



# **Higgs Factory Muon Collider**

**FROM A NEUTRINO FACTORY  
TO A HIGGS FACTORY**

**For SNOWMASS 2001**

**Edited by David B. Cline  
and Gait G. Hanson**

<http://www.physics.ucla.edu/snowmass>

**UCLA Center for Advanced Accelerators**

**For the Muon Collider/Neutrino Factory Collaboration**

## OVERVIEW AND SUMMARY

David B. Cline, University of California, Los Angeles, 90095, USA  
Gail Hanson, Indiana University, Bloomington, IN 47405

### 1. PHYSICS OF A HIGGS FACTORY

The major purpose of the Higgs factory is to find the exact Higgs mass (or masses) and then measure the important parameters, such as the width(s) and the common and rare branching fractions [2,3,4,5]. This concept is based mainly on a relatively low-mass Higgs (below 300 GeV). In the low-mass region (below 150 GeV), the Higgs could well be supersymmetric (SUSY), and the width measurement will be crucial. Above 150 GeV, the Higgs could be more of a standard-model type. However, this will once again lead to the issue of what keeps the scalar system stable, which might be answered by the study of rare decays of the Higgs particles (in progress). In the near future, there could be evidence for the Higgs mass obtained from precise electroweak parameter measurements and later from the LHC. (Recent results from LEP2 are not fully conclusive [6].) This will be a crucial input for the development of the Higgs factory. In addition, if Nature is supersymmetric, there will be additional SUSY-Higgs particles to study and, thus, the Higgs factory concept will include the search for and study of the SUSY Higgses. This is an experimental issue; theory can only take us just so far!

From all we now know about elementary particle physics, the scalar or SUSY scalar sector is the key to future understanding. A complete understanding of this sector is really the goal of the Higgs factory and of nearly all elementary particle physics these days. The Higgs factory is designed to first give the exact Higgs mass using an energy scan and then measure the general properties of the Higgs, such as the total width, largest branching fractions, etc. It would produce  $10^4 \rightarrow 4 \times 10^3$  Higgs/yr and could investigate rare branching modes. If there are more Higgs', the Higgs factory would be used to scan and find and study these in detail as well (see Table 1).

We expect the super-collider LHC to extract the signal from background (i.e., seeing either  $h^0 \rightarrow \gamma b$  or the very rare  $h^0 \rightarrow \gamma b$  in this mass range, since  $h \rightarrow b\bar{b}$  is swamped by hadronic background). However, detectors for the LHC are designed to extract this signal. In this low mass region, the Higgs is also expected to be a fairly narrow resonance and, thus, the signal should stand out clearly from the background from

$$\mu^+ \mu^- \rightarrow b\bar{b}. \quad (1)$$

For masses above 180 GeV, the dominant Higgs decay is

$$h^0 \rightarrow W^+ W^- \text{ or } Z^0 Z^0, \quad (2)$$

and the LHC should easily detect this Higgs particle [4,5]. Thus the  $\mu^+ \mu^-$  collider is better adapted for the low mass region.

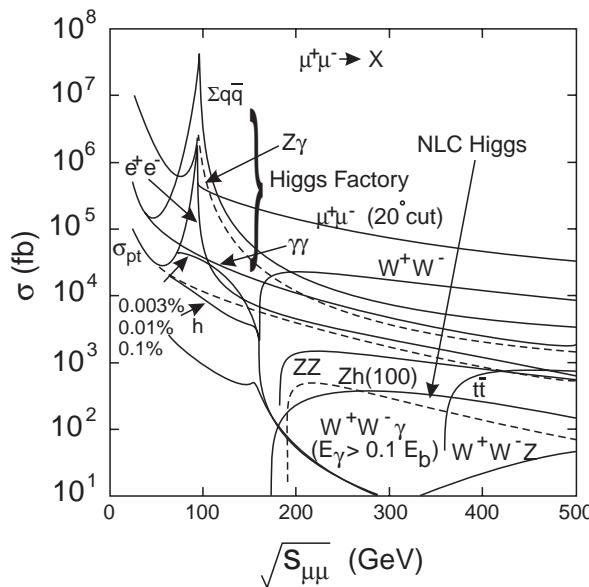
In Fig. 1, we show a comparison of the Higgs factory  $\mu^+ \mu^-$  collider and an  $e^+ e^-$  collider (NLC) that could also study the Higgs [5]. Note the very great differences in cross sections, indicating that the  $e^+ e^-$  collider must have very high luminosity. There is also a possibility to search for CP violation in the Higgs sector as we discussed at a recent UCLA workshop [6]. The machine research reported here came from the following: V. Balbekov, A. R. Fornow, Y. Fukui, A. Garren, C. Johnstone, D. Neuffer, A. Sessler, and D. Summers.

**TABLE 1. Logic of Detailed Study of the Higgs Sector**

---

If particles in the scalar sector are ever discovered, it will be essential to determine their properties, which will give direct information about the nature of the particle and the underlying theory. Three simple examples can be cited:

- Suppose a Higgs-like particle is discovered with mass 120 GeV. This could either be the standard model (SM) Higgs or an SM Higgs. A measurement of the width of the state would presumably tell the difference. However, the SM width is 5 MeV, a formidable measurement!
  - Suppose a Higgs-like particle is discovered with a mass of 160 GeV. This is presumably beyond the MSSM bound, but it could be an NMSSM or an SMHiggs. A measurement of the width could presumably resolve the issue.
  - Suppose a Higgs-like particle of mass 180 GeV is discovered. This is presumably even beyond the NMSSM limits. If this is an SM Higgs, can we learn more by the study of the rare decay modes?
- 



**Figure 1.** The cross sections as a function of energy for  $e^+ e^-$  and  $\mu^+ \mu^-$  interactions producing a Higgs boson and other systems.

## 2. A SCHEME TO CONVERT A NEUTRINO FACTORY TO A HIGGS FACTORY

We consider the possibility of staging the Neutrino Factory [6] Muon Collider program by converting as much as possible of a Neutrino Factory to a Higgs Factory. We take the example of Study II of the MC group to use BNL for the Neutrino Factory [8]. Figure 2 shows this current scheme. We show in Figure 2 that the addition of 3 rings may lead to the required beam properties and cooling for the Higgs Factory. The cooling is a major challenge. In Figure 3 we show the required longitudinal and transverse emittance for the Higgs Factory [6].

## 3. PARAMETERS OF THE HIGGS FACTORY

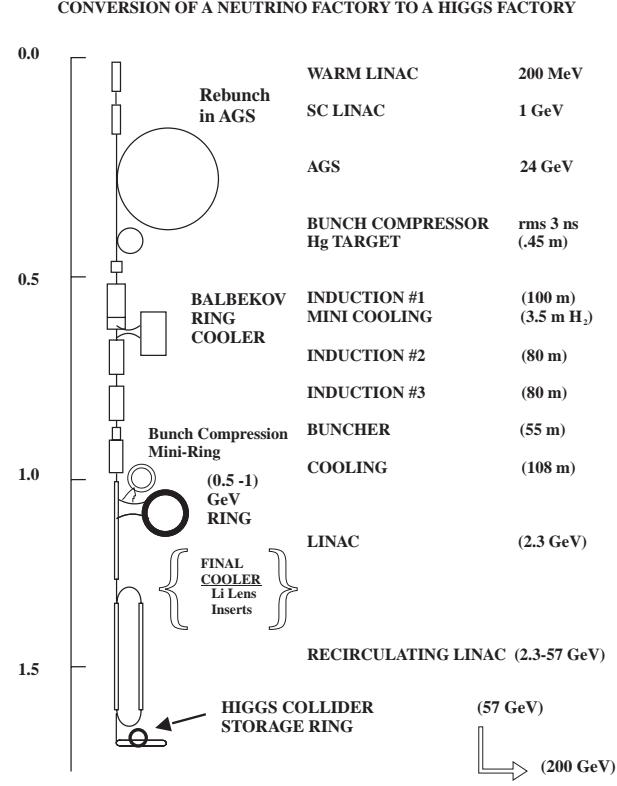
The projected parameters of the Muon Collider Higgs Factory are given in Table 2 [6,7]. The key to achieving these parameters is the final cooler as described in the next section.

## 4. BALBEKOV RING COOLER AND A STORAGE RING FINAL COOLER

The keys to the conversion of a Neutrino Factory to a Higgs Factory are shown in Figure 2. We consider these two rings to be:

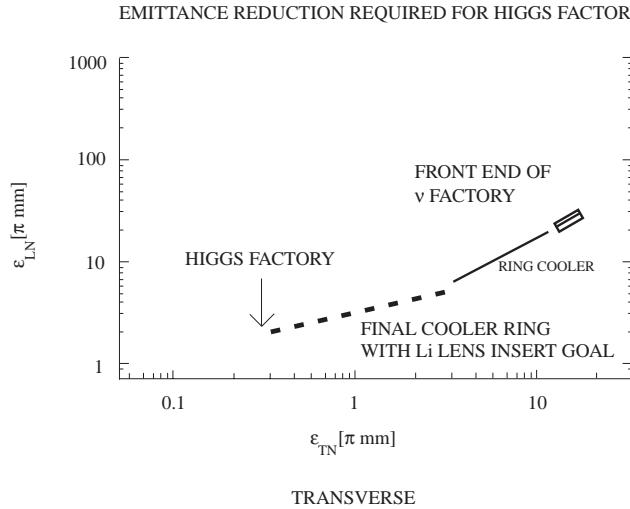
- A Balbekov Cooler Ring [9]

## 2. A Storage Ring Cooler [10,11]



**Figure 2.** Scheme for converting a Neutrino Factory to a Higgs Factory.

The basic concept is that the Balbekov ring reduces the emittance as shown in Figure 3 to a level that allows the beam to be injected into a storage ring cooler (Figure 5). This final cooler ring could have lithium lens inserts or hydrogen wedges as shown in Figure 5.



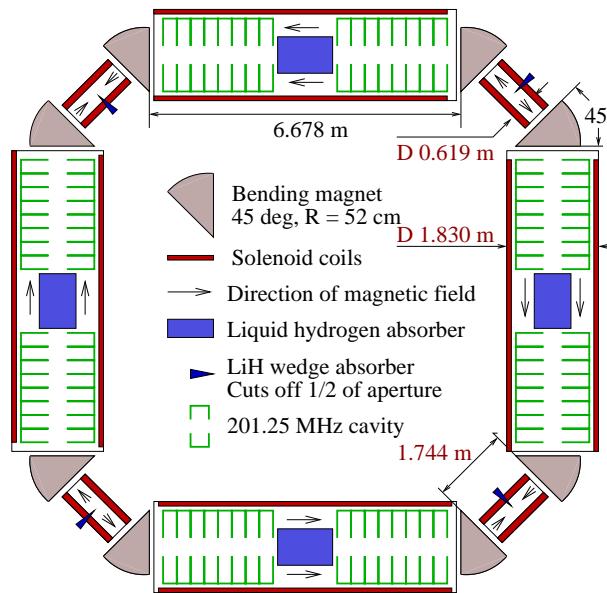
**Figure 3.** Emittance reduction needed for a Higgs Factory.

We believe that Balbekov's design of a low-energy cooling ring shows a very promising approach, and the use of such rings at higher energies may also be very useful. We are presently exploring this possibility [11]. A conceptual example of a cooling module is shown in Figure 5. Four such modules could be placed in each long straight section of a 1 GeV 300 m racetrack shaped storage ring, taking and restoring 240 MeV from the

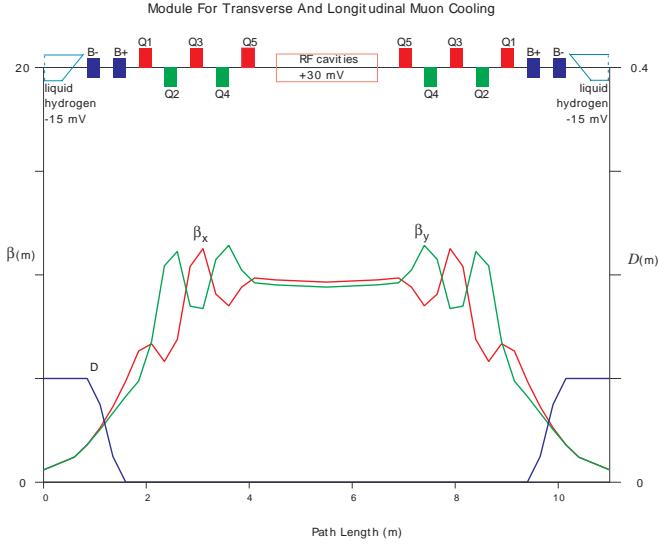
particles each turn. Thus each mode may be damped by a factor of 1/3 in about 6 turns, or one quarter of the muon lifetime. We thank K. Lee and D. MacLaughlan-Dumes for their help.

**TABLE 2.** Baseline parameters for Higgs factory muon collider. Higgs/year assumes a cross-section of  $5 \times 10^4$  fb, Higgs width of 2.7 MeV, 1 year =  $10^7$  s [7].

COM energy (TeV)	0.1
$p$ energy (GeV)	16
$p$ 's/bunch	$5 \times 10^{13}$
Bunches/fill	2
Rep. rate (Hz)	15
$p$ power (MW)	4
$\mu$ /bunch	$4 \times 10^{12}$
$\mu$ power (MW)	1
Wall power (MW)	81
Collider circum. (m)	350
Ave bending field (T)	3
rms $\delta p/p$ (%)	0.01
$\beta^*$ (cm)	9.4
$\sigma_z$ (cm)	9.4
$\sigma_t$ spot ( $\mu$ m)	196
$\sigma_\theta$ IP (mrad)	2.1
Tune shift	0.022
$n_{\text{turns}}$ (effective)	450
Luminosity ( $\text{cm}^{-2} \text{s}^{-1}$ )	$2.2 \times 10^{31}$
Higgs/yr	$4 \times 10^3$



**Figure 4.** The Balbekov cooling ring.



**Figure 5.** Cooling module of a storage ring cooler.

## References

1. P. Igo-Kemenes, talk at LEPC open session (Nov. 3, 2000) on behalf of the LEP Higgs Working Group.
2. Cline, D., *Nucl. Instrum. Methods* **A350**, 24 (1994).
3. V. Barger, M.S. Berger, J. F. Gunion and T. Han, *Phys. Rept.* **286**, 1 (1997), V. Barger, M.S. Berger, J. F. Gunion and T. Han, *Phys. Rev. Lett.* **75**, 1462 (1995).
4. *Physics Potential and Development of  $\mu^+ \mu^-$  Colliders* (Proc., 2nd Wksp., Sausalito, CA, 1994), edited by D. B. Cline, AIP Conference Proceedings 352, New York, 1996.
5. *Physics Potential & Development of  $\mu^+ \mu^-$  Colliders* (Proc. 3rd Intl. Conf., San Francisco, Dec. 1995), edited by D. B. Cline, *Nucl. Phys. B* (PS), **51A** (1996).
6. Higgs Factory Workshop, UCLA, February 28–March 1, 2001 (<http://www.physics.ucla.edu/conferences/HIGGS/higgs.html>).
7. M. Ankenbrandt et al., *Phys. Rev. Special Topics Accel. Beams* **2**, 081001 (1990).
8. BNL Neutrino Factory Feasibility Study II, to be published.
9. V. Balbekov, “Possibility of using a ring accelerator for ionization cooling of muons”, Proc. 1999 part. Acc. Conf., p. 315-7; V. Balbekov, “Ring cooler update”, NuMu note 189, 2001.
10. Al Garren, private communication.
11. Report on Higgs Factory for Snowmass 2001, to be published.

# Physics of Higgs Factories (contribution to Snowmass)

V. Barger<sup>1</sup>, M. S. Berger<sup>2</sup>, J. F. Gunion<sup>3</sup>, and T. Han<sup>1</sup>

<sup>1</sup>*Physics Department, University of Wisconsin, Madison, WI 53706*

<sup>2</sup>*Physics Department, Indiana University, Bloomington, IN 47405*

<sup>3</sup>*Physics Department, University of California, Davis, CA 95616*

**Abstract.** We outline the unique role of a muon collider as a Higgs factory for Higgs boson resonance production in the  $s$ -channel. Physics examples include: the precision measurements of the Higgs mass and total width, and the resulting ability to discriminate between the SM-like Higgs bosons of different models such as between a light SM Higgs boson and the light Higgs boson of the MSSM; the determination of the spin and coupling via the  $h \rightarrow \tau^-\tau^+$  decay mode; differentiation of two nearly degenerate heavy Higgs bosons by an energy scan; and the ability to explore a general extended Higgs sector, possibly with CP-violating couplings. The muon collider Higgs factory could perform measurements that would be highly complementary to Higgs studies at the LHC and LC; it would be likely to play a very crucial role in fully understanding the Higgs sector.

## I INTRODUCTION

A muon collider with c.m. energy centered at the Higgs boson mass offers a unique opportunity to produce Higgs bosons in the  $s$ -channel and thereby measure the Higgs masses, total width and several partial widths to very high precision. In the event that only a SM-like Higgs boson is discovered and its properties measured at the Tevatron, the LHC, and a LC, it may prove essential to build a muon collider to fully explore the Higgs sector. In particular, the very narrow width of a Standard Model (SM) Higgs bosons cannot be measured *directly* at the Large Hadron Collider (LHC) or at a future Linear Collider (LC). Furthermore, there are regions of parameter space for which it will be impossible for either the LHC or a LC to discover the heavier Higgs bosons of supersymmetry or, in the case of a general two-Higgs-doublet or more extended model, Higgs bosons of any mass with small or zero  $VV$  coupling.

The value of a future Higgs factory should be discussed in light of recent experimental data. While by no means definitive, recent experimental results point in promising directions for Higgs factories. First, there is the  $\gtrsim 2\sigma$  statistical evidence from LEP [1–5] for a Higgs boson near  $m_H \simeq 115$  GeV. Such a mass is in the optimal range for study at a Higgs factory and it is for such a low mass that the muon collider factory option would add the most information to data from the LHC and a LC. First, 115 GeV is sufficiently above the  $Z$ -pole that the background from  $Z$  production and decay to  $b\bar{b}$  is not so large, and the mass is sufficiently below the  $WW^*$  threshold that the decay width remains small and the ability of the muon collider to achieve a very narrow beam energy spread can be exploited. Second, it is for masses below 120 GeV that the LC will have difficulty getting a precision measurement of the Higgs to  $WW^*$  branching ratio, resulting in large error for the indirect determination of the total Higgs width. Of course, a Higgs boson in this mass range, and having substantial  $VV$  coupling, is also the most natural interpretation of current precision electroweak data. On the theoretical side, a Higgs mass of  $\sim 115$  GeV is very suggestive of supersymmetry. In the Minimal Supersymmetric Model (MSSM) such a mass is near the theoretical upper limit of  $m_H < 130$  GeV, and would indicate a value of the supersymmetry parameter  $\tan\beta$  substantially above 1 (assuming stop masses  $\lesssim 1$  TeV).

A Higgs with mass  $\sim 115$  GeV in the context of a large- $\tan\beta$  supersymmetry scenario would mesh nicely with recent evidence for an anomalous magnetic moment of the muon [6] that deviates from the Standard Model prediction. The  $2.6\sigma$  discrepancy is naturally accounted for provided  $\tan\beta$  is relatively large (and superparticle

masses are not too heavy). More specifically, a supersymmetric interpretation of this discrepancy with the SM prediction implies the following relationship between the mass scale  $\tilde{m}$  of supersymmetric particles contributing to the one-loop anomalous magnetic moment diagram and  $\tan \beta$  [7],

$$\tan \beta \left( \frac{100 \text{ GeV}}{\tilde{m}} \right)^2 = 3.3 \pm 1.3$$

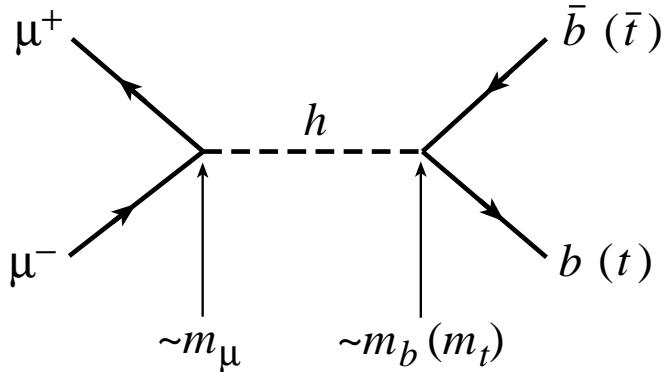
Furthermore, if the anomalous magnetic moment is explained by supersymmetry the value of the Higgs mass parameter  $\mu$  of supersymmetric models has a sign which is consistent with the constraints from the radiative decays,  $b \rightarrow s\gamma$ . Thus, a consistent picture begins to emerge suggesting low-energy supersymmetry with a Higgs boson in the predicted mass range.

While these recent experimental data are not definitive, they do point to an interesting scenario whereby a muon collider might prove essential to the understanding of the Higgs sector of a supersymmetric model. The muon collider could perform at least two measurements crucial for detailing a SUSY Higgs sector: (1) accurately measuring the properties of a light SM-like Higgs boson and distinguishing it from a supersymmetric Higgs bosons, and (2) discovering heavy Higgs bosons of supersymmetry and accurately measuring their properties.

## II MUON COLLIDERS

Muon colliders have a number of unique features that make them attractive candidates for future accelerators [8]. The most important and fundamental of these derive from the large mass of the muon in comparison to that of the electron. This leads to: a) the possibility of extremely narrow beam energy spreads, especially at beam energies below 100 GeV; b) the possibility of accelerators with very high energy; c) the possibility of employing storage rings at high energy; d) the possibility of using decays of accelerated muons to provide a high luminosity source of neutrinos (under active consideration as reviewed elsewhere); e) increased potential for probing physics in which couplings increase with mass (as does the SM  $h_{SM} f\bar{f}$  coupling).

Here our focus is on the Higgs sector. The relatively large mass of the muon compared to the mass of the electron means that the coupling of Higgs bosons to  $\mu^+ \mu^-$  is very much larger than to  $e^+ e^-$ , implying much larger  $s$ -channel Higgs production rates at a muon collider as compared to an electron collider [see Fig. 1]. For Higgs bosons with a very small MeV scale width, such as a light SM Higgs boson, production rates in the  $s$ -channel are further enhanced by the muon collider's ability to achieve beam energy spreads comparable to the tiny Higgs width. In addition, there is little bremsstrahlung, and the beam energy can be tuned to one part in a million through continuous spin-rotation measurements [9]. Due to these important qualitative differences between the two types of machines, only muon colliders can be advocated as potential  $s$ -channel Higgs factories capable of determining the mass and decay width of a Higgs boson to very high precision [10,11]. High rates of Higgs production at  $e^+ e^-$  colliders rely on substantial  $V V$ Higgs coupling for the  $Z + \text{Higgs}$  (Higgs-strahlung) or  $WW \rightarrow \text{Higgs}$  ( $WW$  fusion) reactions. In contrast, a  $\mu^+ \mu^-$  collider can provide a factory for producing a Higgs boson with little or no  $V V$  coupling so long as it has SM-like (or enhanced)  $\mu^+ \mu^-$  couplings.



**FIGURE 1.** Feynman diagram for  $s$ -channel production of a Higgs boson.

Of course, there is a trade-off between small beam energy spread,  $\delta E/E = R$ , and luminosity. Current estimates for yearly integrated luminosities (using  $\mathcal{L} = 1 \times 10^{32} \text{ cm}^{-2} \text{s}^{-1}$  as implying  $L = 1 \text{ fb}^{-1}/\text{yr}$ ) are:

$L_{\text{year}} \gtrsim 0.1, 0.22, 1 \text{ fb}^{-1}$  at  $\sqrt{s} \sim 100 \text{ GeV}$  for beam energy resolutions of  $R = 0.003\%, 0.01\%, 0.1\%$ , respectively;  $L_{\text{year}} \sim 2, 6, 10 \text{ fb}^{-1}$  at  $\sqrt{s} \sim 200, 350, 400 \text{ GeV}$ , respectively, for  $R \sim 0.1\%$ . Despite this, studies show that for small Higgs width the  $s$ -channel production rate (and statistical significance over background) is maximized by choosing  $R$  to be such that  $\sigma_{\sqrt{s}} \lesssim \Gamma_h^{\text{tot}}$ . In particular, in the SM context this corresponds to  $R \sim 0.003\%$  for  $m_{h_{SM}} \lesssim 120 \text{ GeV}$ .

If the  $m_h \sim 115 \text{ GeV}$  LEP signal is real or if the interpretation of the precision electroweak data as an indication of a light Higgs boson (with substantial  $VV$  coupling) is valid,<sup>1)</sup> then both  $e^+e^-$  and  $\mu^+\mu^-$  colliders will be valuable. In this scenario the Higgs boson would have been discovered at a previous higher energy collider (possibly a muon collider running at high energy), and then the Higgs factory would be built with a center-of-mass energy precisely tuned to the Higgs boson mass.<sup>2)</sup> The most likely scenario is that the Higgs boson is discovered at the LHC via gluon fusion ( $gg \rightarrow H$ ) or perhaps earlier at the Tevatron via associated production ( $q\bar{q} \rightarrow WH, t\bar{t}H$ ), and its mass is determined to an accuracy of about 100 MeV. If a linear collider has also observed the Higgs via the Higgs-strahlung process ( $e^+e^- \rightarrow ZH$ ), one might know the Higgs boson mass to better than 50 MeV with an integrated luminosity of  $500 \text{ fb}^{-1}$ . The muon collider would be optimized to run at  $\sqrt{s} \approx m_H$ , and this center-of-mass energy would be varied over a narrow range so as to scan over the Higgs resonance (see Fig. 2 below).

### III HIGGS PRODUCTION

The production of a Higgs boson (generically denoted  $h$ ) in the  $s$ -channel with interesting rates is a unique feature of a muon collider [10,11]. The resonance cross section is

$$\sigma_h(\sqrt{s}) = \frac{4\pi\Gamma(h \rightarrow \mu\bar{\mu})\Gamma(h \rightarrow X)}{(s - m_h^2)^2 + m_h^2(\Gamma_h^{\text{tot}})^2}. \quad (1)$$

In practice, however, there is a Gaussian spread ( $\sigma_{\sqrt{s}}$ ) to the center-of-mass energy and one must compute the effective  $s$ -channel Higgs cross section after convolution assuming some given central value of  $\sqrt{s}$ :

$$\overline{\sigma}_h(\sqrt{s}) = \frac{1}{\sqrt{2\pi}\sigma_{\sqrt{s}}} \int \sigma_h(\sqrt{\hat{s}}) \exp\left[\frac{-\left(\sqrt{\hat{s}} - \sqrt{s}\right)^2}{2\sigma_{\sqrt{s}}^2}\right] d\sqrt{\hat{s}} \underset{\sqrt{\hat{s}}=m_h}{\simeq} \frac{4\pi}{m_h^2} \frac{\text{BF}(h \rightarrow \mu\bar{\mu})\text{BF}(h \rightarrow X)}{\left[1 + \frac{8}{\pi} \left(\frac{\sigma_{\sqrt{s}}}{\Gamma_h^{\text{tot}}}\right)^2\right]^{1/2}}. \quad (2)$$

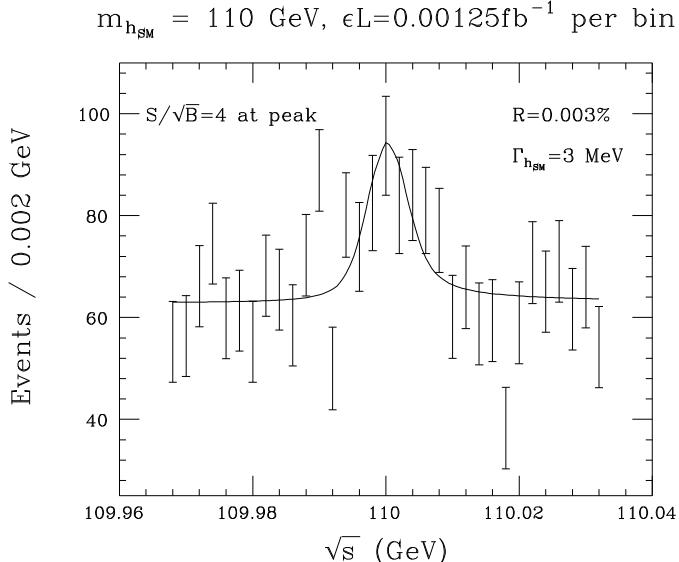
It is convenient to express  $\sigma_{\sqrt{s}}$  in terms of the root-mean-square (rms) Gaussian spread of the energy of an individual beam,  $R$ :

$$\sigma_{\sqrt{s}} = (2 \text{ MeV}) \left( \frac{R}{0.003\%} \right) \left( \frac{\sqrt{s}}{100 \text{ GeV}} \right). \quad (3)$$

From Eq. (1), it is apparent that a resolution  $\sigma_{\sqrt{s}} \lesssim \Gamma_h^{\text{tot}}$  is needed to be sensitive to the Higgs width. Further, Eq. (2) implies that  $\overline{\sigma}_h \propto 1/\sigma_{\sqrt{s}}$  for  $\sigma_{\sqrt{s}} > \Gamma_h^{\text{tot}}$  and that large event rates are only possible if  $\Gamma_h^{\text{tot}}$  is not so large that  $\text{BF}(h \rightarrow \mu\bar{\mu})$  is extremely suppressed. The width of a light SM-like Higgs is very small (e.g. a few MeV for  $m_{h_{SM}} \sim 110 \text{ GeV}$ ), implying the need for  $R$  values as small as  $\sim 0.003\%$  for studying a light SM-like  $h$ . Fig. 2 illustrates the result for the SM Higgs boson of an initial centering scan over  $\sqrt{s}$  values in the vicinity of  $m_{h_{SM}} = 110 \text{ GeV}$ . This figure dramatizes: a) that the beam energy spread must be very small because of the very small  $\Gamma_{h_{SM}}^{\text{tot}}$  (when  $m_{h_{SM}}$  is small enough that the  $WW^*$  decay mode is highly suppressed); b) that we require the very accurate *in situ* determination of the beam energy to one part in a million through the spin precession of the muon noted earlier in order to perform the scan and then center on  $\sqrt{s} = m_{h_{SM}}$  with a high degree of stability.

<sup>1)</sup> Even in a two-doublet extension of the minimal one-doublet SM Higgs sector, parameters can be chosen so that the only light Higgs boson has no  $VV$  coupling and yet good agreement with precision electroweak data maintained [20].

<sup>2)</sup> If the higher energy muon collider has already been constructed, this would simply require construction of a small storage ring tuned to the appropriate energy.



**FIGURE 2.** Number of events and statistical errors in the  $b\bar{b}$  final state as a function of  $\sqrt{s}$  in the vicinity of  $m_{h_{SM}} = 110$  GeV, assuming  $R = 0.003\%$ , and  $\epsilon L = 0.00125 \text{ fb}^{-1}$  at each data point.

If the  $h$  has SM-like couplings to  $WW$ , its width will grow rapidly for  $m_h > 2m_W$  and its  $s$ -channel production cross section will be severely suppressed by the resulting decrease of  $\text{BF}(h \rightarrow \mu\mu)$ . More generally, any  $h$  with SM-like or larger  $h\mu\mu$  coupling will retain a large  $s$ -channel production rate when  $m_h > 2m_W$  only if the  $hWW$  coupling becomes strongly suppressed relative to the  $h_{SM}WW$  coupling.

The general theoretical prediction within supersymmetric models is that the lightest supersymmetric Higgs boson  $h^0$  will be very similar to the  $h_{SM}$  when the other Higgs bosons are heavy. This ‘decoupling limit’ is very likely to arise if the masses of the supersymmetric particles are large (since the Higgs masses and the superparticle masses are typically similar in size for most boundary condition choices). Thus,  $h^0$  rates will be very similar to  $h_{SM}$  rates. In contrast, the heavier Higgs bosons in a typical supersymmetric model decouple from  $VV$  at large mass and remain reasonably narrow. As a result, their  $s$ -channel production rates remain large.

For a SM-like  $h$ , at  $\sqrt{s} = m_h \approx 115$  GeV and  $R = 0.003\%$ , the  $b\bar{b}$  final state rates are

$$\text{signal} \approx 10^4 \text{ events} \times L(\text{fb}^{-1}), \quad (4)$$

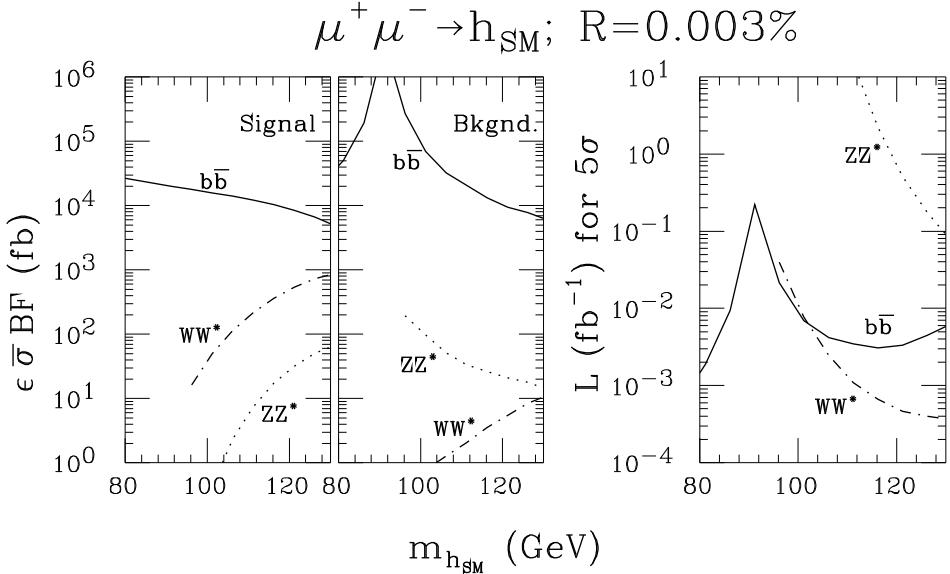
$$\text{background} \approx 10^4 \text{ events} \times L(\text{fb}^{-1}), \quad (5)$$

The SM Higgs cross sections and backgrounds are shown in Fig. 3 for  $R = 0.003\%$  and  $m_{h_{SM}}$  values such that the dominant decay mode is  $b\bar{b}$ .

#### IV THE MUON COLLIDER ROLE

An assessment of the need for a Higgs factory requires that one detail the unique capabilities of a muon collider versus the other possible future accelerators as well as comparing the abilities of all the machines to measure the same Higgs properties. Muon colliders and a Higgs factory in particular would only become operational after the LHC physics program is well-developed and quite possibly after a linear collider program is mature as well. So one important question is the following: if a SM-like Higgs boson and, possibly, important physics beyond the Standard Model have been discovered at the LHC and perhaps studied at a linear collider, what new information could a Higgs factory provide?

The  $s$ -channel production process allows one to determine the mass, total width, and the cross sections  $\bar{\sigma}_h(\mu^+\mu^- \rightarrow h \rightarrow X)$  for several final states  $X$  to very high precision. The Higgs mass, total width and the cross sections can be used to constrain the parameters of the Higgs sector. For example, in the MSSM their precise values will constrain the Higgs sector parameters  $m_{A^0}$  and  $\tan\beta$  (where  $\tan\beta$  is the ratio of the two



**FIGURE 3.** The SM Higgs cross sections and backgrounds in  $b\bar{b}$ ,  $WW^*$  and  $ZZ^*$ . Also shown is the luminosity needed for a 5 standard deviation detection in  $b\bar{b}$ . From Ref. [10].

vacuum expectation values (vevs) of the two Higgs doublets of the MSSM). The main question is whether these constraints will be a valuable addition to LHC and LC constraints.

The expectations for the luminosity available at linear colliders has risen steadily. The most recent studies assume an integrated luminosity of some  $500 \text{ fb}^{-1}$  corresponding to 1-2 years of running at a few  $\times 100 \text{ fb}^{-1}$  per year. This luminosity results in the production of greater than  $10^4$  Higgs bosons per year through the Bjorken Higgs-strahlung process,  $e^+e^- \rightarrow Zh$ , provided the Higgs boson is kinematically accessible. This is comparable or even better than can be achieved with the current machine parameters for a muon collider operating at the Higgs resonance; in fact, recent studies have described high-luminosity linear colliders as “Higgs factories,” though for the purposes of this report, we will reserve this term for muon colliders operating at the  $s$ -channel Higgs resonance.

A linear collider with such high luminosity can certainly perform quite accurate measurements of certain Higgs parameters such as the Higgs mass, couplings to gauge bosons, couplings to heavy quarks, etc. [18]. Precise measurements of the couplings of the Higgs boson to the Standard Model particles is an important test of the mass generation mechanism. In the Standard Model with one Higgs doublet, this coupling is proportional to the particle mass. In the more general case there can be mixing angles present in the couplings. Precision measurements of the couplings can distinguish the Standard Model Higgs boson from the SM-like Higgs boson typically present in a more general model. If deviations are found, their magnitude can be extremely crucial for constraining the parameters of the more general Higgs sector. In particular, it might be possible to estimate the masses of the other Higgs bosons of the extended Higgs sector, thereby allowing a more focused search for them.

**TABLE 1.** Achievable relative uncertainties for a SM-like  $m_h = 110 \text{ GeV}$  for measuring the Higgs boson mass and total width for the LHC, LC ( $500 \text{ fb}^{-1}$ ), and the muon collider ( $0.2 \text{ fb}^{-1}$ ).

	LHC	LC	$\mu^+\mu^-$
$m_h$	$9 \times 10^{-4}$	$3 \times 10^{-4}$	$1 - 3 \times 10^{-6}$
$\Gamma_h^{\text{tot}}$	$> 0.3$	0.17	0.2

The accuracies possible at different colliders for measuring  $m_h$  and  $\Gamma_h^{\text{tot}}$  of a SM-like  $h$  with  $m_h \sim 110$  GeV are given in Table 1. To achieve these accuracies, one first determines the Higgs mass to about 1 MeV by the preliminary scan illustrated in Fig. 2. Then, a dedicated three-point fine scan [10] near the resonance peak using  $L \sim 0.2 \text{ fb}^{-1}$  of integrated luminosity (corresponding to a few years of operation) would be performed. For a SM Higgs boson with a mass sufficiently below the  $WW^*$  threshold, the Higgs total width is very small (of order several MeV), and the only process where it can be measured *directly* is in the  $s$ -channel at a muon collider. Indirect determinations at the LC can have higher accuracy once  $m_h$  is large enough that the  $WW^*$  mode rates can be accurately measured, requiring  $m_h > 120$  GeV. This is because at the LC the total width must be determined indirectly by measuring a partial width and a branching fraction, and then computing the total width,

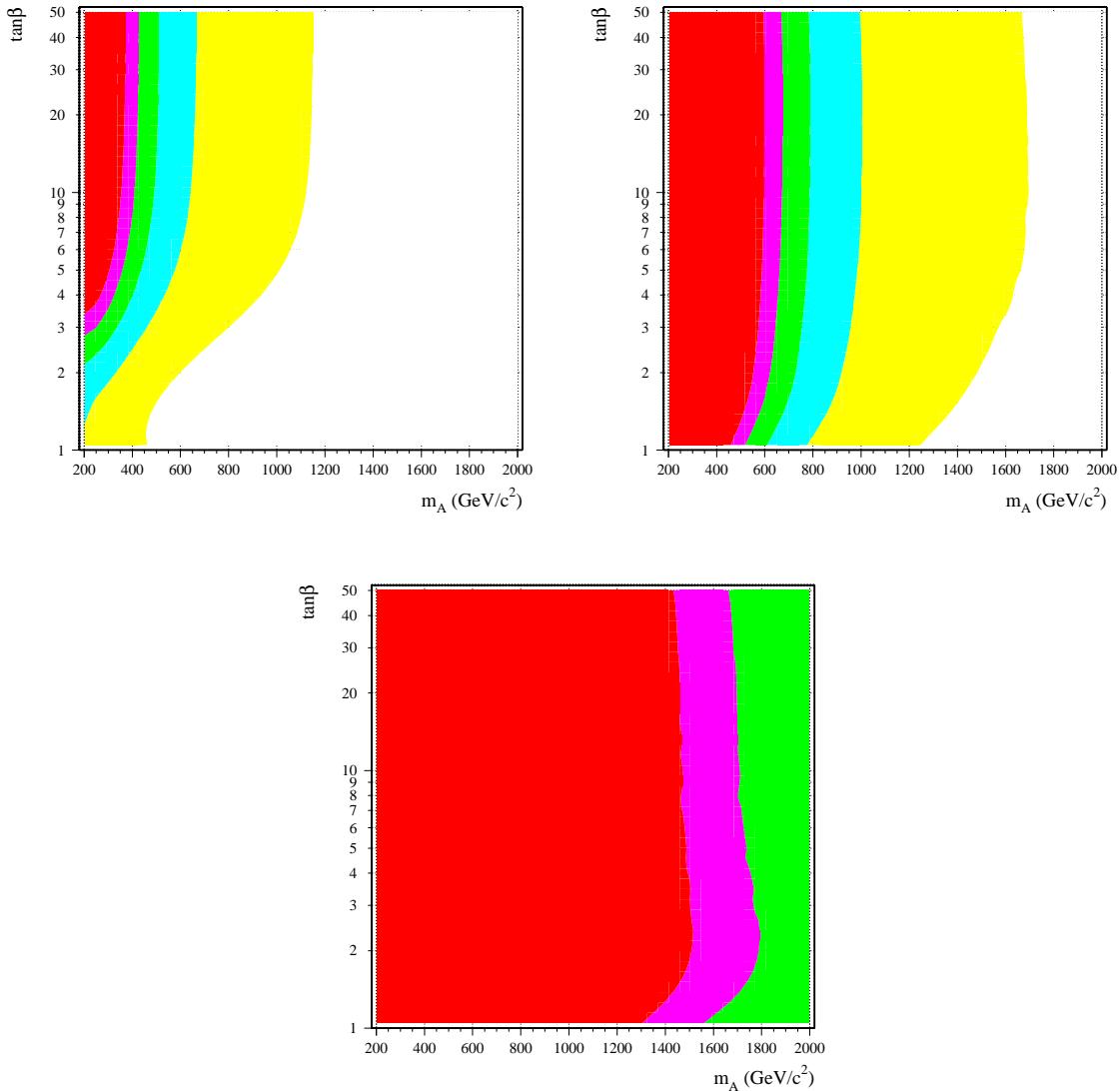
$$\Gamma_{\text{tot}} = \frac{\Gamma(h \rightarrow X)}{BR(h \rightarrow X)}, \quad (6)$$

for some final state  $X$ . For a Higgs boson so light that the  $WW^*$  decay mode is not useful, then the total width measurement would probably require use of the  $h \rightarrow \gamma\gamma$  decays [19]. This would require information from a photon collider as well as the LC and a small error is not possible. For  $m_h \lesssim 115$  GeV, the muon collider can measure the total width of the Higgs boson with greater precision than can be achieved using the indirect  $\gamma\gamma$  mode technique at the LC, and would be a very valuable input for precision tests of the Higgs sector. In particular, since all the couplings of the Standard Model  $h_{SM}$  are known,  $\Gamma_{h_{SM}}^{\text{tot}}$  is precisely predicted. Therefore, the precise determination of  $\Gamma_h^{\text{tot}}$  obtained by this scan would be an important test of the Standard Model, and any deviation would be evidence for a nonstandard Higgs sector (or other new physics).

In fact, a muon collider of limited luminosity can remain more than competitive with LHC + LC for discriminating between the SM  $h_{SM}$  and some SM-like  $h$  even for  $m_h$  values such that the LC obtains a good measurement of  $WW^*$  rates. As it happens, for  $X = b\bar{b}$  there is a fortuitous compensation that results in  $\bar{\sigma}_h(\mu^+\mu^- \rightarrow h \rightarrow b\bar{b})$  being almost completely independent of the somewhat uncertain  $b$  quark mass. Very roughly, larger  $m_b$  means larger  $\text{BF}(h \rightarrow b\bar{b})$  but also larger  $\Gamma_h^{\text{tot}}$ . The latter implies a smaller convoluted cross section  $\bar{\sigma}_h(\mu^+\mu^- \rightarrow h)$  (i.e. before including the branching ratio). Further, larger  $\Gamma_h^{\text{tot}}$  means less damping because of beam energy spread. The result is that  $\bar{\sigma}_h(\mu^+\mu^- \rightarrow h \rightarrow b\bar{b})$  is essentially independent of the input  $m_b$  value (within reasonable limits) [13]. As a result, the precise measurement of  $\bar{\sigma}_h(\mu^+\mu^- \rightarrow h \rightarrow b\bar{b})$  at a muon collider might provide the best single discriminator between the SM Higgs and a SM-like Higgs. This is nicely illustrated in the context of the MSSM. For a Higgs mass of 110 GeV, and assuming a typical soft-supersymmetry-breaking scenario, Fig. 4 shows the resulting excluded regions of  $m_{A^0}$  for the (a) LHC+LC, (b) with a muon collider with  $0.2 \text{ fb}^{-1}$  integrated luminosity, and (c) with a muon collider with  $10 \text{ fb}^{-1}$  integrated luminosity.

