Higgs Factory Front End

Higgs Factory Front End

In this chapter we will consider the beam cooling system for a "Higgs Factory" μ^+–μ^- 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 μ^+–μ^- Collider.

1. Introduction to μ-Cooling

Table 1 shows some suggested parameters [1.1] of Higgs-energy μ^+–μ^- colliders, operating in either high-luminosity or high-resolution (small δE) 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 ~4×10^{12} μ's within transverse emittances of $\varepsilon_{N,rms}$ ~ 10^{-4} m and longitudinal emittance $\varepsilon_{L,rms} \cong 10^{-2}$ 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 $\varepsilon_{N,rms}$ ~ 2×10^{-2} m and $\varepsilon_{L,rms} \cong 1$ 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 ~10^6, and the cooling must be completed before μ 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 ~400 × smaller. To obtain the small δE 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 ~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	
σ_x	10 cm	295 μm	86 μm	
$\sigma_{x'}$	70	2.1	2.1	mr
σ_z	680	14.1	4.1	cm
σ_p / p	0.65	3×10^{-5}	0.0012	
ε_{xN}	~20	0.29	0.085	mm

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

1.1 Cooling process and requirements

The cooling process that will be used is ionization cooling. In ionization cooling (μ-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]:

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

where the first term is the energy-loss cooling effect and the second is the multiple scattering heating term. Here ε_N is the normalized emittance, E is the beam energy, $\beta = v/c$ and γ are the usual kinematic factors, dE/ds is the energy loss rate, θ_{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 (~13.6 MeV).[6] (The normalized emittance is related to the geometric emittance $\varepsilon_\perp$ by $\varepsilon_N = \varepsilon_\perp/(\beta\gamma)$, and the beam size is given by $\sigma_x = (\varepsilon_\perp\beta_\perp)^{1/2}$.)

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:

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

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]:

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

where N_A is Avogadro's number, ρ, 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 MeV cm^2/gm). The ionization constant I is approximately 16 $Z^{0.9}$ eV, and δ is the density effect factor, which is small for low-energy μ's. The energy loss as a function of p_μ is shown in Fig. 2. The derivative is negative (or naturally heating) for E_μ < ~ 0.3 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]:

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

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 γ), 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:

$$\frac{\partial \frac{dE}{ds}}{\partial E} \Rightarrow \left.\frac{\partial \frac{dE}{ds}}{\partial E}\right|_0 + \frac{dE}{ds} \frac{\eta \rho'}{\beta c p \rho_0} = g_L \frac{\frac{dp}{ds}}{p} \tag{5}$$

where ρ'/ρ_0 is the relative change in density with respect to transverse position, ρ_0 is the reference density associated with dE/ds, and η is the dispersion (η = d x /d(Δp/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: $g_x \to (1-\eta\rho'/\rho_0)$, 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 ~2 at p_μ > 0.3 GeV/c. Fig. 4 shows this sum of partition numbers as a function of p_μ; small values of p_μ, where the sum becomes less than ~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_μ > 0.3 GeV/c . This means cooling at higher energy than some of our earlier studies, which are designed at $p_\mu \cong$ 0.2 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:

$$\frac{d\varepsilon_L}{ds} = -\frac{g_L \frac{dp}{ds}}{p} \varepsilon_L + \frac{\beta_L}{2} \frac{d\langle \Delta E_{rms}{}^2 \rangle}{ds} \tag{6}$$

where $\beta_L = \sigma_{ct}^2/\varepsilon_L$ 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:

$$\varepsilon_{N,eq} = \frac{\beta_\perp E_s^2}{2 g_x \beta m_\mu c^2 L_R \frac{dE}{ds}} \tag{7}$$

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_RdE/ds	Density	$g_x\beta\varepsilon_{N,eq}/\beta_\perp$
				MeV/cm	cm	MeV	gm/cm^3	mm-mrad/cm
Hydrogen	H_2	1	1.01	0.292	865	252.6	0.071	37
Lithium	Li	3	6.94	0.848	155	130.8	0.534	71
Lith. Hydride	LiH	3+1	7+1	1.34	102	137	0.9	68
Beryllium	Be	4	9.01	2.98	35.3	105.2	1.848	88
Carbon	C	6	12.01	4.032	18.8	75.8	2.265	122
Aluminum	Al	13	26.98	4.37	8.9	38.9	2.70	238
Copper	Cu	29	63.55	12.90	1.43	18.45	8.96	503
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 $\beta_\perp \cong 0.01$m, which means the transverse emittances can be cooled to $\varepsilon_{\perp,n} \cong 0.0001$ m-rad (in hydrogen or lithium). Similarly a cooling rf bucket at 200 MHz can maintain the beam within a longitudinal emittance of ~0.01m, 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 ± 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 ± 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 × 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^3 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 μ'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 μ-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 ~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 ~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 ~7.5 10^{12} muons/bunch, but with a large normalized transverse emittance ε_{XN} ~15 10^3 mm-mrad and a large normalized longitudinal emittance ε_{ZN} ~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 (β_T) function. Figure 8 shows simulation results from cooling in a complete stage containing ~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 ~5%, to a final value of ~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 ~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 μ-storage ring neutrino factory has somewhat different requirements on the muon beams from those of a μ^+-μ^- collider. The ν-source event rate depends primarily on the number of stored muons and not on the quality of the μ-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 μ-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 ν-Factory precooler

In the neutrino factory studies a proton bunch on the target produces a π beam, which is then allowed to drift while the π's decay into μ's and the beam develops a position-energy correlation with the lower-energy μ's trailing behind the higher energy μ'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 μ's and accelerates low energy μ'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 (~100 m full width). The induction linac system includes a 3.5 m long liquid H_2 absorber. This "mini-cooler" stage provides an ~20% reduction in transverse normalized emittance to $\varepsilon_{t,rms} \cong 0.012$ 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 ~ 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 μm (220), with a radius of 18 (11) cm, for the 2.75 m and 1.65 m lattices, respectively. The muons therefore lose ~12 MeV per lattice cell for the 2.75 m lattices and ~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 ~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 ~3.1× (or 5× 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)	σ_X	$\sigma_{X'}$	σ_P	σ_t(per bunch)	<p>
	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 ν-Factory cooling

We have presented in detail only one example of a cooling scenario for preparing the μ-beams of a ν-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 μ-bunches per primary p-bunch and would be more adaptable to future μ^+-μ^- collider beams.

