

Beam optics ramping in underconstrained lattice design: Application to Electron-Ion Collider hadron storage ring cooling section

Derong Xu^{*}*Brookhaven National Laboratory, Upton, New York 11973, USA*

(Received 30 April 2025; accepted 25 July 2025; published 22 August 2025)

This paper presents the lattice design and optics ramping strategy for the cooling section of the hadron storage ring at the electron-ion collider. The main challenge is that available tuning knobs exceed beam-optics constraints. Independently optimized injection and top-energy optics often yield disconnected solutions, making interpolation impossible. To address this, we propose two new methods. The first is an intermediate-penalty scheme that ensures ramping path continuity by penalizing constraint violations at intermediate states. The second is a continuation-based approach that adapts high-energy optics to low energy, guided by an adaptive weighting scheme to balance injection and ramping constraints. The solutions meet all beam dynamics and hardware limits. The two methods offer a general strategy for ramping in systems where the solution space is under-constrained and the starting and target configurations are far apart.

DOI: 10.1103/86y4-h97k

I. INTRODUCTION

Beam optics ramping is a fundamental task in synchrotron accelerators. The lattice must evolve smoothly between operational states to maintain acceptable optics. Common examples include the beta squeeze in colliders [1], where the beta functions at the interaction point are reduced gradually, and solenoid ramping in detectors [2], which requires continuous optics matching to minimize betatron coupling as the solenoid field increases.

A particularly challenging case arises in the electron-ion collider (EIC), where the cooling section of the hadron storage ring (HSR) must support electron cooling requirements at injection and accommodate different conditions at top energy. The large variation in requirements complicates the construction of a smooth and feasible ramping path.

The EIC, under construction at Brookhaven National Laboratory, will collide polarized electron and hadron beams over a broad energy and species range. It targets a peak luminosity of $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ and average polarization of $\sim 70\%$ [3]. As shown in Fig. 1, the EIC includes multiple key components: the hadron storage ring (HSR), the electron storage ring (ESR), a new injector complex, and two interaction regions.

The HSR, upgraded from Relativistic Heavy Ion Collider (RHIC), will store highly polarized hadron beams with a flat profile, where vertical emittance is 10 times smaller than horizontal. Six helical Siberian snakes [4] will preserve polarization during acceleration and storage. The insertion region at 2 o'clock (IR2) [5] will serve as the cooling section to generate flat beams at injection

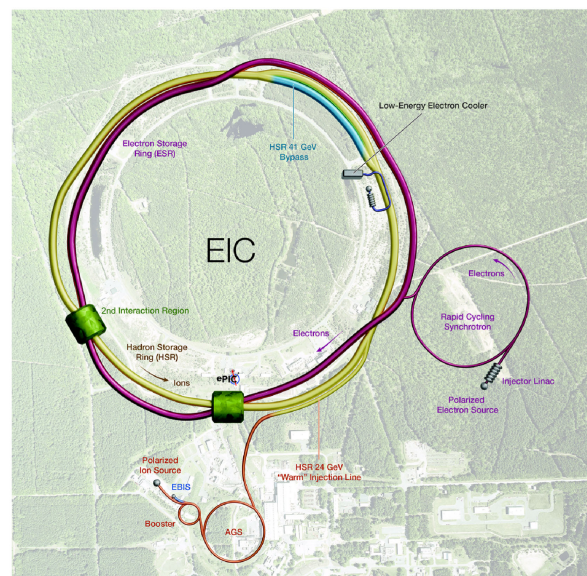


FIG. 1. Conceptual layout of the EIC, showing key components including the HSR, ESR, injector complex, and interaction regions. IR2 serves as the cooling section for generating flat hadron beams and will probably support high-energy cooling in future upgrades.

^{*}Contact author: dxu@bnl.gov

energy and potentially provide high-energy cooling. It will also host one Siberian snake.

This paper is organized as follows. Section II defines the beam optics matching and ramping problem for HSR-IR2, including relevant constraints. Section III presents an intermediate-penalty method for the case without top-energy cooling constraints. Section IV addresses the more constrained case and develops a continuation-based strategy that adapts high-energy optics to satisfy injection requirements. Section V summarizes the results.