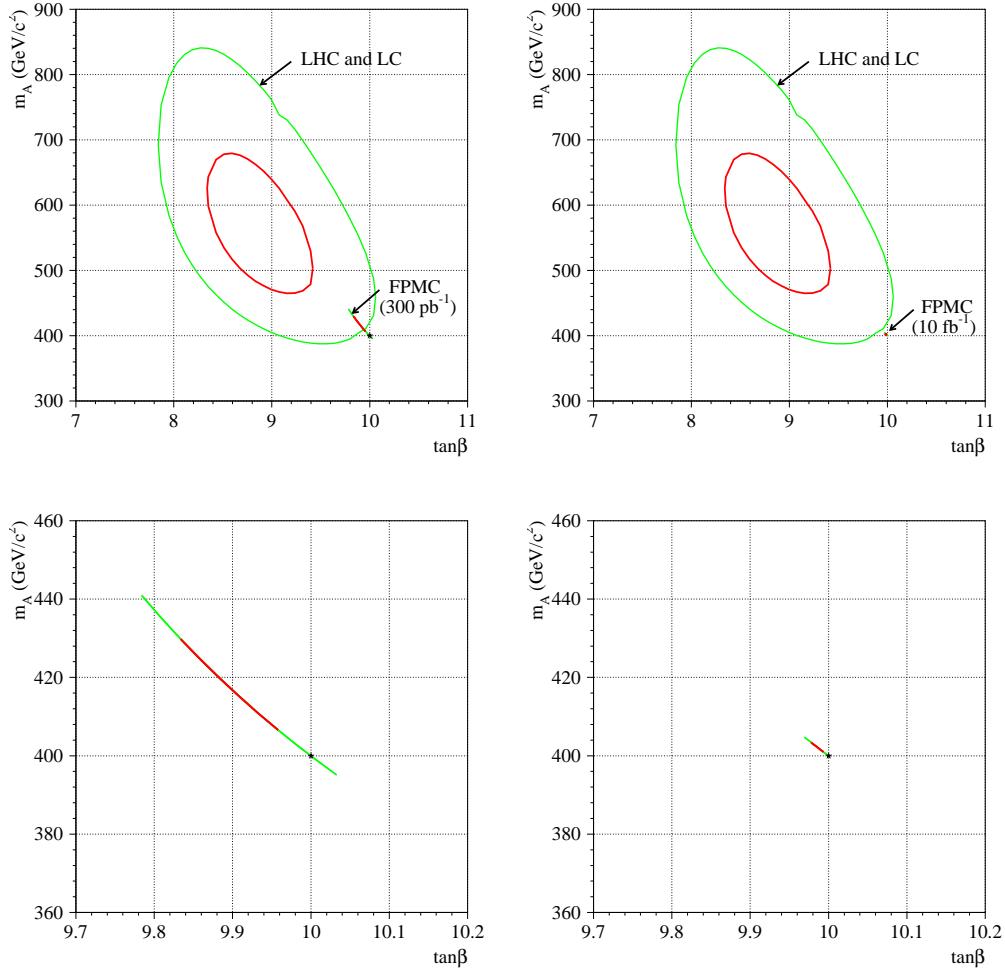
Some comments on these results are appropriate. First, one should note that the measurement of  $\Gamma_h^{\text{tot}}$  ( $\pm 0.5$  MeV, i.e.  $\pm 20\%$ ) at the muon collider is not nearly so powerful a discriminator as the  $\pm 3.5\%$  ( $\pm 0.5\%$ ) measurement of  $\bar{\sigma}_h(\mu^+\mu^- \rightarrow h \rightarrow b\bar{b})$  at  $L = 0.2 \text{ fb}^{-1}$  ( $10 \text{ fb}^{-1}$ ). Second, as  $m_h$  increases and the  $WW^*$  decay mode becomes more prominent, much more accurate determinations of partial width ratios and the total width become possible at LHC+LC and the LHC+LC exclusion regions move rapidly to higher  $m_{A^0}$ , but at best becoming comparable to the  $0.2 \text{ fb}^{-1}$  muon collider exclusion regions. Third, the conclusion that with higher luminosities than the  $0.1 \text{ fb}^{-1}$  per year currently envisioned for the Higgs factory this discriminator would have incredible sensitivity to  $m_{A^0}$  assumes that systematic errors for the absolute cross section will be smaller than the statistical errors. Fourth, we should note that there are high  $\tan\beta$  scenarios in which decoupling sets in very early in  $m_{A^0}$  and no machine would be able to set a lower bound on  $m_{A^0}$ ; in particular, for such scenarios it would be incorrect to conclude that the absence of deviations with respect to SM expectations implies that  $m_{A^0} \sim m_{H^0}$  would be such that  $m_{A^0} + m_{H^0} > \sqrt{s}$  so that  $e^+e^- \rightarrow H^0 A^0$  pair production is forbidden at a  $\sqrt{s} = 500$  GeV LC. Finally, if there was a very light neutralino such that  $h^0 \rightarrow \chi^0 \chi^0$  decays were possible, this would be known ahead of time and the  $\mu^+\mu^- \rightarrow h^0 \rightarrow b\bar{b}$  rate prediction within the SUSY context would have to be corrected to very high precision to account for these additional decays. SUSY loop corrections to the  $b\bar{b}$  coupling might also have to be accounted for to high precision if the SUSY spectrum turns out to be light. But these last two caveats also apply to the LC measurements.

Given the above sensitivity, the next question is the extent to which parameters of the supersymmetric Higgs sector can be determined with some reasonable level of precision. To study this, an input MSSM model was assumed with  $m_{h^0} = 110$  GeV,  $m_{A^0} = 400$  GeV,  $\tan\beta = 10$  and  $A_t = \mu = M_{\text{SUSY}} = 1$  TeV. Various observables were computed as a function of  $m_{A^0}$  and  $\tan\beta$ . Let us for the moment imagine that  $m_{h^0}$  can be



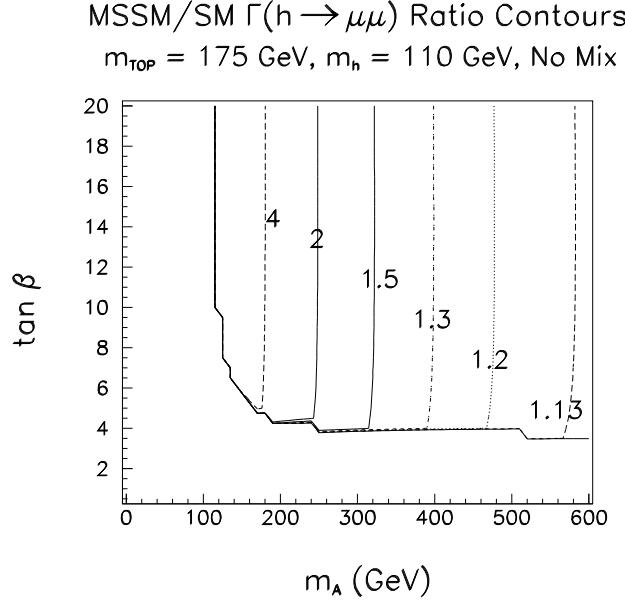
**FIGURE 4.** The  $m_{A^0}$ – $\tan\beta$  discrimination from the measurements at (a) LHC( $300 \text{ fb}^{-1}$ ) + LC( $500 \text{ fb}^{-1}$ ), (b)  $0.2 \text{ fb}^{-1}$  at a muon collider, and (c)  $10 \text{ fb}^{-1}$  at a muon collider. The exclusion regions (starting from the left) are  $> 5\sigma$ ,  $4 - 5\sigma$ ,  $3 - 4\sigma$ ,  $2 - 3\sigma$ , and  $1 - 2\sigma$ . From Ref. [13].

computed theoretically with arbitrary accuracy in terms of the input SUSY model parameters. Were this the case, then the value of  $m_{h^0}$  would determine  $m_{A^0}$  as a function of  $\tan\beta$  (or vice versa) given the fixed SUSY breaking scenario parameters. The  $b\bar{b}$  event rate and, to a lesser extent,  $\Gamma_{h^0}^{\text{tot}}$  determine the location along the line allowed by the fixed value of  $m_{h^0}$ . This line in  $[m_{A^0}, \tan\beta]$  parameter space is illustrated in the lower figures of Fig. 5 for the above sample model [14]. Also shown in these lower figures is the extent to which experimental measurements of  $N(\mu^+\mu^- \rightarrow h^0 \rightarrow b\bar{b})$  and  $\Gamma_{h^0}^{\text{tot}}$  for  $L = 0.1 \text{ fb}^{-1}$  and  $L = 10 \text{ fb}^{-1}$  would restrict the location along this line. The accuracy ( $\pm[0.1 - 0.3] \text{ MeV}$ ) with which  $m_{h^0}$  can be determined experimentally at the muon collider would not significantly broaden this line. For the experimental accuracies of  $\pm 90 \text{ MeV}$  at the LHC and  $\pm 30 \text{ MeV}$  at the LC, the line turns into the ellipses of the upper figures of Fig. 5. Unfortunately, due to the expected level of theoretical uncertainties in the computation of  $m_{h^0}$  the muon collider results are certainly unrealistic and even the LHC+LC ellipses are probably overly optimistic. We estimate that one might eventually be able to achieve a theoretical accuracy of  $\pm 100 \text{ MeV}$  for the  $m_{h^0}$  computation in terms of the model parameters. (Currently, the accuracy of the theoretical computations is  $\sim \pm[2 - 3] \text{ GeV}$ , so that much



**FIGURE 5.** The implications of the  $h^0$  scan for the MSSM  $[m_{A^0}, \tan\beta]$  parameter space, assuming all other SUSY parameters are known. In the lower figures, we illustrate the results that would emerge were there no systematic theoretical uncertainties in the  $m_{h^0}$  computation in terms of input SUSY parameters. The experimental error of  $m_{h^0}$  at a muon collider would not significantly broaden this line. The LH (RH) lower figure shows the extent to which the location along this line would be fixed by  $L = 0.1 \text{ fb}^{-1}$  ( $L = 10 \text{ fb}^{-1}$ ) muon collider measurements of  $N(\mu^+\mu^- \rightarrow h \rightarrow b\bar{b})$  and  $\Gamma_h^{\text{tot}}$ , with the former being the dominant ingredient given its much smaller error. In the upper two figures, the restrictions ( $1$  and  $2\sigma$  ellipses) that would emerge from LHC+LC measurements (including the measurement of  $m_{h^0}$  with accuracy of order  $\pm 30$  MeV) are shown. (Note the much more coarse scale of the upper figures.) These figures are from Ref. [13]. Unfortunately, the systematic error ( $\gtrsim \pm 100$  MeV, at best) expected for the  $m_{h^0}$  computation in terms of the input SUSY parameters will cause the potential muon collider lines of the lower figures to turn into ellipses similar in size to the LHC+LC ellipses and will increase the size of the LHC+LC ellipses significantly.

higher-loop work will be required to reach this level.) This would be comparable to the LHC experimental errors on  $m_{h^0}$ . Thus, the reality may be that LHC+LC ellipses of the upper half of Fig. 5 will be substantially enlarged. In any case, the ellipse sizes in both cases would most probably be determined by the accuracy of the theoretical computation of  $m_{h^0}$  as a function of SUSY parameters. A determination of the allowed elliptical regions including a reasonable level of systematic uncertainty for the  $m_{h^0}$  computation should be made. Despite this systematic uncertainty from the  $m_{h^0}$  computation, it is nonetheless clear that strong constraints would be imposed on the allowed regions in the multi-dimensional MSSM parameter space (that includes  $m_A$  and  $\tan\beta$  and the SUSY-breaking parameters) in order to achieve consistency with the measurements of  $m_{h^0}$ ,  $\bar{\sigma}(\mu^+\mu^- \rightarrow h^0 \rightarrow b\bar{b})$  and  $\Gamma_{h^0}^{\text{tot}}$ .



**FIGURE 6.** Contours in  $(m_{A^0}, \tan \beta)$  parameter space for  $\Gamma(h^0 \rightarrow \mu^+ \mu^-)/\Gamma(h_{SM} \rightarrow \mu^+ \mu^-)$ . We have assumed a no-mixing SUSY scenario and employed  $m_{h^0} = m_{h_{SM}} = 110$  GeV. For maximal mixing, there is little change in the contours — only the size of the allowed range is altered. From [22].

One very important probe of the physics of a light  $h$  that is only possible at a muon collider is the possibility of measuring  $\Gamma(h \rightarrow \mu^+ \mu^-)$ . Typically, the muon collider data must be combined with LC and/or LHC data to extract this very fundamental coupling. If the  $h$  is SM-like then the following determinations are possible.

$$1) \quad \Gamma(h \rightarrow \mu^+ \mu^-) = \frac{[\Gamma(h \rightarrow \mu^+ \mu^-) \text{BF}(h \rightarrow b\bar{b})] \mu C}{\text{BF}(h \rightarrow b\bar{b})_{\text{NLC}}};$$

$$2) \quad \Gamma(h \rightarrow \mu^+ \mu^-) = \frac{[\Gamma(h \rightarrow \mu^+ \mu^-) \text{BF}(h \rightarrow WW^*)] \mu C}{\text{BF}(h \rightarrow WW^*)_{\text{NLC}}};$$

$$3) \quad \Gamma(h \rightarrow \mu^+ \mu^-) = \frac{[\Gamma(h \rightarrow \mu^+ \mu^-) \text{BF}(h \rightarrow ZZ^*)] \mu C \Gamma_h^{\text{tot}}}{\Gamma(h \rightarrow ZZ^*)_{\text{NLC}}};$$

$$4) \quad \Gamma(h \rightarrow \mu^+ \mu^-) = \frac{[\Gamma(h \rightarrow \mu^+ \mu^-) \text{BF}(h \rightarrow WW^*) \Gamma_h^{\text{tot}}] \mu C}{\Gamma(h \rightarrow WW^*)_{\text{NLC}}}.$$

Using the above, a determination of  $\Gamma(h \rightarrow \mu^+ \mu^-)$  with accuracy  $\pm 4\%$  would be possible for an  $L \sim 0.2 \text{ fb}^{-1}$  muon collider run on the  $h$  peak and combining with LC( $200 \text{ fb}^{-1}$ ) data. In the MSSM context, such precision means that one would have  $3\sigma$  or greater difference between the expectation for the  $h_{SM}$  vs. the result for the  $h^0$  if  $m_{A^0} \leq 600$  GeV, assuming  $m_{h^0} \lesssim 135$  GeV (the MSSM upper limit). Further, this is an absolutely direct and model independent determination of  $\Gamma(h^0 \rightarrow \mu^+ \mu^-)$  that for certain has no systematic theoretical uncertainties. Of course, the caveat remains that there are peculiar MSSM parameter choices for which ‘decoupling’ occurs very rapidly and the  $h^0 \rightarrow \mu^+ \mu^-$  coupling would be independent of  $m_{A^0}$ . However, we would know ahead of time from the SUSY spectrum observed at the LHC whether or not such a peculiar scenario was relevant. Finally, we emphasize that the muon collider provides the only accurate probe of this 2nd generation lepton coupling<sup>3</sup> and would thus be one of the best checks of the the SM or MSSM explanation of lepton masses.

<sup>3)</sup> For  $m_h \sim 120$  GeV, estimates given M. Battaglia at the recent ECFA/DESY meeting for the accuracy to which  $\Gamma(h^0 \rightarrow \mu^+ \mu^-)$  can be determined at an  $e^+ e^-$  collider range from  $\pm 30\%$  for  $\sqrt{s} = 350$  GeV running ( $L = 500 \text{ fb}^{-1}$ ) (not much different for  $\sqrt{s} = 800$  GeV and  $L = 1 \text{ ab}^{-1}$ ) to  $\pm 7\%$  for  $\sqrt{s} = 3$  TeV ( $L = 5 \text{ ab}^{-1}$ ).

To summarize, if a Higgs is discovered at the LHC, or possibly earlier at the Fermilab Tevatron, attention will turn to determining whether this Higgs has the properties expected of the Standard Model Higgs. If the Higgs is discovered at the LHC, it is quite possible that supersymmetric states will be discovered concurrently. The next goal for a linear collider or a muon collider will be to better measure the Higgs boson properties to determine if everything is consistent within a supersymmetric framework or consistent with the Standard Model. A Higgs factory of even modest luminosity can provide uniquely powerful constraints on the parameter space of the supersymmetric model via the highly accurate determination of the total rate for  $\mu^+\mu^- \rightarrow h^0 \rightarrow b\bar{b}$  (which has almost zero theoretical systematic uncertainty due to its insensitivity to the unknown  $m_b$  value), the moderately accurate determination of the  $h^0$ 's total width and the remarkably accurate, unique and model-independent determination of the  $h^0\mu^+\mu^-$  coupling constant.

$$\mathbf{V} \quad H \rightarrow \tau^+\tau^-$$

A particularly important channel is the  $\tau^-\tau^+$  final state [15]

$$\mu^-\mu^+ \rightarrow \tau^-\tau^+. \quad (7)$$

In the SM at tree level, this  $s$ -channel process proceeds in two ways, via  $\gamma/Z$  exchange and Higgs boson exchange. The former involves the SM gauge couplings and presents a characteristic  $FB$  (forward-backward in the scattering angle) asymmetry and a  $LR$  (left-right in beam polarization) asymmetry; the latter is governed by the Higgs boson couplings to  $\mu^-\mu^+, \tau^-\tau^+$  proportional to the fermion masses and is isotropic in phase space due to spin-0 exchange. The unambiguous establishment of the  $\tau^-\tau^+$  signal would allow a determination of the relative coupling strength of the Higgs boson to  $b$  and  $\tau$  and thus test the usual assumption of  $\tau - b$  unification. The angular distribution would probe the spin property of the Higgs resonance.

The differential cross section for  $\mu^-\mu^+ \rightarrow \tau^-\tau^+$  via  $s$ -channel Higgs exchange can be expressed as

$$\frac{d\sigma_h(\mu^-\mu^+ \rightarrow h \rightarrow \tau^-\tau^+)}{d\cos\theta} = \frac{1}{2}\bar{\sigma}_h (1 + P_-P_+) \quad (8)$$

where  $\theta$  is the scattering angle between  $\mu^-$  and  $\tau^-$ ,  $P_\mp$  the percentage longitudinal polarizations of the initial  $\mu^\mp$  beams, with  $P = -1$  purely left-handed,  $P = +1$  purely right-handed and  $P = 0$  unpolarized.

The differential cross section for the SM background is given by the  $\gamma/Z$  contributions

$$\frac{d\sigma_{SM}}{d\cos\theta} = \frac{3}{8}\sigma_{QED}A[1 - P_+P_- + (P_+ - P_-)A_{LR}](1 + \cos^2\theta + \frac{8}{3}\cos\theta A_{FB}^{eff}). \quad (9)$$

Here the effective  $FB$  asymmetry factor is

$$A_{FB}^{eff} = \frac{A_{FB} + P_{eff}A_{LR}^{FB}}{1 + P_{eff}A_{LR}}, \quad (10)$$

with the effective polarization

$$P_{eff} = \frac{P_+ - P_-}{1 - P_+P_-}, \quad (11)$$

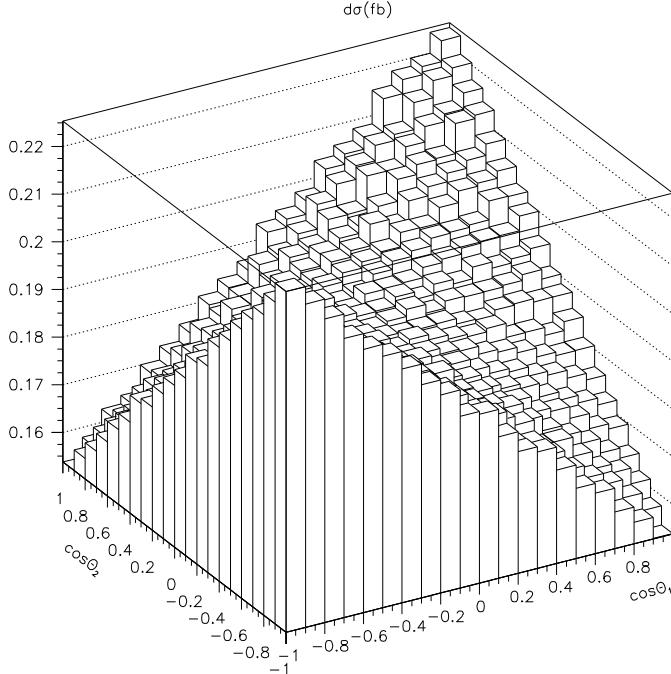
and

$$A_{LR}^{FB} = \frac{\sigma_{LR+RL \rightarrow LR} - \sigma_{LR+RL \rightarrow RL}}{\sigma_{LR+RL \rightarrow LR} + \sigma_{LR+RL \rightarrow RL}}. \quad (12)$$

$A_{FB}, A_{LR}$  are the standard asymmetries. For the case of interest where initial and final state particles are leptons,  $A_{LR} = A_{LR}^{FB}$ .

From the cross section formulas of Eqs. (8) and (9), the enhancement factor of the signal-to-background ratio ( $S/B$ ) due to the beam polarization effects is

$$\frac{S}{B} \sim \frac{1 + P_-P_+}{1 - P_-P_+ + (P_+ - P_-)A_{LR}}. \quad (13)$$



**FIGURE 7.** Double differential distribution for  $\mu^-\mu^+ \rightarrow h \rightarrow \tau^-\tau^+ \rightarrow \rho^-\nu_\tau\rho^+\bar{\nu}_\tau$ .  $\sqrt{s} = m_h = 120$  GeV is assumed. Initial  $\mu^\mp$  beam polarizations are taken to be  $P_- = P_+ = 0.25$ . The Higgs production cross section is convoluted with Gaussian energy distribution for a resolution  $R = 0.05\%$ .

The final state polarization configurations of  $\tau^-\tau^+$  from the Higgs signal and the SM background are very different. There is always a charged track to define a kinematical distribution for the decay. In the  $\tau$ -rest frame, the normalized differential decay rate can be written as

$$\frac{1}{\Gamma} \frac{d\Gamma_i}{d\cos\theta} = \frac{B_i}{2} (a_i + b_i P_\tau \cos\theta) \quad (14)$$

where  $\theta$  is the angle between the momentum direction of the charged decay product in the  $\tau$ -rest frame [16] and the  $\tau$ -momentum direction,  $B_i$  is the branching fraction for a given channel  $i$ , and  $P_\tau = \pm 1$  is the  $\tau$  helicity. For the two-body decay modes,  $a_i$  and  $b_i$  are constant and given by

$$a_\pi = b_\pi = 1, \quad (15)$$

$$a_i = 1 \quad \text{and} \quad b_i = -\frac{m_\tau^2 - 2m_i^2}{m_\tau^2 + 2m_i^2} \quad \text{for } i = \rho, a_1. \quad (16)$$

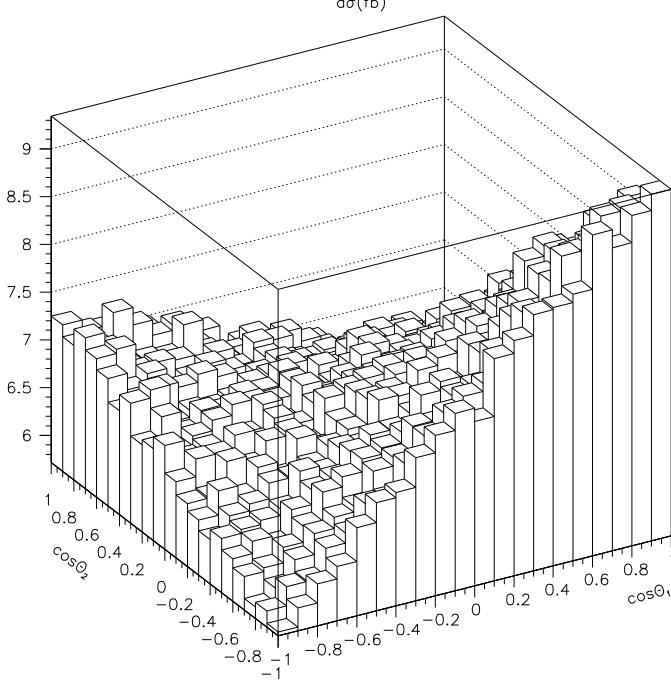
For the three-body leptonic decays, the  $a_{e,\mu}$  and  $b_{e,\mu}$  are not constant for a given three-body kinematical configuration and are obtained by the integration over the energy fraction carried by the invisible neutrinos. One can quantify the event distribution shape by defining a “sensitivity” ratio parameter

$$r_i = \frac{b_i}{a_i}. \quad (17)$$

For the two-body decay modes, the sensitivities are  $r_\pi = 1$ ,  $r_\rho = 0.45$  and  $r_{a_1} = 0.007$ . The  $\tau \rightarrow a_1\nu_\tau$  mode is consequently less useful in connection with the  $\tau$  polarization study. As to the three-body leptonic modes, although experimentally readily identifiable, the energy smearing from the decay makes it hard to reconstruct the  $\tau^-\tau^+$  final state spin correlation.

The differential distribution for the two charged particles  $(i, j)$  in the final state from  $\tau^-\tau^+$  decays respectively can be expressed as

$$\frac{d\sigma}{d\cos\theta_i d\cos\theta_j} \sim \sum_{P_\tau=\pm 1} \frac{B_i B_j}{4} (a_i + b_i P_{\tau^-} \cos\theta_i)(a_j + b_j P_{\tau^+} \cos\theta_j), \quad (18)$$



**FIGURE 8.** Double differential distribution for  $\mu^-\mu^+ \rightarrow \gamma^*/Z^* \rightarrow \tau^-\tau^+ \rightarrow \rho^-\nu_\tau\rho^+\bar{\nu}_\tau$ .  $\sqrt{s} = 120$  GeV is assumed. Initial  $\mu^\mp$  beam polarizations are taken to be  $P_- = P_+ = 0.25$ . The SM production cross section is convoluted with Gaussian energy distribution for a resolution  $R = 0.05\%$ .

where  $\cos\theta_i$  ( $\cos\theta_j$ ) is defined in  $\tau^-$  ( $\tau^+$ ) rest frame as in Eq. (14). For the Higgs signal channel,  $\tau^-\tau^+$  helicities are correlated as  $LL$  ( $P_{\tau^-} = P_{\tau^+} = -1$ ) and  $RR$  ( $P_{\tau^-} = P_{\tau^+} = +1$ ). This yields the spin-correlated differential cross section

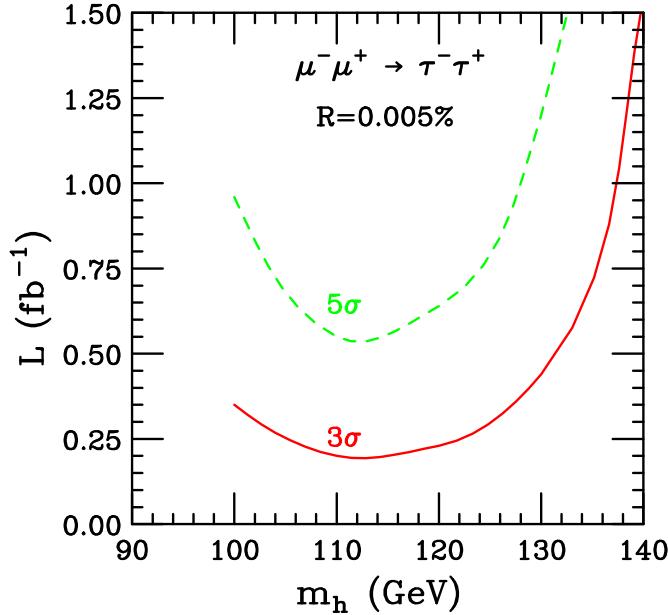
$$\frac{d\sigma_h}{d\cos\theta_i d\cos\theta_j} = (1 + P_- P_+) \sigma_h \frac{B_i B_j}{4} [a_i a_j + b_i b_j \cos\theta_i \cos\theta_j], \quad (19)$$

We expect that the distribution reaches maximum near  $\cos\theta_i = \cos\theta_j = \pm 1$  and minimum near  $\cos\theta_i = -\cos\theta_j = \pm 1$ . How significant the peaks are depends on the sensitivity parameter in Eq. (17). Here we simulate the double differential distribution of Eq. (19) for  $\mu^-\mu^+ \rightarrow h \rightarrow \tau^-\tau^+ \rightarrow \rho^-\nu_\tau\rho^+\bar{\nu}_\tau$  and the result is shown in Fig. 7. Here we take  $\sqrt{s} = m_h = 120$  GeV for illustration. The Higgs production cross section is convoluted with Gaussian energy distribution for a resolution  $R = 0.05\%$ . We see distinctive peaks in the distribution near  $\cos\theta_{\rho^-} = \cos\theta_{\rho^+} = \pm 1$ , as anticipated. In this demonstration, we have taken  $\mu^\mp$  beam polarizations to be  $P_- = P_+ = 25\%$ , which is considered to be natural with little cost to beam luminosity.

In contrast, the SM background via  $\gamma^*/Z^*$  produces  $\tau^-\tau^+$  with helicity correlation of  $LR$  ( $P_{\tau^-} = -P_{\tau^+} = -1$ ) and  $RL$  ( $P_{\tau^-} = -P_{\tau^+} = +1$ ). Furthermore, the numbers of the left-handed and right-handed  $\tau^-$  at a given scattering angle are different because of the left-right asymmetry, so the initial muon beam polarization affects the  $\tau^-\tau^+$  spin correlation non-trivially. Summing over the two polarization combinations in  $\tau^-\tau^+$  decay to particles  $i$  and  $j$ , we have

$$\begin{aligned} \frac{d\sigma_{SM}}{d\cos\theta_i d\cos\theta_j} &= (1 - P_- P_+) \sigma_{SM} (1 + P_{eff} A_{LR}) \times \\ &\quad \frac{B_i B_j}{4} [(a_i a_j - b_i b_j \cos\theta_i \cos\theta_j) + A_{LR}^{eff} (a_i b_j \cos\theta_j - a_j b_i \cos\theta_i)]. \end{aligned} \quad (20)$$

The final state spin correlation for  $\mu^-\mu^+ \rightarrow \gamma^*/Z^* \rightarrow \tau^-\tau^+$  decaying into  $\rho^-\rho^+$  pairs is shown in Fig. 8. The maximum regions near  $\cos\theta_{\rho^-} = -\cos\theta_{\rho^+} = \pm 1$  are clearly visible. Most importantly, the peak regions in Figs. 7 and 8 occur exactly in the opposite positions from the Higgs signal. We also note that the spin correlation from the Higgs signal is symmetric, while that from the background is not. The reason is that



**FIGURE 9.** Integrated luminosity (in  $\text{fb}^{-1}$ ) needed for observing the two-body decay channels  $\tau \rightarrow \rho \nu_\tau$  and  $\tau \rightarrow \pi \nu_\tau$  at  $3\sigma$  (solid) and  $5\sigma$  (dashed) significance. Beam energy resolution  $R = 0.005\%$  and a 25% polarization are assumed.

the effective  $LR$ -asymmetry in the background channel changes the relative weight of the two maxima, which becomes transparent from the last term in Eq. (20).

We next estimate the luminosity needed for signal observation of a given statistical significance. The results are shown in Fig. 9. The integrated luminosity ( $L$  in  $\text{fb}^{-1}$ ) needed for observing the characteristic two-body decay channels  $\tau \rightarrow \rho \nu_\tau$  and  $\tau \rightarrow \pi \nu_\tau$  at  $3\sigma$  (solid) and  $5\sigma$  (dashed) significance is calculated for both signal and SM background with  $\sqrt{s} = m_h$ . Beam energy resolution  $R = 0.005\%$  and a 25%  $\mu^\pm$  beam polarization are assumed.

We estimate the statistical error on the cross section measurement. If we take the statistical error to be given by

$$\epsilon = \frac{\sqrt{S+B}}{S} = \frac{1}{\sqrt{L}} \frac{\sqrt{\sigma_S + \sigma_B}}{\sigma_S}, \quad (21)$$

summing over both  $\rho \nu_\tau$  and  $\pi \nu_\tau$  channels for  $R = 0.005\%$ , a 25% beam polarization with  $1 \text{ fb}^{-1}$  luminosity, we obtain

$$\begin{array}{lccccc} \sqrt{s} = m_h \text{ ( GeV)} & 100 & 110 & 120 & 130 \\ \epsilon (\%) & 27 & 21 & 23 & 32 \end{array} \quad (22)$$

The uncertainties on the cross section measurements determine the extent to which the  $h\tau^-\tau^+$  coupling can be measured.

In summary, we have demonstrated the feasibility of observing the resonant channel  $h \rightarrow \tau^-\tau^+$  at a muon collider. For a narrow resonance like the SM Higgs boson, a good beam energy resolution is crucial for a clear signal. On the other hand, a moderate beam polarization would not help much for the signal identification. The integrated luminosity needed for a signal observation is presented in Fig. 9. Estimated statistical errors for the  $\mu^-\mu^+ \rightarrow h \rightarrow \tau^-\tau^+$  cross section measurement are given in Eq. (22). We emphasized the importance of final state spin correlation to purify the signal of a scalar resonance and to confirm the nature of its spin. It is also important to carefully study the  $\tau^-\tau^+$  channel of a supersymmetric Higgs boson which would allow a determination of the relative coupling strength of the Higgs to  $b$  and  $\tau$ .

## VI HEAVY HIGGS BOSONS

As discussed in the previous section, precision measurements of the light Higgs boson properties might make it possible to detect deviations with respect to expectations for a SM-like Higgs boson that would point to a limited range of allowed masses for the heavier Higgs bosons. This becomes more difficult in the decoupling limit where the differences between a supersymmetric and Standard Model Higgs are smaller. Nevertheless with sufficiently precise measurements of the Higgs branching fractions, it is possible that the heavy Higgs boson masses can be inferred. A muon collider light-Higgs factory might be essential in this process.

In the context of the MSSM,  $m_{A^0}$  can probably<sup>4)</sup> be restricted to within 50 GeV or better if  $m_{A^0} < 500$  GeV. This includes the  $250 - 500$  GeV range of heavy Higgs boson masses for which discovery is not possible via  $H^0 A^0$  pair production at a  $\sqrt{s} = 500$  GeV LC. Further, the  $A^0$  and  $H^0$  cannot be detected in this mass range at either the LHC or LC for a wedge of moderate  $\tan\beta$  values. (For large enough values of  $\tan\beta$  the heavy Higgs bosons are expected to be observable in  $b\bar{b}A^0, b\bar{b}H^0$  production at the LHC via their  $\tau^+\tau^-$  decays and also at the LC.)

A muon collider can fill some, perhaps all of this moderate  $\tan\beta$  wedge. If  $\tan\beta$  is large the  $\mu^+\mu^-H^0$  and  $\mu^+\mu^-A^0$  couplings (proportional to  $\tan\beta$  times a SM-like value) are enhanced, thereby leading to enhanced production rates in  $\mu^+\mu^-$  collisions. The most efficient procedure is to operate the muon collider at maximum energy and produce the  $H^0$  and  $A^0$  (often as overlapping resonances) via the radiative return mechanism. By looking for a peak in the  $b\bar{b}$  final state, the  $H^0$  and  $A^0$  can be discovered and, once discovered, the machine  $\sqrt{s}$  can be set to  $m_{A^0}$  or  $m_{H^0}$  and factory-like precision studies pursued. Note that the  $A^0$  and  $H^0$  are typically broad enough that  $R = 0.1\%$  would be adequate to maximize their  $s$ -channel production rates. In particular,  $\Gamma \sim 30$  MeV if the  $t\bar{t}$  decay channel is not open, and  $\Gamma \sim 3$  GeV if it is. Since  $R = 0.1\%$  is sufficient, much higher luminosity ( $L \sim 2 - 10 \text{ fb}^{-1}/\text{yr}$ ) would be possible as compared to that for  $R = 0.01\% - 0.003\%$  as required for studying the  $h^0$ .

In short, for those portions of parameter space characterized by moderate  $\tan\beta$  and  $m_{A^0} \gtrsim 250$  GeV, which are particularly difficult for both the LHC and the LC, the muon collider would be the only place that these extra Higgs bosons can be discovered and their properties measured very precisely.<sup>5)</sup>

In the MSSM, the heavy Higgs bosons are largely degenerate, especially in the decoupling limit where they are heavy. Large values of  $\tan\beta$  heighten this degeneracy as shown in Fig. 10. A muon collider with sufficient energy resolution might be the only possible means for separating out these states. Examples showing the  $H$  and  $A$  resonances for  $\tan\beta = 5$  and 10 are shown in Fig. 11. For the larger value of  $\tan\beta$  the resonances are clearly overlapping. For the better energy resolution of  $R = 0.01\%$ , the two distinct resonance peaks are still visible, but they are smeared out and merge into one broad peak for  $R = 0.06\%$ .

A precise measurement on the heavy Higgs boson masses could provide a powerful window on radiative corrections in the supersymmetric Higgs sector [23]. Supersymmetry with gauge invariance in the MSSM implies the mass-squared sum rule

$$m_h^2 + m_H^2 = m_A^2 + m_Z^2 + \Delta, \quad (23)$$

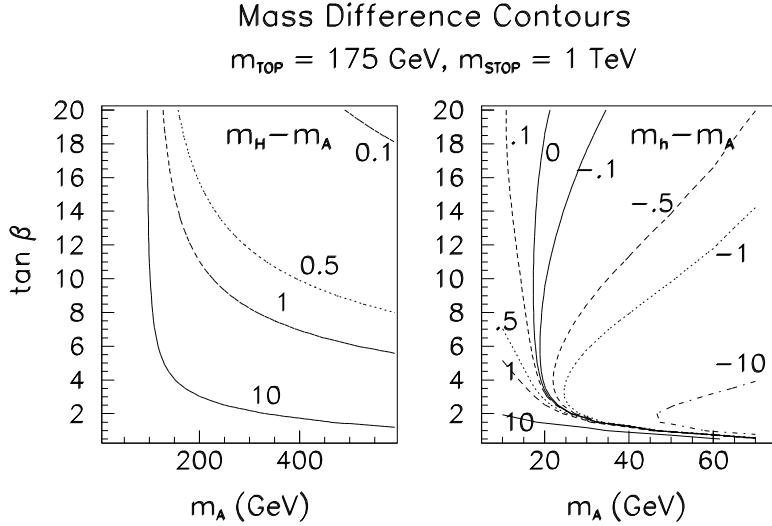
where  $\Delta$  is a calculable radiative correction (the tree-level sum rule results from setting  $\Delta = 0$ ). Solving for the mass difference

$$m_A - m_H = \frac{m_h^2 - m_Z^2 - \Delta}{m_A + m_H}, \quad (24)$$

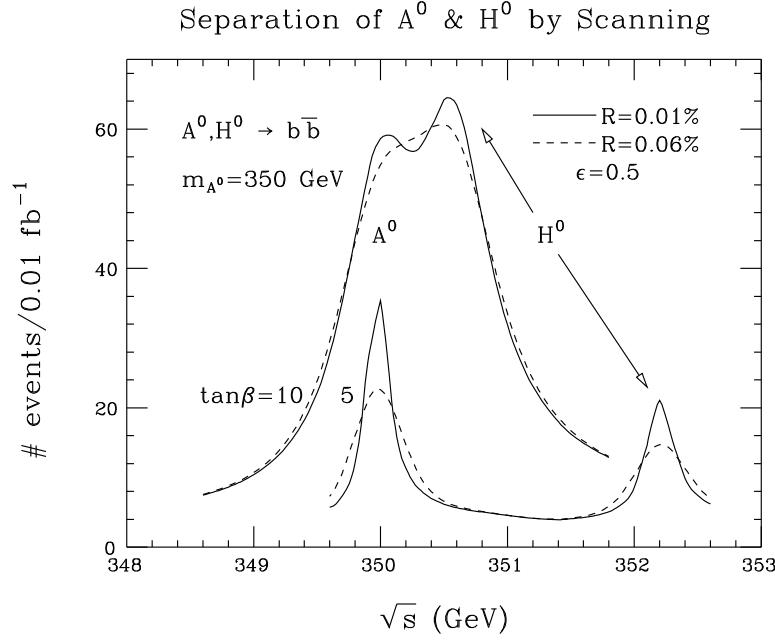
one obtains a formula involving observables that can be precisely measured. For example the error on the  $m_Z$  is just 2.2 MeV from the LEP measurements [21], and the light Higgs mass can be measured to less than an MeV in the  $s$ -channel. The masses of and the mass difference between the heavy Higgs states  $H$  and  $A$  can also be measured precisely by  $s$ -channel production. The ultimate precision that can be obtained on the masses of the  $H$  and  $A$  depends strongly on the masses themselves and  $\tan\beta$ . But a reasonable expectation is that a scan through the resonances should be able to determine the masses and the mass-difference to some tens

<sup>4)</sup> For the peculiar parameter regions with ‘early’ decoupling, mentioned earlier, this would not be possible. However, as noted earlier, the SUSY spectrum observed at the LHC would allow us to determine if we are in such an exceptional region of parameter space.

<sup>5)</sup> The  $\gamma\gamma$  collider option at an LC would also allow  $H^0, A^0$  discovery throughout much of the wedge region [17], but only the muon collider could directly scan for their total widths and determine their  $\mu^+\mu^-$  coupling.



**FIGURE 10.** Contours of  $m_H - m_A$  (in GeV) in the  $(m_A, \tan \beta)$  parameter space. Two-loop/RGE-improved radiative corrections are included taking  $m_t = 175$  GeV,  $m_{\tilde{t}} = 1$  TeV, and neglecting squark mixing.



**FIGURE 11.** Separation of  $A$  and  $H$  signals for  $\tan \beta = 5$  and 10. From Ref. [10].

of MeV [23]. Altogether these mass measurements yield a prediction for the radiative correction  $\Delta$  which is calculable in terms of the self-energy diagrams of the Higgs bosons [24]. To fully exploit this constraint might, however, prove difficult given the notorious difficulty of computing Higgs boson masses to high enough loop order that accuracy better than even a GeV can be achieved.

Finally it will be especially interesting to measure the branching ratios of these heavy Higgs bosons and compare to the theoretical predictions. For  $\tan \beta \gtrsim 5$  the  $H^0, A^0$  decay more often into  $b\bar{b}$  than into  $t\bar{t}$ . There is a substantial range of parameter space where significant numbers of events involving both types of decays will be seen and new type of determination of  $\tan \beta$  will be possible. If supersymmetric particle masses are below  $\sim m_{A^0}/2$ , then the branching ratios for  $A^0, H^0$  decays to the many distinguishable channels provide extremely powerful constraints on the soft-supersymmetry-breaking parameters of the model [25–27].

## VII HIGGS THRESHOLD MEASUREMENT

The mass, width and spin of a SM-like Higgs boson can also be determined by operating either a muon collider or a linear collider at the  $Zh$  production threshold. The rapid rise in the production near the threshold is sensitive to the Higgs mass [28]. Furthermore the spin of the Higgs boson can be determined by examining the rise in the cross section near threshold. However, these measurements require tens of inverse femtobarns to provide a useful measurement of the mass ( $< 100$  MeV). These threshold measurements can be performed at a LC; with  $100 \text{ fb}^{-1}$  of integrated luminosity, an error of less than 100 MeV can be achieved [28] for  $m_h < 150$  GeV. This is comparable to the other methods at energies above threshold. The only means to reduce the experimental error on the Higgs mass further to below 1 MeV is to produce the Higgs in the  $s$ -channel at a muon collider.

The shape of the  $\ell^+\ell^- \rightarrow Zh$  threshold cross section can also be used to determine the spin and to check the CP=+ property of the Higgs [29]. These threshold measurements become of interest for a muon collider in the case where at least a hundred inverse femtobarns of luminosity is available.

## VIII NON-EXOTIC, NON-SUPERSYMMETRIC SM HIGGS SECTOR EXTENSIONS

Although the standard interpretation of precision electroweak data is that there should be a light Higgs boson with SM-like  $VV$  couplings, alternative Higgs sector models can be constructed in which a good fit to the precision data is obtained even though the Higgs boson with large  $VV$  coupling is quite heavy ( $\sim 1$  TeV). The simplest such model [20] is based upon the CP-conserving general two-Higgs-doublet model. The large  $\Delta S > 0$  and  $\Delta T < 0$  coming from the heavy Higgs with large  $VV$  coupling is compensated by an even larger  $\Delta T > 0$  coming from a small ( $\Delta M \sim 1$  GeV is sufficient) mass splitting between the  $H^\pm$  and the other heavy neutral Higgs boson. The result is a shift in the  $\Delta S > 0$ ,  $\Delta T > 0$  direction (relative to the usual  $m_{h_{SM}} \sim 100$  GeV scenario in the SM) that remains well within the current 90% CL ellipse in the  $S, T$  plane. The first signal for this type of scenario would be discovery of a heavy SM-like Higgs boson at the LHC. If such a heavy SM-like Higgs is discovered, Consistency with precision electroweak data would then require the above type of scenario or some other exotic new physics scenario.

Models of this type cannot arise in the supersymmetric context because of constraints on the Higgs self couplings coming from the SUSY structure. They require a special ‘non-decoupling’ form for the potential that could arise in models where the two-doublet Higgs sector is an effective low energy description up to some scale  $\Lambda$  of order 10 TeV or so. For these special potential forms, there is typically also a Higgs boson  $\hat{h}$  ( $\hat{h}$  =decoupled- $h^0$  or  $\hat{h} = A^0$ ) with  $m_{\hat{h}} < 500$  GeV and no tree-level  $VV$  coupling. Its primary decay modes would be to  $b\bar{b}$  or  $t\bar{t}$  (depending upon its mass) and its  $\mu^+\mu^-$  coupling would be proportional to  $\tan\beta$ . For a substantial range of  $\tan\beta$ , this  $\hat{h}$  could not be detected at either the LHC or the LC [20]. In particular, at the LC even the  $e^+e^- \rightarrow Z^* \rightarrow Z\hat{h}\hat{h}$  process (the quartic coupling being of guaranteed strength) would only allow  $\hat{h}$  discovery up to 150 GeV (250 GeV) for  $\sqrt{s} = 500$  GeV (800 GeV) [30].