4. Other Cooling Components

The Status Report and the ν-factory studies relied on ionization cooling in a single-pass cooling channel with absorbers (usually 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 ν-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 μ^+–μ^- 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 μ 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 ~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 μ-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[ ]:

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

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:

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

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_μ = 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)	$\tau(\delta=0.7r)$	β^* 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 μs	1.83
20	8000	0.25	0.1717	0.25	63 μs	1.12
20	16000	0.125	0.041	0.125	15 μs	0.79

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

The long lenses needed to obtain large energy losses (~1 m of Li to obtain ~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 μ cooling was also planned in the μ^+-μ^- Collider R&D program; that R&D has been postponed, however.

Practical difficulties exist in matching the large emittance, large Δp μ-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 from $\varepsilon_{t,rms} = 0.001$ 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 β^* 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 μ-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 $\varepsilon_{x,N}$ cooling from 0.0061 to 0.0039 mm, with $\varepsilon_{y,N}$ unchanged and δp 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 β^* 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 β^* 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 $\varepsilon_{T,N}$ from 0.01 cm to 0.0077 cm, while δp increased from 2 to 4.36 MeV/c, and p_μ was reduced to 68 MeV/c. On the order of 4 such lenses with interlaced reacceleration rf can cool $\varepsilon_{T,N}$ from 2×10^{-2} cm-rad to 0.5×10^{-2} with longitudinal emittance increasing by a factor of ~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 aborbers, 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 μ 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 μ'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 (μ^+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 ~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 μ^- 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 μ^+-μ^- Collider intensities is problematic.[4.10]

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

$$\tau_{\mathrm{cool}} \cong \frac{N}{W} \qquad (10)$$

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 ~ $10^{14}s^{-1}$. It is in principle possible to cool 100 GeV μ's before decay.[4.11] However, practical difficulties are significant.

Both of these methods have the potential of cooling μ'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 μ-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 μ^+–μ^- 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 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 = ½) produced equal focusing in both transverse planes, while adding dispersion in one. The bend angles were 45^o, giving a dispersion $\eta \cong 40$ cm. In order to get emittance exchange in both transverse planes, an 18^o/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 mm^3 to 1350 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 λ I transfer matrix per turn, where λ 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 11×10^4 mm^3 to 24 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 ~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 $4.6\text{x}10^4$ mm^3 to 1.2×10^3 mm^3. Approximately 50% of the muons were lost because of aperture restrictions and decays, as well as 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 mm^3 to 200 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 ν-Factory and/or μ^+–μ^- 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 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-δE 100 GeV Higgs Factory and of a 3 TeV μ^+–μ^- 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 μ^+-μ^- 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 ν-factory cooling system could be extended or enlarged to obtain μ-collider beams. The system described in the ν-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 μ^+–μ^- Collider, but the μ^+–μ^- Collider requires that the beam be concentrated in a small number of bunches (combine ~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 μ^+ and μ^- 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 ν-factory front end. That design uses a single bunch of each charge, has many cooling stages, and uses emittance exchange. The initial μ-beam would be captured and rf-rotated in a low frequency rf system (~30 MHz), and the initial transverse cooling system would also be at ~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 μ-bunches would originate from single proton bunches, with an initial linac based phase-energy rotation that would give a ~30 MHz bunch (full length ~8 m)and rms momentum spread of ~±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 (~30 MHz) rf (similar to the ring cooler of ref. [5.12]), and cooled for ~10 turns, during which the transverse emittance could be reduced to $\varepsilon_{T,N}$ ~ 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 (~200 MHz) for ~10 turns of cooling with a transverse cooling goal of $\varepsilon_{T,N}$ ~ 1 mm-mrad (or less), accompanied with longitudinal cooling by a factor of ~4 (or more?) to $\varepsilon_{L,N}$ ~ 0.5 cm (or less). The beam would be extracted into a (predominantly) linear cooler for (mostly) transverse cooling to $\varepsilon_{T,N}$ ~ 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 ν-Factory front end

It is quite possible that the first high-intensity stored-μ facility would be a ν-Factory, and it will be natural to extend that existing system to collider parameters. A ν-factory μ-storage ring facility has a similar total number of μ's as a collider and the longitudinal emittance per bunch is also similar to that required by the high luminosity HF. However, the μ's are split up into a string of bunches (~50), and these must be combined to obtain the high-intensity bunches needed in a high-luminosity collider. The ν-Factory also has transversely cooled bunches (with $\varepsilon_{\perp} \cong 0.002$ 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 ν-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 ν-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 μ-bunches. These strings must then be cooled and recombined into high-intensity μ^+ and μ^- bunches for the μ^+–μ^- 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 μ^+–μ^- 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 μ^+–μ^- 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 15th, 2000.
[3.2] S. Ozaki, R. Palmer, and M. Zisman, ed., "BNL ν-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, "μ^+–μ^- Colliders", CERN 99-12 (1999).
[4.9] K. Nagamine, "New μ^+-μ^- Cooling for μ^+-μ^- 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
[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. Fernow, 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. Fernow & 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-μ 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 ϕ-δE rotation for the μ^+–μ^- Source”, MuCool Note , October 2000, see also Proc. PAC2001.

Figure 1. Concept of ionization cooling.

