

# Optimization of flat to round transformers with self-fields using adjoint techniques

L. Dovlatyan<sup>✉</sup>,\* B. L. Beaudoin, S. Bernal<sup>✉</sup>, I. Haber<sup>✉</sup>, D. Sutter, and T. M. Antonsen, Jr.

*Institute for Research in Electronics and Applied Physics, University of Maryland,  
College Park, Maryland 20742, USA*



(Received 15 February 2021; accepted 16 March 2022; published 6 April 2022)

A continuous system of moment equations is introduced that models the transverse dynamics of a beam of charged particles as it passes through an arbitrary lattice of quadrupoles and solenoids in the presence of self-fields. Then, figures of merit are introduced specifying system characteristics to be optimized. The resulting model is used to optimize the parameters of the lattice elements of a flat to round transformer with self-fields, as could be applied in electron cooling. Results are shown for a case of no self-fields and two cases with self-fields. The optimization is based on a gradient descent algorithm in which the gradient is calculated using adjoint methods that prove to be very computationally efficient. Two figures of merit are studied and compared: one emphasizing radial force balance in the solenoid, the other emphasizing minimization of transverse beam energy in the solenoid.

DOI: [10.1103/PhysRevAccelBeams.25.044002](https://doi.org/10.1103/PhysRevAccelBeams.25.044002)

## I. INTRODUCTION

Beams of charged particles in accelerators are guided and manipulated by complex systems of magnets whose design optimization requires tracking the trajectories of the particles through the fields of the magnets as well as through the self-fields of the beam. Since the properties of the beam after passing through a lattice of magnetic focusing elements depend in a complicated way on the many parameters describing the lattice and self-fields, optimizing the lattice is a computationally intensive task, and historically, a large body of knowledge and techniques for addressing the optimization problem have been developed.

Traditionally, the design of these magnet systems is carried out using computer codes that calculate the beam particle phase space trajectories in the prescribed lattice of magnets. The process often reduces to the optimization of figures of merit (FOMs) in the high dimensional parameter space characterizing the lattice. Because of the large number of parameters, the efficiency of the optimization algorithm is critical. One class of optimization algorithms, based on calculating the gradient of the FOM in parameter space, becomes computationally prohibitive if the gradient is to be calculated directly (by individually varying each of the parameters) in the high dimensional parameter space.

The number of computations needed to calculate the gradient can be reduced via the introduction of adjoint techniques [1–5]. In this approach an alternate, but related, mathematical problem is introduced in which in a single (or in several) computation(s) the linear dependence of the FOM on all the parameters can be determined. The adjoint approach has previously been applied in circuit theory [1], electromagnetics [2], aerodynamics [3], and accelerator physics [4,5], as well as in other fields.

In this paper, we will illustrate the application of the adjoint approach to the design of flat-to-round (FTR) or round-to-flat (RTF) transformers as have been proposed for use in relativistic electron cooling [6–9]. As the names suggest, these transformers are systems of magnets that will convert an unmagnetized beam that has a high aspect ratio, elliptical spatial cross section, to a round beam in a solenoidal magnetic field, or vice versa. In its simplest form, this conversion is accomplished with a triplet of quadrupoles and a solenoid (see Fig. 1). Parameters that can be varied to optimize this conversion are the positions and strengths of the four magnet elements, including the orientations of the quadrupoles (11 parameters including the location of the first quadrupole). The method we will introduce also allows for a more detailed optimization in terms of the spatial profiles of the magnet elements. In addition, one might have the option to vary the parameters of the incoming beam, but we will take these to be given. The description of the beam that we will use is based on the solution of second moment equations describing the four-dimensional transverse phase space of the beam. There are ten such moments corresponding to the ten independent elements of the matrix of second moments, the so-called

\*levondov@umd.edu

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

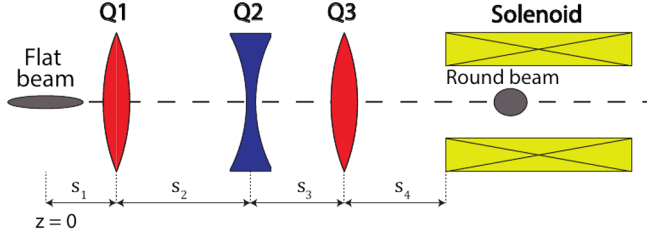


FIG. 1. General geometry of quadrupole triplet for FTR beam transformations.

sigma matrix [10]. The effect of self-fields is included in the calculations under the assumption that the beam maintains an elliptical spatial cross section of uniform density. A future publication will extend the approach to a kinetic or phase space description. The FOM will be a weighted sum of combinations of the moments at the exit of the transformer. Different combinations can be used depending on the desired application. We will find that combining the adjoint calculation with a gradient descent algorithm allows us to effectively optimize these different FOMs.

We chose FTR and RTF transformers as an example because of their multiple applications. The original motivation for the development of circular mode adapters, introduced by Derbenev in 1998 [6], was to improve the rates and efficiency of relativistic electron cooling [6–9] of hadron beams. Here, a high energy electron beam and a hadron beam are made to copropagate during which the beams interact and the hadrons are cooled. High energy electron beam sources are generally electron storage devices, in which the beams tend to become flat through radiation effects and intrabeam scattering. The cooling of the hadron beam is optimized when the electron beam interacting with the hadron beam is magnetized and matches the hadron beam size and shape, which is approximately round, and has a low effective transverse temperature [11]. The adapters provide the means for converting an unmagnetized beam to the “magnetized” state in the cooling solenoid and then converting back to the unmagnetized state at the exit of the solenoid for reinjection into the storage device.

Almost immediately after Derbenev’s initial proposal, it was realized that an RTF adapter following an electron source immersed in a solenoid field could be used to produce a flat, uncoupled beam of high enough emittance ratio and charge to potentially replace the high energy electron damping rings required in proposed designs for high energy linear colliders [12]. Experiments at Fermilab and DESY confirmed the feasibility [13–15]. Since then, the possible use of circular mode adapters for phase space manipulations has been expanded to include generating round beams at select points in circular colliders to compensate for beam-beam effects at high energy and to offset the effects of space charge and the cyclotron component of the ion beam in low energy cooler rings [9]. Additionally, their use has been

explored for injection into high intensity x-ray sources [16,17] or into dielectric laser driven accelerators [18,19].

The FTR and RTF beam transformations typically involve a solenoid and three, 45-degree skew quadrupoles, as depicted in Fig. 1. In the absence of self-fields, the parameters of the solenoid and the quadrupoles can be determined using transfer matrices. The overall length of the triplet, excluding the solenoid, scales inversely with the strength of the solenoidal field and linearly with the momentum of the beam, as reflected in the reference length

$$\beta_s = \frac{2p}{|qB|}, \quad (1)$$

where  $p$  is the momentum,  $q$  is the charge, and  $B$  is the maximum solenoidal field. In symmetrical triplets, the simplest configuration, the two outer quadrupoles have the same strength and polarity and are equidistant from the central one with opposite polarity. In asymmetric triplets [8,9], the two outer quadrupoles have differing strengths and different separations from the inner quadrupole.

Using a thin lens model for the symmetric triplet, the distance  $d$  between the outer thin quadrupoles  $Q_1$ ,  $Q_3$ , and the central one  $Q_2$ , and the inverse focal lengths,  $q_1 = q_3, q_2$ , are, following Wolske, all given in terms of  $\beta_s$  [20]:

$$d = \frac{\beta_s}{2\sqrt{1+\sqrt{2}}}, \quad q_1 = q_3 = \frac{-\sqrt{2}+1}{\beta_s}, \quad q_2 = \frac{2\sqrt{2}}{\beta_s}. \quad (2)$$

In this design, the flat beam (for FTR transformer) is incident directly on the first quadrupole which is located at  $z = 0$ , i.e., no drift space is assumed between the input beam and the first quadrupole and between the last quadrupole and the solenoid ( $s_1 = s_4 = 0$  in Fig. 1). More general equations still apply to thin lenses and no initial and final drift spaces but relax the condition of symmetry still apply to thin lenses and no initial and final drift spaces but relax the condition of symmetry, i.e.,  $q_1 \neq q_3$ ,  $s_2 \neq s_3$  [21,22]. For the triplet described by Eq. (2), an initial horizontal flat beam is transformed into a round beam with a canonical angular momentum of opposite sign to the canonical angular momentum term generated by the solenoid on transport to its center.

As an example, Burov and Danilov [7] described a symmetrical FTR triplet for an electron energy of 500 MeV and solenoid field strength  $B_s = 10.3$  kG; the triplet length is 2.23 m. Any initial triplet thin lens design calculation can use Eq. (2), following a choice of the  $\beta_s$  parameter in Eq. (1). Alternatively, it is possible to scale existing designs, as in Refs. [7] or [8], to a desired electron energy and solenoid field. We will adapt as a starting point for the test model to be used in the following scaled version of the Burov-Danilov symmetrical FTR triplet [7], which includes nonzero drift spaces  $s_1 = s_4$  (Fig. 1), an electron energy of 5 keV, and a solenoid field of 15 G ( $\beta_s = 0.319$  m).

The resulting overall length of the triplet is 0.213 m. From these initial parameters, it is straightforward to obtain a thick lens model that provides the initial conditions for the adjoint optimizations.

The initial FTR optical parameters also provide a computation of the initial beam second moments. If  $M_{\text{st}}$  is the matrix that transports an initial flat beam through the rotated triplet to the center of the solenoid, then for an initial matrix,  $\sum_0$ , the sigma matrix at the center of the solenoid is

$$\sum = M_{\text{st}} \sum_0 M_{\text{st}}^T, \quad (3)$$

where  $M_{\text{st}}^T$  is the transpose [10,23]. The generalized second moment equations derived below for the use in the adjoint optimization replace the elements of  $\sum$  in Eq. (3).

The treatments discussed so far assume no space-charge or self-field effects. Conservation laws have been developed that relate initial and final beam parameters that are independent of space charge [23]. However, the design of a system of quadrupoles needed to transform a flat beam to a cylindrically symmetric beam does depend on space charge and is addressed here. The starting point of the optimizations presented below is the flat to round parameters obtained in the absence of self-fields.

In Sec. II, we will introduce a general formulation based on a set of ten differential equations for the evolution with a distance of the second moments of the beam particle distribution function. The system allows for the profiles of the magnetic fields of the focusing elements to be treated; it allows for the orientations, locations, and the strength of the quadrupoles to be varied; and it treats self-fields in the approximation that the beam maintains an elliptical transverse cross section of constant density. We will also introduce a system of adjoint equations that allows for the efficient calculation of the change in system performance with changing parameters. In Sec. III, we will use the adjoint calculation of the gradient in parameter space to optimize the parameters of a selected model of an FTR transformer under a variety of assumptions and constraints. Here, we will also address the role of constants of motion in determining final values for beam parameters.

## II. BASIC MODEL

In this section, we first present a system of equations that describes the evolution with a distance of the second moments of a charged particle beam distribution in the presence of a combination of transverse forces. These forces include the Lorentz force of a spatially varying solenoidal (axial) magnetic field, the Lorentz force of a superposition of arbitrarily oriented quadrupole magnetic fields, and the electric and magnetic self-force due to the beam's charge and current densities. These moment equations will be used to simulate the propagation of a beam

through a system of magnets, which converts a beam with an elliptical cross section to one with a round cross section. A general figure of merit will be introduced that quantifies how successfully the shape conversion has been made. Subsequently, we will formally perturb this system by making small changes in the parameters defining the focusing forces and assess the changes by evaluating the change in the figure of merit. Such an evaluation would be used in a gradient-based optimization scheme. Then we will introduce an adjoint system of equations that will allow one to calculate compactly the changes in the system due to changes in the focusing parameters.

We note that the basic system of moment equations is well studied, and codes that solve for the moments exist in the community [24]. However, our goal is to linearize the system with respect to magnet parameters including the resulting changes to the self-fields, introduce an adjoint set of equations that efficiently calculate gradients of figures of merit, and perform optimizations. This requires presenting the base equations, their linearization, and the adjoint system for completeness. Accordingly, we have relegated much of the detail to the Appendixes.

The moments we consider are averages of products of all possible pairs of variables describing the transverse displacement of beam particles and the rate of change of the transverse displacement with distance. These moments correspond to the 16 elements of the 4 by 4 transverse sigma matrix mentioned in Eq. (3). Due to the symmetry of this matrix, only ten elements are independent. Thus, our governing system consists of ten moment evolution equations. Although there are ten separate moment equations, we show that, as expected, there are accompanying conservation laws relating the moments. The evolution with a distance of the sigma matrix is normally treated by matrix multiplication with individual matrices representing focusing elements and drift spaces. We choose to deal with differential equations for the continuous moments, as this allows us to introduce adjoint equations that include self-field effects and spatial profiles of focusing fields.

