A Practical High-Energy High-Luminosity $\mu^+\mu^-$ Collider

Robert B. Palmer, SLAC, Stanford University, Stanford CA 94305 and
Brookhaven National Laboratory, Box 5000, Upton NY 11973

David V. Neuffer, CEBAF, 12000 Jefferson Avenue, Newport News VA 23606

Juan Gallardo, Brookhaven National Laboratory, Box 5000, Upton NY 11973

Abstract. We present a candidate design for a high-energy high-luminosity $\mu^+\mu^-$ collider, with $E_{cm} = 4$ TeV, $L = 3 \times 10^{34} \text{cm}^{-2}\text{s}^{-1}$, using only existing technology. The design uses a rapid-cycling medium-energy proton synchrotron, which produces proton beam pulses which are focused onto two $\pi$-producing targets, with two $\pi$-decay transport lines producing $\mu^+$'s and $\mu^-$'s. The $\mu$'s are collected, rf-rotated, cooled and compressed into a recirculating linac for acceleration, and then transferred into a storage ring collider. The keys to high luminosity are maximal $\mu$ collection and cooling; innovations with these goals are presented, and future plans for collider development are discussed. This example demonstrates a novel high-energy collider type, which will permit exploration of elementary particle physics at energy frontiers beyond the reach of currently existing and proposed electron and hadron colliders.

INTRODUCTION

Lepton($e^+e^-$) colliders have the valuable property of producing simple, single-particle interactions, and this property is essential in the exploration of new particle states. Extension of $e^+e^-$ colliders to multi-TeV energies is performance-constrained by radiation and “beamstrahlung” effects, which increase as $(E/m_e)^4$, and cost-constrained by the need for two full-energy linacs. However, muons (heavy electrons, with $m_\mu = 200m_e$) have negligible beamstrahlung and can be accelerated and stored in rings. The liabilities of muons are that they decay, with a lifetime of $2.2 \times 10^{-6}$ $E_\mu/m_\mu$ s, and that they are created through decay into a diffuse phase space. But that phase space can be reduced by ionization cooling, and the lifetime is sufficient for storage-ring collisions. (At 2 TeV, $\tau_\mu = 0.044$ s.) We present the first practical design for a high-energy high-luminosity $\mu^+\mu^-$ collider, with an energy of $E_{cm} = 2E_\mu = 4$ TeV, and a luminosity of $L = 3 \times 10^{34} \text{cm}^{-2}\text{s}^{-1}$, which uses only existing technical capabilities.

The possibility of muon ($\mu^+-\mu^-$) colliders has been introduced by Skrinsky et al. and Neuffer. More recently, several mini-workshops have greatly increased the level of discussion, stimulating the present developments. In this paper we introduce improvements, and develop a complete scenario for a high-luminosity high-energy collider. Table 1 shows parameters for the candidate...
DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, make any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.
DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.
design, which is displayed graphically in fig. 1. The design consists of a muon source, a muon collection, cooling and compression system, a recirculating linac system for acceleration, and a full-energy collider with detectors for multturn high-luminosity collisions.

**MUON PRODUCTION**

The $\mu$-source driver is a high-intensity rapid-cycling (10Hz) synchrotron at KAON\(^8\) proposal parameters (30 GeV), which produces beam which is formed into two short bunches of $3 \times 10^{13}$ protons. Combination of the accelerated beam bunches into two short bunches at full energy (possibly in a separate extraction ring) simplifies the subsequent longitudinal phase space manipulations. The two proton bunches are extracted into separate lines for $\mu^+$ and $\mu^-$ production. (Separate lines permit use of higher-acceptance, zero-dispersion $\pi \rightarrow \mu$ capture lines.) Each bunch collides into a target, producing $\pi$'s ($\sim$1 $\pi$ interacting per proton) over a broad energy and angular range ($E_\pi = 0$—4 GeV, $p_\perp < 0.5$ GeV/c). The target is followed by $Li$ lenses, which collect the $\pi$'s into a large-aperture high-acceptance transport line (an $r = 0.15m$, $B = 4T$ FODO transport with a 0.8m period in a current scenario), designed to accept a large energy width (2±1 GeV) and have a large transverse acceptance ($p_\perp < 0.4$ GeV). This array is sufficiently long (~300 m) to insure $\pi \rightarrow \mu$ decay, plus debunching, in which the energy-dependent particle speeds spread the beam longitudinally to a full width of ~6 m, while reducing the local momentum spread.

