final states (i.e., final states that allow efficient cuts to eliminate backgrounds while keeping a high signal efficiency), pair-produced new particles could be discovered up to the production threshold with mass scale $m \lesssim \sqrt{\frac{\kappa}{2}}$, at a high energy muon collider. For higgsinos, in the worst case scenario, the mass splitting between states in the multiplet is due to the Standard Model radiative correction. In this case, the final state consists only soft radiation from transitions within the multiplet, and a more dedicated search is needed, as described in the previous section on dark matter searches. Once on-shell decays within the multiplet are allowed given larger mass splitting, the rule of thumb above applies. In summary, a 10 TeV muon collider with 10 ab$^{-1}$ data could probe a higgsino up to 1 TeV in the worst-case small splitting scenario, and 5 TeV in the large splitting limit, which corresponds to percent level and per mille level fine-tuning respectively.

Next we turn to stops, $t\tilde{t}$, the superpartners of the top quark. Away from the limit in which the lightest stop is almost degenerate in mass with one of its decay products, the final states of stop pair production are usually distinctive, e.g., $tt\tilde{t} \to tt\tilde{t}$ missing energy. It is estimated that the mass reach scales as $m_{\tilde{t}} \sim 0.9 \times \sqrt{s}/2$, consistent with the rule of thumb [12]. At the HL-LHC, the expected 2$\sigma$ exclusion reach for stops is about 1.7 TeV. To compete with the LHC, a muon collider with $\sqrt{s} \gtrsim 4$ TeV is needed. For the benchmark 10 TeV center of mass energy and 10 ab$^{-1}$ data, stops with masses up to 4.5 TeV could be reached, associated with a $\sim 0.4\%$ level fine-tuning. More importantly, the stop mass is crucial to the supersymmetric prediction of the Higgs mass. It is well known that to obtain a 125 GeV Higgs in the minimal supersymmetric Standard Model with no significant left- and right-handed stop mixing, stops have to be in the mass range of $(5-10)$ TeV at large $\tan \beta \gg 1$ (for a review, see [51]). While in principle they could be lighter with large mixing,
it is important to understand what is needed to cover the parameter space implied by the Higgs mass. To test this prediction thoroughly, we need the muon collider to operate at \( \gtrsim 20 \) TeV. For small \( \tan \beta \lesssim 5 \), stop masses could be significantly larger, \( \sim (100 - 10^4) \) TeV, to accommodate the Higgs mass, in well-motivated models. Though it is a tall order to probe these scales even at future colliders, the models also predict significantly lighter electroweak supersymmetric states, such as electroweakinos. In particular, it is difficult to construct models that only have the stops as the lightest particles. Taking into account the Higgs mass, in predictive frameworks for SUSY masses that include the transmission of SUSY breaking effects, typically sleptons and electroweakinos are often in the \( \sim 1 \) TeV to few TeV range. These provide accessible motivated targets for a high-energy muon colliders in the \( O(10) \) TeV range. An example of the reach as compared to the HL-LHC and FCC-hh is shown in Figure 9 as reproduced from [16].

![Figure 9](image_url)

Figure 9: Figure from [16] showing the discovery reach of a top partner and several supersymmetric particles for a 10, 14, and 30 TeV muon collider shown as horizontal lines. The lightly shaded and darker bars correspond to the 95\% C.L mass reach of the HL-LHC and FCC-hh as determined for the European Strategy Update briefing book.

Lastly we want to comment briefly on the gluinos. Since the gluino does not carry electroweak quantum numbers, its production channel such as \( \mu^+ \mu^- \rightarrow \tilde{g} \tilde{q} q \) is of high order with a suppressed cross section. Though the gluino search will lag behind the stop and higgsino searches, a decent reach is still achievable: a 10 TeV muon collider with 10 ab\(^{-1}\) of data could match the gluino reach at HL-LHC.

In summary, a 10 TeV muon collider with 10 ab\(^{-1}\) data could be sensitive to per cent to per mille level electroweak fine tuning in supersymmetric models, probing important states closely tied with the Higgs potential such as higgsinos and stops.

In the other avenue to solve the big hierarchy problem, the compositeness scenario, fermionic top partners should exist at a scale \( m_T \sim y_t f \) with \( y_t \) the top Yukawa coupling and \( f \) the Goldstone symmetry breaking scale, parametrically below the compositeness scale \( m_* \sim g_* f \) with \( g_* \) the strong coupling, at which a plethora of bound states appear [33]. The rule of thumb will apply to the direct reach of the fermionic top partners as well. Similar to supersymmetry, a 10 TeV muon collider could detect or exclude top partners up to a few TeV, testing the scenario with per cent to per mille fine tuning. More importantly, the composite scenarios lead to sizable deviations in indirect electroweak observables, such as the Higgs couplings. As discussed in [25], a high-energy muon collider enjoys an advantage over all other future colliders in testing the Higgs composite
scales through constraining dimension-six operators such as

\[ O_H = \frac{1}{2} \left[ \partial_\mu |H|^2 \right]^2, \quad O_W = \frac{ig}{2} \left( H^+ \sigma^a D^\mu H \right) D^\nu W^a_{\mu \nu}, \]

(8)
due to a combination of high energy and clean backgrounds. In particular, a 10 TeV muon collider could probe the compositeness scale through the indirect precision measurements as high as \( m_* \sim 45 \text{ TeV} \) with 10 ab\(^{-1}\) data [25].

Solutions to the little hierarchy problem could evade the LHC searches if the partner states are neutral under the Standard Model and related to the visible sector only via a discrete symmetry, as realized in neutral naturalness models, e.g., twin Higgs [34]. A high-energy muon collider, however, is well suited to tackle the subtle signals in such scenarios through either precision measurements or direct searches. More concretely, the muon collider could measure the Higgs coupling deviations due to the scalar mixings, probe exotic Higgs decays with displaced vertices, and search for the Standard Model singlet partners, as well as the radial modes associated with spontaneous breaking of the associated discrete symmetry. In terms of direct searches, searching for the radial modes turns out to be a more promising strategy compared to the other partners. It has been shown that in a twin-Higgs setup, direct searches for the extra scalar at a 10 TeV muon collider could probe most of the parameter space corresponding to a \% level deviations in Higgs couplings, and even explore regions with smaller deviations which will be difficult to be observed in precision measurements [12]. In Figure 10, adapted from [12], we explicitly show the reach for both a neutral naturalness interpretation and more generally the Higgs portal as compared to the HL-LHC and FCC-hh projections.

![Figure 10](image-url)

Figure 10: Figures from [12], the left hand plot shows the exclusion limits for a massive scalar singlet \( \phi \) that mixes with SM Higgs with mixing angle \( \gamma \). The various colored curves correspond to different muon collider COM energies, while the expected limits at HL-LHC (solid) and a FCC-hh (dashed) are shown as black lines for comparison. The thin dashed lines indicate two possible scalings of the mixing angle in realistic models with fixed coupling. The left hand plot represents a more generic statement than that of naturalness, and is a more general illustration of the powerful Higgs portal reach of a muon collider. The right hand plot uses the same limits as the LH plot re-interpreted in terms of the reach on the sigma-model scale \( f \) in the context of a Twin Higgs model.

### 3.6 Complementary Probes

Some of the strongest constraints on new physics come from searches for flavor or CP violation, or other accidental or approximate symmetries of the SM. These include, for example, searches for electron and neutron electric dipole moments (EDMs), lepton flavor-violating (LFV) muon decays, and measurements of Kaon (or other meson) mixing. The sensitivity of many of these searches, in particular for EDMs and LFV decays or transitions, is set to improve by several orders of magnitude in the coming decades [52, 53, 54, 55]. Furthermore, new data from Cosmology, such as measurements of a stochastic gravitational wave background [56, 57] or primordial non-Gaussianities in Large Scale Structure and high-redshift 21-cm maps [58], may also provide hints of new physics [59, 60].

Given the broad scope of these indirect searches for new physics, there is ample reason for optimism that one or more of these experiments might detect a signal of BSM physics with the
increased sensitivity available. In fact, as we will discuss below, several potential signals of new physics already exist in the data. However, these indirect probes can only give a signal that new physics exists; they cannot say much about it. Furthermore, as has been seen on many occasions, indirect signals of BSM physics that at first were thought to be incontrovertible have become swamped by new backgrounds or systematics and remain ambiguous at best. For this, pushing the energy frontier along with these probes is essential, so that a discovery in a complementary, low-energy experiment can be quickly followed up by in depth studies at the energy scale of the new physics. Fortunately, for motivated extensions of the Standard Model that could lead to such signatures, the relevant scale is in the $O(1-10)$ TeV range, which is amenable to collider exploration.

A high-energy muon collider provides an ideal laboratory for complementing these low energy probes. Colliding muons at high energies can not only extend the reach to new electroweak states in the multi-TeV range, but also study processes that are closely related to those measured in low-energy probes at a different energy scale. Muon colliders are therefore well-suited for following up a signal in a complementary low-energy experiment with unique capabilities for discovering the new states and characterizing the structure of BSM physics. To demonstrate this complementary nature of high energy muon colliders, we focus on two different types of examples. The first falls under the class of future precision probes that may point to new physics in Section 3.6.1. There are of course numerous other exciting avenues in experiments that reside out of the energy frontier but have complementary ways to study them at muon colliders [12]. The second example we focus on is the fact that there are existing anomalies that may turn out to be a true sign of new physics. For example the numerous flavor anomalies, the recent CDF $W$-mass measurement, and the Fermilab g-2 measurement. Remarkably all the current hints in data are well suited to a high energy muon collider. As an example to demonstrate this we focus on the recent muon g-2 measurement, which clearly has connections to muon colliders in Section 3.6.2.

### 3.6.1 Flavor and CP violation

To demonstrate the complementary nature of high energy muon colliders, we look at the example of contributions to the electron EDM from loops of new, electroweak-charged particles. If these particles couple to the Higgs, they will produce a 2-loop EDM via Barr-Zee diagrams [61] (see Fig. 21 of ref. [12]), whose size is of order

$$d_e \sim \sin(\delta_{\text{CP}}) \frac{e m_e}{M^2} \left(\frac{\alpha}{4\pi}\right)^2 \sim 10^{-32} \, e \text{cm} \sin(\delta_{\text{CP}}) \times \left(\frac{20 \, \text{TeV}}{M}\right)^2,$$

where $\delta_{\text{CP}}$ is a CP-violating phase and $M$ is the mass scale of the new particles. The current bound from the ACME collaboration is $|d_e| \leq 1.1 \times 10^{-29} \, e \text{cm}$ at 90% confidence [62]. It is clear then, that a discovery of an EDM within the next two orders of magnitude in sensitivity would point to the few TeV scale for new physics. The particles responsible for the EDM could thus be pair-produced at a multi-TeV muon collider, and moreover, the large rates of Higgs and Higgs-pair production at these energies would also allow for precise measurements of their couplings to the Higgs. This allows the possibility of not only discovering the new particles responsible for an EDM, but measuring their interactions that contribute in loops directly. Given the limited information about the structure of new physics available in a single low-energy measurement, these complementary searches are essential for confirming our understanding of any discovery of new physics.

As a second example, consider the lepton flavor violating processes $\mu \rightarrow e\gamma$, $\mu \rightarrow 3e$ and $\mu \rightarrow e$ conversion in atomic nuclei. In motivated extensions of the SM, these processes arise from loops involving new states that are not flavor eigenstates. For instance, in models of supersymmetry, the slepton mass eigenstates will in general be mixtures of different flavors, but may be nearly degenerate, leading to a “super-GIM” mechanism that suppresses the LFV signatures and allows the new states to lie at the few TeV scale. As discussed in ref. [63], this situation is a highly-motivated target for the next generation of low-energy LFV experiments, but the physics responsible for any signature can also be studied in detail at a muon collider. Once again, a high-energy lepton collider would not only pair-produce the new states responsible, but would also allow for detailed studies of the flavor violation by measuring e.g., $\mu^+\mu^- \rightarrow \mu^+e^- + \text{missing momentum mediated by intermediate sleptons}$. For motivated models of supersymmetry breaking, a 3 TeV muon collider could probe all of the parameter space that will be explored by the Mu2e experiment, while a 10 TeV collider
would have reach comparable to the most optimistic future low-energy experimental proposals, as shown in Fig. 11. Alternatively, a muon collider could study lepton flavor-violating contact interactions, such as the $\tau_3\mu$ operator studied in ref. [12]. In either situation, the importance of searching for LFV interactions at the TeV scale is clear.

Another suite of indirect probes, which already shows some hints of a deviation from the SM, are tests of lepton flavor universality, e.g., in heavy meson decays. The most recent measurements of $R_{K^{(*)}}$, the $B$-meson decay ratios to muons over electrons, at LHCb [64], show an interesting discrepancy with the SM prediction. While lepton flavor universality is not a fundamental property of the SM, it is respected by the gauge interactions, and any deviation at the level currently being tested would be an unequivocal signal of new physics at the few to $10$ s of TeV scale, depending on the precise structure of the new interactions. A number of studies [65, 66, 67, 68, 69, 70, 27, 71, 72, 73, 74, 75, 76], summarized in refs. [17] and [16], demonstrate that a muon collider with energies at least $10$ TeV would decisively test the most plausible explanations for this anomaly, whether it be a new gauge force, a model involving leptoquarks, or some other new physics leading to $b\bar{s}\mu\mu$ contact interactions at the $10$s of TeV scale.

As these examples exemplify, while the indirect reach of precision probes can naively extend to very high scales, there are well-understood, natural ways of suppressing these processes that bring these targets to the TeV scale. Furthermore, given the hierarchy problem and the unexplained patterns of masses and mixing angles that exist even in the Standard Model, it is reasonable to expect that there might be new physics with interesting flavor structure lying near the TeV scale.

A broad program in precision physics is vital in searching for these new states, and a muon collider that extends the energy reach as high as possible is an ideal tool to supplement this strategy.

### 3.6.2 $g-2$ Anomaly

As an example of how muon colliders provide complementary sensitivity to existing anomalies, we investigate the overlap with the recent muon $g-2$ measurement. The persistent discrepancy between the theoretical and experimental values for $(g-2)_\mu$ has recently increased to $4.2\sigma$ between the recent measurement from the Fermilab $g-2$ experiment [77, 78, 79, 80] and the consensus value from the Muon $g-2$ Theory Initiative [81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105]. This observable may point to new physics beyond the SM, perhaps specific to the muon sector. A muon collider offers unique opportunities to probe the new physics of $(g-2)_\mu$ that no other experiment can offer, for two related reasons.

Most obviously, if a new SM singlet (either scalar or vector) resolves the $(g-2)_\mu$ anomaly, in the worst-case scenario where the singlet only couples to the muon, only a muon beam can
Figure 12: Required luminosity to yield a $2\sigma$ exclusion of singlets responsible for $(g-2)_\mu$. A 125 GeV muon collider (left) has superior sensitivity to 1 TeV (right) and to 10 TeV (not shown) through the combination of direct singlet production when kinematically accessible and deviations from muonic Bhabha scattering for heavier singlets, even considering the lower luminosity. Based on the analysis of Ref. [68].

probe the singlet coupling responsible for $(g-2)_\mu$. The maximum singlet mass is at the TeV scale, and to match the observed deviation in $(g-2)_\mu$, the coupling must increase with mass; thus, even if the singlet is too heavy to produce on-shell, its contribution to muonic Bhabha scattering may be probed directly at a muon collider (Fig. 12). In a preliminary analysis, Ref. [68] provides sensitivity estimates for both mono-photon searches for on-shell singlets (independent of the singlet decay channel) and deviations in Bhabha scattering. Ref. [27] persuasively shows that only a low-energy muon collider can fully probe the $(g-2)_\mu$ singlet parameter space above 1 GeV, providing a strong BSM physics motivation for a staged muon collider program.

If the new physics of $(g-2)_\mu$ comes instead from 1-loop contributions from electroweak-charged states, there must be at least two new particles (either two fermions and a scalar, or two scalars and a fermion), and a careful consideration of the possible SU(2) representations shows that there must always be at least one new charged particle [66]. A muon collider is once again the ideal vehicle for discovery of such states because of its strong sensitivity to electroweak physics and the availability of the full center-of-mass energy. The maximum mass scale of the new physics, assuming only perturbative unitarity, is around 100 TeV, but such corners of parameter space require extreme fine-tunings in both the muon mass and the flavor sector, let alone an additional fine-tuning of the Higgs mass. Ref. [68] phrased this optimistically as a “no-lose theorem” for $(g-2)_\mu$ at a TeV-scale muon collider: if no new charged particles are discovered, we will learn something profound about the role of fine-tuning in the SM. The situation is even better, though, because for new physics above the energy of the muon collider, the effective operator which yields $(g-2)_\mu$ will contribute a large excess rate for $\mu^+\mu^- \rightarrow h\gamma$ [67]. This process is essentially background-free in the SM, and thus a detailed analysis of the reducible background rate from e.g. Z’s faking Higgses at a muon collider will be needed to determine the true sensitivity. Regardless, the prospects for confirming the new physics of $(g-2)_\mu$ at all energy scales above 1 GeV are extremely strong.