The underlying assumption that will be made in deriving the moment equations is that the beam particles' trajectories are well described by the paraxial equations of motion in which transverse forces are linear in the particles' displacements from the axis or are linearly proportional to a particle's transverse velocity. The final equations will describe the beam's evolution in the Larmor, or rotating, frame defined by the applied solenoidal magnetic field. We begin the calculation in the lab frame, transform it to the Larmor frame, and complete the derivation in the Larmor frame. Beam quantities in the lab frame can always be recovered from the Larmor frame values by the application of rotation transformations. These are given in Appendix B.

To start, we write equations for the evolution of the transverse particle displacements in a Cartesian coordinate system in the lab frame

$$x_1'' = k_x + k_\Omega y_1' + \frac{1}{2} k_\Omega' y_1, \quad (4a)$$

$$y_1'' = k_y - k_\Omega x_1' - \frac{1}{2} k_\Omega' x_1. \quad (4b)$$

Here,  $(x_1, y_1)$  are a particle's transverse displacements in the lab frame, for example,  $x_1$  is the horizontal displacement from the centerline and  $y_1$  is the vertical displacement. A prime denotes differentiation with respect to the axial coordinate  $z$ , which measures distance along the centerline. The quantity

$$k_\Omega(z) = \frac{qB_z(z)}{mc\gamma v_z} \quad (5)$$

is the spatial gyration rate due to the axial (solenoidal) magnetic field. The quantities in Eq. (5) are as follows:  $q$  and  $m$  are the particle charge and mass,  $B_z$  is the axial component of the solenoidal magnetic field strength,  $c$  is the speed of light,  $v_z$  is the axial velocity, and  $\gamma = (1 - v_z^2/c^2)^{-1/2}$ . (If the beam consists of electrons,  $q = -e$ , and if  $B_z > 0$ , then  $k_\Omega < 0$ .) The quantities  $k_x$ ,  $k_y$  represent the transverse forces from the quadrupoles and space charge and will be defined subsequently. The second terms in Eqs. (4a) and (4b) describe the Lorentz force due to the axial component of the solenoidal field and the transverse velocity, and the third terms describe the Lorentz force due to the radial component of the solenoidal field and the axial velocity.

Our next step is to transform the variables to their Cartesian representation in a rotating frame

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \quad (6)$$

where  $\phi(z)$  is an axially dependent rotation phase that we choose to satisfy

$$\phi' = -\frac{k_\Omega}{2} \quad (7)$$

and which defines the Larmor frame. Substituting Eqs. (6) and (7) into (4a) and (4b) and multiplying by the inverse of the rotation matrix appearing in Eq. (6) result in the system

$$\begin{pmatrix} x'' \\ y'' \end{pmatrix} + \left(\frac{k_\Omega}{2}\right)^2 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{bmatrix} \begin{pmatrix} k_x \\ k_y \end{pmatrix}. \quad (8)$$

We wish to obtain equations for the average of products of the four variables  $(x, x', y, y')$ . There are 16 ordered products. However, the order does not matter, leaving ten independent products. We choose combinations of these products that distinguish  $x$ - $y$  symmetric and nonsymmetric motion. For spatial moments, we choose

$$\mathbf{Q} = \begin{pmatrix} Q_+ \\ Q_- \\ Q_x \end{pmatrix} = \begin{pmatrix} \langle x^2 + y^2 \rangle / 2 \\ \langle x^2 - y^2 \rangle / 2 \\ \langle xy \rangle \end{pmatrix}. \quad (9)$$

Here, the angle brackets imply average over the beam distribution function. Accompanying these spatial moments are momentum like moments

$$\mathbf{P} = \frac{d}{dz} \mathbf{Q} = \begin{pmatrix} P_+ \\ P_- \\ P_x \end{pmatrix} = \begin{pmatrix} \langle xx' + yy' \rangle \\ \langle xx' - yy' \rangle \\ \langle yx' + xy' \rangle \end{pmatrix}. \quad (10)$$

The angular momentum completes this group of momentum like moments,

$$L = \langle xy' - yx' \rangle. \quad (11)$$

Note, that  $L$  is the angular momentum in the rotating frame. The group of ten moments is completed by three energy-like moments

$$\mathbf{E} = \begin{pmatrix} E_+ \\ E_- \\ E_x \end{pmatrix} = \begin{pmatrix} \langle x'^2 + y'^2 \rangle \\ \langle x'^2 - y'^2 \rangle \\ 2\langle y'x' \rangle \end{pmatrix}. \quad (12)$$

The evolution of these moments is determined by the following system of equations:

$$\frac{d}{dz} \mathbf{Q} = \mathbf{P}, \quad (13a)$$

$$\frac{d}{dz} \mathbf{P} = \mathbf{E} + \mathbf{O} \cdot \mathbf{Q}, \quad (13b)$$

$$\frac{d}{dz} \mathbf{E} = \mathbf{O} \cdot \mathbf{P} + N\mathbf{L}, \quad (13c)$$

$$\frac{d}{dz} L = -N^\dagger \cdot \mathbf{Q}. \quad (13d)$$

Here, the matrix  $\mathbf{O}$  and vector  $N$  are defined as follows:

$$\begin{aligned} \mathbf{O} = & -\frac{k_\Omega^2}{2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + 2 \sum_{\text{quads}} K_q \begin{bmatrix} 0 & c_q & -s_q \\ c_q & 0 & 0 \\ -s_q & 0 & 0 \end{bmatrix} \\ & + \frac{\Lambda}{Q_\Delta} \begin{bmatrix} 1 & c_\alpha & s_\alpha \\ c_\alpha & 1 & 0 \\ s_\alpha & 1 & 1 \end{bmatrix} \end{aligned} \quad (14)$$



and

$$N = 2 \sum_{\text{quads}} K_q \begin{pmatrix} 0 \\ s_q \\ c_q \end{pmatrix} - \frac{\Lambda}{Q_\Delta} \begin{pmatrix} 0 \\ s_\alpha \\ -c_\alpha \end{pmatrix}. \quad (15)$$

The matrices defined in Eqs. (14) and (15) are continuous functions of axial distance,  $z$ . The first matrix in Eq. (14) is due to the solenoidal field, the second is due to the quadrupoles, and the third matrix is due to the self-fields. The first vector in Eq. (15) is due to the quadrupoles and the second is due to the self-fields. In Eqs. (14) and (15), the following expressions and notations have been introduced. Each quadrupole magnet has field strength given in the lab frame by

$$\begin{aligned} B_{qx} &= B'_q(z) [\sin(2\psi_q)x_1 - \cos(2\psi_q)y_1] \\ B_{qy} &= -B'_q(z) [\cos(2\psi_q)x_1 + \sin(2\psi_q)y_1], \end{aligned} \quad (16)$$

where  $B'_q(z)$  defines the strength and axial profile of the quadrupole field and the angle  $\psi_q$  defines the orientation. As shown in Appendix A, this leads to the displayed contributions to the  $\mathbf{O}$  matrix and  $\mathbf{N}$  vector where

$$K_q(z) = \frac{qB'_q(z)}{mc\gamma v_z}. \quad (17)$$

If  $K_q > 0$ , the magnet will be defocusing in the lab frame when  $\theta = \psi_q, \psi_q + \pi$ , and focusing when  $\theta = \psi_q + \pi/2, \psi_q - \pi/2$ . Here,  $\theta$  is the angle a particle's transverse displacement makes with respect to the  $x_1$  axis. If  $K_q < 0$ , these are reversed. In the Larmor frame, the orientation of the quadrupole is characterized by the variables

$$(s_q, c_q) = [\sin(2\phi - 2\psi_q), \cos(2\phi - 2\psi_q)], \quad (18)$$

where  $\phi(z)$  is the angle resulting from the integration of Eq. (5). We note that with one group of quadrupoles placed separately from the solenoid, it is always possible to define the phase such that it vanishes in the region of the quadrupoles.

The self-field term is calculated in the Larmor frame in Appendix B. Here, it is assumed that the beam charge density distribution is spatially uniform inside an ellipse with differing major and minor radii and tilted at an angle in the Larmor frame. The values of the major and minor radii and the angle are determined by the three spatial moments ( $Q_+$ ,  $Q_-$ ,  $Q_x$ ). The result is the following:

$$Q_\Delta = [Q_+^2 - (Q_-^2 + Q_x^2)]^{1/2}, \quad (19a)$$

$$c_\alpha = -\frac{Q_-}{(Q_+ + Q_\Delta)}, \quad (19b)$$

$$s_\alpha = -\frac{Q_x}{(Q_+ + Q_\Delta)}. \quad (19c)$$

The strength of the self-fields is measured by the beam current parameter

$$\Lambda = \frac{cqZ_0I}{4\pi m\nu_z^3\gamma_0^3}, \quad (20)$$

with  $Z_0 = 377$  Ohms and  $I$  is the beam current.

Equations (13a)–(13d) can be combined to show the following conservation relation:

$$\frac{d}{dz} \left[ \mathbf{E} \cdot \mathbf{Q} + \frac{L^2}{2} - \frac{1}{2} \mathbf{P} \cdot \mathbf{P} \right] = 0. \quad (21)$$

This does not represent the conservation of a four-dimensional emittance,  $\epsilon_{4D} = \det[\Sigma]$ . Rather, it is the conserved quantity  $I_2(\Sigma)$  given by Eq. (12) of [24] and attributed to [25]. The implications of these constants will be discussed subsequently.

We compare the solution of the moment equations with the moments calculated by the PIC code Warp [26] in Fig. 2. For this simulation, we consider a symmetric triplet with parameters selected to convert a flat beam to a round one in the presence of a space charge. The space charge is evaluated along the entire path of the beam. Parameters for the simulation are given in Table I. The large pipe radius in the list of parameters is chosen to minimize the contributions from image charge forces on the simulations in order to compare the results with the moment equations.

Plotted in Fig. 2 are the computed values of  $x_{\text{rms}}^2 = \langle x^2 \rangle = Q_+ + Q_-$  and  $y_{\text{rms}}^2 = \langle y^2 \rangle = Q_+ - Q_-$  as functions of  $z$  for three values of beam current. The solutions of the moment equations are solid lines, and the PIC values are

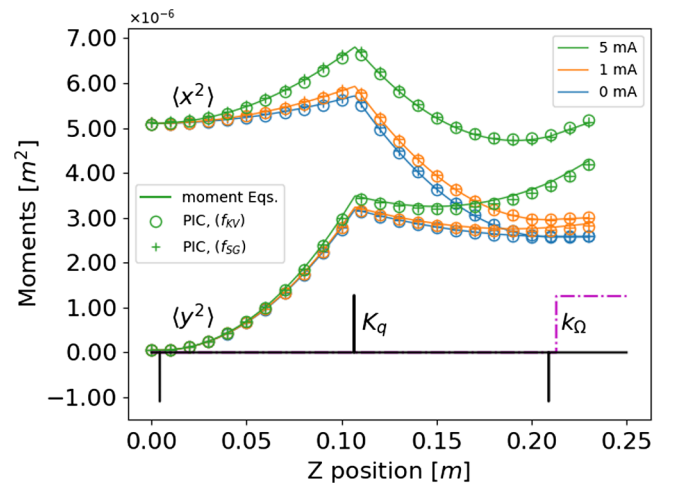


FIG. 2. Comparison of solutions of the moment equations (solid lines) with solutions from the PIC code Warp (circular and cross symbols). PIC simulations were run with an initial Kapchinskij-Vladimirskij (K-V) beam distribution (circles) [27] as well as a semi-Gaussian (crosses) distribution using 40K particles. Comparisons are for beams with 0, 1, and 5 mA current.

TABLE I. PIC code simulation and FTR lattice parameters.

Parameter	Value
Beam distribution	K-V, semi-Gaussian
Particle number	40 000
Longitudinal step size	0.1 [mm]
Grid cell number	2048 × 2048
Pipe radius	250.0 [mm]
Beam energy	5 keV
Quadrupole 1 location	0.0043 [m]
Quadrupole 2 location	0.1066 [m]
Quadrupole 3 location	0.2090 [m]
Quadrupole 1,3 strength	-18.236 [T/m]
Quadrupole 2 strength	21.364 [T/m]
Quadrupole 1,2,3 length	0.1 [mm]
Solenoid location	0.2133
Solenoid strength	15 × 10 <sup>-4</sup> [T]
Initial $\sqrt{\langle x^2 \rangle}, \sqrt{\langle y^2 \rangle}$	2.2581, 0.2258 [mm]
Initial rms $\epsilon_x, \epsilon_y^a$	16, 0.16 [mm mrad]

<sup>a</sup>Where  $\epsilon_x = \sqrt{2\langle x^2 \rangle \langle x'^2 \rangle - \langle xx' \rangle^2}$ ,  $\epsilon_y = \sqrt{2\langle y^2 \rangle \langle y'^2 \rangle - \langle yy' \rangle^2}$ .

shown as circles and crosses. For the PIC case, two different distributions are used: a K-V [27] distribution function:

$$f_{KV}(x, x', y, y') = \frac{1}{4\pi^2 A} \delta\left(\frac{x^2}{\langle x^2 \rangle} + \frac{y^2}{\langle y^2 \rangle} + \frac{x'^2}{\langle x'^2 \rangle} + \frac{y'^2}{\langle y'^2 \rangle} - 4\right), \quad (22)$$

where  $A = \sqrt{\langle x^2 \rangle \langle y^2 \rangle \langle x'^2 \rangle \langle y'^2 \rangle}$  and  $\delta$  is the Dirac delta function, as well as a semi-Gaussian distribution function:

$$f_{SG}(x, x', y, y') = \frac{1}{8\pi^2 A} \exp\left[-\frac{1}{2}\left(\frac{x'^2}{\langle x'^2 \rangle} + \frac{y'^2}{\langle y'^2 \rangle}\right)\right], \quad (23a)$$

when  $x^2/\langle x^2 \rangle + y^2/\langle y^2 \rangle < 4$  otherwise  $f_{SG} = 0$ . The semi-Gaussian distribution is uniform in space over the cross section of the beam and Gaussian in velocity, representing a distribution function from a thermionic cathode. As can be seen in Fig. 2, in the zero current case, the beam becomes round at the end of the triplet, whereas in the presence of self-fields, the roundness is spoiled. In the next section, we will show how the optimizations of the triplet parameters can recover the round state of the beam.