This is followed by a nonlinear rf system (3 harmonics are sufficient) which flattens the momentum spread. We have conservatively assumed that rf gradients at these frequencies (~10—30 MHz) are limited to less than ~2 MeV/m; this implies a total rf debuncher length of ~1.4 km. (Higher gradients would permit a shorter debuncher. Also, an induction linac rather than resonant-cavity rf could be used, as suggested by Barletta.)\(^9\) The resulting $\mu$-beam is then matched into a beam cooling system. Figure 2 shows a schematic overview of the production and collection system.

A Monte Carlo program (MCM - Monte Carlo Muon) has simulated the muon production and cooling. The generation of $\pi$'s in the target is calculated using a thermodynamic model or the Wang distribution.\(^{10}\) The $\pi$'s are tracked through decay to $\mu$'s, and phase-energy rotation, into the cooling system. We obtain ~0.15 captured $\mu$'s per initial proton, with $\varepsilon_N = 0.01$ m-rad, an rms bunch length of 3m, and energy width of 0.15 GeV with an average energy of 1 GeV.

The $\mu$ capture efficiency (0.15$\mu/p$) is larger than estimated in previous scenarios\(^2\),\(^3\),\(^4\), and this is a result of the use of a high acceptance decay transport with a larger momentum acceptance. The transport is followed by a linac-based rf rotation, which reduces the momentum spread to a level acceptable by the
subsequent transport and cooling system, while lengthening the beam bunch.\textsuperscript{11} Previously, Noble\textsuperscript{4} has noted that $\mu^+\mu^-$ collider luminosity increases with the energy spread acceptance $\Delta E_A$ by as much as $\Delta E_A^4$, with factors accumulating from the production and the decay acceptances of both beams. The high-acceptance transport plus rf rotation has increased that acceptance by almost an order of magnitude above previous estimates. Also, for the first time, the production has been directly calculated in a realistic simulation, and not simply estimated.

**BEAM COOLING**

For collider intensities, the phase-space volume must be reduced by beam-cooling and the beam size compressed, within the $\mu$ lifetime. Much of the needed compression is obtained through adiabatic damping in acceleration from GeV-scale $\mu$ collection to TeV-scale collisions. Beam cooling is obtained by "ionization cooling" of muons ("$\mu$-cooling"), in which beam transverse and longitudinal energy losses in passing through a material medium are followed by coherent reacceleration, resulting in beam phase-space cooling.\textsuperscript{2,3,12} (Ionization cooling is not practical for protons and electrons because of nuclear scattering ($p$'s) and bremsstrahlung ($e$'s) effects, but is for $\mu$'s and the necessary energy losses are easily obtained within the $\mu$ lifetime.) In this section we present the equations for $\mu$-cooling, use these to deduce optimal cooling conditions, and generate a practical cooling scenario.

The equation for transverse cooling is:

\[
\frac{d\epsilon_N}{ds} = -\frac{dE}{ds} E_{\mu} \frac{\beta_\perp}{2} \left(0.014\right)^2 \frac{1}{E_{\mu} m_\mu L_R}
\]

(with energies in GeV), where $\epsilon_N$ is the normalized emittance, $\beta_\perp$ is the betatron function at the absorber, $dE/ds$ is the energy loss, and $L_R$ is the material radiation length. The first term in this equation is the coherent cooling term and the second term is heating due to multiple scattering. This heating term is minimized if $\beta_\perp$ is small (strong-focusing) and $L_R$ is large (a low-Z absorber).

The equation for energy cooling is:
Energy-cooling requires that $\frac{\partial}{\partial E} \left( \frac{dE}{ds} \right) > 0$. The energy loss function, $dE/\rho$, is rapidly decreasing with energy for $E < 0.2$ GeV (and therefore heating), but is slightly increasing (cooling) for $E > 0.3$ GeV. This natural cooling is ineffective; but $\frac{\partial}{\partial E} \left( \frac{dE}{ds} \right)$ can be increased by placing a transverse variation in absorber density or a wedge absorber where position is energy-dependent. (This variation is used in two modes: a weak variation to balance cooling rates, or a thick wedge to transfer phase space.) The sum of cooling rates is invariant:

$$\frac{1}{E_{cool_x}} + \frac{1}{E_{cool_y}} + \frac{1}{E_{cool_{\Delta E}}} = \text{Constant} \approx \frac{2}{E_\mu},$$

where $E_{cool}$ is the total energy loss needed to obtain an e-folding of cooling and $E_\mu$ is the $\mu$ energy.