3.7 Future Theory Development for Simulation

A multi-TeV muon collider will provide the unique opportunity of performing precision measurements, similar to those foreseen at $e^+e^-$ colliders and to explore very high-energy phenomena like at a $pp$ collider. Therefore, not only high-precision predictions will be needed for specific observables, but also fully exclusive simulations that can account for multi-particle QCD final state radiation and electroweak radiation from the initial as well as final states. Multi-TeV muon collisions will give the possibility to explore, for the first time in an accelerator context, SM interactions in a regime where EW symmetry is basically restored, i.e., where the weak scale $v$ is negligible with respect to the energies involved and a new set of phenomena, such for example multi weak boson emissions, may become dominant. Such SM yet novel effects will have an impact in the search and sensitivity to New Physics.
As discussed above, the physics at muon colliders at high energy can be schematically divided into two broad classes. In s-channel annihilation all collider energy flows into the high $Q^2$ interaction, making it possible to create new heavy states (with a reach in mass that is given by the kinematical limit) or very energetic SM final states. From the prediction/simulation point of view, this regime offers challenges qualitatively similar to usual $e^+e^-$ experiments, at least for what concerns QED and QCD. On the other hand, in the multi-TeV regime virtual and real emissions of soft and soft-collinear EW radiation can significantly impact cross-section predictions and therefore affect measurements. Production of multi-boson final states is enhanced due to large logs, leading to multi-jet/lepton signatures that can be used to reconstruct the original "weak partons" entering in the hard interaction, similarly to how QCD radiation patterns are used to associate jets to gluons or quarks [106, 107].

In weak-boson fusion, the emission low-virtuality vector bosons from the initial state muons, becomes dominant turning the muon collider into a high-luminosity vector bosons collider. This regime poses novel theoretical challenges both for precision as well as for Monte Carlo event generation. Even at the lowest order in pertubation theory, the large scale separation between $\sqrt{s}$ and $\gamma$, the necessity of including both QED and weak radiation respecting EW gauge invariance, and the relevance of multi-boson final states, make fixed-order event generation challenging for standard tools such as WHIZARD [108] or MadGraph5_aMC@NLO [15, 109]. In addition, with EW radiation becoming dominant makes a systematic theoretical reformulation of problems in terms of EW PDF's form the initial state and possible fragmentation functions for the final states necessary. Preliminary studies, see e.g., [110, 111], show that EW resummation effects can be significant at a multi-TeV muon collider.

Considering initial state and final state effects together, and noting the non-abelian nature of EW interactions, one could be lead to think that techniques developed over decades for QCD resummation would provide a suitable path to correctly account for these phenomena. This is far from being the case. At low energy, EW symmetry is broken and physical states display weak charges (at low energy physical states are QCD neutral) giving rise to very different experimental signatures, e.g., a $Z$ vs a $W$ boson. This fact also impacts the very definition of IR safety and inclusive observables that should be employed in EW calculations and the matching between an EW-symmetric forward/backward evolution to EW-broken initial/final states. In this context, even though far from having achieved an accurate understanding and implementation of the corresponding physics, significant progress has been made over the years. From the seminal work on the effective vector boson approximation [112, 113, 114, 115] to the more recent progress on resummation [116, 117, 118, 119, 107, 120], including the formal definitions and proof of factorization [121, 122, 123, 124, 125, 126]. The first implementations of EW showering have also become available with different degrees of accuracy [127, 128, 129]. The muon collider project will bring further motivation to the theoretical community towards developing a systematic understanding of EW radiation.

4 Accelerator

4.1 General Introduction

The concept of a muon collider is not new. Muon storage rings were mentioned in the literature in 1965 [7] and concepts for a muon collider and for the required muon cooling were developed in the 1970s and 1980s. A muon collider collaboration was formed in the U.S. in the 1990s which delivered a design study in 1999 [8]. In 2000 the Neutrino Factory and Muon Collider Collaboration (NFMCC) was formed [9] which set out to perform a multi-year R&D program aimed at validating the critical design concepts for the Neutrino Factory (NF) and the Muon Collider (MC). The Muon Accelerator Program (MAP) [10] was a follow-on (approved in 2011) program to the NFMCC and was tasked to assess the feasibility of the technologies required for the construction of the NF and the MC.

Although muons offer many potential physics benefits, their use brings substantial complications as well. Indeed, if intense muon beams were easy to produce, they would already be available. Firstly, muons are created as a tertiary beam. The commonly proposed production scheme uses a proton beam to bombard a high-Z target. This produces pions, which are captured in a solenoidal decay channel, where they decay to muons. To produce an acceptably large sample of muons, a multi-MW proton beam is required. A target system capable of tolerating such an intense beam
is a substantial challenge. The capture and decay process just described gives rise to a muon beam having a large energy spread and a large transverse phase space. The large transverse phase space has several implications: (1) It favors the use of solenoidal focusing in the lower energy portions of the facility, as opposed to the more conventional quadrupole focusing. A solenoid focuses in both planes simultaneously, avoiding the excessively large beam size in one plane when using an alternating gradient quadrupole channel. (2) It requires a rapid mechanism for reducing the emittance to more tractable values. The second major challenge of muon beams is due to the short lifetime of the muon, only 2.2 $\mu$s at rest. Clearly, the short lifetime puts a premium on very rapid beam manipulations. A fast emittance cooling technique, “ionization cooling,” is needed to reduce the transverse emittance of the muon beam, along with a very rapid acceleration system. The ionization cooling technique \cite{130, 131, 132, 133} requires high-gradient normal conducting RF cavities due to the need to immerse the cavities in a strong solenoidal magnetic field. Finally, the decays of the muons lead to potentially severe backgrounds in the detector of a Muon Collider. There are also a number of challenges related to the magnet requirements. In the target area, the initial capture magnet is a 20 T solenoid design. In the cooling channel, large aperture magnets having relatively low field, up to 2–3 T and 1.5-m diameter, are utilized. As the beam emittance is reduced through cooling, higher field solenoids with lower diameter bores are needed. In the final cooling stages of a Muon Collider, very high strength solenoids, up to $\sim 35$–50 T, are required. In the acceleration system, solenoids with very low fringe fields are needed to permit operation of nearby superconducting RF cavities. In the acceleration system and collider ring, special split-midplane or shielded dipoles are needed to accommodate the high heat load from muon decay electrons.

MAP developed the concept shown in Fig. 13. The proton complex produces a short, high intensity proton pulse that hits a target and produces pions. The decay channel guides the pions and collects the muons produced in their decay into a buncher and phase rotator system to form a muon beam. Several cooling stages then reduce the longitudinal and transverse emittance of the beam using a sequence of absorbers and RF cavities in a high magnetic field. A linac and two recirculating linacs accelerate the beams to 60 GeV. One or more rings accelerate the beams to the final energy. As the beam is accelerated, the lifetime in the lab frame increases due to relativistic time dilation so later stage accelerators have proportionally more time for acceleration, so that rapid-cycling or fast pulsed synchrotrons can be used. Fixed-Field Alternating-gradient (FFA) accelerators are an interesting alternative. Finally the two single bunch beams are injected at full energy into the collider ring to produce collisions at two interaction points.

The MAP Collaboration initiated its study with an evaluation of the feasibility of the key

---

**Figure 13:** A conceptual scheme for the Muon Collider and the NuMax Neutrino Factory. Different stages of the muon production, cooling and acceleration are shown. Various options of accelerating muon bunches to the desired energy (RCS, RLA, etc) are also indicated. Parts of the complex up to initial cooling are common to the Muon Collider and the Neutrino Factory.
subsystems required to deliver an energy-frontier collider. Several issues were identified as part of the MAP Feasibility Assessment that had the greatest potential to prevent the realization of a viable muon collider concept. These included:

- operation of RF cavities in high magnetic fields in the front end and cooling channel;
- development of a 6D cooling lattice design consistent with realistic magnet, absorber, and RF cavity specifications;
- a direct demonstration and measurement of the ionization-cooling process;
- development of very-high-field solenoids to achieve the emittance goals of the Final Cooling system;
- demonstration of fast-ramping magnets to enable RCS capability for acceleration to the TeV scale.

While other machine design and engineering conceptual efforts were pursued to develop the overall definition of a muon collider facility, research in the above feasibility areas received the greatest attention as part of the MAP effort. An important outcome of MAP was that progress in each of the above areas was sufficient to suggest that there exists a viable path forward.

4.2 Feasibility statement

4.2.1 Proton Source

A muon collider requires a high intensity proton source that produces intense pulses of protons that can be focused onto a target to produce pions that decay into muons. The general parameter requirements developed by MAP are a multi-MW source of multi-GeV protons producing pulses at a 5–15 Hz rate. Multi-MW proton sources have been and are being produced for spallation neutron sources and neutrino sources (SNS, ESS, J-PARC, Fermilab). Proton sources of the required intensity are within the capabilities of existing technology. The R&D challenge is to adapt and extend existing facilities to muon collider requirements, or design an optimized new source.

In the MAP program, an 8 GeV CW SRF linac capable of 4 MW was presented as the baseline proton source, which accelerated H\(^-\) beam into an accumulator ring using charge exchange injection. This was based upon the Project X design \[134\]. The initial part of that design is being constructed at Fermilab as PIP-II \[135\], an 0.8 GeV linac capable of 1.6 MW in CW operation. As discussed in Snowmass white papers, PIP-II can be extended to higher energies using a linac-based upgrade \[136\] or a linac plus RCS system \[137\]. These upgrades could be developed into proton drivers for a muon collider.

4.2.2 Accumulation and Compression

The high power beam from the proton source must be collected into short intense bunches on target for a collider. In the MAP scenario, 8 GeV H\(^-\) beam from a 4 MW Linac was accumulated in a small number of bunches in an Accumulator Ring (AR) at 15 Hz rep rate \[138\]. The bunches are then transferred to a Compressor Ring (CR) for compression to \(\sim 1\) m rms bunch lengths. Simultaneous delivery of all of the bunches from the CR onto the target can be obtained by directing each bunch through separate transport lengths in a “trombone” transport configuration onto the target. In Ref. \[138\], separate lattices of \(\sim 300\) m circumference for the Accumulator and Compressor are presented, along with simulations of bunch compression. A single ring combining accumulation and compression functions could also be considered.

The general concept can be implemented with other high power proton source configurations. Output beam from a suitable RCS or FFA could also be bunched and compressed in a ring transport. An essential part of the proton source R&D will be adapting the source to obtain the compressed bunches needed for the collider scenario.

4.2.3 Targets

The proposed R&D criteria for the target system for muon colliders are high pion/muon production yields and high tolerances of the thermal stress caused by beam impact. This is the extended
goal of high power target (HPT) systems for future neutrino facilities, such as the Long Baseline Neutrino Facility (LBNF) at Fermilab and the upgraded Tokai to Kamioka (T2K) facility at J-PARC. Neutrino target technology has matured to the 1 megawatt (MW) level which will be in operation at Fermilab and J-PARC in the next decade. The real challenge for the neutrino community is to produce the next-generation 2.4 MW LBNF target system. A key difference between MC and neutrino target systems is in the thermal stress on the system, because of the differing time structures of the incident proton beams. The proton beam for the neutrino target is multi-bunched within a microsecond (the bunch length is a few to a few hundred nanoseconds) and the beam repetition rate is around 1 Hz so as to generate a high integrated neutrino flux at a neutrino detector, while the proton beam for the MC will be single bunched with a 1-m-long rms bunch length and a beam repetition rate of 5–15 Hz to generate a high luminosity at an interaction region. As a result, the instantaneous thermal stress in the MC target system is several orders of magnitude higher than in the neutrino target system. Moreover, the repetition rate of the thermal stress cycle on the MC target is an order of magnitude higher than that of the neutrino target.

The international collaboration, Radiation Damage In Accelerator Target Environments (RADIATE), has been carrying out R&D on target materials for high power beam facilities since 2012 [139]. Various sorts of material response (mechanical, chemical and physical) have been studied using high energy beams and radiation. They propose to extend the study of novel materials for multi-MW beam facilities [140]. These materials include high-entropy alloys (HEA), electro-spin nanofiber materials, SiC-coated graphite and SiC-SiC composites, toughened fine-grained recrystallized (TFGR) tungsten, dual-phase titanium alloys, and advanced graphitic materials. In addition, liquid and highly granular targets, such as fluidized tungsten powder, will be studied. The MC target design effort will be part of the HPT material science.

Previously observed particle production measurements have been compiled as a “big data” set. This is used for calibrating particle simulation codes such as MCNP, FLUKA, MARS, Geant4, and PHITS. These simulation codes compute a radiation map and a precise production yield of secondary and tertiary particles and their time evolution in phase space in the beam transport system. The code is also used to estimate the instantaneous thermal stress due to beam impact. Time domain heat propagation in the target material is modeled by applying another simulation code, such as Finite Element Analysis (FEA) or a fluid dynamics program. Often, the interface among these different types of code is an issue. Further development of particle simulation codes is proposed in the white paper [141]. In addition, the application of artificial intelligence (AI) is considered for optimization of the dimensions and shape of the target [140]. The MC target design group will be part of this simulation R&D effort.

The target area and the front end consist of a series of large-aperture high-field solenoids. The (most challenging) 20 T target solenoid is composed of a 15 T field, 2 m aperture superconducting outsert and a 5 T field, 0.3 m aperture normal conducting insert. It is followed by 12 m long decay channel solenoids with tapered apertures from 2 m to ~0.6 m and field from 20 T to 2.5 T. The SC solenoid design faces several challenges including the severe radiation environment, high field, large stored energy and magnetic forces, etc. Given the field strengths, the outer SC solenoid may use hybrid Nb₃Sn inner and Nb-Ti outer coils. Detailed studies may be needed to develop reinforced cryogenically stabilized superconducting cables and validate a magnet cooling design that can withstand the high heat deposition. Experience with large Nb-Ti detector solenoids and the Nb₃Sn 13 T ITER central solenoid may be used. Possibilities for use of HTS cables need also to be studied.

Fermilab proposes the hot cell facility at the Target System Integration Building (TSIB) for the post-irradiation examinations (PIE) [142]. There is strong need for a new beam irradiation facility for target material R&D on the Fermilab site. PIP-II is a versatile machine that generates sufficient beam power to serve several HEP programs as well as beam irradiation tests [143]. A possible alternative is an upgraded Booster, which has sufficient beam power to serve both the beam irradiation facility as well as an ionization-cooling demonstrator [144].

4.2.4 Ionization Cooling

A complete scheme for cooling a muon beam sufficiently for use in a Muon Collider has been previously defined by MAP. The proposed scheme consists of a sequence of steps: first, an ionization cooling channel reduces the 6D emittance of the incoming bunch train until it can be injected into a bunch-merging system. The single muon bunches, one of each sign, are then sent through a second 6D cooling channel in which the transverse emittance is reduced as much as possible and
the longitudinal emittance is cooled to a value below that needed for the collider. If necessary, the
beam can then be sent through a final 4D cooling channel using high-field solenoids that further
reduces the transverse emittance while allowing the longitudinal emittance to grow.

An important outcome of MAP was that sufficient progress was made in each of the above
areas to suggest that there exists a viable path forward. For instance, a 6D cooling lattice was
designed that incorporated reasonable physical assumptions [145]; a final cooling channel design,
which implemented the constraint of a 30 T maximum solenoid field, came within a factor of two
of meeting the transverse emittance goal for a high energy collider and current development efforts
appear poised to deliver another factor of 1.5 improvement. In parallel, the MICE collaboration
built and operated a section of a solenoidal cooling channel and demonstrated the ionization cooling
of muons using both liquid hydrogen and lithium hydride absorbers [146]. This demonstration of
ionization cooling is an important advance in the development of high-brightness muon beams.

Together with these MAP Studies, since the end of MAP some relevant technology R&D has
continued to progress, enhancing the promise of muon colliders as an avenue of study. In particular,
new studies are required to leverage the now increased limits of solenoids and RF cavities, which
theory suggests should give an improved cooling channel performance.

4.2.5 RF in Magnetic Field

Because muons have a finite lifetime and are produced within a very large phase space, making
a compact and efficient muon ionization cooling channel is essential, in which high gradient RF
cavities must be placed in a strong magnetic field to compensate quickly for the muons’ ionization
energy loss. However, early RF experimental studies at the MuCool Test Area (MTA) at Fermilab
showed that the probability of RF breakdown events is significantly enhanced when cavities are
operated in external magnetic fields. As a result, the maximum achievable RF gradients were
seriously degraded. The impact on RF system performance is mainly caused by enhanced effects
of multipacting (MP) and field emission due to the external magnetic field, where field emission
electrons are focused by the magnetic field into beamlets that have current density high enough
to damage the cavity surface material and eventually lead to RF breakdown [147]. To mitigate
breakdown in external magnetic fields, two types of RF cavities were proposed and experimentally
studied in MAP: (A) vacuum and (B) high pressure gas filled cavities.

(A) Vacuum RF cavity: A low-Z, refractory, high electrical and thermal conductivity material—
beryllium—is chosen to terminate a conventional open iris RF cavity, taking advantage of
the penetrating nature of muons. As a result, the cavity has nearly a factor of two higher
cavity shunt impedance—a factor of two saving in peak RF power; moreover, beryllium can
withstand high peak surface field without surface damage from the magnetically focused
field emission electrons. In addition, the cavity inner surface was post-processed using the
techniques developed for SRF cavities, which include electropolishing (EP), dry-ice cleaning,
and assembly in clean rooms. The experimental RF program included cavities with beryllium
windows at two frequencies, 201 MHz and 805 MHz. The required RF performance was
demonstrated in the available experimental conditions at the MTA:

- a 201 MHz prototype cavity with thin beryllium windows for the Muon Ionization
  Cooling Experiment (MICE) achieved 10.3 MV/m in the fringe field of a 5 T solenoid
  field [148];
- an 805 MHz beryllium-wall modular cavity achieved 50 MV/m within a 3 T solenoid
  field [149].