The muon collider could be the key to discovering such a  $\hat{h}$ . By running at high energy, the radiative return tail for  $E_{\mu^+\mu^-}$  might result in production of a detectable number of events. In particular, if  $\tan\beta > 5$ , operation at maximal  $\sqrt{s}$  with  $R = 0.1\%$  would guarantee that the  $\hat{h}$  would be detected as a  $4\sigma$  or higher bump in the bremsstrahlung tail of the  $m_{b\bar{b}}$  distribution after 3 to 4 years of running. Alternatively, a scan could be performed to look for the  $\hat{h}$ . The scan procedure depends upon how  $\Gamma_{\hat{h}}^{\text{tot}}$  depends on  $m_{\hat{h}}$  in that one must always have  $R$  such that  $\sigma_{\sqrt{s}} \lesssim \Gamma_{\hat{h}}^{\text{tot}}$ ; the luminosity expected for the required  $R$  must then be employed. Further, one must use steps of size  $\Gamma_{\hat{h}}^{\text{tot}} \sim \sigma_{\sqrt{s}}$ . For  $2m_t > m_{\hat{h}} > 150$  GeV,  $\hat{h} \rightarrow b\bar{b}$  and  $\Gamma_{\hat{h}}^{\text{tot}} \sim 0.05 - 0.1$  GeV unless  $\tan\beta < 1$ . For  $m_{\hat{h}} > 2m_t$ ,  $\Gamma_{\hat{h}}^{\text{tot}}$  rises to at least 1 GeV. As result, it would be possible to employ  $R = 0.05 - 0.1\%$  or so for  $m_{\hat{h}} < 2m_t$  rising to  $R = 0.5 - 1\%$  for  $m_{\hat{h}} > 2m_t$ . In a 3 – 4 year program, using earlier quoted nominal yearly  $L$ ’s for such  $R$ ’s as function of  $\sqrt{s}$ , we could imagine devoting:

- $L = 0.003 \text{ fb}^{-1}$  to 2000 points separated by 0.1 GeV in  $\sqrt{s} = 150 - 350$  GeV range — the total luminosity required would be  $L = 4 \text{ fb}^{-1}$  or about 3 years of operation. One would find ( $4\sigma$  level) the  $\hat{h}$  in the  $b\bar{b}$  state if  $\tan\beta \gtrsim 4 - 5$ .

- $L = 0.03 \text{ fb}^{-1}$  to each of 100 points separated by 0.5 GeV in the  $\sqrt{s} = 350 - 400$  GeV range — the corresponding total luminosity used is  $L = 3 \text{ fb}^{-1}$  or about 1/2 year of operation. For  $\tan\beta > 6 (< 6)$ , one would find the  $\hat{h}$  in  $b\bar{b}$  ( $t\bar{t}$ ) final state.
- $L = 0.01 \text{ fb}^{-1}$  to each of 100 points separated by 1 GeV in the  $\sqrt{s} = 400 - 500$  GeV range — the total luminosity employed would be  $L = 1 \text{ fb}^{-1}$ , or about 1/10 year. For  $\tan\beta > 7 (< 8)$ , one would detect the  $\hat{h}$  in the  $b\bar{b}$  ( $t\bar{t}$ ) final state.

In this way, the muon collier would detect the  $\hat{h}$  if  $m_{\hat{h}} < 2m_t$  and  $\tan\beta \gtrsim 5$  or if  $m_{\hat{h}} > 2m_t$  for any  $\tan\beta$ . Once discovered,  $\sqrt{s} = m_{\hat{h}}$  could be chosen for the muon collider and it would be possible to study the  $\hat{h}$  properties in detail.

## IX CP VIOLATION

A muon collider can probe the CP properties of a Higgs boson produced in the  $s$ -channel. One can measure correlations in the  $\tau^+\tau^-$  final state or, if the Higgs boson is sufficiently heavy, in the  $t\bar{t}$  final state [32,33]. In the MSSM at tree-level the Higgs states  $h^0$ ,  $H^0$ , and  $A^0$  are CP eigenstates, but it has been noted recently that sizable CP violation is possible in the MSSM Higgs sector through loop corrections involving the third generation squarks [34,35]. As noted earlier, in the MSSM the two heavy neutral Higgs bosons ( $H^0$  being CP-even and  $A^0$  being CP-odd) are almost degenerate with a mass splitting comparable or less than their widths. If there are CP-violating phases in the neutral Higgs potential, these will cause these CP eigenstates to mix. The resulting mass splitting between the eigenstates can be larger than their widths. The excellent mass resolution at the muon collider would make it possible separate the masses of the  $H^0$  and  $A^0$  bosons. The measured mass difference could be combined with the mass sum rule to provide a powerful probe of this physics. As already noted, various CP asymmetries in the  $t\bar{t}$  final state can be observed as well, and a muon collider is an ideal place to look for these effects [32,33].

The most ideal means for determining the CP nature of a Higgs boson at the muon collider is to employ transversely polarized muons. For  $h$  production at a muon collider with muon coupling given by the form  $\overline{\mu}(a + ib\gamma_5)\mu h$ , the cross section takes the form

$$\begin{aligned}\overline{\sigma}_h(\zeta) &= \overline{\sigma}_h^0 \left( 1 + P_L^+ P_L^- + P_T^+ P_T^- \left[ \frac{a^2 - b^2}{a^2 + b^2} \cos\zeta - \frac{2ab}{a^2 + b^2} \sin\zeta \right] \right) \\ &= \overline{\sigma}_h^0 [1 + P_L^+ P_L^- + P_T^+ P_T^- \cos(2\delta + \zeta)] ,\end{aligned}\quad (25)$$

where  $\overline{\sigma}_h^0$  is the polarization average convoluted cross section,  $\delta \equiv \tan^{-1} \frac{b}{a}$ ,  $P_T$  ( $P_L$ ) is the degree of transverse (longitudinal) polarization, and  $\zeta$  is the angle of the  $\mu^+$  transverse polarization relative to that of the  $\mu^-$  as measured using the direction of the  $\mu^-$ 's momentum as the  $\hat{z}$  axis. Of course, if there is no  $P_T$  there would be sensitivity to  $\overline{\sigma}_h^0 \propto a^2 + b^2$  only. Only the  $\sin\zeta$  term is truly CP-violating, but the  $\cos\zeta$  term also provides significant sensitivity to  $a/b$ . Ideally, one would isolate  $\frac{a^2 - b^2}{a^2 + b^2}$  and  $\frac{-2ab}{a^2 + b^2}$  by running at fixed  $\zeta = 0, \pi/2, \pi, 3\pi/2$  and measuring the asymmetries (taking  $P_T^+ = P_T^- \equiv P_T$  and  $P_L^\pm = 0$ )

$$\begin{aligned}\mathcal{A}_I &\equiv \frac{\overline{\sigma}_h(\zeta = 0) - \overline{\sigma}_h(\zeta = \pi)}{\overline{\sigma}_h(\zeta = 0) + \overline{\sigma}_h(\zeta = \pi)} = P_T^2 \frac{a^2 - b^2}{a^2 + b^2} = P_T^2 \cos 2\delta , \\ \mathcal{A}_{II} &\equiv \frac{\overline{\sigma}_h(\zeta = \pi/2) - \overline{\sigma}_h(\zeta = -\pi/2)}{\overline{\sigma}_h(\zeta = \pi/2) + \overline{\sigma}_h(\zeta = -\pi/2)} = -P_T^2 \frac{2ab}{a^2 + b^2} = -P_T^2 \sin 2\delta .\end{aligned}$$

If  $a^2 + b^2$  is already well determined, and the background is known, then the fractional error in these asymmetries can be approximated as  $\frac{\delta\mathcal{A}}{\mathcal{A}} \propto P_T^2 \sqrt{L}$ , which points to the need for the highest possible transverse polarization, even if some sacrifice in  $L$  is required.

Of course, in reality the precession of the muon spin in a storage ring makes running at fixed  $\zeta$  impossible. A detailed study is required [36]. We attempt a brief outline. Taking  $\vec{B} = -B\hat{y}$  we may write

$$s_{\mu^-} = P_H^- [\gamma(\beta, \hat{z}) \cos\theta^- - (0, \hat{x}) \sin\theta^-] + P_V^- (0, \hat{y}), \quad s_{\mu^+} = P_H^+ [\gamma(\beta, -\hat{z}) \cos\theta^+ - (0, \hat{x}) \sin\theta^+] + P_V^+ (0, \hat{y}). \quad (26)$$

Here,  $\hat{z}$  is the direction of the  $\mu^-$  instantaneous momentum,  $P_H$  ( $P_V$ ) is the horizontal (vertical, i.e.  $\hat{y}$ ) degree of polarization,  $P_H^\pm \cos \theta^\pm = P_L^\pm$ , and  $\sqrt{[P_H^\pm \sin \theta^\pm]^2 + [P_V^\pm]^2} = P_T^\pm$ . For any setup for initial insertion into the storage ring,  $\theta^\pm$  can be computed as functions of the turn number  $N_T$  (counting starting with  $N_T = 1$  the first time the bunch passes the IP). (For example, if the  $\mu^\pm$  beams enter the storage ring with  $\hat{P}_H^\pm = \hat{p}_{\mu^\pm}$ , then  $\theta^\pm(N_T) = \omega(N_T - 1/2)$ , where  $\omega = 2\pi\gamma\frac{g_\mu - 2}{2}$ , with  $\gamma = E/m_\mu$ .) As a function of  $\theta^-$  and  $\theta^+$ , defining  $c_- \equiv \cos \theta^-$  etc.,

$$\frac{\bar{\sigma}_h(\theta^+, \theta^-)}{\bar{\sigma}_h^0} = (1 + P_H^+ P_H^- c_+ c_-) + \cos 2\delta (P_V^+ P_V^- + P_H^+ P_H^- s_+ s_-) + \sin 2\delta (P_H^- P_V^+ s_- - P_H^+ P_V^- s_+). \quad (27)$$

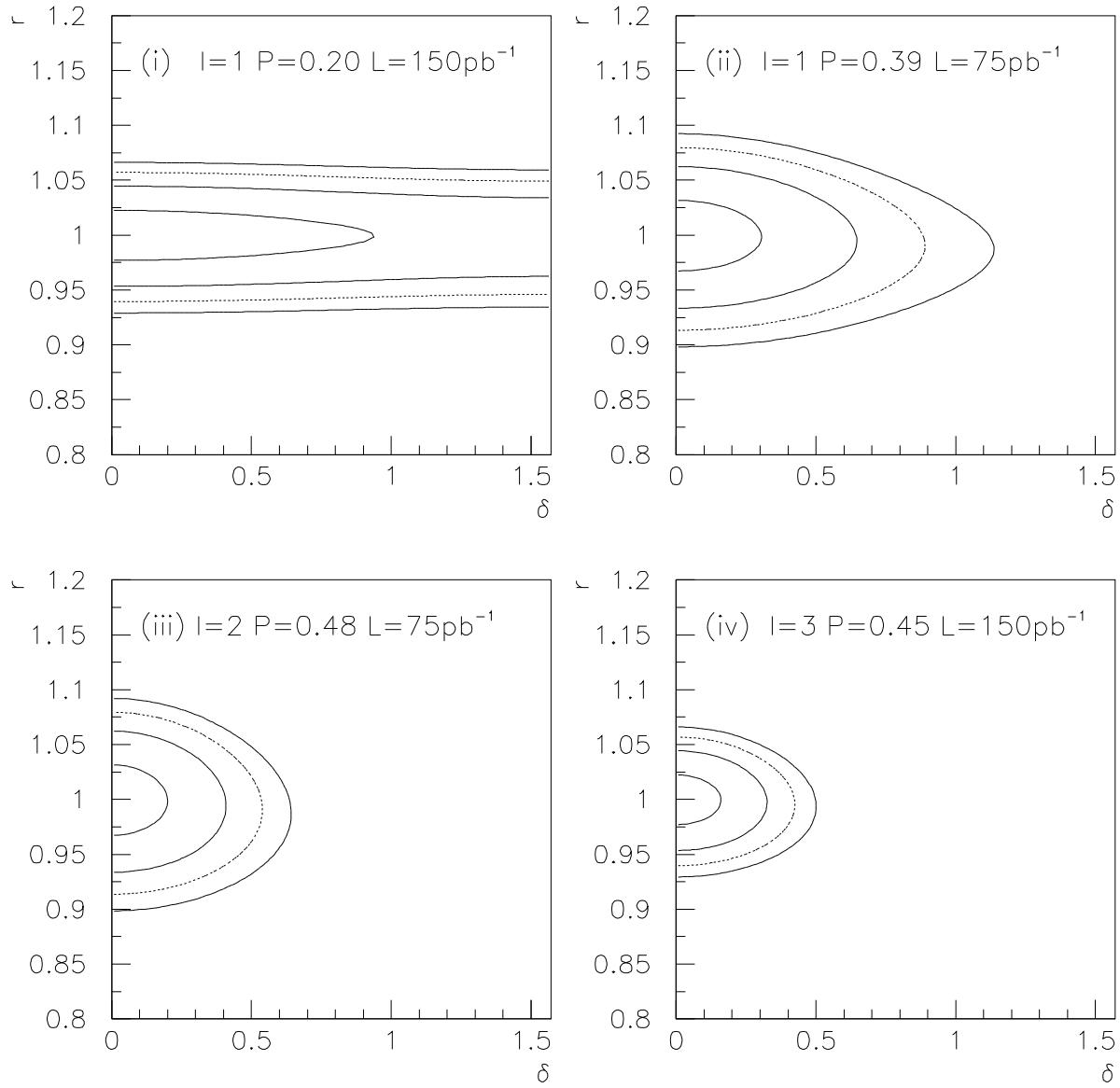
This formula shows that by following the dependence of  $\bar{\sigma}_h(\theta^+, \theta^-)$  on  $N_T$ , one can extract values for  $\cos 2\delta$  and  $\sin 2\delta$ . In practice, it is best to run in several configurations. To approximate the  $\zeta = 0$  configuration, one would choose  $P_H^+ = P_H^- = P_H = 0.05$ ,  $\theta^- = \theta^+$ ,  $P_V^+ = P_V^- = \sqrt{P^2 - P_H^2}$ . To approximate the  $\zeta = \pi$  configuration, choose  $P_H^+ = P_H^- = P_H = 0.05$ ,  $\theta^- = \theta^+ + \pi$ ,  $P_V^- = -P_V^+ = -\sqrt{P^2 - P_H^2}$ . To emphasize the  $\zeta = \pi/2$  and  $\zeta = 3\pi/2$  configurations over many turns of the bunches, we choose  $P_H^- = P$  ( $P_V^- = 0$ ),  $P_H^+ = P_H = 0.05$  and  $P_V^+ = \sqrt{P^2 - P_H^2}$ . To obtain an accurate measurement of  $\delta$ , it is necessary to develop a strategy for maximizing  $\langle P_T^2 \rangle \sqrt{L}$  by selecting only energetic muons to accelerate and combining bunches. Lack of space prevents a detailed description.

To gain a quantitative understanding of how successful such a strategy for determining the CP-nature of the  $h$  can be, let us define  $(\hat{a}, \hat{b}) = (a, b)/(gm_\mu/2m_W)$  and give contours at  $\Delta\chi^2 = 1, 4, 6.635, 9$  in the  $\delta = \tan^{-1} \frac{\hat{b}}{\hat{a}}$ ,  $r = \sqrt{\hat{a}^2 + \hat{b}^2}$  parameter space. We define  $I$  as the proton source intensity enhancement relative to the standard value implicit for the earlier-given benchmark luminosities. We compare four cases: (i) the case of  $P = 0.2$ ,  $L = 0.15 \text{ fb}^{-1}$ , which corresponds to  $I = 1$  and the polarization level naturally achieved without any special selection against slow muons; (ii) we maintain the same proton intensity,  $I = 1$ , select faster muons to the extent that it becomes possible to merge neighboring muon bunches, leading to  $P^m(I = 1) \sim 0.39$  and  $L = 0.075 \text{ fb}^{-1}$ ; (iii) we increase the proton source intensity by a factor of two,  $I = 2$ , while selecting faster muons and merging the bunches, corresponding to  $P^m(I = 2) \sim 0.48$  and  $L = 0.075 \text{ fb}^{-1}$ ; finally (iv) we employ  $I = 3$  and use so-called ‘just-full bunches’, corresponding to  $P^f(I = 3) \sim 0.45$  and  $L = 0.15 \text{ fb}^{-1}$ . Results in the case of a SM Higgs boson with  $m_{h_{SM}} = 130 \text{ GeV}$  are presented in Fig. 12. One sees that a 30% ( $1\sigma$ ) measurement of  $\hat{b}/\hat{a}$  is possible without increased proton source intensity, using the simple technique of selecting fast muons and performing bunch merging. An  $\lesssim 20\%$  measurement would require a moderately enhanced proton source intensity.

After studying a number of cases, the overall conclusion of [36] is that this procedure will provide a good CP determination (superior to other techniques) provided one merges bunches and compensates for the loss of luminosity associated with selecting only energetic muons (so as to achieve high average polarization) by having a proton source that is at least two times as intense as that needed for the studies discussed in previous sections (that do not require large transverse polarization).

## X CONCLUSIONS AND PLANNING FOR FUTURE FACILITIES

Around 2006 the LHC will begin taking data, hopefully revealing the path that particle physics will take in the next century. At the moment there are a few experimental hints suggesting that a Higgs boson might be just around the corner, and there are intriguing indications from the anomalous magnetic moment of the muon that supersymmetric particles may be easily detected at the LHC. This scenario would present a strong argument for the construction of a LC to study this interesting physics which would be at a scale light enough to be probed. A muon collider could play a crucial role in several ways. First, a  $s$ -channel light-Higgs factory would provide crucial precision measurements of the  $h^0$  properties, including the only accurate measurement of its  $\mu^+\mu^-$  coupling. Deviations of these properties with respect to expectations for the SM Higgs boson can, in turn, impose critical constraints on the masses of heavier Higgs bosons and other SUSY parameters. Among other things, the heavier Higgs bosons might be shown to definitely lie within reach of muon collider  $s$ -channel production. Further, it could be that the heavier  $H^0$  and  $A^0$  cannot be detected at the LHC or LC (a scenario that arises, in the MSSM for example, for moderate  $\tan\beta$  values and  $m_{A^0} \sim m_{H^0} \gtrsim 250 \text{ GeV}$ ). Since their detection in  $s$ -channel production at the muon collider would be relatively certain, the muon collider would



**FIGURE 12.** Contours at  $\Delta\chi^2 = 1, 4, 6.635, 9$  for the  $\hat{a}$  and  $\hat{b}$  measurement for a SM Higgs ( $\hat{a} = 1, \hat{b} = 0$ ) with  $m_{h_{SM}} = 110$  GeV for the four luminosity/bunch-merging options outlined in the text. Here,  $\delta = \tan^{-1} \frac{\hat{b}}{\hat{a}}$  and  $r = \sqrt{\hat{a}^2 + \hat{b}^2}$ . For small  $\delta$ ,  $\frac{\hat{b}}{\hat{a}} \sim \delta$ .

be an essential component in elucidating the full physics of the Higgs sector. Further, there are even (non-supersymmetric) scenarios in which one only sees a SM-like Higgs as the LHC and LC probe scales below a TeV, but yet muon collider Higgs factory studies would reveal additional Higgs bosons. Using  $s$ -channel Higgs production, a muon collider would also provide particularly powerful possibilities for studying the CP nature of the Higgs boson(s) that are found. Such CP determination might be absolutely crucial to a full understanding of the Higgs sector. Finally, one should not forget that the muon collider might prove to be the best approach

to achieving the highest energies possible in the least amount of time. Construction of a Higgs factory would be a vital link in the path to high energy.

## REFERENCES

1. R. Barate *et al.* [ALEPH Collaboration], Phys. Lett. B **495**, 1 (2000) [hep-ex/0011045].
2. P. Abreu *et al.* [DELPHI Collaboration], Phys. Lett. B **499**, 23 (2001) [hep-ex/0102036].
3. M. Acciarri *et al.* [L3 Collaboration], Phys. Lett. B **495**, 18 (2000) [hep-ex/0011043].
4. G. Abbiendi *et al.* [OPAL Collaboration], Phys. Lett. B **499**, 38 (2001) [hep-ex/0101014]. [1–5]
5. A. N. Okpara, hep-ph/0105151.
6. H. N. Brown *et al.* [Muon g-2 Collaboration], Phys. Rev. Lett. **86**, 2227 (2001) [hep-ex/0102017].
7. A. Czarnecki and W. J. Marciano, hep-ph/0102122.
8. C. M. Ankenbrandt *et al.*, Phys. Rev. ST Accel. Beams **2**, 081001 (1999) [physics/9901022].
9. R. Raja and A. Tollestrup, Phys. Rev. D **58**, 013005 (1998) [hep-ex/9801004].
10. V. Barger, M. S. Berger, J. F. Gunion and T. Han, Phys. Rept. **286**, 1 (1997) [hep-ph/9602415].
11. V. Barger, M. S. Berger, J. F. Gunion and T. Han, Phys. Rev. Lett. **75**, 1462 (1995) [hep-ph/9504330].
12. P. Chankowski *et al.*, Phys. Lett. **B496** (2000) 195 [hep-ph/0009271].
13. B. Autin et.al., CERN-99-02.
14. W. J. Murray, hep-ph/0104268.
15. V. Barger, T. Han and C.G. Zhou, Phys. Lett. **B480**, 140 (2000).
16. B. K. Bullock, K. Hagiwara and A. D. Martin, Nucl. Phys. **B395**, 499 (1993).
17. D. Asner, J. Gronberg, J. Gunion and T. Hill, in preparation. For a summary, see the  $\gamma\gamma$  collider section of the Snowmass 2001 Proceedings.
18. M. Battaglia and K. Desch, hep-ph/0101165.
19. J. F. Gunion, L. Poggioli, R. Van Kooten, C. Kao and P. Rowson, hep-ph/9703330.
20. P. Chankowski, T. Farris, B. Grzadkowski, J. F. Gunion, J. Kalinowski and M. Krawczyk, Phys. Lett. B **496**, 195 (2000) [hep-ph/0009271].
21. D. E. Groom *et al.* [Particle Data Group Collaboration], Eur. Phys. J. C **15**, 1 (2000).
22. J. F. Gunion, “Detecting and studying Higgs bosons,” in *Perspectives on Higgs Physics, II*, ed. G.L. Kane, World Scientific Publishing (1997), hep-ph/9705282.
23. M. S. Berger, hep-ph/0105128.
24. M. S. Berger, Phys. Rev. D **41**, 225 (1990).
25. J.F. Gunion and J. Kelly, Phys. Rev. **D56** (1997) 1730 [hep-ph/9610495].
26. J.F. Gunion and J. Kelly, hep-ph/9610421.
27. J.L. Feng and T. Moroi, Phys. Rev. **D56** (1997) 5962 [hep-ph/9612333].
28. V. Barger, M. S. Berger, J. F. Gunion and T. Han, Phys. Rev. Lett. **78**, 3991 (1997) [hep-ph/9612279].
29. D. J. Miller, S. Y. Choi, B. Eberle, M. M. Muhlleitner and P. M. Zerwas, Phys. Lett. B **505**, 149 (2001) [hep-ph/0102023].
30. T. Farris and J.F. Gunion, in preparation.
31. J.F. Gunion, in “Transparency Book”, Higgs Factory Workshop, Feb. 28 – March 1, 2001, UCLA.
32. B. Grzadkowski and J. F. Gunion, Phys. Lett. B **350**, 218 (1995) [hep-ph/9501339].
33. E. Asakawa, S. Y. Choi and J. S. Lee, Phys. Rev. D **63**, 015012 (2001) [hep-ph/0005118].
34. A. Pilaftsis, Phys. Lett. B **435**, 88 (1998) [hep-ph/9805373].
35. A. Pilaftsis, Phys. Rev. D **58**, 096010 (1998) [hep-ph/9803297].
36. B. Grzadkowski, J. F. Gunion and J. Pliszka, Nucl. Phys. B **583**, 49 (2000) [hep-ph/0003091].

# Extending Luminosity in $\mu^+ \text{-} \mu^-$ Collider Scenarios

D. Neuffer

In Table 1 we present proposed parameters for  $\mu^+ \text{-} \mu^-$  colliders as developed for the 1999 status report[1]. These parameters are used as starting points for the discussion of the potential of future  $\mu^+ \text{-} \mu^-$  colliders. These parameters and their luminosity numbers were developed under the constraints that:

- The parameters were within reasonable reach of a collider facility, using feasible parameters of a proton source, target and  $\pi \rightarrow \mu$  collection system, ionization cooling, recirculating-linac acceleration and a collider ring.
- The parameters would provide at least the minimal luminosity needed for the exploration of new physics in the proposed colliders.
- The physics reach of the proposed facility would be comparable (or better, in Higgs physics) than that of other concurrently proposed facilities, such as  $e^+ \text{-} e^-$  colliders.

Since then the posted potential luminosities of proposed  $e^+ \text{-} e^-$  colliders have increased through incremental improvements, without any fundamental changes in the collider approaches. We would now like to explore the potential of extending the parameters of the 1999 baseline scenarios to higher luminosity; firstly, under the consideration that the same general design approaches are used but that the operating parameters are incrementally extended to increase luminosity, and, secondly, under the possibilities that substantially different approaches are used for some facility components to obtain dramatically increased luminosities. In this section we only address the possibility of incremental improvements.

The formula for luminosity is:

$$L = \frac{f_0 n_s n_b \gamma_\mu N_\mu^2}{4\pi \epsilon_N \beta^*}, \quad [1]$$

where  $n_s$  is the luminosity lifetime (in turns) in the collider,  $n_b$  is the number of colliding bunches in each beam,  $N_\mu$  is the number of muons per bunch,  $\gamma_\mu = (E_\mu/m_\mu)$ ,  $\epsilon_N$  is the normalized emittance, and  $\beta^*$  is the collider focusing parameter, with the beam size at collisions given by  $\sigma^2 = \epsilon_N \beta^* / \gamma_\mu$ .

An important parameter that limits luminosity in colliders is the beam-beam tune shift (per collision), which is:

$$\Delta v = \frac{3r_\mu N_\mu}{2\epsilon_N \gamma_\mu^2} \quad [2]$$

where  $r_\mu$  is the muon electromagnetic radius. In the status report, it was assumed that we were limited to  $\Delta v < 0.05$ , which is the limit in medium-energy circular  $e^+ \text{-} e^-$  colliders. However, in a  $\mu^+ \text{-} \mu^-$  collider, beam storage is required for only  $\sim 1000$  turns and the beams are continuously decaying. Simulations show that  $\Delta v$  as large as 0.015 or more can be used without difficulty.[2]

Assuming that  $\Delta v$  is a controlling limitation, luminosity can be written as:

$$L = \frac{f_0 n_S n_B \gamma_\mu^3 N_\mu \Delta v}{6\pi\beta^*} . \quad [3]$$

In this section we will now consider the implications of increasing L from the values in Table 1, considering eqs [1-3].

Increasing  $\Delta v$  implies increasing  $N_\mu$  or decreasing  $\epsilon_N$ . The total number of muons is  $N_{tot} = n_B N_\mu$ . If  $n_B > 1$ ,  $N_\mu$  can be increased by combining bunches (without increasing  $N_{tot}$ ). The high-energy scenarios of table 1 (0.4 and 3 TeV) have two bunches (each) of  $\mu^+$  and  $\mu^-$ , so these bunches could be combined, doubling the luminosity. Reducing  $\epsilon_N$  requires more transverse cooling (or emittance exchange). The beams in the 0.1 TeV scenarios can probably receive further transverse cooling; cooling to  $\epsilon_N = 10^{-4}$  m-rad should be possible without increasing the longitudinal emittance.

The total number of muons  $N_{tot} = n_B N_\mu$  can be increased by increasing the power of the initial proton beam on the production target, and/or by increasing the efficiency of the  $\mu$  collection and cooling system. The 4 MW proton source is already at fairly high intensity, but successful design and/or operation at that intensity will suggest incremental improvements and an increase to  $\sim 10$  MW will eventually be possible. In this increase it is uncertain how much one would rely on more protons per pulse and how much would be obtained by increasing the pulsing frequency.

The number of storage turns  $n_s$  is set by the collider circumference and beam energy; it could be slightly increased by using higher-field bending magnets to shorten the circumference.

The other remaining parameter is  $\beta^*$ , the focussing strength at the interaction point. This parameter has several constraints related to beam parameters, and substantial reduction in  $\beta^*$  may require added beam cooling. In general it is limited by the “hourglass effect” to approximately the bunch length and in the luminosity calculations we have assumed it to be equal to the bunch length. This limit could be reduced by reducing the bunch length, which implies a lower longitudinal emittance and/or increasing collision energy spreads, which are limited by the lattice chromaticity acceptances.  $\beta^*$  can actually be reduced below  $\sigma_z$ , with continuing luminosity increase.

In Table S-1, we display the baseline luminosity parameters of the 1999 status report. In table S-2 we generate some “upgraded” luminosity numbers. The upgraded numbers are obtained by increasing the number of muons by 2.5 times, which implies more proton power ( $\sim 2 \times$ ) and greater collection efficiency ( $\sim 1.25 \times$ ), and by incremental improvements ( $\sim 0.7 \times$  emittance and  $\beta^*$  reduction) in cooling and focusing. We have not enforced the assumption that 6-D emittances must be the same; this is probably realistic since the final emittance exchanges for minimizing  $\epsilon_t$  are expected to increase 6-D emittances. In the tables we assume  $\beta^* = \sigma_z$  so that lowering  $\beta^*$  requires reduced longitudinal emittance; this is a bit pessimistic since  $\beta^*$  could be reduced to  $\sim$ half of  $\sigma_z$  and luminosity would increase by  $\sim 50\%$ , offsetting additional cooling requirements. We have also not decreased collider circumferences, which would proportionately increase luminosity. However, the power, focusing, and cooling upgrades would increase luminosity by an order of magnitude.

As a caveat, the “upgraded” luminosities would require improvements on  $\mu^+ - \mu^-$  collider technologies which are not yet fully established, and are certainly less established than the status

report numbers. However the present discussion indicates that the status report luminosities are not absolute limits, and would be subject to incremental improvement in an operating collider.

An intriguing question is whether there are “advanced” technologies which could increase  $\mu^+ \text{-} \mu^-$  collider performance beyond the status report baselines. Among these are “low-energy” cooling for stopped (or nearly stopped) muons, optical stochastic cooling, beam-beam force neutralization at the interaction point, plasma lens focusing (in the cooling or in the collider), etc. Possible gains by such methods are discussed by B. King et al.[[1](#)]

**Table S-1 Initial  $\mu^+ - \mu^-$  Collider parameters (“Higgs factory” and 3TeV Colliders) [ref. 1]**

Beam properties at capture	100GeV high-resolution Higgs Factory	100GeV high-luminosity Higgs Factory	3TeV High-Energy Collider
<b>Collision Energy</b>			
E <sub>μ</sub> /beam	100	100	3000 GeV
σ <sub>x</sub>	10 cm	295 μm	3.2 μm
σ <sub>x'</sub>	70	2.1	1.1 mr
σ <sub>z</sub>	680	14.1	0.3 cm
σ <sub>p</sub> / p	0.65	3 × 10 <sup>-5</sup>	0.0012 0.0016
ε <sub>xN</sub>	~20	0.29	0.05 mm
ε <sub>zN</sub>	~1000	2.02	70 mm
ε <sub>6N</sub>	~4 × 10 <sup>5</sup>	0.17	0.17 mm <sup>3</sup>
β*	14.1	4.1	0.3 cm
Δv	0.015	0.051	0.044
n <sub>tuens</sub>	450	450	785
F	15	15	15 Hz
N <sub>b</sub>	1	1	2 Bunch/beam/spill
N	4 × 10 <sup>12</sup>	4 × 10 <sup>12</sup>	2 × 10 <sup>12</sup> Particles/bunch
L	10 <sup>31</sup>	1.2 × 10 <sup>32</sup>	7 × 10 <sup>34</sup> cm <sup>-2</sup> ·s <sup>-1</sup>

**Table S-2 “High-Luminosity”  $\mu^+ - \mu^-$  Collider parameters (“Higgs factory” and 3TeV Colliders)**

Beam properties at capture	100GeV high-resolution Higgs Factory	100GeV high-luminosity Higgs Factory	3TeV High-Energy Collider
<b>Collision Energy</b>			
E <sub>μ</sub> /beam	100	100	3000 GeV
σ <sub>x</sub>	10 cm	160 μm	2.2 μm
σ <sub>x'</sub>	70	1.6	2.6 1.1 mr
σ <sub>z</sub>	680	10	2 0.2 cm
σ <sub>p</sub> / p	0.65	3 × 10 <sup>-5</sup>	0.0011 0.0016
ε <sub>xN</sub>	~20	0.12	0.065 0.035 mm
ε <sub>zN</sub>	~1000	1.4	10 45 mm
ε <sub>6N</sub>	~4 × 10 <sup>5</sup>	0.02	0.045 0.055 mm <sup>3</sup>
Δv	0.044	0.083	0.157
F	15	15	15 Hz
N <sub>b</sub>	2	2	2 Bunch/beam/spill
N	5 × 10 <sup>12</sup>	5 × 10 <sup>12</sup>	5 × 10 <sup>12</sup> Particles/bunch
L	1.1 × 10 <sup>32</sup>	1.0 × 10 <sup>33</sup>	1 × 10 <sup>36</sup> cm <sup>-2</sup> ·s <sup>-1</sup>

## **References**

- [1] C. M. Ankenbrandt et al., Phys. Rev. STAB 2, 081001 (1999).
- [2] M. Furman, BF-19/CBP-Note-169/LBL-38563, 1996
- [3] B. King, ed. “High-energy  $\mu^+ - \mu^-$  Colliders,” in progress, (2001).

# Neutrino Factory Parameters from Feasibility Studies 1 and 2

D. Neuffer et al.

A  $\mu$ -storage ring neutrino factory has somewhat different requirements on the muon beams from those of a  $\mu^+ \mu^-$  collider. The  $\nu$ -source event rate depends primarily on the number of stored muons and not on the quality of the  $\mu$ -beam; therefore the beam needs only to be cooled sufficiently to be within the acceptance of the accelerator and storage ring, and not to a minimal emittance for high luminosity. Also the  $\mu$ -beam need not be confined within single bunches, but can be distributed in a string of bunches.

The MC collaboration has produced two detailed Neutrino Factory design studies[1,2]. Both studies have a similar muon collection and cooling system. A schematic layout of the Study 1 scenario is shown in Fig. 1a. Following the target there is a 50 m long drift, a 100 m long induction linac for phase rotation, a mini-cooling stage, a 17 m long buncher and a 140 m long cooling section. The Study 2 scenario layout is shown in Fig. 1b. It has a 18 m drift following the target, a 108 m induction linac, a mini-cooling stage, an additional 200 m of induction linac and drift to complete the phase rotation, a 55 m buncher and a 108 m cooling section, for a total length of 540 m, about 200m longer than Study 1. Most of the following description will concentrate on the more recent Study 2 case, which is an improvement of Study 1.

## I. $\nu$ -Factory precooler

In the neutrino factory studies a proton bunch on the target produces a  $\pi$  beam, which is then allowed to drift while the  $\pi$ 's decay into  $\mu$ 's and the beam develops a position-energy correlation with the lower-energy  $\mu$ 's trailing behind the higher energy  $\mu$ 's. The energy spread of the muons is very large, much larger in fact than the acceptance of the following cooling stage and accelerators. Therefore this drift is followed by an induction linac system that decelerates high-energy  $\mu$ 's and accelerates low energy  $\mu$ 's, reducing the energy spread. The voltage pulse across the gaps in the induction linac cells can be tailored to match the time-energy correlation of the incoming beam bunch. In Study 2 the induction linac is broken up into two sections in order to reduce the distortion in the resulting longitudinal phase space of the muons. The fractional energy spread of the beam after the induction linacs is reduced to 3.7%. However, the rms bunch width grows to 27 m ( $\sim$ 100 m full width). The induction linac system includes a 3.5 m long liquid H<sub>2</sub> absorber. This "mini-cooler" stage provides an  $\sim$ 20% reduction in transverse normalized emittance to  $\epsilon_{t,\text{rms}} \approx 0.012$  m.

After the induction linacs, the muons are distributed continuously over a bunch length of around 100 m. It is then necessary to form the muons into a train of bunches for cooling and subsequent acceleration, as well as to match the beam transversely into the focusing lattice used for cooling. Thus the beam is transported into a 201 MHz rf buncher section, which forms the beam into about 70 bunches. The transverse and longitudinal functions of this section are performed sequentially for design simplicity. First an 11 m long magnetic lattice section is used to match the beam from the approximately uniform solenoidal field used in the induction linacs to the so-called "super-FOFO", or sFOFO, focussing lattice used in the remainder of the front end. This is followed by the 55 m long rf buncher, which consists of rf cavity sections interspersed with drift regions.

The buncher magnetic lattice is identical to that used in the first cooling section. It contains rf cavities in selected lattice cells and no absorbers. The main rf frequency is chosen to be 201.25 MHz in the front end, so that the beam would fit radially inside the cavity aperture and because

power sources and other technical components are available at this frequency. The 201.25 MHz cavities are placed at the high beta regions in this lattice, just as in the cooling section. Maximum bunching efficiency was obtained by breaking the region into three rf stages separated by drift regions. Second harmonic cavities (402.5 MHz) are added at the entrance and exit of the first and second stages to linearize the shape of the rf pulse. The buncher encompasses 20 lattice cells, each 2.75 m long. By the end of the buncher, most, but not all, particles are within the 201.25 MHz buckets. About 25% are outside the bucket and are lost relatively rapidly, and another 25% are lost more slowly as the longitudinal emittance rises from straggling and the negative slope of the energy loss with energy.

## II. sFOFO Cooling System

In the Neutrino Factory the rms transverse emittance of the muon beam emerging from the induction linac must be reduced to  $\sim 2$  mm-rad (normalized) in order to fit into the downstream accelerators, and be contained in the storage rings. Ionization cooling is currently our only feasible option. The cooling channel described below is based on extensive theoretical studies and computer simulations.

Solenoidal fields are used for focusing; however, energy-loss cooling within a constant (or same-sign) field leads to an increasing beam angular momentum through the cooling channel. The solenoidal field must flip sign, while maintaining good focusing throughout the beam transport and low  $\beta_{\perp}$  at the absorbers. One of the simplest solutions is to vary the field sinusoidally; this is the “FOFO” lattice. The cooling system of Study 2 is composed of “sFOFO” or “super-FOFO” lattice cells. (See Fig. 2.) Each of these cells includes an absorber for energy-loss cooling, an rf cavity of beam reacceleration and solenoids for transverse focusing, with the focusing designed to minimize beam size in the absorbers. The sFOFO lattice uses alternating solenoids like the FOFO, but is a bit more complicated. As in the FOFO case, the longitudinal B-field vanishes at the  $\beta_{\perp,\min}$  position, located at the center of the absorber. This is accomplished by placing two short, strong “focussing” coils about the absorber, running in opposite polarities. The field is decreased and flattened outside the absorber region, due to a “coupling” coil located around the linac.

The Study 2 cooling channel operates at a nominal momentum of 200 MeV/c. There are six sections with steadily decreasing  $\beta_{\perp,\min}$ . In the first three the lattice half period length is 2.75 m, and in the last three sections this half period length is 1.65 m. The matching sections between these sections also consist of cooling cells, which differ from the regular cooling sections only by the currents (except for matching between different cell lengths, where the length is also changed).

Each lattice half-period includes a multicell linac, and to increase the useful gradient of the accelerating cavities, the cell irises are covered with a foil or grid. The baseline design calls for thin, pre-stressed beryllium foils with thicknesses that increase with radius. An accelerating gradient of  $E = 16$  MV/m is required in each linac.

The absorber material is liquid hydrogen ( $LH_2$ ). The length of these absorbers is 35 cm for the 2.75 m lattices and 21 cm for the 1.65 m lattices, respectively. The  $LH_2$  vessels must also be equipped with thin aluminum windows. Their thickness is 360  $\mu$ m (220), with a radius of 18 (11) cm, for the 2.75 m and 1.65 m lattices, respectively. The muons therefore lose  $\sim 12$  MeV per lattice cell for the 2.75 m lattices and  $\sim 7$  MeV for the 1.65 m lattices.

The complete Study 2 cooling system contains 16 2.75m cells, 36 1.65m cells, and a 4.4 m matching section between them for a total length of  $\sim$ 108 m. Complete descriptions of the system and detailed simulations of its cooling performance are presented in Study 2 [2] and summarized below.

### III. Simulation results

An important accomplishment in the collaboration has been the development of the simulation codes ICOOL [3] and DPGeant [4], which include the full complexity of the absorber + rf + solenoid system, including all materials and their properties, magnets defined in terms of coils, currents, and positions (rather than actual focusing fields), rf defined with complete cavity fields, etc. The goal is to establish cooling systems which function when described in full complexity. These simulations confirm that the Study 1 and 2 cooling systems will perform as planned. In this section we describe these simulation results.

The  $\beta_{\perp,\min}$  function, calculated at the absorber centers using the beam second-order moments calculated in Geant4 simulations, is shown in Fig. 3. This function is reduced with each new section of the cooling lattice. The transverse and longitudinal emittances as calculated through the cooling system are shown in Fig. 4. Emittances are computed using diagonalized covariance matrices. The emittance values are corrected for correlations between the variables, including the strong momentum-transverse amplitude correlation. At the end of the cooling channel a transverse emittance of 2.2 mm rad is reached. The longitudinal emittance shows an initial rise and fall as particles not within the rf bucket increase in amplitude and are later lost, and then an approach to an asymptotic value set by the bucket size. The longitudinal emittance should rise due to straggling and the negative slope of energy loss with energy, but, since the rf bucket is already full, we see a steady loss of particles instead of an emittance growth.

Despite the overall particle loss, the numbers of particles within the accelerator acceptance increases, as seen in Fig. 5. The gain in muons within the accelerator acceptance of 150 mm due to cooling is  $\sim 3.1 \times$  (or  $5 \times$  if the Study 1 acceptances were used). If the particle loss from longitudinal emittance growth could be eliminated, as should be the case if emittance exchange were used, then these gains could double.

**Table 1 Beam characteristics summary**

Location (end of)	$\sigma_x$	$\sigma_{x'}$	$\sigma_p$	$\sigma_i$ (per bunch)	$\langle p \rangle$
	Cm	mrad	MeV/c	ns	MeV/c
Induction linac	8.6	95	113		260
Matching section	5.8	114	113		260
Buncher	5.3	107	111	0.84	256
2.75 m cooler	3.0	91	70	0.55	226
1.65 m cooler	1.8	102	30	0.51	207

The rms beam characteristics in the buncher and cooler sections are summarized in Table 1. The beam is symmetric in this lattice, so the y properties are similar to those in x. We see that the size steadily decreases as we proceed down the channel. The angular divergence is kept constant for maximum cooling efficiency. The momentum spread of the entire beam is still large after the

induction linac, but this includes very low and high energy muons that do not get transmitted through the subsequent sFOFO lattice. The decrease in energy spread is due to particle losses, since there is no longitudinal cooling or emittance exchange in this lattice. These losses could be controlled by adding some longitudinal cooling to the channel.

#### **IV. Alternatives for v-Factory cooling**

We have presented in detail only one example of a cooling scenario for preparing the  $\mu$ -beams of a v-factory, the baseline cooling scenario for Feasibility Study 2. Other cooling scenarios could be used and future studies will explore alternative configurations, either by optimizing the present proposal or developing a substantially different but superior system. However any cooling scenario would also require: absorbers for energy-loss, acceleration for longitudinal energy recovery, and a transport lattice with strong focusing of the beam into the absorbers.

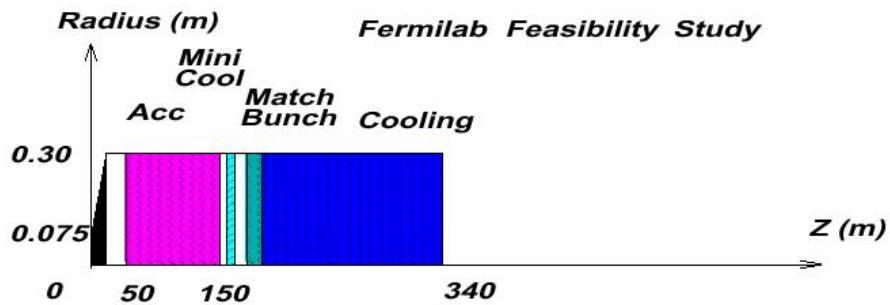
The liquid hydrogen absorbers were chosen because hydrogen has the least multiple scattering; however other low-Z material (LiH, Li, Be, ...) could also be used and would avoid the mechanical difficulties of handling liquid hydrogen, at the cost of more scattering. It is likely that such denser materials may be necessary for emittance exchange wedges.

The sFOFO focusing system was used here, but other lattices could be used. An attractive alternative is the “double-flip” scenario [5], which has long constant or same-sign field sections with only two changes in sign. This is a simpler lattice, but it requires more field volume for the same focusing effect as the sFOFO. Lattices that incorporate energy cooling could also be preferable (see below).

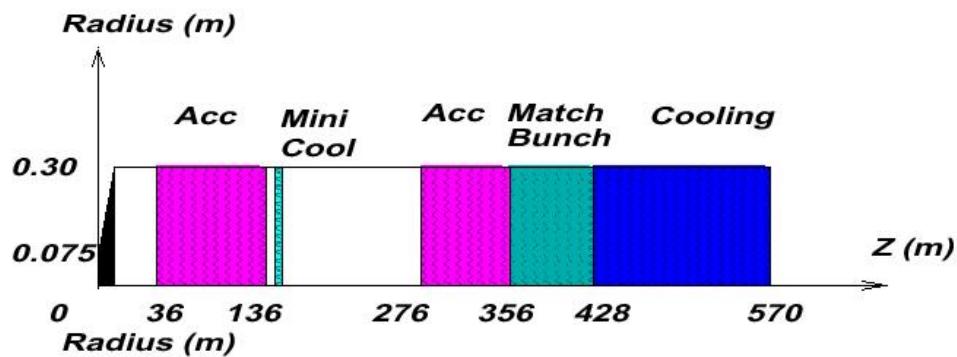
The ~200 MHz rf system was based on the perception that 200 MHz rf would be available and affordable. A low-frequency system (40-80 MHz) for capture and cooling has been proposed at CERN [6] and has some preferable properties. It would develop fewer  $\mu$ -bunches per primary p-bunch and would be more adaptable to future  $\mu^+ \mu^-$  collider beams.