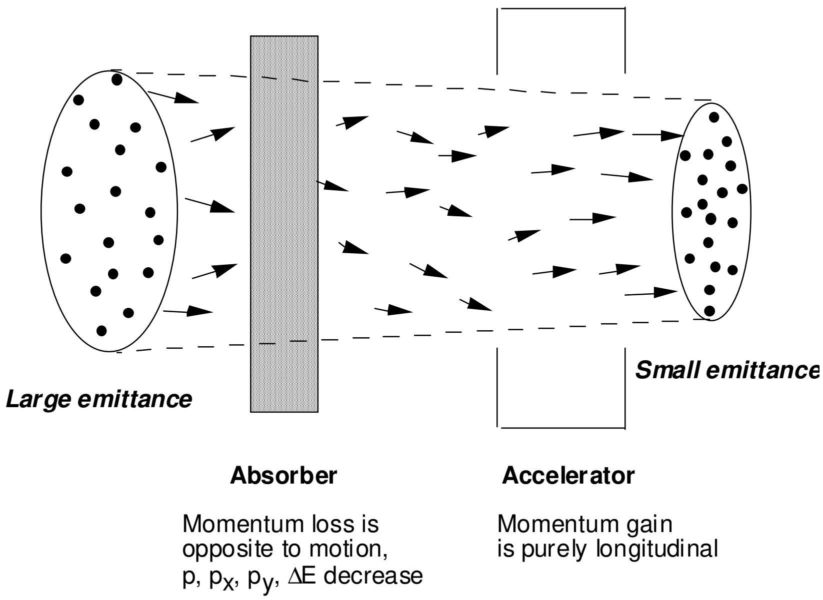


Figure 2. (dE/dx)/ρ (MeV/(gm/cm^2)) as a function of muon momentum P_μ for various atoms. Note that this function is heating (negative slope) for P_μ < ~0.350 GeV/c and becomes strongly heating (steep slope) for P_μ < 0.200 GeV/c.

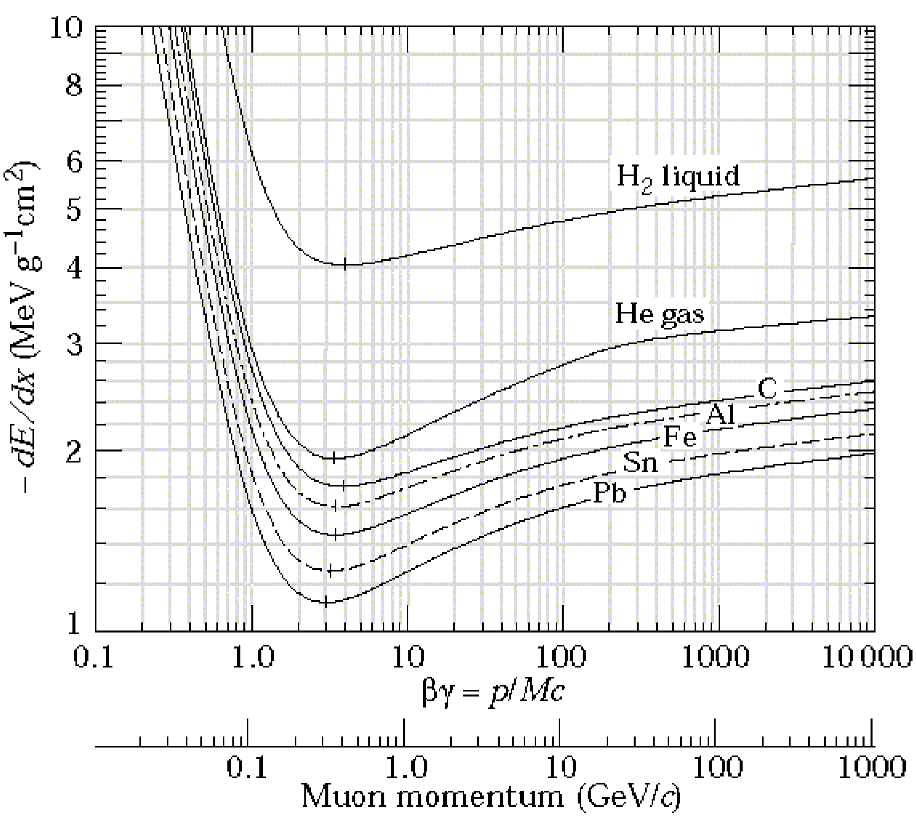


Figure 3. Overview of an emittance exchange section, in which longitudinal emittance is reduced by using a wedge absorber at nonzero dispersion.

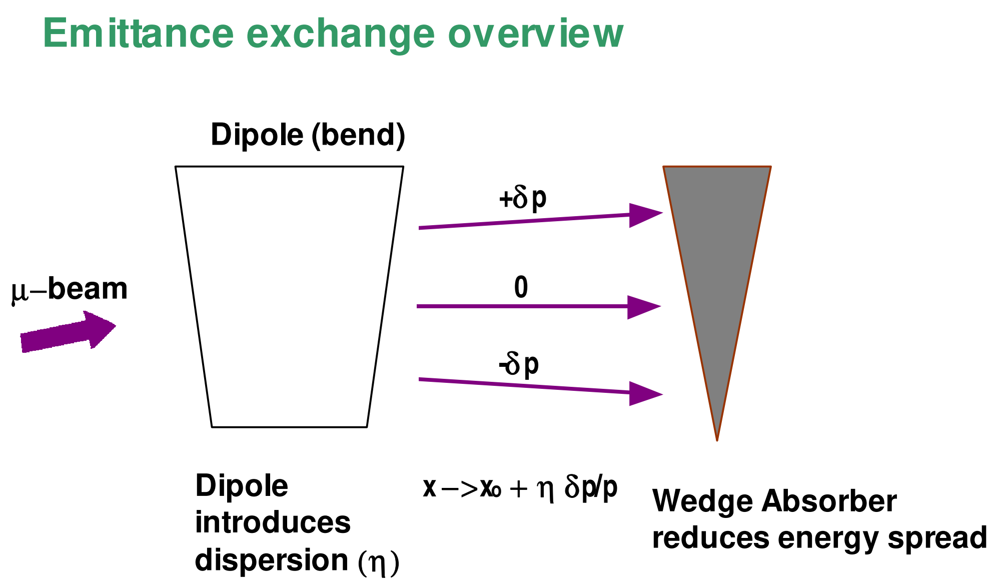


Figure 4. The sum of the cooling partition numbers $\Sigma_g = (g_x + g_y + g_L)$ as a function of momentum P_μ (0—500 MeV/c). g_x and g_y are naturally 1 while g_L becomes strongly negative for $P_\mu < 200$ MeV/c. Σ_g remains greater than 0, which means that ionization loss remains intrinsically cooling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