The success of these experimental demonstrations indicates the importance of cavity design
(taking external magnetic field into consideration ), cavity surface finish, and understanding
of RF breakdown in magnetic field. Although significant progress was made in understanding
RF breakdown in magnetic field, there are still open issues that need to be further explored,
such as the RF gradient limit versus resonant frequency and cavity stored energy, cavity
frequency versus external magnetic field, etc. Our RF breakdown physics model predicts
that the breakdown probability depends on the temperature and material strength of the
cavity. With the recent advancement in C3 technology, a cryogenically cooled copper cavity
may have potential as the cavity option for muon ionization cooling channels in cases in
which the required performance outweighs system electrical efficiency. Nevertheless, design,
engineering, and high-power testing of a practical muon ionization cooling channel with RF remains one of the most important R&D tasks for a future muon collider.

(B) High pressure gas filled cavity: A dense hydrogen-gas filled RF cavity is proposed as the second method to overcome electrical RF breakdown in strong magnetic fields. Field emission electrons in the RF cavity are diffused by the dense gas via Coulomb scattering and are slowed down and prevented from cascading. Experimental studies confirmed that the maximum accelerating gradient in a gas filled RF cavity is indeed independent of the external magnetic field [150]. Moreover, dense hydrogen is an ideal ionization cooling medium because of its long radiation length and large ionization energy loss rate and has already been chosen as the preferred energy-absorber material [151, 152]. In addition, the gas-plasma-beam interactions have been studied experimentally by using an intense proton beam at Fermilab and numerically by a particle-in-cell plasma simulation [153]. This physics model of beam-induced-plasma RF loading has been developed and verified by experiments [152]. A plasma is created by the incident beam. The studies show that this beam-induced plasma is thermalized by interactions with neutral gas on a picosecond time-scale, and the electron density in the cold plasma can be minimized by adding a small amount of electronegative dopant, in order to reduce plasma loading of the cavities. The collective dynamics neutralizes the space charge in the incident beam. As a result, the incident beam is focused by a self-induced azimuthal field. This is a new plasma focusing mechanism that can be used to bring about extra focusing of muons [154, 155].

4.2.6 Reaching the Final Collider Emittances

The cooling channels described thus far cannot reach the transverse emittance of around 25 µm (normalized) required for the beams in the collider ring. A “final cooling” system first described by R. Palmer involving an alternation between uniform, high-field solenoids and acceleration was designed and simulated during the MAP program [156]. The beamline does not so much cool the beam as “exchange” emittance to reduce the transverse emittance at the cost of a longitudinal emittance increase. The design demonstrated the feasibility of the concept, but its performance was somewhat lower than expected. There are a number of research directions that could lead to improvements in that design. One direction likely to yield some improvement would be to perform more extensive studies of the optimization of that channel. Due to the nonadiabatic nature of the system and the low energies it reaches, maintaining the beam emittance even in the conservative parts of the system is challenging, particularly in the longitudinal plane.

In the final cooling stage, a series of 50 mm aperture solenoids with magnetic field above 30 T (ideally in the range of 50 to 60 T) will be needed. These solenoids require hybrid coils with HTS and LTS sections. The key challenges of this ultrahigh-field solenoid are related to choices of the HTS material and cable, forces and stresses in the coil, large stored energy, and quench management. The parameters of this solenoid go significantly beyond available technology, and it will require considerable R&D and demonstration, including studies and measurements of effects that are specific to ultrahigh fields.

The design in [156] assumed a maximum solenoid field of 30 T; a higher field would be expected to produce lower emittances. Hybrid superconducting–resistive solenoids exist with fields as high as 45 T [157], purely superconducting solenoids exist with fields as high as 32 T [158], and an R&D project toward a 40 T purely superconducting solenoid is funded [159]. Advances in high-temperature superconducting magnet design lead to hopes of achieving even higher fields. We should examine the impact that such higher field magnets would have on the performance of this channel. Such magnets can also be used to improve the performance of the cooling channel upstream of this system [160]. The impact of improved input emittances on this design should be studied, both on the channel’s overall performance and to provide input on any properties of the input distribution that would improve performance.

There are other ideas for replacing all or part of this last solenoid channel, or improving the beam characteristics entering that channel. One of the most frequently discussed is parametric resonance ionization cooling (PIC) [161], which relies on enhancing the beam divergence by inducing a half-integer resonance at an ionization absorber. Therefore, it does not require a strong focusing field. While in principle this channel can provide very small transverse emittances, the practical implementation faces challenges in managing chromatic and nonlinear aberrations [161]. We believe there are still possibilities for managing these aberrations, and the potential performance of PIC
warrants further theoretical studies.

There are other ideas for reaching the required collider emittances that are less well-developed, such as plasma focusing [162] and the slicing and recombining of beam distributions at higher energies [163].

4.2.7 Acceleration

Acceleration to moderate energies (10s of GeV) can be accomplished by the relatively conventional means of linacs and recirculating linear accelerators (RLAs). The MAP program proposed concrete designs for accelerating to 5 GeV, and suggested how an RLA could accelerate to 63 GeV [164, 165, 166, 167]. What was not studied in detail was acceleration from the final cooling systems to an energy of around 200 MeV. We expect this to use systems similar to those used in the cooling system, but without the requirements of high magnetic fields.

Reaching higher energies would ideally be done via pulsed synchrotrons. Rapid cycling synchrotrons are extensively utilized in existing accelerators, but the requirements for acceleration in muon colliders differ in two important aspects: short (ms-scale) acceleration cycles with a relatively low repetition rate (around 10 Hz), thus “pulsed” rather than “rapid cycling”; and the requirement for a high average bend field. To achieve a high average bend field, fixed field superconducting magnets are interleaved with bipolar pulsed iron dipoles, denoted a “hybrid” pulsed synchrotron.

Pulsed iron magnets can achieve fields of 1.5 T with manageable power consumption [168]; they could reach somewhat higher fields at the cost of increased power losses. Conventional copper coils can be used and make a relatively small contribution to the power losses, and small magnets have been built demonstrating the feasibility of the required short pulsed times [169]. HTS coils can be used instead if that prove more cost-effective, and have demonstrated significant ramping rates [170]. If that technology could be adapted to create pulsed fields well beyond what could be achieved with iron-dominated magnets, that could significantly improve the achievable energy and reduce the number of stages of pulsed synchrotron acceleration. Providing power to the pulsed magnets can be challenging and costly [171]. A single power supply can likely deliver only a small number of magnets. The peak power delivered is high and, to have reasonable average power requirements, a very large fraction of the energy delivered to the magnets in a pulse must be recovered. To maintain good beam properties, the pulse shape must be well-controlled and reproducible. No detailed design for such a power supply currently exists.

Fixed field alternating gradient accelerators (FFAs) have long been under consideration for muon acceleration. They allow acceleration over a reasonable energy range without the used of pulsed magnets. Earlier estimates sometimes found that they appeared less cost-effective to alternatives [172]; other considerations such as longitudinal emittance growth [173] and the cost of pulsed power supplies may cause one to reconsider that decision. Furthermore, more advanced FFA designs, such as vertical FFAs [174] or nonlinear FFAs [175, 176], could increase their energy range and make FFAs more favorable alternatives. For a neutrino factory, nonlinear FFAs were ruled out due to dynamic aperture considerations for the large transverse-emittance beams that were used there [175], but for a muon collider the transverse emittances are significantly smaller, and thus the nonlinear fields may not be a concern. A further direction of research would be to combine the pulsed synchrotron and FFA concepts, attempting to expand the capabilities of both technologies.

The impact of collective effects in acceleration systems must be assessed [177]. Accelerators are not duplicated for each bunch sign, but rather the two charges counter-rotate in each system, thereby colliding twice per turn. Those beams also pass through RF cavities, leaving significant wakes due to the high bunch charges. Since high-frequency RF is desirable for cost and efficiency, the energy extracted by a bunch from a cavity on each pass can be a significant fraction of the cavity’s stored energy, which would be a new way of operating an RF cavity.

4.2.8 Collider Ring

This section briefly summarizes the submitted white papers [178, 179] and our previous studies of Muon Collider (MC) Storage Ring (SR) and Interaction Regions (IR) [8, 180, 181, 182, 183] which brought together in a coherent form the results of the previous studies on a Muon Collider and presented a design concept of the 3 and 6 TeV MC optics, the conceptual design and parameters of large-aperture high-field superconducting (SC) magnets, and a preliminary design and analysis of the protection system to substantially reduce radiation loads on the magnets as well as particle
backgrounds in the collider detector. The SC magnets and detector protection considerations impose strict limitations on the lattice choice, hence the designs of the collider optics, SR and IR magnets and Machine-Detector Interface (MDI) are closely intertwined.

For a 3 TeV Muon Collider with $\beta^*$ of 5 mm it was possible to achieve the target design parameters with either a triplet or a quadruplet Final Focus (FF) with moderate strength quadrupoles.

The momentum compaction factor for a standalone arc cell is $\alpha_c = -0.004$. Each arc consists of six such cells and two dispersion suppressors. The betatron phase advance is 300° in both planes to ensure cancellation of spherical aberrations. Though the sextupoles for chromaticity correction are interleaved they are too weak to noticeably affect the dynamic aperture.

In the 6 TeV c.o.m. MC, as a first approximation we use the IR design with $\beta^*$ of 3 mm described in [184], whereas for the arcs we rescale the arc cell design of the 3 TeV collider [184]. The concepts of the 6 TeV IR layout and the 3 TeV arc cell are shown in Fig. 14. The next step toward a consistent design is to build the complete ring including utility sections for $\beta^*$ tuning and RF cavities. Chromaticity correction, crucial for achieving sufficient dynamic aperture and large momentum acceptance in the presence of very small $\beta^*$, will be based on the concepts developed in the previous studies [184, 185].

The baseline 3 TeV MC is based on 150 mm aperture dipoles and combined function magnets with the nominal dipole fields up to 10.5 T and field gradient up to 85 T/m in the arc, and focusing quadrupoles with field gradients up to 250 T/m and apertures from 80 to 180 mm and 180 mm aperture 8 T dipoles 8 T in IRs (triplet FF). These magnets are based on traditional cos-theta coil geometry and Nb$_3$Sn superconductor and were used to provide realistic field maps for the analysis and optimization of the arc lattice and IR design, as well as for studies of beam dynamics and magnet protection against radiation. The 11–12 T field level has been demonstrated in Nb$_3$Sn dipoles and quadrupoles developed for HL-LHC. The record field of 14.5 T was achieved in 60-mm and 100-mm aperture Nb$_3$Sn dipole models.

The 6 TeV IR design assumed use of HTS technology to achieve 20 T nominal operation fields in dipoles and 16 T pole tip fields in quadrupoles. The magnet inner bore diameter is required to be at least 5$\sigma_{\text{max}}$ plus 3 cm for absorbers or 24 cm. The flexible momentum compaction arc cells compensate the large positive contribution of the IR to $\alpha_c$. To mitigate neutrino induced radiation as well as energy deposition by decay electrons, the arc cell quadrupoles are combined-function magnets with dipole fields of about 9 T.

Figure 14: Top: IR layout of the 6 TeV c.o.m collider with 5σ envelopes and quadrupole apertures. In cyan are shown the quadrupoles with up to 5 T dipole component. Bottom: 3 TeV arc cell concept.
An important issue to address is the stress management in SC magnet coils to avoid substantial degradation or even damage of the brittle Nb\textsubscript{3}Sn and/or HTS superconductor. Stress management concepts for shell-type coils are being studied experimentally for high-field accelerator magnets based on LTS (Nb\textsubscript{3}Sn) \cite{186} and HTS (Bi2212 and ReBCO) cables \cite{187, 188}.

In the assumed IR designs, the dipoles close to the Interaction Point (IP) and tungsten masks in each interconnect region (needed to protect the magnets) help in reducing background particle fluxes in the detector by a substantial factor. The tungsten nozzles in the 6 to 600 cm region from the IP, assisted by the detector solenoid field, allow trapping of most of the decay electrons close to the IP as well as most of the incoherent $e^+e^-$ pairs generated in the IP. Analysis shows that with sophisticated tungsten, iron, concrete and borated polyethylene shielding in the MDI region, a total reduction of background loads between 2 and 3 orders of magnitude can be achieved.

An optimized MDI from the thorough MARS15 \cite{189} Monte Carlo simulation study of \cite{180} is shown in Fig. 15 with a sophisticated tungsten nozzle clad with borated polyethylene on each side of the interaction point, massive steel/concrete shielding around, tungsten masks in magnet interconnect regions, and steel and tungsten liners inside each IR magnet.

All the most critical concerns of the MC optics design, magnets, and radiation protection have not only been conceptually resolved but also addressed in sufficient detail for a 3–6 TeV c.o.m. muon collider which can be considered either for the International Muon Collider Collaboration or as a Fermilab site-filler. We call for establishing in the DOE OHEP a corresponding R&D program for 2023–2030. The R&D program would include extension of these studies to 10 TeV and higher energies.

A low-energy medium-luminosity MC has also been studied \cite{179, 180} and discussed as a possible Higgs Factory (HF). It was shown that electrons from muon decays will deposit more than 300 kW in the HF SC magnets. This imposes significant challenges on SC magnets used in the HF SR and IR. Magnet design concepts were proposed that provide high operating gradient and magnetic field in a large aperture to accommodate the large size of muon beams (due to the expected large transverse emittance), as well as a cooling system to intercept the large heat deposition from the showers induced by decay electrons. The distribution of heat deposition in the MC SR lattice elements requires large-aperture magnets to accommodate thick high-Z absorbers to protect the SC coils. Based on the developed MARS15 \cite{189} model and intensive simulations, a sophisticated radiation protection system was designed for the collider SR and IR to bring the peak power density in the superconducting coils below the quench limit and reduce the dynamic heat deposition in the cold mass by a factor of 100. The system consists of tight tungsten masks in the magnet interconnect regions and elliptical tungsten liners in the magnet aperture optimized individually for each magnet. These also reduce the background particle fluxes in the collider detector.
### 4.2.9 Neutrino Flux

One of the important issues [178] to consider when designing a muon collider is a potentially elevated flux of charged particles generated by neutrino interactions with soil and building materials at very large distances from MC [8, 190, 191, 192, 193]. Intense highly collimated neutrino beams, created from muon decays in the ring and various straight sections of a high-energy MC, can cause—to the surprise of many—radiation problems even at very large distances from the machine. The more energetic decay neutrinos emanate radially outward from the collider ring at angles with respect to the muon direction of order $\theta_\nu = 1/\gamma_\mu = m_\mu/E_\mu \approx 10^{-4}/E_\mu [\text{TeV}]$.

![Diagram of Intense Neutrino Fluxes](image)

**Figure 16:** Intense highly collimated neutrino fluxes around MC [191]

Neutrino flux and dose per neutrino at a given location from muon colliders grow with muon energy—keeping all other MC parameters the same—roughly as $E_\mu^3$ due to (each responsible for a factor of $E_\mu$):

1. Increase with energy of the neutrino cross section
2. Growth of total energy deposited
3. Collimation of the decay neutrinos

This will strongly impact siting issues and cost of a high energy muon collider and needs to be taken seriously in evaluating long-term averaged neutrino flux and resulting dose. Developed in [193], a weighted neutrino-interaction generator for the MARS Monte Carlo code [189] permitted detailed simulations of the interactions with matter of neutrinos and of their progeny in and around an MC, capable of modeling neutrinos in the energy range from 10 MeV to 10 TeV.

The model [193] serves to represent energies and angles of the particles emanating from a simulated interaction. These particles, along with the showers initiated by them, are then further processed by the MARS15 code [189] which calculates, e.g., energy deposition, absorbed, and effective dose as a function of location in a user specified geometry model. Effective dose—caused by charged particles from neutrino interactions—is calculated with particle- and energy-dependent quality factors taken into account. Muon and electron neutrinos and their antiparticles are included and distinguished throughout, which are represented in the decays from MC in roughly equal amounts.

The MARS15 model identifies charged and neutral current deep inelastic neutrino and antineutrino interactions with nuclei as the dominant channels forming the main contributions to the dose from neutrino interactions. Besides that, the model accurately describes neutrino-nucleon elastic and quasi-elastic scattering, interactions with atomic electrons, and coherent elastic scattering. In the latter, a Pauli form-factor of quarks—topological fluctuations of the QCD vacuum—is included (as a weight) to discourage small $|q^2|$ insufficient to liberate a nucleon or promote the nucleus to an excited state.

Extremely low interaction and scattering probabilities mean that neutrinos travel essentially in a straight line and survive over enormous distances. Much like neutrons and gammas, neutrinos...
by themselves cause little or no biological damage but instead create charged particles which in
turn deposit their energy in tissue to be interpreted as dose “due to neutrinos.” “Neutrino” dose
is thus by charged particles generated by neutrinos upstream of a human. Therefore, radiological
neutrino-induced effects around MC cause:

- Small effects for anyone above the ground and/or in an above-ground building;
- Noticeable effects inside a basement swimming pool;
- Unacceptably high effects for, e.g., a person lying in a basement room for extended periods.