## References

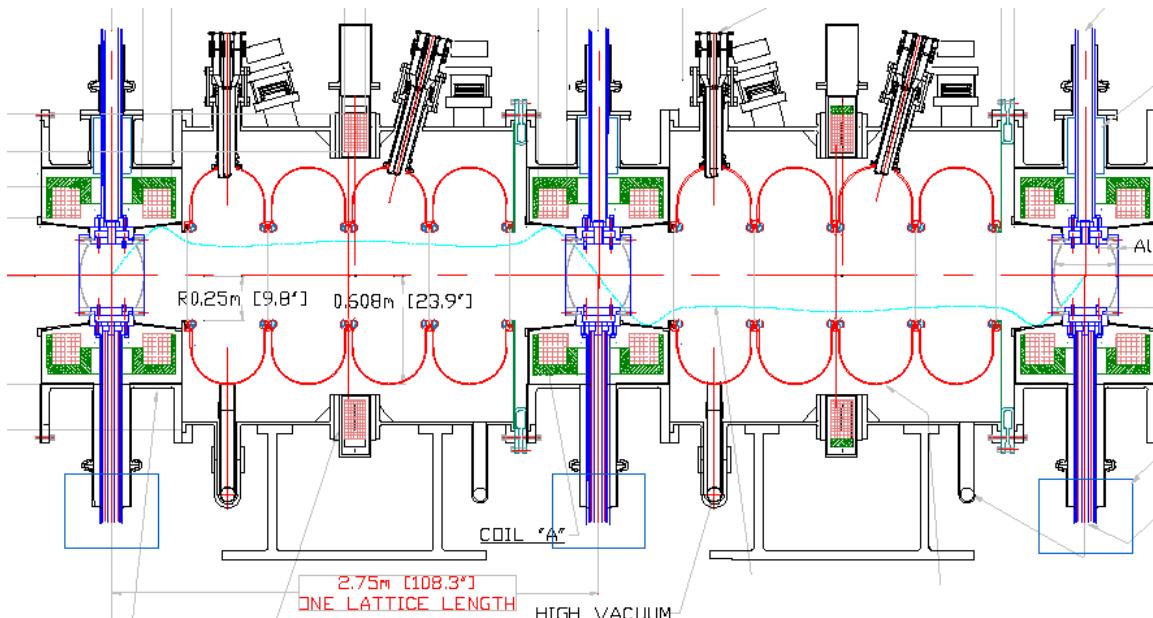
- [1] N. Holtkamp and D. Finley, ed. “A Feasibility Study of a Neutrino Source Based on a Muon Storage Ring”, Fermilab, April 15<sup>th</sup>, 2000.
- [2] S. Ozaki, R. Palmer, and M. Zisman, ed., “BNL  $\nu$ -Factory Feasibility Study”, April, 2001.
- [3] R. Fernow, “ICOOL: a simulation code for ionization cooling of muon beams,” Proc. 1999 Particle Accelerator Conference, New York, p. 3020-2.
- [4] The DPGEANT and GEANT4 codes have been developed at FNAL under the direction of P. Lebrun.
- [5] V. Balbekov, “Comparison of an alternate solenoid and long solenoid cooling channels”, NuMu note 130, 2000.
- [6] A. Lombardi, “a 40-80 MHz system for phase rotation and cooling”, CERN–NUFACT Note 34, 1 August 2000.



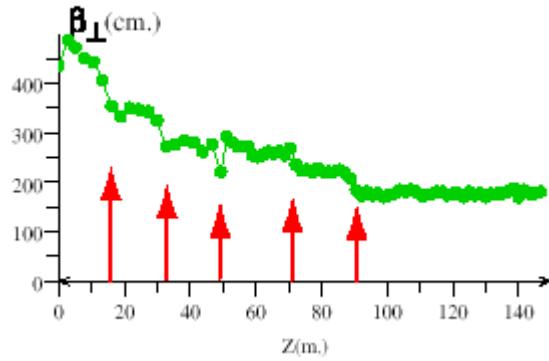
**Figure 1a.** Study 1 layout of  $\mu$ -capture and cooling transport.



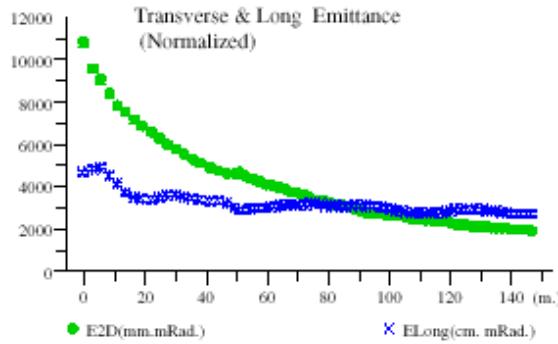
**Figure 1b.** Study 2 layout of  $\mu$ -capture and cooling transport



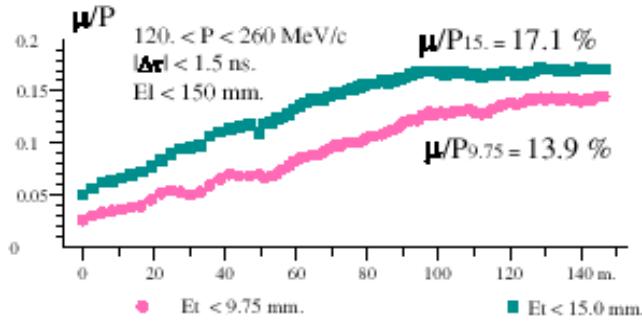
**Figure 2.** Cross-section of two Study 2 2.75m sFOFO lattice cooling cells, showing two 4-cell rf cavities, 3 Liquid hydrogen absorbers, and magnetic coils for focusing around the absorbers and rf cavities.



**Figure 3.** The  $\beta_{\perp,\min}$  function in mm at cavity centers, as calculated from the second order moments of a beam in GEANT4 simulations, for the entire sFOFO cooling channel. The 5 arrows indicate the beginning of new lattice sections. (In this figure the cooling channel is extended beyond the Study2 reference length of 108 m to 144 m by adding 1.65 m cells.)



**Figure 4.** The transverse and longitudinal normalized emittances along the cooling channel.



**Figure 5.** The muon to proton yield ratio for the two emittance cuts, showing that the particle density in the center of the phase space increases as the beam is cooled in the channel. The two curves give the number of particles within the baseline longitudinal and transverse acceptances. The upper line represents the values for the accelerator parameters in this Study. The lower line, given for comparison, gives the values for the acceptances used in Feasibility Study 1.

# Proton Driver Section

W. Chou

Fermilab has recently completed a design study for a Proton Driver. It describes in detail the design of a new brighter booster, called the Proton Driver, as a complete functional replacement for the present Booster. The design report has been published as a technical memo FERMILAB-TM-2136. It can also be downloaded from the web: <http://www-bd.fnal.gov/pdriver/reports.html>.

The Proton Driver is a rapid-cycling (15 Hz), high-intensity ( $3 \times 10^{13}$  protons per pulse), 1 MW 16-GeV synchrotron. It serves a number of purposes in the Fermilab hadron program. In the near term, it replaces the present Booster and increases the proton beam intensity in the Main Injector by a factor of four, thereby providing an upgrade path for NuMI and other 120 GeV fixed target programs. It also opens the avenue for new physics programs based on its stand-alone capabilities as a source of intense proton beams. The beam power of the Proton Driver is a factor of twenty higher than that of the present Booster. It can be employed for the production of high-intensity secondary particle beams of pions, kaons, neutrons and neutrinos. In the long term, the Proton Driver can serve a neutrino factory and a muon collider by generating intense short muon bunches from a target. The design also allows an upgrade path to a 4 MW proton source by adding a 600 MeV linac and a 3 GeV Pre-Booster at some late time (called Phase II). To meet the requirement of a muon collider, such as a Higgs factory, the number of bunches in Phase II will be reduced to four.

The main parameters of the Proton Driver in Phase I and Phase II are listed in Table 1. As a comparison, the parameters of the present proton source are also listed.

**Table 1.** Parameters of the Proton Driver: Present, Phase I and Phase II

Parameters	Present	Phase I (MI, v-Fact)	Phase II (μ-Coll)
<b>Linac</b> (operating at 15 Hz)			
Kinetic energy (MeV)	400	400	1000
Peak current (mA)	40	60	80
Pulse length (μs)	25	90	200
H <sup>-</sup> per pulse	$6.3 \times 10^{12}$	$3.4 \times 10^{13}$	$1 \times 10^{14}$
Average beam current (μA)	15	81	240
Beam power (kW)	6	32	240
<b>Pre-Booster</b> (operating at 15 Hz)			
Extraction kinetic energy (GeV)			3
Protons per bunch			$2.5 \times 10^{13}$
Number of bunches			4
Total number of protons			$1 \times 10^{14}$
Normalized transverse emittance (mm-mrad)			$200 \pi$
Longitudinal emittance (eV-s)			2
RF frequency (MHz)			7.5
Average beam current (μA)			240
Target beam power (MW)			720
<b>Booster</b> (operating at 15 Hz)			
Extraction kinetic energy (GeV)	8	16	16
Protons per bunch	$6 \times 10^{10}$	$1.7 \times 10^{12}$	$2.5 \times 10^{13}$
Number of bunches	84	18	4
Total number of protons	$5 \times 10^{12}$	$3 \times 10^{13}$	$1 \times 10^{14}$
Normalized transverse emittance (mm-mrad)	$15 \pi$	$60 \pi$	$200 \pi$
Longitudinal emittance (eV-s)	0.1	0.4	2
RF frequency (MHz)	53	7.5	7.5
Extracted bunch length $\sigma_t$ (ns)	0.2	1	1
Average beam current (μA)	12	72	240
Target beam power (MW)	0.1	1.2	4

# PARTICLE PRODUCTION AND RADIATION ENVIRONMENT AT A NEUTRINO FACTORY TARGET STATION<sup>1</sup>

N. V. Mokhov<sup>2</sup>, FNAL, Batavia, IL 60510, USA

## Abstract.

Efficient production and collection of a large number of muons is needed to make a neutrino factory based on a muon storage ring viable. The results of extensive MARS simulations are reported for Megawatt proton beams on a carbon rod and mercury jet in a 20-T hybrid solenoid, followed by a matching section and decay channel. Beam energy and power in a 2 to 30 GeV range, beam spot size, beam and target tilt angle, target material and dimensions, and capture system parameters are optimized to get maximum muon yields at the end of the decay channel. Other particles transported down the beam line are also studied for the purpose of beam instrumentation. Prompt and residual radiation distributions are calculated and analysis of target integrity, quench stability and dynamic heat load to the superconducting coils, radiation damage and activation of materials near the beam is performed. Absorption of showers in the direction of a primary beam is considered.

## I MARS MODELING

To achieve adequate parameters of a neutrino factory based on a muon storage ring [1,2] it is necessary to produce and collect large numbers of muons. The system starts with a proton beam impinging on a thick target sitting in a high-field solenoid (20 T, about 1-m long, aperture radius  $R_a=7.5$  cm), followed by a matching section and a solenoidal decay channel (1.25 T, 50-100 m in length,  $R_a=30$  cm) which collects muons resulting from pion decay. Optimization of beam, target and solenoid parameters was done over the years with the MARS code [3] for a  $\mu^+\mu^-$  collider and a neutrino factory (see bibliography in Ref. [4,5]).

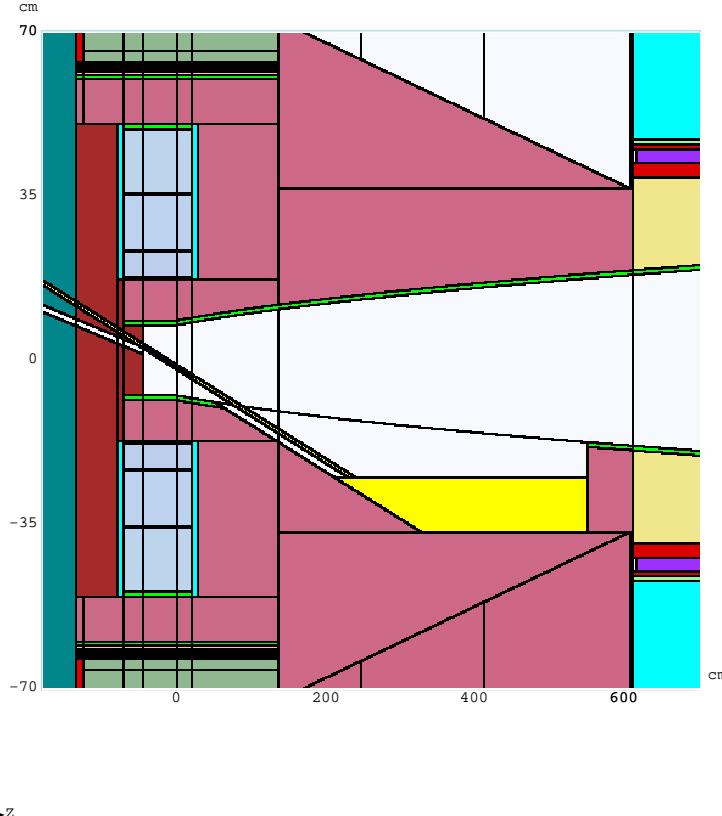
List of targetry issues includes  $\pi/\mu$  production, other particles transported down the beamline, superconducting (SC) coil quench stability, heat loads, radiation damage and activation of materials near the beam, spent proton beam, and numerous shielding issues from prompt radiation to ground-water activation. All these

---

<sup>1)</sup> Work supported by the Universities Research Association, Inc., under contract DE-AC02-76CH03000 with the U. S. Department of Energy.

<sup>2)</sup> mokhov@fnal.gov

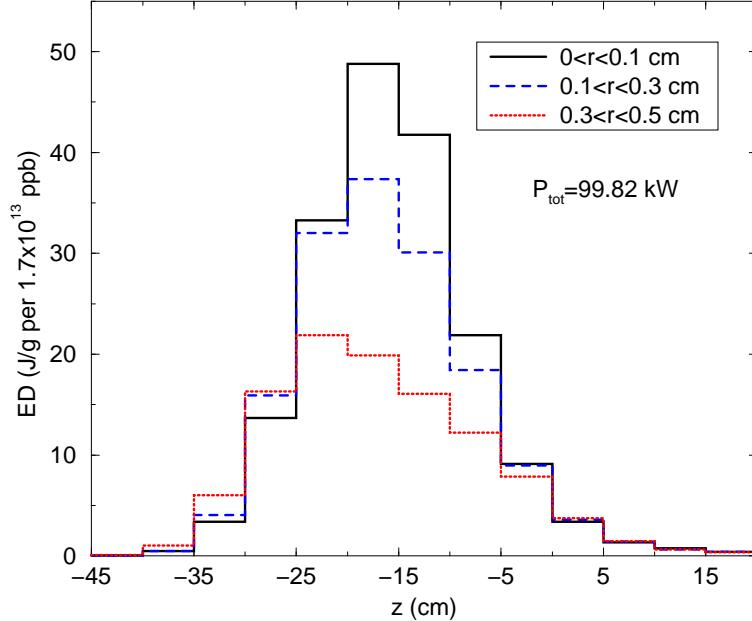
issues were addressed in detailed MARS simulations. Realistic 3-D geometry together with material and magnetic field distributions based on the solenoid magnet design optimization have been implemented into MARS. Graphite (C) and mercury (Hg) tilted targets were studied. A two interaction length target (80 cm for C of radius  $R_T=7.5$  mm and 30 cm for Hg of  $R_T=5$  mm) is found to be optimal in most cases, keeping  $R_T \geq 2.5 \sigma_{x,y}$ , where  $\sigma_{x,y}$  are the beam RMS spot sizes.



**FIGURE 1.** A fragment of the MARS model of target/capture system with tilted proton beam and mercury jet.

The optimized configuration for the Study-2 [2], designed for a 1 MW proton beam of 24 GeV energy (upgradable to 4 MW), is shown in Fig. 1. The beam intensity is  $1.7 \times 10^{13} \text{ ppb} \times 6 \times 2.5 \text{ Hz} = 2.55 \times 10^{14} \text{ p/s}$ , resulting in  $5.1 \times 10^{21} \text{ p/yr}$  at  $2 \times 10^7 \text{ s/yr}$ . The model was optimized for  $-2 < z < 36 \text{ m}$ ,  $r < 1.8 \text{ m}$ . It includes sophisticated coil shielding: water-cooled tungsten-carbide balls at  $z < 6 \text{ m}$  and water-cooled copper at  $z > 6 \text{ m}$ . A proton beam ( $\sigma_x = \sigma_y = 1.5 \text{ mm}$ ,  $\sigma_z = 3 \text{ ns}$ , 67 mrad) interacts with a 5 mm radius mercury jet tilted by 100 mrad, which is ejected from the nozzle at  $z = -60 \text{ cm}$ , crosses the z-axis at  $z = 0 \text{ cm}$ , and hits a mercury pool at  $z = 220 \text{ cm}$ ,  $x = -25 \text{ cm}$ . With such a beam-jet crossing, about 97% of protons have a probability to interact with target material, generating pions and resulting in significant energy deposition in material (Fig. 2) that can at some conditions destroy solid or liquid target. A 8-cm wide mercury pool ( $210 < z < 550 \text{ cm}$ ) is a core

of a sophisticated spent beam absorber. A 2-mm beryllium window at  $z=610$  cm withstands beam-induced heating (with appropriate cooling), but its lifetime is an issue because the absorbed dose in its center reaches tens of GGy/yr.

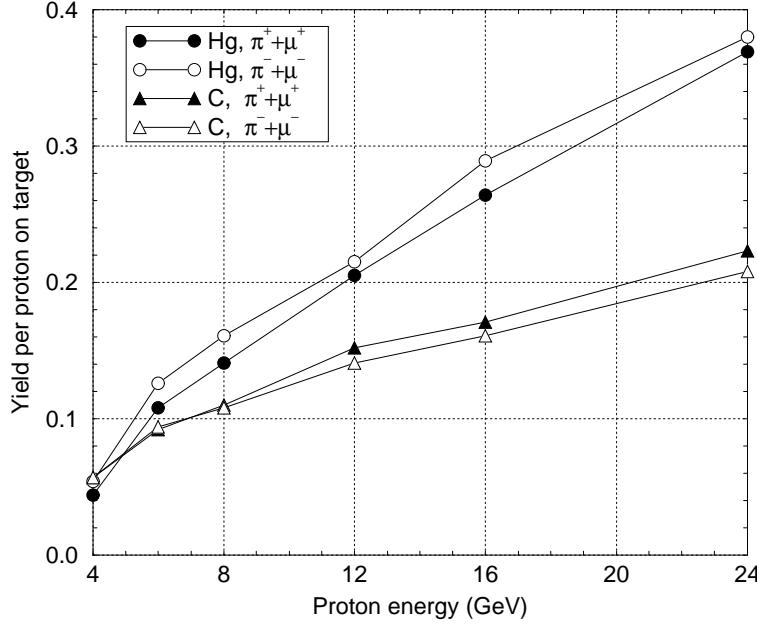


**FIGURE 2.** Longitudinal profiles of the energy density deposited in the mercury jet target in three radial regions.

## II PARTICLE PRODUCTION

Detailed optimizations were performed for the particle yield  $Y$ , defined as a sum of the numbers of  $\pi$ ,  $K$  and  $\mu$  of a given sign and energy interval at the downstream end of the considered system. It turns out that for proton energies  $E_p$  from a few GeV to about 30 GeV, the shape of the low energy spectrum of such a sum is energy-independent and peaks around  $E=130$  MeV, where  $E$  is  $\pi/\mu$  kinetic energy. For the given parameters, the interval of 30 MeV  $< E <$  230 MeV around the spectrum maximum is considered as the one to be captured by a phase rotation system. The yield  $Y$  grows with  $E_p$ , is almost material-independent at low energies and grows with target  $A$  at high energies, being almost a factor of two higher for Hg than for C at  $E_p=16\text{--}30$  GeV (Fig. 3). It is interesting that the yield per beam power, i.e.,  $Y/E_p$  has a broad maximum around 6 GeV. For a 1 to 2 GeV proton beam (CERN, SNS), the optimal target material, from the pion production point of view, is carbon with significantly lower  $\pi^-$  production compared to  $\pi^+$ . To avoid absorption of spiraling pions by target material, the target and beam are tilted by an angle  $\alpha$  with respect to the solenoid axis. The yield is higher by up to 30% for the tilted target with a broad maximum around  $\alpha=100$  mrad. Maximum yield occurs at target radius  $R_T=7.5$  mm for C and  $R_T=5$  mm for Hg targets with

$R_T = 3.5\sigma_{x,y}$  and  $R_T = 4\sigma_{x,y}$  conditions for the beam spot size, respectively. The baseline criterion  $R_T = 2.5\sigma_{x,y}$  reduces the yield by about 10% for the graphite target, but is more optimal from the energy deposition point of view. The use of a realistic 3-D magnetic field map in simulations results in the reduction of the  $\pi+\mu$ -yield in the decay channel by about 7% for C and by 10-14% for Hg targets, compared with a simple-minded  $B_z(r, z)$  model.



**FIGURE 3.**  $\pi+\mu$  yield from Hg and C targets *vs* proton energy.

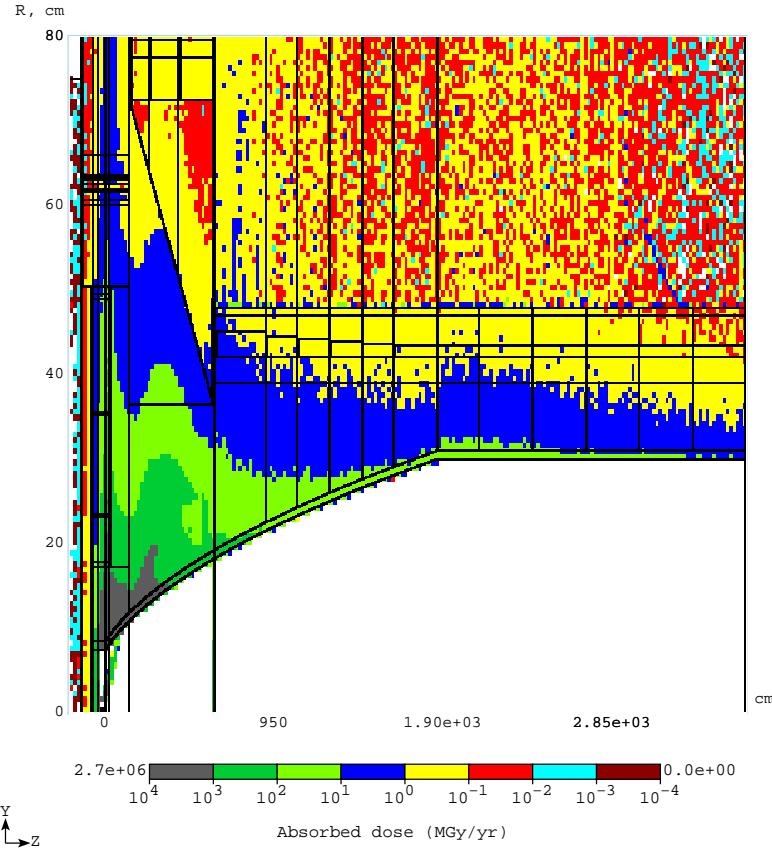
The optimized results for the yield per a proton on target, for Study-1 (16 GeV on C) are  $Y_{\pi^++\mu^+} = 0.18$  and  $Y_{\pi^-+\mu^-} = 0.15$  at  $z=9$  m, and for Study-2 (24 GeV on Hg, more realistic geometry and field) are  $Y_{\pi^++\mu^+} = 0.40$  and  $Y_{\pi^-+\mu^-} = 0.39$  at  $z=36$  m. There are substantial fluxes of accompanying particles in the system, which should be taken into account in designing beam instrumentation. In the aperture of the Study-2 channel, at the end of the matching region ( $z=18.6$  m), the numbers of particles per proton are 1.03 ( $\mu$ ), 1.15 ( $p + \pi^\pm$ ), 0.07 ( $e^\pm$ ), 0.02 ( $n$ ) and 0.46 ( $\gamma$ ).

### III RADIATION FIELDS

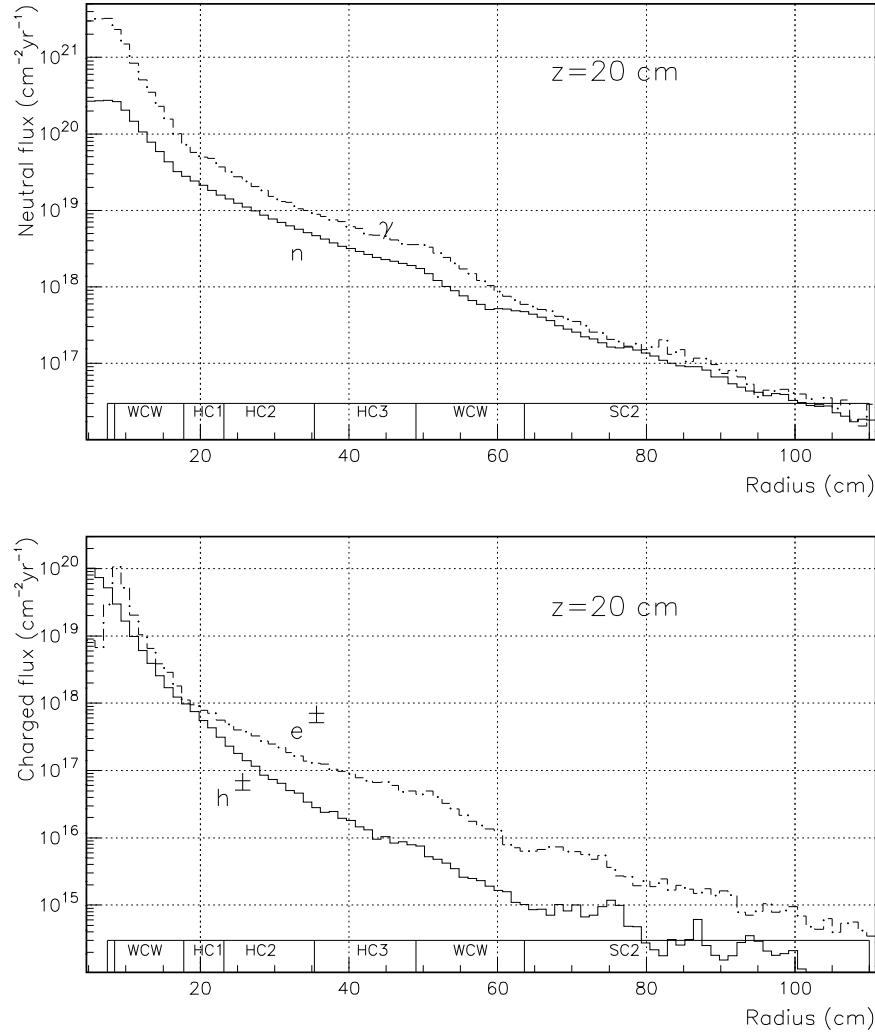
Hadronic and electromagnetic showers are induced in the target and capturing system, resulting in particle fluxes and accumulated dose in system components which can deteriorate their performance rapidly. The SC coils are to be adequately protected to provide their short and long term operation. A carefully designed coil shielding consists of two parts (Fig. 1): 1) at  $z < 6$  m it is made of tungsten-carbide balls (80% filling factor) cooled by circulating water (WCW), placed in front of the SC coils SC1-SC2 in the 20-T region and SC3-SC6 in the matching section, and

surrounds the resistive coils and the spent beam absorber; 2) at  $z>6$  m it is made of copper (70% filling factor) cooled by circulating water, and protects the potted SC7-SC12 coils in the matching section and further in the straight 1.25-T decay channel (SC13). The calculations show that it does an excellent job in protecting the SC coils against radiation.

The hottest regions in the system are the one at the downstream end of the target at the transition from the 20-T region to a matching section and at a primary beam dump at  $z\approx 4$  m (Figs. 1 and 4). The shielding reduces the peak power density to less than 0.3 mW/g (below the quench limit) in these two regions as well as in the entire system. The shielding provides also acceptable integrated levels of the absorbed dose (Fig. 4 and Table 1) and particle fluxes (Fig. 5) in the hottest spots, equalizing these to even lower levels in the rest of the system. As Table 1 shows, estimated lifetimes of the critical components are quite satisfactory. The component lifetimes are four times shorter for a 4 MW beam. In the Study-1 design [1,5], the annual hadron flux in a stationary graphite target is  $\sim 5 \times 10^{21} \text{ cm}^{-2}$  which corresponds to several month lifetime. The annual hadron flux ( $E>0.1$  MeV) and dose in the hottest spot of the inner resistive coil are  $1.2 \times 10^{20} \text{ cm}^{-2}$  and  $3 \times 10^{10} \text{ Gy}$ , respectively.

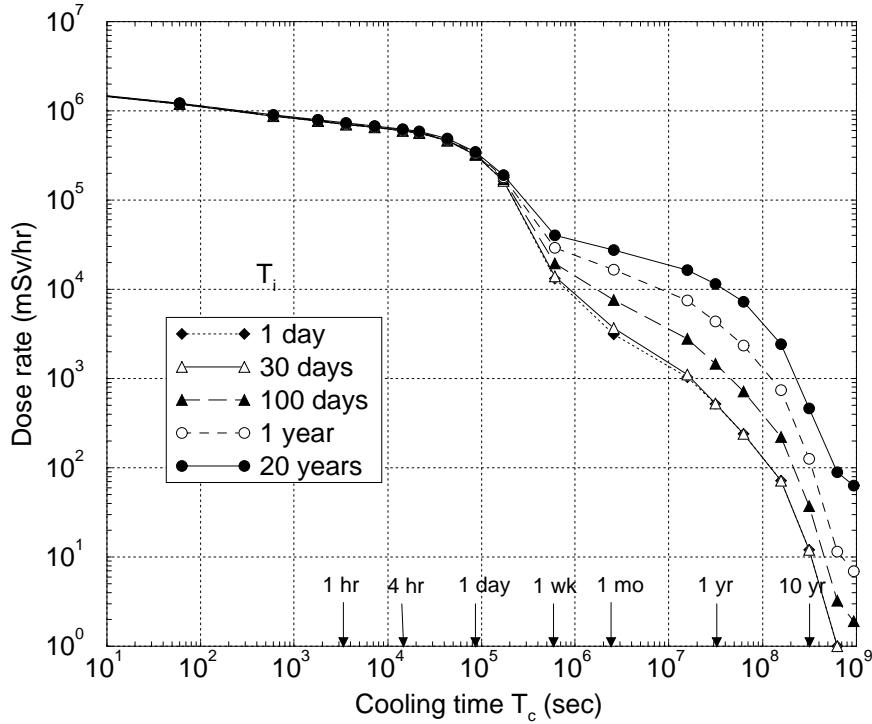


**FIGURE 4.** Absorbed radiation dose (MGy/yr) in target/capture system components.



**FIGURE 5.** Radial distribution of neutral (top) and charged (bottom) particle fluxes ( $\text{cm}^{-2}\text{yr}^{-1}$ ) in 20-T solenoid components at the downstream end of the target.

Heat loads to the main components of the Study-2 design, calculated for a 1 MW beam (0.979 MW to be exact), are shown in Table 2. About 12% of the beam power are deposited in mercury (jet plus pool), 50% in the coil shielding, 1% in resistive hollow conductor, and only about 0.1% in the high-field and potted SC coils. About 20% dissipate in other components and leak from the system. As Fig. 6 shows, the inner shielding becomes extremely radioactive, with residual dose rate up to 1 kSv/hr. This will require remote control and robotics for the inner parts of the system. It drops by two orders of magnitude after several weeks. The residual dose outside the cryostat is significantly lower, of the order of 100 mSv/hr. Radiation shielding needed is about 2 m of steel followed by concrete blocks to protect ground water followed by several meters of concrete and dirt to provide personnel protection.



**FIGURE 6.** Residual dose rate (mSv/hr) in the innermost tungsten-carbide shielding around the target *vs* cooling time for several irradiation times.

## REFERENCES

1. N. Holtkamp and D. Finley, eds., *A Feasibility Study of a Neutrino Source Based on a Muon Storage Ring*, Fermilab-Pub-00/108-E, 2000.
2. <http://www.cap.bnl.gov/mumu/studyii/FS2-report.html>.
3. N. V. Mokhov, “The MARS Code System User’s Guide”, Fermilab-FN-628, 1995; N. V. Mokhov and O. E. Krivosheev, “MARS Code Status”, Fermilab-Conf-00/181, 2000. <http://www-ap.fnal.gov/MARS/>.
4. N.V. Mokhov, “Particle Production for a Muon Storage Ring: I. Targetry and  $\pi/\mu$  Yield”, Neutrino Factory-2000, Monterey, CA, May 2000; Fermilab-Conf-00/208, 2000.
5. N.V. Mokhov, “Particle Production for a Muon Storage Ring: II. Radiation Loads”, *ibid*, Fermilab-Conf-00/209, 2000.

**TABLE 1.** Maximum radiation doses per  $2 \times 10^7$  s/yr and 1 MW lifetimes of some components of the target system.

Component	Dose/yr (MGy)	Limit (MGy)	Life (yr)
Inner shielding	$5 \times 10^4$	$10^6$	20
Hg containment	$2 \times 10^3$	$10^5$	50
Hollow conductor	$1 \times 10^3$	$10^5$	100
Superconducting coil	6	$10^2$	16

**TABLE 2.** Power dissipation in the main target/capture system components.

Component	Total heat load (kW)
Mercury	119.181
1-cm inner vessel	113.873
WCW shielding	489.118
Cu-water shielding	12.939
Hollow conductor	9.910
SC1-SC2	1.256
SC3-SC13	1.385

# Higgs Factory Front End

D. Neuffer, R. Fornow, and V. Balbekov

## 0. General Introduction

### 1. Introduction to muon cooling

- .1 cooling process and requirements
- .2 longitudinal cooling and emittance exchange
- .3 cooling considerations

### 2. Summary of status report (SR) front end scenario

- .1 phase rotation linac
- .2 cooling channel
- .3 bent solenoid emittance exchange example

### 3. Summary of neutrino factory (NF) front end scenario

- .1 precooling
- .2 sFOFO cooling system
- .3 simulation results
- .4 alternatives for NF cooling

### 4. Other Cooling components

- .1 Ring Cooler
- .2 Li lens
- .3 Final cooling and emittance exchange techniques
- .4 Other methods

### 5. Emittance exchange

- .1 standalone lattice sections
- .2 small dispersion superimposed on transverse cooling lattice
- .3 ring coolers
- .4 helical capture
- .5 bunch stacking
- .6 rf cavity modes

### 6. Simulation summary

### 7. Scenarios of front ends for $\mu^+ - \mu^-$ Colliders

- .1 SR based solution
- .2 Ring Cooler based scenarios
- .3 recycle maximum amount of NF front end
- .4 use adiabatic buncher

### 8. Status and plans

## Higgs Factory Front End

In this chapter we will consider the beam cooling system for a “Higgs Factory”  $\mu^+ - \mu^-$  collider. (HF). This could be either a completely new system, or could be based on a “neutrino factory” (NF) cooling system, and the modifications necessary to transform the front end of a NF into a system suitable for use as a HF are discussed. By “front end” we mean those systems that come after the pion collection solenoids and before the accelerators which take the HF beams to full collider energies. The front end has two parts. The first, which we call the “precooler”, includes phase rotation, any mini-cooling systems, and bunching into the rf-based cooling systems. The phase rotation is used to reduce the extremely large initial energy spread of the muon beam. The “mini-cooling” is a short energy-absorber section in the phase rotation system that cools transversely while reducing the momentum of the muon beam. Bunching is required to match the longitudinal characteristics of the beam into the rf-cavity-based cooling system. The second part of the front end is the cooling channel itself, which is used to make a large reduction in the normalized emittance of the muon beam, obtaining the small emittances needed for a high luminosity  $\mu^+ - \mu^-$  Collider.

### 1. Introduction to $\mu$ -Cooling

Table 1 shows some suggested parameters [1.1] of Higgs-energy  $\mu^+ - \mu^-$  colliders, operating in either high-luminosity or high-resolution (small  $\delta E$ ) modes. The reference energy of the collider is 50 GeV/beam, close to half the expected Higgs particle energy of 115 GeV. The collider energy can be rechosen at any value with a suitable physics goal; the collider cooling requirements would be similar. According to this table, the high-luminosity collider needs bunches of  $\sim 4 \times 10^{12}$   $\mu$ 's within transverse emittances of  $\epsilon_{N,\text{rms}} \sim 10^{-4}$  m and longitudinal emittance  $\epsilon_{L,\text{rms}} \approx 10^{-2}$  m. These emittances are substantially smaller than that of the muon bunches produced from pion decay, as described in the targetry section. The transverse and longitudinal emittances of these production beams are more like  $\epsilon_{N,\text{rms}} \sim 2 \times 10^{-2}$  m and  $\epsilon_{L,\text{rms}} \approx 1$  m. Thus we need to cool in each of the emittances by, roughly, a factor of 100 or, in 6-D emittance, by a factor of  $\sim 10^6$ , and the cooling must be completed before  $\mu$  decay. The only cooling method that can provide sufficiently fast cooling is ionization cooling, and the final beam parameters are within the expected capabilities of ionization cooling.

Table 1 also shows parameters for a “high-resolution” Higgs factory, where the energy spread is a factor of  $\sim 400 \times$  smaller. To obtain the small  $\delta E$  from the same 6-D emittance beam, the longitudinal emittance is decreased by an order of magnitude from the “high-luminosity” case, while each transverse emittance is increased by a factor of  $(10)^{1/2}$ , and the bunch length is increased by a factor of  $\sim 4$ . Emittance exchange techniques to obtain final collision parameters, and to switch between high-luminosity and high-resolution parameters are discussed below.

Table 1 Higgs factory requirements

Beam properties at capture	50 GeV high-resolution Higgs Factory	50 GeV high-luminosity Higgs Factory	
$\sigma_x$	10 cm	295 $\mu$ m	86 $\mu$ m
$\sigma_{x'}$	70	2.1	2.1 mr
$\sigma_z$	680	14.1	4.1 cm
$\sigma_p / p$	0.65	$3 \times 10^{-5}$	0.0012
$\epsilon_{xN}$	$\sim 20$	0.29	0.085 mm

$\epsilon_{zN}$	$\sim 1000$	2.02	24	mm
$\epsilon_{6N}$	$\sim 2 \times 10^5$	0.17	0.17	$\text{mm}^3$
f		15	15	Hz
$N_b$		2	2	Bunch/spill
N		$4 \times 10^{12}$	$4 \times 10^{12}$	Particles/bunch
L		$10^{31}$	$1.2 \times 10^{32}$	$\text{cm}^{-2}\text{s}^{-1}$

## 1.1 Cooling process and requirements

The cooling process that will be used is ionization cooling. In ionization cooling ( $\mu$ -cooling), particles pass through a material medium and lose energy (momentum) through ionization interactions, and this is followed by beam reacceleration in rf cavities.(see Figure 1) The losses are parallel to the particle motion, and therefore include transverse and longitudinal momentum losses; the reacceleration restores only longitudinal momentum. The loss of transverse momentum reduces particle emittances, cooling the beam. However, the random process of multiple scattering in the material medium increases the rms beam divergence, adding a heating term, which must be controlled in a complete cooling system. This cooling method is not very practical for protons, which would have frequent nuclear interactions, or electrons, which would have bremsstrahlung, but is practical for muons, and cooling rates compatible with muon lifetimes are possible.

The differential equation for rms transverse cooling is [1.2-6]:

$$\frac{d\epsilon_N}{ds} = -\frac{1}{\beta^2 E} \frac{dE}{ds} \epsilon_N + \frac{\beta \gamma \beta_\perp}{2} \frac{d\langle \theta_{\text{rms}}^2 \rangle}{ds} = -\frac{1}{\beta^2 E} \frac{dE}{ds} \epsilon_N + \frac{\beta_\perp E_s^2}{2\beta^3 m_\mu c^2 L_R E} \quad (1)$$

where the first term is the energy-loss cooling effect and the second is the multiple scattering heating term. Here  $\epsilon_N$  is the normalized emittance, E is the beam energy,  $\beta = v/c$  and  $\gamma$  are the usual kinematic factors,  $dE/ds$  is the energy loss rate,  $\theta_{\text{rms}}$  is the rms multiple scattering angle,  $L_R$  is the material radiation length,  $\beta_\perp$  is the betatron function, and  $E_s$  is the characteristic scattering energy ( $\sim 13.6$  MeV).[6] (The normalized emittance is related to the geometric emittance  $\epsilon_\perp$  by  $\epsilon_N = \epsilon_\perp / (\beta \gamma)$ , and the beam size is given by  $\sigma_x = (\epsilon_\perp \beta_\perp)^{1/2}$ .)

## 1.2 Longitudinal Cooling and Emittance Exchange

Cooling to collider intensities requires longitudinal cooling, which is difficult since ionization cooling does not directly provide longitudinal cooling. The equation for longitudinal cooling with energy loss is:

$$\frac{d\sigma_E^2}{ds} = -2 \frac{\partial \frac{dE}{ds}}{\partial E} \sigma_E^2 + \frac{d\langle \Delta E_{\text{rms}}^2 \rangle}{ds} \quad (2)$$

in which the first term is the cooling term and the second is the heating term caused by random fluctuations in the particle energy. Beam cooling can occur if the derivative  $\partial(dE/ds)/\partial E > 0$ . This energy loss can be estimated by the Bethe-Bloch equation[1.7]:

$$\frac{dE}{ds} = 4\pi N_A r_e^2 m_e c^2 \rho \frac{Z}{A} \left[ \frac{1}{\beta^2} \ln \left( \frac{2m_e c^2 \gamma^2 \beta^2}{I} \right) - 1 - \frac{\delta}{2\beta^2} \right] \quad (3)$$

where  $N_A$  is Avogadro's number,  $\rho$ ,  $A$  and  $Z$  are the density, atomic weight and number of the absorbing material,  $m_e$  and  $r_e$  are the mass and classical radius of the electron,  $(4\pi N_A r_e^2 m_e c^2 = 0.3071 \text{ MeV cm}^2/\text{gm})$ . The ionization constant  $I$  is approximately  $16 Z^{0.9} \text{ eV}$ , and  $\delta$  is the density effect factor, which is small for low-energy  $\mu$ 's. The energy loss as a function of  $p_\mu$  is shown in Fig. 2. The derivative is negative (or naturally heating) for  $E_\mu < \sim 0.3 \text{ GeV}$ , and is only slightly positive for higher energies. In the long-pathlength Gaussian-distribution limit, the second term in Eq. 2 is given approximately by[1.8]:

$$\frac{d\langle \Delta E_{\text{rms}}^2 \rangle}{ds} = 4\pi(r_e m_e c^2)^2 n_e \gamma^2 \left(1 - \frac{\beta^2}{2}\right) \approx 0.157 \rho \frac{Z}{A} \gamma^2 \left(1 - \frac{\beta^2}{2}\right) (\text{MeV})^2 \text{cm}^2 / \text{gm}, \quad (4)$$

where  $n_e$  is the electron density in the material ( $n_e = N_A \rho Z/A$ ). This expression increases rapidly with higher energy (larger  $\gamma$ ), opposing the cooling process. After adding this energy straggling, ionization cooling does not naturally provide adequate longitudinal cooling.

However, the longitudinal cooling term can be enhanced by placing the absorbers where transverse position depends upon energy (nonzero dispersion) and where the absorber density or thickness also depends upon energy, such as in a wedge absorber.[1.9, 1.10](see Fig. 3) In that case the cooling derivative can be rewritten as:

$$\frac{\partial \frac{dE}{ds}}{\partial E} \Rightarrow \frac{\partial \frac{dE}{ds}}{\partial E} \Bigg|_0 + \frac{dE}{ds} \frac{\eta \rho'}{\beta c p p_0} = g_L \frac{dp}{ds} \quad (5)$$

where  $\rho'/\rho_0$  is the relative change in density with respect to transverse position,  $\rho_0$  is the reference density associated with  $dE/ds$ , and  $\eta$  is the dispersion ( $\eta = d x / d(\Delta p/p)$ ). We have introduced the partition number  $g_L$  to describe the cooling rate related to the mean momentum loss, and the wedge configuration increases the longitudinal partition number by  $\eta \rho'/\rho_0$ . It also decreases the corresponding transverse partition number by the same amount:  $g_x \rightarrow (1 - \eta \rho'/\rho_0)$ , which decreases the transverse cooling. The sum of the cooling rates or partition numbers (over  $x$ ,  $y$ , and  $L$ ) remains constant; a similar invariant sum of cooling rates, with emittance exchange from radiation at nonzero dispersion, occurs in radiation damping of electrons. In ionization cooling, however, there is an energy dependence of this sum of partition numbers, due to the energy dependence of the natural energy loss. This sum of partition numbers is  $\sim 2$  at  $p_\mu > 0.3 \text{ GeV}/c$ . Fig. 4 shows this sum of partition numbers as a function of  $p_\mu$ ; small values of  $p_\mu$ , where the sum becomes less than  $\sim 1.5$  should be avoided in cooling.

Emittance exchange methods to obtain longitudinal cooling are discussed in more detail below. The intrinsic difficulties in obtaining longitudinal cooling indicate that it is very desirable to avoid longitudinal heating effects, if possible, since any heating must be later removed by added cooling. It is therefore desirable that transverse cooling sections avoid longitudinal heating ( $g_L < 0$ ), which can be avoided by cooling at  $p_\mu > 0.3 \text{ GeV}/c$ . This means cooling at higher energy than some of our earlier studies, which are designed at  $p_\mu \approx 0.2 \text{ GeV}/c$ , which would then require proportionately stronger focusing fields to achieve equivalent  $\beta_\perp$ .