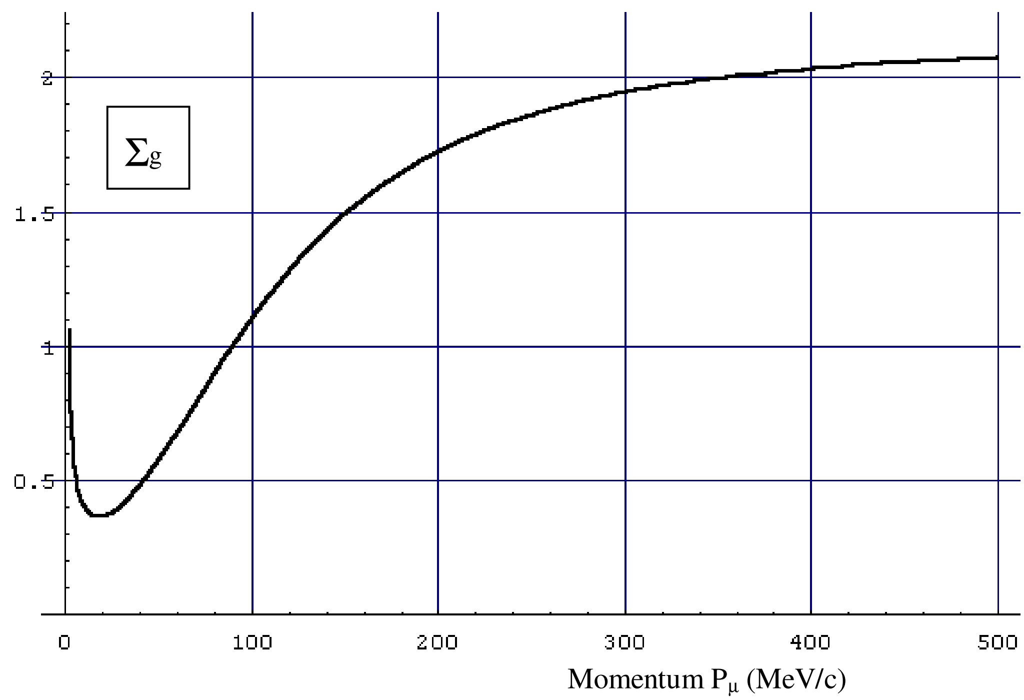


Figure 5. SR baseline cooling scenario

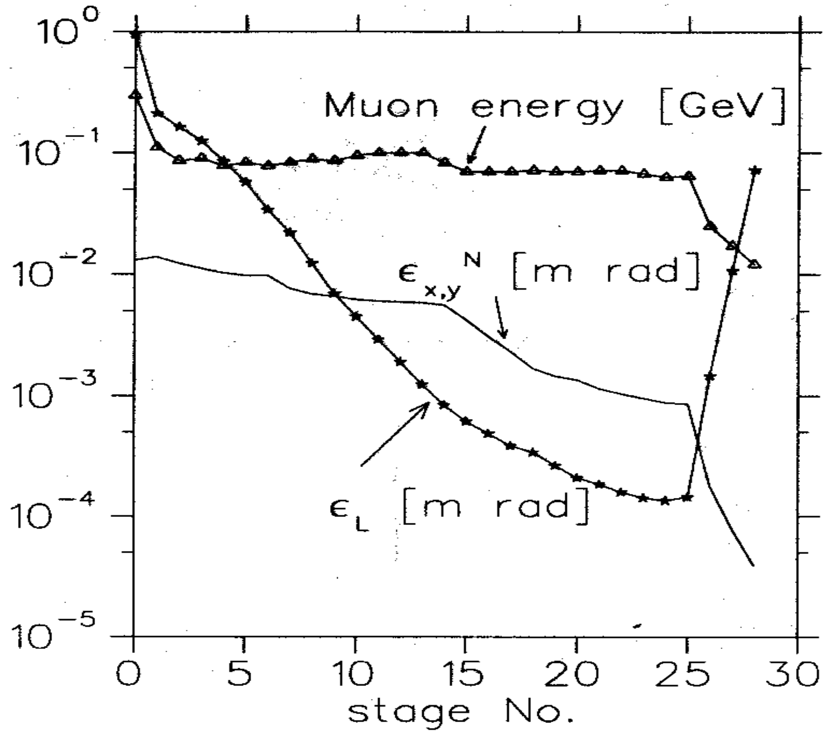
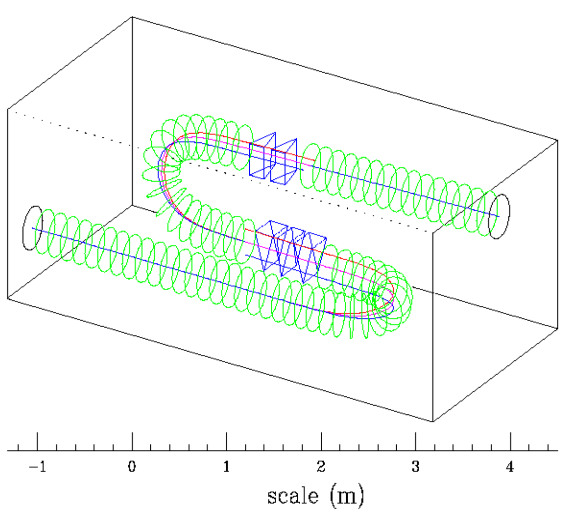
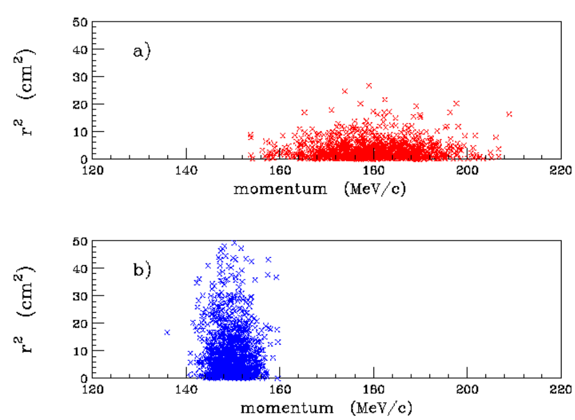