We now imagine that Eqs. (13a)–(13d) have been solved for a given set of initial conditions,  $(\mathbf{Q}, \mathbf{P}, \mathbf{E}, L)|_{z_i}$ , and magnetic field parameters and profiles, which we label with a vector  $\mathbf{a}$ . An assessment of the configuration can be made on the basis of a figure of merit that depends on the final values of the moments,

$$F(\mathbf{Q}, \mathbf{P}, \mathbf{E}, L, \mathbf{a})|_{z_f}. \quad (24)$$

For a flat to round transition, the figure of merit might be a function which when minimized forces the components

$Q_x$ ,  $Q_y$ , and  $\mathbf{P}$  all to zero, signifying the beam is round and its radius will remain constant. Alternatively, the FOM might force the beam to be cylindrically symmetric in the solenoid with no rotation or minimum second moment  $E_{+, \text{Lab}} = \langle x'^2 + y'^2 \rangle$ . This will be discussed in more detail in the next section. We refer to a general solution just described (not necessarily the minimizing one) as a base case.

If a small change is made in the focusing parameters,  $\delta\mathbf{a}$ , there will be a commensurately small change in the moments as functions of  $z$  and a small change in the figure of merit from the corresponding base case. The small changes in the moments will satisfy a linearization of Eqs. (13)–(15). We denote the small changes with a superscript  $(X)$ , and we refer to them as true changes. These satisfy

$$\frac{d}{dz} \delta\mathbf{Q}^{(X)} = \delta\mathbf{P}^{(X)} \quad (25a)$$

$$\frac{d}{dz} \delta\mathbf{P}^{(X)} = \delta\mathbf{E}^{(X)} + \mathbf{O} \cdot \delta\mathbf{Q}^{(X)} + \delta\mathbf{O}^{(X)} \cdot \mathbf{Q} \quad (25b)$$

$$\frac{d}{dz} \delta\mathbf{E}^{(X)} = \mathbf{O} \cdot \delta\mathbf{P}^{(X)} + N\delta L^{(X)} + \delta\mathbf{O}^{(X)} \cdot \mathbf{P} + \delta N^{(X)} L \quad (25c)$$

$$\frac{d}{dz} \delta L^{(X)} = -N^\dagger \cdot \delta\mathbf{Q}^{(X)} - \delta N^{\dagger(X)} \cdot \mathbf{Q}. \quad (25d)$$

The resulting change in the figure of merit is

$$\delta F^{(X)} = \left[ \delta\mathbf{Q}^{(X)} \frac{\partial F}{\partial \mathbf{Q}} + \delta\mathbf{P}^{(X)} \frac{\partial F}{\partial \mathbf{P}} + \delta\mathbf{E}^{(X)} \frac{\partial F}{\partial \mathbf{E}} + \delta L^{(X)} \frac{\partial F}{\partial L} \right]_{z_f} + \delta\mathbf{a} \frac{\partial F}{\partial \mathbf{a}} \Big|_{Q, P, E, L}. \quad (26)$$

Again, the superscript  $(X)$  signifies that it is a true perturbation, i.e., the result of changing the focusing system: the solenoidal and quadrupole magnetic fields.

The change in the FOM due to changes in parameters has two types of contribution: contributions due to the explicit dependence of the FOM on parameters, the last term on the right of Eq. (26), and contributions due to the implicit dependence on parameters due to changes in the moments.

The perturbed matrices  $\mathbf{O}^{(X)}, \mathbf{N}^{(X)}$  appearing in Eqs. (25b)–(25d) are the result of perturbations to matrices  $\mathbf{O}$  and  $\mathbf{N}$ . These matrices are perturbed due to two effects. One effect is the set of perturbations due to changes in the solenoidal and quadrupole magnetic fields that appear explicitly in the definitions of  $\mathbf{O}$  and  $\mathbf{N}$ . These are being varied in order to optimize the configuration. A second effect is due to changes in the self-field terms. These changes are expressed through the changes in the spatial moments,  $\delta\mathbf{Q}^{(X)}$  in the definitions of Eqs. (19a)–(19c). We write the changes to these matrices as a sum of contributions from explicit changes in the focusing configuration and implicit changes in the self-fields

$$\delta \mathbf{O}^{(X)} \cdot \mathbf{Q} = \delta \mathbf{O}_{Q,B}^{(X)} \cdot \mathbf{Q} + \mathbf{M}_Q \cdot \delta \mathbf{Q}^{(X)}, \quad (27a)$$

$$\delta \mathbf{O}^{(X)} \cdot \mathbf{P} = \delta \mathbf{O}_{Q,B}^{(X)} \cdot \mathbf{P} + \mathbf{M}_P \cdot \delta \mathbf{Q}^{(X)}, \quad (27b)$$

$$\delta \mathbf{N}^{(X)} = \delta \mathbf{N}_{Q,B}^{(X)} + \mathbf{M}_N \cdot \delta \mathbf{Q}^{(X)}. \quad (27c)$$

Here, the variables with subscripts  $Q, B$  are the contributions to the changes from the explicit dependence of the matrices on the magnetic focusing parameters. The matrices  $\mathbf{M}_Q, \mathbf{M}_P, \mathbf{M}_N$  are the contributions from the self-field terms in the matrices  $\mathbf{M}, \mathbf{N}$ , which depend on  $\mathbf{Q}$  as given by Eqs. (14), (15), and (19). Formulas for these matrices are given in Appendix B.

Changing each parameter will result in a change to the figure of merit. This can be thought of as defining a gradient of the figure of merit in parameter space. Direct evaluation of the change in the figure of merit would require resolving the system of ten coupled equations  $M_p$  times where  $M_p$  is the number of parameters that could be varied. We can avoid this burden by introducing an adjoint system of equations. Consider a second linear perturbation to the base case, denoted by a superscript  $(Y)$ . These are the adjoint perturbation equations that we will actually solve:

$$\frac{d}{dz} \delta \mathbf{Q}^{(Y)} = \delta \mathbf{P}^{(Y)}, \quad (28a)$$

$$\frac{d}{dz} \delta \mathbf{P}^{(Y)} = \delta \mathbf{E}^{(Y)} + \mathbf{O} \cdot \delta \mathbf{Q}^{(Y)}, \quad (28b)$$

$$\frac{d}{dz} \delta \mathbf{E}^{(Y)} = \mathbf{O} \cdot \delta \mathbf{P}^{(Y)} + \mathbf{N} \delta L^{(Y)} + \delta \dot{\mathbf{E}}^{(Y)}, \quad (28c)$$

$$\frac{d}{dz} \delta L = -\mathbf{N}^\dagger \cdot \delta \mathbf{Q}^{(Y)}. \quad (28d)$$

In Eqs. (28)–(28d), we use the base case (unperturbed)  $\mathbf{O}$  and  $\mathbf{N}$  matrices. However, we have added a term in Eq. (28c),  $\delta \dot{\mathbf{E}}^{(Y)}$ , which will be chosen, as described subsequently, so as to achieve a desired cancellation of the self-field contributions to the changes in  $\mathbf{O}$  and  $\mathbf{N}$ .

We next form the following combination of true  $(X)$  and adjoint  $(Y)$  variables,

$$\epsilon \equiv \delta \mathbf{P}^{(Y)} \cdot \delta \mathbf{P}^{(X)} - \delta \mathbf{Q}^{(X)} \cdot \delta \mathbf{E}^{(Y)} - \delta \mathbf{Q}^{(Y)} \cdot \delta \mathbf{E}^{(X)} - \delta L^{(Y)} \delta L^{(X)} \quad (29)$$

and differentiate it with respect to  $z$ . We evaluate the individual terms in the derivative of Eq. (29) using the product rule and Eqs. (25) and (28). Noting numerous cancellations, we arrive at

$$\begin{aligned} \frac{d}{dz} \epsilon &= \delta \mathbf{P}^{(Y)} \cdot \delta \mathbf{O}^{(X)} \cdot \mathbf{Q} + \delta L^{(Y)} \delta \mathbf{N}^{(X)} \cdot \mathbf{Q} - \delta \mathbf{Q}^{(X)} \cdot \delta \dot{\mathbf{E}}^{(Y)} \\ &\quad - \delta \mathbf{Q}^{(Y)} \cdot \delta \mathbf{O}^{(X)} \cdot \mathbf{P} - \delta \mathbf{Q}^{(Y)} \cdot \delta \mathbf{N}^{(X)} L. \end{aligned} \quad (30)$$

Let us now consider the quantities  $\delta \mathbf{N}^{(X)}$  and  $\delta \mathbf{O}^{(X)}$  appearing in Eq. (30). We write them according to the separation defined in Eq. (25). We then have

$$\begin{aligned} \frac{d}{dz} \epsilon &= \delta \mathbf{P}^{(Y)} \cdot \mathbf{M}_Q \cdot \delta \mathbf{Q}^{(X)} + \delta \mathbf{P}^{(Y)} \cdot \delta \mathbf{O}_{Q,B}^{(X)} \cdot \mathbf{Q} \\ &\quad - \delta \mathbf{Q}^{(X)} \cdot \delta \dot{\mathbf{E}}^{(Y)} + \delta L^{(Y)} \mathbf{Q} \cdot \mathbf{M}_N \cdot \delta \mathbf{Q}^{(X)} \\ &\quad + \delta L^{(Y)} \mathbf{Q} \cdot \delta \mathbf{N}_{Q,B}^{(X)} - \delta \mathbf{Q}^{(Y)} \cdot \delta \mathbf{O}_{Q,B}^{(X)} \cdot \mathbf{P} \\ &\quad - \delta \mathbf{Q}^{(Y)} \cdot \mathbf{M}_P \cdot \delta \mathbf{Q}^{(X)} - \delta \mathbf{Q}^{(Y)} \cdot \delta \mathbf{N}_{Q,B}^{(X)} L \\ &\quad - \delta \mathbf{Q}^{(Y)} \cdot \mathbf{M}_N \cdot \delta \mathbf{Q}^{(X)} L. \end{aligned} \quad (31)$$

The next step is to pick  $\delta \dot{\mathbf{E}}^{(Y)}$  to cancel all terms proportional to the unknown  $\delta \mathbf{Q}^{(X)}$  in Eq. (31),

$$\begin{aligned} \delta \dot{\mathbf{E}}^{(Y)} &= \delta \mathbf{P}^{(Y)} \cdot \mathbf{M}_Q + \delta L^{(Y)} \mathbf{Q} \cdot \mathbf{M}_N \\ &\quad - \delta \mathbf{Q}^{(Y)} \cdot \mathbf{M}_P - \delta \mathbf{Q}^{(Y)} \cdot \mathbf{M}_N L. \end{aligned} \quad (32)$$

This relation is taken to define the quantity  $\delta \dot{\mathbf{E}}^{(Y)}$  that enters the adjoint equations in Eq. (28c). This leaves for the adjoint relation

$$\begin{aligned} \frac{d}{dz} \epsilon &= \delta \mathbf{P}^{(Y)} \cdot \delta \mathbf{O}_{Q,B}^{(X)} \cdot \mathbf{Q} + \delta L^{(Y)} \mathbf{Q} \cdot \delta \mathbf{N}_{Q,B}^{(X)} \\ &\quad - \delta \mathbf{Q}^{(Y)} \cdot \delta \mathbf{O}_{Q,B}^{(X)} \cdot \mathbf{P} - \delta \mathbf{Q}^{(Y)} \cdot \delta \mathbf{N}_{Q,B}^{(X)} L. \end{aligned} \quad (33)$$