Eq. 2 is the expression for energy spread cooling. The equation for longitudinal emittance cooling, similar to the transverse cooling equations, is:

$$\frac{d\epsilon_L}{ds} = -\frac{g_L}{p} \frac{dp}{ds} \epsilon_L + \frac{\beta_L}{2} \frac{d\langle \Delta E_{\text{rms}}^2 \rangle}{ds} \quad (6)$$

where  $\beta_L = \sigma_{ct}^2/\epsilon_L$  is the longitudinal focusing function, which depends on the rf bunching wavelength and voltage.

### 1.3 Cooling considerations

Some general considerations on the conditions for cooling, and the required absorbers and beam transports, can be developed from Eqs. 1 to 6. From Eq. 1 we find an equilibrium emittance from setting the derivative to zero:

$$\epsilon_{N,eq} = \frac{\beta_\perp E_s^2}{2g_x \beta m_\mu c^2 L_R \frac{dE}{ds}} \quad (7)$$

This represents the minimal obtainable emittance for a given material and focusing parameter  $\beta_\perp$ . From this expression, obtaining small emittance implies having small  $\beta_\perp$  (strong focussing), as well as large  $L_R dE/ds$  (small multiple scattering) at the absorber. Table 2 displays parameters of typical cooling materials; large  $L_R dE/ds$  implies light elements (H, Li, Be, ..) for the absorber material.

**Table 2: Material Properties for Ionization Cooling [1.7]**

Material	Z	A	dE/ds (min.)	$L_R$	$L_R dE/ds$	Density	$g_x \beta \epsilon_{N,eq} / \beta_\perp$
			MeV/cm	cm	MeV	gm/cm <sup>3</sup>	mm-mrad/cm
Hydrogen	H <sub>2</sub>	1	1.01	0.292	865	252.6	0.071
Lithium	Li	3	6.94	0.848	155	130.8	0.534
Lith. Hydride	LiH	3+1	7+1	1.34	102	137	0.9
Beryllium	Be	4	9.01	2.98	35.3	105.2	1.848
Carbon	C	6	12.01	4.032	18.8	75.8	2.265
Aluminum	Al	13	26.98	4.37	8.9	38.9	2.70
Copper	Cu	29	63.55	12.90	1.43	18.45	8.96
Tungsten	W	74	183.85	22.1	0.35	7.73	19.3
							1200

From consideration of minimum- $\beta_\perp$  focusing conditions (such as in a Li lens, see below), we expect to be able to obtain  $\beta_\perp \approx 0.01\text{m}$ , which means the transverse emittances can be cooled to  $\epsilon_{\perp,n} \approx 0.0001 \text{ m-rad}$  (in hydrogen or lithium). Similarly a cooling rf bucket at 200 MHz can maintain the beam within a longitudinal emittance of  $\sim 0.01\text{m}$ , and smaller emittances could be obtained with higher frequency rf systems. Thus, the collider emittance goals are within conceptual reach of ionization cooling. However a complete cooling scenario taking the beam from production emittances to cooling emittances must be developed, and some possible approaches will be discussed below.

## 2. The Status Report Front End Scenario

A complete cooling scenario to collider requirements has not yet been completely worked out. However, a general approach and many of the individual components have been described in a status report [1.1]. We summarize here some important features of the front end used in that design.

### 2.1 Phase rotation linac

The pions, and the muons into which they decay, are produced with a momentum distribution with an rms spread ( $\delta p/p$ ) of approximately 100% about a maximum around 200 MeV/c. It would be difficult to handle such a wide energy spread in any subsequent system. A linac is thus introduced along the decay channel, with frequencies and phases chosen to decelerate the fast particles and accelerate the slow ones, i.e., to phase rotate the muon bunch. Several studies have been made of the design of this system, using differing ranges of rf frequency, delivering different final muon momenta, and differing final bunch lengths. In all

cases, muon capture efficiencies close to 0.3 muons per proton can be obtained. Until the ionization cooling section is fully designed, an optimal capture optics is not yet defined, and will depend on future rf cavity and solenoidal focusing development. The SR presented two capture configurations: a low-energy and a high-energy example.

The low energy example captures muons at a mean kinetic energy of 130 MeV. Four linacs are used with frequencies varying from 30 to 60 MHz and gradients up to 5 MV/m.[1.1] These gradients are relatively high for continuous low frequency systems, but far below the maximum surface fields that have been achieved in short pulses. The significant challenge will be the development of affordable and sufficiently high-power low-frequency rf sources. The example has been simulated by the Monte Carlo programs MCM and ARC, starting from pion production by 24 GeV protons on a copper target of 1 cm radius at an angle of 150 mrad.[1.1] A uniform solenoidal field of 1.25 T was used in the phase rotation section, and the rf was approximated by a series of kicks. A final bunch selection was defined by a kinetic energy window of  $130 \pm 70$  MeV and a bunch length ( $c\tau$ ) 8 m long. Within these cuts, the rms energy spread of the selected muons is 16.5%, the rms  $c\tau$  is 1.7 m, and there are 0.39 muons per incident proton. A tighter acceptance cut at an energy of  $130 \pm 35$  MeV and bunch  $c\tau$  length of 6 m gave an rms energy spread of 11.7%, rms  $c\tau$  of 1.3 m, and contained 0.31 muons per incident proton.

In the high-energy example the captured muons have a mean kinetic energy close to 320 MeV. It is based on a Monte Carlo study using MCM and MARS that uses pions created by 16 GeV protons on a 36 cm long, 1 cm radius coaxial gallium target. The phase rotation system consisted of an 80 m long, 5 T solenoidal decay channel with cavities of frequency in the 30-90 MHz range and acceleration gradients of 4—18 MV/m. A total of 0.33 muons per proton fall within the cut: 6 m  $\times$  300 MeV. The rms bunch length inside the cut is 148 cm and rms energy spread is 62 MeV. The normalized six dimensional (6-D) emittance is 217 cm<sup>3</sup> and the transverse part is 1.86 cm.

Protons on the target produce pions of both signs, and a solenoid will capture both, but the subsequent rf systems will have opposite effects on each sign. The proposed baseline approach uses two separate proton bunches to create separate positive and negative pion bunches and accepts the loss of the wrong-sign pions/muons during phase rotation. If the pions can be charge separated with limited loss before the phase rotation cavities are reached, both signs can be captured from each primary bunch, resulting in more  $\mu$ 's and therefore higher luminosity.

## 2.2 Cooling channel

Following the rf phase rotation the long bunches are injected into a cooling system, designed as a sequence of cooling stages. Each stage consists of a succession of the following components:

- Transverse cooling sections using energy-absorber materials in a strong focusing (low- $\beta_{\perp}$ ) environment, alternating with linear accelerators
- Emittance exchange in lattices that generate dispersion, with absorbing wedges to reduce momentum spread
- Matching sections to optimize the transmission and cooling parameters of the following section

For the SR cooling scenario, the  $\mu$ -beams have a central kinetic energy close to 100 MeV, which was chosen as an apparent optimal cooling energy. At higher energies, weaker focusing raises the heating term from Coulomb scattering, and more acceleration is required for a given amount of cooling. At lower energies, the beam divergence become large, and the rise of  $dE/dx$  with falling energy causes a greater increase in energy spread. There is an advantage, initially, in using a

somewhat higher energy to reduce the beam dimensions and bucket length; and at the end, the energy can be dropped to attain the lowest transverse emittances at the expense of longitudinal heating. Each of the design transverse cooling stages lower the 6-D emittance by a factor of about 2. Since the required total 6-D cooling is  $\sim 10^6$ , about 20 such stages are required.

We have performed initial analytical calculations for complete cooling systems for the Higgs factory and for the cooling systems for a high energy collider. These calculations are based on theoretical models (rms cooling equations) of the expected cooling performance. They give an indication of the system dimensions, magnet strengths, rf frequencies and gradients, and beam parameters that will be required in a cooling system. The SR calculations of a collider cooling scenario indicate that the required cooling for a Higgs factory could be achieved in 25 stages, while the high energy collider would require an additional 3 stages. Emittances and energies as a function of stage are shown in Fig. 5. The sequence can be considered to consist of 3 parts.

For the first 12 stages the primary effort is to cool in the longitudinal direction in order to reduce the bunch lengths and allow higher frequency rf to be employed. Some transverse cooling is also needed to reduce the transverse dimensions of the beam and allow it to fit through the smaller irises in higher frequency cavities. In this example, for the first stage, an energy of 300 MeV was used. Emittance exchanges are used at the beginning of the system to reduce the longitudinal emittance; however, these exchanges also increase the size of the beam, and complete design of the necessary transport and low-frequency rf remains a major challenge. In later stages the kinetic energy is reduced closer to 100 MeV. Solenoid focusing was assumed in all of these stages, with an initial field of the order of 1 T rising to about 3 T at the end.

In the second part (in this example, stages 13 - 25) the 6-D emittance is reduced as far as possible, with simultaneous transverse and longitudinal cooling. For the case of a low momentum spread Higgs collider, the required beam parameters are achieved at stage 25 of the SR scenario, and the third part (last three stages) is not required. In the SR scenario, an 80 MeV central energy was used for stages 13-25. Solenoid focusing was used in all but the last two of these, where lithium lenses were assumed.

For the higher luminosity and higher energy colliders, the third section is needed. Further reduction in transverse emittance is required, but this can be obtained without reduction of the 6-D phase space, by allowing the longitudinal phase space to grow. This exchange of emittances is, in this example, achieved by reducing the energy to near 10 MeV in two long lithium lens cooling stages. The same effect could probably be achieved at similar energy, by using a hydrogen absorber with solenoid focusing. It might also be possible by using low-energy wedges.

The total length of the system would be of the order of 600 m, and the total acceleration required would be approximately 6 GeV. The fraction of muons transmitted through the cooling system is estimated to be  $\sim 60\%$ . It must be emphasized that this sequence was initially derived without detailed simulation of the individual stages. It serves however to guide the choice of stages to study in detail.

Three transverse cooling stages from the SR scenario were designed and simulated in detail. The first uses 1.25 T solenoids to cool the very large emittance beam coming from the phase rotation channel. The muon beam at the end of the decay channel is very intense, with  $\sim 7.5 \cdot 10^{12}$  muons/bunch, but with a large normalized transverse emittance  $\epsilon_{xN} \sim 15 \cdot 10^3$  mm-mrad and a large normalized longitudinal emittance  $\epsilon_{zN} \sim 612$  mm. The second example would lie toward the end of a full cooling sequence and uses 15 T solenoids. The third example, using 31 T solenoids, meets the requirements for the high resolution Higgs factory and could be the final cooling stage for this machine.

The baseline solution for emittance exchange involved the use of bent solenoids to generate dispersion and wedges of hydrogen or LiH to reduce the energy spread. A simulated example was given for exchange that would be needed after the 15 T transverse cooling case. (see Fig. 6)

Figure 7 shows a 2m cooling cell from one of the coooling stages. A schematic cross-section of the cell with absorbers, rf cavities and solenoid coils is shown, along with the magnetic fields and resulting focusing betatron ( $\beta_T$ ) function. Figure 8 shows simulation results from cooling in a complete stage containing  $\sim 13$  such cells.

A lithium lens solution may prove more economical for the final stages, and might allow even lower emittances to be obtained. In this case, the lithium lens serves simultaneously to maintain the low  $\beta_{\perp}$  and provide  $dE/dx$  for cooling. Similar lenses, with surface fields of 10 T, were developed at Novosibirsk and have been used, at low repetition rates, as focusing elements at FNAL and CERN. Lenses for the cooling application, which would operate at 15 Hz, would need to employ flowing liquid lithium to provide adequate thermal cooling. Higher surface fields would also be desirable.

### 2.3 Bent solenoid emittance exchange example

In addition to the transverse cooling section, each cooling stage includes an emittance exchange section for reduction of longitudinal emittance. As an example of such a stage, we have studied a system that exchanges longitudinal and transverse emittance using dispersion in a large acceptance channel, with low-Z wedge absorbers in the region of dispersion. In a bent solenoid, there is a drift perpendicular to the bend plane of the center of the Larmor circular orbit, which is proportional to the particle's momentum. In our example we have added a uniform dipole field over the bend to cancel this drift for particles with the reference momentum. Particles with momenta differing from the reference momentum then spread out spatially, giving the required dispersion (0.4 m). The momentum spread is reduced by introducing liquid hydrogen wedges. (The hydrogen wedges would be contained by thin beryllium or aluminum foils, but these were not included in this simulation.)

After one bend and one set of wedges, the beam is asymmetric in cross section, since the emittance exchange has occurred in a single plane. Symmetry is restored by a following bend and wedge system rotated by 90 degrees with respect to the first. Figure 6 shows a representation of the two bends and wedges. The total solenoid length was 8.5 m. The beam tube outside diameter is 20 cm, and the minimum bend radius is 34 cm. Note that this is only a preliminary design. No rf was included in this configuration and the growth of the bunch length passing through the system was ignored.

The solenoid bend curvature is exactly that given by the trajectory of a reference particle (equal in momentum to the average momenta) in the given transverse fields. The actual shape of the bend turns out to be very important. Discontinuities in the bend radius can excite perturbations, which increase the transverse emittance.

The simulations were performed using the program ICOOL. The maximum beam radius is 10 cm. Transmission was 100%. The fractional momentum spread decreases from an initial value of  $\sim 5\%$ , to a final value of  $\sim 2.2\%$ . At the same time, since this is an emittance exchange, the transverse beam area grows. The area increases not only in the regions of bends, but also in the regions of wedges. This is probably due to failures in betatron matching. The dispersion is clearly observed after both bends. It is removed, with a corresponding decrease in momentum spread, after both set of wedges. Figure 6 shows a scatterplot of the square of the particle radii vs. their longitudinal momenta, (a) at the start, and (b) at the end of the emittance exchange section. The decrease in momentum spread and rise in beam area are clearly evident. Although this example demonstrates a factor of  $\sim 3$  reduction in the longitudinal momentum spread, there is a 37% increase in the 5-D phase space. These simulations must be extended to include rf, so that the full 6-D emittance behavior can be studied.

### 3. Neutrino Factory Front End

A  $\mu$ -storage ring neutrino factory has somewhat different requirements on the muon beams from those of a  $\mu^+ \mu^-$  collider. The  $\nu$ -source event rate depends primarily on the number of stored muons and not on the quality of the  $\mu$ -beam; therefore the beam needs only to be cooled sufficiently to be within the acceptance of the accelerator and storage ring, and not to a minimal emittance for high luminosity. Also the  $\mu$ -beam need not be confined within single bunches, but can be distributed in a string of bunches.

The MC collaboration has produced two detailed Neutrino Factory design studies[3,1,2]. Both studies have a similar muon collection and cooling system. A schematic layout of the study 1 scenario is shown in Fig. 9A. Following the target there is a 50 m long drift, a 100 m long induction linac for phase rotation, a mini-cooling stage, a 17 m long buncher and a 140 m long cooling section. The study 2 scenario layout is shown in Fig. 9B. It has a 18 m drift following the target, a 108 m induction linac, a mini-cooling stage, an additional 200 m of induction linac and drift to complete the phase rotation, a 55 m buncher and a 108 m cooling section, for a total length of 540 m, about 200m longer than study 1. Most of the following description will concentrate on the more recent study 2 case, which is an improvement of study 1.

#### 3.1 $\nu$ -Factory precooler

In the neutrino factory studies a proton bunch on the target produces a  $\pi$  beam, which is then allowed to drift while the  $\pi$ 's decay into  $\mu$ 's and the beam develops a position-energy correlation with the lower-energy  $\mu$ 's trailing behind the higher energy  $\mu$ 's. The energy spread of the muons is very large, much larger in fact than the acceptance of the following cooling stage and accelerators. Therefore this drift is followed by an induction linac system that decelerates high-energy  $\mu$ 's and accelerates low energy  $\mu$ 's, reducing the energy spread. The voltage pulse across the gaps in the induction linac cells can be tailored to match the time-energy correlation of the incoming beam bunch. In study 2 the induction linac is broken up into two sections in order to reduce the distortion in the resulting longitudinal phase space of the muons. The fractional energy spread of the beam after the induction linacs is reduced to 3.7%. However, the rms bunch width grows to 27 m ( $\sim$ 100 m full width). The induction linac system includes a 3.5 m long liquid  $H_2$  absorber. This "mini-cooler" stage provides an  $\sim$ 20% reduction in transverse normalized emittance to  $\epsilon_{t,rms} \cong 0.012$  m.

After the induction linacs, the muons are distributed continuously over a bunch length of around 100 m. It is then necessary to form the muons into a train of bunches for cooling and subsequent acceleration, as well as to match the beam transversely into the focusing lattice used for cooling. Thus the beam is transported into a 201 MHz rf buncher section, which forms the beam into about 70 bunches. The transverse and longitudinal functions of this section are performed sequentially for design simplicity. First an 11 m long magnetic lattice section is used to match the beam from the approximately uniform solenoidal field used in the induction linacs to the so-called "super-FOFO", or sFOFO, focussing lattice used in the remainder of the front end. This is followed by the 55 m long rf buncher, which consists of rf cavity sections interspersed with drift regions.

The buncher magnetic lattice is identical to that used in the first cooling section. It contains rf cavities in selected lattice cells and no absorbers. The main rf frequency is chosen to be 201.25 MHz in the front end, so that the beam would fit radially inside the cavity aperture and because power sources and other technical components are available at this frequency. The 201.25 MHz cavities are placed at the high beta regions in this lattice, just as in the cooling section. Maximum

bunching efficiency was obtained by breaking the region into three rf stages separated by drift regions. Second harmonic cavities (402.5 MHz) are added at the entrance and exit of the first and second stages to linearize the shape of the rf pulse. The buncher encompasses 20 lattice cells, each 2.75 m long. By the end of the buncher, most, but not all, particles are within the 201.25 MHz buckets. About 25% are outside the bucket and are lost relatively rapidly, and another 25% are lost more slowly as the longitudinal emittance rises from straggling and the negative slope of the energy loss with energy.

### 3.2 sFOFO Cooling System

In the Neutrino Factory the rms transverse emittance of the muon beam emerging from the induction linac must be reduced to  $\sim 2$  mm-rad (normalized) in order to fit into the downstream accelerators, and be contained in the storage rings. Ionization cooling is currently our only feasible option. The cooling channel described below is based on extensive theoretical studies and computer simulations.

Solenoidal fields are used for focusing; however, energy-loss cooling within a constant (or same-sign) field leads to an increasing beam angular momentum through the cooling channel. The solenoidal field must flip sign, while maintaining good focusing throughout the beam transport and low  $\beta_{\perp}$  at the absorbers. One of the simplest solutions is to vary the field sinusoidally; this is the “FOFO” lattice. The cooling system of study 2 is composed of “sFOFO” or “super-FOFO” lattice cells. (see Fig. 10) Each of these cells includes an absorber for energy-loss cooling, an rf cavity of beam reacceleration and solenoids for transverse focusing, with the focusing designed to minimize beam size in the absorbers. The sFOFO lattice uses alternating solenoids like the FOFO, but is a bit more complicated. As in the FOFO case, the longitudinal B-field vanishes at the  $\beta_{\perp,\min}$  position, located at the center of the absorber. This is accomplished by placing two short, strong “focussing” coils about the absorber, running in opposite polarities. The field is decreased and flattened outside the absorber region, due to a “coupling” coil located around the linac.

The study 2 cooling channel operates at a nominal momentum of 200 MeV/c. There are six sections with steadily decreasing  $\beta_{\perp,\min}$ . In the first three the lattice half period length is 2.75 m, and in the last three sections this half period length is 1.65 m. The matching sections between these sections also consist of cooling cells, which differ from the regular cooling sections only by the currents (except for matching between different cell lengths, where the length is also changed).

Each lattice half-period includes a multicell linac, and to increase the useful gradient of the accelerating cavities, the cell irises are covered with a foil or grid. The baseline design calls for thin, pre-stressed beryllium foils with thicknesses that increase with radius. An accelerating gradient of  $E = 16$  MV/m is required in each linac.

The absorber material is liquid hydrogen ( $LH_2$ ). The length of these absorbers is 35 cm for the 2.75 m lattices and 21 cm for the 1.65 m lattices, respectively. The  $LH_2$  vessels must also be equipped with thin aluminum windows. Their thickness is 360  $\mu m$  (220), with a radius of 18 (11) cm, for the 2.75 m and 1.65 m lattices, respectively. The muons therefore lose  $\sim 12$  MeV per lattice cell for the 2.75 m lattices and  $\sim 7$  MeV for the 1.65 m lattices.

The complete study2 cooling system contains 16 2.75m cells, 36 1.65m cells, and a 4.4 m matching section between them for a total length of  $\sim 108$  m. Complete descriptions of the system and detailed simulations of its cooling performance are presented in study 2 [FS2] and summarized below.

### 3.3 Simulation results

An important accomplishment in the collaboration has been the development of the simulation codes ICOOL [3.3] and DPGeant [3.4], which include the full complexity of the absorber + rf + solenoid system, including all materials and their properties, magnets defined in terms of coils, currents, and positions (rather than actual focusing fields), rf defined with complete cavity fields, etc. The goal is to establish cooling systems which function when described in full complexity. These simulations confirm that the study 1 and 2 cooling systems will perform as planned. In this section we describe these simulation results.

The  $\beta_{\perp,\min}$  function, calculated at the absorber centers using the beam second-order moments calculated in Geant4 simulations, is shown in Fig. 11. This function is reduced with each new section of the cooling lattice. The transverse and longitudinal emittances as calculated through the cooling system are shown in Fig. 12. Emittances are computed using diagonalized covariance matrices. The emittance values are corrected for correlations between the variables, including the strong momentum-transverse amplitude correlation. At the end of the cooling channel a transverse emittance of 2.2 mm rad is reached. The longitudinal emittance shows an initial rise and fall as particles not within the rf bucket increase in amplitude and are later lost, and then an approach to an asymptotic value set by the bucket size. The longitudinal emittance should rise due to straggling and the negative slope of energy loss with energy, but, since the rf bucket is already full, we see a steady loss of particles instead of an emittance growth.

Despite the overall particle loss, the numbers of particles within the accelerator acceptance increases, as seen in Fig. 13. The gain in muons within the accelerator acceptance of 150 mm due to cooling is  $\sim 3.1 \times$  (or  $5 \times$  if the study 1 acceptances were used). If the particle loss from longitudinal emittance growth could be eliminated, as should be the case if emittance exchange were used, then these gains could double.

**Table 3 Beam characteristics summary**

Location (end of)	$\sigma_x$	$\sigma_{x'}$	$\sigma_p$	$\sigma_t$ (per bunch)	$\langle p \rangle$
	Cm	mrad	MeV/c	ns	MeV/c
Induction linac	8.6	95	113		260
Matching section	5.8	114	113		260
Buncher	5.3	107	111	0.84	256
2.75 m cooler	3.0	91	70	0.55	226
1.65 m cooler	1.8	102	30	0.51	207

The rms beam characteristics in the buncher and cooler sections are summarized in Table 3. The beam is symmetric in this lattice, so the y properties are similar to those in x. We see that the size steadily decreases as we proceed down the channel. The angular divergence is kept constant for maximum cooling efficiency. The momentum spread of the entire beam is still large after the induction linac, but this includes very low and high energy muons that do not get transmitted through the subsequent sFOFO lattice. The decrease in energy spread is due to particle losses, since there is no longitudinal cooling or emittance exchange in this lattice. These losses could be controlled by adding some longitudinal cooling to the channel.

### 3.4 Alternatives for v-Factory cooling

We have presented in detail only one example of a cooling scenario for preparing the  $\mu$ -beams of a v-factory, the baseline cooling scenario for feasibility study 2. Other cooling scenarios could be used and future studies will explore alternative configurations, either by

optimizing the present proposal or developing a substantially different but superior system. However any cooling scenario would also require: absorbers for energy-loss, acceleration for longitudinal energy recovery, and a transport lattice with strong focusing of the beam into the absorbers.

The liquid hydrogen absorbers were chosen because hydrogen has the least multiple scattering; however other low-Z material (LiH, Li, Be, ...) could also be used and would avoid the mechanical difficulties of handling liquid hydrogen, at the cost of more scattering. It is likely that such denser materials may be necessary for emittance exchange wedges.

The sFOFO focusing system was used here, but other lattices could be used. An attractive alternative is the “double-flip” scenario [3.5], which has long constant or same-sign field sections with only two changes in sign. This is a simpler lattice, but it requires more field volume for the same focusing effect as the sFOFO. Lattices that incorporate energy cooling could also be preferable (see below).

The ~200 MHz rf system was based on the perception that 200 MHz rf would be available and affordable. A low-frequency system (40-80 MHz) for capture and cooling has been proposed at CERN [3.6] and has some preferable properties. It would develop fewer  $\mu^+$ - $\mu^-$  bunches per primary p-bunch and would be more adaptable to future  $\mu^+$ - $\mu^-$  collider beams.

## 4. Other Cooling Components

The Status Report and the v-factory studies relied on ionization cooling in a single-pass cooling channel with absorbers (usually  $\text{LH}_2$ ) periodically placed within an rf linac structure with strong solenoidal focussing. Various solenoidal focusing lattices have been studied, including FOFO, sFOFO, “single-flip”, “double-flip”, etc., and analytical methods for describing ionization cooling in solenoidal focussing systems have been developed, and all of these have the same general structure. The cooling systems require a very long (nearly) linear structure (~100 m for the v-factory and ~600 m for a collider) of rather expensive components. Also it is limited in performance by the focusing limitations of solenoid focusing, and integration with longitudinal cooling is not yet developed. In this section we discuss some variations in cooling which may avoid some of these difficulties. The following section will address the longitudinal cooling issues.

### Ring Coolers

It appears inefficient to use a single-pass linac-based structure for cooling; it would be more efficient if the beam could pass several turns through the same cooling structure, obtaining much more cooling from a given structure than a single pass device. A  $\mu^+$ - $\mu^-$  collider system may require recirculating cooling systems to be affordable. Balbekov has presented several explicit ring cooler designs that are able to obtain cooling in 6-D phase space by large factors in ~10 turns of circulation. Two of these ring designs are shown in Figs. 14 and 15. This concept has the important advantage in that the cooling hardware is reused for several turns of cooling. The ring cooler designs also have cooling systems which cooled longitudinally as well as transversely, obtain cooling by factors of 3—10 in each dimension. Also since the bunch lengths are naturally decreased in the Ring Cooler, matching to a higher-frequency succeeding cooling device is relatively easy. These designs are described in more detail in the following section on emittance exchange and longitudinal cooling.

The major unsolved problem in the Ring Cooler concept is how to inject and extract  $\mu$  bunches, without beam loss or emittance dilutions. The cooling lattices are packed with focusing, acceleration and energy-loss elements and there is no free space for such elements; one would like injection/extraction kickers that overlap ring recirculating magnets. The solution could be a large-aperture fast kicker, similar to that used at CERN for the Antiproton accumulator. A rise time of  $\sim 50$  ns or less would be required. The problem is in inserting the kicker hardware into the ring without degrading the ring cooler performance.

### Li lens cooling

A particularly attractive configuration for  $\mu$ -cooling is obtained by passing the beam through a conducting light-metal rod (such as a Li lens shown in Fig. 16), which acts simultaneously as a focusing element and as an energy-loss absorber. [4.1, 4.2] A high current passing through the conductor provides an azimuthal magnetic field given by[ ]:

$$B_\theta = \frac{\mu_0 I r}{2\pi R_c^2} \quad (8)$$

where  $R_c$  is the rod radius and  $I$  is the total current in the rod. This azimuthal magnetic field combines with the longitudinal velocity to obtain a radial focusing force. The matched focusing  $\beta_\perp$  for a Li lens is:

$$\beta_\perp = \sqrt{\frac{B\rho}{B'}} = \sqrt{\frac{p_\mu}{eB'}} \quad (9)$$

where  $B' = dB/dr = \mu_0 I / (2\pi R_c^2)$ . Li lenses can provide quite strong focusing and are used for short, strong-focusing collection lenses. A Li lens with  $B = 20$  T at a radius  $R_c = 2$  mm is possible and this would give a matched  $\beta_\perp$  of 1 cm for  $p_\mu = 300$  MeV/c muons.

Some parameters of Li lenses considered for cooling are tabulated in Table 4. In this table the Li lens lengths have been standardized at 1 m. A sequence of lenses of increasing strength is tabulated as examples of possible parameters.

**Table 4: Li lens parameters**

B(T)	B' (T/m)	Radius (cm)	Power/m (10Hz)	I (MA)	$\tau(\delta=0.7r)$	$\beta^*$ at $P=0.3$ GeV/c
10	1000	1	0.68 MW	0.50	1 ms	3.16 cm
15	3000	0.5	0.383	0.375	250 $\mu$ s	1.83
20	8000	0.25	0.1717	0.25	63 $\mu$ s	1.12
20	16000	0.125	0.041	0.125	15 $\mu$ s	0.79

Li lenses can be used to extend the cooling to small emittances. As discussed above, lenses which can focus to  $\beta^* = 1$  cm or less can cool  $\mu$ -beams to  $\epsilon_T \approx 10^{-4}$  or less. Fig. 17 shows simulation results of cooling through a sequence of 12 lenses (including 2 emittance exchange segments), with  $\epsilon_{T,N}$  reduced from  $10^{-2}$  to  $0.86 \times 10^{-4}$  m-rad.[4.3]

The long lenses needed to obtain large energy losses ( $\sim 1$  m of Li to obtain  $\sim 100$  MeV of energy loss), and the high repetition rates of collider scenarios imply large power requirements and large power deposition associated with higher frequency operation (5—15 Hz) would melt Li. Liquid Li lenses are also desirable because of the brittleness of solid Li lenses. A replacement liquid lens is being built for the Fermilab antiproton source. A longer, higher gradient liquid Li lens testing the limits of that technology for  $\mu$  cooling was also planned in the  $\mu^+ \mu^-$  Collider R&D program; that R&D has been postponed, however.

Practical difficulties exist in matching the large emittance, large  $\Delta p$   $\mu$ -beams into and out of Li lenses, as well as in matching the beams into reaccelerating rf buckets with minimal dilution and losses. While initial attempts were unsuccessful, some solutions to this problem have been demonstrated. A simulation by Spentzouris and Neuffer [4.4] considered a 2-lens system with 800 MHz rf, which reduced transverse emittances by a factor of 2 with small longitudinal heating and mismatch effects. V. Balbekov [4.5] simulated a sequence of 5 lenses, with intervening 800 MHz rf and 2 dipole/wedge coolers, which cooled transversely from  $\epsilon_{t,rms} = 0.001$  to  $0.0002$  m-rad and with longitudinal emittance increasing from 2 to 3 mm-rad. Results are shown in Fig. 18. These final parameters are close to Higgs Collider goals. These used 800 MHz rf; the greater acceptances of 200 MHz rf systems would make these solutions even easier. Optics + cooling scenario optimization remain a research topic[4.6], and the practical limits on Li lens field strengths, lengths and repetition rates are not established.

### Final Cooling and Emittance Exchange Techniques

In the final cooling sections, cooling system parameters can be extended to extreme values to obtain collider beam conditions. In these final sections, the transverse emittances are reduced to minimal values, while allowing nontrivial longitudinal emittance growth, or the longitudinal emittances are minimized with transverse emittance increase. Some techniques for obtaining these final “emittance exchanges” are described in this section.[see 1.1, 4.7, 4.8]

In order to obtain minimal-transverse-emittance beams in the final cooling stages, the beams are run at very low energies, so that  $\beta^*$  can be minimized, and an emittance exchange between transverse and longitudinal cooling is generated. Two methods that can achieve this have been suggested:

1. an “anti-wedge” absorber which increases energy spread while reducing transverse emittances. In an “anti-wedge” configuration, the  $\mu$ -beam passes through a wedge absorber at non-zero dispersion, but the wedge is oriented so that the low-energy portion of the beam passes through more material than the high-energy portion. The net effect is an increase in energy spread and longitudinal emittance, with a decrease in the dispersion-plane transverse emittance. A low-energy beam permits large emittance exchange in short wedges, with relatively small transverse heating. In a simulated example, a beam with  $p_\mu = 77$  MeV/c was passed through a 0.8 cm,  $\tan\theta = 1$  wedge at dispersion  $\eta = -0.105$  m obtaining  $\epsilon_{x,N}$  cooling from 0.0061 to 0.0039 mm, with  $\epsilon_{y,N}$  unchanged and  $\delta p$  increased from 1 to 1.76 MeV/c
2. cooling at low energies in a Li lens. At low energies (low momenta) the Li lens can focus to very small  $\beta^*$  and relatively short lengths of absorber can cool the beam to small transverse emittances. For example, a 2000 T/m Li lens with a 75 MeV/c beam produces a  $\beta^*$  of 0.35 cm. However at low energies, longitudinal energy loss is strongly antidamping, and the 6-D cooling is at best stationary. The net effect is a strong transverse cooling situation with large longitudinal antidamping, which is a large emittance exchange. In a simulated example, a 100 MeV/c beam was tracked through a 14 cm,  $B' = 10000$  T/m lens, and cooled  $\epsilon_{T,N}$  from 0.01 cm to 0.0077 cm, while  $\delta p$  increased from 2 to 4.36 MeV/c, and  $p_\mu$  was reduced to 68 MeV/c. On the order of 4 such lenses with interlaced reacceleration rf can cool  $\epsilon_{T,N}$  from  $2 \times 10^{-2}$  cm-rad to  $0.5 \times 10^{-2}$  with longitudinal emittance increasing by a factor of  $\sim 10$ .

In both of the simulated examples 6-D emittance increased; the longitudinal heating effects were greater than the transverse cooling effects.

## Other Cooling Methods

To date, only ionization cooling using magnetic and/or Li lenses for focusing, and reaccelerating rf with low-Z absorbers, with the cooling of medium energy muons (100—400 MeV kinetic energies), is believed to be within reach of presently available technology and to provide cooling fast enough to avoid  $\mu$  substantial decay. Only ionization cooling is included in the baseline Higgs collider scenarios. Other cooling methods may be considered and could eventually become practical. These methods include:

1) Low energy cooling methods: Here the general technique is to stop (or nearly stop) the muons within a material, which gives very cold  $\mu$ 's. The difficulty then is in separating them from the material into a compressed, accelerable bunch before they decay.

For positive muons, the bunches can be stopped in a hot tungsten foil, where they combine with atomic electrons to form muonium ( $\mu^+e^-$  atoms). The muonium atoms evaporate from the foil, where intense laser light pulses resonantly excite and ionize the atoms, and the resulting cloud of muons can then be electromagnetically trapped and accelerated. The process has been implemented at the level of a few per second by Nagamine et al., and intensity upgrades to  $\sim 10^{10}/s$  are being considered.[4.9]

For negative muons, a sequence of tungsten foils can be used to obtain a very low-energy  $\mu^-$  beam, which can then be cooled with "frictional cooling". This is ionization energy-loss cooling at kinetic energies that are small enough that energy cooling is naturally damping (< 20 keV). The frictional cooling process has been demonstrated at PSI, but extrapolation to  $\mu^+\mu^-$  Collider intensities is problematic.[4.10]

2) Optical stochastic cooling: Stochastic cooling has a natural cooling time set by:

$$\tau_{\text{cool}} \equiv \frac{N}{W} \quad (10)$$

where  $W$  is the bandwidth of the cooling system (pickup and kicker) and  $N$  is the number of particles. In optical stochastic cooling the pickup and kicker are magnetic wigglers producing light near optical frequencies, with  $W \sim 10^{14}s^{-1}$ . It is in principle possible to cool 100 GeV  $\mu$ 's before decay.[4.11] However, practical difficulties are significant.

Both of these methods have the potential of cooling  $\mu$ 's to emittances much smaller than the limitations of ionization cooling. They could be used to increase luminosity beyond the current Higgs factory specifications or be applied to later, higher-energy collider scenarios. It is, of course, conceivable that still other methods may be developed and applied to the problem of  $\mu$ -cooling, and these methods may include some of the concepts we have presented, as well as yet to-be-invented components.

## 5. Emittance Exchange Development

Developing a practical method of implementing emittance exchange is an essential requirement for building a Higgs Factory or any other  $\mu^+\mu^-$  collider. A number of schemes have been proposed, some of which are summarized below:

- standalone lattice sections
- small dispersion superimposed on transverse cooling lattice

- ring coolers
- helical capture of unbunched beam
- bunch stacking
- special rf cavity modes

These and other emittance exchange ideas are summarized on a webpage [5.1] devoted to emittance exchange efforts and in the proceedings of a workshop [5.2] held in September 2000.

None of these concepts has yet been developed to the degree of detail that has been obtained for the feasibility study transverse cooling scenarios. In particular, detailed simulations (using ICOOL and Geant4), including all of the underlying physics of ionization cooling and emittance exchange, with integration of the cooling segment into a complete scenario, have not yet been accomplished. While analytical tools for the understanding of ionization cooling in radially symmetric, solenoidal-focusing systems have been established [5.3,4], we must now add dispersion and include nonsymmetric transport and absorbers with the goal of obtaining simultaneous transverse and longitudinal cooling in all dimensions. This much more complex problem has not yet been completely developed. (Solenoids, which provide radial focussing, add beam rotation as well as amplitude-energy correlations in non-axially-symmetric optics. The optics is particularly difficult when the fields are defined by coil locations rather than field strengths.)

### 5.1 Standalone lattice sections

In a “standalone” lattice section, a large amount of emittance exchange is done in a section of lattice isolated from the transverse cooling section. One would cool transversely until the longitudinal heating became unacceptable. Then a pure emittance exchange section would be inserted to reduce the longitudinal emittance back down to an acceptable level. The process could continue through as many stages as needed to achieve the final emittance requirements.

The first implementation of this idea, described above, was presented in the Status Report [1.1,5.5], where the emittance exchange is achieved by using bent solenoids to generate dispersion within a focusing channel, with wedges placed at high-dispersion points. The goal was to achieve a factor of 3 reduction in momentum spread, with a corresponding increase in transverse emittance. Simulations of this scheme showed good exchange in 5-D emittance, that is, between transverse emittance and energy spread (the system did not contain rf cavities and the bunch length was ignored). Subsequent attempts to add 800 MHz rf cavities and track the longitudinal motion were not very successful [5.6], largely due to emittance dilution in the longitudinal motion and longitudinal-transverse couplings.

Recent developments would use lower-frequency rf to ease the longitudinal matching problem, and smaller dispersion to reduce the uncorrected correlations among the phase space variables. A modified version of this scheme uses separated, nearly isochronous regions of the lattice to introduce the dispersion [5.7]. There is then no rf in the dispersive region. This scheme uses smaller dispersion and aims for a smaller amount of exchange in each stage.

### 5.2 Small dispersion superimposed on a transverse cooling lattice

The idea here is to take a successful transverse cooling lattice and superimpose a small amount of dispersion on the lattice. The small dispersion could come from dipoles, bent solenoids or helical dipoles. The dispersion is assumed to be small enough that it does not greatly perturb the transverse cooling behavior.

Calculations of the expected performance from adding dipole fields to a sFOFO transverse cooling lattice were done by Palmer [5.8]. He put a dipole over the rf cavity in the middle of the lattice cell. A gradient-dipole field (with gradient index  $n = \frac{1}{2}$ ) produced equal focusing in both transverse planes, while adding dispersion in one. The bend angles were  $45^\circ$ , giving a dispersion  $\eta \approx 40$  cm. In order to get emittance exchange in both transverse planes, an  $18^\circ/\text{cell}$  helical twist is introduced into the lattice. Variations of this idea have also been studied by others [5.2].

Another scheme [5.9] uses a rotating dipole to generate dispersion in a single flip transverse cooling channel. A simulation was made of a 72 m long channel with a 0.3 T rotating dipole in a 5 T solenoid. LiH wedge absorbers were spaced periodically down the channel. The beam was reaccelerated using 201 MHz rf cavities. Including nonlinearities and an initial momentum-transverse amplitude correlation, the 6-D emittance was reduced from  $5300 \text{ mm}^3$  to  $1350 \text{ mm}^3$ . The beam transmission was 81%. Another simulation of the same scheme, using Geant4 [5.10], obtained similar answers.

### 5.3 Ring coolers

The ring cooler uses some of the dispersion from the bending dipoles together with wedges to incorporate emittance exchange. A number of ring designs have been proposed by Valeri Balbekov.

(1) The first design [5.11] used a ring with a  $\lambda I$  transfer matrix per turn, where  $\lambda$  is the cooling factor. In this case the variables in 6-D phase space are independent. It used two bending sections with wedge absorbers. The dipoles had an  $n=0.5$  quadrupole gradient superimposed. The straight sections had rf cavities and LiH absorbers for transverse cooling. Skew quadrupoles were used to control dispersion in the straight sections. The beam was injected at 225 MeV/c and circulated at 9.3 MHz revolution frequency. Cooling takes place primarily through a reduction in transverse size and bunch length. The simulation consisted of a mix of tracking in the absorbers and matrix transport. The 6-D emittance was reduced from  $11 \times 10^4 \text{ mm}^3$  to  $24 \text{ mm}^3$ . Approximately 25% of the muons were lost because of aperture restrictions and 25% from decays.

(2) In the second design [5.12] alternating direction solenoids were incorporated in a racetrack ring. Solenoids focus the beam leaving the LiH absorbers into the rf cavities. The arcs contain a bent solenoid superimposed on the gradient dipoles. The ring has a circumference of 42.5 m. The muon momentum is  $\sim 330$  MeV/c and the revolution frequency is 6.2 MHz. Transverse and longitudinal nonlinearities significantly reduced the predicted linear cooling performance. Including nonlinearities the 6-D emittance was reduced from  $4.6 \times 10^4 \text{ mm}^3$  to  $1.2 \times 10^3 \text{ mm}^3$ . Approximately 50% of the muons were lost because of aperture restrictions and decays, as well as mismatch from the initial beam.

(3) A higher frequency ring has also been developed [5.13]. This is a 32.6 m circumference ring with 201 MHz rf and liquid hydrogen absorbers. There are 8 dipoles in the arcs with superimposed solenoid fields and LiH wedges. The simulations still lack realistic fringe fields around the dipole magnets and any method of injection or extraction. The 6-D emittance of a pre-bunched beam was reduced from  $2800 \text{ mm}^3$  to  $200 \text{ mm}^3$  in 10 turns. The transmission was 60%.

### 5.4 Helical capture

Y. Derbenev [5.14] has proposed a “sweeping” method to reduce the energy spread of an initially unbunched muon beam. Dispersion is created using a helically-rotating dipole field.

Wedges must be placed periodically along the channel. Preliminary simulations [5.15] show that the method works in principle with ideal beams, but does not work with the large-emittance muon beams collected from the production target.

### 5.5 Bunch stacking and transverse cooling

If trains of bunches are used in the cooling sections, some form of bunch stacking must be provided before the beam reaches the collider ring. One scheme [5.16] proposed synchronizing the bunches in time by separating them with a transverse deflector into individual time delay lines. The bunches are then recombined by merging them in transverse momentum space. One important issue is the emittance dilution caused by the stacking process. Bunch combinations of this type dilute phase space by at least a factor of 2, which implies that proportionately more cooling will be required after bunch combination.

Preliminary simulation work on this idea has begun [5.17]. A kicker magnet inside a solenoid was used to create different temporal paths for 10 bunches. A sector magnet channel was used for stacking the time-synchronized bunches.

### 5.6 rf cavity modes

There have been proposals to use rf cavity modes to reduce the energy spread in the beam [5.18]. The cavity must be placed in a dispersion region. The beam can be sent through a normal accelerating cavity off-axis, such that there is a transverse variation of accelerating field. Any such exchange that does occur is believed to take place through non-linear processes only [5.19].

### 5.7 emittance exchange overview and plans

Following completion of the neutrino factory design study 2, the collaboration will resume more intensive studies of the emittance exchange methods discussed above, and will develop the most promising of these into engineering designs that may be included in v-Factory and/or  $\mu^+ - \mu^-$  Collider designs.

## 6 Simulation summary

It is worthwhile at this point to summarize how much of the required cooling effort has been simulated in detail so far. Figure 19 is a plot of normalized transverse versus longitudinal phase space, and it displays the initial and final emittances of some cooling schemes and simulation results

The beam collected from the target is shown at START in the upper right corner. The contour for a constant 6-dimensional normalized emittance of  $0.17 \text{ mm}^3$ , which is the final emittance goal of the SR collider scenarios [1.1], is shown by the dotted line in the lower left corner. The emittance specifications of a low- $\delta E$  100 GeV Higgs Factory and of a 3 TeV  $\mu^+ - \mu^-$  Collider are shown as points on that contour. The emittance specifications of a high-luminosity Higgs factory would be also be on that contour, roughly halfway between the reference points. The general goal of a collider cooling system is to obtain a scenario that takes the beam from the START parameters to this collider contour, and develop a complete simulation of that scenario. The solid line connecting the starting point to the Higgs Factory shows the proposed baseline SR cooling scenario [1.1]. (This is a scenario cooling path and is not yet completely simulated.)

Simulations of cooling devices follow trajectories within this diagram. The study 2 neutrino factory goal is the square marked NFPJK. Three neutrino factory simulation results are shown as lines on the chart: FS1FO3, a 3 T FOFO cooling lattice and FS1FL5, the 5 T single flip solution from Feasibility Study 1; and FS2SF5, the 5 T sFOFO solution from Feasibility Study 2. These simulations all start at a lower longitudinal emittance than is produced at the target. This is because the neutrino factory beam is split into a bunch train and the longitudinal emittance displayed is that of each bunch in the train. (The Higgs factory cannot use a bunch train unless a stacking ring is available to later recombine the bunches.) These solutions also end at a lower longitudinal emittance than their initial point; this results from beam losses in the tails of the longitudinal distributions and not from actual cooling. Two simulations are shown that include emittance exchange and thus provide some real longitudinal cooling. VBHEL5 uses a helical dipole to generate dispersion in a 5 T single flip solenoid channel. The RING COOLER uses the dipole fields to generate dispersion and incorporates wedges in the ring.

The three simulations marked SRAS are alternating solenoid lattice solutions from the Status Report. These solutions have good transmission and show the characteristic behavior of cooling transversely while heating longitudinally, since the simulations included no longitudinal cooling. They show transverse cooling over much of the desired range, with the 31 T solution ending very near the collider cooling requirement. The SRLi10 simulation is the cooling in a single 10 T surface field lithium lens. It also nearly reaches the target final emittances; a 15 T lens would probably achieve them.

