

# Optimization and Beam Control in Large-emittance Accelerators: Neutrino Factories

C. Johnstone

*Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, IL 60510, USA. Email: [cjj@fnal.gov](mailto:cjj@fnal.gov).*

M. Berz

*Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, USA. Email: [berz@msu.edu](mailto:berz@msu.edu).*

D. Errede and K. Makino

*Department of Physics, University of Illinois at Urbana-Champaign, 1110 W. Green Street, Urbana, IL 61801-3080, USA. Email: [derrede@uiuc.edu](mailto:derrede@uiuc.edu), [makino@uiuc.edu](mailto:makino@uiuc.edu).*

## **Abstract**

Schemes for intense sources of high-energy muons require collection, rf capture, and transport of particle beams with unprecedented emittances, both longitudinally and transversely. These large emittances must be reduced or “cooled” both in size and in energy spread before the muons can be efficiently accelerated. Therefore, formation of muon beams sufficiently intense to drive a Neutrino Factory or Muon Collider requires multi-stage preparation. Further, because of the large beam phase space which must be successfully controlled, accelerated, and transported, the major stages that comprise such a facility: proton driver, production, capture, phase rotation, cooling, acceleration, and storage are complex and strongly interlinked. Each of the stages must be consecutively matched and simultaneously optimized with upstream and downstream systems, meeting challenges not only technically in the optics and component design, but also in the modeling of both new and extended components. One design for transverse cooling, for example, employs meter-diameter solenoids to maintain strong focusing—300-500 mr beam divergences—across ultra-large momentum ranges,  $\geq \pm 20\% \delta p/p$ , defying conventional approximations to the dynamics and field representation. To now, the interplay of the different systems and staging strategies has not been formally addressed. This work discusses two basic, but different approaches to a Neutrino Factory and how the staging strategy depends on beam parameters and method of acceleration.

## **Introduction**

The important stages in the U.S. scheme for a Muon Collider[1] and Neutrino Factory[2] are diagrammed in Figure 1 (left and right, respectively). Muons are created via the decay of pions, and pions are produced by directing an intense beam of protons onto a production target. The initial stage of a muon facility is considered to be a proton driver capable of delivering an ultra-short (3 ns long), high-intensity ( $10^{14}$  p/pulse) beam. Collection, capture and bunching of pions and muons following the production target are the next major systems. First, a 50 m long channel allows the majority of pions with momentum less than  $\sim 1$  GeV/c to decay into muons; the muons are then captured and bunched into 200 MHz rf buckets. Just after production, the captured

particle distribution exhibits an enormous rms momentum spread of  $\pm 55\%$ . The tremendous energy spread of the muons is reduced through phase rotation in an induction linac or, more recently, in consecutive rf stations with varying frequencies. The combined bunching and phase rotation process produces a train of approximately a hundred 200 MHz bunches with a final rms value that is about  $\pm 10\%$  in  $\delta p/p$  and 10 cm in bunch length. (Both the induction linac and the rf-based bunching and rotation schemes produce similar final momentum spreads and bunch lengths so the criterion in choice of technique is merely cost.) The transverse emittance successfully captured is approximately 16 cm-rad (full, normalized) at a central momentum of 200 MeV/c (mainly due to the strong, large-aperture solenoid surrounding the production target). These large emittances cannot be efficiently accelerated so a “cooling stage” for emittance reduction precedes acceleration. To be effective, a cooling channel must be able to accept not only the large transverse emittances (implying large apertures common to both magnetic and rf components), but also a large (full) momentum spread of at least  $\pm 20\%$   $\delta p/p$ . The acceleration and storage rings--albeit nonconventional due to the large admittance and rapid cycle requirements that are imposed by large emittances and short muon lifetimes--represent the final stages of these facilities.