In the long-pathlength Gaussian-distribution limit, the heating term or energy straggling term is given by:

$$\frac{d(E\Delta E)^2}{ds} = 4\pi(r m e^2)^2 \rho \gamma^2(1-\beta^2/2),$$

where $N_0$ is Avogadro's number and $\rho$ is the density. Since this increases as $\gamma^2$, and the cooling system size scales as $\gamma$, cooling at low energies is desired.

To obtain energy cooling and to minimize energy straggling, we require cooling at low relativistic energies ($E_\mu \sim 300$ MeV). For optimum transverse cooling, the ideal absorber is itself a strong focussing lens which maintains small beam size over extended lengths, and a low-Z material. In this design, we use Be($Z=4$) or Li($Z=3$) current-carrying rods, where the high current provides strong radial focussing. For Be, $Z=4$, $A=9$, $dE_\mu/\rho = 3$ MeV/cm, and $\rho = 1.85$ gm/cm$^3$.

The beam cooling system reduces transverse emittances by more than two orders of magnitude (from 0.01 to $3 \times 10^{-5}$ m-rad), and reduces longitudinal emittance by more than an order of magnitude. This cooling is obtained in a series of cooling cells, with the initial cells reducing the energy toward the cooling optimum of 300 MeV. A typical cooling cell consists of a focusing cooling rod.
(~0.7 m long for Be, ~2.1 m for Li) which reduces the central energy by ~200 MeV, followed by a ~200 MeV linac (20—40 m at 10—5 MeV/m), with optical matching sections (~40—60 m total cell length) (see fig. 3). Small angle bends introduce a dispersion (position-dependence on energy), and wedge absorbers (or density gradients) introduce an energy loss dependence on beam energy. Bends are also used to provide path length dependence on momentum, in order to compress the bunch lengths. The cell parameters are adjusted to optimal transverse and longitudinal cooling rates, and cooling by a 6-D factor of ~4 is obtained in each cell. ~15–20 such cells (~800 m) are needed in the complete machine.

From equation 1, we find a limit to transverse cooling when the multiple scattering balances the cooling, at $\varepsilon_n = 10^{-2} \beta_\perp$ for Be. The value of $\beta_\perp$ in the Be rod is limited by the peak focusing field to $\beta_\perp \sim 0.01 m$, obtaining $\varepsilon_n \sim 10^4 m$-rad. This is a factor of ~3 above the emittance goal of Table 1. The additional factor can be obtained by cooling more than necessary longitudinally and exchanging phase-space with transverse dimensions in a thick wedge absorber. MCM simulations have demonstrated that the desired cooling and phase-space exchange can be obtained.

In the present scenario, we cool only with ionization cooling in conducting Be (or Li) rods, along with phase space exchange, and that is sufficient for high luminosity. However, other techniques (such as ionization cooling in focusing transports or rings, or using plasma lenses, or high-frequency "optical" stochastic cooling) may permit improvements, and are being studied.

**ACCELERATION AND COLLISIONS**

Following cooling and initial bunch compression to ~1–3m bunch lengths, the beams are accelerated to full energy (2 TeV). A full-energy linac would work, but it would be costly and does not use our ability to recirculate $\mu$'s. A recirculating linac (RLA) like CEBAF can accelerate beam to full energy in 10—20 recirculations, using only 200—100 GeV of linac, but requiring 20—40 return arcs. The $\mu$-bunches would be compressed on each of the return arcs, to a length of 0.003m at full energy. A cascade of RLA stages (i. e., 1—10, 10—100 and 100—2000 GeV), with rf frequency increasing as bunch length decreases, may be used. Rapid-cycling synchrotrons, or hybrid devices, are also possible. (A low-cost scenario requiring only 20 GeV of rf, using an injector and three RLA stages with rapid-cycling in the last stage, has also been developed.) The cooling and acceleration cycle is timed so that less than ~half the initial $\mu$'s decay. (In Table 1, we allow a factor of 3 in total losses.)

After acceleration, the $\mu^+$ and $\mu^-$ bunches are injected into the 2-TeV superconducting storage ring (~1-km radius), with collisions in one or two low-$\beta^*$ interaction areas. The beam size at collision is $r = (\varepsilon_n \beta_\perp \gamma)^{1/2} \sim 2\mu$m, similar to
hadron collider values. The bunches circulate for \( \sim 300B \) turns before decay, where \( B \) is the mean bending field in T. (This is 150B luminosity turns, a factor of two smaller since both beams decay.) The design is restricted by \( \mu \) decay within the rings (\( \mu \to e \nu \nu \)), which produces 1/3-energy electrons which radiate and travel to the inside of the ring dipoles. This energy could be intercepted by a liner inside the magnets, or specially designed C-dipoles could be used and the electrons intercepted in an external absorber. The design constraints may limit \( B \); we have chosen \( B = 6T \) (900 turns) in the present case. \( \mu \)-decays in the interaction areas will also provide some background levels in detectors. The limitations in detector design are being studied.