Integrating over  $z$  from initial to final point

$$\begin{aligned} &(\delta \mathbf{P}^{(Y)} \cdot \delta \mathbf{P}^{(X)} - \delta \mathbf{Q}^{(X)} \cdot \delta \mathbf{E}^{(Y)} - \delta \mathbf{Q}^{(Y)} \cdot \delta \mathbf{E}^{(X)} - \delta L^{(Y)} \delta L^{(X)}) \Big|_{z=z_i}^{z=z_f} \\ &= \int_{z_i}^{z_f} dz \{ \delta \mathbf{P}^{(Y)} \cdot \delta \mathbf{O}_{Q,B}^{(X)} \cdot \mathbf{Q} + \delta L^{(Y)} \mathbf{Q} \cdot \delta \mathbf{N}_{Q,B}^{(X)} - \delta \mathbf{Q}^{(Y)} \cdot \delta \mathbf{O}_{Q,B}^{(X)} \cdot \mathbf{P} - \delta \mathbf{Q}^{(Y)} \cdot \delta \mathbf{N}_{Q,B}^{(X)} L \}. \end{aligned} \quad (34)$$

We are now in a position to describe the utility of the adjoint approach. Suppose we integrate the base case equations, Eqs. (13a)–(13d), forward from  $z_i$  to  $z_f$ , and

we then integrate the adjoint equations, Eqs. (28a)–(28d), backward in  $z$ , starting at  $z_f$ . Further, suppose we take the conditions on the adjoint variables at  $z_f$  to be

$$\delta \mathbf{P}^{(Y)}(z_f) = \left. \frac{\partial F}{\partial \mathbf{P}} \right|_{z_f}, \quad (35a)$$

$$-\delta \mathbf{E}^{(Y)}(z_f) = \left. \frac{\partial F}{\partial \mathbf{Q}} \right|_{z_f}, \quad (35b)$$

$$-\delta \mathbf{Q}^{(Y)}(z_f) = \left. \frac{\partial F}{\partial \mathbf{E}} \right|_{z_f}, \quad (35c)$$

$$\delta L^{(Y)}(z_f) = \left. \frac{\partial F}{\partial L} \right|_{z_f}, \quad (35d)$$

where  $F$  is the figure of merit. Then Eqs. (26) and (34) give for the change in the figure of merit due to perturbation ( $X$ )

$$\begin{aligned} \delta F^{(X)} = & (\delta \mathbf{P}^{(Y)} \cdot \delta \mathbf{P}^{(X)} - \delta \mathbf{Q}^{(X)} \cdot \delta \mathbf{E}^{(Y)} \\ & - \delta \mathbf{Q}^{(Y)} \cdot \delta \mathbf{E}^{(X)} - \delta L^{(Y)} \delta L^{(X)})|_{z=z_i} \\ & + \int_{z_i}^{z_f} dz \{ \delta \mathbf{P}^{(Y)} \cdot \delta \mathbf{O}_{Q,B}^{(X)} \cdot \mathbf{Q} + \delta L^{(Y)} \mathbf{Q} \cdot \delta \mathbf{N}_{Q,B}^{(X)} \} \\ & + \int_{z_i}^{z_f} dz \{ -\delta \mathbf{Q}^{(Y)} \cdot \delta \mathbf{O}_{Q,B}^{(X)} \cdot \mathbf{P} - \delta \mathbf{Q}^{(Y)} \cdot \delta \mathbf{N}_{Q,B}^{(X)} L \}. \end{aligned} \quad (36)$$

Equation (36) can now be used to evaluate the changes in the figure of merit ( $F$ ) due to arbitrary, small changes in the initial conditions of the true solution as contained in the first term on the right of Eq. (36). Or it can be used to evaluate changes in the FOM  $F$  due to arbitrary, small changes in the strength or profile of the focusing magnetic fields as contained in the integral term on the right of Eq. (36). These changes can be evaluated by a single integral once the adjoint solution is found without having to resolve the coupled system of ten equations for each possible change in parameters.

We illustrate the utility of the adjoint method in calculating the gradient of an FOM with respect to variations of parameters in Fig. 3. Here we show the dependence on parameters of a particular FOM that will be introduced in the next section. The dependence on parameter values is calculated two ways: first by directly varying the parameter and second by using Eq. (36) to compute the gradient. The top plot in Fig. 3 shows the variation of the FOM with respect to changes in the strengths of the three quadrupoles, while the bottom plot in Fig. 3 shows the dependence on the solenoidal field. The solid straight lines have slopes that are predicted by the adjoint method, while the dashed curved lines show the effect of directly varying the parameter and plotting the change in the FOM. As can be seen, the solid lines are tangent to the dashed lines for vanishing perturbations indicating that the gradient has been calculated correctly. This method of calculating the gradient will be used in the next section to optimize an FTR transformer.

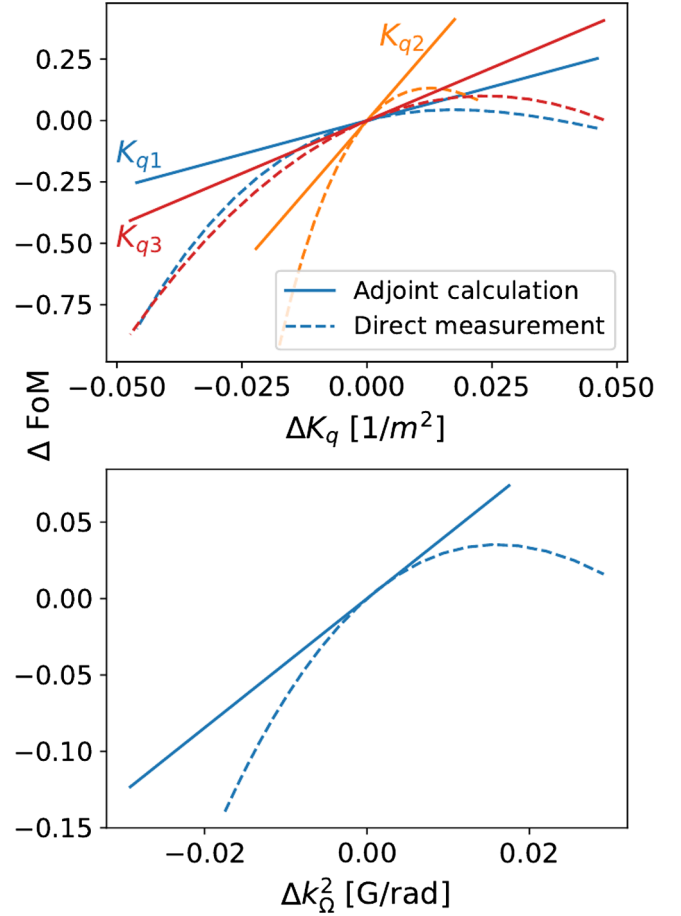


FIG. 3. Comparison of changes in a figure of merit calculated using the adjoint method (straight lines) and calculated directly (curved lines).

Before using the adjoint method to optimize a configuration, we discuss the implications of the two conservation laws following Eq. (21). The first conserved quantity we consider is the four-dimensional emittance that is conserved as a consequence of the incompressible, Hamiltonian nature of the particle motion in the phase space  $(x, x', y, y')$ . This leads to the constancy of  $\epsilon_{4D} = \det[\Sigma]$ . If we evaluate this quantity upstream for the entering flat beam, assuming all cross moments involving products of  $x$  and  $y$  are zero, and evaluate it in the solenoid assuming the beam has become cylindrically symmetric (all cross moments are zero and  $x$  and  $y$  moments are equal), we find

$$\begin{aligned} 4\text{Det}[\Sigma_T] &= 4\langle xx \rangle_u \langle yy \rangle_u \langle \dot{x} \dot{x} \rangle_u \langle \dot{y} \dot{y} \rangle_u \\ &= (Q_{+s} E_{+s} - \frac{1}{2} P_{+s}^2 - \frac{1}{2} L^2)^2 \\ &\equiv \epsilon_{ts}^2. \end{aligned} \quad (37)$$

Here the subscript  $u$  refers to upstream and subscript  $s$  refers to the solenoid. Similarly, evaluating the constant defined in Eq. (21), first upstream and then in the solenoid,



$$\begin{aligned} (\langle xx \rangle_u \langle \dot{x} \dot{x} \rangle_u + \langle yy \rangle_u \langle \dot{y} \dot{y} \rangle_u) &= Q_{+s} E_{+s} + \frac{L^2}{2} - \frac{1}{2} P_{+s}^2 \\ &= \epsilon_{ts} + L^2. \end{aligned} \quad (38)$$

Combining these, we can solve for the value of  $L$  in the solenoid [23]

$$L = \pm \left( \sqrt{\langle xx \rangle_u \langle \dot{x} \dot{x} \rangle_u} - \sqrt{\langle yy \rangle_u \langle \dot{y} \dot{y} \rangle_u} \right). \quad (39)$$

Here the sign of  $L$  is determined by the orientation of the quadrupoles that turn the flat beam into a cylindrically symmetric beam. They can be oriented at  $\pm 45^\circ$ . This follows by noting that changing the orientation of the quadrupoles from  $+$  to  $-45^\circ$  is equivalent to reversing the sign of the magnetic fields in the quadrupoles. If in addition, one reverses the sign of the solenoidal magnetic field, a configuration that successfully transforms a flat beam to a round beam with one sign of angular momentum, will transform a flat beam to a round beam with the opposite sign of angular momentum.

We have verified that relations Eqs. (37) and (38) are satisfied by solutions of our moment equations when they apply. That is when the beam is cylindrically symmetric in the solenoid. We emphasize that the strength, location, and orientation of the quadrupoles need to be adjusted in the presence of space charge to make the beam cylindrically symmetric in the solenoid.

### III. OPTIMIZATION

The goal of the optimization is to minimize a given FOM based on the values of the moments at a location ( $z = z_f$ ) inside the solenoid using the adjoint approach. We first choose the following function of the beam moments to be our FOM:

$$\begin{aligned} F &= \frac{1}{2} \left[ |\mathbf{P}|^2 + k_0^2 (Q_-^2 + Q_x^2) + k_0^{-2} (E_-^2 + E_x^2) \right. \\ &\quad \left. \times k_0^{-2} (E_+ - \frac{1}{2} k_\Omega^2 Q_+ + \Lambda)^2 + (2E_+ Q_+ - L^2)^2 \right]. \end{aligned} \quad (40)$$

It is a sum of terms quadratic in the moments, each of which should be as small as possible in an FTR transformer that leads to a matched beam in the solenoid. The quantity  $k_0$  is introduced so that each term has the same units. Here  $k_0$  is a scaling parameter approximately equal to the inverse of the lattice length. This makes all the terms roughly comparable in magnitude. The choice of each term is made as follows. We would like the beam in the solenoid to be round,  $Q_- = Q_x = 0$ , and we would like all the nonzero moments to be independent of  $z$ . This implies for the spatial moments  $d\mathbf{Q}/dz = \mathbf{P} = 0$ . If we set the components of  $d\mathbf{P}/dz = 0$ , we find from Eq. (13b),  $E_- = E_x = 0$  is required as well as  $E_+ - k_\Omega^2 Q_+/2 + \Lambda = 0$  (radial force balance). Finally, the

last term is designed to force the trajectories to be as laminar as possible and the rotation to be rigid in the solenoid. It can be motivated as follows. If the solenoid particles have a radially varying mean rotation rate  $\Omega(r)$  in the Larmor frame, then

$$E_{+0} = \langle \delta x'^2 + \delta y'^2 \rangle + \langle \Omega^2 r^2 \rangle, \quad (41)$$

where  $\delta x', \delta y'$  are deviations from the mean rotation. The angular momentum in this case is  $L = \langle \Omega r^2 \rangle$ . We thus have, by virtue of the Schwarz inequality,

$$E_+ \geq \langle \Omega^2 r^2 \rangle \geq \langle \Omega r^2 \rangle^2 / \langle r^2 \rangle = L^2 / (2Q_+). \quad (42)$$

The second inequality becomes an equality if the rotation is rigid.

The FOM in Eq. (40) is a general function in that it meets the requirements for an FTR transformation. In practice, an FOM will also contain extra constraint conditions pertinent to the system being optimized. This can include beam size limits based on a physical aperture size, magnet sizes, locations based on available beamline space, etc. The extra constraints can be incorporated within the FOM or as constraints on the input parameters used for the optimization.

We start with a design based on the symmetric triplet with parameters given by Eqs. (1) and (2). It can be shown from analytic solutions of the moment equations in the thin lens approximations that the entrance and exit conditions in Table II can be achieved for a symmetric triplet oriented at  $45^\circ$  with respect to the long dimension of the incident flat beam. The strengths of the quadrupoles needed to achieve these parameters are given by

$$\begin{aligned} d \int_{-\infty}^{\infty} 2K_{1,3}(z) s_{1,3} dz &= -\sqrt{1 + \sqrt{2}} \\ d \int_{-\infty}^{\infty} 2K_2(z) s_2 dz &= 2\sqrt{2} / \sqrt{1 + \sqrt{2}}. \end{aligned} \quad (43)$$

TABLE II. Entrance and exit conditions for the moments in a thin lens approximation with a symmetric triplet and quadrupoles oriented at  $45^\circ$  rotations with respect to the longitudinal direction.