Dose to a human body strongly depends on neutrino energy and presence of material immediately
upstream of that body. Total whole-body effective dose in a bare seated person (non-equilibrium)
is lower than that in one embedded in infinite soil (equilibrium). The whole-body dose is a factor
of 2 lower than the maximum dose, because a neutrino flux footprint could be smaller than typical
human dimensions. The equilibrium dose is achieved after 3–4 m of soil or concrete at all neutrino
energies considered here.

Instead of providing shielding, the presence of soil/concrete upstream enhances the dose by a
factor of up to 1000 in the TeV region as compared to the case with no shielding. The dose can
be higher than the annual off-site regulatory limits, e.g., 10 mrem/year at Fermilab. Contrary to
conventional irradiation, the use of high-Z shielding in front of a “human” can increase neutrino-
related dose by a factor of ten compared to low-Z shielding at low neutrino energies, while the
values converge in the TeV energy range. Fig. 17 (left) shows MARS15 calculated dose around the
2, 3 and 4 TeV c.o.m. MC rings in the orbit plane with 1.2 × 10^{21} decays per year vs distance in
soil from the ring center, while Fig. 17 (right) shows the annual dose as a function of soil thickness
downstream of a 0.5 m drift with 2.6 × 10^{16} decays/yr of a 1.5 TeV muon beam. One sees that for
the 4 TeV c.o.m. MC, the dose drops below the regulatory limit at a radial distance of 60 km from
the ring center, and at 55 km downstream of the 0.5 m drift with 1.5 TeV muon beam decaying
there.

![Figure 17](image)

Figure 17: Left: Dose in the orbit plane vs radial distance from the ring center of 2, 3 and 4 TeV
c.o.m. MC. Right: Dose downstream of a 0.5 m drift with 2.6 × 10^{16} decays/yr of a 1.5 TeV muon
beam there vs. distance in soil downstream of the drift. Red line shows annual off-site regulatory
limit at Fermilab, that is, 10 mrem/year = 0.1 mSv/year [193].

The first obvious way to reduce neutrino fluxes and mitigate radiological issues at large distances
from the collider ring is to place the ring deep underground. MARS-calculated depths D along
with distances R (see Fig. 16) are shown in Table 1 [193] for five c.o.m. energies from 0.5 to 4 TeV
and two annual off-site limits 1 mSv = 100 mrem and 0.1 mSv = 10 mrem (Fermilab), all for N
muon decays/year. Results here were obtained assuming suppressed contribution from the field-
free regions. There is also the regulatory question whether delivering an off-site dose above the
limit at any depth underground or height above it is permissible.

Note that simplified expressions derived in [8, 191] give noticeably more conservative results
compared to those from MARS full Monte Carlo [193]. For example, for the 3 TeV case, depth to
Table 1: Required depths \(D\), radial distances \(R\) and off-site dose limits for several scenarios of muon colliders [193]

<table>
<thead>
<tr>
<th>E.c.o.m. TeV</th>
<th>(N \times 10^{21})</th>
<th>0.5</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 mSv/yr</td>
<td>(R) (km)</td>
<td>0.4</td>
<td>1.1</td>
<td>6.5</td>
<td>12</td>
<td>18</td>
</tr>
<tr>
<td>1 mSv/yr</td>
<td>(D) (m)</td>
<td>(\leq 1)</td>
<td>(\leq 1)</td>
<td>3.3</td>
<td>11</td>
<td>25</td>
</tr>
<tr>
<td>0.1 mSv/yr</td>
<td>(R) (km)</td>
<td>1.2</td>
<td>3.2</td>
<td>21</td>
<td>37</td>
<td>57</td>
</tr>
<tr>
<td>0.1 mSv/yr</td>
<td>(D) (m)</td>
<td>(\leq 1)</td>
<td>(\leq 1)</td>
<td>34</td>
<td>107</td>
<td>254</td>
</tr>
<tr>
<td>0.01 mSv/yr</td>
<td>(D) (m)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>300</td>
</tr>
</tbody>
</table>

stay within 0.01 mSv/yr (1% of the DOE limit) is 300 m (see Table 1) compared to the analytical depth of 500 m [8].

The second, “exotic,” mitigation looked at in the early days of high-energy muon collider studies was to build it at an isolated site such as a desert or mountain region or remote island with no earth or rock immediately after a muon vector exits the Earth’s surface.

The third mitigation technique, investigated in MARS simulations [193], was to minimize the field-free regions in the collider (see Fig 17). It was found that the presence of a field of even a fraction of 1 tesla is enough to reduce the dose at large distances from an MC to a below-limit level. The application of such a field over all RF and other components seems possible. The MC straight sections could also be shortened by using continuous combined function magnets.

The fourth mitigation technique was proposed in [190] and studied in great detail in [193]. Fast beam wobbling by a systematic time-varying vertical wave field in the ring would drastically disperse the strongly-directed neutrino flux in the orbit plane outside the MC ring. The dose reduction is quite impressive as shown in Fig. 18 for the wave fields of 0.05, 0.1, 0.2 and 0.4 T for a 4 TeV c.o.m. energy MC. Without such a wave field, the radial distance in the orbit plane to reduce the neutrino induced dose to the offsite limit of 0.1 mSv/yr is 57 km, while a very modest wave field of 0.05 T reduces that distance to 14 km, and the effect rapidly increases with the field value and/or combination with other mitigation measures described earlier in this paper. A recently proposed alternative technique with large-stroke high-resolution magnet movers (being refined at CERN) has a high mitigation potential as well and aims to reduce the dose to levels comparable to those at the LHC.

The fifth way would be the reduction of muon beam current with better cooling—e.g., optical stochastic cooling might reduce the emittances by several orders of magnitude, thus greatly reducing the muon beam currents—and neutrino fluxes—while keeping the luminosity the same. The focusing strength could be further increased by the use of plasma or other exotic focusing methods at the IP.

### 4.3 European Accelerator Roadmap

In 2020, the European LDG formed a muon beam panel and charged it with delivering input to the European Accelerator R&D Roadmap covering the development and evaluation of a muon collider option. In parallel, CERN formed IMCC to assess feasibility of building a high energy muon collider, identify critical challenges, and develop an R&D program aimed at addressing them. Besides the accelerator, the effort includes development of the MDI, detector concepts, and an evaluation of the physics potential.

The IMCC held four "community meetings" in 2020 and 2021 to develop the scope and the plan of work to be done between now and the next ESPPU. R&D objectives have been identified in several key areas, including muon production and cooling, neutrino induced radiation mitigation, MDI studies and optimization, and the high energy complex. Technologically, the design imposes challenging requirements on the high power targets where short proton bunch length and frequency may compromise the target’s lifetime and integrity, on the high-field solenoidal magnets used in the production, collection and cooling of the muons, as well as on the specs of fast-ramping and fixed-field magnets used in the accelerator and collider rings. The ionization cooling system is a novel concept and requires careful studies for optimal integration of the absorber and RF stations inside of high magnetic fields. Successful demonstration of a partial muon cooling system is therefore crucial for the design verification. Currently rough dimensions of the facility have been identified and siting at CERN is being actively explored. An alternative siting at Fermilab is possible and is
Figure 18: Neutrino flux and dose reduction around the 4 TeV c.o.m. muon collider using beam wobbling induced by wave field of 0, 0.05, 0.1, 0.2 and 0.4 tesla. The red line is the Fermilab offsite annual limit which is reached at 57 km with no wave field and at 14 km with 0.4 T wave field.

A technically limited timeline developed by IMCC is shown in Fig. 19 and discussed in greater detail in [195]. A muon collider with a center-of-mass energy around 3 TeV could be delivered on a timeline compatible with the end of operation of the HL-LHC. The muon collider R&D program will consist of an initial phase followed by the conceptual and the technical design phases. The initial phase will establish the potential of the muon collider and the required R&D program for the subsequent phases.

The performance and cost of the facility would be established in detail. A program of test stands and prototyping of equipment would be performed over a five-year period, including a cooling cell prototype and the possibility of beam tests in a cooling demonstrator. This program is expected to be consistent with the development of high field solenoid and dipole magnets that could be exploited for both the final stages of cooling and the collider ring development. A technical design phase would follow in the early 2030s with a continuing program focusing on prototyping and preserves development before production for construction begins in the mid-2030s, to enable delivery of a 3 TeV collider by 2045. The program is flexible, in order to match the prioritization and timescales defined by the next ESPPU, P5, and equivalent processes.

Based on the MAP design, target parameter sets have been defined for the collider as a starting point. The parameter sets have a luminosity-to-beam-power ratio that increases with energy. They are based on using the same muon source for all energies and a limited degradation of transverse and longitudinal emittance with energy. The design of the technical components to achieve this goal are a key element of the study. It is important to emphasize that a 10 TeV lepton collider poses a number of key challenges:

- The collider can potentially produce a high neutrino flux that might lead to increased levels of radiation far from the collider. This would need to be mitigated and is a prime concern for the high-energy option.
- The Machine Detector Interface (MDI) might limit the physics reach due to beam-induced background, and the detector and machine need to be simultaneously optimized.
- The collider ring and the acceleration system that follows the muon cooling can limit the
energy reach. These systems have not been studied for 10 TeV or higher energy. The collider ring design impacts the neutrino flux and MDI.

- The production of a high-quality muon beam is required in order to achieve the desired luminosity. Optimization and improved integration are required to achieve the performance goal, while maintaining low power consumption and cost. The source performance also impacts the high-energy design.

4.4 R&D Priorities and possible US contributions

Below we list possible areas of research with emphasis on areas wherein the U.S. can contribute:

- **Proton driver**: Existing facilities can be upgraded to provide multi-MW beams for a muon collider or neutrino factory. Fermilab’s PIP-H program will be capable of delivering beam power up to 1.2 MW. Several proposals are under development for either expanding the superconducting proton linac (PIP-III) or combining the existing linac with an RCS to increase the beam power to > 2 MW. The ESS MW proton linear accelerator can also be upgraded and extended to demonstrate the generation of nanosecond-scale beams with very high-charge ($10^{15}$) proton pulses that can be used for the generation of the muon pulses required for a muon collider. The R&D program will explore these and other source options.

- **Target**: The design criteria of the target system are high pion/muon production yields and high tolerances of the thermal stress caused by the beam impact.
  - Extending the RaDIATE effort to study high stress tolerant material. Optimizing dimensions of the target system to produce high pion/muon yields as well as minimizing thermal stress on the target material by use of a new target material. Applying AI algorithms to find the best mix of target compounds to fulfill the design criteria.
  - Collaborating with the MagLab and other national institutions to develop a radiation robust high field magnet system for capturing pions/muons and transporting them to the cooling channel.
  - Utilizing the TSIB hot lab and other PIE facilities to develop target material science. Propose the needed beam irradiation facility: intense proton beam facilities will be
considered for the beam irradiation facility, such as PIP-II and the upgraded-Booster accelerator complex. These accelerators are also good candidates for the cooling demonstrator.

**Cooling:** Improving the cooling performance is a primary goal of the cooling design R&D. Depending upon the future target system, decay, bunching, and phase rotation (called the “front end”) systems, the subsequent 6D cooling channel must be optimized. Improving cooling can significantly relax the beam requirements, reducing the primary proton beam power, the beam induced background at the collider detector, and the neutrino flux. Research on integration of AI techniques can aid in making the channels shorter and perhaps identify new parameter sets for improved cooling. The FOFO Snake [196], which cools both muon charge signs simultaneously, should be further explored. Further simulation studies should be made of PIC [161, 197] to determine whether it can achieve acceptable performance or to rule out its utility. Further improvements should be considered in the high-field solenoid cooling channel to reach the final collider emittances, incorporating both more extensive simulation studies and expected improvements from more advanced magnet technologies.

- Investigating the space charge effect in the cooling channel, especially in the later cooling channel.
- Engineering efforts to integrate high gradient RF cavities into high field magnet coils. Designing modular beam components including beam instrumentation.
- Investigating the cooling performance with imperfect beam optics, especially in the later cooling channel.
- Investigating the influence of ionizing particles in the cooling channel, especially for the initial cooling channel.
- Investigating the beam-plasma interactions, especially the plasma focusing effect.

**Acceleration:** To better understand acceleration for a muon collider, a near-term research program should include:

- Creating a first design of acceleration from a final cooling stage to the initial energy of the MAP acceleration scenario, with goals of having a more general description of the acceleration scenario, an approximate idea of achievable muon transmission, and studying emittance preservation, particularly in the longitudinal plane.
- Designing a detailed pulsed acceleration lattice design that would fit on the Fermilab site. The goal would be to have sufficient detail in the design to be able to specify with reasonable confidence the maximum energy that could be achieved for a muon collider on the Fermilab site given certain technology choices.
- Study FFA acceleration designs, both for moderate and maximum energies, with particular emphasis on advanced designs (such as nonlinear and vertical FFAs), to understand if and for what parameter ranges they may be more favorable than other types of accelerators.
- A study of the collective effects in acceleration systems, to assess their impact on the beam distribution, and identify any required mitigations and their impacts.
- Designing and testing an approximately full-scale pulsed magnet similar to a magnet that would be used for pulsed muon acceleration, along with its power supply. It should reach fields required for muon acceleration (at least 1.5 T) and ramp at the desired rate (around 1 ms). The goals would be to confirm the expected magnet performance, our ability to control the field variation with time, and to understand the engineering and cost issues for the system, particularly the power supply.

**Collider Ring:** The TF and EF groups can investigate the physics cases at 600 GeV, 3 TeV, and 10 TeV center of mass energy. A new collider lattice must be designed. Possible solutions to mitigate neutrino flux are to situate the collider at ~ 100 m depth, add magnets, or move the lattice over time. Improving cooling can significantly relax these requirements.

**Magnets:** The following key MC magnet systems will require focused R&D efforts:
Muon cooling. MAP considered muon beam cooling with magnetic fields up to 30 T. This field level has been demonstrated with commercial 29 T MRI magnets. The record field of 32 T achieved in a superconducting solenoid with bore diameter similar to that needed for cooling allows extending its design and parameters to the MC level. Experimental demonstration of the final cooling solenoid will be needed.

Muon acceleration. Recent tests of the HTS-based 0.5-m long two-aperture SC dipole at Fermilab have shown record-high field ramp rate of 300 T/s at 10 Hz repetition rate and 0.5 T field amplitude. Based on this result, a possible upgrade of this magnet design to MC requirements of 2 T amplitude in a 10 mm beam gap with dB/dt up to 1000 T/s for the rapid cycling acceleration dipoles has been proposed and needs to be supported.

Collider ring. The collider ring for 6 TeV or higher c.o.m. energy MC requires 24 cm aperture, 20 T nominal field arc dipoles with large operating margin in a high radiation environment. The US-MDP is working on designs and technologies of accelerator magnets based on low temperature and high temperature superconductors. In 4–5 years the program plans experimental demonstration of a 120-mm aperture, 10–12 T Nb$_3$Sn dipole, and development of conceptual and engineering designs of a 20 T hybrid HTS/Nb$_3$Sn dipole with 50-mm aperture, which will be fabricated and tested. Combining these two results would pave the way towards the large-aperture (>150 mm diameter) high-field (> 20 T with large margin) dipoles and quadrupoles needed for the HE MC.

**RF cavity:** Significant achievements have been made on the RF cavities in magnetic field for muon ionization cooling under MAP. However, there is still a considerable amount of R&D needed to make and test realistic cavities for the current cooling channel designs, which includes:

- Design a compact multiple-cavity module, with practical frequency tuning and RF power feed systems. In the current cooling channel design, the voltage in each segment requires multiple cavities or one linac with multiple cells. To achieve the required strong solenoidal fields, the cavity transverse dimension should be kept as compact as possible. The RF tuning mechanism and RF power coupling feed must fit into the tight space, and more importantly be immune to the strong magnetic field background. Novel ideas, such as distributed coupling structure [198], should be explored and studied. The RF module design is an integral part of the cooling channel, and comprehensive system engineering efforts are required to include all accessory components, such as RF coupler, RF feed-through probe, cryogenic system and diagnostics for the cooling channel during the early design phase.

- Study of the beryllium windows for RF cavities at various cooling stages: The thickness of beryllium windows is determined by the acceptable RF frequency tuning range due to window deformation from the thermal stress induced by RF heating and scattering effects. The Be window should be kept as thin as possible, and especially becomes extremely challenging at the later stage of the cooling channel where the scattering effect from beryllium may become one of the dominant heating effects on muon emittance.

- Explore RF cavities at other frequencies: RF breakdown scales with frequency, as expressed by the empirical Kilpatrick limit; optimization study of cavity frequency choice, in consideration of RF breakdown limit (performance), cavity dimensions, and surrounding magnets may provide additional benefit for a compact cooling channel design.

- Cavity operation at cryogenic (LN$_2$) temperature: A recent experiment shows that a copper cavity operated at cryogenic temperature is more resilient to RF breakdown than at room temperature [199]. Although the experiment was carried out without a strong external magnetic field, it is expected that it will withstand higher RF breakdown even in the presence of a strong magnetic field. In addition, both copper and beryllium have a factor 2–3 improvement in electrical conductivity at LN$_2$ temperature, a significant and potential peak RF power savings for a given RF gradient.

- Investigation of beam loading and wakefield effects in vacuum RF cavities: Theoretical and experimental study to include collective beam loading and wakefield effects and
mitigation methods and implementation of higher-order-mode (HOM) damping schemes if needed.