**COMMENTS AND CONCLUSIONS**

We have presented a candidate design for a high-energy high-luminosity \( \mu^+\mu^- \) collider. The critical features of the scenario (\( \pi \)-collection and decay, phase space compression and cooling) have been modeled with Monte Carlo simulations, as well as by analytical methods. The design uses practical components and concepts within existing technical capabilities, and with requirements within the size and scope of existing facilities (i.e., Fermilab, CERN, HERA). Since we have confined ourselves to existing technology, we can initiate cost estimates, and costs similar to similar-sized facilities are expected.

The 4 TeV energy was set as a benchmark goal for the high-energy frontier. The \( \mu^+\mu^- \) collider concept naturally increases in luminosity with energy. One factor of \( E_{\mu} \) increase results from emittance adiabatic damping. If the injector size/cost is allowed to increase as \( E_{\mu} \), then beam intensities increase as \( E_{\mu}^2 \), and \( L \) increases by at least another factor of \( E_{\mu} \). Smaller interaction-region \( \beta' \) should also be possible through longitudinal adiabatic damping, permitting some further enhancement. This increase of luminosity with energy by at least a factor of \( E_{\mu}^2 \) should be followed up to the 100 TeV scale, beyond which \( \mu \) synchrotron radiation becomes large. Radiation damping would then initially permit further improvements.

Lower energy machines (100+ GeV Higgs, top factories, etc.) are also possible, as proposed by D. Cline.\(^{16}\) These require specific physics motivations, but could result in important discoveries.

The present scenario is simply a first proof-of-principle calculation of the design concept, using realistic components. Much further optimization and design and concept development is needed. We are forming and recruiting a collaboration for further development of the \( \mu^+\mu^- \) collider concept. (Interested readers are invited to contact one of the authors for further discussion.) We next describe some of the topics needing further development.

The scenario for production and collection of muons needs further study and
improvement. More detailed, and hopefully more accurate, production calculations should be implemented. Many options for targeting, collection and transport, as well as variations on the rf debunching and compression scenario, should be explored. A critical parameter is the rf-debuncher gradient; a higher gradient would permit a more compact facility. Variations on initial proton and π collection energies should be considered, including the possibility of low-energy stopped π sources, and optimizations should be obtained. (H. Daniel\textsuperscript{17} has investigated μ-cooling at low (thermal) energies.)

The ionization cooling scenario has only been developed in broad outline form, and needs to be optimized in full detail, and verified by more complete simulations. The detailed design is strongly dependent on rf gradient and linear and nonlinear optics optimization.

The bunch-compression and acceleration scenario also needs to be optimized and simulated. Variations such as rapid-cycling synchrotrons and fixed-field alternating gradient (FFAG) machines should be considered. Complete lattices are needed. A complete lattice is also needed for the full-energy collider, integrating the interaction-regions with the high-field arcs, and including modifications needed to accomodate μ-decay in the arcs and near the IR's. Tracking of the complete lattice for a muon lifetime should be completed, and any instability limits should be identified.

It would be desirable to collide polarized beams. Muons are naturally polarized when produced in π-decay, however, we do not yet know if that polarization can be maintained through a debunching, cooling, acceleration and collision cycle. Further study is needed.

The luminosity estimate of the present scenario should not be considered as an absolute limit. Luminosity can be readily increased by reducing losses, or by increasing the source repetition rate (increasing the source to above KAON-class intensity).

To initiate practical implementation, some experimental explorations are needed. A first demonstration of ionization cooling is essential. A collaboration is being formed (and recruited; call if interested) to plan an initial cooling experiment. In this experiment, a diffuse muon beam at ~0.5 GeV would be focussed into a Li (or Be) rod, in a configuration similar to the cooling cell of Fig. 3. Since ionization cooling is intensity-independent, only a low-intensity external beam is needed. Beam formation optics, including strong (quad or solenoid) focussing into the Li rod, a high-current power supply for lens activation of the rod, and accurate beam diagnostics are also required.

Other experiments may be needed. More accurate measurements of π-production under the same conditions used in the collider would be useful, and could be important in developing an accurately optimized design scenario. rf acceleration development would also be desirable, both in the low-frequency rf
systems needed in the debuncher and cooling and in the high-frequency rf needed in the accelerator and collider.