$z = 0$	$z = z_f = 2d$
$Q_+ = Q_+(0)$	$Q_+ = Q_+(0)$
$Q_- = Q_-(0)$	$Q_- = 0$
$Q_x = 0$	$Q_x = 0$
$P_+ = 0$	$P_+ = 0$
$P_- = 0$	$P_- = 0$
$P_x = 0$	$P_x = 0$
$E_+ = Q_+(0)/[d^2(2 + 2\sqrt{2})]$	$E_+ = Q_+(0)/[d^2(2 + 2\sqrt{2})]$
$E_- = Q_-(0)/[d^2(2 + 2\sqrt{2})]$	$E_- = 0$
$E_x = 0$	$E_x = 0$
$L = 0$	$L = 2Q_-(0)/(d\sqrt{1 + \sqrt{2}})$

The values given in Eq. (43) are in agreement with those expressed in Eq. (2). It is interesting to note that for the 45-degree triplet in the absence of space charge, the ten moments break into three independent groups, each with a conserved quantity

$$E_- Q_- + L^2/2 - P_-^2/2 = J_1$$

$$(E_+ \pm E_x)(Q_- \pm Q_x) - (P_+ \pm P_x)^2/2 = J_{\pm}. \quad (44)$$

Also note there are seven quantities that vanish at the exit but only four adjustable parameters, the values of  $E_{\pm}$  and the strengths of the inner and outer quadrupoles displayed in Eq. (43). The ability to satisfy what seems like three extra conditions then results from the symmetry leading to Eq. (44).

We now optimize the parameters describing the magnetic fields by using the FOM in Eq. (40). We consider the FOM to be a function of the elements of a list of parameters  $\mathbf{a}$ . The elements of the list include the strengths, locations, and orientations of the quadrupoles as well as the strength and location of the solenoid. In the case of no self-fields, the strength and location of the solenoid, along with the beam energy, would be enough to optimize the system. With self-fields, the initial matching of the beam through the three quadrupoles no longer becomes trivial and requires the quadrupole parameters to be included in the optimization, resulting in a total of 11 parameters. We use a simple steepest descent algorithm in which we calculate the gradient of  $F$  in the space of parameters  $\mathbf{a}$ ,  $\nabla_{\mathbf{a}} F(\mathbf{a})$  and then adjust the values of the parameters by moving from the current set of parameters along the line of steepest descent.

Generally, we take the parameters  $\mathbf{a}$  to be a set of dimensionless multipliers; each one multiplying the initial value of its corresponding dimensional quantity. To calculate the gradient, we set the conditions at  $z = z_f$  for the adjoint variables in Eqs. (35a)–(35d) according to the current values of the parameters,  $\mathbf{a}_n$ . This then allows the solution for the adjoint variables and the integration in Eq. (36) to be carried out in order to find the variation in the FOM due to a perturbation in any of the focusing magnets in the system (contained within the  $\mathbf{O}$  and  $\mathbf{N}$  matrices). We then update the parameters according to

$$\mathbf{a}_{n+1} = \mathbf{a}_n - \gamma \nabla_{\mathbf{a}} F(\mathbf{a})|_{\mathbf{a}_n}. \quad (45)$$

The step size  $\gamma$  is adjusted iteratively according to a simple algorithm based on successive values of  $F(\mathbf{a}_{n+1})$ . Once the values of the parameters  $\mathbf{a}_{n+1}$  are determined, the gradient is recomputed and Eq. (45) is reapplied. The procedure is repeated until the FOM stops decreasing.

A termination criterion is used in cases where the FOM does not reach a minimum fast enough. The criterion has the optimization terminate if the FOM drops below a specified threshold level. A relative tolerance level of around  $10^{-7}$  is the cutoff point in the optimizations.

TABLE III. Initial moment values when solving the equations from Eq. (13).

Initial moments ( $z = 0$ )	Value
$Q_+$	$2.58 \times 10^{-6}$
$Q_-$	$2.52 \times 10^{-6}$
$E_+$	$5.07 \times 10^{-5}$
$E_-$	$4.97 \times 10^{-5}$

At this level, the trade-off in computation time for further reduction of FOM is too great. In our case, the threshold is acceptable enough for the lattice configurations being investigated. Depending on the lattice accuracy requirements for various experimental setups, the relative tolerance thresholds would need to be adjusted accordingly.

As a first example of the technique, three optimizations are run to optimize an FTR lattice in the presence of self-fields. A 5 keV beam is used with 0 (no self-fields), 1, and 5 mA beam currents. For these currents the self-field parameter takes on values  $\Lambda = 0, 2.13 \times 10^{-5}$ , and  $1.06 \times 10^{-4}$ . The importance of self-fields is estimated from Eq. (20) to be quantified by the value of  $d^2\Lambda/Q_+(0) = 0.0, 0.104$ , and  $0.518$ . The moment values in Table III are used to create a horizontally flat beam with an initial emittance ratio of 100/1.

The lattice consists of a quadrupole triplet with a solenoid at the end. A set of 11 input parameters is tuned for the optimization: the positions and magnet strengths of the three quadrupoles and solenoid along with the rotation angle of each quadrupole. An initial guess at parameters for the 0 mA case is used as the starting values for the optimization. The final optimized results for the 0 mA run are then used as initial values for the optimizations with added self-fields. We can think of the added self-fields as perturbations on the 0 mA solution. Thus, the best starting values for the self-field optimizations are the results obtained from the 0 mA, no self-fields, solution. Results are shown in Fig. 4.

Figure 4 shows that by using the adjoint technique, we are able to find and optimize solutions with increasingly large self-field forces. Each panel shows the profiles of the  $\langle x^2 \rangle$  and  $\langle y^2 \rangle$  moments of the beam as well as the location and strength of the quadrupole and solenoidal magnetic fields. The three columns in Fig. 4 correspond to three different beam currents, 0, 1, and 5 mA. The top row shows the spatial dependence of the moments with the initial magnet parameters selected based on the thin lens and no space-charge approximation.

The middle row shows the same information after the magnetic fields have been optimized. The figure of merit is applied at  $z = 0.3$  m, which is on the right of the panels in the first two rows. In these optimized cases, the beam appears to be round and its moments are constant in the solenoid as demanded by the FOM.

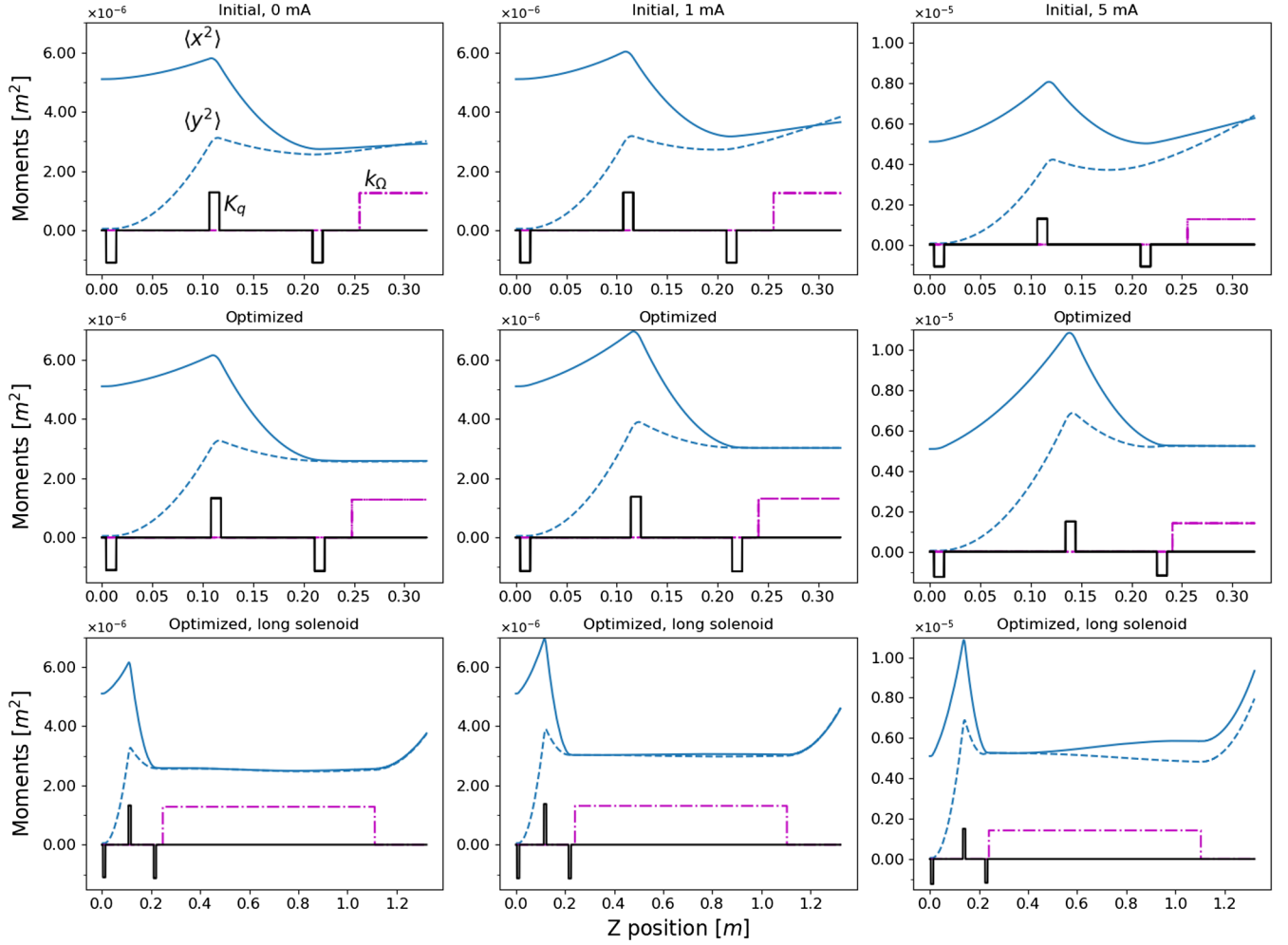


FIG. 4. rms beam size,  $\langle x^2 \rangle$ ,  $\langle y^2 \rangle$ , plotted as a function of the longitudinal coordinate through an FTR transformer. The top row shows the initial results with no optimization. The middle row shows the optimized results. The bottom row shows the same optimized results as the middle row, but integrates the beam motion through a much longer solenoid ( $z$  position). Scaled quadrupole ( $K_q$ ) and solenoid ( $k_\Omega$ ) field profiles are also displayed for reference. Results are for a 5 keV beam.

The bottom row shows the evolution of the moments if the propagation of the beam is continued through a longer solenoid, past the point at  $z = 0.3$  m where the FOM is evaluated. In the 0 and 1 mA cases, the beam parameters do not vary after this point and the beam is round indicating the FOM has been minimized to zero. However, in the 5 mA case, it is seen that the beam is not perfectly round in the solenoid. This is a consequence of our having restricted the solenoid and third quadrupole to not overlap. When we run the optimizer with no restriction, it finds a configuration where the beam is perfectly round in the solenoid. However, in this case, the quadrupole and solenoid fields overlap. To avoid this overlap, we do not allow the start location of the solenoidal field to cross to the left of the third quadrupole.

While the optimizations in Fig. 4 use hardedge magnet models, the adjoint equations are not limited to such profiles. The magnet strengths,  $k_\Omega$  and  $K_q$ , are continuous

variables in the moment equations, and as such, any profile can be used. This may be useful in analyzing certain designs that cannot necessarily be represented in matrix form for single particle tracking or would take too long to simulate in PIC tracking codes. Figure 5 shows an optimization result using Gaussian like (actually  $\cos^2$ ) magnet profiles similar to the quadrupole fields available in University of Maryland Electron Ring's magnets [28]. The solenoid field also has a fringe field edge included in its profile. Various properties of the profiles, such as width, height, etc., can be used as optimization parameters for a given FOM.