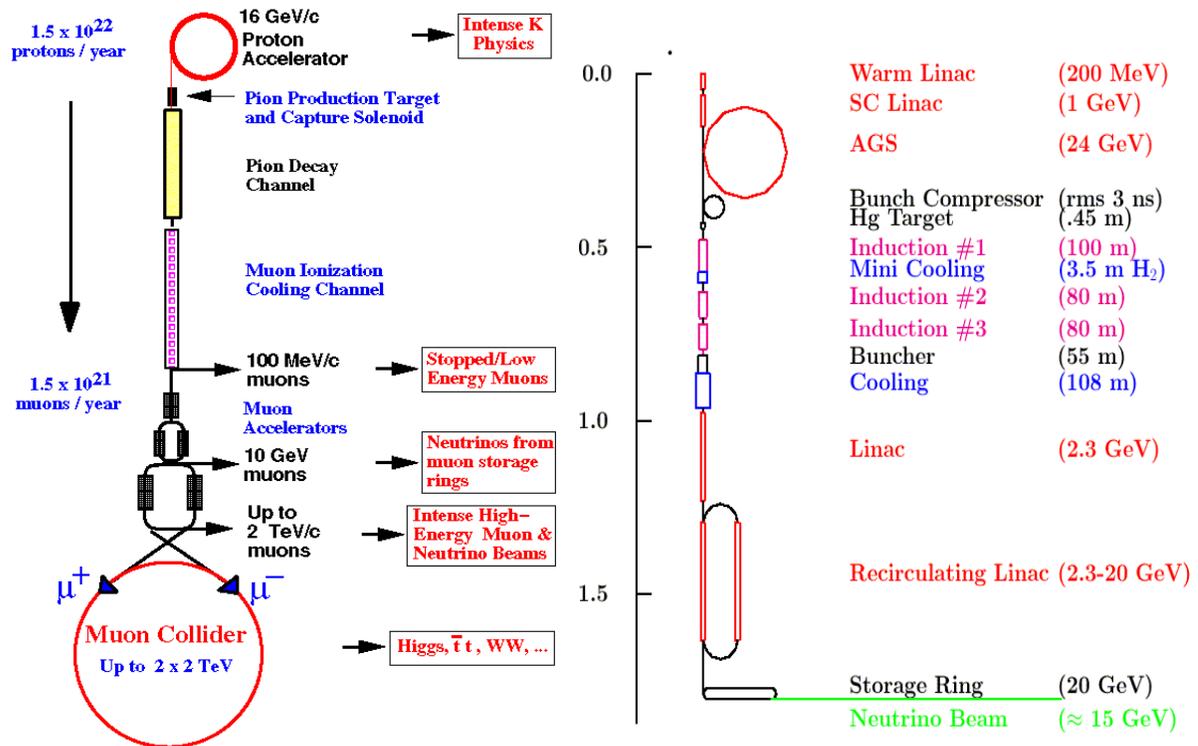


Figure 1. “Schematics of a muon collider[1] (left) and a neutrino factory[2] (right).”

In all of the scenarios developed for Neutrino Factories and Muon Colliders, the captured transverse and longitudinal beam emittances are unprecedented. In comparison with high-energy hadron facilities, the transverse emittance is a factor of 1000 larger and the longitudinal emittance, is 20-100x larger even after bunching and phase rotation. Acceleration and collision of intense muon beams becomes impractical without a significant reduction, or cooling, of incipient emittances—transversely by a factor of 2.5 to 10 for a Neutrino Factory[2][3] and at least a factor of 1000 for a muon collider[1]. In the former the required emittance reduction is tailored to the conditions for acceleration and in the latter for the collider ring. The challenge, then, in the design of these facilities lies in the large beam emittances further complicated by the short muon lifetime, or timescale on which these facilities must operate.

## ***Acceleration in a Neutrino Factory***

In a Neutrino Factory, the ability of, or limits to, accelerating large-emittance beams determines the specifications which upstream systems must meet, particularly the cooling. The downstream storage rings and experiments are presently not the limiting constraint. Acceleration proves, then, not only a difficult stage to develop, it becomes a pivotal one in the path to this facility. To further complicate issues, acceleration must occur rapidly because of potentially heavy losses from decay. Linear accelerators are the optimal choice in this respect, but, above a few GeV, they become prohibitively expensive. Conventional synchrotrons cannot be used because normal conducting magnets cannot readily cycle in the ramping times [1] required by muon decay, nor do they support ultra-large beam emittances. In the past, the U.S. baseline[2,3] relied on recirculating linacs (RLAs) with separate, fixed-field arcs for each acceleration turn. Separate arcs allow control over the pathlength as a function of energy, allowing traversal times to be matched to the rf phase requirements for stable acceleration. Alternative approaches have focused on adapting the Fixed-Field Alternating Gradient (FFAG) accelerator first developed and tested at MURA[4], primarily because of its inherently large longitudinal acceptance. The Japanese approach (KEK)[5], for example, supports a radial-sector FFAG accelerator, but only in the context of a single-muon bunch and low frequency, broadband rf. Recent breakthroughs[6] have resulted in a new design for a FFAG accelerator that can support a high-frequency bunch train, which applies to the U.S. scenario.

Initial transverse cooling (often termed precooling) is unavoidable in any scheme because of otherwise exorbitant component apertures, power levels, and, hence, cost. However, it is the choice of accelerator that predetermines the longitudinal emittance that can be accelerated, and, therefore, the degree of, or even the need for, longitudinal cooling. At a very fundamental level the overall design or staging of a muon facility depends on the method of acceleration chosen, and, for RLAs versus FFAGs, the longitudinal acceptance of the two machines embodies the most significant difference between the two scenarios. Since the RLA approach has been discussed extensively in past feasibility studies and the intent of this paper is to focus on larger-acceptance accelerators, the RLA requirements will be discussed only briefly. Instead this paper focuses on the FFAG approach, addressing the relaxed cooling requirements and re-optimization of the stages relative to the RLA. A number of references to FFAG acceleration can be found in reference 6 for a more complete overview.

The acceleration stage is composed of two subsystems: a linear preaccelerator and a recirculating-beam accelerator (in this case either the RLA or the FFAG machine). The preaccelerator will not be discussed in any detail here, but certain assumptions in its parameters are required to extrapolate to the next stage of acceleration.

### **Preaccelerator**

A linear accelerator is required for the preacceleration stage in order to bring the low-energy muon beam delivered by the cooling stage up to an energy where it becomes feasible to inject into a re-circulating-beam accelerator. Efficient injection into a re-circulating accelerator precludes extended transverse beam sizes, and, given the ultra-large transverse emittances captured upstream ( $\epsilon_n(\text{full}) \approx 16\pi$  cm-rad), the size at injection depends strongly on the transverse reduction factor of the cooling stage combined with the output energy of the preacceleration stage. To achieve both efficient injection and reasonable component apertures in the second acceleration system, a 2.5 GeV preaccelerator has been assumed which, even with modest transverse cooling factors, brings the beam size to less than or approximately  $\pm 10$  cm (full) at this energy.