II. PROBLEM

A flat hadron beam is essential for achieving the EIC's peak luminosity goal. According to the EIC design, the injection and top-energy optics must satisfy distinct optics requirements, and a smooth ramping path is required to preserve the emittance ratio during acceleration.

A. EIC cooling strategy

Synchrotron radiation is negligible in the EIC HSR. Without external cooling, beam emittance is determined by the source and is similar in both transverse planes. The original EIC strategy included staged cooling: a low-energy electron cooler (LEC) is used at injection to create a flat hadron beam [6], followed by strong hadron cooling (SHC) at top energy. Two methods were considered for the SHC—microbunched electron cooling [7] and electron cooling via a ring cooler (RC) [8]. However, luminosity studies show that the SHC increases integrated luminosity by only a factor of 2 [9]. The current EIC baseline therefore defers high-energy cooling while preserving layout flexibility for a future upgrade.

Once the flat hadron beam is established at injection, it must be accelerated to top energy through a ramping process. Although a large transverse emittance ratio has been demonstrated at RHIC [10], preserving beam flatness during ramping has not been experimentally verified. This motivates careful optics design to suppress vertical emittance growth. Accurate control of orbit, coupling, and Twiss functions is essential. In general, the ramping should be adiabatic, with continuously matched optics to minimize emittance dilution. The inclusion of the SHC strongly impacts ramping design. In the absence of high-energy

cooling, there are no specific optics constraints at top energy, allowing greater flexibility. In contrast, if high-energy cooling is implemented, the IR2 optics must satisfy strict conditions set by the chosen cooling method, complicating the construction of a smooth, matched ramping path. This paper uses the ring cooler as an example to compare ramping strategies with and without top-energy cooling constraints.

B. Constraints and formulation

The HSR-IR2 layout must satisfy geometric constraints. A primary goal is to allocate a long drift for low-energy electron cooling while accommodating a Siberian snake and remaining compatible with the RHIC tunnel and magnet inventory. The layout should reuse existing RHIC superconducting magnets to reduce cost and simplify integration. It also respects the modularity of RHIC magnet families, while allowing flexibility through the module rearrangement.

The configuration shown in Fig. 2 meets these requirements and serves as the geometric basis for the beam optics matching discussed in the following sections. Additional layout details are provided in Appendix A.

The beam optics must satisfy Twiss and dispersion matching conditions at arc interfaces, limits on beam size and aperture, and quadrupole strength bounds. These constraints are summarized in Appendix B.

Given these conditions, the optics matching problem at a given energy is formulated as solving a system of nonlinear equations under box and inequality constraints:

$$\mathbf{f}(\mathbf{k}) = \mathbf{0} \quad \text{subject to} \quad \mathbf{k}_l \leq \mathbf{k} \leq \mathbf{k}_u, \quad \mathbf{c}(\mathbf{k}) \leq \mathbf{0}, \quad (1)$$

where $\mathbf{k} \in \mathbb{R}^n$ is the vector of quadrupole strengths, $\mathbf{f}(\mathbf{k}) \in \mathbb{R}^m$ represents the exact matching conditions (e.g., Twiss and dispersion functions), $\mathbf{c}(\mathbf{k})$ encodes soft constraints such as aperture or beam size limits, and $\mathbf{k}_l, \mathbf{k}_u$ are the hardware-imposed lower and upper bounds on $\mathbf{k}$.

In this paper, the equality constraints are normalized on a comparable scale:

$$F_i = \frac{f_i - \bar{f}_i}{\max(\epsilon, |\bar{f}_i|)}, \quad (2)$$

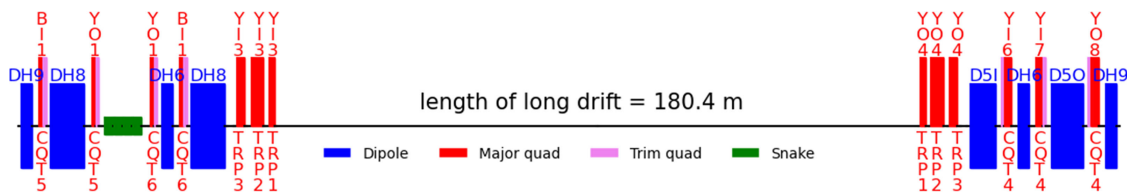


FIG. 2. Optimized HSR-IR2 layout used for lattice design. This configuration achieves the longest possible usable drift section. The placement of dipoles, quadrupoles, and the Siberian snake is shown schematically. Quadrupole and dipole names are labeled near their respective magnet blocks. These names follow RHIC naming conventions; for further details, refer to the RHIC Configuration Manual [11].

where f_i is the quantity to be matched, $\bar{f}_i$ stands for the target value, and ϵ is a user-defined floor to prevent overweighting small targets; we use $\epsilon = 0.01$.