Much study and development is needed and this initial scenario will be greatly changed and (hopefully) improved before implementation. However, we believe that the improvements reported here have transformed the \( \mu^+\mu^- \) collider concept into a real possibility, which will permit exploration of elementary particle physics at energy frontiers beyond the reach of currently existing and proposed electron and hadron colliders.

**Acknowledgments**

We acknowledge extremely important contributions from our colleagues, especially D. Cline, F. Mills, A. Chao, A. Ruggiero, A. Sessler, S. Chattopadhyay, J. D. Bjorken, W. Barletta, D. Douglas, R. Noble, D. Winn, S. O'Day, Y. Y. Lee, I. Stumer and C. Taylor.

**References**

Table 1: Parameter list for 4 TeV $\mu^+\mu^-$ Collider

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy per beam</td>
<td>$E_{\mu}$</td>
<td>2 TeV</td>
</tr>
<tr>
<td>Luminosity</td>
<td>$L=f_0 n_s n_b N_{\mu}^2/4\pi\alpha^2$</td>
<td>$3\times10^{34}$ cm$^{-2}$s$^{-1}$</td>
</tr>
</tbody>
</table>

**Source Parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proton energy</td>
<td>30 GeV</td>
</tr>
<tr>
<td>Protons/pulse</td>
<td>$2\times10^{13}$</td>
</tr>
<tr>
<td>Pulse rate</td>
<td>10 Hz</td>
</tr>
<tr>
<td>$\mu$-production acceptance</td>
<td>.15</td>
</tr>
<tr>
<td>$\mu$-survival allowance</td>
<td>.33</td>
</tr>
</tbody>
</table>

**Collider Parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of $\mu$ /bunch</td>
<td>$1.5\times10^{12}$</td>
</tr>
<tr>
<td>Number of bunches</td>
<td>1</td>
</tr>
<tr>
<td>Storage turns</td>
<td>900</td>
</tr>
<tr>
<td>Normalized emittance</td>
<td>$3\times10^{-5}$ m-rad</td>
</tr>
<tr>
<td>$\mu$-beam emittance</td>
<td>$1.5\times10^{-6}$ m-rad</td>
</tr>
<tr>
<td>Interaction focus</td>
<td>0.3 cm</td>
</tr>
<tr>
<td>Beam size at interaction</td>
<td>2.1 $\mu$m</td>
</tr>
</tbody>
</table>
Figure 1: Overview of the $\mu^+\mu^-$ collider system, showing a muon ($\mu$) source based on a high-intensity rapid-cycling proton synchrotron, with the protons producing pions ($\pi$'s) in a target, and the $\mu$'s are collected from subsequent $\pi$ decay. The source is followed by a $\mu$-cooling system, and an accelerating system of recirculating linac(s) and/or rapid-cycling synchrotron(s), feeding $\mu^+$ and $\mu^-$ bunches into a superconducting storage-ring collider for multiturn high-energy collisions. The entire process cycles at 10 Hz.
Figure 2: Overview of the $\pi \to \mu$ production transport line. Each proton bunch collides into a target, producing large numbers of $\pi$'s ($\sim 1 \pi$/interacting p) over a broad energy and angular range ($E_\pi = 0\text{--}4 \text{ GeV}, p_\perp < 0.5 \text{ GeV/c}$). The target is followed by a Li lens, which collects the $\pi$'s into a transport line, which accepts a large energy width ($2\times 1 \text{ GeV}$) and has a large transverse acceptance ($p_\perp < 0.4 \text{ GeV/c}$). This line is sufficiently long to insure $\pi \to \mu$ decay, plus debunching, in which the energy-dependent particle speeds spread the beam longitudinally, while a nonlinear rf system flattens the momentum spread. The $\mu$-beam is then matched into a beam-cooling system.

Figure 3: Overview of a $\mu$-cooling (ionization cooling) cell. A typical cooling cell consists of a focusing cooling rod which reduces the central energy from 400 to 200 MeV, followed by a $\sim 200 \text{ MeV}$ linac, with appropriate optical matching. Cooling by a 3-D factor of $\sim 4$ is obtained in each cell, and 15--20 such cells ($\sim 1 \text{ km}$) are needed in the complete machine. For energy cooling, we introduce a dispersion (position-dependence on momentum), and use an absorber with a density gradient or wedge shape. This arrangement permits enhancement of energy-cooling$^1$, and the parameters are adjusted to obtain optimal transverse and longitudinal cooling.