The emittance reduction factors which follow will be quoted relative to normalized units for convenience, but they are calculated based on a 2.5-GeV preaccelerator (which in turn determines the injection energy of the next machine), along with the lattice design and practical assumptions for component apertures in the RLA and the FFAG cases. It is further assumed that a preaccelerator can be designed that is capable of linearly accelerating both large transverse and large longitudinal emittances, especially in the case of a FFAG accelerator. (Linear acceleration implies that the effective normalized transverse emittance and the absolute momentum spread remain reasonable through the acceleration process and do not “blow-up” significantly.)

## **Recirculating Accelerators**

### ***Recirculating Linear Accelerator (RLA)***

In an RLA, the beam is injected into a linac, accelerated, and returned by separate, fixed-field arcs on each acceleration turn, thereby achieving multiple passes of acceleration through the same linac. At the exit of each linac, the beam is sorted by energy and directed into a separate arc for transport on each acceleration pass. At the end of each arc, the trajectories from all arcs are recombined for acceleration in the opposing linac. The difficulty in the design of the recirculating linacs lies in directing a beam with both a large transverse emittance and a large momentum spread into separate arcs on each acceleration pass. Clearly, to separate cleanly in a passive magnetic system (the only option for the microsecond circulation times), the energy width must be less than the energy difference between consecutive acceleration passes. A large transverse beam size implies a further increase in the distance required for achieving clean separation, promoting an unavoidable correlation between longitudinal and transverse acceptance.

Further complications arise from the requirements on longitudinal motion and the need for chromatic correction (or the inclusion of sextupoles) in the arcs. The full transverse acceptance achieved so far in the recirculating linacs lies between  $1.5\pi$  and  $2\pi$  cm-rad. The momentum spread that can be practically accelerated to extraction energy appears to be  $\pm 1\%$  for 400 MeV of acceleration per turn until a  $\pm 10\%$  limit is reached, a consequence of chromatic correction. As a result, only 4 acceleration turns have been reported for RLAs: with a  $\sim \pm 5\%$   $\delta p/p$  for a 3-11 GeV RLA[7] and  $\sim \pm 10\%$  for a 3-20 GeV RLA[2].

Since the full initial normalized beam emittance, as defined by the present bunching scheme, is about  $16\pi$  cm-rad, this incipient emittance must be reduced or cooled by a factor of 8-10 transversely and at least 2-4 longitudinally before acceleration can be accomplished in an RLA designed for a Neutrino Factory. A higher longitudinal cooling factor is required if the transverse cooling stage is capable of cooling  $> \pm 20\%$   $\delta p/p$ .

### ***Fixed Field Alternating Gradient (FFAG) Accelerators***

Alternatively, a circular accelerator system can be devised with magnetic fields that remain constant during acceleration by adopting an alternating gradient focusing lattice. The arcs of such machines, composed of large aperture magnets, can be designed to accommodate the large energy range in acceleration. The beam centroid orbit is not fixed as in a ramped machine, but rather moves across the magnet aperture during acceleration,. Lattices have been developed which can contain an energy change of at least a factor of four.

There are several classifications of FFAGs which refer to the lattice and momentum dependence of the optics. Since the concern here is with machine admittance, only this general property will be advanced for the different machines. The so-called *scaling* FFAG accelerators, such as the radial or spiral sector, display an almost unlimited momentum acceptance, but transverse acceptance remains restricted. Another approach to FFAGs, referred to as *nonscaling*, employs only linear magnetic elements (quadrupoles and dipoles). Although the transverse optics changes

slowly with energy, these FFAGs demonstrate both strong momentum acceptance and unlimited dynamic aperture (limited only by the physical restrictions of the components).

The *scaling* FFAG designs have successfully achieved dynamic apertures of a few centimeter-radians (full), thereby mitigating the transverse cooling requirements by about a factor of 2 relative to the RLA and yielding an overall transverse reduction factor of  $\sim 5$ . *Nonscaling* FFAG designs exist with component apertures of 30-40 (16-30) cm horizontally (vertically) that accept a value of  $6.4\pi$  cm-rad for the full normalized emittance at the 2.5 GeV injection energy. Their demonstrated emittance reduction factor is modest: only 2.5 in cooling prior to acceleration. Another advantage in both the *scaling* and *nonscaling* cases is total elimination of longitudinal cooling. One drawback is that component apertures are at least 10-20 cm larger horizontally than in the RLAs. This is to accommodate both shifts of the beam centroid during acceleration, and less transverse cooling (beam sizes are larger by 40-100%). (Beta functions are similar in the RLA and the FFAG.)

There is not a significant enhancement, however, in the number of acceleration turns that can be supported over the RLA, even though a large number of turns ( $>20$ ) is not desirable because of muon decay, especially at the lower energies. The reason for this is that in a FFAG accelerator orbit length unavoidably changes with energy; this change can be substantial and can result in a significant phase-slip relative to the rf waveform (unless low-frequency rf is employed). The phase slip accumulates on a per turn basis and eventually precludes acceleration to the extraction energy. This effect limits the number of turns that can be supported under conditions of rapid acceleration when the rf phase cannot be adjusted on a corresponding timescale. Consequently, a dramatic reduction in rf voltage is not gained using the FFAG, but there has been some improvement evidenced in recent work[6].