Certain lattice configurations require the quadrupole triplet to maintain a fixed 45-degree transverse rotation angle. While such an FTR solution is possible for a beam with no self-fields, it is not necessarily true when self-fields are added [29]. Keeping the quadrupole rotation angles fixed, a solution can still be found for beams with little to no space charge. However, our optimizations show that as

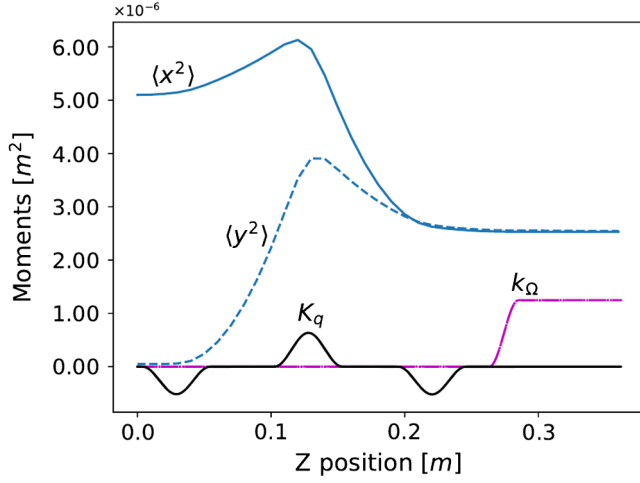


FIG. 5. rms beam size vs longitudinal position for an optimized result that uses non-hardedge magnet profiles. A beam with 5 keV energy and zero current is used.

the space-charge forces grow, finding an FTR solution becomes progressively more difficult without allowing the quadrupole rotation angles to vary in the optimization, effectively introducing three more tuning parameters that can be used.

Figure 6 demonstrates the effects of fixed vs varied quadrupole rotation angles. A set of optimizations was run using fixed and varied quadrupole rotation angles for three values of space charge, 0, 1, and 5 mA beams. The results show that as the space-charge force increases, the optimizations made with fixed rotation angles do not have constant moments in the solenoid. In the case of varied rotation angles, the moments are able to remain constant in the solenoid. This shows that in the presence of self-fields

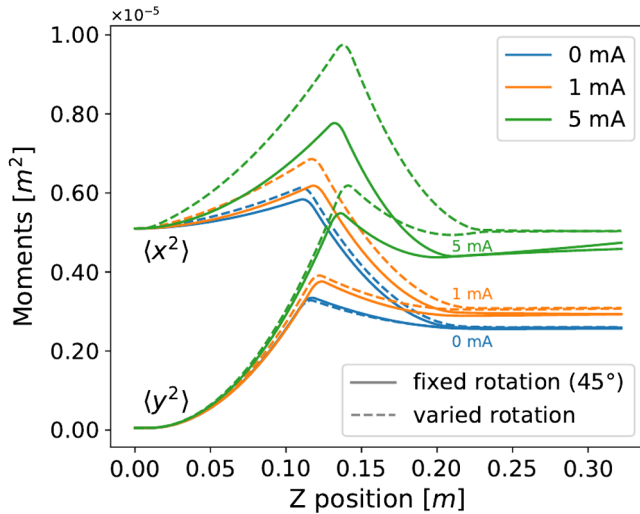


FIG. 6. rms beam size vs longitudinal position for optimizations that used varied quadrupole rotation angles (dashed lines) as well as fixed 45-degree rotation angles (solid lines). Optimizations are run for 0, 1, and 5 mA beam currents with 5 keV beam energy.

the quadrupole triplet matching outside the solenoid is no longer trivial and requires adjusting of rotation angles to properly match the beam into the solenoid.

Another interesting note in Fig. 6 is that fixed rotation angle optimizations result in overall smaller beam sizes in the solenoid than varied rotation angle optimizations. All the optimizations were run with the same initial beam conditions (beam size, emittance, etc.). We also observed that tuning the initial conditions did not significantly improve the optimized results in the cases of fixed rotation angles. Our results, through a set of numerical optimizations, show that, in agreement with [29], with increasing space-charge forces the quadrupole rotations angles need to be tuned in order to meet FTR transformation conditions.

All optimizations up to this point have used the same initial beam conditions displayed in Table II. These initial conditions are needed in order to achieve a successful FTR transformation, in particular, the values of  $E_+$  and  $E_-$ , which depend on  $Q_+$  and  $Q_-$ . However, some level of deviation from these initial conditions is expected in experiments. It might be difficult to exactly produce a beam with the required initial conditions. Figure 7 shows that it is possible to correct mismatched parameters using the optimizer. The three panels correspond to three different current values. Each panel displays three pairs of profiles for the spatial second moments. The profiles marked with the circles are the optimized profiles from the middle row of Fig. 4. These were constructed with initial values for  $E_+$  and  $E_-$  given in Table II. We then changed the  $E_{+,-}$  moment values by 50%. This resulted in the pairs of red curves which indicate the beam is no longer round in the solenoid. The optimizer was then run with the perturbed moments and the beam was brought back to round as shown by the blue curves.

The FOM in Eq. (40) is constructed as a sum of terms each of which can be made to vanish. As mentioned, the terms are selected to make the moments independent of axial distance in the solenoid. The second to last term enforces radial force balance so that  $P_+$  is constant. The last term enforces rigid rotation. This term has had little effect on the optimizations shown so far, so we omit it from further considerations. If the beam is in radial force balance it must acquire a mean rotation to balance the radial self-fields. This would degrade the cooling properties for a copropagating beam. To show the robustness of the optimization technique and explore a trade-off between cooling potential and uniformity of moments in the solenoid, we modify the FOM by adding a new fifth term that reflects the transverse energy of the beam particles in the lab frame. The new FOM is

$$F_{\text{total}} = F_1 + F_2 + F_3 + \epsilon_1 F_4 + \epsilon_2 F_5, \quad (46)$$

where



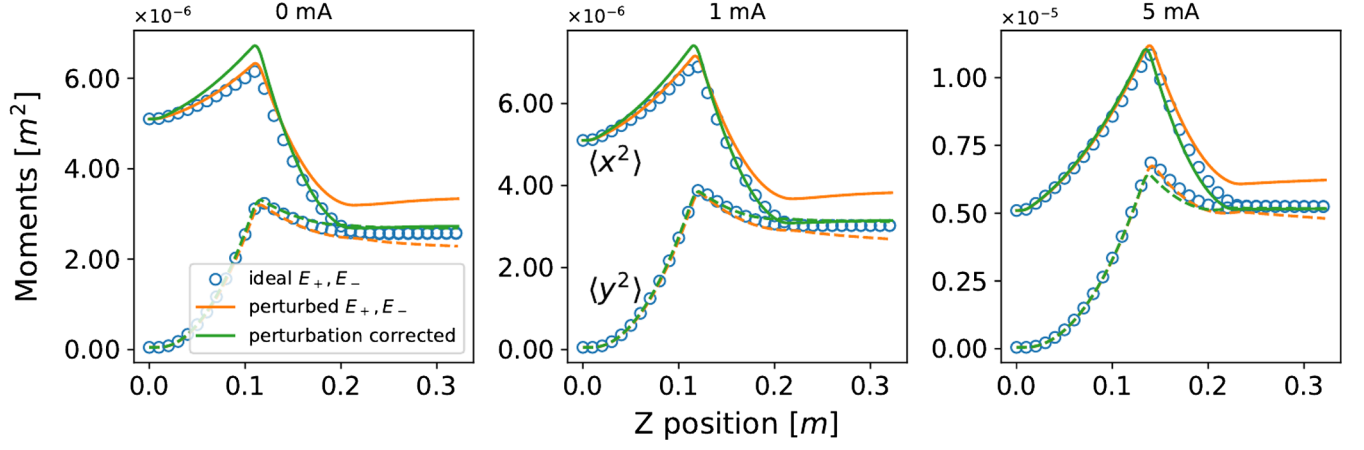


FIG. 7. rms beam size is plotted for optimized results with ideal initial conditions (blue circles). Initial conditions are purposely perturbed to break the optimized results (orange line). Magnet parameters are reoptimized to correct for the perturbed initial conditions (green line).

$$F_1 = \frac{1}{2} |\mathbf{P}|^2, \quad (47a)$$

$$F_2 = \frac{1}{2} k_0^2 (Q_-^2 + Q_x^2), \quad (47b)$$

$$F_3 = \frac{1}{2} k_0^{-2} (E_-^2 + E_x^2), \quad (47c)$$

$$F_4 = \frac{1}{2} k_0^{-2} \left( E_+ - \frac{1}{2} k_\Omega^2 Q_+ + \Lambda \right)^2, \quad (47d)$$

$$F_5 = \frac{1}{2} k_0^{-2} E_{+lab}^2 \quad (47e)$$

and  $E_{+lab}$  is the  $E_+$  moment from Eq. (12) represented in the lab frame (vs the Larmor frame) as

$$E_{+lab} = \left( E_+ + \frac{1}{2} k_\Omega^2 Q_+ - k_\Omega L \right). \quad (48)$$

The terms  $F_1$ – $F_3$  measure the degree of axial uniformity and the degree of cylindrical symmetry of the beam in the solenoid. The fourth term,  $F_4$ , in the FOM still represents radial force balance. These terms can, in principle, be minimized to zero. The fifth term,  $F_5$ , which is new, is positive definite and cannot be minimized to zero; it represents the beam's transverse kinetic energy in the lab frame and should be minimized if the considered application is copropagating hadron beam cooling. In principle, one would like to minimize the integral of the transverse energy over the length of the solenoid rather than in just one plane as done here. Further, one would also want to account for the radial size and distribution of both the hadron beam and the cooling beam. However, this is beyond the scope of

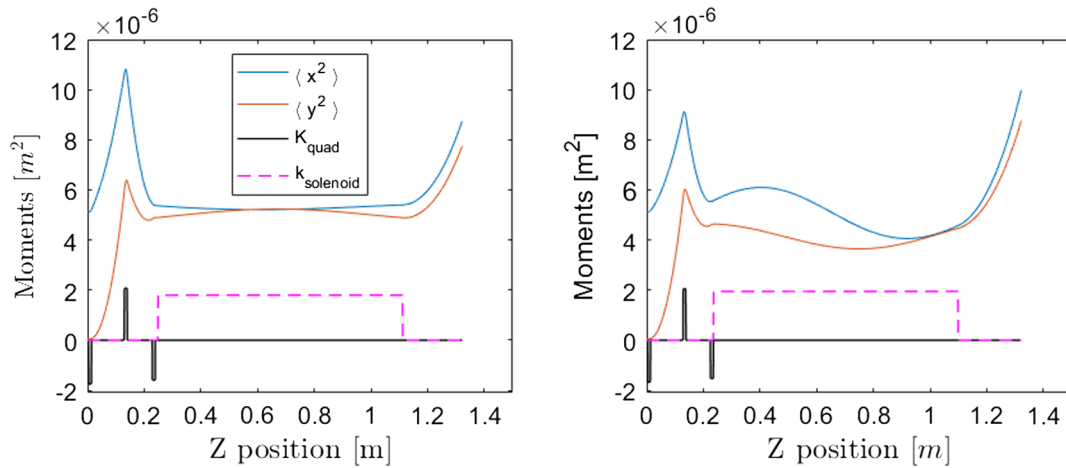


FIG. 8. rms beam size is plotted for two different optimization using the FOM from Eq. (46). The left side uses weight values of  $\epsilon_1, \epsilon_2 = 1, 0$  while the right side uses  $\epsilon_1, \epsilon_2 = 0, 1$ . A 5-mA beam current and the same set of initial conditions are used for both optimizations. The optimizations minimized the FOM from Eq. (46) at roughly the center of the solenoid ( $Z = 0.722$  m).

TABLE IV. Optimization results for minimizing components  $F_4$  and  $F_5$  from the FOM in Eq. (46).

FOM	$\epsilon_1, \epsilon_2 = 1, 0$	$\epsilon_1, \epsilon_2 = 0, 1$
$F_1$	$0.0005 \times 10^{-10}$	$0.0812 \times 10^{-10}$
$F_2$	$0.0000 \times 10^{-10}$	$0.1212 \times 10^{-10}$
$F_3$	$0.0026 \times 10^{-10}$	$0.0172 \times 10^{-10}$
$F_4$	$0.0003 \times 10^{-10}$	$0.0267 \times 10^{-10}$
$F_5$	$0.3100 \times 10^{-10}$	$0.1986 \times 10^{-10}$

the present paper, and we will simply minimize the transverse energy in a single plane. Further, a pair of weighting parameters labeled  $\epsilon_1, \epsilon_2$  is included in the FOM in order to trade off between minimizing the beam's radial oscillations, as measured by  $F_4$ , and transverse energy, as measured by  $F_5$ , within the solenoid.

Using the new FOM in Eq. (46) along with a 5-mA high space-charge beam, an optimization is performed using weighting parameters  $\epsilon_1, \epsilon_2$  to prioritize either minimizing the beam's radial oscillations or its transverse energy within the solenoid. The results are shown in Fig. 8. On the left side of Fig. 8, the weight parameters are set to  $\epsilon_1, \epsilon_2 = 1, 0$  which forces the optimization to focus on minimizing the beam's radial oscillations, creating a relatively flat, oscillation free, beam within the solenoid. On the right side of Fig. 8, the weight parameters are set to  $\epsilon_1, \epsilon_2 = 0, 1$  which directs the optimization to instead minimize the beam's transverse energy. Table IV shows the calculated FOM component values post optimization with the indicated weight parameters. That is, the two optimizations were performed with weight parameters set to  $\epsilon_1, \epsilon_2 = (1, 0)$  and  $(0, 1)$ . Using the optimized magnet parameters, the weight parameters were reset to  $\epsilon_1, \epsilon_2 = (1, 1)$  and each FOM component was calculated; these are the values shown in Table IV. The table shows that the transverse energy,  $F_5$ , is lower (by a third) for the optimization of  $\epsilon_1, \epsilon_2 = (0, 1)$  than for the  $\epsilon_1, \epsilon_2 = (1, 0)$  optimization as expected. The beam's radial oscillations,  $F_4$ , along with the other FOM components, are larger in the  $\epsilon_1, \epsilon_2 = (0, 1)$  optimization, again as expected. For the optimization case of  $\epsilon_1, \epsilon_2 = (1, 0)$ , the first four components of the FOM are very small, but the transverse energy term,  $F_5$ , is bigger than the first four in this case.