Similarly, inequality constraints are handled using a one-sided penalty formulation. Specifically,

$$F_{m+i} = \frac{\max(0, c_i - \bar{c}_i)}{\max(\epsilon, |\bar{c}_i|)}. \quad (3)$$

The matching is performed by minimizing the objective vector $\mathbf{F}$ using the Levenberg-Marquardt algorithm [12], or minimizing $\|\mathbf{F}\|$ via global search methods such as differential evolution [13].

The variable space $\mathbf{k}$ consists of $n = 20$ tunable knobs. The number of hard matching conditions, i.e., independent components of $\mathbf{f}(\mathbf{k})$, is typically $m = 9$ or 12, making the system under-determined. This redundancy allows multiple valid solutions at fixed energies, offering flexibility for optics matching.

However, it complicates ramping. Solutions at injection and top energy may occupy disconnected volumes of k space. In addition, the bounds $\mathbf{k}_l$ and $\mathbf{k}_u$ are energy dependent; a solution feasible at one energy may fall outside the feasible range at next energy step. As a result, the optimization could fail to converge or yield discontinuous solutions, disrupting the continuity of the ramping path. These challenges motivate the need for a robust ramping strategy.

III. INTERMEDIATE-PENALTY METHOD

If the high-energy cooling is not included, the beam optics at the center of the long drift is only constrained at injection. According to Table II, the LEC requires the beta functions to lie within the range $100 \text{ m} \leq \beta_{x,y}^* \leq 200 \text{ m}$, where the superscript “*” denotes at the center of the long drift. However, the aperture constraints listed in Table III strongly favor solutions near the lower bound of this range. Consequently, the optimizer is naturally driven toward $\beta_{x,y}^* \approx 100 \text{ m}$.

The LEC also allows a range for the horizontal dispersion and its derivative, as shown in Table II. However, due to the large beta functions in this region, the optics are effectively close to a dispersion-free condition. Therefore, the matching constraints are simplified by imposing:

$$\beta_{x,y}^* = 100 \text{ m}, \quad \alpha_{x,y}^* = 0, \quad \eta_x^* = 0, \quad \eta_x'^* = 0. \quad (4)$$

Figure 3 shows the matched optics solution along the IR2 region. The Twiss parameters β_x , β_y , η_x , and η_x' at the center of the long drift, as well as at the region boundaries, satisfy the target values specified in Eq. (4) and Table II. The aperture requirements listed in Table III are also fulfilled throughout the region.

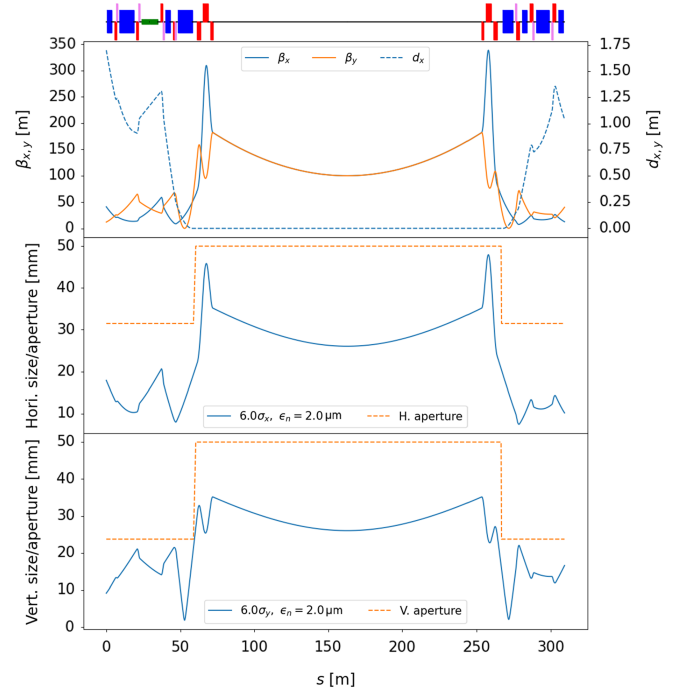


FIG. 3. Injection beam optics without RC constraints. Top: β -functions and horizontal dispersion; middle: horizontal beam size and aperture limit; bottom: vertical beam size and aperture limit.