## Summary

Using references to the acceleration schemes discussed above, a simplistic view of cooling for a Neutrino Factory is illustrated in Figure 2. By way of explanation, the technique of ionization cooling permits reduction of transverse emittances (4D phase space), or beam sizes, to levels acceptable for injection into accelerators with large momentum acceptance, or into ring coolers[4]. Ring coolers are multi-turn cooling channels designed to further reduce the transverse plus longitudinal emittance (6D phase space) to the smaller values required by a Neutrino Factory RLA or for a Muon Collider.

Elimination of all longitudinal cooling, involving further muon decay, inevitable transmission losses, injection and extraction difficulties, and general R&D issues associated with the optics and component designs of advanced cooling channels, makes a persuasive argument to adopt the FFAG as the acceleration stage. For an FFAG, it appears sufficient to consider only a simple cooling stage: a straight or linear channel for transverse cooling (bends are required only for emittance exchange or longitudinal cooling). Such a linear channel must bring the initial, precooled emittance of  $16\pi$  cm-rad down to  $3.2\pi$  cm-rad for the *scaling* FFAG and  $6.4\pi$  cm-rad for the *nonscaling* FFAG accelerator. (It should be noted that a linear transverse precooler stage is relevant not only for the FFAG scenario, it extends to RLAs and Muon Colliders as well. For the latter, ring coolers can be inserted, for example, as they also require “precooled” beam since the present design cannot accept large transverse emittances.) The following sections deal with optimizing a cooling stage for FFAG acceleration.

## Cooling Staging Relative to Acceleration

$\epsilon_{\text{normalized}}$ (full)/per plane			
16 cm-r	6.4 cm-r	3.2 cm-r	1.6 cm-r

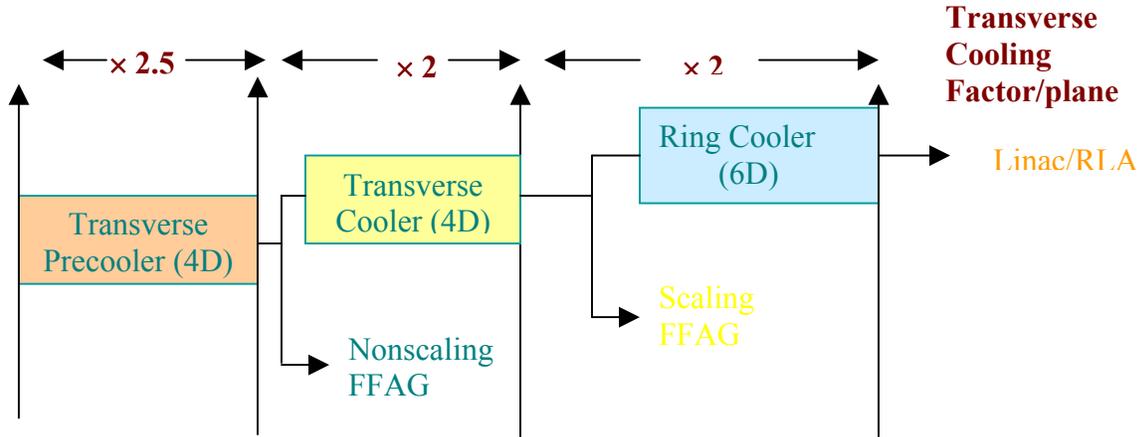


Figure 2. Rough schematic of staged cooling relative to the acceleration model for a Neutrino Factory.

### The Cooling Stage

Emittance reduction, or ionization cooling, occurs because the muon beam loses momentum in all directions when traversing a target, or so-called absorber, and this energy loss can be replaced solely in the longitudinal direction by re-acceleration in an rf cavity; thus decreasing the beam's divergence for a given transverse dimension. The designs of stable optical configurations for cooling channels are particularly challenging because the straightforward cooling dynamics described above compete with the stochastic processes in the absorber; predominately multiple, or Coulomb, scattering which re-heat the beam. A net cooling effect can be achieved only if the cooling terms surpass the reheating ones, a state achieved through proper optics design in a cooling channel. The equation which follows represents the emittance increase in an absorber due to multiple scattering and, in the presence of cooling (reacceleration by an rf cavity), the minimum emittance achievable for a specific channel design. From an optics standpoint, it is clear from equation 1 that the lower the average beta across the target, the proportionally lower the emittance increase from re-heating (multiple scattering), and therefore, the lower the achievable equilibrium emittance

$$e_{N,\min} = \frac{\beta_{\perp} (14 \text{ MeV})^2}{(2\beta m_{\mu} L_R \cdot dE/ds)} \quad (1)$$

where  $\beta_{\perp}$  is the transverse beta function at the absorber,  $\beta$  the relativistic velocity,  $m_{\mu}$  the mass of the muon,  $L_R$  the radiation length of the absorber material, and  $dE/ds$  the energy lost per meter in the absorber.

Therefore, the performance of a cooling channel depends not only on the beta functions at the absorber and on their constancy across a large momentum range, but, as important, on the ratio of the starting emittance to the equilibrium emittance; the ratio needs only to be larger than  $\sim 1.5$  for effective cooling. This observation indicates that the extremely low beta functions required in the latter stages of cooling are not a prerequisite during the early stages and may actually be less optimal from both a technical and nonlinear standpoint. Relaxing the low-beta conditions at the absorber impacts tremendously the design, stability and strength of the elements used in the

upstream cooling. This observation will later be used to optimize the design of the cooling channels and adapt to the mode of acceleration chosen.