Figure 6. SR emittance exchange example. The figure on the left shows a schematic view of the bent solenoid system with wedges. Simulation results are shown on the right: (a) initial beam distribution of r^2 versus p; (b) final distribution of r^2 versus p.





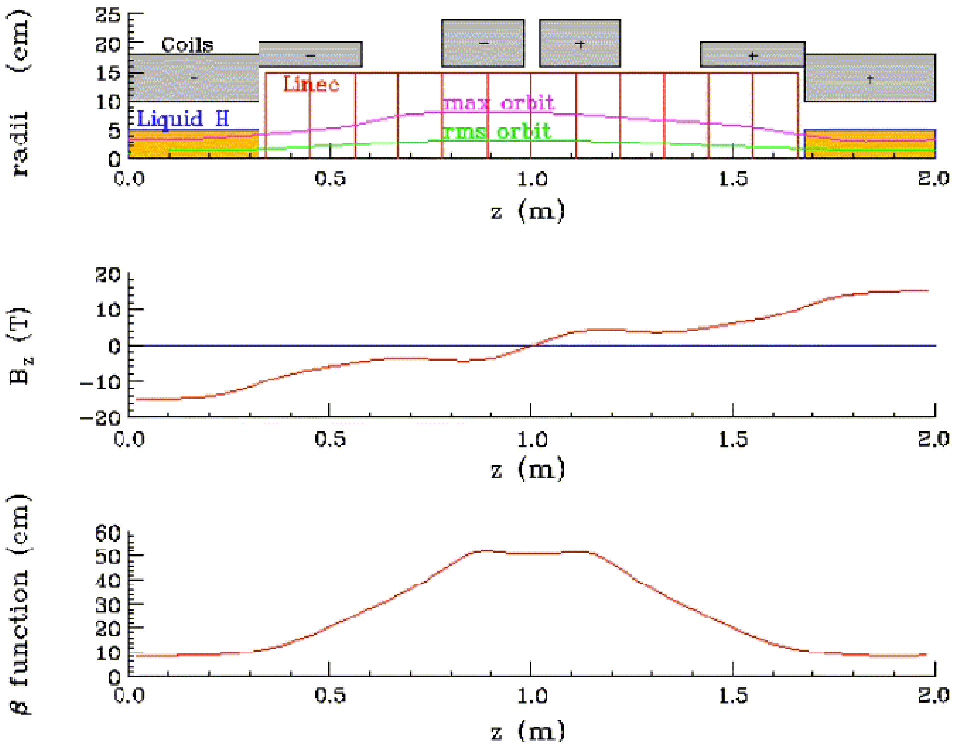


Figure 7. Status report cooling cell. The upper plot is a radial cross-section of a schematic view of a 2 m cooling shell showing hydrogen absorbers, a 12-cell 800 MHz rf cavity and focusing coils around the absorber and rf. The second plot is the solenoidal magnetic field and the third plot is the focusing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

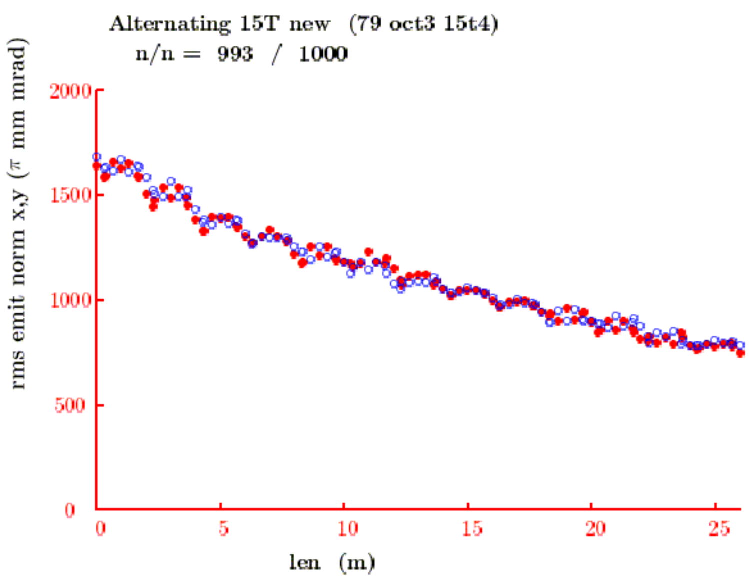


Figure 8. Simulation results showing transverse cooling in $\varepsilon_{T,N}$ from 1.7 mm to 0.8 mm in 13 cells (~1section) of the Status Report cooling channel.

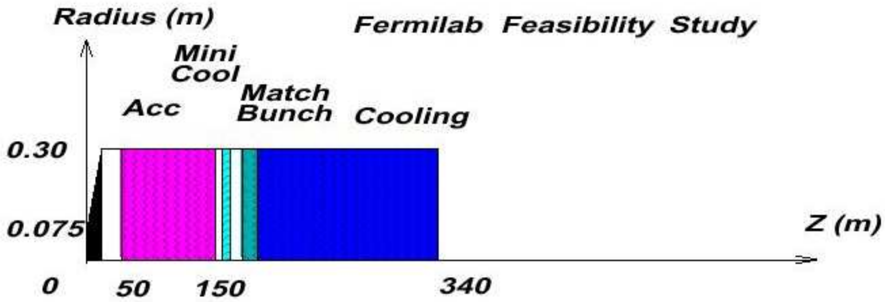


Figure 9a. Study 1 layout of μ-capture and cooling transport.