Although the beam optics can be successfully matched at injection energy, the resulting solution introduces challenges for constructing a smooth ramping path. In particular, many trim quadrupoles operate with opposite polarity to their neighboring main quadrupoles. This configuration effectively forms local doublets, where the trim and main quadrupoles act in opposition to control the horizontal and vertical beam sizes simultaneously.

However, these quadrupole pairs are positioned very close to each other, resulting in nearly identical phase advances. From the optics perspective, they function almost as a single element, and to achieve any significant differential effect, both quadrupoles must operate at high strength to produce a small net focusing.

In the matched solution shown in Fig. 3, the strongest quadrupole pair operates at normalized focusing strengths of 0.19378 m^{-2} and -0.18541 m^{-2} , respectively. These values are several times larger than the hardware-imposed limits at top energy. It is therefore impossible to ramp these quadrupoles directly with increasing energy.

Although the beam optics at top energy is not strictly constrained in this scenario, it is beneficial to impose additional reasonable conditions to ensure the optics is well behaved and avoids extreme configurations.

As always, the Twiss parameters at both boundaries listed in Table II must be matched. In addition, the following conditions are introduced:

$$\alpha_{x,y}^* = 0, \quad \eta_x^* = 0, \quad \beta_{x,y} < 400 \text{ m}. \quad (5)$$

Here, the conditions $\alpha_{x,y}^* = 0$ and $\eta_x^* = 0$ ensure symmetry of the beta and dispersion functions across the long drift. The constraint $\beta_{x,y} < 400$ m is consistent with the injection beam optics in Fig. 3 and maintains a reasonable upper bound on the beta functions.

After obtaining the matched optics at top energy, a common strategy for constructing an energy ramp is to interpolate between the injection and top-energy solutions. This is the approach used in the RHIC ramping system, where step stones store set points for each device, and cubic interpolation is applied between stones to define intermediate values [14].

Another example is the LHC beta squeeze at the top energy. As shown in [15], the LHC uses linear interpolation with parabolic round-offs of magnet currents between matched optics points.

However, this strategy implicitly assumes that the matched optics at both ends lie on a continuous solution branch in the variable space. In our case, the constraints defined in Eq. (5) do not guarantee that the injection and top-energy configurations are connected in $\mathbf{k}$ space. Therefore, direct interpolation between them often fails to produce a feasible ramping path.

Figure 4 presents an example where the top-energy optics is successfully matched, but the solution is not continuously connected to the injection optics obtained earlier. The top plot shows the Twiss parameters at top energy, satisfying both the matching conditions in Eq. (5) and the boundary conditions listed in Table II.

In the bottom plot of Fig. 4, we evaluate the fitness function at intermediate ramping fractions by linearly interpolating the quadrupole strengths between injection and top configurations. Specifically, the interpolated strength vector is defined as

$$\mathbf{k}(t) = t \cdot \mathbf{k}_{\text{top}} + (1 - t) \cdot \mathbf{k}_{\text{inj}}, \quad 0 \leq t \leq 1, \quad (6)$$

where $t = 0$ corresponds to injection configuration $\mathbf{k}_{\text{inj}}$ and $t = 1$ to top configuration $\mathbf{k}_{\text{top}}$. The fitness is evaluated as the squared two-norm of the objective vector: $\|\mathbf{F}\|^2$.

When t is close to 0 or 1, the fitness remains near 0, as the interpolated quadrupole configuration stays close to the respective matched solutions. However, the fitness increases to a large number in the middle between them. As shown in Eqs. (2) and (3), the objective and constraint terms have been normalized, yet the fitness at intermediate state exceeds 10^5 , indicating a severe violation of the matching conditions.