Assuming that a full transverse acceptance of  $1.5\pi$  to  $2\pi$  cm-rad, as required for an RLA, corresponds to  $2.5\sigma$  of a Gaussian beam profile, the rms normalized emittance demanded from transverse cooling is  $2.4\pi$  to  $3.2\pi$  mm-rad. This degree of cooling (Equation 1) corresponds to a cooling channel with a low-beta of  $< 0.5$  m at the absorber. In the FFAGs, the relaxed transverse conditions allow betas at the absorber to increase to  $\leq 1$  and  $\leq 2.5$  m for *scaling* and *nonscaling* machines, respectively.

In momentum the practical range for ionization cooling extends roughly from  $\sim 150$  to  $\sim 600$  GeV/c. Application of all or part of this range in the specification and design of a cooling channel depends on the output emittance requirements for acceleration. For RLAs with their restricted momentum acceptance, a transverse cooling channel with a 155-245 MeV/c acceptance based on large-aperture solenoids and followed by longitudinal cooling is the accepted solution. In the case of FFAG acceleration one argues for a cooling channel which can accept and cool over as much of this momentum range as possible. When combined with the relaxed transverse specifications, a simple, quadrupole-based channel is well-suited to FFAG acceleration. The concepts for this channel will be developed in the following sections. Such a channel could also serve as a precooling stage for RLA acceleration.

### **Linear Quadrupole Precooler**

When beta functions at the absorber exceed or approach one meter, the focusing strength is dramatically reduced and the absorber no longer has to be located at the lowest or a very-low beta point, allowing more flexibility in the choice of optical structure and focusing elements. This observation represents the basis for designing a competitive cooling channel based on normal-conducting quadrupoles in a simple lens, or FODO-cell configuration. Its application is solely as an upstream stage of cooling and, being a linear channel with no bends, serves to reduce the large transverse beam size in preparation for acceleration in an FFAG or for injection into ring coolers.

### ***Comparison of Optical Structures***

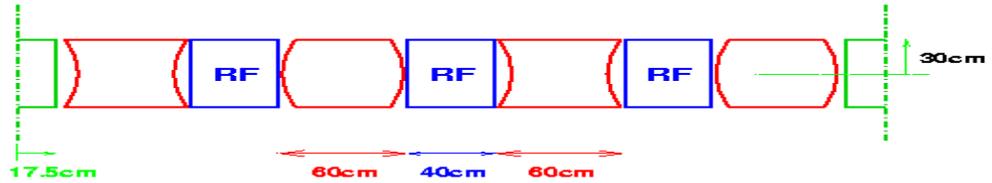
With a sufficiently-relaxed beta at the absorber, one can consider a short, alternating quadrupole lens structure. The advantages of a short FODO cell structure over a doublet or triplet quadrupole telescope are primarily in the acceptance and stability of optical parameters over a tremendous chromatic range. The dynamical range in telescope structures is about  $\pm 5\%$   $\delta p/p$ , beyond which there is no closed-orbit solution. The limited momentum acceptance of the triplet/doublet quadrupole channels restrict their implementation to after longitudinal, or momentum, cooling has occurred and are not considered further here. However, in standard (implying repetitive) FODO-cell optics, the minimum beta in one plane is located at the maximum beta in the other. A minimum beta or beam size cannot be established simultaneously in both planes, and, therefore, the absorber cannot be located at the lowest beta point in this type of channel. The smallest beta for both planes combined is found halfway between the quadrupoles, at the “crossing point” in  $\beta_x$  and  $\beta_y$ . Due to this limitation, the valid application of a FODO-based cooling channel is just after capture and phase rotation.

### ***FODO-based Quadrupole Cooling Cell***

For a short FODO cell, the average beta in both planes is equal and lies between 1 and 2 meters for normal conducting quadrupoles and short ( $\sim 0.5$  m) spacing between them. The value of the beta functions at the crossing point is unusually stable over a large momentum range: from  $-20\%$  to almost  $+100\%$  if the phase advance is adjusted properly. The optics rationale for its design and stability will be discussed after presenting the physical parameters chosen for this channel.

The physical parameters chosen for this channel shown in Figure 3 were initially chosen to be comparable, or competitive with the sFOFO channel[2]. The aperture of the quadrupole was chosen somewhat conservatively—its length is constrained to be equal to its aperture—in order that the quadrupole field profile and therefore the optics are not fringe-field dominated.

## Construction of FODO Quadrupole Cooling Cell



### COOLING CELL PHYSICAL PARAMETERS:

Quad Length	0.6 m
Quad bore	0.6 m
Poletip Field	~1 T
Interquad space	0.4 - 0.5 m
Absorber length	0.35 m *
RF cavity length	0.4 - 0.7 m*
Total cooling cell length	2 m

Muons ( $p = 155 \text{ MeV}/c$  to  $400 \text{ MeV}/c$ )  
 Liquid H Absorber:  $-dE/dx = -12 \text{ MeV}/35 \text{ cms}$   
 RF Cavities: Energy gain  $+12 \text{ MeV}/\text{Cell}$   
 to compensate the loss in the absorber

\*Lengths over 0.4 m can be accommodated within the quadrupole ends

Figure 2. Physical parameters and layout of the FODO-based quadrupole cooling channel.