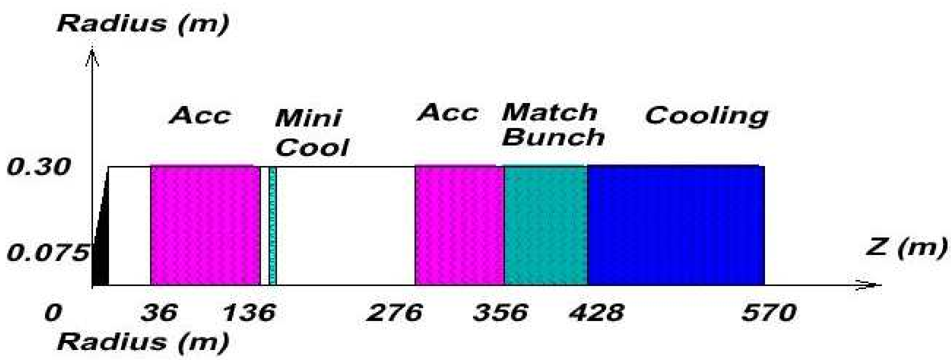


Figure 9b. Study 2 layout of μ-capture and cooling transport

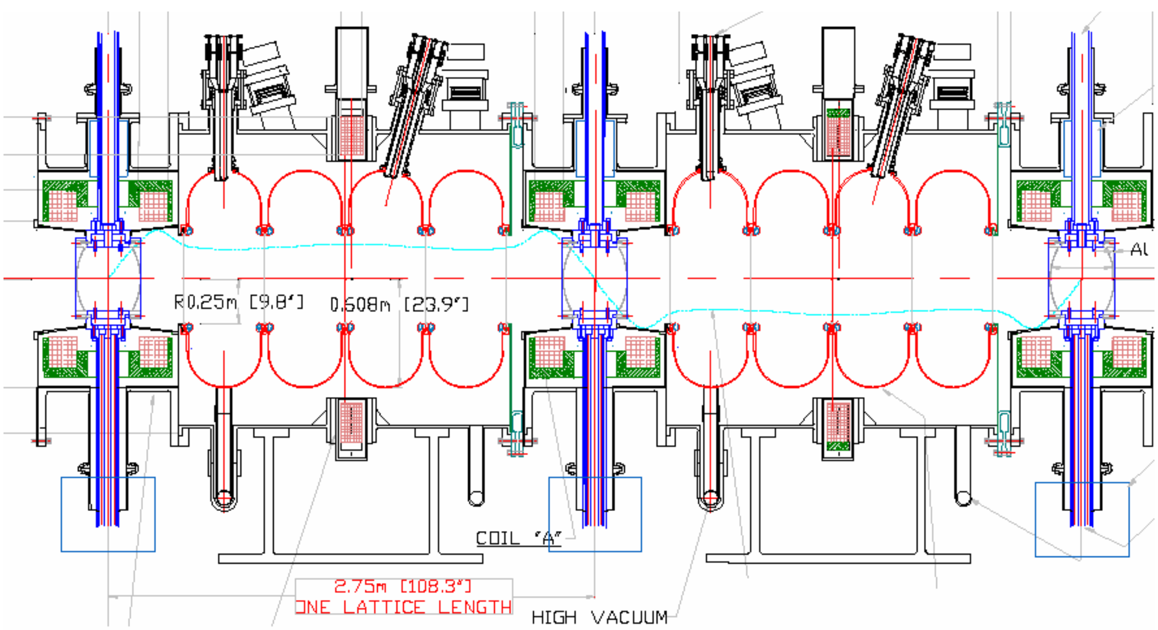


Figure 10. Cross-section of two Study 2 2.75m sFOFO lattice cooling cells, showing two 4-cell rf cavities, 3 Liquid hydrogen absorbers, and magnetic coils for focusing around the absorbers and rf cavities.

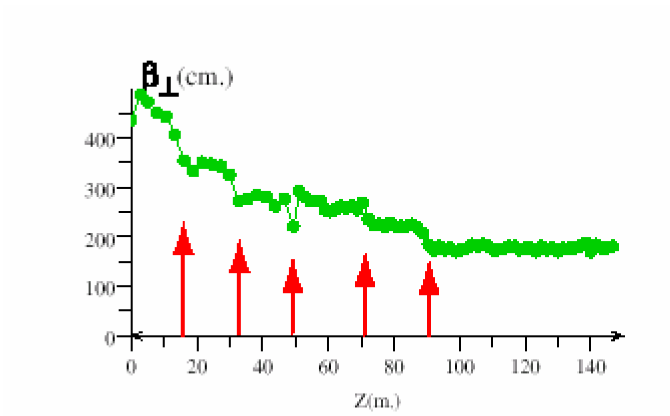


Figure 11. The $\beta_{\perp,\mathrm{min}}$ function in mm at cavity centers, as calculated from the second order moments of a beam in GEANT4 simulations, for the entire sFOFO cooling channel. The 5 arrows indicate the beginning of new lattice sections. (In this figure the cooling channel is extended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

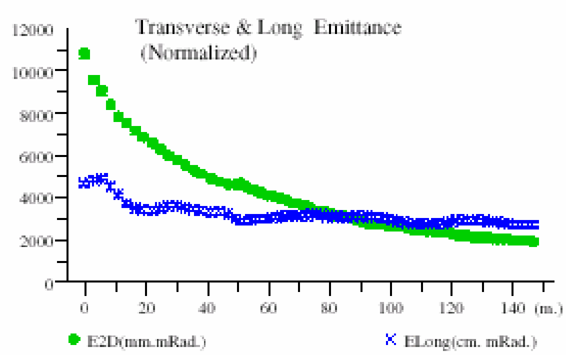


Figure 12. The transverse and longitudinal normalized emittances along the cooling channel.