Three points should be obvious from this summary. (1) Many different simulations have shown that transverse cooling should be possible over the required range of transverse emittances. (2) Emittance exchange is a critical technology that is necessary to tie the transverse cooling sections together in such a way that we can follow the baseline scenario down to the Higgs Factory goal. (3) Sufficient scenario development with cooling simulations has not been done yet to completely cover the desired cooling range. Therefore, there is still a great deal of simulation work that needs to be done before we have a self-consistent plan, listing the various cooling sections in a single complete scenario that satisfies the Higgs Factory requirements.

## 7. Scenarios of front ends for $\mu^+ - \mu^-$ Colliders

In this section we describe some paths toward complete collider cooling scenarios, based on our existing cooling and simulation experience. A neutrino factory front end is not directly usable in a muon collider because (1) the neutrino factory muon beam is spread out over a long series of bunches; (2) the normalized emittance of the beam is much larger than that required for the collider; and (3) no provision for emittance exchange is included. We consider in the following several options for converting an existing neutrino factory facility so that it is suitable for use in a Higgs factory. Any option will require significant alterations of the existing facilities. Fortunately, the research since the SR does suggest some potential new approaches for phase rotation and cooling.

An important focus of future research will be to determine whether the v-factory cooling system could be extended or enlarged to obtain  $\mu$ -collider beams. The system described in the v-Factory Feasibility study could provide a substantial amount of the needed transverse cooling. A following linac cooler with stronger focusing, including Li lens focusing, could readily be added. The longitudinal emittance per bunch is similar to that required for the high luminosity  $\mu^+ - \mu^-$  Collider, but the  $\mu^+ - \mu^-$  Collider requires that the beam be concentrated in a small number of bunches (combine  $\sim 70$  bunches to 1—4 bunches), so a bunch combiner system with beam cooling would also be needed. Some concepts toward the required bunch combination are

discussed below. Some longitudinal cooling is required, at least to the level of avoiding longitudinal emittance dilution in the cooling channels and/or in enabling bunch combination. (see below). Also separate or combined cooling channels which can simultaneously obtain  $\mu^+$  and  $\mu^-$  bunches would be required.

### 7.1 Status Report-based Solution

One possible path is to return to the 1999 status report collider scenario [1.1], without explicitly including the v-factory front end. That design uses a single bunch of each charge, has many cooling stages, and uses emittance exchange. The initial  $\mu$ -beam would be captured and rf-rotated in a low frequency rf system ( $\sim 30$  MHz), and the initial transverse cooling system would also be at  $\sim 30$  MHz, with following cooling systems at higher frequencies. These low-frequency cooling systems must also include a lot of longitudinal cooling, since the shorter bunches needed for increased rf frequency are obtained through longitudinal cooling. Since this option would not use any of the existing front end facilities, it is likely the most expensive option.

### 7.2 Ring-Cooler based scenarios

The key difficulties in the SR scenario are the high cost of a single-pass linear cooling system and the awkward inclusion of emittance exchange with many rebunchings and rf frequency changes. These difficulties can be reduced by inclusion of “ring coolers” to provide much of the necessary cooling.

The scenario outline would be similar to the SR scenario. The  $\mu$ -bunches would originate from single proton bunches, with an initial linac based phase-energy rotation that would give a  $\sim 30$  MHz bunch (full length  $\sim 8$  m) and rms momentum spread of  $\sim \pm 15\%$ . This section may or may not include a “minicooling” absorber or an initial wedge absorber to provide some initial cooling. These single bunches would then be injected into a ring cooler with low-frequency ( $\sim 30$  MHz) rf (similar to the ring cooler of ref. [5.12]), and cooled for  $\sim 10$  turns, during which the transverse emittance could be reduced to  $\epsilon_{T,N} \sim 4$  mm-rad and longitudinal emittance would be reduced by an order of magnitude, and bunch lengths would be reduced by at least a factor of 5. The bunches would be kicked out and, after perhaps a single matching/cooling stage, inserted into a second ring cooler with higher-frequency rf ( $\sim 200$  MHz) for  $\sim 10$  turns of cooling with a transverse cooling goal of  $\epsilon_{T,N} \sim 1$  mm-mrad (or less), accompanied with longitudinal cooling by a factor of  $\sim 4$  (or more?) to  $\epsilon_{L,N} \sim 0.5$  cm (or less). The beam would be extracted into a (predominantly) linear cooler for (mostly) transverse cooling to  $\epsilon_{T,N} \sim 0.2$  mm-mrad (or less). This linear cooling may be Li-lens based, similar to the cooling in ref. [4.5]. From there some final cooling and wedge/antiwedge emittance-exchange stages could bring the bunches to collider requirements.

### 7.3 Recycle maximum amount of v-Factory front end

It is quite possible that the first high-intensity stored- $\mu$  facility would be a v-Factory, and it will be natural to extend that existing system to collider parameters. A v-factory  $\mu$ -storage ring facility has a similar total number of  $\mu$ 's as a collider and the longitudinal emittance per bunch is also similar to that required by the high luminosity HF. However, the  $\mu$ 's are split up into a string of bunches ( $\sim 50$ ), and these must be combined to obtain the high-intensity bunches needed in a

high-luminosity collider. The v-Factory also has transversely cooled bunches (with  $\epsilon_{\perp} \approx 0.002$  m-rad), and transverse cooling by another order of magnitude is required for collider luminosities.

If we keep the induction linacs after the target for phase rotation, we obtain a single, very long bunch. Then we would probably use the existing NF buncher system also and make a long train of bunches. Part of the existing NF cooling channel could also be used as the first stage of cooling. This would then have to be followed by a series of new emittance exchange and transverse cooling stages. Part of this cooling and exchange could possibly be done using cooling rings. A new stacking ring would be needed after the cooling to coalesce the bunch train into a single bunch. It is likely that an additional stage of cooling will be needed after the stacking ring to counteract any emittance growth in the stacking process.

In summary a scenario based on a v-source then requires three key additional components: a bunch combiner, with cooling to obtain the same emittance within a single bunch of that of one of the ~50 separate bunches, a second-stage transverse cooler to reduce transverse emittances by an order of magnitude and emittance exchange. A bunch combiner concept was discussed above. The second-stage transverse cooler is in principle possible, and an example of its implementation would be an extension of the 5-Li lens cooler of Balbekov [4.5].

#### 7.4 Replace NF precooler with adiabatic buncher

The adiabatic buncher [7.1] uses a long drift followed by a series of rf cavities with sequentially varying frequencies (about ~200 MHz) after the production target to form the long bunch which is then rotated to form a string of bunches, which is then rf-rotated by a high frequency rf system to form a string of bunches of equal energies, similar to the beam after the induction linac + buncher of v-Factory studies 1 and 2. The adiabatic buncher is an alternative to these NF systems. It may be much cheaper and offers performance similar to that of the induction linacs and buncher in the NF studies, except that the same system would form strings of both positive and negative  $\mu$ -bunches. These strings must then be cooled and recombined into high-intensity  $\mu^+$  and  $\mu^-$  bunches for the  $\mu^+ - \mu^-$  collider. It would thus need the additional transverse cooling and emittance-exchange stages, and stacking ring and final cooling, as in option 3 above.

## 8 Status and Plans

In this chapter we have discussed a variety of paths toward a  $\mu^+ - \mu^-$  Collider cooling system. Initially more simulation and analytical exploration is needed in exploring these possible paths. This must be followed by hardware R&D on key components to determine their practicality, as well as performance limits. The key R&D questions are:

- longitudinal cooling: We need to determine which of the many potential cooling methods we have discussed are most effective and practical.
- low-frequency rf: Many potential scenarios require high-gradient low-frequency rf.
- ring cooler: While the general cooling capabilities of a ring cooler look very promising, further simulations of a complete system are needed, including injection-extraction and all fringe field effects.
- bunch recombiner: Simulations are needed showing the stacking efficiency and emittance dilution.

- fast kickers: Both the ring cooler and bunch recombiner scenarios require advances in the technology of fast kickers to inject and extract large-phase-space muon bunches with minimal dilution.
- Li lens cooling: The practical possibilities and limitations of Li lens system must be determined.

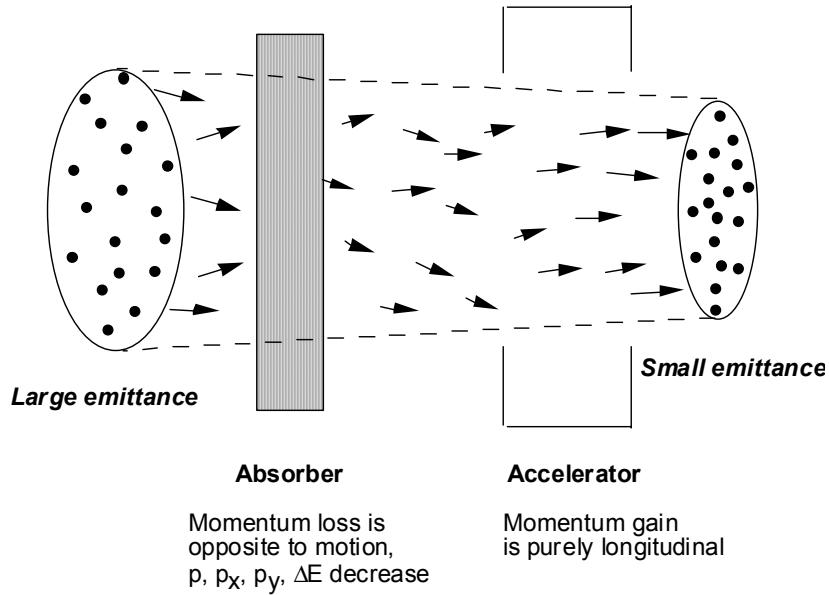
The progress in developing these technologies will determine which of them could be implemented in a future  $\mu^+ - \mu^-$  collider cooling scenario.

## References

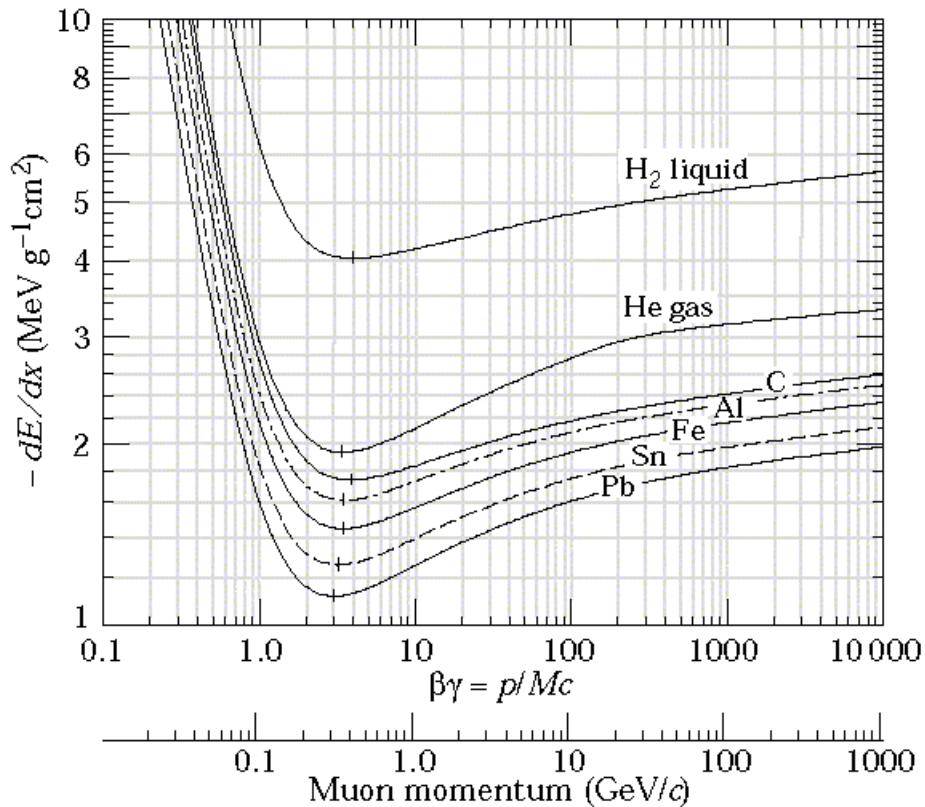
- [1.1] C. M. Ankenbrandt et al., Phys. Rev. Special Topics Accel. Beams 2, 081001 (1999).
- [1.2] A. N. Skrinsky and V.V. Parkhomchuk, Sov. J. Nucl. Physics **12**, 3(1981)
- [1.3] E. A. Perevedentsev and A. N. Skrinsky, Proc. 12th Int. Conf. on High Energy Accel., 485 (1983).
- [1.4] D. Neuffer, Particle Accelerators **14**, 75 (1983)
- [1.5] D. Neuffer, Proc. 12th Int. Conf. on High Energy Accelerators, 481 (1983)
- [1.6] D. Neuffer, Nucl. Inst. and Meth. **A350**, 27 (1994)
- [1.7] Particle Data Group – C. Caso et al., “Review of Particle Physics”, European Physical Journal **C3** (1998).
- [1.8] U. Fano, Ann. Rev. Nucl. Sci. **13**, 1 (1963).
- [1.9] D. Neuffer and A. Van Ginneken, Nucl. Inst. and Meth. **A403**, 1 (1998); see also refs. [1.2-6]
- [1.10] D. Neuffer, “Phase Space Exchange in Thick Wedge Absorbers for Ionization Cooling”, **AIP Conf. Proc. 441**, 270 (1998).
  
- [3.1] N. Holtkamp and D. Finley, ed. “A Feasibility Study of a Neutrino Source Based on a Muon Storage Ring”, Fermilab, April 15<sup>th</sup>, 2000.
- [3.2] S. Ozaki, R. Palmer, and M. Zisman, ed., “BNL v-Factory Feasibility Study”, April, 2001.
- [3.3] R. Fernow, “ICOOL: a simulation code for ionization cooling of muon beams,” Proc. 1999 Particle Accelerator Conference, New York, p. 3020-2.
- [3.4] The DPGEANT and GEANT4 codes have been developed at FNAL under the direction of P. Lebrun.
- [3.5] V. Balbekov, “Comparison of an alternate solenoid and long solenoid cooling channels”, NuMu note 130, 2000.
- [3.6] A. Lombardi, “a 40-80 MHz system for phase rotation and cooling”, CERN–NUFACT Note 34, 1 August 2000.
  
- [4.1] B. F. Bayanov et al., Nucl. Inst. And Meth. NIM 190, 9 (1981)
- [4.2] G. I. Silvestrov, “Lithium Lenses for Muon Colliders”, **AIP Conf. Proc. 372**, 168 (1996).
- [4.3] D. Neuffer and A. Van Ginneken, Proc. 1997 PAC, p. 417, Vancouver BC (1997).
- [4.4] P. Spentzouris and D. Neuffer, Proc. 1999 PAC, New York, NY, p. 3083 (1999).
- [4.5] V. Balbekov, 1999 PAC, New York, NY, p. 3146 (1999).
- [4.6] G. I. Silvestrov, A. N. Skrinsky, T. Vsevolozhskaya, Proc. 1999 PAC, New York, NY, p. 3091 (1999).
- [4.7]  $\mu^+ \mu^-$  Collider - A Feasibility Study, BNL-52503, Fermi-Lab-Conf.-96-092, LBNL-38946 (1996), presented at the Snowmass 96 workshop (1997).
- [4.8] D. Neuffer, “ $\mu^+ - \mu^-$  Colliders”, CERN 99-12 (1999).
- [4.9] K. Nagamine, “New  $\mu^+ - \mu^-$  Cooling for  $\mu^+ - \mu^-$  Colliders and Possible Realization at JHF/KEK”, AIP Conf. Proc. **441**, p. 295 (1998).

- [4.10] M. Muhlbauer et al., “Frictional Cooling : Latest experimental results and first application”, Nucl. Phys. B (Proc. Suppl.) **51A**, 135 (1996).
- [4.11] A. A. Mikhailichenko and M. S. Zolotorev, **PRL** **71**, 4146 (1993).
- [5.1] [http://needmore.physics.indiana.edu/~gail/emittance\\_exchange.html](http://needmore.physics.indiana.edu/~gail/emittance_exchange.html)
- [5.2] <http://www.cap.bnl.gov/mumu/exchange/>
- [5.3] G. Penn & J. Wurtele, Beam envelope equations for cooling of muons in solenoid fields, Phys. Rev. Lett. 85:764-7, 2000.
- [5.4] K. Kim & C. Wang, Formulas for transverse ionization cooling in solenoidal focusing channels, Phys. Rev. Lett. 85:760-3, 2000.
- [5.5] R. Palmer, Emittance exchange & full system, Particle Accelerator School, Vanderbilt, TN, 1999; <http://pubweb.bnl.gov/people/palmer/course/9exch.ps>
- [5.6] R. Fornow, Emittance exchange using thick bent solenoid elements, NuMu note 166, 2000.
- [5.7] Chun-xi Wang, Solution to emittance exchange?: design principles and a possible solution, BNL Emittance Exchange Workshop 2000; <http://www.cap.bnl.gov/mumu/exchange/Wang1.pdf>
- [5.8] R. Palmer, Emittance exchange with super FOFO lattice, 8/17/00; <http://pubweb.bnl.gov/people/palmer/nu/exchange/exch2.ps>
- [5.9] V. Balbekov, “Helical channel studies and simulations”, NuMu note 146, 2000.
- [5.10] D. Elvira, P. Lebrun & P. Spentzouris, “Simulation of helical cooling channel using Geant4”, NuMu note 193, 2001.
- [5.11] V. Balbekov & A. van Ginneken, “Ring cooler for muon collider”, AIP Conf. Proc. 441:310-3, 1998.
- [5.12] V. Balbekov, “Possibility of using a ring accelerator for ionization cooling of muons”, Proc. 1999 Part. Acc. Conf., p. 315-7.
- [5.13] V. Balbekov, “Ring cooler update”, NuMu note 189, 2001.
- [5.14] Y. Derbenev, “Conceptual studies on ionization cooling of muon beam”, NuMu note 185, 2000; “Sweeping method to cool a large energy spread of initial muon beam”, NuMu note 135, 2000; “Resume on sweeping cooling and cooling agenda in general”, NuMu note 133, 2000.
- [5.15] R. Fornow & J. Gallardo, “Emittance exchange using the helical sweeping method,” NuMu note 187, 2001.
- [5.16] C. Kim, “An emittance exchange idea using transverse bunch stacking,” NuMu note 70, 1999.
- [5.17] Y. Fukui, “Stacking mini- $\mu$  bunches for emittance exchange,” BNL Emittance Exchange Workshop, 2000; <http://www.cap.bnl.gov/mumu/exchange/Fukui3.pdf>
- [5.18] Kirk McDonald, “Compression of beam energy via off-axis traversal of an rf cavity”, Princeton note / $\mu\mu$ /97-1.
- [5.19] E. Courant, “Conservation of phase space in Hamiltonian systems and particle beams”, in R. Marshak (ed), Perspectives on Modern Physics, Wiley, 1966, p. 257-60.
- [7-1] D. Neuffer, “A High-frequency Buncher and  $\phi$ - $\delta E$  rotation for the  $\mu^+ - \mu^-$  Source”, MuCool Note , October 2000, see also Proc. PAC2001.

**Figure 1.** Concept of ionization cooling.

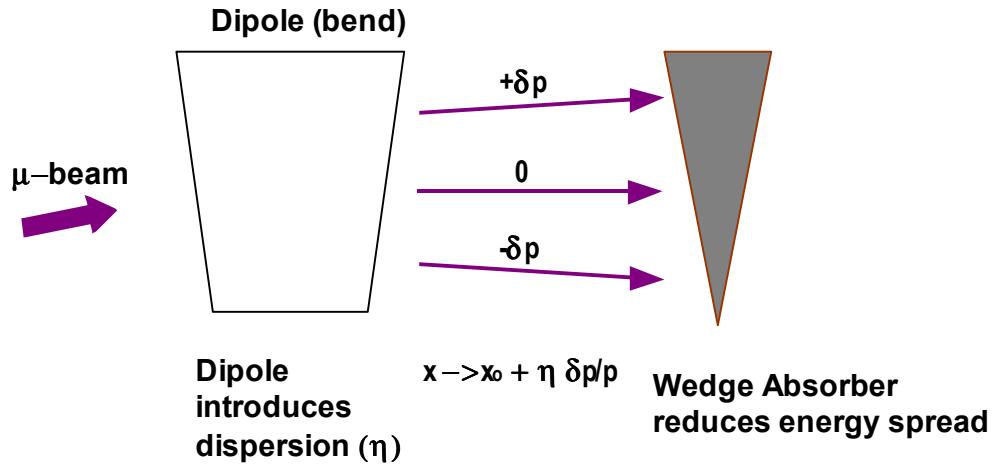


**Figure 2.**  $(dE/dx)/\rho$  (MeV/(gm/cm<sup>2</sup>)) as a function of muon momentum  $P_\mu$  for various atoms. Note that this function is heating (negative slope) for  $P_\mu < \sim 0.350$  GeV/c and becomes strongly heating (steep slope) for  $P_\mu < 0.200$  GeV/c.

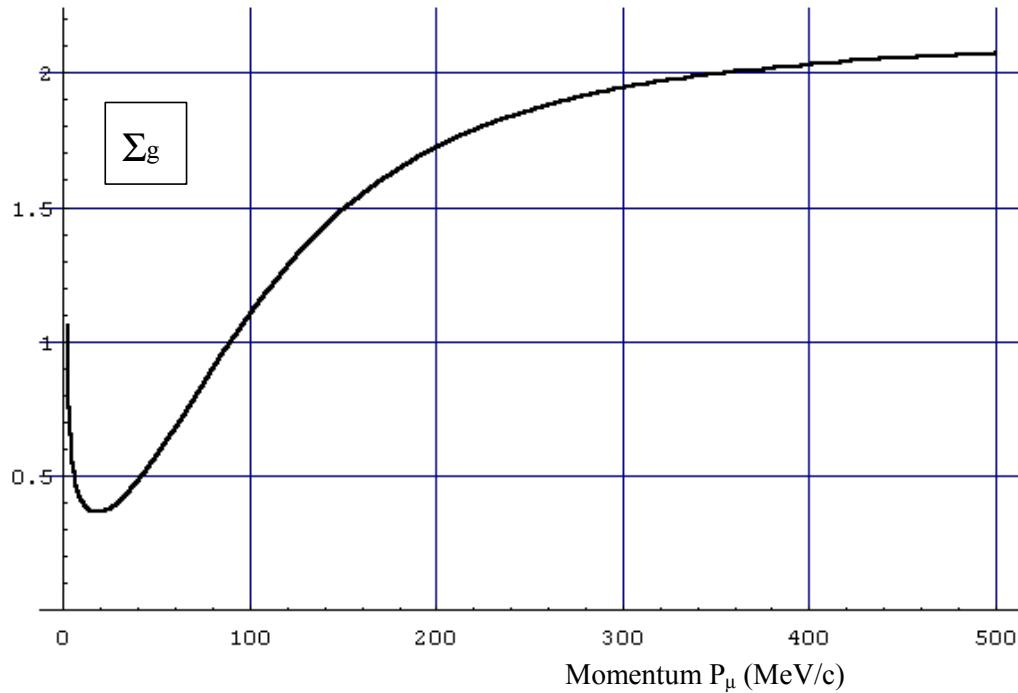


**Figure 3.** Overview of an emittance exchange section, in which longitudinal emittance is reduced by using a wedge absorber at nonzero dispersion.

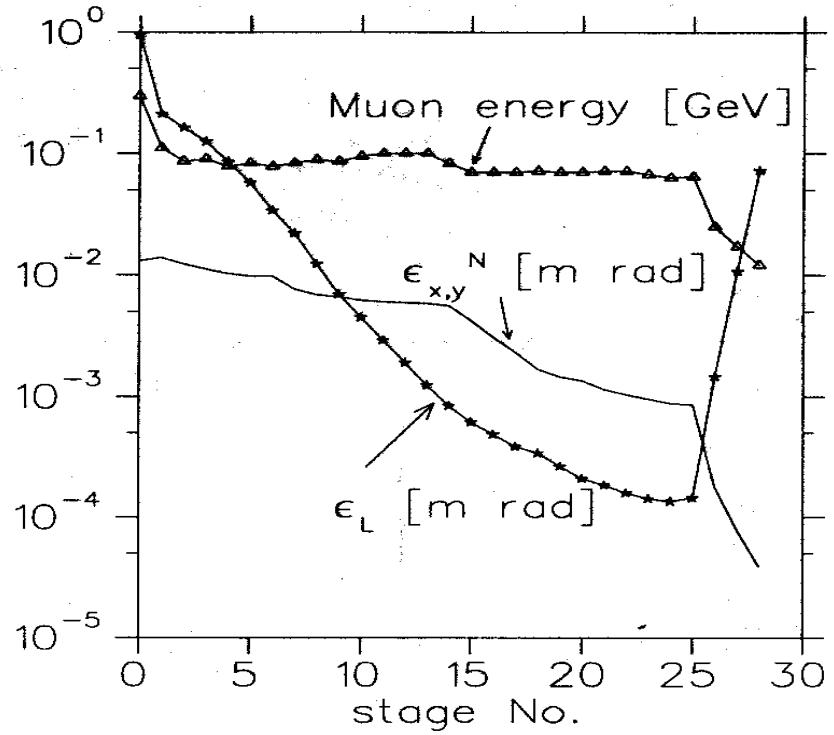
## Emittance exchange overview



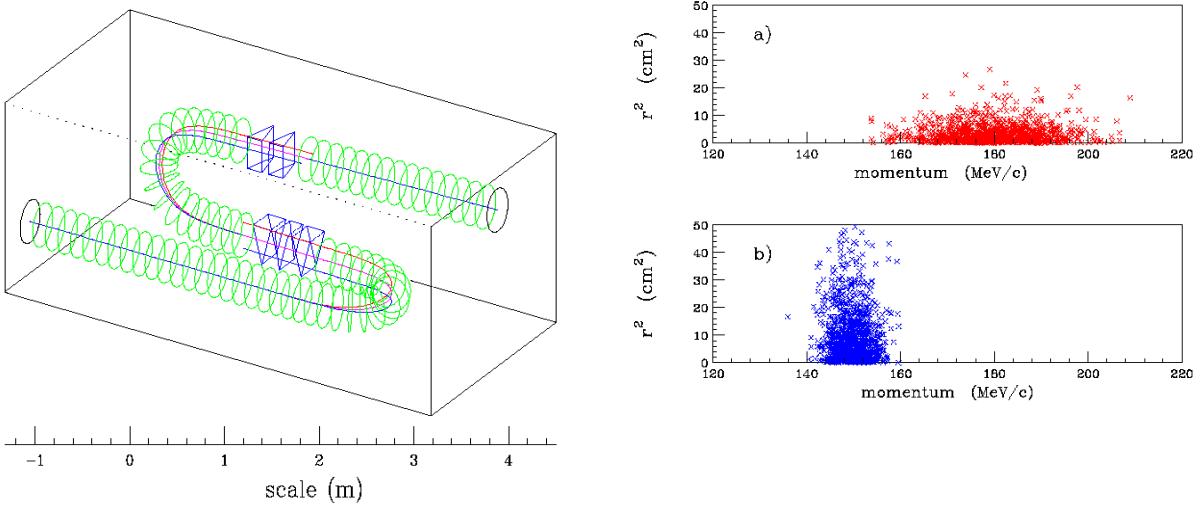
**Figure 4.** The sum of the cooling partition numbers  $\Sigma_g = (g_x + g_y + g_L)$  as a function of momentum  $P_\mu$  (0—500 MeV/c).  $g_x$  and  $g_y$  are naturally 1 while  $g_L$  becomes strongly negative for  $P_\mu < 200$  MeV/c.  $\Sigma_g$  remains greater than 0, which means that ionization loss remains intrinsically cooling at low momenta.

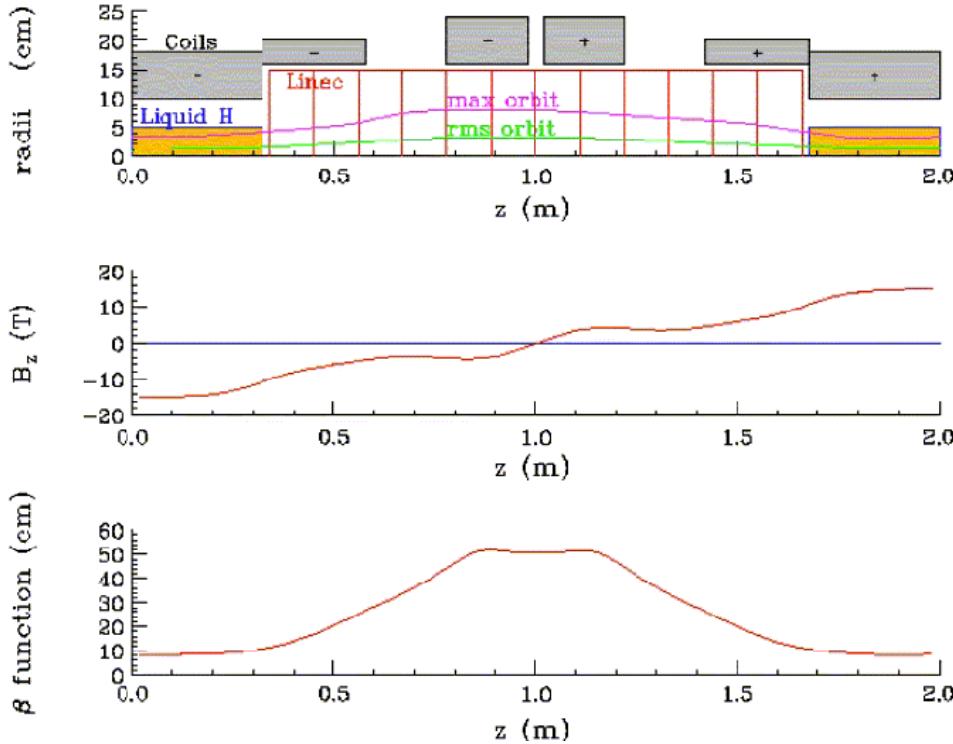


**Figure 5.** SR baseline cooling scenario

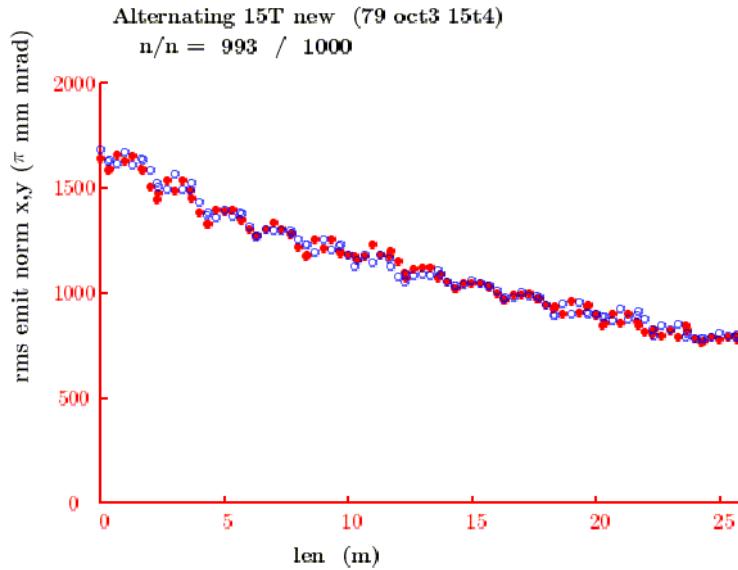


**Figure 6.** SR emittance exchange example. The figure on the left shows a schematic view of the bent solenoid system with wedges. Simulation results are shown on the right: (a) initial beam distribution of  $r^2$  versus  $p$ ; (b) final distribution of  $r^2$  versus  $p$ .

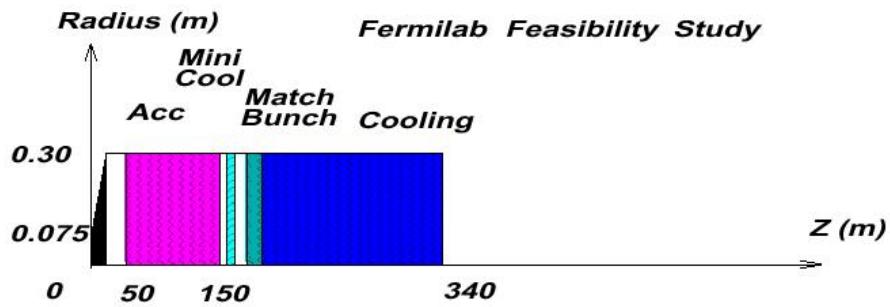




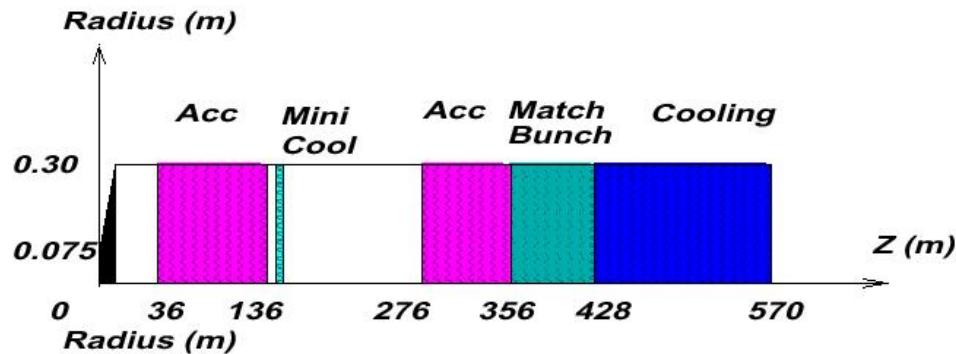
**Figure 7.** Status report cooling cell. The upper plot is a radial cross-section of a schematic view of a 2 m cooling shell showing hydrogen absorbers, a 12-cell 800 MHz rf cavity and focusing coils around the absorber and rf. The second plot is the solenoidal magnetic field and the third plot is the focusing function  $\beta^*$ , showing  $\beta^* \equiv 10\text{cm}$  in the absorbers.



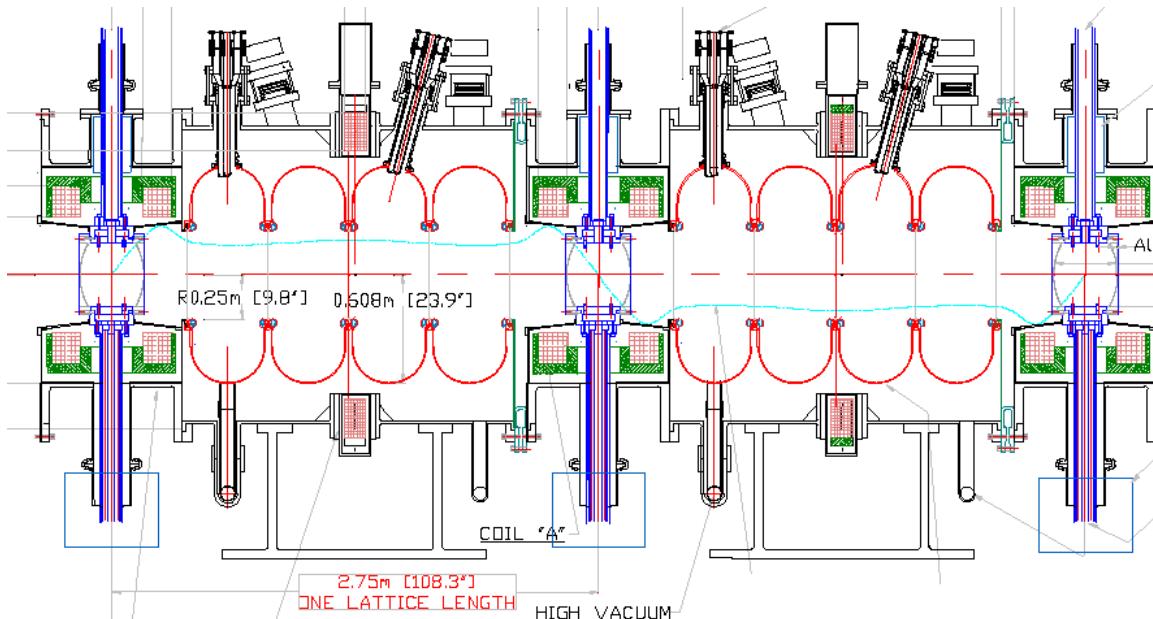
**Figure 8.** Simulation results showing transverse cooling in  $\epsilon_{T,N}$  from 1.7 mm to 0.8 mm in 13 cells ( $\sim 1$  section) of the Status Report cooling channel.



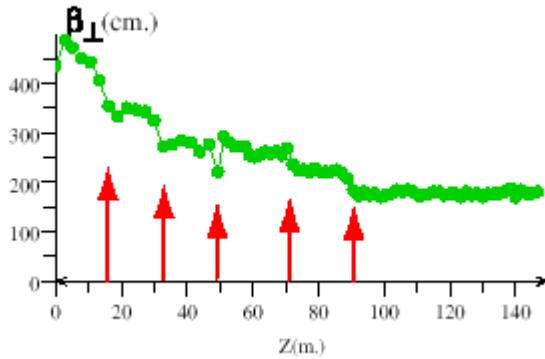
**Figure 9a.** Study 1 layout of  $\mu$ -capture and cooling transport.



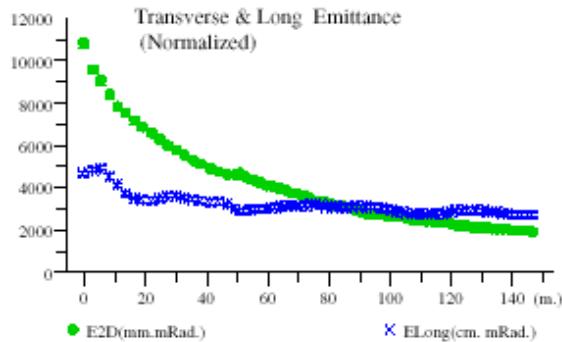
**Figure 9b.** Study 2 layout of  $\mu$ -capture and cooling transport



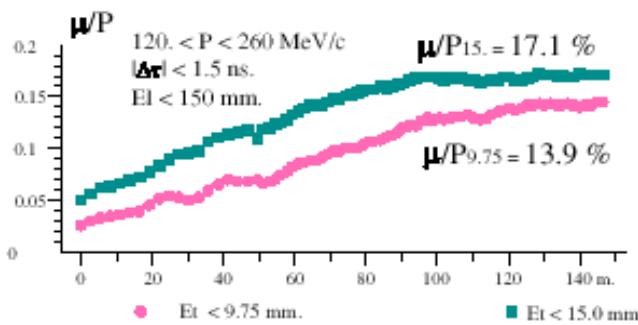
**Figure 10.** Cross-section of two Study 2 2.75m sFOFO lattice cooling cells, showing two 4-cell rf cavities, 3 Liquid hydrogen absorbers, and magnetic coils for focusing around the absorbers and rf cavities.



**Figure 11.** The  $\beta_{L,\min}$  function in mm at cavity centers, as calculated from the second order moments of a beam in GEANT4 simulations, for the entire sFOFO cooling channel. The 5 arrows indicate the beginning of new lattice sections. (In this figure the cooling channel is extended beyond the study2 reference length of 108 m to 144 m by adding 1.65 m cells.)

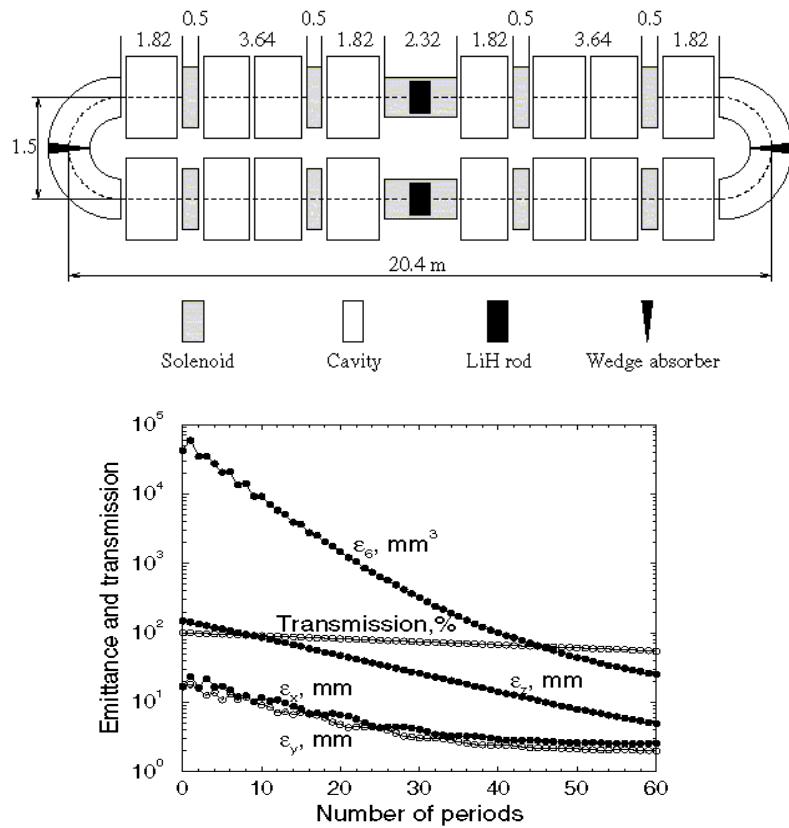


**Figure 12.** The transverse and longitudinal normalized emittances along the cooling channel.

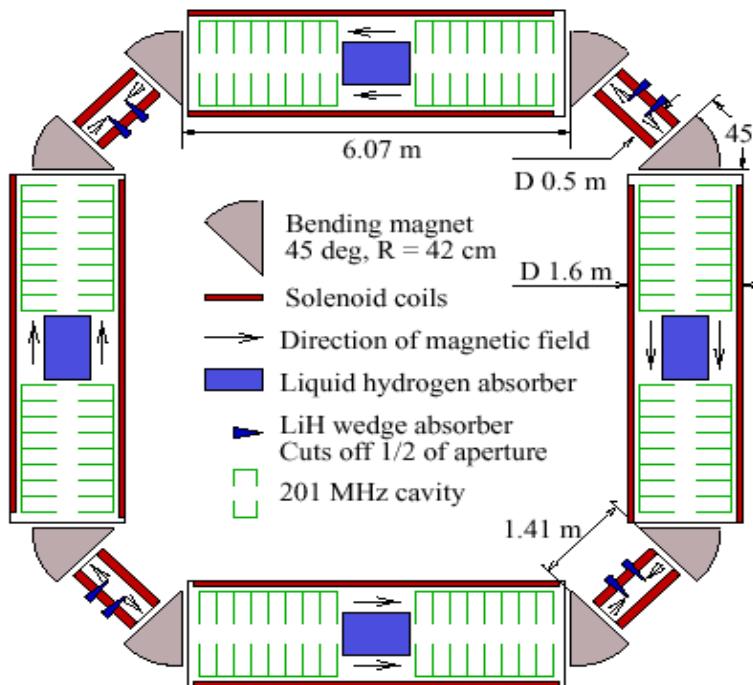


**Figure 13.** The muon to proton yield ratio for the two emittance cuts, showing that the particle density in the center of the phase space increases as the beam is cooled in the channel. The two curves give the number of particles within the baseline longitudinal and transverse acceptances. The upper line represents the values for the accelerator parameters in this study. The lower line, given for comparison, gives the values for the acceptances used in Feasibility Study 1.

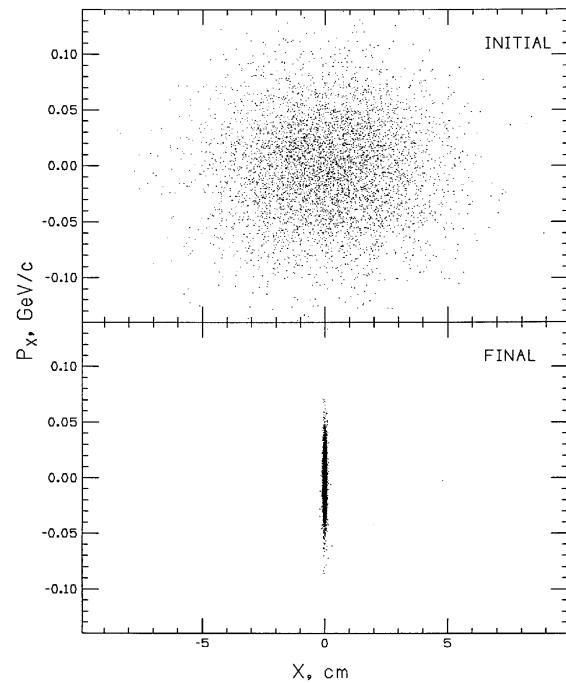
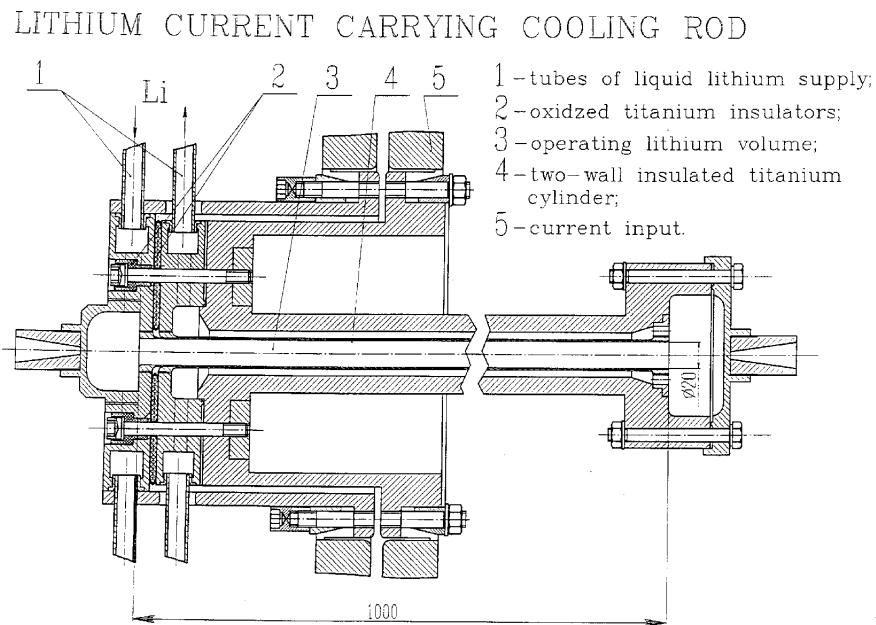
**Figure 14.** Ring cooler A, with simulation results showing cooling over 60 periods (30 turns).