The average beta achieved at the absorber in this channel is 1.6 m at 200 MeV/c (this is the defined central momentum of the sFOFO channel[2]). This absorber beta yields a design equilibrium emittance (rms, normalized) of  $6.8\pi \text{ mm-rad}$ , or a practical rms final beam emittance of  $10.2\pi \text{ mm-mrad}$ . Assuming a  $2.5\sigma$  Gaussian, the full final transverse beam emittance is then  $6.4\pi \text{ cm-rad}$ , or a factor of 2.5 below the  $16\pi \text{ cm-rad}$  emittance coming from the upstream bunching stage.

### *Momentum Performance of a FODO-based Cooling Channel*

With the extreme demands placed on momentum performance, it is instructive to examine the FODO cell under the precepts of thin-lens conditions. First, it is useful to choose a reference momentum,  $p_0$ , and study the phase advance as a function of momentum relative to the reference in order to evaluate performance limits. For such a study, it is only practical to assign a working point, or initial cell phase advance, to this reference momentum and one which is centrally located between stability limits:  $0^\circ$  and  $180^\circ$ . Clearly  $90^\circ$  is an obvious choice, hopefully optimizing the momentum reach of the channel. This choice of phase advance was applied to a  $p_0$  of 200 MeV/c, a value chosen to be compatible with current cooling channel designs. The phase advance dependence,  $\varphi$ , on momentum can now be obtained in the thin-lens approximation, starting with

$$\sin \frac{\varphi}{2} = 1/\kappa; \quad \text{with } \kappa = f/L \text{ for a thin lens}$$

where  $f$  is the focal length of  $\frac{1}{2}$  of a full quadrupole, and  $L$  is the length of a half cell from quadrupole center to center (see, for example references listed in [7]). Since  $f = 1/(kl)$  with  $kl$  being the *strength*, ( $m^{-2}$ )  $\times$  *length* (m) of a half-size quadrupole; then

$$\sin \frac{\varphi}{2} = \frac{0.3B'l}{p} L; \quad \text{since } k = \frac{0.3B'l}{p}$$

if  $B'$  is the quadrupole gradient in  $T/m$  and  $p$  is the momentum in  $GeV/c$ . Selecting  $\varphi = 90^\circ$  at  $p_0$ , the reference momentum implies further:

$$\sin \frac{\varphi}{2} = \frac{1}{\sqrt{2}} \frac{p_0}{p}; \quad \text{with a clear lower limit of stability, } p = \frac{1}{\sqrt{2}} p_0 \quad (2)$$

Differentiating gives:

$$\frac{1}{2} \cos \frac{\varphi}{2} d\varphi = -\frac{p_0}{\sqrt{2} p^2} dp \quad \text{or} \quad \frac{d\varphi}{dp} = \frac{\sqrt{2} p_0}{p^2 \sqrt{1 - p_0^2 / 2 p^2}} \quad (3)$$

Notice that for a  $p_0$  of 200 MeV/c, the above analysis (equation 2) gives a lower momentum cutoff for the channel of  $\sim 140$  MeV/c and, at large  $p$ , the phase advance varies more and more slowly, as  $1/p^2$ . The results of this analysis are graphed in Figure 3 clearly demonstrating the large play in momentum of the simple-lens FODO cell. When compared with calculations, an almost constant factor of 0.8 was needed to translate the changes in phase advance from the thin-lens model to ones accurate for the channel as designed. "Thick" quadrupoles actually extend the momentum reach of the channel beyond the thin-lens prediction.

The slow variation in phase advance does not set restrictions on the length of the cell, but the variation of the peak beta function with momentum does. Using the definitions above, the peak beta function for a FODO cell is given by:

$$\beta_{\max} = L \frac{\kappa(\kappa + 1)}{(\kappa^2 - 1)^{1/2}}; \quad \frac{d\beta_{\max}}{dp} = L(\kappa^2 - \kappa - 1) \frac{d\kappa}{dp} = 0 \quad \text{for a minimum} \quad (4)$$

In the above equation (4),  $(\kappa^2 - \kappa - 1)$  can only be set to 0 locally (at  $\sim 76^\circ$ ), but this does not guarantee stability in the beta function over a large range in momentum. The only approach that minimizes  $d\beta_{\max}/dp$  over a broad spectrum is to let  $L$  approach 0. No drift between quadrupoles is optimal, but the choice of a short drift of  $\sim 0.5$  m (which corresponds here to a half-cell length of 1 m) intentionally slows the variation of the maximum beam size with energy and at the same time insures a more feasible technical channel design. (Absorbers and rf cavities are not installed inside magnet apertures.) The variation of the maximum beta with momentum for this design is shown below.