- Further investigation of the gas–plasma–beam interactions in gas-filled RF cavities: Study the plasma chemistry and learn how to control the cold plasma dynamics. In order to take advantage of the extra plasma focusing effects, a new type of plasma focusing experiment should be conducted by using intense beams, and simulated using particle-in-cell codes.

### 4.5 Fermilab Site Option

The idea of having a MC as a potential “site filler” for Fermilab dates back to the early 2000’s, when parameters for a 4 TeV machine were presented. More recently, using higher field magnets and higher-gradient acceleration, the parameter space towards a 10 TeV Muon Collider concept that would fit within the Fermilab site has been identified and a first design concept has been developed. A schematic layout of this configuration is shown in Figure 20. The concept begins with use of PIP-II as the initial part of the proton source. The PIP-II linac would be extended to
higher energy and followed by either a higher energy linac leading into proton accumulation and bunching rings or a rapid-cycling synchrotron or FFA (fixed-field accelerator) ring. The goal would be to produce intense $\approx 10$ GeV proton pulses at $\sim 5$ Hz and $\sim 2$ MW onto a pion production target. This is followed by muon collection (from $\pi$ decay) and bunching that leads into 6D muon cooling channels, obtaining minimal emittance beams. The collection and cooling channel would be $\sim 1$–$2$ km long.

Muon acceleration is achieved in three stages: (1) A Linac (up to 5 GeV) first that is followed by a Recirculating Linac (up to 65 GeV). This energy would be sufficient for a Higgs Factory [200]. (2) This is followed by a set of two Rapid Cycling Synchrotrons that can fit into the Tevatron ring tunnel and are capable of delivering an energy up to 1 TeV. (The first RCS would accelerate to $\sim 300$ GeV, using normal-conducting magnets. The second would be a hybrid high-field RCS.) (3) A final RCS ring that has a radius of 2.65 km (site filler) and can bring the energy up to 5 TeV. (This is a hybrid RCS ring with $\sim 16$ T magnets interlaced with cycling $\pm 4$ T magnets. If cycling is limited to $\pm 2$ T, two rings are required.) The acceleration will use superconducting RF cavities at frequencies of 650 MHz and 1300 MHz.

The 5 TeV beams would be injected into a 10 TeV collider ring using high field bending magnets. Based on extrapolations from Ref. [201] the 10 TeV collider is expected to have a radius of 1.65 km.

It is important to note that, given the 3 accelerator stages, staging is possible and operation at 125 GeV, 1 TeV, and 3 TeV can be envisioned as intermediate states. Figure 20 shows a schematic view of the collider in its various stages.

5 Detectors

5.1 General Introduction

A circular $\mu^+\mu^-$ collider is a particularly attractive option for the future of energy frontier exploration. Such a machine has the potential to deliver a vast physics program in a relatively compact and power efficient accelerator complex. However, on the experimental side, a great physics potential is accompanied by unprecedented technological challenges, due to the fact that muons are unstable particles. Their decay products interact with the machine elements and produce an intense flux of background particles that eventually reach the detector and may significantly degrade its performance.

The physics program includes a precision component consisting of the exploration of the electroweak sector of the SM and a broad spectrum of searches for BSM physics. In order to achieve these physics goals, experiments at a Muon Collider need to be able to reconstruct products of $\mu^+\mu^-$ collisions with required performance. The performance demands identification and reconstruction of charged leptons, photons, and jets with high efficiency and good energy/momentum resolution. This in turn places stringent requirements on the performance of tracking and calorimeter reconstruction. Identification of displaced vertices originating in decays of heavy flavour mesons is also of major interest, in particular for the Higgs program and for BSM signatures involving heavy flavor.

The unstable nature of muons (lifetime $\tau_\mu \approx 2 \mu$s at rest) results in a significant fraction of muons to decay, with their decay products producing a large flux of secondary and tertiary particles after interacting with accelerator and detector elements. Such a flux of particles is referred to as Beam-induced background (BIB). This results in a unique experimental environment [183], with a related set of challenges. One of the biggest challenges for a muon collider detector is to successfully disentangle the products of the $\mu^+\mu^-$ collisions from an intense BIB coming primarily from the muon decay and shower products. The detectors and event reconstruction techniques need to be designed to cope with the presence of the BIB. In this sense, considerations and challenges related to the detector design at a muon collider are somewhat distinct from those at $pp$ and $e^+e^-$ colliders.

The goal of this section is to demonstrate that high quality physics is achievable despite challenges originating from the BIB. This is possible due to significant advancements in detector technologies during the last decade, motivated in large by requirements for the HL-LHC experimental upgrades. A comprehensive review of promising technologies for a muon collider can be found in Ref. [202]. Major breakthroughs have also been made in development of dedicated BIB suppression and physics object reconstruction techniques. This work has been carried out within the context of IMCC and Snowmass studies and is summarized in Ref. [203].

Here we first discuss the environment in which the detectors are expected to operate and
draw some comparisons with High Luminosity LHC. We present the current simulated detector
configuration and highlight the expected performance. We also briefly describe technologies that
have a potential to match challenging specifications of a muon collider detector. The current
detector configuration was adopted from CLIC and is optimized for 3 TeV center-of-mass energy;
the design needs to be updated for future higher energy studies. We then present two very distinct
physics studies (H \rightarrow bb and Dark Matter with disappearing track) and compare results at the full
and fast simulation levels. The goal of these comparisons is to demonstrate that impact of the BIB
can be mitigated and the residual differences between full- and fast- simulation are small. Finally,
we outline a path forward for future improvements.

5.2 Environment

The expected characteristics of BIB depend on the beam properties, accelerator lattice, interaction
region as well as detector design. Detailed simulation studies have been performed \cite{204} using the
MARS15 \cite{205} software and, more recently \cite{206, 207, 202}, using a combination of Linebuilder \cite{208}
and FLUKA \cite{209, 210}. The two simulations are found to give compatible results. Muons decays
are simulated within \pm 200 m from the interaction point (IP) in a collider ring with the parameters
summarised in Table 2. In particular, these studies were performed for a collider with center-of-
mass energy of 1.5 TeV. The expected distance between bunches is such that the effect of nearby
bunches is negligible.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center-of-mass energy</td>
<td>\sqrt{s}</td>
<td>1.5 TeV</td>
</tr>
<tr>
<td>Muons per bunch</td>
<td>N_µ</td>
<td>2 \cdot 10^{12}</td>
</tr>
<tr>
<td>Normalised transverse emittance</td>
<td>\epsilon_{TN}</td>
<td>25 \pi \mu m rad</td>
</tr>
<tr>
<td>Normalised longitudinal emittance</td>
<td>\epsilon_{LN}</td>
<td>7.5 MeV m</td>
</tr>
<tr>
<td>IP relative energy spread</td>
<td>\delta_E</td>
<td>0.1 %</td>
</tr>
<tr>
<td>IP beta function</td>
<td>\beta_{x,y}</td>
<td>1 cm</td>
</tr>
<tr>
<td>IP transverse beam size</td>
<td>\sigma_{x,y}</td>
<td>6 \mu m</td>
</tr>
<tr>
<td>IP longitudinal beam size</td>
<td>\sigma_z</td>
<td>10 mm</td>
</tr>
</tbody>
</table>

Table 2: Representative set of muon collider parameters used in the detailed simulation presented
in [207].

Particles from BIB can deposit a huge amount of energy in the detector, if not shielded properly.
For this reason an essential part of the machine detector interface at a Muon Collider is a pair of
tungsten (W) nozzles cladded with borated polyethylene (BCH), which reduce the rate and energy
of BIB particles reaching the detector by several orders of magnitude. Such a nozzle also limits the
acceptance of the detector to polar angles \theta > 10^\circ. The flux of particles surviving the shielding and
entering the detector arise partially through shower products of BIB particles exiting the nozzle
and through back-scattering of particles from one beam into the nozzle on the opposite side of the
IP with respect to the direction the beam is arriving from. The result is a diffuse background of
mostly low-momentum and out-of-time photons, neutrons, and electrons/positrons; as shown in
Figure 21, their flux reduces by 2 to 5 orders of magnitude for energies above 100 MeV and the
expected time structure extends to several hundreds of ns. A smaller flux of muons and charged
hadrons is expected but their flux is smaller by several orders of magnitude.

The amount of expected radiation from BIB has been estimated using the FLUKA simulation
mentioned above, using a simplified detector geometry. Detailed maps are available in Ref. \cite{202}
for both expected dose and 1 MeV neutron equivalent fluence. Table 3 provides an approximate
comparison of the expected doses and fluences in the muon collider environment when compared
to high-radiation environments of high-energy colliders such as HL-LHC. The expected HL-LHC
doses and fluences are taken from Ref. \cite{211}; such calculations include a very detailed simulation
of detector elements and a safety factor that is not included in the other numbers of this table,
but should not change qualitatively the comparison. We also re-normalized the expected dose and
fluence to a single year of data-taking at the ultimate performance of the HL-LHC accelerator.

The flux of BIB particles can be further reduced exploiting the different timing, direction
and energy spectrum compared to the products of the main \mu^+\mu^- collisions. The effect of such
requirements is explored in Ref. \cite{203, 202, 212} and summarized in the following Sections 5.3, 5.4.
**Figure 21:** Kinematic properties of BIB particles entering the detector region: momentum (left) and arrival time with respect to the bunch crossing (right).

<table>
<thead>
<tr>
<th></th>
<th>Muon Collider</th>
<th>Maximum Dose (Mrad)</th>
<th>Maximum Fluence (1 MeV-neq/cm$^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R = 22 mm</td>
<td>10</td>
<td>$10^{15}$</td>
<td>$10^{14}$</td>
</tr>
<tr>
<td>R = 1500 mm</td>
<td>100</td>
<td>$10^{15}$</td>
<td>$10^{13}$</td>
</tr>
</tbody>
</table>

Table 3: Rough comparison of expected dose and 1 MeV neutron equivalent per cm$^2$ fluence per year of data-taking for a muon collider environment as well as for HL-LHC. The numbers are reported for different radii (R), loosely corresponding to possible positions for a first tracking layer, and before the entrance of an electromagnetic calorimeter. See text for details and caveats.

FLUKA simulations of higher center-of-mass energy muon collider configurations is still under development. A very preliminary comparison of the multiplicity of particles entering the detector volume using FLUKA and MARS15 and for beam energies up to 5.0 TeV (corresponding to 10 TeV COM energy) is shown in Fig. 22. The numbers are obtained with MDI optimized for 750 GeV beam energy. These and further preliminary results reported in Ref. [207] show energy evolution of two main effects: muon lifetime becomes longer in the laboratory frame, leading to fewer decays per meter, while the average initial energy of muon decay products is larger, leading to a higher multiplicity of particles generated by shower in material. It turns out that the two effects roughly cancel out resulting in similar multiplicity, spatial, and time distributions of BIB particles entering the detector. Optimization of MDI at 3 and 10 TeV should results in further reduction of the BIB.

**5.3 Current Configuration**

In order to benchmark the realistic physics expectations of a muon collider, a full GEANT model of a detector based on the CLIC detector has been developed. The detector model was adapted from the post-CDR CLIC detector model, CLICdet [213]. It should be of course noted, that since the time scale for the R&D and construction of a muon collider facility requires several decades it should be assumed that detector technology and performance will naturally evolve during that time period and advancements in detector technology are almost guaranteed to improve the performance that a real detector could obtain in the future. Nonetheless, in order to make progress on understanding the physics potential and background challenges of such a machine a detector with minimal requirements on technology R&D has been simulated in detail [203]. The rendering of the detector geometry is presented in Fig. 23. The detector model uses the basic cylindrical layout of an electron collider detector but makes important modifications to the machine detector interface (MDI) and tracking detectors to shield and optimize the layout for the large beam induced backgrounds.

The detector is comprised of the typical cylindrical layout of a collider detector with a silicon inner tracker, electromagnetic and hadronic calorimeters enclosed by a strong solenoid magnet.
Figure 22: Comparison of multiplicities of different types of particles produced in each bunch crossing by the BIB at different energies. The uncertainties are about 20%. The MDI optimized for 750 GeV is used for the 1.5 and 5.0 TeV columns. Vast majority of these particles have very low momenta.

<table>
<thead>
<tr>
<th>beam energy [GeV]</th>
<th>MARS15</th>
<th>MARS15</th>
<th>FLUKA</th>
<th>FLUKA</th>
<th>FLUKA</th>
</tr>
</thead>
<tbody>
<tr>
<td>μ decay length [m]</td>
<td>3.9 x 10³</td>
<td>46.7 x 10⁵</td>
<td>46.7 x 10⁵</td>
<td>93.5 x 10⁵</td>
<td>311.7 x 10⁵</td>
</tr>
<tr>
<td>μ decays/m per beam (for 2x10⁻¹² μbunch)</td>
<td>51.3 x 10³</td>
<td>4.3 x 10⁵</td>
<td>4.3 x 10⁵</td>
<td>2.1 x 10⁵</td>
<td>0.84 x 10⁵</td>
</tr>
<tr>
<td>photons/BX (Eγ &gt; 0.1 MeV)</td>
<td>170 x 10⁵</td>
<td>86 x 10⁵</td>
<td>51 x 10⁵</td>
<td>70 x 10⁵</td>
<td>116 x 10⁵</td>
</tr>
<tr>
<td>neutrons/BX (Eγ &gt; 1 meV)</td>
<td>65 x 10⁵</td>
<td>76 x 10⁵</td>
<td>110 x 10⁵</td>
<td>91 x 10⁵</td>
<td>89 x 10⁵</td>
</tr>
<tr>
<td>e⁻/BX (Ee⁻ &gt; 0.1 MeV)</td>
<td>1.3 x 10⁹</td>
<td>0.75 x 10⁸</td>
<td>0.86 x 10⁸</td>
<td>1.1 x 10⁸</td>
<td>0.95 x 10⁸</td>
</tr>
<tr>
<td>charged hadrons/BX (Eγ &gt; 0.1 MeV)</td>
<td>0.011 x 10⁶</td>
<td>0.032 x 10⁶</td>
<td>0.017 x 10⁶</td>
<td>0.020 x 10⁶</td>
<td>0.034 x 10⁶</td>
</tr>
<tr>
<td>muons/BX (Eγ &gt; 0.1 MeV)</td>
<td>0.0012 x 10⁶</td>
<td>0.0015 x 10⁶</td>
<td>0.0031 x 10⁵</td>
<td>0.0033 x 10⁵</td>
<td>0.0030 x 10⁶</td>
</tr>
</tbody>
</table>

Figure 23: Rendering of the muon collider detector geometry used for the presented simulation studies, including the cone-shaped shielding nozzles (cyan) and the beryllium beampipe (violet). Shown are the R-Z cross sections of the full detector geometry (left) and two zoomed-in portions: up to ECAL (top right) and up to Vertex Detector (bottom right). Muon Detector (violet and green) surrounds the solenoid (cyan), which encloses the HCAL (magenta), ECAL (yellow) and the Tracking Detector (green and black).
Table 4: Rough comparison of expected hit density at a muon collider detector compared to ATLAS ITk in HL-LHC. The numbers are reported for different radii, corresponding to possible positions for the first and second layers of a pixelated vertex detector and first layer of the silicon strip detector.

<table>
<thead>
<tr>
<th>Detector Layer</th>
<th>ITk Hit Density [mm$^{-2}$]</th>
<th>Muon Col. Hit Density [mm$^{-2}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pixel Layer 0</td>
<td>0.643</td>
<td>3.68</td>
</tr>
<tr>
<td>Pixel Layer 1</td>
<td>0.22</td>
<td>0.51</td>
</tr>
<tr>
<td>Strip Layer 1</td>
<td>0.003</td>
<td>0.03</td>
</tr>
</tbody>
</table>

with an iron return yoke with muon chambers interleaved. The detector features an all-silicon Tracking Detector with three subsystems: a pixel Vertex Detector from 30 mm to 104 mm radius, a macropixel Inner Tracker from 127 mm to 550 mm in radius, and a microstrip detector from 819 mm to 1486 mm. The spatial resolution is presumed to be 5µm by 5µm for the Vertex Detector and 7µm by 90µm for the Trackers with timing resolutions of 30 (60) ps for the Vertex (Tracking) Detectors. Such requirements are not achievable by current technology, but are the subject of intense R&D and are thought to be achievable in the relatively near future, as outlined in Section 5.4. Immediately after the Outer Tracker is a silicon-tungsten (SiW) ECAL made of 40 layers of 1.9 mm thick absorber plates with 5 × 5 mm$^2$ silicon sensor pads extending to a radius of about 1700 mm. This is followed by the HCAL comprised of 60 layers of iron and plastic scintillating tiles with a total of 7.5 interaction lengths. A 3.57 T solenoid magnet surrounds the previous elements comprised of 3-module 4 layer aluminum coil inside a liquid helium cryostat. The final element are 6 layers of resistive plate chambers (RPCs) inside a large iron return yoke which extends out to 6.5 meters in radius. Design details including the placement of various components, exact material composition, and estimate of the timing and position resolution are given in [203] and [213].

The most distinctive addition to the CLIC design is modification of the MDI to account and shield the large beam induced backgrounds. The MDI includes the placement of two large tungsten nozzle shields. These are cones of tungsten originating at the center of the detector and limit the forward coverage. In order to improve the shielding borated polyethylene is included in the outer portion of the cone in the inner detector and cylinders of iron, concrete and borated polyethylene surround the beam pipe in the MDI approaching from both sides of the detector with layout and dimensions as shown in [203].