Using the modified FOM in Eq. (46), the optimization results in Fig. 8 and Table IV demonstrate the ability to trade off minimization of different properties of the beam within the solenoid. Depending on a given experimental tolerance, a pair of weight parameters can be used to create a balance between the beam's radial oscillations and transverse energy within the solenoid. Such a method can be applied to any set of terms in a given FOM.

#### IV. CONCLUSION

In conclusion, this paper presents the adjoint approach to calculating the gradient with respect to magnet parameters

of figures of merit characterizing accelerator lattices. The gradient is then used in the steepest descent algorithm to optimize a flat-to-round lattice transformer in the presence of space-charge forces. The optimization is developed for a model of the system consisting of a set of continuous differential equations for the ten moments describing the 4D phase space in the paraxial limit. The optimization is achieved by tuning a set of quadrupole and solenoid magnet parameters within the lattice.

The adjoint approach and gradient descent algorithm are able to successfully optimize various transformer systems. The optimization can account for space-charge forces, continuous variation of magnetic field profiles, and for mismatches in beam parameters on entry to the transformer. In using the adjoint approach to the problem, optimizations are performed at significantly reduced computational costs. The system of moment equations can be solved twice and the gradients with respect to all parameters are determined as integrals of the adjoint solution. Direct determination of the gradient would require as many simulations as there are parameters.

While the results shown here are based on a reduced model of the charged particle dynamics, namely a solution of the moment equations assuming linear restoring forces, the adjoint method can also be applied to a particle description (see [4] and [5]). In further work, we plan to generalize the optimization procedure to the particle description of the beam dynamics and to the optimization of circular lattices. A future challenge is to include dissipative effects such as synchrotron radiation which are important for high energy accelerators.

#### ACKNOWLEDGMENTS

This work was supported by DOE-HEP Award No. DE-SC0010301.

#### APPENDIX A: COUPLED MOMENT EQUATIONS

In this Appendix, we outline the steps leading to Eqs. (13)–(15). We start with the definitions of the moments in Eqs. (9)–(12). Equation (13), the derivative of  $\mathbf{Q}$  with respect to distance  $z$ , follows directly from the definition of the components of  $\mathbf{P}$  in Eq. (10). Differentiating the vector  $\mathbf{P}$  and angular momentum  $L$  with respect to  $z$  gives

$$P'_+ = E_+ + \langle xx'' + yy'' \rangle, \quad (\text{A1a})$$

$$P'_- = E_- + \langle xx'' - yy'' \rangle, \quad (\text{A1b})$$

$$P'_x = E_x + \langle xy'' + yx'' \rangle, \quad (\text{A1c})$$

$$L' = \langle xy'' - yx'' \rangle, \quad (\text{A1d})$$

where the moments  $E$  are defined in Eq. (12). Differentiating the moments  $E$  with respect to axial distance gives

$$E'_+ = 2\langle x'x'' + y'y'' \rangle, \quad (\text{A2a})$$

$$E'_- = 2\langle x'x'' - y'y'' \rangle, \quad (\text{A2b})$$

$$E'_x = 2\langle x'y'' + y'x'' \rangle. \quad (\text{A2c})$$

Evaluation of each of the averages on the right sides of Eqs. (A1) and (A2) requires inserting expressions for the second derivatives of  $x$  and  $y$  in the Larmor frame. For this we use Eq. (9).

The first contribution to the averages in Eqs. (A1) and (A2) comes from the solenoidal field contribution to Eq. (9),

$$\begin{pmatrix} x''_B \\ y''_B \end{pmatrix} = -\left(\frac{k_Q}{2}\right)^2 \begin{pmatrix} x \\ y \end{pmatrix}. \quad (\text{A3})$$

This when inserted in Eqs. (A1) and (A2) gives rise to the first term in the expression for the matrix  $O$  in Eq. (14).

The second contribution to the averages in Eqs. (A1) and (A2) comes from the quadrupole fields. To evaluate those terms we write the second derivatives in the Larmor frame in the following way:

$$\begin{pmatrix} x''_Q \\ y''_Q \end{pmatrix} = \mathbf{R}^{-1}(\phi) \sum_{\text{Quads}} K_Q \begin{pmatrix} \cos 2\psi & \sin 2\psi \\ \sin 2\psi & -\cos 2\psi \end{pmatrix} \mathbf{R}(\phi) \begin{pmatrix} x \\ y \end{pmatrix}. \quad (\text{A4})$$

Here the Larmor variables,  $x$  and  $y$ , are first transformed back to the lab frame, where Eq. (17) describes the quadrupole fields. The transverse Lorentz force is computed in the lab frame, and the resulting acceleration is transformed to the Larmor frame. Multiplying the three matrices in Eq. (A4) gives

$$\begin{aligned} & \mathbf{R}^{-1}(\phi) \sum_{\text{Quads}} K_Q \begin{pmatrix} \cos 2\psi & \sin 2\psi \\ \sin 2\psi & -\cos 2\psi \end{pmatrix} \mathbf{R}(\phi) \\ &= -\sum_{\text{Quads}} K_Q \begin{bmatrix} -\cos(2\phi - 2\psi) & \sin(2\phi - 2\psi) \\ \sin(2\phi - 2\psi) & \cos(2\phi - 2\psi) \end{bmatrix}. \end{aligned} \quad (\text{A5})$$

Inserting Eq. (A5) into Eq. (A4) and then inserting the resulting second derivatives in the expressions in Eqs. (A1) and (A2) for the rates of change of the moments gives rise to the second contribution to the matrix  $O$  in Eq. (14) and the first contribution to the vector  $N$  in Eq. (15).

The evaluation of the self-field force is done in what we label the beam frame. In the beam frame the beam is assumed to have an elliptical shape with major and minor semiaxes radii  $a$  and  $b$ , aligned with the  $x_2$  and  $y_2$

directions respectively. The  $x_2$  direction makes an angle  $\alpha$  with respect to the  $x$ -axis in the Larmor frame. The transformation from beam frame to Larmor frame coordinates is thus given by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{R}(\phi) \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}. \quad (\text{A6})$$

The beam is assumed to have uniform density within an elliptical cross section in the beam frame. Thus, the beam frame moments satisfy

$$\langle x_2^2 \rangle = a^2/4, \quad \langle y_2^2 \rangle = b^2/4, \quad \langle x_2 y_2 \rangle = 0. \quad (\text{A7})$$

This implies through application of Eq. (A6) for the Larmor frame moments

$$\begin{aligned} \langle x^2 \rangle &= \frac{1}{8} [(a^2 + b^2) + \cos(2\alpha)(a^2 - b^2)] \\ \langle y^2 \rangle &= \frac{1}{8} [(a^2 + b^2) - \cos(2\alpha)(a^2 - b^2)] \\ \langle xy \rangle &= \frac{1}{8} (a^2 - b^2) \sin(2\alpha). \end{aligned} \quad (\text{A8})$$

The Larmor frame values for the elements of  $Q$  are then related to the radii  $a$  and  $b$  and the angle  $\alpha$  by

$$\begin{aligned} Q_+ &= \frac{1}{2} \langle x^2 + y^2 \rangle = \frac{1}{8} (a^2 + b^2) \\ Q_- &= \frac{1}{2} \langle x^2 - y^2 \rangle = \frac{\cos(2\alpha)}{8} (a^2 - b^2) \\ Q_x &= \langle xy \rangle = \frac{\sin(2\alpha)}{8} (a^2 - b^2). \end{aligned} \quad (\text{A9})$$

## APPENDIX B: SELF-FIELDS AND TRANSFORMATIONS

The effective electric potential due to a beam with elliptical cross section far from conducting boundaries is given by

$$\begin{aligned} \Phi &= -\frac{2\Lambda}{(a+b)} \left[ \frac{x_2^2}{a} + \frac{y_2^2}{b} \right] \\ &= -\frac{2\Lambda}{ab} \left\{ \frac{1}{2} (x^2 + y^2) + \frac{b-a}{b+a} \left[ \frac{\cos(2\alpha)}{2} (x^2 - y^2) \right. \right. \\ &\quad \left. \left. + \sin(2\alpha)xy \right] \right\}, \end{aligned} \quad (\text{B1})$$

where the first expression is in the beam frame and the second expression has been transformed to the Larmor frame. The current strength parameter  $\Lambda$  is defined in Eq. (20). The relation between the effective potential and the space-charge potential  $\phi_{sc}$  is  $\Phi = q\phi_{sc}/(m\nu_z^2\gamma^3)$ .

One of the three powers of  $\gamma$  in the denominator of the effective potential accounts for the effective relativistic mass increase of the beam particles, and the other two powers account for the self-magnetic field. The accelerations due to self-fields then follow from

$$\begin{pmatrix} x''_{\Lambda} \\ y''_{\Lambda} \end{pmatrix} = - \begin{pmatrix} \partial/\partial x \\ \partial/\partial y \end{pmatrix} \Phi(x, y). \quad (\text{B2})$$

When Eq. (B2) is inserted in Eqs. (A1) and (A2), and the parameters  $a$ ,  $b$ , and  $\alpha$  are expressed in terms of the moments  $\mathbf{Q}$  using Eq. (A9), the result is the self-field contributions to the matrix  $\mathbf{O}$  and vector  $\mathbf{N}$  appearing in Eqs. (14) and (15) along with definitions in Eq. (19).

Equations (13a)–(13d) describe the evolution of the moments in the Larmor frame. If the values of the moments in the lab frame are desired, these can be recovered by the following transformations

$$\mathbf{Q}_{\text{lab}} = \mathbf{R}_1 \cdot \mathbf{Q} \quad (\text{B3a})$$

$$\mathbf{P}_{\text{lab}} = \mathbf{R}_1 \cdot \mathbf{P} + 2\phi' \mathbf{R}_2 \cdot \mathbf{Q} \quad (\text{B3b})$$

$$\mathbf{E}_{\text{lab}} = \mathbf{R}_1 \cdot \mathbf{E} + 2\phi' \mathbf{R}_2 \cdot \mathbf{P} + 2\phi'^2 \mathbf{R}_3 \cdot \mathbf{Q} + 2\phi' \mathbf{1}_+ L \quad (\text{B3c})$$

$$L_{\text{lab}} = L + 2\phi' Q_+. \quad (\text{B3d})$$

Here the following matrices are defined:

$$\mathbf{R}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(2\phi) & -\sin(2\phi) \\ 0 & \sin(2\phi) & \cos(2\phi) \end{bmatrix} \quad (\text{B4a})$$

$$\mathbf{R}_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\sin(2\phi) & -\cos(2\phi) \\ 0 & \cos(2\phi) & -\sin(2\phi) \end{bmatrix} \quad (\text{B4b})$$

$$\mathbf{R}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\cos(2\phi) & \sin(2\phi) \\ 0 & -\sin(2\phi) & -\cos(2\phi) \end{bmatrix} \quad (\text{B4c})$$

$$\mathbf{1}_+ = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \quad (\text{B4d})$$

The adjoint equation (28c) has an added term  $\delta \dot{\mathbf{E}}^{(Y)}$ , which is added to cancel out the terms in Eq. (30) that come from the dependence of the matrix  $\mathbf{O}$  and vector  $\mathbf{N}$  on the changes in the self-fields. These changes are proportional to changes in the moments  $\delta \mathbf{Q}^{(X)}$ . To represent these changes in the  $\mathbf{O}$  matrix we first separate the changes due to the self-fields ( $\Lambda$ ) from the changes in the magnetic focusing

parameters ( $\mathbf{Q}$ ,  $B$ ),  $\delta \mathbf{O}^{(X)} = \delta \mathbf{O}_{Q,B}^{(X)} + \delta \mathbf{O}_{\Lambda}^{(X)}$ . We then construct matrices  $\mathbf{M}_P$  and  $\mathbf{M}_Q$  that satisfy

$$\delta \mathbf{O}_{\Lambda}^{(X)} \cdot \mathbf{P} = \mathbf{M}_P \cdot \delta \mathbf{Q}^{(X)} \quad (\text{B5a})$$

$$\delta \mathbf{O}_{\Lambda}^{(X)} \cdot \mathbf{Q} = \mathbf{M}_Q \cdot \delta \mathbf{Q}^{(X)}. \quad (\text{B5b})$$

The vector  $\mathbf{N}$  is treated similarly.