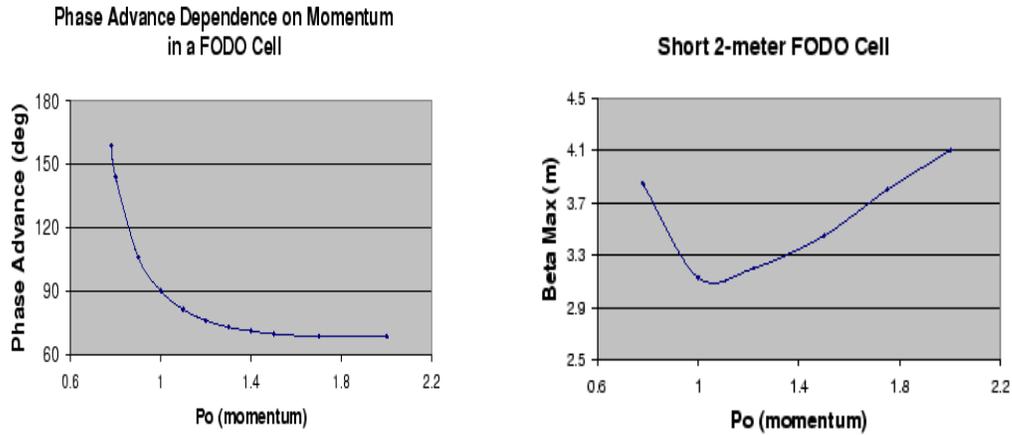


Figure 3. On the left is the phase advance of a FODO cell plotted with respect to an arbitrary momentum,  $p_0$ , whose phase advance has been set to  $90^\circ$ . On the right is the variation of the peak beta function relative to  $p_0$  for a half-cell length of 1 m.

When the momentum dependence of the average beta at the absorber was studied, the change was also small up to 300 MeV/c due to the slowly varying peak beta values as a function of momentum. The minimum beta value is a strong function of energy, but not the peak beta, indicating a large, constant geometrical acceptance as a function of energy for a given quadrupole aperture. (By comparison, a constant normalized, rather than geometrical, acceptance appears to be more characteristic of solenoidal confinement.)

#### ***The simulation and results***

Both the sFOFO and quadrupole cooling channels were fully modeled and tracked to high-order using the code COSY[9]. The simulation included:

- full nonlinear terms
- with full solenoidal[10] and quadrupole fringe fields[11,12] (including different models);
- multiple scattering (absorbers + windows)
- energy loss including straggling and spin
- $dE/dx$  as a function of energy
- 200 MHz sinusoidal rf

Simulations were initiated using a standard Enge function fall-off[13] for the quadrupole end-field profile which is known to represent fringe field effects to an accuracy sufficient for the initial analysis.

$$F(z) = \frac{1}{1 + \exp(a_1 + a_2(z/D) + \dots a_6(z/D)^5)} \quad (5)$$

where  $z$  is the distance perpendicular to the effective field boundary (for multipoles this is the arc length along the reference trajectory),  $D$  is the full aperture of the element and  $a_1 - a_6$  are the Enge coefficients which depend primarily on the geometric details of the element and are adjusted to reproduce measured or design-computed data. The simulations were performed with “as-built” or measured quadrupole end fields of existing large-aperture quadrupoles[14].

For the tracking, particles were launched in 2 cm steps along both axes and along the diagonal starting at the center of the absorber. Without cooling (absorber and rf cavity), the ideal solenoidal channel—with no field errors—exhibited an average 50% larger dynamic aperture over its stable range in momentum acceptance,  $p=155\text{-}250$  MeV/c than the quadrupole channel with full fringe fields. (Without fringe fields, the quadrupole channel has the larger transverse acceptance.) With cooling turned on, however, the transmission losses in the quadrupole channel exclusive of muon decay were almost negligible—less than 1%. (This transmission corresponds to an rms bunch length of 7.5 cm, a  $\sigma_E$  of 12 MeV,  $\phi_s=60^\circ$ ,  $\Delta\phi_s=\pm 54^\circ$ , which corresponds to the 200 MHz rf bucket being about half filled.) Since the cooling rapidly reduces the emittance, the dynamic aperture is almost not relevant because a beam that fills the entire quadrupole aperture is cooled and is not lost even in the presence of fringe fields.

The predicted cooling behavior was observed; if a Gaussian distribution is launched which fills the quadrupole aperture, then the final rms of the distribution was found to be near 6.8 mm-rad (normalized) for this specific channel (Figure 9). The longitudinal losses of the quadrupole channel appear to be less than the solenoidal channel, the reason being the absence of longitudinal-transverse correlations that plague solenoids.

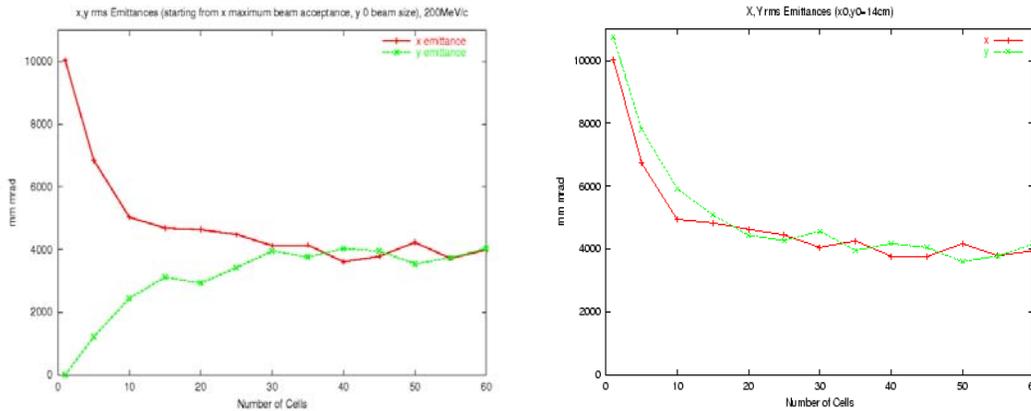


Figure 9. Emittance reduction as a function of number of cooling cells. Left plot shows this starting emittance in x, no emittance in y with corresponding emittance growth up to the equilibrium emittance. Plot on right are particles launched in x,y Gaussian distributions, but along the diagonal.