5.4 Feasibility Statement

Since the original proposals of colliding muons to study physics at high center of mass energies more than fifty years ago, there have been multiple advances in detector technologies that represent a quantum leap forward in experimentalists’ ability to perform precision measurements in high occupancy environments.

In addition, there has been a technological revolution in high throughput data processing, which has been readily integrated into modern particle physics experiments, namely, Graphical Processing Units (GPUs) and Field Programmable Gate Arrays (FPGAs).

5.4.1 Tracker

Closest to the beamline, the tracker suffers from the highest density of BIB. At the track level, the BIB is easy to separate from collision products: it is predominantly low-energy, and does not originate at the interaction point. However, the number of hits produced by the BIB is many orders of magnitude larger than that of the collision products, which creates a challenge for tracker technology, requiring high granularity silicon detectors to avoid saturation. Figure 24 shows the hit density per bunch crossing throughout the tracker which reaches 1000 hits/cm$^2$ in the vertex detector. A comparison of the expected hit density (after applying loose timing selection) with ATLAS ITk detector is shown in Table 4. One can see that the density is approximately an order of magnitude larger than in ATLAS in HL-LHC.

Over the past decade several transformative new developments in silicon tracking detector technology have been proposed. The introduction of Low Gain Avalanche Diodes (LGAD) and similar devices with intrinsic amplification have improved particle tracking time resolutions by a factor...
of ten, while AC coupled LGADs achieve few-micron position resolution [215, 216, 217]. CMOS-based sensors have the potential to lower cost/cm² by a factor of 5-10 with excellent resolution. In addition, the ability to integrate small pixels with 3D integration or double-sided LGADs can add angular information to each data point. Finally, low power radiation hard extensions of optical transceivers and advanced power delivery systems developed for the HL-LHC will allow for a low mass radiation-hard tracking system.

These capabilities are enabling for tracking in the high beam-induced background environment of the Muon Collider. Much of the background is low energy, out of time, and non-pointing. These emerging technologies can reduce the majority of these backgrounds at the sensor level by factors of 10-100 by requiring in-time hits and pixel cluster patterns and energies consistent with tracks emerging from the primary vertex. This would reduce the data load and associated power and mass, simplify triggering, and possibly allow for full tracker event readout every crossing.

Various handles to reduce the BIB can be explored for both on- and off-detector filtering. Possible filtering schemes are described in [202] and quickly summarized here:

- **Timing**: Removing out of time hits appears to reduce the data load by a factor of 3. Timing information will eventually be needed in the reconstruction, but it makes sense to apply initial filtering on-detector. The time spread expected from the finite size of the interaction region is approximately 20–30 ps.

- **Clustering**: Clustering reduces the number of pixel groups read out. This requires more on-detector processing and results in more bits per cluster and a higher power budget, but can reduce the number of hits read out. Selection requirements can also be applied to the cluster shape. The effectiveness needs to be assessed for each BIB cluster type.

- **Energy Deposition**: Each of the backgrounds has a characteristic energy deposition signature. For example neutrons have low, localized energy deposit. This should be a useful requirement to make on-detector.

- **Correlation Between Layers**: This is a powerful handle for background rejection. However, implementation may be complex and costly, doubling the number of channels. For on-detector filtering, it also requires transfer of data between layers in a very busy environment.

Figure 24: Average hit density per bunch crossing in the tracker as a function of the detector layer [214].
**Figure 25**: Performance of a BIB-reduction algorithm at the hit level of the vertex detector, after a fiducial timing window selection of [-250, +300] ps. An ideal resolution case is shown in black, while the other colors represent a range of timing and angular resolutions. Different hue ribbons represent different detector capabilities, while the surface of the ribbon represents different choices for cut values [212].

- **Local Track Angle**: Track angle measurement can be made in a single detector if the thickness/pitch ratio distributes the signal over several pixels. This avoids the complexity of inter-detector connections and could provide a monolithic solution [218, 219].

- **Pulse Shape**: Signals from BIB can come with a variety of angles and may not give the deposit profile and pulse shape of a typical MIP. Appropriate pulse processing, such as multiple sampling, RC-CR filters, zero crossing, or delay line clipping can be used to further reduce the data load.

The basic trade-offs are between the complexity, power, and mass needed to implement a on-detector filter, and the benefit of reduced data rate. Currently, the simulated tracker relies primarily on using precision timing information to suppress the BIB. Studies of the other handles outlined above are ongoing and demonstrate good potential for further improvements.

It is evident that developments in 4D tracking are an essential tool for reducing the BIB. New detector technologies, such as the LGADs being used in the ATLAS and CMS Phase-2 upgrades, can provide 20-30 ps per hit timing resolution. Simple timing window requirements that take advantage of this tens of picosecond timing can reduce the background by a factor of three or more [214].

The double-layered tracker strategy currently being employed by the CMS Experiment for the HL-LHC [220] provides extra handles for BIB reduction. Double-layer correlation significantly reduces backgrounds, and can be used to make measurements of local track angles. These measurements can be used to reject particles emanating from the BIB-blocking tungsten nozzles.

Figure 25 shows potential hit-level BIB reduction from a combination of angular and timing variables designed to isolate collision products. On top of a fiducial timing window that already substantially reduces this background, these requirements have the potential to eliminate another order of magnitude of BIB hits in the vertex detector, before entering the tracking stage, substantially reducing the combinatoric problem introduced by the BIB. More complex analysis of pulse shapes and hit cluster characteristics could potentially reduce this background even further.
5.4.2 Calorimetry

A calorimeter design for a muon collider must be able to disentangle photons and neutrons produced by beam muon decays upstream of the interaction point (IP) from prompt, high-energy electromagnetic and hadronic showers produced by the $\mu^+\mu^-$ collision at the IP. The beam-induced background (BIB) consists primarily of soft photons ($p_\gamma \approx 2 \text{ MeV}$) and moderately energetic neutrons ($p_n \approx 500 \text{ MeV}$) [221]. The occupancy of the showers produced by these particles in the calorimeter is high. A large fraction of the BIB can be rejected with a short readout window, of order a few ns or less, around the nominal beam crossing. This requires fast signal formation in the front ends. With high lateral granularity, advanced pattern recognition may be used to reject the non-pointing background of low-energy BIB photon and neutron showers. Finally, the calorimeter active materials and front end electronics need to withstand the 1-MeV-equivalent neutron fluence (a bit less than expected at HL-LHC) and total ionizing dose (worse than expected at HL-LHC) of this BIB.

The amount of ambient diffuse energy in the calorimeters is a quantity that provides a measure of energy deposited by the beam background that needs to be corrected for during reconstruction. To estimate this quantity we deployed a grid-median background estimation technique that divides the calorimeter into an arbitrary grid and sums up the energy within the grid. We find that the average ECAL+HCAL ambient energy is 50 GeV per unit area when integrated over the entire detector rapidity region. This number is similar to approximately 40 GeV at HL-LHC.

In the last decade, advances have been made in the areas of precision timing and high granularity that pave the way towards a workable muon collider calorimeter design. Cell sizes as small as 13 mm$^2$ have been realized in prototype silicon sensors for the SiD detector, one of two validated detector designs for the ILC. These prototypes have been exposed to cosmic rays and test beam electrons and shown to function properly [222]. Similar sensors are being used for the HGCal endcap calorimeter upgrade of the CMS detector for the HL-LHC, but with cell sizes of $\approx 0.5$ and $\approx 1.1$ cm$^2$. Prototyping of these sensors is well advanced, with positron test beams indicating that an EM energy resolution of $22\%/\sqrt{E\text{[GeV]}} \oplus 0.6\%$ is achievable [223]. This is not as good as homogeneous crystal calorimeters, but demonstrates reasonable performance with a particle flow technology that is optimized for jet energy resolution and BIB rejection.

For hadronic calorimetry, compact arrays of scintillator or crystal tiles glued directly over PCBs holding silicon photomultipliers (SiPMs) for readout have been shown to be a scalable, cost-effective active material choice. The CALICE Collaboration, whose mission is to develop particle flow calorimetry technologies for any of the linear collider options, has already demonstrated in the early 2000s that hadronic energy resolutions of $\approx 60\%/\sqrt{E\text{[GeV]}}$ are possible with the so-called “SiPM-on-tile” technology [224]. A 2018 pion test beam exercised the feasibility of constructing and operating a large-scale hadronic calorimeter slice—the 72 cm $\times$ 72 cm $\times$ 38-layer prototype alone consisted of 22k $3\times3$ cm$^2$ channels [225]. Even smaller cell sizes of $\approx 1$ cm$^2$ with binary readout have been realized using resistive plate chambers [226].

Important advances have also been made in precision timing and high-speed front ends. As mentioned above, silicon low-gain avalanche detectors (LGADs) have achieved 30 ps timing resolution for minimum-ionizing particles in test beams. Discrimination of pulses arriving 5 ns apart and single-cell time resolution of 20 ps has been demonstrated in PbF$_2$ crystals with SiPM readout for the Muon g-2 electromagnetic calorimeter [227]. Many scintillators with ultra-fast decay times of less than 1-10 ns, for example BaF$_2$, have been identified and characterized for light yield, timing performance, and radiation hardness [228]. Candidate inorganic scintillators, both in crystal and ceramic form, that survive up to 1-100 Mrad with light loss no worse than 70% have been identified [229].

In [214], four key features are identified that will enable good energy measurements at a muon collider experiment:

- **High granularity** to reduce the overlap of BIB particles in the same calorimeter cell. The overlap can produce hits with an energy similar to the signal, making harder to distinguish it from the BIB;

- **Good timing** to reduce the out-of-time component of the BIB. An acquisition time window of about $\Delta t = 300$ ps could be applied to remove most of the BIB, while preserving most of the signal. This means that a time resolution in the order of $\sigma_t = 100$ ps (from $\Delta t \approx 3\sigma_t$) should be achieved. The CMS BTL based on LYSO:Ce bars with SiPM readout on each end integrated a precision 30 ps timing layer with a high resolution crystal calorimeter [230];
• **Longitudinal segmentation:** the energy profile in the longitudinal direction is different between the signal and the BIB, hence a segmentation of the calorimeter can help in distinguishing the signal showers from the fake showers produced by the BIB;

• **Good energy resolution** of \( \frac{105}{\sqrt{E}} \) in the ECAL system for photons and a jet energy resolution of \( \frac{37}{E} \) are expected to be enough to obtain good physics performance. These have been demonstrated for conceptual particle flow calorimeters.

Stemming from the ILC R&D program, two major approaches are being studied to exploit sampling calorimeters and improve upon the current generation of collider experiments: multi-readout (dual or triple) [231, 232] and particle flow [233] calorimetry. The first approach focuses on reducing the fluctuations in the hadronic shower reconstruction, which are the main responsible for the deterioration in the determination of the jet energy. This goal is achieved by measuring independently the electromagnetic and the non-electromagnetic components of a hadronic shower, thus allowing to correct event-by-event for the different response of the calorimeter to various particle species. Notable examples of of the Dual Readout calorimeters can be found in the work of the DREAM/RD52 collaboration [234] and the proposed IDEA [235, 236] detector for FCCee/CepC.

The second approach focuses on the reconstruction of the four-momenta of every particle recorded by the detector. This method exploits tracking information and requires a detector with extreme granularity, combined with powerful reconstruction algorithms aimed at resolving each particle’s trajectory through the whole detector. Silicon sensors can be used as active elements to achieve a high channel granularity and longitudinal segmentation. State-of-the-art silicon sensors can sustain the high radiation dose of the expected BIB. Analogous technologies are being adopted by CMS HGCal [237] and considered by the CLIC collaboration. HGCal implements a precise timing measurement in these sensors (<100 ps), making the approach usable at a muon collider.

An alternative hybrid approach implements dual-readout methods with a segmented crystal ECAL and fiber HCAL to further improve the particle flow performance of jets with intrinsically high resolution calorimetry. The fine traverse granularity, longitudinal segmentation and precision timing provide additional handles to suppress BIB, increase particle identification and tracking matching with individual calorimeter hits [238]. Further development toward test beam evaluations is being organized as part of the CalVision Detector R&D [239].

Comparisons of performance of different calorimeter technologies in the muon collider environment have not yet been done. A comparison for FCC-ee can be found in Ref. [240].

### 5.4.3 Readout and Computation

With a single bunch collider operation scheme, beam crossing frequency is defined by the beam energy and the size of the collider ring. Collisions are expected to happen at the maximum rate of 100 kHz, corresponding to the minimum time between crossings of 10 µs.

The intense beam induced backgrounds that will be seen at a Muon Collider will present real computational challenges. Recent advances in real-time particle reconstruction techniques for reducing the amount of saved data for offline processing. At any experiment where the bulk of events contain interesting signal, using a trigger system to down-select events is undesirable. In order to reduce the data size two methods could be employed: implementation of selections based on timing windows matching particles to the primary vertex and processing detector signals in real-time and saving only a subset of the high-level reconstructed event information computed in the final software trigger stage is recorded and sent to permanent storage. The former is already implemented in many detectors where signals are only stored within a window of the primary interaction. The latter requires real-time reconstruction and has been employed at LHCB since Run-2.

Initial loose data filtering can happen on-detector. The basic trade-offs are between the complexity, power, and mass needed to implement a on-detector filter, and the benefit of reduced data rate. Overly aggressive front-end filtering schemes can introduce irrecoverable inefficiencies and biases in the data and limit acceptance for certain BSM physics signatures, such as these of long-lived particles.

Preliminary estimates of data rates based on simulation [202] indicate that a streaming DAQ architecture can provide an attractive solution for a future Muon Collider experiment. With improvements to the tracking speed, such a solution can likely be realized with technologies available today. Future advancements (e.g. higher speed optical links, fast processors, etc) are likely to
result in a smaller and/or more performing DAQ system. Work should be invested in improving HLT reconstruction algorithms and exploiting hardware acceleration schemes with the aim to bring per-event processing time down to a few second level.

5.5 Simulated Performance

A set of reconstruction algorithms were developed and tuned as described in detail in Ref. [203]. As an example, Figure 26 shows that charged particles can be successfully reconstructed in the tracker system with momentum resolutions comparable or better than the ones expected for HL-LHC. The developed algorithms also show that running a full track reconstruction in such a busy environment is feasible in a reasonable amount of time (minutes) without huge optimization efforts, proving that the large combinatorics will not be a significant problem in the tracker of a muon collider detector that satisfy the requirements outlined above. It should be noted that time information was used merely in measurement pre-selection and it is not actively used in the track reconstruction at this stage; using that information in track reconstruction (as noted in Section 5.4) is expected to further improve performance.

Figure 27 (left) shows the output of a simple secondary-vertex reconstruction algorithm, proving that separation between b, charm and light quarks is achievable. The performance of such a simple secondary-vertex tagger is comparable to the one of analogous algorithms in use at the LHC in terms of efficiency and rejection for each flavor.

Jet reconstruction is also affected by unique challenges due to the diffuse BIB component. While on a calorimeter-level this is analogous to a very high level of pile-up, the additional tracking information is useful to disentangle the main collision products, together with an average subtraction of the expected BIB energy deposition. Figure 27 (right) shows the expected jet energy resolution from a relatively simple particle-flow based jet reconstruction algorithm with minimal optimization on BIB energy subtraction. The resolution is comparable to the ones achieved by LHC when using analogous simple calibration algorithms.

The reconstruction of muons, electrons and photons have been implemented and assessed as well in Ref. [203], and shown to be viable. While more advanced studies will surely be needed to establish a full set of expected performance, the current studies present, for the first time, a complete picture that builds confidence in the ability to extract satisfactory performance in the muon collider environment for the basic high-pt objects needed for the physics program outlined earlier in this report.

The aim of the these reconstruction algorithms is, at this stage, to show what minimal performance is surely achievable, while the ultimate performance is almost certain to be vastly better than the projections shown.

5.6 Fast- to Full- Simulation comparisons

5.6.1 $H \rightarrow bb$ Cross Section

In this study the full simulation of the Muon Collider experiment is used to determine the statistical sensitivity on the measurement of the $H \rightarrow bb$ cross section at $\sqrt{s} = 3$ TeV, assuming an integrated luminosity of 1 ab$^{-1}$.

The signal $\mu^+\mu^- \rightarrow X H(\rightarrow bb)$ has been generated with WHIZARD, where $X$ are two muons or neutrinos. The background $\mu^+\mu^- \rightarrow X q\bar{q}$ has been generated with WHIZARD as well, considering diagrams not mediated by the Higgs bosons, and with $q = b, c$. Contributions from light jets are considered negligible, since a heavy flavour tagging technique is applied as explained later. A number of 10k signal events and 10k background events are generated in this way.

The detector configuration and simulation is described in Sec. 5.3. The beam-induced background is included. Jets are reconstructed using a Particle Flow algorithm for selecting tracks and calorimeter clusters, and the $k_t$ algorithm with radius $R = 0.5$ is used for the clustering. In order to select jets in the region with the best performance in term of reconstruction efficiency and jet energy resolutions, the requirements $p_T > 40$ GeV and $|\eta| < 2.5$ are applied.