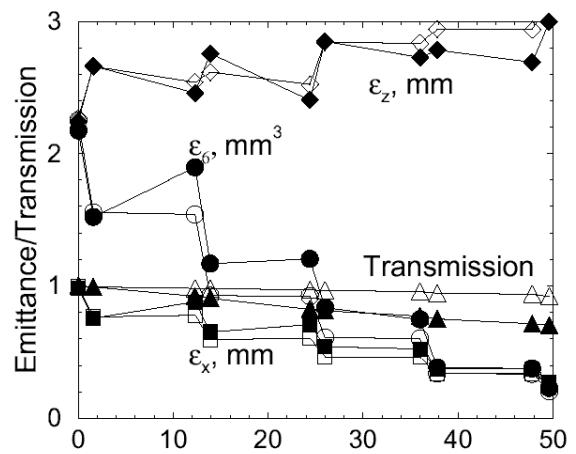
**Figure 15.** Ring Cooler B



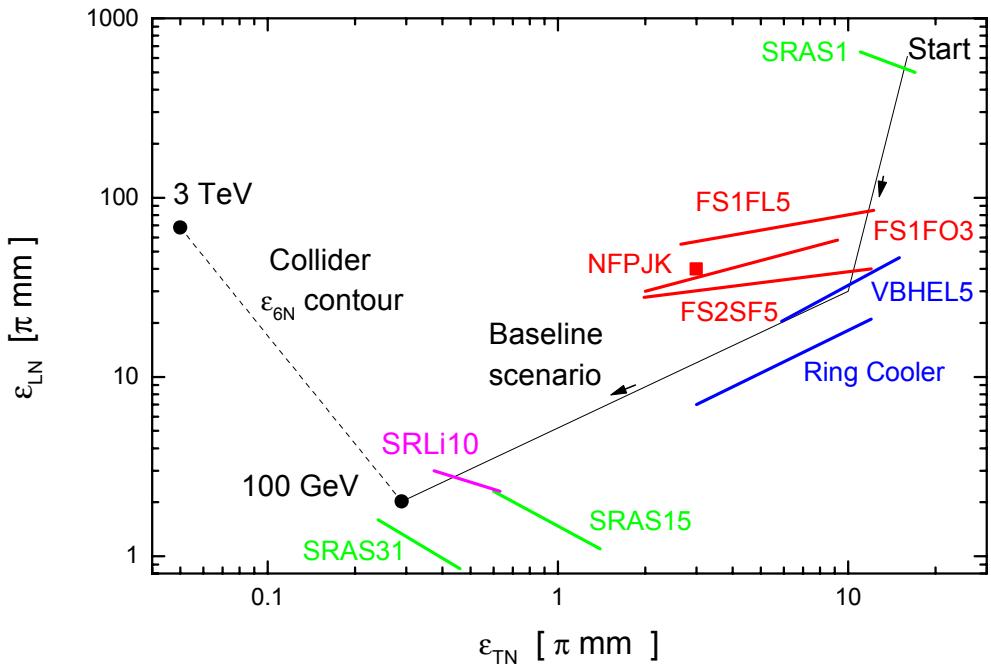
**Figure 16.** Liquid Li lens schematic



**Figure 17.** Transverse phase space reduction from a series of 12 Li lenses.



**Figure 18 –** Simulation of a sequence of Li lenses near the end of cooling



**Figure 19.** Cooling Summary: The dashed line shows collider cooling goals in emittance space, and the solid line shows a possible path from initial beam parameters [Start] to that goal. Segments of cooling that have been designed and simulated in some detail are shown in color.

# Muon Acceleration from 20 to 60 GeV for a Higgs Factory

D. J. Summers, University of Mississippi–Oxford

What is required to accelerate a  $\mu^+$  and a  $\mu^-$  particle bunch from 20 to 60 GeV? A neutrino factory as outlined in the recent Brookhaven study provides 20 GeV muons which have enough energy to explore CP violation in the lepton sector. Further acceleration to 60 GeV may be enough to reach a low mass Higgs as suggested by theory and recent measurements at LEP.

To reduce the RF cost, consider twenty passes through a two GeV Linac and see if enough muons survive decay. A single continuous Linac with teardrop shaped arcs at each end is adopted. Muon decay losses are minimized; muons pass through shorter arcs when their gamma boost is low. To minimize magnet cost  $45^0$  turns are used with short straight sections to line up the arcs. For each teardrop, the length added to the curved sections by the two straight sections is  $(4 - 2\sqrt{2})/2\pi = 18.6\%$ .

Take a muon lifetime of  $2.2 \times 10^{-6}$  seconds, 1.8 Tesla dipoles, a 70% dipole packing fraction, and a 133 meter long 2 GeV Linac with 15 MV/meter. The total magnet bore length required is 7000 meters, 11% longer than the Fermilab Tevatron. Muon survival after twenty passes through the 2 GeV Linac is 95.5%. Squaring this percentage the luminosity is 91.8% of what it would be in a Higgs factory if there had been no decay loss in accelerating the muons from 20 to 60 GeV.

The magnet cell length may have to be short to provide good acceptance for the muons in the arcs. An alternating gradient design where the magnet lamination change shape within a magnet avoids magnet ends and makes it easier to consider superconducting wire rather than copper. The magnets do have to be at full field constantly, so power consumption is an issue.

Finally note that in a dogbone geometry, muons can orbit clockwise in one end and counterclockwise in the other end, minimizing dispersion. This may help to preserve polarization. If muons are 100% polarized, the  $\mu^+\mu^- \rightarrow$  Higgs cross section doubles (versus the case of zero polarization).

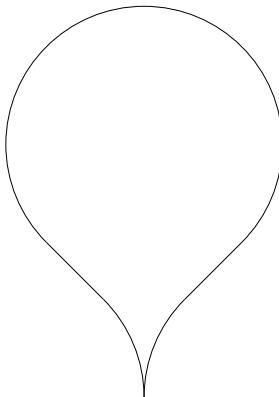


Figure 1: Sets of teardrop shaped magnet arcs each with 1.8 Tesla iron dipoles are used at each end of the 2 GeV Linac.

# 50 on 50-GeV Muon Collider Storage Ring

C. Johnstone\* and A. Garren†

\*FNAL,<sup>1</sup>

Batavia, IL 60555, USA

†UCLA, Los Angeles, CA 90024, USA

**Abstract.** Two modes are being considered for a 50 on 50-GeV muon collider: one being a high-luminosity ring with broad momentum acceptance ( $dp/p$  of  $\sim 0.12\%$ , rms) and the other lower luminosity with narrow momentum acceptance ( $dp/p$  of  $\sim 0.003\%$ , rms), or Higgs Factory. To reach the design luminosities, the value of beta at collision in the two rings must be 4 cm and 14 cm, respectively. In addition, the bunch length must be held comparable to the value of the collision beta to avoid luminosity dilution due to the hour-glass effect. To assist the rf system in preventing the bunch from spreading in time, the constraint of isochronicity is also imposed on the lattice. Finally, the circumference must be kept as small as possible to minimize luminosity degradation due to muon decay. Two lattice designs will be presented which meet all of these conditions. Furthermore, the high-luminosity and Higgs Factory lattice designs have been successfully merged into one physical ring with mutual components; the only difference being a short chicane required to match dispersion and floor coordinates from one lattice into the other.

## INTRODUCTION

After one  $\mu^+$  bunch and one  $\mu^-$  bunch have been accelerated to collision energy, the two bunches are injected into the collider ring, which is a fixed-field storage ring. Two cases are being considered for a 50 on 50-GeV collider: a ring with broad momentum acceptance ( $dp/p_{rms}$  of  $\pm 1.2\%$ ) and high luminosity, and one with a much narrower momentum acceptance ( $dp/p_{rms}$  of  $\pm 0.003\%$ ) and lower luminosity. The narrow-band machine is intended to resolve the width of the Higgs mass to high precision.

The two operational modes for the 100-GeV collider require different machine optics. The following sections discuss collider lattices for both the broad momentum application and the monochromatic mode.

---

<sup>1)</sup> Work supported by the U.S. Department of Energy under contract No. DE-AC02-76HO3000.

## DESIGN CRITERIA

Stringent criteria have been imposed on the collider lattice designs in order to attain the specified luminosities. The first and most difficult criterion to satisfy is provision of an Interaction Region (IR) with extremely low  $\beta^*$  values at the collision point consistant with acceptable dynamic aperture. The required  $\beta^*$  values for the 100-GeV collider are 4 cm for the broad momentum-width case and 14 cm for the narrow-width case. These  $\beta^*$  values were tailored to match the longitudinal bunch lengths in order to avoid luminosity dilution from the hour-glass effect. Final-focus designs must also provide collimators and background sweep dipoles, and other provisions for protecting the magnets and detectors from muon-decay electrons. Effective schemes have been incorporated into the current lattices.

Another difficult constraint imposed on the lattice is that of isochronicity. A high degree of isochronicity is required in order to maintain the short bunch structure without excessive rf voltage. A final criterion especially important in the lower-energy colliders, is that the ring circumference be as small as feasible in order to minimize luminosity degradation through decay of the muons. To achieve small circumference requires high fields in the bending magnets as well as a compact, high dipole packing-fraction design. (To meet the small circumference demand, 8 T poletip fields have been assumed for all superconducting magnets.)

Some of these criteria conflict with one another. For example, the small value of  $\beta^*$  leads to large peak beta values in the final-focus quadrupoles and correspondingly large linear chromaticities in the IR. For the high-luminosity machine, local correction of the linear part of the IR chromaticity is required to achieve adequate momentum acceptance. Efficient chromatic correction in turn requires large positive values for dispersion in the correction sextupoles. Because of the short circumference condition, high dipole packing fractions must be maintained not only in the arcs, but in the local Chromatic Correction Section (CC) as well. One consequence of the high dipole concentration in the CC is that a small momentum compaction becomes difficult to maintain because of the large number of dipoles in regions of high positive dispersion, in conflict with the need for isochronicity. Control over the momentum compaction is achieved through appropriate design of the arcs. The following sections discuss a base ring design which approaches the limit of compactness for a 50-GeV collider lattice under isochronous conditions and with strong local chromatic correction.

## OVERVIEW

For the 100 GeV CoM collider, two operating modes are contemplated: a high-luminosity case with broad momentum acceptance to accommodate a beam with a  $\delta p/p$  of  $\pm 0.12\%$  (rms), and one with a much narrower momentum acceptance and lower luminosity for a beam with  $\delta p/p$  of  $\pm 0.003\%$  (rms). For the broad momentum acceptance case,  $\beta^*$  must be 4 cm and for the narrow momentum acceptance case,

14 cm. In either case, the bunch length must be held comparable to the value of  $\beta^*$ .

The 100-GeV CoM ring has a roughly racetrack design with two circular arcs separated by an experimental insertion on one side, and a utility insertion for injection, extraction, and beam scraping on the other. The experimental insertion includes the interaction region (IR) followed by a local chromatic correction section and a matching section. The chromatic correction section is optimized to correct the ring's linear chromaticity, which is almost completely generated by the low beta quadrupoles in the IR. In designs of  $e^+e^-$  colliders, it has been found that local chromatic correction of the final focus is essential [1], as was found to be the case here.

Two 100 GeV lattice designs have been made; these are described below. The design has two optics modes: one mode has a  $\beta^*$  value of 4 cm with small transverse and large momentum acceptance; a second mode has a  $\beta^*$  value of 14 cm with large transverse and small, approximately monochromatic, momentum acceptance. Both lattices were merged into one physical, highly compact ring design with a total circumference of only about 345 m. The arc modules account for only about a quarter of the ring circumference.

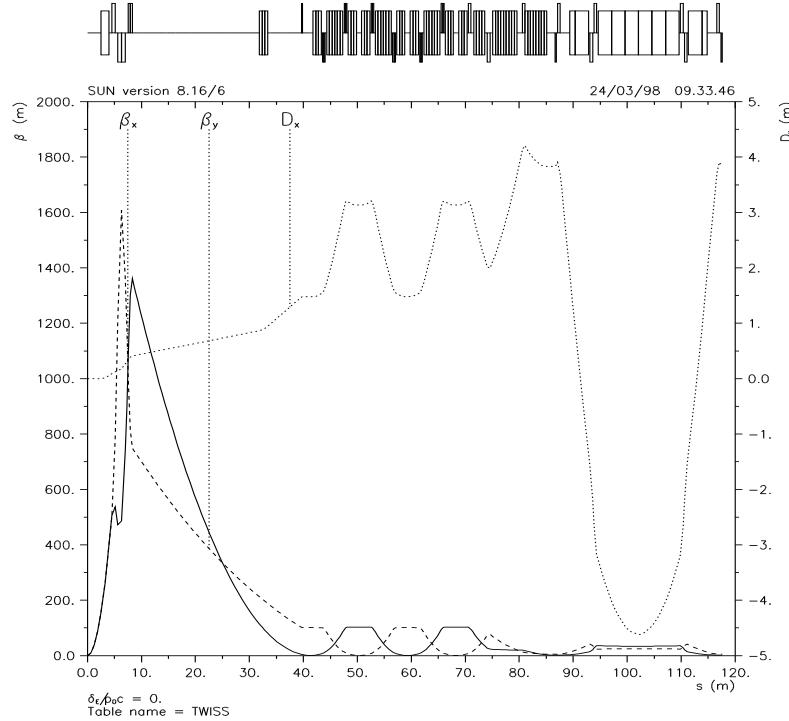
## The Interaction Region

Because of the dynamics of the cooling process,  $\mu^+$  and  $\mu^-$  emerge from the cooling stage with roughly equal emittances. Initially unequal  $\beta^*$ 's, or elliptical beams, were explored at the collision point. From an optics standpoint, elliptical beams are more manageable and less nonlinear than round beams in the design of Interaction Regions. Using a  $\beta^*$  ratio of 1:4 for the horizontal to vertical (factor of 2 in the relative beam sizes), however, causes a decrease in the luminosity of a factor of 2 and this was felt to be unacceptable. Therefore, the condition of round beams at the Interaction Point (IP) has been imposed in all current collider designs.

The need for different collision modes in the 100-GeV machine led to an Interaction Region design with two optics modes: one with broad momentum acceptance ( $dp/p$  of 0.12%, rms) and a collision  $\beta$  of 4 cm (Fig. 1), and the other basically monochromatic ( $dp/p$  of 0.003%, rms) and a larger collision  $\beta$  of 14 cm (Fig. 2). The low beta function values at the IP are mainly produced by three strong superconducting quadrupoles in the Final Focus Telescope (FFT) with pole-tip fields of 8 T. Because of significant, large-angle backgrounds from muon decay, a background-sweep dipole is included in the final-focus telescope and placed near the IP to protect the detector and the low- $\beta$  quadrupoles [2]. It was found that this sweep dipole, 2.5 m long with an 8 T field, provides sufficient background suppression. The first quadrupole is located 5 m away from the interaction point, and the beta functions reach a maximum value of 1.5 km in the final focus telescope, when the maxima of the beta functions in both planes are equalized. For this maximum beta value, the quadrupole apertures must be at least 11 cm in radius to accommodate

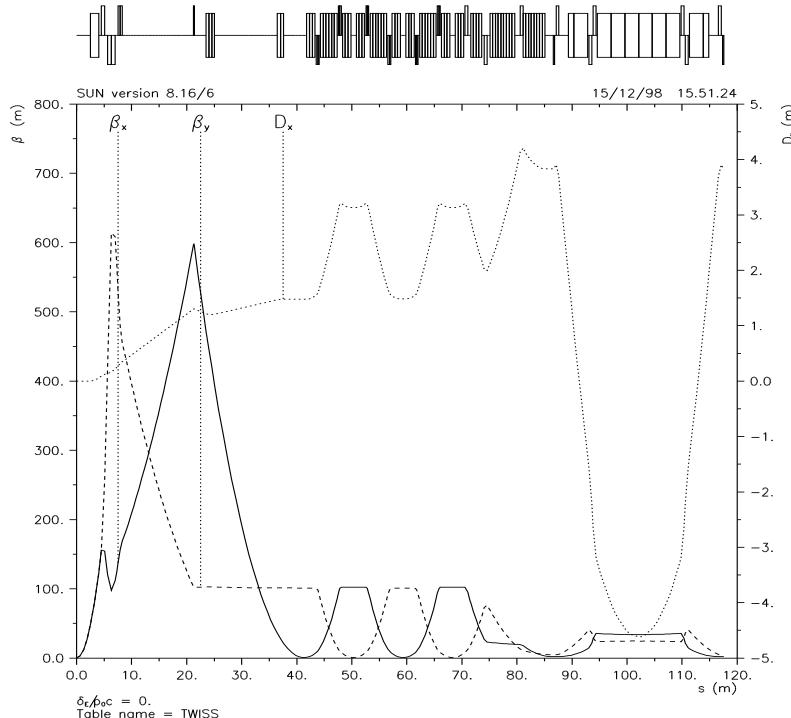
$5\sigma$  of a  $90 \pi$  mm mrad, 50-GeV muon beam (normalized rms emittance) plus a 2 to 3 cm thick tungsten liner [3]. The natural chromaticity of this interaction region is about  $-60$ .

The proximity of the final-focus quadrupoles to the IP determines the maximum beta and this value combined with the quadrupole strengths and lengths determine the natural chromaticity and, ultimately, the nonlinear behavior of the lattice. With pole tip fields reaching 8T, the final-focus triplet in the 100-GeV collider remains short: quadrupole lengths range from .6 to 1.5 m. With such short quadrupoles, the peak beam size in the 100-GeV machine and, therefore, the natural chromaticity of its interaction region is almost completely a property of the IP to quadrupole spacing.



**FIGURE 1.** 4 cm  $\beta^*$  Mode showing half of the IR, local, chromatic correction, and one of three arc modules.

The optimum design of a very low-beta IR is to make the imaging as point to parallel as is practical to soften chromatic aberrations. The less the applied chromatic correction, the larger, in general, is the dynamic aperture. In the 100-GeV machine, circumference constraints require the IP to be imaged in a short distance; implying stronger than optimal focussing from the high-beta triplet. The IP image distance can be reduced by as much as 35 meters on either side of the IP; or about a 30% decrease in the ring circumference. The stronger quadrupole strengths do increase the linear chromaticity of the IR from about 60 to 85 in the vertical with little effect on the horizontal (assuming the triplet powering is



**FIGURE 2.** 14 cm  $\beta^*$  Mode showing half of the IR, local, chromatic correction, and one of three arc modules.

FDF). In practice, the demagnification is about halfway between a compact and an optimal, or soft-focussing IR. Some deterioration in dynamic aperture is evident with stronger focussing, although studies of high-order and phase dependencies are underway and careful tuning appears to meliorate these effects.

Initially, the powering of the triplet was chosen such that the vertical apertures in the near dipoles were minimized. This requires placing the vertical high-beta peak at the center of the triplet, so that the triplet sequence is FDF. This has the disadvantage in that the local chromatic correction is not as efficient (the higher the dispersion, the more efficient the correction). Higher values of dispersion are usually obtained at peaks in the horizontal beta function than in vertical beta peaks. The plane corrected first should be the one with the highest chromaticity; in this case the vertical. If the dispersion is lower, then the chromatic correction, even with  $\pi$  pairs of sextupoles, is not as efficient and generates stronger nonlinearities. These nonlinearities propagate and appear to be enhanced by the sextupoles of the opposite plane and can be correlated to an observed decrease in dynamic aperture in this plane.

In a test lattice, the triplet was powered in a DFD configuration out of concern for the dynamic aperture. The plane with the highest chromaticity and the highest achievable dispersion at the sextupoles was corrected closest to the source, effecting a more efficient chromatic correction. Nonlinear terms were amplified less

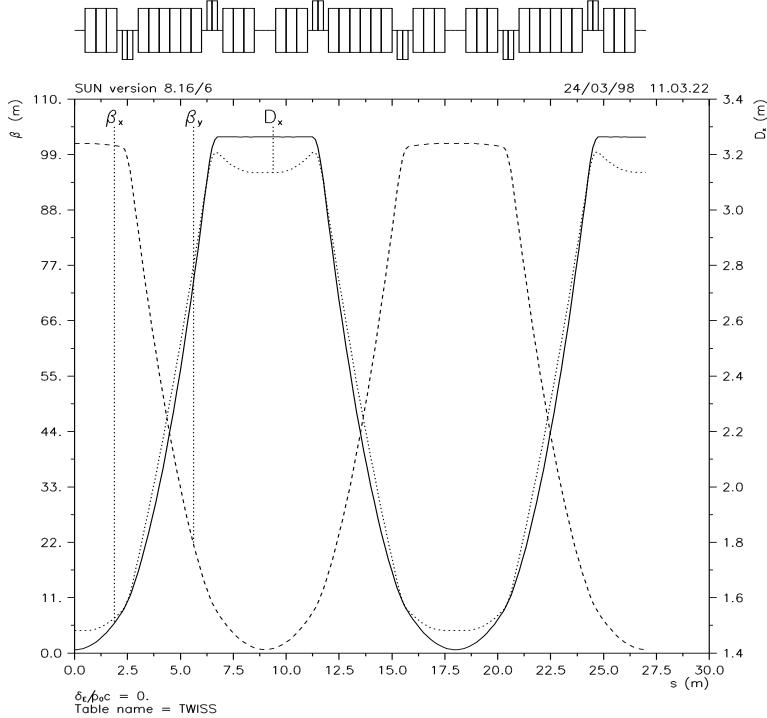
by sextupoles in the opposite plane as was evidenced by a slight improvement in dynamic aperture. A questionable consequence of installing the horizontal high-dispersion peak nearest the IP was the unavoidable application of reverse bends to create a dispersion plateau ( $D'=0$ ) after a defocussing quadrupole. (These reverse bends are not needed if vertical chromaticity correction is performed first because a dispersion plateau can follow a focussing quadrupole.) The net increase in circumference due to reverse bends and less efficient dipole packing in general brought the circumference up by at least 50 m; making the circumference more than 400 m when injection and scraping are included. The loss in muon lifetime was felt to outweigh the small advantage to the optics of the ring. The final triplet powering remains as FDF with the vertical chromaticity beign corrected closest to the IP.

## Chromatic Correction

Local chromatic correction of the muon collider interaction region is required to achieve broad momentum acceptance. With such a large aperture in the final-focus quadrupoles, adding dispersion to the final focus is not reasonable and therefore chromatic correction must take place in a specialized section. The basic approach developed by Brown [1] and others is implemented in the Chromatic Correction Region (CC) used here. The CC contains two pairs of sextupoles, one pair for each transverse plane, all located at locations with high dispersion. The sextupoles of each pair are located at positions of equal, high beta value in the plane (horizontal or vertical) whose chromaticity is to be corrected, and low beta value in the other plane. Moreover, the two sextupoles of each pair are separated by a betatron phase advance of  $\pi$ , and each sextupole has a phase separation of  $(2n + 1)\frac{\pi}{2}$  from the IP, where  $n$  is an integer. The result of this arrangement is that the geometric aberrations of each sextupole is canceled by its companion while the chromaticity corrections add.

An innovative module was developed specifically for chromatic correction (Fig. 3) and implemented first in the 4-TeV muon collider [4]. Its characteristics include a high-dispersion and high-beta plateau in one plane coincident with a deep minimum in beta in the opposite plane. The high-beta plateaus alternate between planes, with the single intervening deep minimum establishing a  $\pi$  phase advance between plateaus in the same plane. The sextupoles of each pair are centered about a minimum in the opposite plane ( $\beta_{min} < 1$ ), which provides chromatic correction with minimal cross correlation between the planes. A further advantage to locating the opposite plane's minimum at the center of the sextupole, is that this point is  $\frac{\pi}{2}$  away from, or “out of phase” with, the source of chromatic effects in the final focus quadrupoles; i.e. the plane not being chromatically corrected is treated like the IP in terms of phase to eliminate a second order chromatic aberration generated by an “opposite-plane” sextupole.

In this lattice example, the CC was optimized to be as short as possible with a smooth transition designed to place the first chromatic correction sextupole at



**FIGURE 3.** The Chromatic Correction Module.

the same phase as the high-beta point [4]. The  $\beta_{\max}$  is only 100 m and the  $\beta_{\min} = 0.7$  m, giving a  $\beta_{\text{ratio}}$  between planes of about 150, so the dynamic aperture is not compromised by a large amplitude-dependent tuneshift.

This large beta ratio, combined with the opposite-plane phasing, allows the sextupoles for the opposite planes to be interleaved, without significantly increasing the nonlinearity of the lattice. In fact, interleaving improved lattice performance compared to that of a non-interleaved correction scheme, due to a shortening of the chromatic correction section, which lowers its chromaticity contribution. The use of somewhat shallower beta-minima with less variation in beta through the sextupoles were also applied to soften the chromatic aberrations, although this caused a slight violation of the exact  $\pi$  phase advance separation between sextupole partners. The retention of an exact  $\pi$  phase advance difference between sextupoles was found to be less important to the dynamic aperture than elimination of minima with  $\beta_{\min} < 0.5$  m.

This module, specifically optimized to perform chromatic correction, is particularly powerful in that it can accommodate long sextupoles without beta and phase changes taking place in the plane being corrected. However, because of finite element lengths and changes in the phase advance between sextupole as a function of energy, a tuneshift with amplitude is unavoidable, and depends most sensitively on the beta amplitude in the sextupole, but also on the length of the sextupole and

the tune of the ring. Ultimately, a tuneshift with amplitude constricts the dynamic aperture and a tradeoff exists between momentum acceptance and transverse dynamic aperture. Lattice parameters, especially the beta values at the sextupoles and the phase advance around the ring, must be carefully tuned to optimize both acceptances simultaneously.

For the narrow band acceptance, local chromatic correction; i.e. the sextupoles are turned off. The momentum acceptance narrows to about a  $\delta p/p$  of about  $\pm 0.2\%$ , while the transverse dynamic aperture increases rapidly to over  $10\sigma$  at the central momentum.

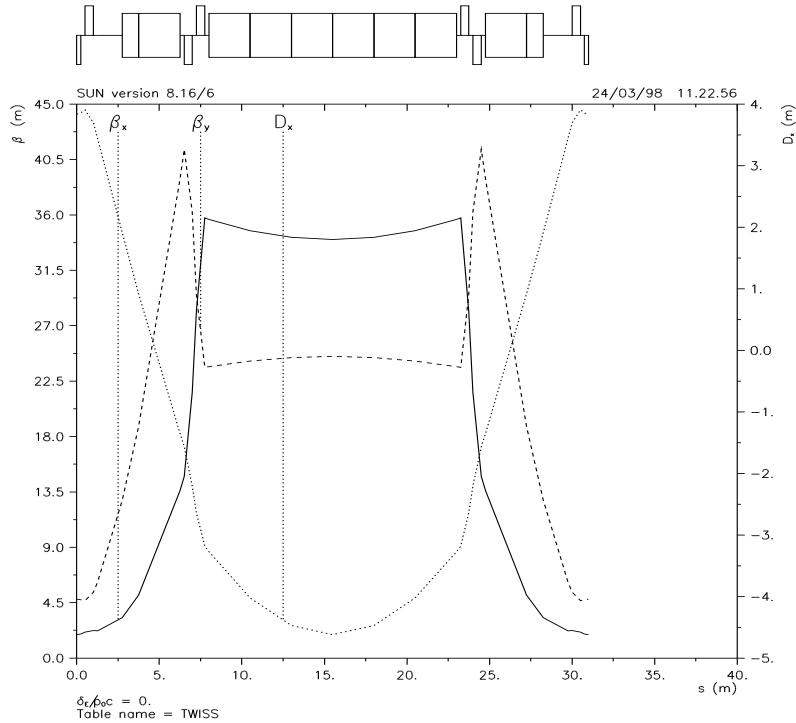
## The Arc

The arc module is shown in Fig. 4. It has the small beta functions characteristic of FODO cells, yet a large, almost separate, variability in the momentum compaction of the module which is a characteristic associated with the flexible momentum compaction module [5]. The small beta functions are achieved through the use of a doublet focusing structure which produces a low beta simultaneously in both planes. At the dual minima, a strong focusing quadrupole is placed to control the derivative of dispersion with little impact on the beta functions. (The center defocusing quadrupole is used only to clip the point of highest dispersion.) Ultimately a dispersion derivative can be generated which is negative enough to drive the dispersion negative through the doublet and the intervening waist. Negative values of momentum compaction as low as  $\alpha = -0.13$  have been achieved, and  $\gamma_t = 2 i$ , has been achieved with modest values of the beta function.

The entire ring was designed to control momentum compaction, even in the match section with connects the CC to the arc. This careful attention to momentum compaction for the isochronicity condition resulted in a circumference which is just under 350 m, as opposed to rings which were greater than the 400 m characteristic of earlier designs. The total momentum compaction contributions of the IR, CC, and matching sections is about 0.04. The total length of these parts is 173 m, while that of the the momentum-compaction-correcting arc is 93 m. From these numbers, it follows that this arc must and does have a negative momentum compaction of about  $-0.09$  in order to offset the positive contributions from the rest of the ring.

## A rf system

The rf requirements depend on the momentum compaction of the lattice and on the parameters of the muon bunch. For the case of very low momentum spread, synchrotron motion is negligible and the rf system is used solely to correct an energy spread generated through the impedance of the machine. For the cases of higher momentum spreads, there are two approaches. One is to make the momentum compaction zero to high order through lattice design. Then the synchrotron motion can be eliminated, and the rf is again only needed to compensate the induced energy



**FIGURE 4.** The 100-GeV CoM collider arc: a new flexible momentum compaction module.

spread correction. Alternatively, if some momentum compaction is retained, then a more powerful rf system is needed to maintain the specified short bunches. In either case, rf quadrupoles will be required to generate BNS damping of the transverse head-tail instability.

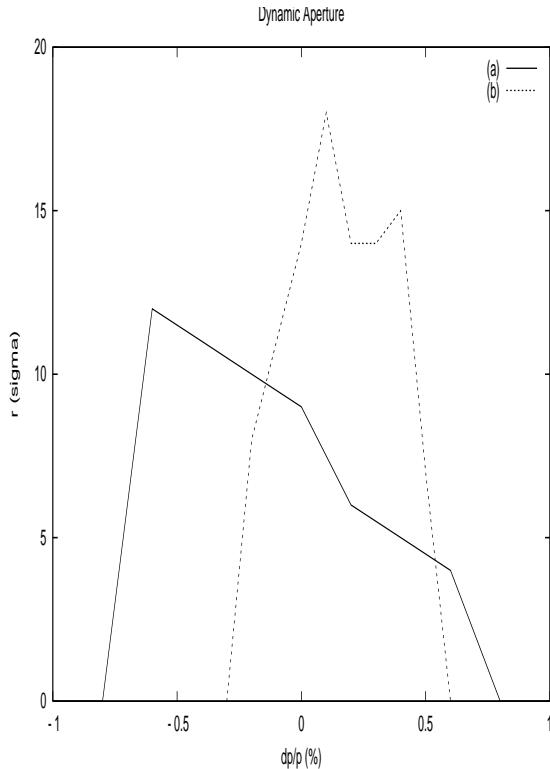
## PERFORMANCE

A very preliminary calculation of the dynamic aperture without optimization of the lattice nor inclusion of errors and end effects is given in Fig. 4. One would expect that simply turning off the chromatic correction sextupoles in the 4 cm  $\beta^*$  mode would result in a linear lattice with a large transverse aperture. With only linear elements, the 4 cm  $\beta^*$  optics showed to be strongly nonlinear with limited on-momentum dynamic acceptance.

A normal form analysis using COSY INFINITY showed that the tune-shift-with-amplitude was large, which was the source of the strong nonlinearity in the seemingly linear lattice. To locate the source of this nonlinearity, a lattice consisting of the original IR and arcs only (no CC), was studied. Numerical studies showed that a similar dynamic aperture and tune-shift-with-amplitude terms. This ruled out the possibility that the dynamic aperture was limited by the low beta points in the local chromatic correction section and points to the IR as the source of

the nonlinearity. (The findings were verified by S. Ohnuma who used a Runge-Kutta integrator to track through the IR and a linear matrix for the rest of the lattice.) Further analytical study using perturbation theory showed that the first-order contribution to the tune shift with amplitude is proportional to  $\gamma_{x,y}^2$  and  $\gamma_x \gamma_y$ , which are large in this IR. These terms come from the nonlinear terms of  $p_x/p_0$  and  $p_y/p_0$ , which, to the first order, equal the angular divergence of a particle. As a demonstration, a comparison to the LHC low-beta IR was done. Taking into account only the drift from the IP to the first quadrupole, the horizontal detuning at  $10\sigma$  of the present IR ( $\beta^* = 4$  cm) is 0.01, whereas the detuning of the entire LHC lattice is below 1e-4. This also explains the fact that the on-momentum aperture of the wide momentum spread mode remains roughly constant among different versions despite various correction attempts.

It was therefore concluded and later shown that the dynamic aperture of the more relaxed  $\beta^*$  of 14 cm would not have the same strong nonlinearities due to the reduced angular terms. In fact, the tune shift with amplitude was less by an order of magnitude; hence the large transverse acceptance shown in Fig. 5 (dashed line).



**FIGURE 5.** A preliminary dynamic aperture for the 4 cm  $\beta^*$  mode where  $\sigma$  (rms) =  $82\mu\text{m}$  (solid line) and the 14 cm  $\beta^*$  mode where  $\sigma$  (rms) =  $281\mu\text{m}$  (dashed line).

## REFERENCES

1. K. Brown, "A Conceptual Design of Final Focus Systems for Linear Colliders," PUB-4159, (1987).
2. C. Johnstone and N. Mokhov, FERMILAB-Conf-96-366(1996).
3. A. Drozhdin, et al, AIP Conf. Proc. **441**, p. 242 (1998).
4. C. Johnstone and A. Garren, "An IR and Chromatic Correction Design for a 2-TeV Muon Collider," *Proceedings of the Workshop on New Directions for High-Energy Physics, Snowmass 96, Jun.-Jul., 1996*, pp. 222.
5. K. Ng, S. Lee, and D. Trbojevic, FNAL-FN595 (Fermilab, 1992).

Table 1: Baseline parameters for Higgs factory muon collider. Higgs/year assumes a cross section of  $5 \times 10^4$  fb, Higgs width of 2.7 MeV, 1 year =  $10^7$  s. (From “Status of Muon Collider Research and Development and Future Plans,” Muon Collider Collaboration, Charles M. Ankenbrandt *et al.*, *Phys. Rev. ST Accel. Beams* **2**, 081001 (1999).)

COM energy (TeV)	0.1		
$p$ energy (GeV)	16		
$p$ 's/bunch	$5 \times 10^{13}$		
Bunches/fill	2		
Rep. rate (Hz)	15		
$p$ power (MW)	4		
$\mu$ /bunch	$4 \times 10^{12}$		
$\mu$ power (MW)	1		
Wall power (MW)	81		
Collider circum. (m)	350		
Ave bending field (T)	3		
rms $\delta p/p$ (%)	0.12	0.01	0.003
6D $\epsilon_{6,N}$ ( $\pi m$ ) <sup>3</sup>	$1.7 \times 10^{-10}$	$1.7 \times 10^{-10}$	$1.7 \times 10^{-10}$
rms $\epsilon_n$ ( $\pi$ mm mrad)	85	195	290
$\beta^*$ (cm)	4.1	9.4	14.1
$\sigma_z$ (cm)	4.1	9.4	14.1
$\sigma_r$ spot ( $\mu m$ )	86	196	294
$\sigma_\theta$ IP (mrad)	2.1	2.1	2.1
Tune shift	0.051	0.022	0.015
$n_{\text{turns}}$ (effective)	450	450	450
Luminosity ( $\text{cm}^{-2} \text{ s}^{-1}$ )	$1.2 \times 10^{32}$	$2.2 \times 10^{31}$	$10^{31}$
Higgs/yr	$1.9 \times 10^3$	$4 \times 10^3$	$3.9 \times 10^3$

# **Cost Drivers for the Conversion of the Neutrino Factory to the Higgs Factory**

M. Zisman

We consider here an entry level Muon Collider to operate as a Higgs Factory. No cost estimate has yet been prepared for this stage, so we mention here only the obvious “cost drivers.” First, the initial muon beam must be prepared as a single bunch of each charge. This may involve an additional ring for the proton driver to coalesce proton bunches into a single pulse. The cooling will have to be significantly augmented. First, a much lower transverse emittance is needed, and second, it will be necessary to provide emittance exchange to maintain a reasonable transmission of the muons. The additional cooling will permit going to smaller solenoids and higher frequency rf systems (402.5 or perhaps 805 MHz), which should lower the incremental cost somewhat. Next, we will need considerably more acceleration, though with smaller energy acceptance and aperture requirements than at present. Lastly, we will need a very low  $\beta^*$  lattice for the storage ring, along with mitigation of the potentially copious background levels near the interaction point. In this case the detector is, in effect, part of the Collider and cannot be ignored in terms of its cost impact.

Of the items mentioned, it is likely that the additional cooling and the additional acceleration are the most significant cost drivers. Future work will define the system requirements better and permit a cost estimate of the same type provided for Studies-I and -II.

# Bunch Stacking Scheme for the Higgs Factory based on the Neutrino Factory Design

Yasuo Fukui, Alpher Garren, David Cline, Ping He (UCLA)  
March 15, 2001

## Abstract

A scheme for a Higgs Factory is designed to generate a high intensity  $\mu^+$  bunch and  $\mu^-$  bunch in order to create the high luminosity in a collider ring. Three modification need to be done to the Neutrino Factory Feasibility Study I(Fermilab) and II(BNL), (1) combine multiple primary proton bunches into one, (2) add a 6 dimensional ionization cooling ring, and (3) add two rings to stack muon mini-bunches transversely.

## 1 Introduction

The Higgs Factory will be the first  $\mu^+\mu^-$  collider to be built. Compared to a Neutrino Factory where muons decay in a straight section of a storage ring, the Higgs Factory requires more 6 dimensional phase space cooling of the muon beam by a factor of  $10^{4-5}$ . The Feasibility Studies I(Fermilab) and II(BNL) showed that those designs are feasible to build a neutrino Factory [3, 4].

In order to create the high luminosity in a collider, single high intensity  $\mu^+/\mu^-$  beam bunch will be transported into a collider where the many muon mini-bunch trains in the Neutrino Factory has to be combined into a single bunch.

The  $\mu^+\mu^-$  collider has been under designing for several year, and many reports have been made by the Neutrino Factory and the Muon Collider Collaboration. [1, 2] The on-going BNL targetry experiment E951 will reveal the feasibility of the proposed mercury jet target with the input proton beam power of multi MW. Several R&D efforts on the components of the muon ionization cooling, 201 MHz and 805 MHz RF cavities, super conducting solenoid coils, and induction linac modules. Design of a cooling demonstration experiment is going on at Fermilab with the model of a cooling channel in a ring.

## 2 Goals

Table 1 shows the parameters of the general  $\mu^+\mu^-$  colliders including a Higgs Factory which was given in Reference [2].

6 dimensional normalized emittance is  $1.7 \times 10^{-10} \pi \text{mm-mrad}$  and the RMS beam length is 9.4 cm, RMS  $\Delta p/p$  is 0.01 % in a Higgs Factory with the center of mass energy of 100 GeV. The proton beam power is 4 MW, and the luminosity is  $2.2 \times 10^{31} \text{cm}^{-2}\text{s}^{-1}$ .

### 3 Steps from a Neutrino Factory to a Higgs factory

#### 3.1 Proton Bunches

The requirement of the proton bunches is that a single proton bunch with the maximum available intensity hits target(s) within the  $\gamma$  times the life time of the muons so that single  $\mu^+/\mu^-$  bunch can be collided in a collider ring. Reducing the number of proton bunches in the proton driver will reduce the effort to combine the muon long bunches later on. In case of the Neutrino Factory, the number of primary proton bunches was not restricted, because the muons can be injected into the decay storage ring any time, and the only issue was the total number of the decays of the muons in the straight section.

The high power proton driver design is given in References [1, 2]. The primary proton power is estimated to be around 4 MW. Using the liquid metal target is the probable option to handle the severe energy deposition inside the target.

#### 3.2 Ring Cooler

In order to add 6 dimensional phase space cooling to that in the neutrino Factory design, it is proposed to use a 6 dimensional ionization cooling ring which consists of cooling lattices, bending magnets, and wedge absorbers. The ionization cooling lattice is made of super conducting solenoid, RF cavities, and liquid Hydrogen absorber, The wedge absorbers are placed in high dispersion area so that  $\Delta p/p$  in the longitudinal phase space is exchanged into the transverse phase space. The circumference of the ring is around 35 m. Depending on the initial normalized emittance of the muons, the cooling ring is expected to generate the 6 dimensional phase space cooling factor of 30-100. [5]

Figure 1 shows a schematic diagram of the components of the Higgs Factory. The additional cooling ring is placed after the phase rotation channel and bunching section and before the linear cooling channel.

#### 3.3 Bunch Stacking Rings

Stacking muon mini-bunches coming out of the 201 MHz RF on top of each other naturally do the emittance exchange from the longitudinal phase space to the transverse phase space, besides the fact that getting a single muon bunch is the absolute must in the Higgs Factory.

Emittance exchange scheme by using transverse bunch stacking was first proposed by C. H. Kim [6]. The idea was then tested in a simulation by using a solenoid as a delay channel with pulse-by-pulse transverse bending which change the path length of beam bunches in the solenoid. Figure 2 shows a schematic diagram of the transverse bunch stacking in solenoid delay channels and in a bunch stacking ring. Figure 3 shows transverse(left plots) and longitudinal(right plots) phase space distribution before and after getting through the bunch stacking solenoid channel which was simulated by using ICOOL simulation code with a set of beam bunch parameters. In

the Neutrino factory design, around 100 muon mini-bunches are created through the bunching channel and the ionization cooling channel with 201 MHz RF cavities.

In using one of the transverse bunch stacking model, a numerical simulation was performed with a set of muon beam parameters. Figure 3 shows the transverse and longitudinal phase space of 10 muon mini-bunches before and after going through bunch by bunch delay channels inside a straight solenoid channel. Muon mini-bunches are simulated to have one  $2\pi$  phase Larmor turn by using the ICOOL simulation code. [7] The transverse phase space, in position and angle, was enlarged and the longitudinal phase space in time was made smaller in this simulation.

We propose the transverse bunch stacking scheme by using an 1 GeV 320 m ring with Lithium lens for the transverse phase space cooling, which contain all the 100 muon mini-bunches. by using a Fast Pulsed kicker magnet/EM filed, we then inject each mini bunches into an 1 GeV 35 m smaller ring with Lithium lens element for the transverse cooling to stack the mini-bunches top on top each other so that we can get a single muon bunch. Figure 4 top plot shows the  $\beta_x$ ,  $\beta_y$ ,  $\eta$ , and the lattice component in the 320 m storage ring where a muon long bunch is contained with the 201 MHz RF structure. Figure 4 bottom plot shows the  $\beta_x$ ,  $\beta_y$ ,  $\eta$ , and the lattice component in the 35 m storage ring where a muon single bunch is generated which then transferred in to the subsequent accelerator section. SYNCH simulation code was used to design the synchrotron rings with 320 m and 35 m circumferences and to analyze beam orbits in those rings. [9] Although the exact beam injection/extraction have not been computer-simulated, injecting muon mini-bunches from the 320 m ring into the 35 m ring is a challenging task. [8]

Table 1: Baseline parameters for high- and low-energy muon colliders. Higgs/year assumes a cross section  $\sigma = 5 \times 10^4$  fb; a Higgs width  $\Gamma = 2.7$  MeV; 1 year =  $10^7$  s.

CoM energy	TeV	3	0.4	0.1		
$p$ energy	GeV	16	16	16		
$p$ 's/bunch		$2.5 \times 10^{13}$	$2.5 \times 10^{13}$		$5 \times 10^{13}$	
Bunches/fill		4	4	2		
Rep. rate	Hz	15	15	15		
$p$ power	MW	4	4	4		
$\mu$ /bunch		$2 \times 10^{12}$	$2 \times 10^{12}$		$4 \times 10^{12}$	
$\mu$ power	MW	28	4	1		
Wall power	MW	204	120	81		
Collider circum.	m	6000	1000	350		
Ave bending field	T	5.2	4.7	3		
Rms $\Delta p/p$	%	0.16	0.14	0.12	0.01	0.003
6-D $\epsilon_{6,N}$	$(\pi m)^3$	$1.7 \times 10^{-10}$				
Rms $\epsilon_n$	$\pi$ mm-mrad	50	50	85	195	290
$\beta^*$	cm	0.3	2.6	4.1	9.4	14.1
$\sigma_z$	cm	0.3	2.6	4.1	9.4	14.1
$\sigma_r$ spot	$\mu$ m	3.2	26	86	196	294
$\sigma_\theta$ IP	mrad	1.1	1.0	2.1	2.1	2.1
Tune shift		0.044	0.044	0.051	0.022	0.015
$n_{\text{turns}}$ (effective)		785	700	450	450	450
Luminosity	$\text{cm}^{-2}\text{s}^{-1}$	$7 \times 10^{34}$	$10^{33}$	$1.2 \times 10^{32}$	$2.2 \times 10^{31}$	$10^{31}$
Higgs/year				$1.9 \times 10^3$	$4 \times 10^3$	$3.9 \times 10^3$

## 4 Conclusion

A scheme for a Higgs Factory has been designed by applying three major upgrades in the Neutrino Factory designs in the Feasibility studies I and II. The first is to use a single high intensity primary proton beams which are generated in a 4 MW proton driver. The second is to add the 6 dimensional ionization cooling ring in between the phase rotation/bunching channel and the straight ionization cooling channel. The third is the transverse muon mini-bunch stacking by using a 320m and 35 m synchrotron rings with Lithium lens for the transverse ionization cooling. A high intensity muon bunch will be generated there which then is accelerated and injected into a Higgs Collider storage ring. We need more detailed numerical simulation of the beam from the target to the collision point in the Higgs collider, where the goals are to obtain the high luminosity in the Higgs Factory.