As mentioned it is important to calibrate the expanded momentum reach of the quadrupole cooling channel. The important point to note here is that because the geometrical acceptance is constant, the normalized one is increasing (the relativistic velocity is not changing significantly). Hence, the absorber beta can be allowed to track the increase in normalized emittance acceptance. When benchmarked against the full cooling simulation performed at a  $p_0$  of 200 MeV/c, the channel still cools at 400 MeV/c. The momentum reach of the cooling in the quadrupole channel appears to be significantly larger than in the sFOFO, which extends from 155 to 245 MeV/c.

## Conclusion

In this paper, a first pass at optimizing the stages in a Neutrino Factory has been presented. Clearly the staging and optimization are critically dependent on the choice and format of accelerator. It has been demonstrated that possibly the simplest scenario is a *nonscaling* FFA machine coupled to a linear (both geometrically and optically) cooling channel constructed from the simplest quadrupole lens system. Detailed simulations show that the channel cools efficiently

and much beyond the momentum range of a sFOFO[2] cooling channel with similar magnetic apertures, or from approximately 155-400 MeV/c in the former compared with 155-250 MeV/c in the latter. Applying different—both assumed and measured fringe fields—to represent the quadrupole elements fully has been an integral part of the simulations and ensures feasibility in quadrupole design and performance. An important concluding observation is that such a channel cools effectively over a large variation in the fringe-field profile. Extensive simulations are currently in progress on the *nonscaling* FFAG and results are also encouraging. This cooling/acceleration scheme potentially represents the baseline scenario for the next U.S. feasibility study.

## References

- [1] The  $\mu^+ \mu^-$  Collaboration, “ $\mu^+ \mu^-$  Collider: A Feasibility Study”, BNL-52503, Fermi-Conf-96-/092, LBNL-38946, July, 1996.
- [2] S. Ozaki, R. Palmer, M. Zisman, and J. Gallardo, eds. “Feasibility Study II of a Muon-based Neutrino Source”, BNL-52623, June, 2001, available at <http://www.cap.bnl.gov/mumu/studii/FS2-report.html>.
- [3] 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). Fermilab-Pub-02-149-E, Jul 2002, submitted to PRSTA.
- [4] K. Symon, MURA-KRS-6,(MURA-43)(1954) ; K. Symon, et al, Phys. Rev. 103 p.1837 (1956).
- [5] S. Machida, et al, “Beam Optics Design of an FFAG Synchrotron”, (MOP1B20), submitted to EPAC 2000, and Y. Sato, et al, “Development of a FFAG Proton Synchrotron”, (MOP1B21), submitted to EPAC 2000.
- [6] C. Johnstone and S. Koscielniak, “FFAGs for Rapid Acceleration”, to be published in NIM A and “Recent Progress on FFAGs for Rapid Acceleration”, Proc. Of the APS Snowmass 2001, The Future of Particle Physics, Snowmass, CO June28-July 21, 2001.
- [7] C. Johnstone, D. Neuffer, and E. Forest in “The CERN Neutrino Factory Study”, Chapter 6, available at <http://www.cdj.home.cern.ch/cdj/yrep/yr.html>. C. Johnstone, D. Neuffer, and E. Forest, “A 3-11 GeV Recirculating Linac for Muon Acceleration: Design and Simulation” Proc. Of the 2001 Particle Accelerator Conference, Chicago, IL, June 18-22, 2001.
- [8] Muon Collider/Neutrino Factory Collaboration, “Recent Progress in Neutrino Factory and Muon Collider Research within the Muon Collaboration”, submitted to Phy. Rev. ST Accel Beams, <http://www.cap.bnl.gov/mumu/pubs/prstab-030422/prstab.pdf>.
- [9] M. Berz and K. Makino, COSY INFINITY Version 8.1 - user's guide and reference manual. Technical Report MSUHEP-20704, Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, 2001. See also <http://cosy.pa.msu.edu>.
- [10] K. Makino, M. Berz, D. Errede, and C. Johnstone, “High-Order Map Treatment of Superimposed Cavities, Absorbers, and Magnetic Multipole, and Solenoid Fields”, to be published Proc. Charged Particle Optics Conference-6, Univ. of Maryland, Oct. 22-25, 2002.
- [11] M. Berz, “Modern Map Methods in Particle Beam Physics”, Academic Press, San Diego, 1999.

G. Hoffstatter and M. Berz. "Symplectic Scaling of Transfer Maps including Fringe Fields", Physical Review E, 54,4, 5664, 1996.

G. H. Hoffstatter. "Rigorous Bounds on Survival Times in Circular Accelerators and Efficient Computation of Fringe-Field Transfer Maps". PhD thesis, Michigan State University, East Lansing, Michigan, USA, 1994.

[12] S. Kowalski and H. Enge, RAYTRACE. Technical report, MIT, Cambridge Massachusetts, 1985.

[13]B. Erdelyi, Simplectic Approximation of Hamiltonian Flows and Accurate Simulation of Fringe Field Effects", PhD thesis, Michigan State University, East Lansing, Michigan, USA, 2001. <http://bt.pa.msu.edu/cgi-bin/display.pl?name=erdelyiphd>

[14]C. Johnstone, M. Berz, D. Errede, and K. Makino, "Muon Beam Ionization Cooling in a Linear Quadrupole Channel", to be published Proc. Charged Particle Optics Conference-6, Univ. of Maryland, Oct. 22-25, 2002.