The heavy flavour identification efficiencies and misidentification rate have been determined with independent samples of $bb$, $cc$ and light jets. Secondary vertices (SVs) are reconstructed using tracks in the jet cones, and a jet is tagged as heavy flavour if at least one SV is found. The light jet misidentification rate is found negligible, and for this reason light jets have not been included in the background. The tagging efficiencies have been determined as a function of the jet $p_T$ and
Figure 26: Tracking efficiency for single muon in events with beam-induced background versus transverse momentum (top) and resolution versus $\theta$ (bottom) from [203]. The loss of efficiency at small angle comes from the effects of the nozzle used to limit the beam-induced background flux into the detector. Work to improve tracking algorithm and to recover the loss of efficiency is ongoing.
Figure 27: Distribution of the secondary vertex proper lifetime for b, c and light tagged jets (top). Distributions are normalized to the unit area. Jet transverse momentum ($p_T$) resolution as a function of the $p_T$ of the jet for $b$-, $c$-, and light jets (bottom). Differences between flavors are mainly due to different angular distributions of jets.
polar angle \( (\theta) \). These efficiencies are then applied to the reconstructed \( H \to b\bar{b} \) and background samples with a reweighting technique.

The dijet invariant mass distributions for signal and background are then fitted with double-gaussian pdfs, in order to obtain the signal and background models. The number of expected signal and background events is determined by considering the WHIZARD cross sections, the total selection efficiency and the integrated luminosity. In particular, about 59.5k \( H \to b\bar{b} \) events and 65.4k background events are expected to be collected with 1 ab\(^{-1}\).

The signal and background invariant mass models and the expected number of events are used to generate pseudo-data. The pseudo-data are then fitted with the invariant mass models, by using an unbinned maximum likelihood fit, and by letting the signal and background yields float. In this way the measured \( H \to b\bar{b} \) yield is extracted. A result of one of this fits is shown in Fig. 28.

The uncertainty on the signal yield obtained from the fit is \( \sigma/\sigma_b \). Several pseudo-experiments (1k) have been performed to check the stability of this result, and to rule out possible biases. This uncertainty can be taken as the statistical uncertainty on the measurement of the \( H \to b\bar{b} \) cross-section.

![Graph showing dijet invariant mass distribution](image)

Figure 28: Result of the dijet invariant mass fit used to extract the \( H \to b\bar{b} \) yield and uncertainty. Pseudo-data are obtained by exploiting the Muon Collider simulation at \( \sqrt{s} = 3 \) TeV, and assuming an integrated luminosity of 1 ab\(^{-1}\).

The sensitivity in this channel, among others, has been studied in \cite{18} including physics backgrounds using fast simulation at both 3 TeV and 10 TeV. Here we briefly summarise the analysis strategy used there at 3 TeV for comparison with the above full simulation results. Events were generated using MADGRAPH5 [241] and showered with PYTHIA8 [242], with detector reconstruction and performance approximated by using the muon collider detector card included in the latest DELPHES [243] releases. Jets were clustered using the exclusive Valencia jet clustering algorithm with \( R = 0.5 \) and \( N = 2 \). After applying a \( p_T \) correction factor, preselection cuts of \( p_T > 40 \) GeV and \( |\eta| < 2.5 \) were applied.

A flat flavour tagging efficiency of 50% was taken for \( b \)-jets, with a \( c \)-jet misidentification rate of roughly 1-3%, and negligible light-jet mistagging rate. Events with two \( b \)-tagged jets passing the preselection cuts were then subject to a cut on the reconstructed dijet invariant mass of \( 100 < m_{jj} < 150 \) GeV. The sensitivity to the channel was then estimated using \( \Delta \sigma/\sigma = \sqrt{S+B}/S \), where \( S \) and \( B \) are the number of surviving signal and background events, respectively. The obtained precision is 0.76% for the total VBF \( \mu^+\mu^- \to X H(\to b\bar{b}) \) signal.

The results in \cite{18} include some diboson backgrounds such as \( \mu^+\mu^- \to \mu^\pm\nu_\mu W^\mp H \), so the precision presented there is not directly comparable to the full simulation result above. In order to obtain a closer comparison, we consider a modification of the results such that we only include the \( \mu^+\mu^- \to Xq\bar{q} \) background while still considering the total \( \mu^+\mu^- \to X H(\to b\bar{b}) \) as the signal. Doing this yields a precision of 0.73% in fast simulation, to be compared to the 0.75% obtained in
the full simulation fit above.

It is worth mentioning that the largest impact of the BIB on this channel’s analyses is the worsened jet energy resolution, which results in a larger overlap of the $H$ and $Z$ peaks, as can be seen in Figure 5 of [18], where using a worse jet energy resolution than the Delphes default in fast simulation reproduced a very similar distribution to Figure 28. Of course, the fast simulation result uses less sophisticated flavour tagging and a simpler analysis strategy, so the comparison is not quite one to one. Nevertheless, we find the sensitivity for this channel in fast simulation and full simulation to be very similar, regardless of the presence of the BIB.

5.6.2 Dark Matter with Disappearing Track

Long-lived particles (LLPs) appear in a variety of models and yield a large range of signatures at colliders [244, 245]. Depending on the LLP quantum numbers and lifetime, these can span from LLP decay products appearing in the detector volume, even outside of the beam crossings, to metastable particles with anomalous ionisation disappearing after a short distance.

The higgsino is among the most compelling dark matter candidates, with tight connections to the naturalness of the weak scale, which could lead to charged LLPs ($\tilde{\chi}^\pm$) being produced in particle collisions and then decaying in the volume of the tracking detectors (e.g. decay lengths between 1 mm and 500 mm).

Searches at the LHC are actively targeting this scenario [246, 247, 248, 249, 250], but are not expected to cover the relic favoured mass of 1.1 TeV [251, 252, 253]. Different studies have assessed the reach of a muon collider in the search for these particles with fast [41, 42] and full [254] simulation. This document reviews these two results and provides comparison of the expected reach from each approach as a function of the main parameters of the model.

The production of pairs of electroweakinos at a muon collider operating at multi-TeV centre-of-mass energies proceeds mainly via an s-channel photon or off-shell Z-boson, with other processes, such as vector boson fusion, being subdominant [41]. The decay products of the $\tilde{\chi}^\pm$ are assumed to be undetectable.

In the work described in Ref. [41, 42], events are selected with an ISR photon with minimal $p_T$ of 25 GeV, and one or two disappearing tracks. We assume the reconstruction probability of a signal event with one disappearing track is

$$
e_{\tilde{\chi}}(\cos\theta, \gamma, d_T^{\text{min}}, d_T^{\text{max}}) = \exp\left(\frac{-d_T^{\text{min}}}{\beta_T\gamma c\tau}\right) - \exp\left(\frac{-d_T^{\text{max}}}{\beta_T\gamma c\tau}\right),$$

with $d_T^{\text{min}} = 5$ cm, $d_T^{\text{max}} = 17$ cm and $|\eta| < 1.5$

where $\gamma = E_{\tilde{\chi}}/m_{\tilde{\chi}}$ and $\beta_T = \sqrt{1 - E_{\tilde{\chi}}/m_{\tilde{\chi}}}$, which is the transverse velocity in the lab frame. The minimal transverse displacement of 5 cm represents the minimal track reconstruction requirement (of two hits) for a typical muon collider detector design with pixel layers. A maximal transverse displacement of 17 cm to capture the “disappearing” signature and comparable to the detector simulation-based study. In principle, longer tracks are easier to be reconstructed and separated from background. Future detector design studies could further optimize the layout for Higgsino-like short tracks, e.g. moving pixel layers closer to the beam spot, that could greatly improve the muon collider sensitivities.

The study described in Ref. [255] investigates the prospects for such a search exploiting a detector simulation based on Geant 4 [256] for the modelling of the response of the tracking detectors, which are crucial in the estimation of the backgrounds. The simulated events were overlaid with beam-induced background events simulated with the MARS15 software [257].

The analysis strategy relies on requiring one (SR$_{11}$) or two (SR$_{22}$) disappearing tracks in each event in addition to a 25 GeV ISR photon. Additional requirements are imposed on the transverse momentum and angular direction of the reconstructed tracklet and on the distance between the two tracklets along the beam axis in the case of events with two candidates. The disappearing track candidates are required to be reconstructed from at least four silicon detector hits (corresponding to a radius of 5.1 cm), and to have no associated hits beyond the first layer of the inner tracker (corresponding to a radius of 12.7 cm). This veto condition is relaxed to the middle layer of the outer tracker (corresponding to a radius of 115.3 cm) for one of the two candidates in SR$_{22}$.

The results of the fast and full simulation approaches are compared in Figure 29 in terms of expected discovery sensitivity. The expected backgrounds are taken from the full detector
Figure 29: Expected sensitivity using $10 \text{ ab}^{-1}$ of 10 TeV $\mu^+\mu^-$ collision data as a function of the $\tilde{\chi}^\pm$ mass and lifetime. The contours represent the $5\sigma$ discovery expectation. The solid lines show the predictions from full simulation, while the dashed lines the predictions from the fast simulation.

5.7 R&D Priorities for Muon Collider Detectors

Optimization of the detector for muon collider environment is far from complete. Significant improvements are possible in all aspects of the detector. Some of the items that need early study are:

- **MDI**: Beam induced backgrounds are significant at the muon collider. Current BIB simulations are done with $E_{cm} = 1.5$ TeV, with some limited extrapolation to $E_{cm} = 3.0$ TeV. Proper simulation of target energy of $E_{cm} = 10$ TeV is amongst the most important items for detector optimization. Options for the absorber material, shape and size determination need to be made with both background mitigation for physics object reconstruction point of view and the lifetime of the sensitive material of the planned detectors. The work requires close collaboration with accelerator lattice designers.
• **Magnet**: What is the size of the magnet and the field? The answer depends on the $E_{cm}$. Additional dependence is on the size and shape of the nozzle used to mitigate the beam induced backgrounds.

• **Tracker**: The location of the inner layers, number of tracking layers needed, segmentation size of the pixels and strips, etc. Of particular concern is the mitigation of single low energy BIB particle hits, which can result in very large data size and excessively large combinatoric problem for track reconstruction. One possible way to mitigate the BIB is to discriminate single-hits against multi-layer correlated hits characteristic of real higher momentum tracks. The optimal way to reduce this background needs to be identified. The study requires detailed simulation of the expected detector conditions, with overlapping simulated events of interest. Cost, radiation tolerance, power and cooling optimization will place constraints on number of electronics channels and sophistication of electronics needed.

The ability to have precision timing information available as hit level information would improve the ability to separate hits from particles from the hard scattering interaction from BIB particles and allow filtering at the readout and trigger level. Silicon sensor with 20-30 picosecond resolution are being studied and developed currently [216]. Continued development of these technologies to be able to construct future detector with picosecond level timing resolution would be expected to vastly decrease the readout requirements and reduce combinatoric confusion in tracking.

• **Calorimeter**: Photons, electrons, taus and jets need to be reconstructed with good resolution and efficiency from few 10s of GeV to TeV scale in order to do physics at the muon collider. In addition to the usual trade-off between the electromagnetic resolution and the hadronic resolution, mitigation of fake jet background, from the BIB background deposits, is essential. Segmentation of the calorimeter in all three dimensions and signal arrival time is needed both for particle-flow particle reconstruction to improve the energy resolutions using tracking of charged-hadrons and BIB mitigation. Cost drivers are both calorimeter materials quality, from signal yield, radiation damage, calibration stability, etc. points of view, and the electronics needed for processing the signals.

Different types of high granularity calorimeters with good energy resolution and precise timing information are being developed and constructed. Continued development of novel calorimeters with picosecond level timing resolution would improve rejection of BIB particles and improve photon, electron, and jet reconstruction by rejecting energy deposits from non-collision sources.

• **Muon system**: The challenge of measuring momentum of TeV-muons is to be studied. The bend needed for TeV muons has significant cost implications due to the magnet size and field. Secondly, bremsstrahlung from high momentum muons begins to degrade the ability to measure the momentum well. Perhaps, a secondary lower momentum but larger sized magnet may be appropriate but will have implications on the size of the experimental cavern and the size of the muon detectors that are needed. Deep underground location for a 10-TeV machine is another cost consideration.

Micropattern gas detectors with tens of picosecond timing resolution and integrated readout electronics are currently an item of high priority in the gas detector community [258]. The continued development of detectors with excellent spatial and timing resolution which do not use greenhouse gases are critical to the performance of such a detector.

• **Data acquisition system**: The true event rate at the muon collider is low enough, and the time between collisions is long enough, to be able to operate triggerlessly provided the data volumes are under control. However, much investigation is needed to tame the data volumes by dropping the BIB hits, while not compromising on the signal hit efficiencies. The power and cooling needed for on-detector electronics is also of concern. As well the continued development of integrated electronics, fast high-bandwidth readout to handle large data volumes should also be pursued.

Studies are needed in all of the areas outlined. While generic detector R&D can provide innovative solutions, dedicated studies of the muon collider environment in simulation is the key to make improvements in the design of the detector, which is currently adopted from the CLIC detector.
The nature of the problem requires detailed layout of the detector and simulation in GEANT, requiring talented and knowledgeable experts providing core framework support, as well as a group of students and postdocs.

While technologies needed for a muon collider detector appear feasible, dedicated R&D efforts are necessary to make them mature and cost-effective by the time the accelerator technology is ready. The detector R&D thrusts of timing, on-detector filtering, and radiation hardness are aligned with work needed for FCC and other future collider detectors. There is clearly a general need for R&D on the simultaneous optimization of mass, speed, power and cooling, and mechanics for the muon collider and other future detector systems. It is therefore imperative to incorporate these needs into the existing and future detector R&D programs.

6 Synergies

6.1 Neutrino Frontier

A muon collider offers a unique opportunity to deepen our understanding of the neutrino sector. A neutrino beam sourced by muon decay could have significant impact in several aspects of neutrinos physics: standard and beyond standard oscillation physics; searches for novel states; BSM effects in the neutrino sector; electroweak precision physics; neutrino-nucleus interaction modeling; among others. Since historically a neutrino experiment in which the beam comes from muon decays has been referred to as the Neutrino Factory, we will also use the term hereafter, even when we will not be focusing on oscillation physics. In what follows, we will briefly describe some of these opportunities and explain what would be the role of a neutrino factory in leveraging those. Further details on the physics case of a neutrino factory and its synergy with a muon collider can be found in Ref. [259].

First, let us consider uncertainties related to neutrino production and detection. In beam neutrino experiments, neutrinos are produced primarily from the decay of mesons, which are in turn produced by a beam of protons impinging on a target. The uncertainties related to meson production will therefore propagate to the neutrino fluxes. On top of that, neutrinos are detected by interacting with nuclei, due to the larger cross section compared to neutrino-electron scattering. Currently, flux and cross section uncertainties are among the dominant limiting factors in the determination of neutrino oscillation parameters [260, 261]. Moreover, neutrino event generators do not describe well exclusive final states and differential cross sections, which are both crucial to the reconstruction of the incoming neutrino energy [262, 263]. Neutrino experiments mitigate these issues by leveraging a two-detector configuration: a near detector which is sensitive to the unoscillated neutrino spectrum, and thus measures the initial flux and cross section; and a far detector which measures the oscillated neutrino spectrum. Nevertheless, the near detector is not sufficient to determine both neutrino flux and cross section, and the uncertainties associated to those will be the dominant uncertainties in future measurements of neutrino experiments such as DUNE and HK [264, 265].

Having a well known neutrino flux is the key strength of a neutrino factory. This would allow precise determination of neutrino-nucleus interaction cross sections, which could significantly improve measurements at DUNE and HK — CP violation, precision determination of mixing angles and mass splittings [264, 265], nonstandard neutrino interactions [266], sterile neutrinos [267], Earth tomography with atmospheric neutrinos [268], etc. To achieve this, one would need a neutrino spectrum that overlaps with DUNE and/or HK spectra (around 1-5 GeV and 0.2-1 GeV, respectively), and near detectors with the same chemical composition (argon and water, respectively). Note that a better determination of neutrino-nucleus interaction cross sections could be used to improve oscillation measurements even after the experiments took data, as long as data preservation plans are put in place. On top of that, searches for new states in neutrino experiments, such as heavy neutral leptons, light axions, dark matter, and dark neutrinos [269, 270, 271, 268, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282], typically need to deal with backgrounds originating in neutrino-nucleus interactions. A better determination of these cross sections would translate into smaller background uncertainties, improving past and current BSM searches.

At higher neutrino energies, above 10 GeV or even 100 GeV, neutrino events would be dominantly observed by deep inelastic scattering cross section (DIS). Measuring DIS in neutrino-nucleus interactions could be relevant to the understanding of high energy atmospheric neutrinos and perhaps even to high energy cosmogenic neutrinos, as observed by the IceCube experiment [283, 284].
In particular, nuclear parton distribution functions still have large uncertainties [285, 286, 287], and a measurement of neutrino-nucleus interactions with well understood neutrino fluxes could improve our knowledge of these nuclear pdfs. DIS events would also allow to revisit the NuTeV determination of the weak mixing angle with improved theoretical descriptions of QCD and nuclear effects, possibly clarifying the NuTeV anomaly [288]. Besides, a high energy neutrino factory would be an excellent environment to probe generic new physics in neutrino interactions via an effective field theory approach. This would only require a near detector, and as much statistics as one can obtain. It has been shown that experiments like FASERν can be competitive to LHC and meson decay observables when probing certain dimension-6 EFT operators [289, 290]. In fact, a high energy muon collider could host a neutrino experiment similar to FASERν [291, 292], with much smaller flux uncertainties compared to a forward detector at a hadron collider.