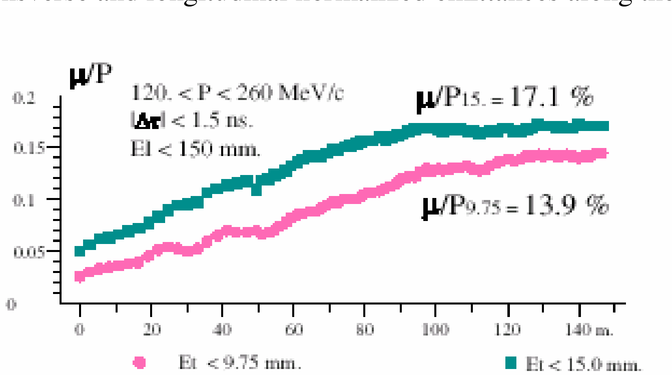


Figure 13. The muon to proton yield ratio for the two emittance cuts, showing that the particle density in the center of the phase space increases as the beam is cooled in the channel. The two curves give the number of particles within the baseline longitudinal and transverse acceptances. The upper line 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).

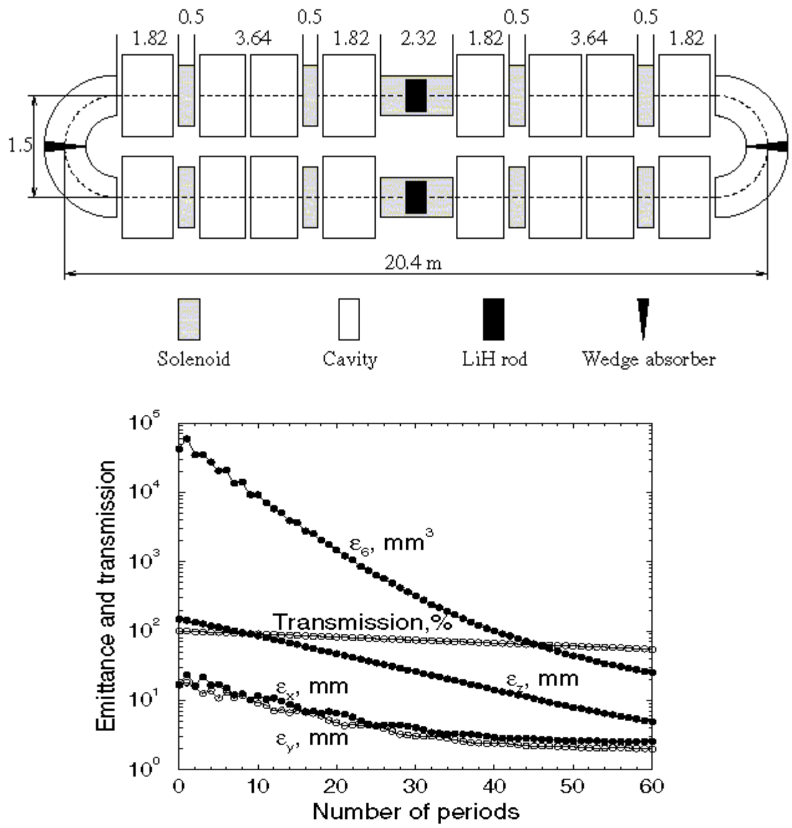


Figure 15. Ring Cooler B

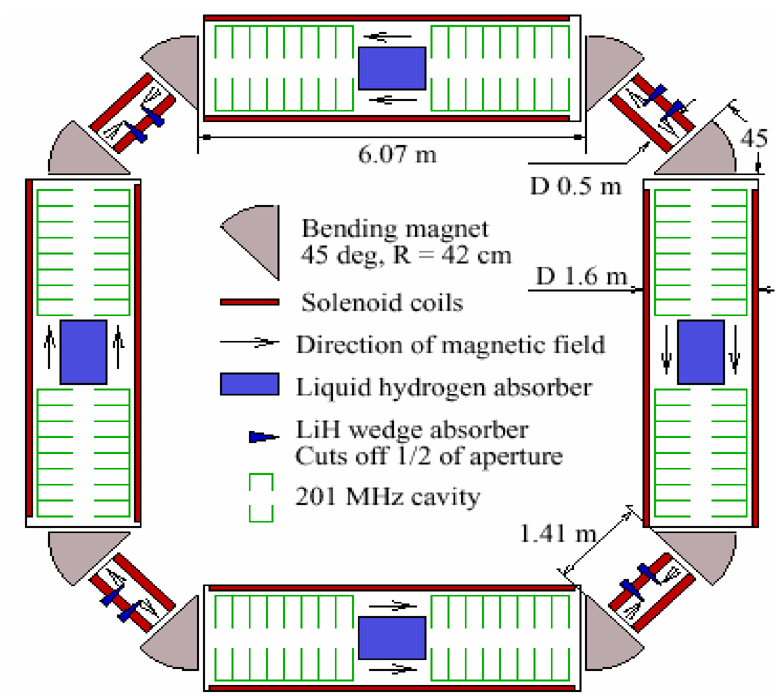
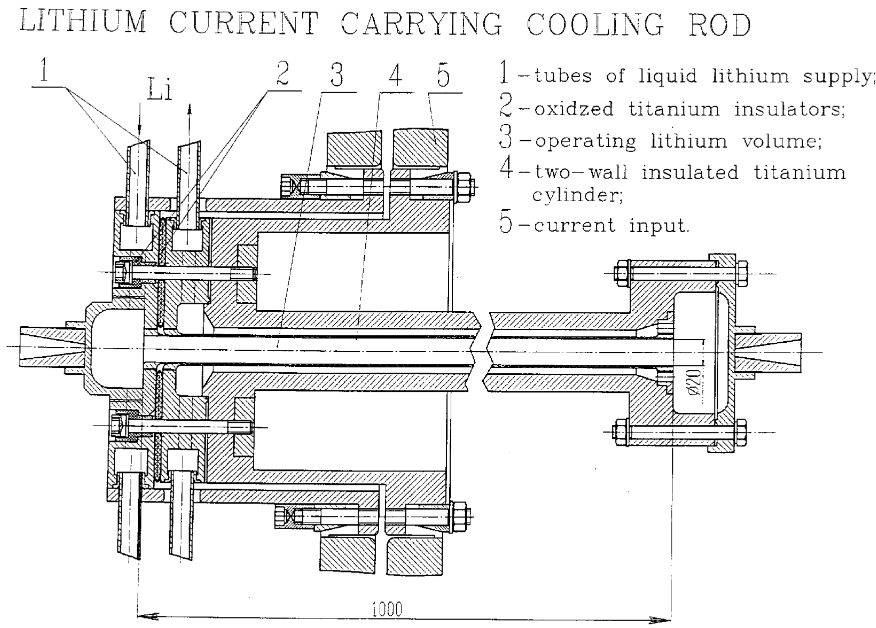