## Acknowledgments

The authors wish to thank A Sessler, M. Zisman, J. Wurtele, W. Fawley, G. Penn, D. Neuffer, and V. Balbekov, for the useful discussions, R. B. Palmer, H. Kirk, R. Fernow and J. Gallardo, and , for their help and encouragement.

## References

- [1] R. B. Palmer et al.,  
“ $\mu^+\mu^-$  Collider: a Feasibility Study”  
Snowmass96 workshop proceedings, BNL-52503, Fermilab-Conf.-96.092, LBNL-38946
- [2] C. M. Ankenbrandt et al.,  
“Status of muon collider research and development and future plans”  
Phys. Rev. ST Accel. Beams 2, 081001
- [3] T. Anderson et al.,  
“A Feasibility Study of a Neutrino Source Based on a Muon Storage Ring”  
Fermilab report, March, 2000
- [4] B. Palmer et al.,  
“A BNL Feasibility Study II Report”  
to be submitted, May, 2001
- [5] V.Balbekov  
“Ring Cooler Update”  
Muon Collider Note 190, February, 2001
- [6] C. H. Kim  
“An Emittance Exchange Idea Using Transverse Bunch Stacking” Muon Collider Note 70,  
December, 1999
- [7] R. Fernow  
“A Simulation Code for Ionization Cooling of Muon Beams”  
submitted to PAC99

- [8] G. Schroder  
“Fast Pulsed Magnet Systems”  
CERN-SL-98-017(BT)
- [9] A. Garren et al.,  
“Synch(A Computer System for Synchrotron Design and Orbit Analysis) User’s Guide”  
SSSL-MAN-0030, 1993

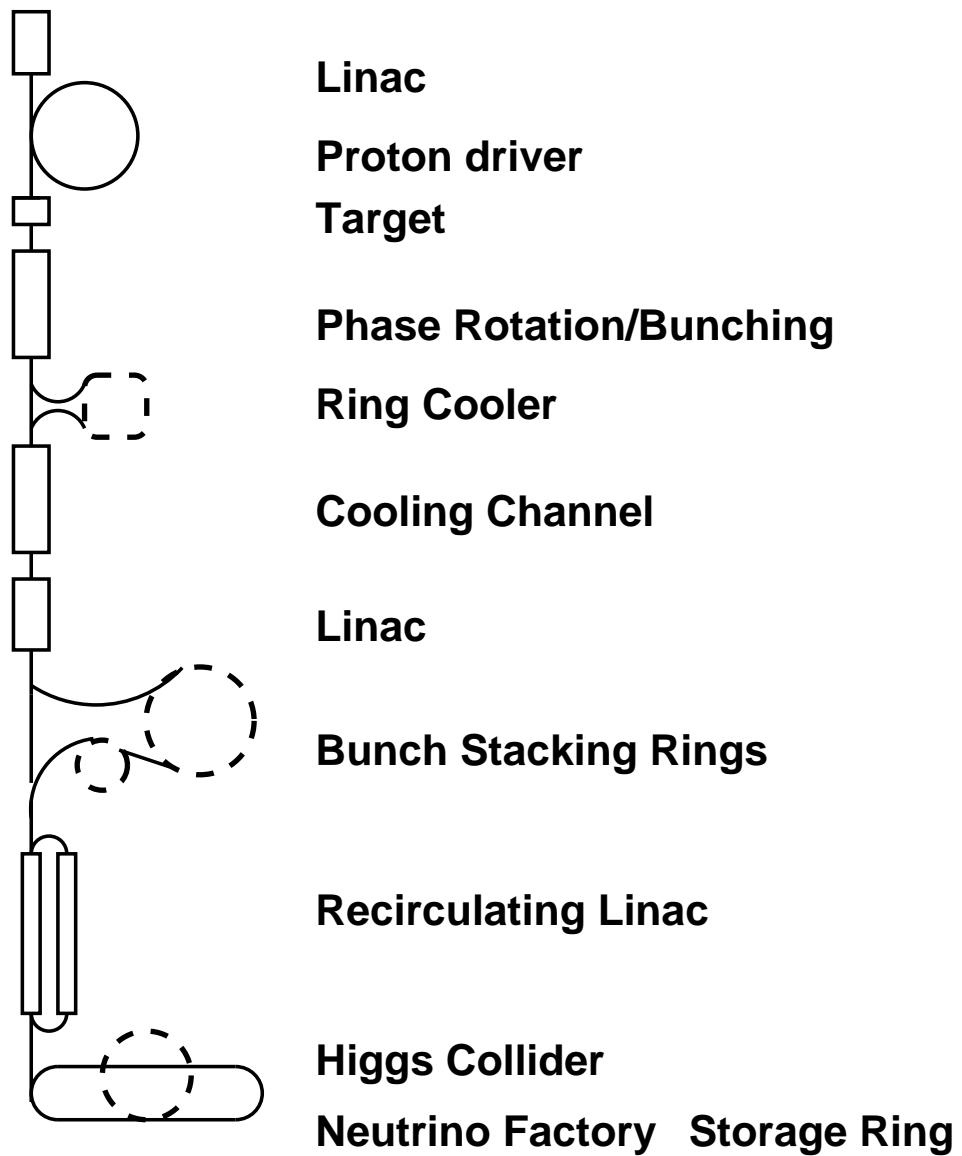


Figure 1: Schematic Diagram of the Higgs Factory and the Neutrino Factory

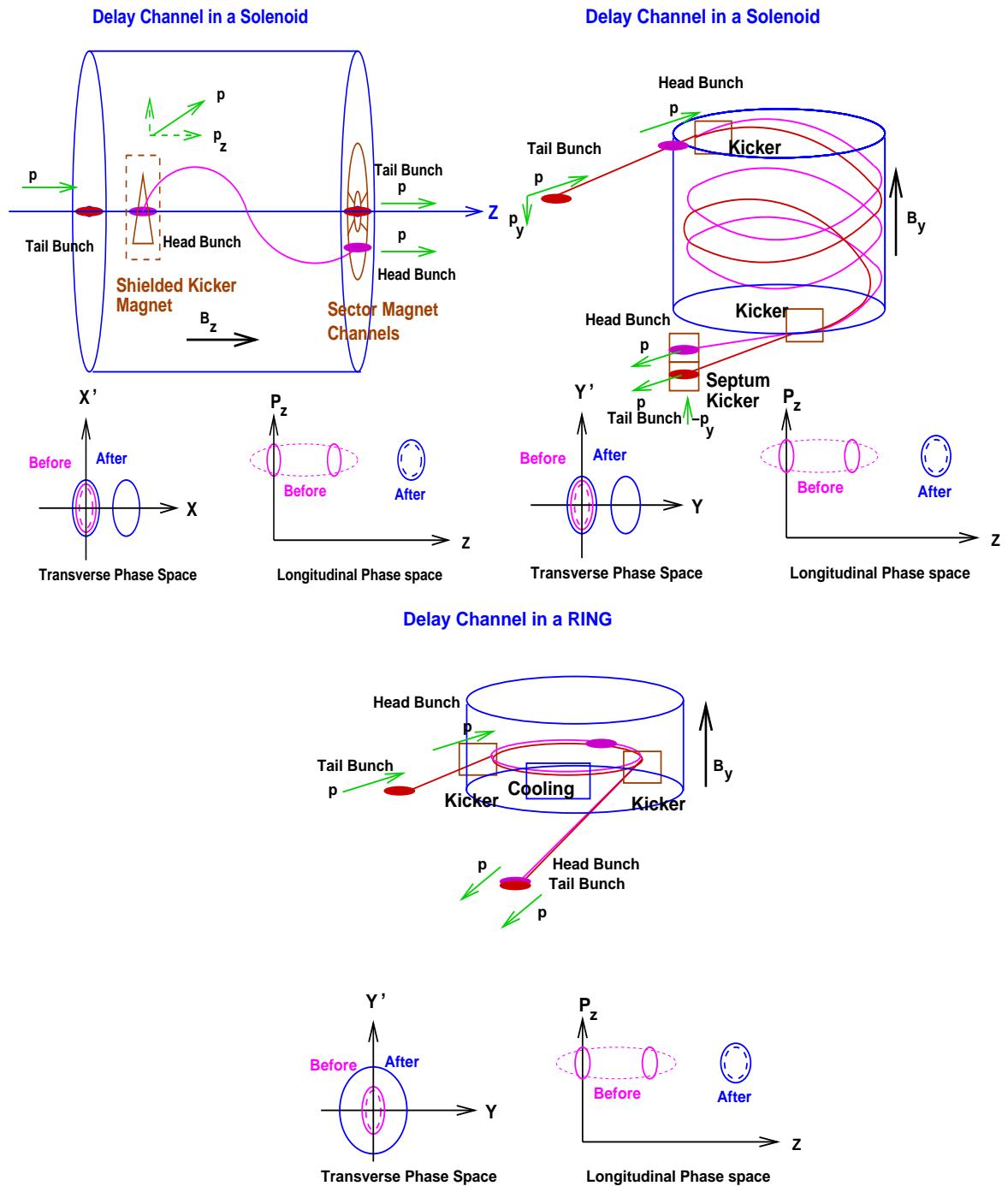


Figure 2: Schematic Diagram of the transverse bunch stacking in solenoid delay channels and in a bunch stacking ring

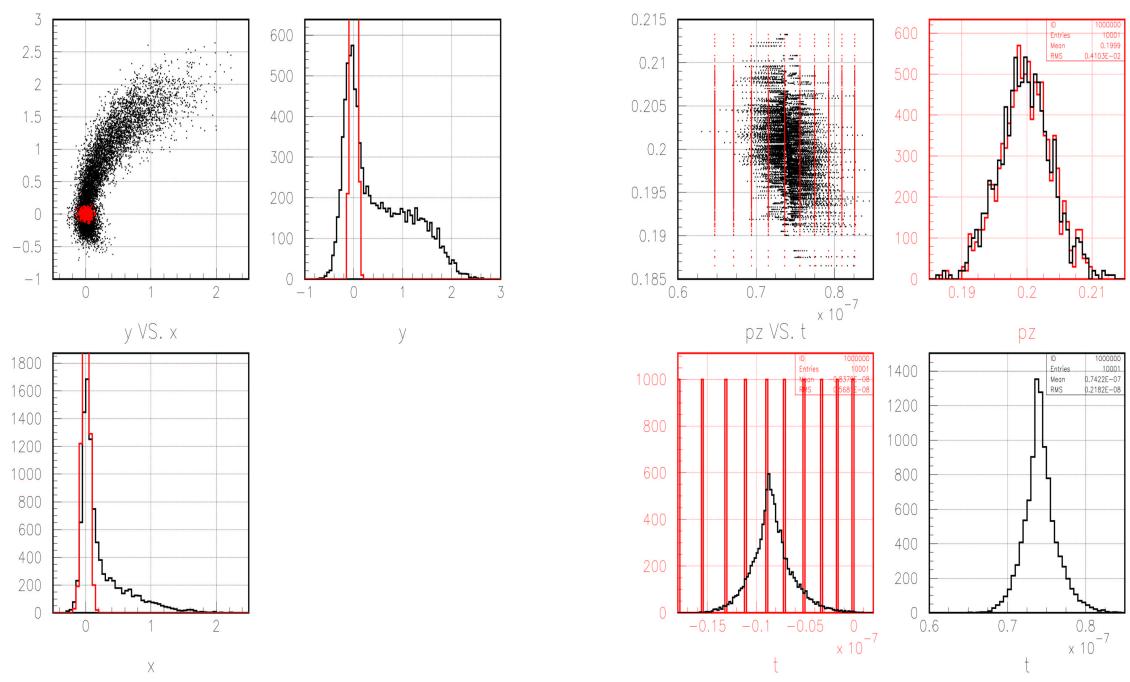


Figure 3: Transverse and longitudinal phase space (ICOOL simulation) on 10 bunch stacking

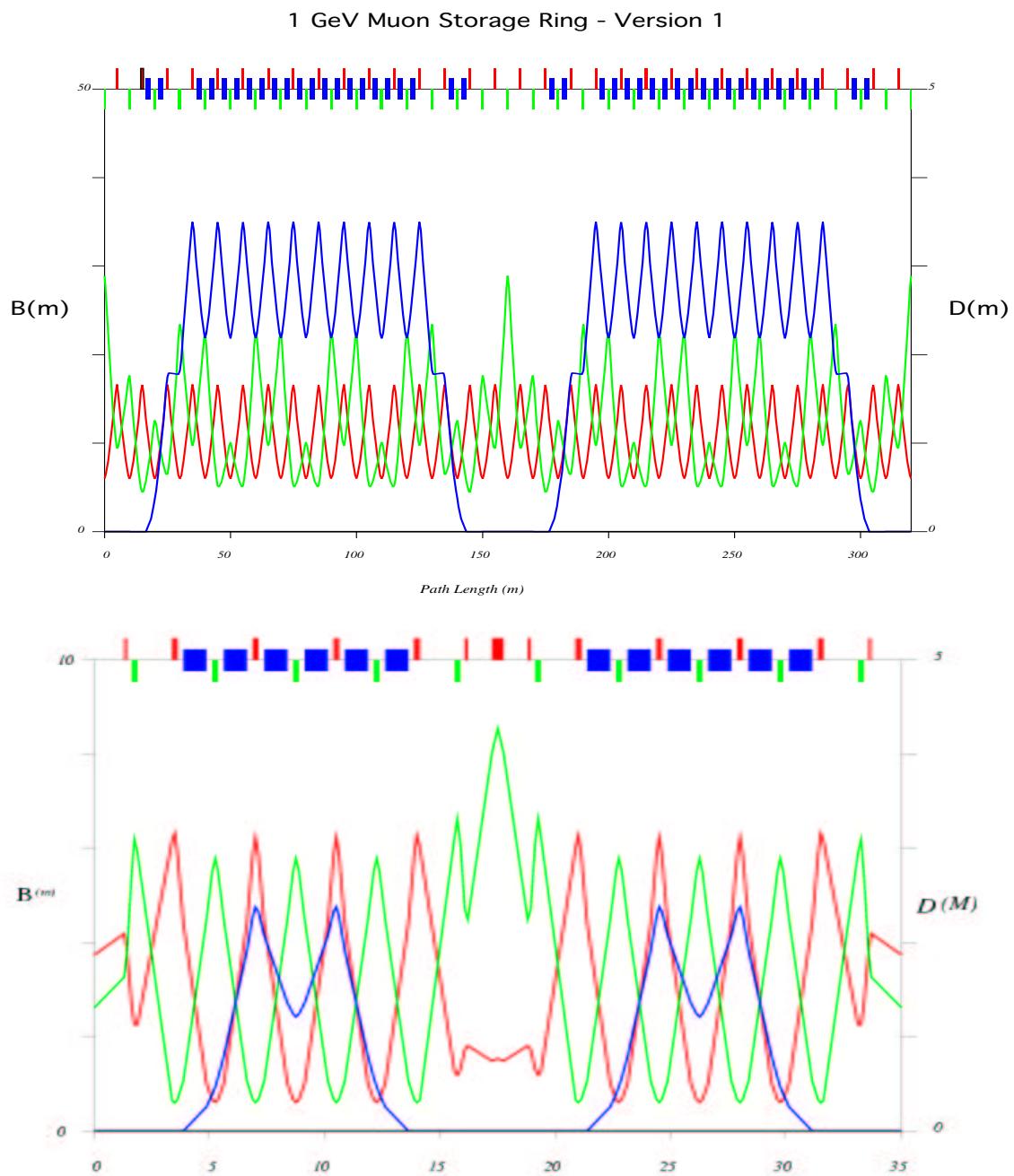


Figure 4:  $\beta_x$ ,  $\beta_y$ ,  $\eta$ , and the lattice component in 320 m ring (top) and in the 35 m ring (bottom)

# Muon Bunch Compressor Based on a Low RF Ring Cooler

V. Balbekov, FNAL, Batavia, IL 60510, USA

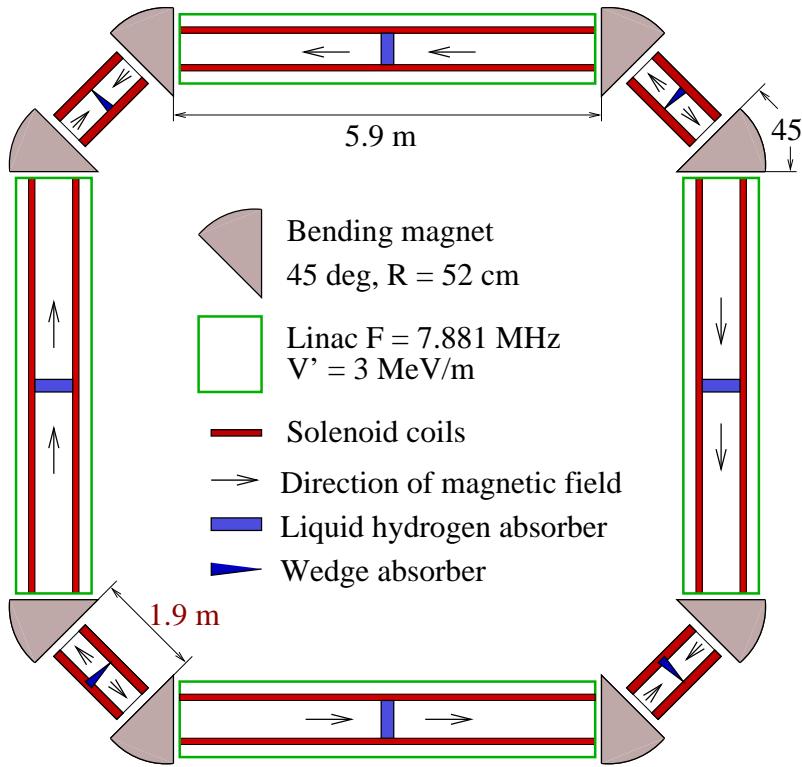
## I INTRODUCTION

One of the most serious problems concerning the construction of a Higgs factory (or any other muon collider) is the creation of very short high intensity muon bunches [1]. A precooling part including pion producing target, decay channel, and low RF phase rotation system, can provide a muon bunch of 6-10 m length in the best case [2], whereas a 200 MHz cooling channel requires at least 10-15 times shorter in bunch length. A strong emittance exchange combined with the beam cooling appears to be the most reasonable method to this end. A ring cooler proposed in Ref. [3]– [4] is considered in this paper. From a technical standpoint, a low frequency and high gradient accelerating system is the critical part in this scheme. Assumed is an 8 MHz and 3 MeV/m RF system that provides capture in the bucket of 10 m length, acceleration, and—with appropriate absorbers—a reasonable cooling rate / emittance exchange at modest beam loss caused by muon decays. Longitudinal cooling factor 6-40 is achievable in this system, depending on used approximations. Transverse nonlinearity of bending magnets, as well as dependence of the revolution frequency on transverse momentum, are the most serious causes of the degradation. These and other effects are investigated including some measures to improve the cooler performance.

## II DESCRIPTION OF THE COOLER

The ring cooler described in the paper [4] is taken as a basis for the designed bunch compressor. Schematic and parameters of the compressor are given in Fig. 1 and Table 1. It consists of  $8 \times 45^\circ$  dipoles, 4 short straight sections (SS), and 4 long SS. The long SS containing RF cavities and liquid hydrogen absorbers are intended for transverse cooling of muons. Wedge absorbers for an emittance exchange are placed in the short SS where there is large dispersion. Thus, there are 4 periods, each including the bending part and the straight section.

Layout of the bending part is displayed in Fig. 2. It consists of 2 bending magnets and 2 solenoids with opposite direction fields. Besides forming a circular orbit, this part provides transverse focussing and dispersion required for the emittance

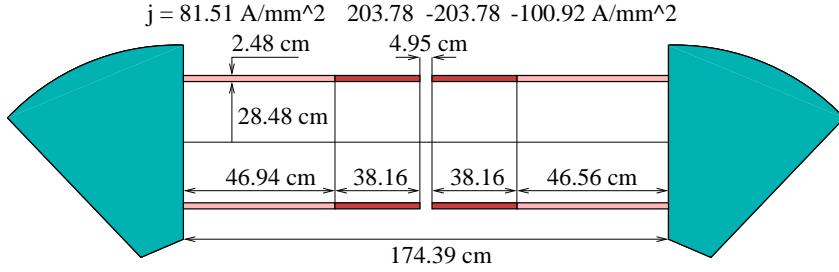


**FIGURE 1.** Schematic of the ring cooler

**TABLE 1.** Parameters of the ring cooler

1.	Circumference	34.475
2.	Nominal energy (total)	250 MeV
3.	Number of bending magnets	8
4.	Bending angle	45°
5.	Bending radius	52 cm
6.	Bending field	1.453 T
7.	Normalized field gradient	0.5
8.	Length of short SS	1.902 m
9.	Length of long SS	5.900 m
10.	Axial field of long solenoid	2.055 T
11.	$\beta$ -function at nominal energy	0.735 m
12.	Revolution/RF frequency	7.881 MHz
13.	RF harmonic number	1
15.	Accelerating gradient	3 MeV/m
16.	Synchronous phase	30°
17.	Main absorber	LH, 30 cm
18.	Wedge absorber	$dE/dy = 0.2 \text{ MeV/cm}$

exchange. Therefore, gradient magnets with a normalized field gradient of 0.5 are used here, providing beta-function of  $R\sqrt{2} \simeq 73.5$  cm in both directions. To get the same beta-function at the dipole, the field at the solenoid is then constrained to  $|B_{\text{solenoid}}| = B_{\text{dipole}}\sqrt{2} \simeq 2.05$  T. Of course, for the alternate solenoid used in this section  $|B_{\text{solenoid}}|$  cannot be constant. However, real magnetic field plotted in Fig. 3 has the same matrix as an “ideal” solenoid with  $|B| = \text{const}$  and instantaneous field flip, i.e. it provides a perfect matching at least for equilibrium particles. Note that the bending magnets are considered as magnetic mirrors at the calculation of



**FIGURE 2.** Layout of the short straight section

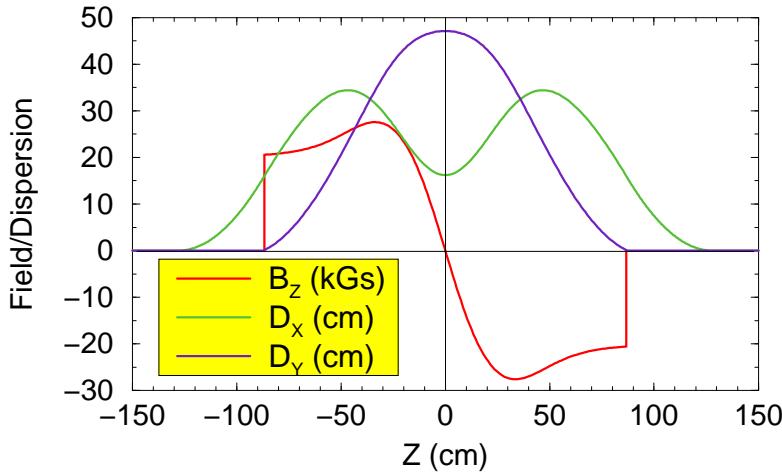
the solenoid field. Transverse field of the bending magnets is introduced analytically including nonlinear part to satisfy Maxwell equations. Calculation of 3D magnetic field for both solenoids and bending magnets has to be performed as a next step.

Two-component dispersion function  $D_{x,y}$  of this section is also plotted in Fig. 3. With an appropriate choice of the solenoid length, the field is localized only in bending part, providing a dispersion-free long SS. The function  $D_x$  is the same in all short SS, while  $D_y$  changes in sign in any subsequent section. Vertical wedge absorbers are placed in the center of the bending sections for emittance exchange. Direction of the wedges also changes in any subsequent section. Material of the wedge absorbers is LiH, and the gradient of ionization energy loss is 0.2 MeV/cm corresponding to the wedge angle of  $6.8^\circ$ .

The long SS are designed very schematically because an external view and construction of low frequency accelerating system are unclear at present. It is assumed that the acceleration is performed with a 7.881 MHz traveling wave having a gradient of 3 MeV/m. There are two linacs of 280 cm length in each long SS, and a liquid hydrogen absorber of 30 cm length is placed at the center of SS providing an energy loss of muons 8.95 MeV for transverse cooling. A homogeneous solenoidal field of 2.055 T strength is used for transverse focusing. It produces the same  $\beta = 73.5$  cm in all sections including absorber areas. This is acceptable because transverse cooling is not a design goal but longitudinal cooling is achieved through emittance exchange, and the transverse emittance remains almost constant at a rather large value.

### III SIMULATION

Cooling of a single bunch with Gaussian distribution was studied for first for estimation of the cooler performance. Initial parameters of the beam w.r.t. the center of the long SS are given in the Table 2. The simulation was performed with varying values of the parameters to estimate the role of all the factors and to lay down a way to improve the system. 10000 muons was used for all simulations.



**FIGURE 3.** Magnetic field and dispersion at the short straight section

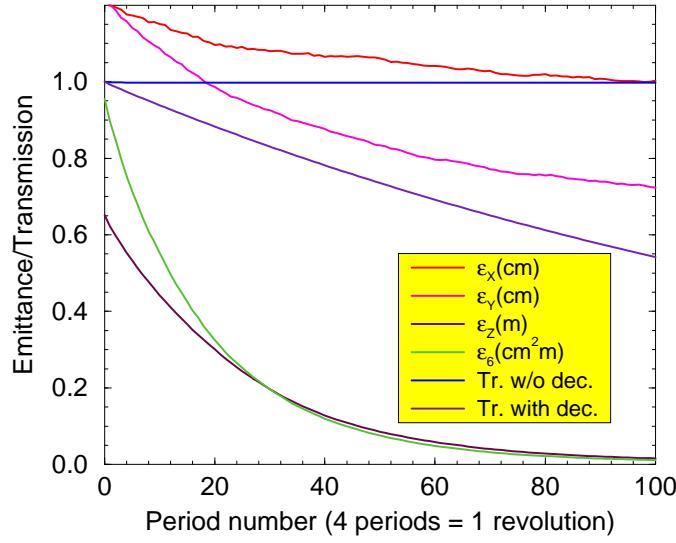
**TABLE 2.** R.m.s. sizes of the injected beam

Horizontal size (cm)	6.40
Vertical size (cm)	6.40
Longitudinal size (cm)	200.
Horizontal momentum (MeV/c)	20.0
Vertical momentum (MeV/c)	20.0
Energy spread (MeV)	35.0
Horizontal emittance (cm)	1.21
Vertical emittance (cm)	1.21
Longitudinal emittance (cm)	66.3
3D-emittance (cm <sup>3</sup> )	97.1

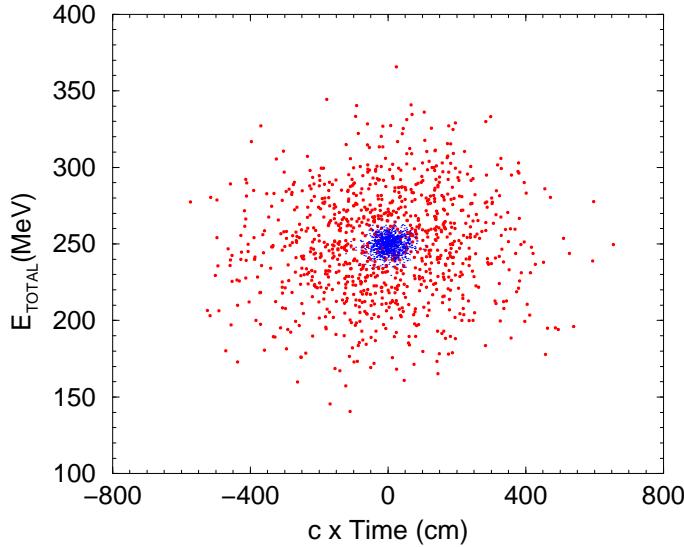
## A Linear approximation without chromaticity

Chromatic and nonlinear effects are ignored in this simulation except for the non-linearity of the accelerating field. (It is a conventional approximations for “normal” accelerators.) The results are presented in Fig. 4 and Fig. 5. Fig. 4 shows the evolution of the bunch emittance and transmission. By 25 turns, the longitudinal emittance decreases from 66.3 cm to 1.6 cm (cooling factor 41), and 6D emittance – from 97.1 cm<sup>3</sup> to 1.2 cm<sup>3</sup> (cooling factor 81). Transmission is 54%, and the only cause of the particle loss is muon decay.

Fig. 5 represents the longitudinal phase space at the injection (red points) and after 25 turns (blue spot at the center). After an appropriate matching the bunch is quite acceptable for a high RF cooler (f.e. 200 MHz).



**FIGURE 4.** Evolution of emittance and transmission at the cooling (linear approximation without chromaticity).



**FIGURE 5.** Longitudinal phase space before (red) and after the cooling (blue).

## B Linear approximation with chromaticity

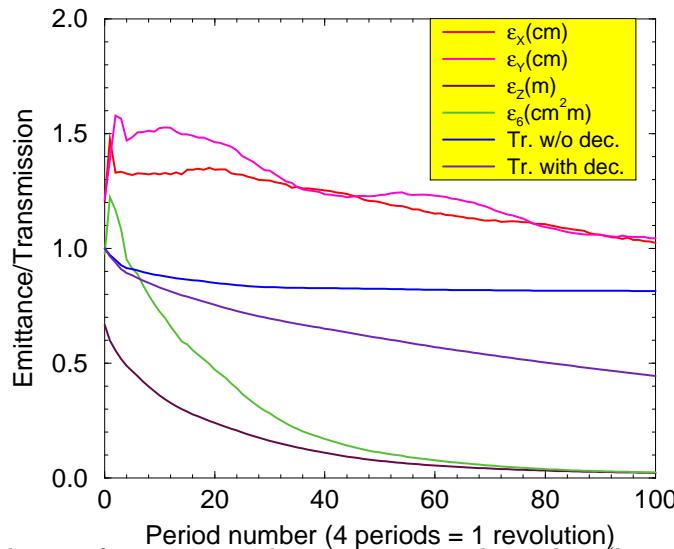
All chromatic effects are taken into account in this simulation and transverse motion is considered in paraxial approximation as before. The results presented in

Fig. 6 and Fig. 7 are somewhat worse than the previous ones. Now, the longitudinal r.m.s. emittance decreases after 25 turns from 66.3 cm to 2.3 cm (cooling factor 29), and 6D emittance – from 97.1 cm<sup>3</sup> to 2.4 cm<sup>3</sup> (cooling factor 40). Transmission is 44% with decay and 81% without decay. Thus, 19% of particles are lost in this approximation because of the machine imperfection.

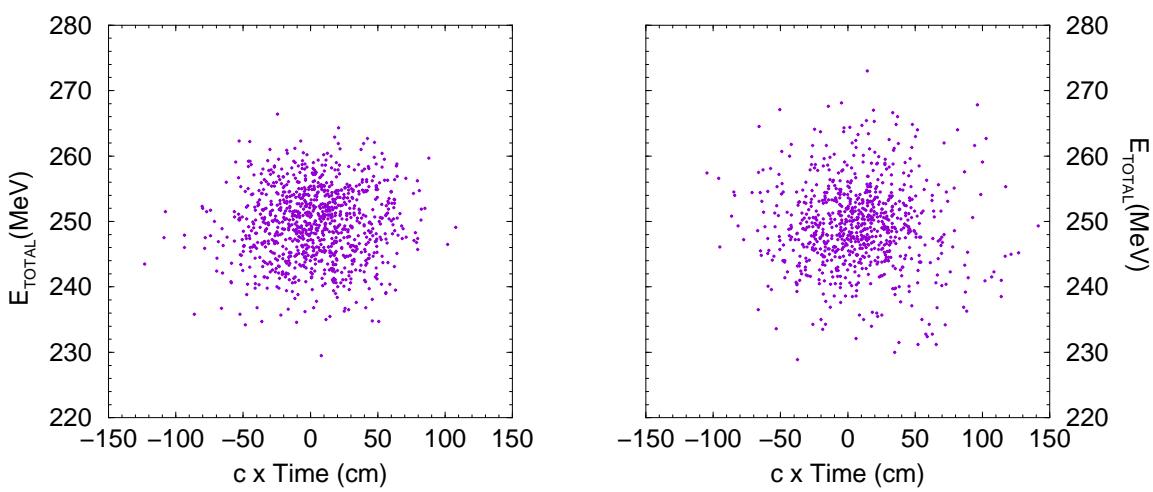
Longitudinal phase space after 25 turns is shown in the right in Fig. 7. For comparison, the same is plotted for previous case – without chromaticity (left). It is seen that the chromaticity causes an additional beam halo which can explain the increased emittance. Possibly, the distinction will be not so marked if an acceptance cut is applied. This problem has to be studied further specifically.

Probably, rather high particle loss in this case is an effect of linear betatron resonances at non-equilibrium energies. In support of this the dependence of  $\beta$ -function on total energy is plotted in Fig. 8. Note that in the used linear approximation  $\beta_x = \beta_y$  because the bending magnets produce the same focusing in both directions. Several resonances are seen in Fig. 8 with the most serious one at  $E = 236$  MeV. Additionally, resonance excitation of dispersion has to be taken into account. Actually, the long straight sections are dispersion-free only in linear approximation in  $\Delta p/p$  what would mean an achromatic system. The dispersion function with higher order corrections is non-zero in long SS which is shown in Fig. 9. Dependence of dispersion on energy at the center of the long SS is plotted in Fig. 8 demonstrating a resonance behavior also (another components of the dispersion are zero in this point because of symmetry of the system.)

One can suppose on this basis that a suppression both the resonances and the nonlinear dispersion is a way to improve the bunch compressor.



**FIGURE 6.** Evolution of emittance and transmission at the cooling (linear approximation with chromaticity).

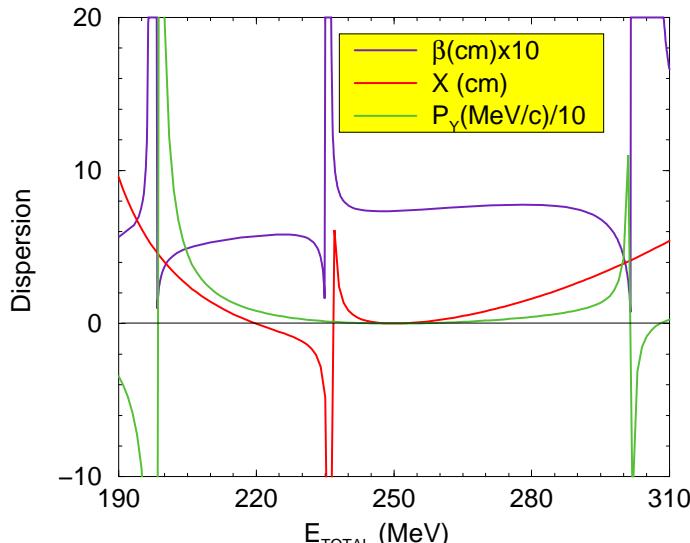


**FIGURE 7.** Longitudinal phase space after 25 turns. Comparison of the cooling without chromaticity (left) and with chromaticity (right).

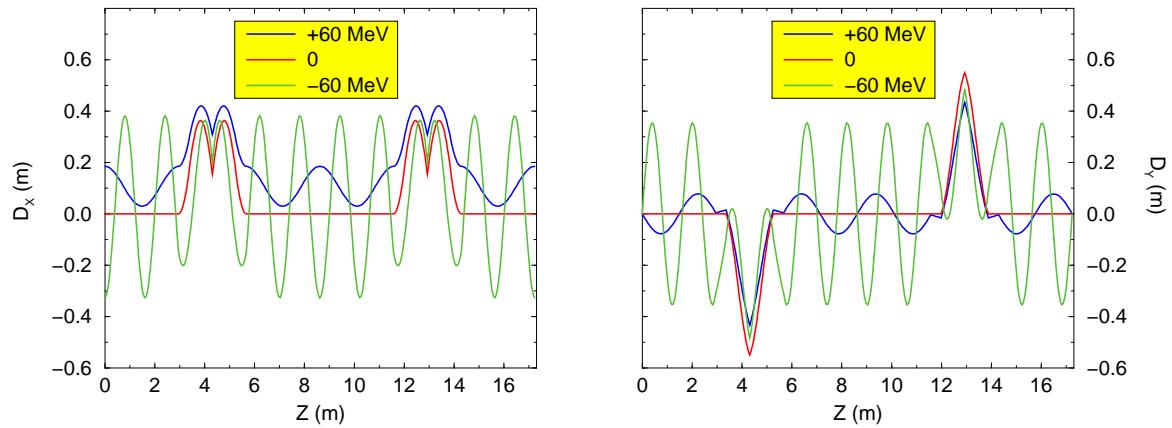
### C Transverse nonlinearity and chromaticity

This simulation is performed with all chromatic and nonlinear effects included *except for dependence of the revolution frequency on transverse momentum of a particle*. Beam parameters in dependence on number of turns are plotted in Fig. 10. The longitudinal cooling is almost the same as in the previous simulation: r.m.s. emittance decreases after 25 turns from 66.3 cm to 2.0 cm (cooling factor 33). However, an increase in transverse emittance by a factor of about 1.3 is observed now; as a result, 6D emittance changes from  $97.1 \text{ cm}^3$  to  $4.9 \text{ cm}^3$  (cooling factor 20). Transmission is also less in this case having only 30% with decay and 55% without decay.

Because the only new factor is the transverse nonlinearity, probably nonlinear betatron resonances are the cause for this degradation. An additional investigation



**FIGURE 8.** Dependence of  $\beta$ -function and dispersion on energy at the center of long SS.



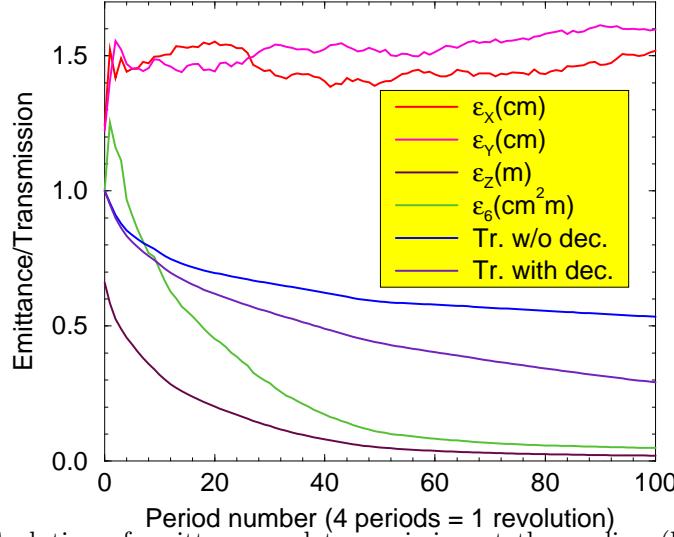
**FIGURE 9.** Dispersion function vs Z at various energy

has demonstrated that the bending magnets are mainly responsible for the resonances while the solenoid nonlinearity is almost negligible (in long SS, it is only kinematic effect because of dependence  $p_z(p_t)$  at given energy). Therefore, some investigation, and if possible, correction of the dipole nonlinearities are required.

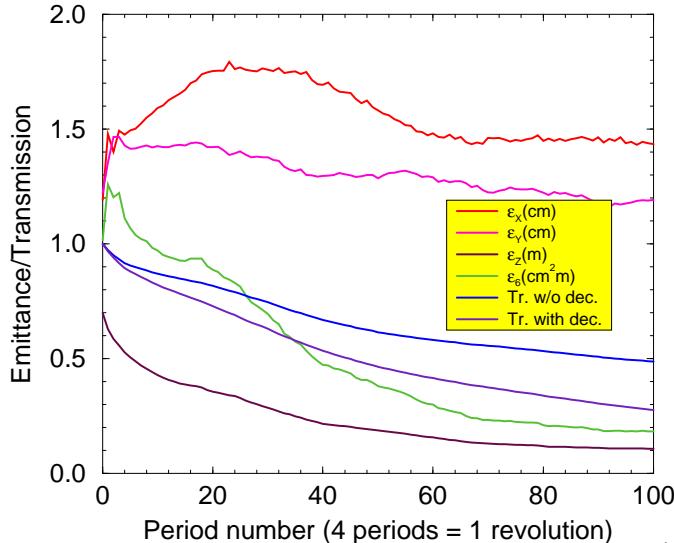
## D Full simulation

The latest effect that has to be taken into account is dependence of revolution frequency on transverse momentum at a given energy. Corresponding simulation is performed at the same conditions as before; however, after the generation of Gaussian distribution the following correlation is introduced:

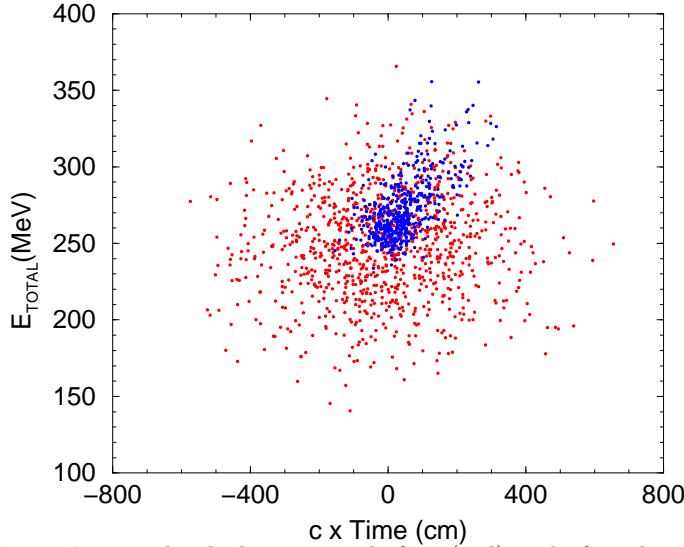
$$E = E_{ref} \sqrt{1 + \frac{1}{2} \left[ \left( \frac{p_t}{mc} \right)^2 + \left( \frac{eBr}{2c} \right)^2 \right]_{random}} + \Delta E_{random},$$



**FIGURE 10.** Evolution of emittance and transmission at the cooling (linear approximation with chromaticity).



**FIGURE 11.** Evolution of emittance and transmission at the cooling (linear approximation with chromaticity).



**FIGURE 12.** Longitudinal phase space before (red) and after the cooling (blue).

where  $E_{ref} = 250$  MeV. It is assumed that such a correlation should appear at a bunching at the ring which is not yet considered. The results presented in Fig. 11 and Fig. 12 are significantly worse than all the previous cases because of the deterioration of longitudinal characteristics. Now the longitudinal r.m.s. emittance decreases only to 10.8 cm and 6D emittance – to  $18.4 \text{ cm}^3$ . The cooling factor is about 6 for both cases because transverse emittance is finally almost the same as in the beginning. However, there is a considerable growth of the horizontal emittance in the first part of the cooling accompanied by more particle loss. Transmission after 25 turns is 28% with decay and 49% without decay.

Longitudinal phase space after 25 turns is shown in Fig. 12 by blue points. The main difference from Fig. 5 is a large energy spread caused by transverse momentum - energy correlation, which does not decrease at the cooling because the transverse

momentum is almost constant. Apparently, transverse cooling should be added to this system to improve the situation. Preliminary investigation have demonstrated that a simple decrease of the wedge angle by factor 2 provides a small transverse cooling and increases the total cooling factor to  $\sim 10$ . However, the considered machine is not designed for the transverse cooling because of large  $\beta$ -function at absorbers. Therefore, more radical changes are required including an adiabatic increase of magnetic field at song solenoids.

## IV CONCLUSION

It is shown that a low RF ring cooler is capable to decrease a longitudinal emittance of a single bunch from 60-70 cm to 2-3 cm which satisfies the requirements of the Higgs factory. There are several factors causing a degradation of the cooler performances, and the most serious of them is the dependence of the revolution frequency on the transverse momentum. An additional energy spread arising at bunching is not suppressed in the considered version because r.m.s. transverse momentum of the beam is about constant at the cooling. In principle, this spread is reversible, and it should vanish after a transverse cooling. However, a modification of the cooler is required including an adiabatic decrease of beta-function.

Another serious factor is nonlinearity of transverse motion in the bending magnets, and efforts have to be undertaken to weaken this. Chromatic effects are less dangerous, though a suppression of linear betatron resonances at non-equilibrium energies, as well as nonlinear in  $\Delta p/p$  dispersion, should be investigated more.

## REFERENCES

1. C. Ankenbrandt, M. Attac, et al., “Status of Muon Collider Research and Development and Future Plans”, Phys. Rev. ST Accel. Beams 2, 081881 (1999).
2. A. Van Ginneken and D. Neuffer, “Muon Collection Channel”, FERMILAB-Pub-98/296 (1998); A. Van Ginneken, “Targetry and Collection Problems”, McNote 0032 (1999).
3. V. I. Balbekov and A. Van Ginneken, “Ring Cooler for Muon Collider”, In the book “Physics Potential and Development of  $\mu^+\mu^-$  Colliders”, AIP Conf. Proc. 441, p.309 (1997); V.Balbekov, “ Possibility of Using a Ring Accelerator for Ionization Cooling of Muons”, PAC1999, V.1, p.315.
4. V.Balbekov, S. Geer, et al., “Muon Ring Cooler for the Mucool Experiment”, PAC2001.