We focus now on  $\mathbf{M}_P$ . We start by forming the vector

$$\mathbf{R}_P = \mathbf{O}_{\Lambda} \cdot \mathbf{P} = \frac{\Lambda}{Q_{\Delta}} \begin{pmatrix} P_+ - \frac{Q_-}{Q_{\Delta} + Q_+} P_- - \frac{Q_x}{Q_{\Delta} + Q_+} P_x \\ -\frac{Q_-}{Q_{\Delta} + Q_+} P_+ + P_- \\ -\frac{Q_x}{Q_{\Delta} + Q_+} P_+ + P_x \end{pmatrix}. \quad (\text{B6})$$

The elements of the matrix  $\mathbf{M}_P$  are generated by differentiating each element of vector  $\mathbf{R}_P$  with respect to each element of vector  $\mathbf{Q}$ ,

$$\mathbf{M}_P = \left[ \frac{\partial \mathbf{R}_P}{\partial \mathbf{Q}} \right]^T. \quad (\text{B7})$$

To keep track of terms, we decompose  $\mathbf{M}_P$  according to

$$\mathbf{M}_P = \mathbf{V}_1 \mathbf{U}_1^T + \mathbf{V}_2 \mathbf{U}_2^T + \mathbf{V}_3 \mathbf{U}_3^T + \mathbf{V}_4 \mathbf{U}_4^T, \quad (\text{B8})$$

where

$$\mathbf{V}_1 = -\frac{\Lambda}{Q_{\Delta}^2} \begin{pmatrix} P_+ - \frac{Q_-}{Q_{\Delta} + Q_+} P_- - \frac{Q_x}{Q_{\Delta} + Q_+} P_x \\ -\frac{Q_-}{Q_{\Delta} + Q_+} P_+ + P_- \\ -\frac{Q_x}{Q_{\Delta} + Q_+} P_+ + P_x \end{pmatrix} \quad (\text{B9a})$$

$$\mathbf{U}_1^T = \left( \frac{\partial Q_{\Delta}}{\partial Q_+}, \frac{\partial Q_{\Delta}}{\partial Q_-}, \frac{\partial Q_{\Delta}}{\partial Q_x} \right) = \frac{1}{Q_{\Delta}} (Q_+ \quad -Q_- \quad -Q_x) \quad (\text{B9b})$$

$$\mathbf{V}_2 = \frac{\Lambda}{Q_{\Delta}(Q_{\Delta} + Q_+)^2} \begin{pmatrix} Q_- P_- + Q_x P_x \\ Q_- P_+ \\ Q_x P_+ \end{pmatrix} \quad (\text{B9c})$$

$$\begin{aligned} \mathbf{U}_2^T &= \left( \frac{\partial(Q_{\Delta} + Q_+)}{\partial Q_+}, \frac{\partial(Q_{\Delta} + Q_+)}{\partial Q_-}, \frac{\partial(Q_{\Delta} + Q_+)}{\partial Q_x} \right) \\ &= \mathbf{U}_1^T + (1 \quad 0 \quad 0) \end{aligned} \quad (\text{B9d})$$

$$\mathbf{V}_3 = -\frac{\Lambda}{Q_{\Delta}(Q_{\Delta} + Q_+)} \begin{pmatrix} P_- \\ P_+ \\ 0 \end{pmatrix}, \quad (\text{B9e})$$

$$\mathbf{V}_4 = -\frac{\Lambda}{Q_{\Delta}(Q_{\Delta} + Q_+)} \begin{pmatrix} P_x \\ 0 \\ P_+ \end{pmatrix} \quad (\text{B9e})$$



$$U_3^T = (0 \quad 1 \quad 0), \quad U_4^T = (0 \quad 0 \quad 1). \quad (\text{B9f})$$

The matrices  $M_Q$  and  $M_N$  are constructed in the same way.

### APPENDIX C: SCALING PROPERTIES

Finally, we discuss the scaling properties of our system of moment equations. Let us suppose that we have found a solution to the moment equations in the form of functions

$$Q_0(z_0), P_0(z_0), E_0(z_0), O_0(z_0), N_0(z_0), L_0(z_0). \quad (\text{C1})$$

It can then be verified that a scaled solution that is also satisfying the moment equations is

$$\begin{aligned} Q(z) &= \lambda Q_0(\epsilon z) \\ P(z) &= \lambda \epsilon P_0(\epsilon z) \\ E(z) &= \lambda \epsilon^2 E_0(\epsilon z) \\ O(z) &= \epsilon^2 O_0(\epsilon z) \\ N(z) &= \epsilon^2 N_0(\epsilon z) \\ L(z) &= \lambda \epsilon L_0(\epsilon z). \end{aligned} \quad (\text{C2})$$

Here  $\epsilon$ ,  $\lambda$  are arbitrary constants allowing for separate scaling of the axial dependence and size of the beam. If  $\epsilon > 1$ , the scaled solution is shorter in spatial length than the original solution, and if  $\epsilon < 1$ , the scaled solution is longer than the original. If  $\lambda > 1$ , the scaled beam is larger. Of course this is subject to the requirement that the transverse forces remain linear in displacements.

Looking at the expressions for the matrices  $O$ ,  $N$  in Eq. (C2), there are a number of conditions that must be satisfied to give the required  $\epsilon^2$  scaling. The solenoidal field contribution requires

$$k_\Omega(z) = \epsilon k_{\Omega 0}(\epsilon z). \quad (\text{C3})$$

This means a shorter solution requires a stronger solenoidal field. The scaled solution gives for the phase

$$\phi(z) = \phi_0(\epsilon z). \quad (\text{C4})$$

Thus, the values of the phase in the locations of the quadrupoles are preserved under the scaling. For the quadrupoles, strict application of the scaling gives

$$K_q(z) = \epsilon^2 K_{q0}(\epsilon z). \quad (\text{C5})$$

However, in the thin lens approximation only the integrated value of the quadrupole field matters:

$$K = \int dz K_q(z) = \epsilon \int \epsilon dz K_{q0}(\epsilon z) = \epsilon K_0. \quad (\text{C6})$$

In this approximation the strength of the quadrupole field also scales inversely with length.

The self-field contribution to the  $O$  and  $N$  matrices can also be preserved. We note that these contributions scale as

$$O, \quad N \propto \Lambda/Q. \quad (\text{C7})$$

So to preserve the solution we require

$$O \propto \Lambda/Q = \Lambda/(\epsilon \lambda Q_0) = \epsilon^2 O_0 \propto \epsilon^2 \Lambda_0/Q_0. \quad (\text{C8})$$

As a result, the current parameter scales as

$$\Lambda = \lambda \epsilon^3 \Lambda_0. \quad (\text{C9})$$

Thus, scaling length or amplitude requires changing beam current to maintain space-charge influence.

- 
- [1] S. Director and R. Rohrer, The generalized adjoint network and network sensitivities, *IEEE Trans. Circuit Theory* **16**, 318 (1969).
  - [2] N. K. Nikolova, M. H. Bakr, and J. W. Bandler, Sensitivity analysis of network parameters with electromagnetic frequency-domain simulators, *IEEE Trans. Microwave Theory Tech.* **54**, 670 (2006).
  - [3] A. Jameson, Optimum aerodynamic design using CFD and control theory, *CFD Rev.* **3**, 926 (1995).
  - [4] T. M. Antonsen, D. Chernin, and J. J. Petillo, Adjoint approach to beam optics sensitivity based on Hamiltonian particle dynamics, *Phys. Plasmas* **26**, 013109 (2019).
  - [5] T. M. Antonsen, B. L. Beaudoin, L. Dovlatyan, and I. Haber, Adjoint approach to accelerator lattice design, in *Proceedings of NAPAC'19*, North American Particle Accelerator Conference No. 4 (JACoW Publishing, Geneva, Switzerland, 2019), pp. 376–378.
  - [6] Ya. Derbenev, Adapting optics for high-energy electron cooling, Technical Report No. UM-HE-98-04-A, 1998.
  - [7] A. V. Burov and V. V. Danilov, An insertion to eliminate horizontal temperature of high-energy electron beam, Technical Report No. FERMILAB-TM-2043, 1998.
  - [8] A. V. Burov and S. Nagaitsev, Courant-Snyder parameters of beam adapters, Technical Report No. FERMILAB-TM-2114, 2000.
  - [9] A. V. Burov, S. Nagaitsev, and Ya. Derbenev, Circular modes, beam adapters, and their applications in beam optics, *Phys. Rev. E* **66**, 016503 (2002).
  - [10] A. W. Chao and M. Tigner, *Handbook of Accelerator Physics and Engineering* (World Scientific, Singapore, 1999).
  - [11] Ya. Derbenev, Advanced optical concepts for electron cooling, *Nucl. Instrum. Methods Phys. Res., Sect. A* **441**, 223 (2000).
  - [12] R. Brinkmann, Ya. Derbenev, and K. Flottmann, A flat beam electron source for linear colliders, Technical Report No. DESY-TESLA-99-09, 1999.

- [13] R. Brinkmann, Y. Derbenev, and K. Flöttmann, A low emittance, flat-beam electron source for linear colliders, *Phys. Rev. ST Accel. Beams* **4**, 053501 (2001).
- [14] D. Edwards *et al.*, The flat beam experiment at the FNAL photoinjector, in *Proceedings of the 20th International Linac Conference, LINAC-2000, Monterey, CA, 2000* (SLAC, Menlo Park, CA, 2000), eConf C000821, MOB13.
- [15] D. Edwards *et al.*, Status of flat electron beam production, in *Proceedings of the Particle Accelerator Conference, Chicago, IL, 2001* (IEEE, New York, 2001), Vol. 1, pp. 73–75.
- [16] S. Wang, I. Corlett, S. Lidia, J. Staples, and A. Zholents, Flat beam production in low energy injectors, in *Proceedings of the 2003 Particle Accelerator Conference, Portland, OR* (IEEE, New York, 2003), Vol. 5, pp. 3198–3200.
- [17] S. Lidia, Emittance compensation studies of photoinjector beams with angular momentum, in *Proceedings of the 2003 Particle Accelerator Conference, Portland, OR* (Ref. [16]), Vol. 3, pp. 2089–2091.
- [18] A. Ody, P. Musumeci, J. Maxson, D. Cesar, R. England, and K. Wootton, Flat electron beam sources for DLA accelerators, *Nucl. Instrum. Methods Phys. Res., Sect. A* **865**, 75 (2017).
- [19] F. W. Cropp *et al.*, Maximizing 2-D beam brightness using the round to flat beam transformation in the ultralow charge regime, in *Proceedings of NAPAC'19*, North American Particle Accelerator Conference No. 4 (Ref. [5]), pp. 986–989.
- [20] A. Wolski, Alternative approach to general coupled linear optics, *Phys. Rev. ST Accel. Beams* **9**, 024001 (2006).
- [21] E. Thrane *et al.*, Photoinjector production of a flat electron beam, in *Proceedings of the 21st International Linac Conference, Gyeongju, Korea, 2002* (Pohang Accelerator Laboratory, Pohang, Korea, 2002), pp. 308–310.
- [22] Y. Sun, Angular-momentum-dominated electron beams and flat-beam generation, Ph.D. thesis, Chicago University, 2005.
- [23] K. Kim, Round-to-flat transformation of angular-momentum-dominated beams, *Phys. Rev. ST Accel. Beams* **6**, 104002 (2003).
- [24] G. Rangarajan, F. Neri, and A. Dragt, Generalized emittance invariants, in *Proceedings of the 1989 IEEE Particle Accelerator Conference, Chicago, IL* (IEEE, New York, 1989), Vol. 2, pp. 1280–1282.
- [25] V. A. Lebedev and S. A. Bogacz, Betatron motion with coupling of horizontal and vertical degrees of freedom, *J. Instrum.* **5**, P10010 (2010).
- [26] A. Friedman, R. H. Cohen, D. P. Grote, S. M. Lund, W. M. Sharp, J.-L. Vay, I. Haber, and R. A. Kishek, Computational methods in the warp code framework for kinetic simulations of particle beams and plasmas, *IEEE Trans. Plasma Sci.* **42**, 1321 (2014).
- [27] I. M. Kapchinsky and V. V. Vladimirovsky, in *Proceedings of the International Conference on High Energy Accelerators* (CERN, Geneva, Switzerland, 1959), p. 274.
- [28] W. W. Zhang, S. Bernal, H. Li, T. Godlove, R. A. Kishek, P. G. O'Shea, M. Reiser, V. Yun, and M. Venturini, Design and field measurements of printed-circuit quadrupoles and dipoles, *Phys. Rev. ST Accel. Beams* **3**, 122401 (2000).
- [29] S. B. Moroch, T. W. Koeth, and B. E. Carlsten, Space-charge effects in low-energy flat-beam transforms, *arXiv: 2102.13567*.