### 6.2 Intensity Frontier

A muon collider facility would also provide opportunities to search for weakly coupled new particles below the weak scale. Given a high energy muon beam, an economical extension to the collider facility is a beam dump experiment, where the muon beam is directed onto a dense target with a detector placed at the end of a long decay volume. Similar extensions have been proposed for the LHC and other future colliders [293, 292, 294]. With a dense target, the cross section for forward production of new states with masses \( m < E_{\text{beam}} \) can be very large, compensating the weak couplings to SM states.

The precise reach for new, weakly-coupled particles depends on the details of the experimental setup, such as the target material, shielding, the length of the decay volume, and the detector, as well as the number of muons on target. These considerations must be optimized in concert with the design of the collider facility to maximize the reach while accounting for the cost and collider facility constraints. Given the sensitivity to these choices, and the immense opportunity, it is necessary to consider this optimization in the design of a muon collider facility itself.

A first estimate of the potential reach, however, was provided in ref. [295], assuming a 1.5 TeV muon beam, with variable choices of the number of muons on target and reasonable choices for the other experimental details. The results of this study are reproduced in Fig. 30, considering two benchmark models: a kinetically-mixed dark photon, and an \( L_\mu - L_\tau \) gauge boson. In the former case, we see that the muon collider reach extends to higher masses and larger mixing angles than any other proposed experiment. This is due to the larger boost that these particles are produced with from a high-energy lepton beam, which compensates their shorter lifetimes at these masses and mixing angles so that they can survive to the detector. An even higher energy beam in e.g., a 10 TeV collider facility, could extend the reach even further. In the \( L_\mu - L_\tau \) gauge boson case, a muon beam dump at these energies would probe entirely disparate parameter space from any other imagined experiment. This is both due to the boost, and from the unique setup involving a muon beam as opposed to an electron or proton. Exploration of the reach for other dark sectors, particularly those benefiting from the muon-specific opportunities, is deserving of further study.
Figure 31: Kinematic coverage of $Q^2$ and $x$ in deep inelastic lepton-proton scattering for two muon-ion collider design options and for the EIC at BNL, HERA at DESY, and the LHeC and FCC-eh options at CERN, each at their maximum beam energies. The inelasticity ($y$) range is assumed to be $0.01 < y < 0.95$ (hatched areas). The long dashed lines indicate the saturation scale as a function of $x$ in the proton from the GBW model [301].

Finally, while we have emphasized the unique opportunities of a high energy muon beam above, other opportunities for probing light, weakly coupled dark sectors exist at intermediate stages of a muon collider facility. The reach of muon beam experiments with lower energy muon beams has been studied in other contexts [296, 297, 298], demonstrating their capability for probing new physics with muon-specific couplings such as the $L_\mu - L_\tau$ gauge boson. Staged approaches to the development of a muon collider facility may present the perfect opportunity for an economical, low-energy muon beam dump experiment to be performed in parallel, and potentially in combination with neutrino beam experiments discussed in the preceding Section. Ultimately, it is important that the full scope of possibilities for exploring new physics at a muon collider facility be explored, so that no stone is left unturned.

6.3 Muon-Ion Collider

A possible scientific target for an intermediate step toward the ultimate development of a multi-TeV muon collider is that of a muon-proton and muon-nucleus collider facility, referred to as a mu-ion collider, as discussed in Refs. [299, 300]. Such a facility could utilize the existing hadron accelerator infrastructure at laboratories such as BNL (as an upgrade to the planned Electron-Ion Collider), CERN (using the LHC), or Fermilab while seeding, or leveraging, the development of a high energy and high intensity muon storage ring at the same site. A muon-proton center-of-mass energy of up to 1 TeV at BNL (6.5 TeV at CERN) can be achieved when a 1 TeV (1.5 TeV) muon beam is brought into collision with a 0.275 TeV (7 TeV) proton beam at that facility. Such a mu-ion collider would enable deep inelastic scattering measurements in completely new regimes at low parton momentum fraction $x$ and high squared four-momentum transfer $Q^2$, as illustrated in Fig. 31, which will further elucidate the structure of the proton and of nuclei as well as provide precision QCD and electroweak measurements complementary to those done at lepton and hadron colliders. The coverage of a mu-ion collider at BNL is nearly identical with that of the proposed Large Hadron electron Collider (LHeC) at CERN, which would use the LHC and a 50 GeV electron beam, while the coverage of a muon-LHC collider at CERN (LHmuC) would significantly exceed that of the FCC-eh option of a 50 TeV proton beam colliding with a 50 GeV electron beam.
The maximum muon beam energy at BNL is taken to be 1 TeV assuming that 11 T dipoles are used in the arcs of the RHIC ring (290 m radius). In terms of staging scenarios for the muon beam energy, we note that any energy above 93 GeV leads to a center-of-mass energy above that of the HERA ep collider. In addition, beam polarization is a unique capability at the BNL facility. A muon beam energy of 1.5 TeV at CERN comes simply from assuming that one beam of a 3 TeV $\mu^+\mu^-$ collider constructed there is brought into collision with the LHC.

The estimate on the maximum achievable luminosity at the mu-ion collider is $4.7 \times 10^{33}$ cm$^{-2}$s$^{-1}$ (2.8 $\times 10^{33}$ cm$^{-2}$s$^{-1}$) for a collider at BNL (CERN). For the muon beam, the proposed parameters of the proton driver scheme from Ref. [302, 303] are taken. The muon bunch repetition frequency is taken to be 12–15 Hz. The BNL proton beam parameters are assumed to be those achieved at RHIC [304] or foreseen to be achieved at EIC [305], and the CERN beam parameters correspond to those of the LHC [306].

The scientific potential of the mu-ion collider is similar to the proposed LHeC, but with complementary scattering kinematics and complementary sensitivity to BSM processes with a muon beam as opposed to an electron beam. For deep inelastic structure function measurements, the mu-ion collider can probe parton momentum fractions $x$ in the proton as small as $10^{-6}$ ($10^{-8}$) for collisions at BNL (CERN), and thus should be sensitive to gluon saturation effects [301]. The corresponding maximum reach in $Q^2$ is $10^6$ GeV$^2$ ($4 \times 10^7$ GeV$^2$), which is well above the electroweak scale where neutral-current and charged-current scattering cross sections become similar. The maximum reach in $Q^2$ of the BNL mu-ion collider can be achieved with 10 fb$^{-1}$ of integrated luminosity, and is well within the machine estimation. However, to probe the highest $Q^2$ reach of the LHmuC would require $\approx 1000$ fb$^{-1}$, which is somewhat disfavored in the beam assumptions. However, the cross sections at lower $Q^2$ and correspondingly lower $x$ are many orders of magnitude larger, and thus enable a science program in that regime with significantly less luminosity by the same order. Finally, the mu-ion collider also could provide polarization of both beams (when utilizing the BNL facility) for spin structure measurements, and provide lepton-proton and antilepton-proton collisions with similar luminosity (although switching between the two would require flipping the polarity of magnets along the muon beam line).

Higgs boson production also opens up at the mu-ion collider through vector boson fusion, with a cross section of 80 fb for $\mu^-p$ collisions at the BNL mu-ion collider and 1700 fb at the LHmuC. The Higgs boson decay products tend to be more central in the detector for collisions at the mu-ion collider than for those at the LHeC, given its more symmetric collision. For a $H \rightarrow b\bar{b}$ measurement at the mu-ion collider, a statistical uncertainty on the signal of about 3% is possible from a 10-year data set (400 fb$^{-1}$). Another decay channel with significant yield that is potentially measurable is $H \rightarrow \tau^+\tau^-$. We note that $H \rightarrow gg$ would be an interesting target as well. The $H \rightarrow c\bar{c}$ channel would open up with the higher cross sections available at higher energies, such as with the LHmuC.

The production of other Standard Model particles at a mu-ion collider also are of interest. In total, the inclusive $W^\pm$ production cross section in $\mu^-p$ collisions for the mu-ion collider at BNL is about 20 pb, yielding $2.1 \times 10^4$ leptonic $W \rightarrow \ell\nu$ decays into each lepton flavor for 10 fb$^{-1}$ of integrated luminosity. This increases by an order of magnitude for the LHmuC configuration. Both single and pair production of top quarks is also possible, the former of which directly targets the CKM matrix element $V_{tb}$. Precision mass measurements of the W boson and t quark also may be viable, depending on the measurement capabilities of a mu-ion collider experiment.

The mu-ion collider also offers interesting sensitivity to BSM processes. For example, the s-channel and t-channel exchange of a new particle such as a leptoquark or Z$'$ boson is possible, and $\mu$-$p$ scattering measurements would be particularly relevant to test BSM models that propose to explain the potential anomaly in the muon $g-2$ measurement and in the deviations from lepton flavor universality in B meson decays. While certain leptoquark cross sections at the mu-ion collider have been reported in [300], an analysis of the constraint sensitivity still remains to be done.

The mu-ion collider concept expands the science opportunities at a facility featuring a high-energy, high-intensity muon accelerator and collider. When the muon beam is brought into collision with a hadron beam, a unique frontier in both particle and nuclear physics is opened up at high $Q^2$ and low $x$. The communities involved in the funding and construction of such a facility overall could therefore increase as well, sharing the overall cost burden. Additionally, much of the science research can be done with modest goals for the initial muon beam energy and luminosity, making it an ideal scientific target for a muon collider demonstrator that is on the path toward the development of an ultimate $\mathcal{O}(10+)$ TeV $\mu^+\mu^-$ collider. The mu-ion collider is very amenable to
a staged approach to the maximum energy and intensity of the muon beam. Costs for the mu-ion collider component of the facility can be further minimized by reusing existing hadron beam and collider infrastructure at an appropriately chosen site.

The kinematics of the collisions and the experimental needs of a mu-ion collider are similar to those for a muon collider experiment. The asymmetric nature of the collisions, however, leads to an emphasis of instrumentation along the downstream muon beam direction at small scattering angles, although the rest of the final state tends to occupy the central region of the experiment as would $\mu^+\mu^-$ annihilations. One shielding cone may be sufficient to reduce the beam induced background for a mu-ion collider experiment (on the incoming muon side, to be studied), which would potentially open up the instrumentation room at small angles on the downstream side. As the scattered muon peaks strongly in the muon beam direction for low-$Q^2$ and for vector boson fusion processes, a dedicated muon spectrometer along the beamline would be needed. But such a spectrometer design also may prove useful for experiments at a $\mu^+\mu^-$ collider to tag the nature of VBF processes.

6.4 Applications outside HEP

Various industry applications of high intensity muon beams have been mentioned in literature. For example, muon spin rotation/relaxation/resonance ($\mu$SR) is a collection of methods that use the muon’s spin to examine structural and dynamical processes in bulk materials at the atomic scale. With $\mu$SR beams of polarized muons are shot into a material. The muons’ spins precess around the local magnetic fields in the material. One can then examine how the internal magnetic fields of different materials have affected the muons’ spins by observing the directions in which the positrons produced in decays of the muons are emitted.

Due to a strong penetrating power of muons, cosmic-ray muons have been widely used for imaging of large scale objects. The muon scattering angle in such an imaging process depends not only on the amount of material but also on the muon energy. A large disadvantage of the cosmic ray based muon imaging is that the flux is low and the muons have a wide spread in energy. On the other hand, artificially produced and accelerated muon beams can have high flux and be monochromatic, enabling higher resolution imaging in less time.

These applications could clearly benefit from better quality muon beam. Because a muon beam is generated as a tertiary beam and initially has a large phase-space volume, cooling techniques are necessary to reduce the phase space volume, thus increasing the quality of the muon beam. Further advancements are also necessary to achieve muon acceleration for uses across numerous industry applications.

7 Path Forward

The most fruitful path forward towards the development of conceptual design of the Muon Collider and associated detector in the USA, will involve coordination with the IMCC efforts. A US MCC needs to be formed with a goal towards preparing a proposal for the US funding agencies, assuming identification of the Muon Collider as an important path forward for the US energy frontier community. The US collaborations of the LHC experiments and the accelerator form a good example. Given the necessity of a strong coupling of the Muon Collider accelerator and detector communities, it seems most suitable to follow the IMCC model with a unified US MCC group, rather than work individually at various institutes. We anticipate that individual US institutions joining the IMCC will also be members of a new US MCC, elect its leadership and represent the combined interest of the US muon collider community to the US funding agencies.

7.1 Engagement in IMCC

To foster the Muon Collider concept an International Muon Collider Collaboration (IMCC) has been initiated after the recommendation of the update of the European Strategy for Particle Physics, initially hosted at CERN. The collaboration will address the muon collider challenges and develop the concept and technologies in the coming years in order to be able to gauge if the investment into a full conceptual design and demonstration program is scientifically justified. This will allow future Strategy Processes in the different regions to make informed decisions. The collaboration welcomes active participation from the United States and encourages interested
institutions to sign a Memorandum of Cooperation. Many US institutions are interested in joining
the collaboration but have been waiting for the outcome of Snowmass/P5. An explicit endorsement
by P5 and a collaborative agreement between CERN and DOE would allow for direct participation
of US institutions in IMCC and for direct contributions from US to the program.

7.2 Contributions to Physics Studies
The Muon Smashers Guide and the studies from the Pittsburgh particle physics community have
made a strong impact in the burgeoning interest in the 10-TeV scale physics. The case for develop-
ing this physics case, with more realism and rigour in event generation and simulation, and addition
of other physics cases of interest is very much of interest. A strong collaboration of phenomenol-
ogists and experimenters is necessary for this effort to be fruitful. We anticipate that there will
be contributions which will establish the prospects for the measurements of both standard model
higgs parameters and exploration of new physics sectors. Furthermore, studies of various staging
scenarios will be carried out together by the accelerator, experimental, and theoretical communities
in order to develop the most comprehensive integrated program.

7.3 Contributions to Detector R&D
The US groups should participate in determining the detector parameters for achieving the best
physics results. The choice of detector technologies, size, overall geometry and segmentation are
yet to be optimized. The initial studies from the IMCC, based on adaptation of the CLIC detector,
is a good start, and appears suitable. Perhaps, a fresh ground up start with a second detector
concept may also be useful in understanding the details of what is important and make a com-
parative evaluation to form a viable detector proposal. Of particular importance is mitigation of
beam induced background. Presently, BIB simulation program is tricky and time consuming and
restricted to lower energy colliders. A close collaboration of FNAL based beam dynamics simulation
experts and postdocs/students from US universities could yield more realistic studies for 10
TeV collider.

7.4 Contributions to Accelerator R&D
The US based MAP collaboration has made very significant contributions to the Muon Collider
concept. The IMCC has essentially taken over the direction of Muon Collider R&D lately because
of the US divestment from the MAP program. However, some of the expertise at BNL and FNAL
remains and can quickly be reconstituted to bootstrap a substantial R&D team. The rebuilding of
the team will enable us to participate in the CERN based demonstrator efforts which are recently
funded. Perhaps, some of the pieces of equipment used for MAP R&D could be resurrected to
contribute towards the IMCC demonstrator. The reconstituted accelerator R&D team will also
give us the opportunity to prepare the US demonstrator proposals, once the P5 process identifies
the project priority and the DOE the scale of project budgets.

7.5 Explore US options
The reconstituted US Muon Collaboration with the MAP participants and its expansion to include
the expanding group of scientists with a new-found interest in the Muon Collider, provides an
ideal opportunity to make designs of the collider for potential US siting. FNAL could be an ideal
site for a Muon Collider with a center-of-mass energy reach at the desirable 10-TeV scale. The
synergy with the existing/planned accelerator complex and neutrino physics program at FNAL
is an additional stimulus for such investment of effort. A set of Muon Collider design options,
with potential siting at FNAL, could be made for discussions at the IMCC and the international
committees to eventually form a global consensus decision on siting and selection of the Muon
Collider. Having a pre-CDR document summarizing design for the FNAL-sited Muon Collider
in time for the next Snowmass is a good goal. The preparation of such a document will require
substantial, yet affordable, investment. Such an investment will reinvigorate the US high-energy
collider community and enable much needed global progress towards the next energy frontier.
References


M. Davier et al. “Reevaluation of the hadronic vacuum polarisation contributions to the
A. Czarnecki, W. J. Marciano, and A. Vainshtein. “Refinements in electroweak contribu-
A. Keshavarzi, D. Nomura, and T. Teubner. “Muon $g - 2$ and $\alpha(M_Z^2)$: a new data-based
G. Colangelo, M. Hoferichter, and P. Stoffer. “Two-pion contribution to hadronic vacuum
A. Kurz et al. “Hadronic contribution to the muon anomalous magnetic moment revisited”. In:
M. Davier et al. “A new evaluation of the hadronic vacuum polarisation contributions to the
A. Keshavarzi, D. Nomura, and T. Teubner. “$g - 2$ of charged leptons, $\alpha(M_Z^2)$, and the hyperfine splitting of muonium”. In:
A. Kurz et al. “Hadronic contribution to the muon anomalous magnetic moment to next-to-next-to-leading order”. In:
K. Melnikov and A. Vainshtein. “Hadronic light-by-light scattering contribution to the muon anomalous magnetic moment revisited”. In:
P. Masjuan and P. Sanchez-Puertas. “Pseudoscalar-pole contribution to the $(g_\mu - 2)$: a rational approach”. In:


V. V. Kashikhin et al. “High-Field Combined-Function Magnets for a 1.5 x 1.5 TeV Muon Collider Storage Ring”. In: Conf. Proc. C (2012). Ed. by V. Suller. FERMILAB-CONF-12-148-APC-TD.