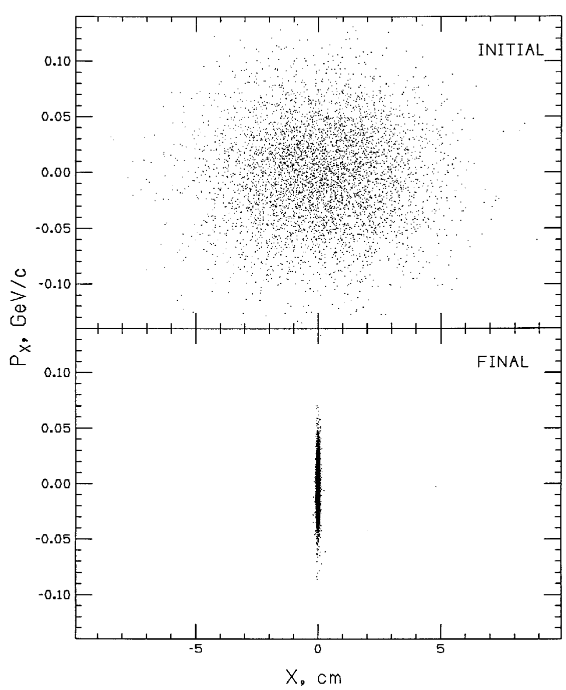


Figure 16. Liquid Li lens schematic





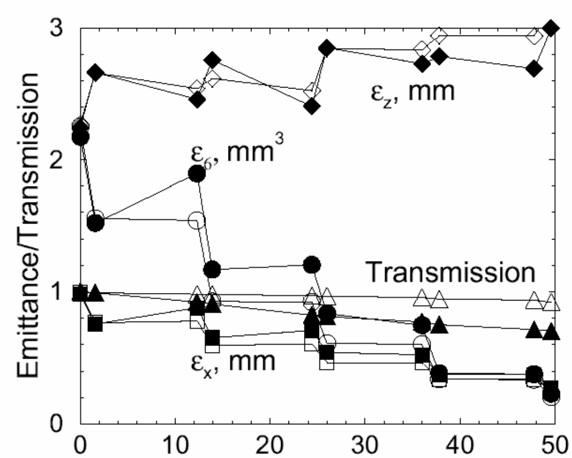


Figure 18 – Simulation of a sequence of Li lenses near the end of cooling

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

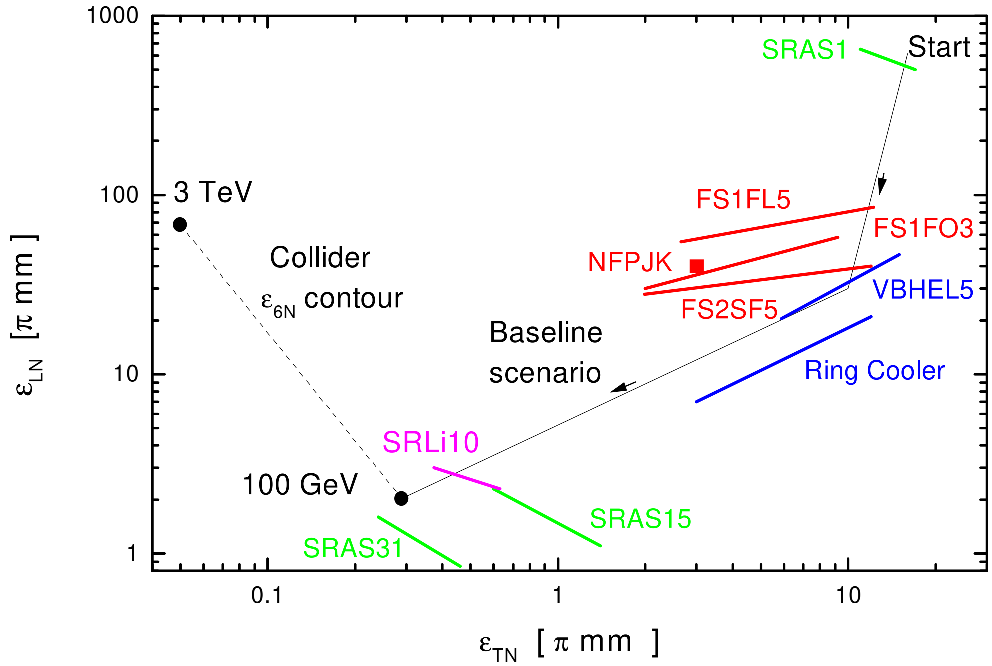


Figure 19. Cooling Summary: The dashed line shows collider cooling goals in emittance space, and the solid line shows a possible path from initial beam parameters [Start] to that goal. Segments of cooling that have been designed and simulated in some detail are shown in color.

.