This large increase demonstrates that the interpolated quadrupole configuration passes through a region of parameter space where the optics constraints cannot be satisfied. Importantly, this behavior persists across different interpolation schemes, including smooth alternatives such as cubic splines. It suggests that the injection and

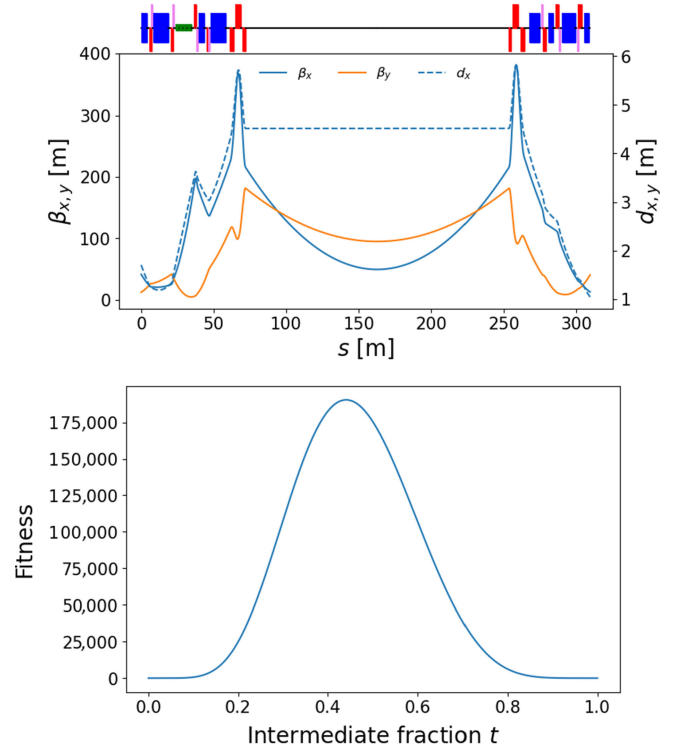


FIG. 4. Matched top-energy (110 GeV) beam optics configuration that cannot be connected to the injection solution through interpolation. Top: matched β -functions and horizontal dispersion at top energy, satisfying the constraints in Eq. (5) and Table II. Bottom: fitness evaluated at intermediate ramping fractions using linear interpolation in quadrupole strength, as defined in Eq. (6). The large fitness in the middle of the ramp indicates that this top-energy solution is not feasible.

top-energy optics lie on disconnected branches of the feasible solution manifold in $\mathbf{k}$ space.

Even if another valid ramping path exists, it would be located far from the region defined by the two endpoints and require entirely different quadrupole settings. Given the hardware-imposed range constraints, such a trajectory would be physically inaccessible. Therefore, this matched top-energy configuration in Fig. 4 is infeasible.

To ensure that the injection and top-energy optics configurations can be connected through a feasible and continuous ramping path, we modify the objective function for optimizing the top-energy solution by adding a penalty term at an intermediate ramping fraction $t = 0.5$. Specifically, the optimization is expressed as

$$\mathbf{k}_{\text{mid}} = \frac{\mathbf{k}_{\text{inj}} + \mathbf{k}_{\text{top}}}{2}$$

$$\text{minimize}_{\mathbf{k}_{\text{top}}} \left\{ \|\mathbf{F}(\mathbf{k}_{\text{top}})\|^2 + \max \left(0, \frac{\|\mathbf{F}(\mathbf{k}_{\text{mid}})\| - \delta}{\delta} \right)^2 \right\}, \quad (7)$$

where δ is a user-defined tolerance; we set $\delta = m$ equal to the number of hard matching constraints defined in Eq. (5) and Table II.

The first term in Eq. (7) tries to fulfill the constraints $\mathbf{F}(\mathbf{k}) = \mathbf{0}$ at top energy while the second penalty term ensures that the interpolated path $\mathbf{k}(t)$ never strays too far from the constraints.

The penalty becomes active only when the fitness at the midpoint exceeds the tolerance δ . It promotes smooth variation of $\|\mathbf{F}[\mathbf{k}(t)]\|$ along the ramp. A large midpoint violation implies a sharp change in constraint violation with respect to t and suggests that the feasible region is disconnected along the interpolation path.

This approach effectively biases the optimizer toward configurations that support a continuous and physically realistic ramping trajectory. Figure 5 shows one such top-energy solution, where the interpolated quadrupole settings yield low fitness values at all intermediate steps, confirming the existence of a smooth and valid ramp.

Based on the feasible solution shown in Fig. 5, a set of stones is selected to serve as anchor points in the ramping

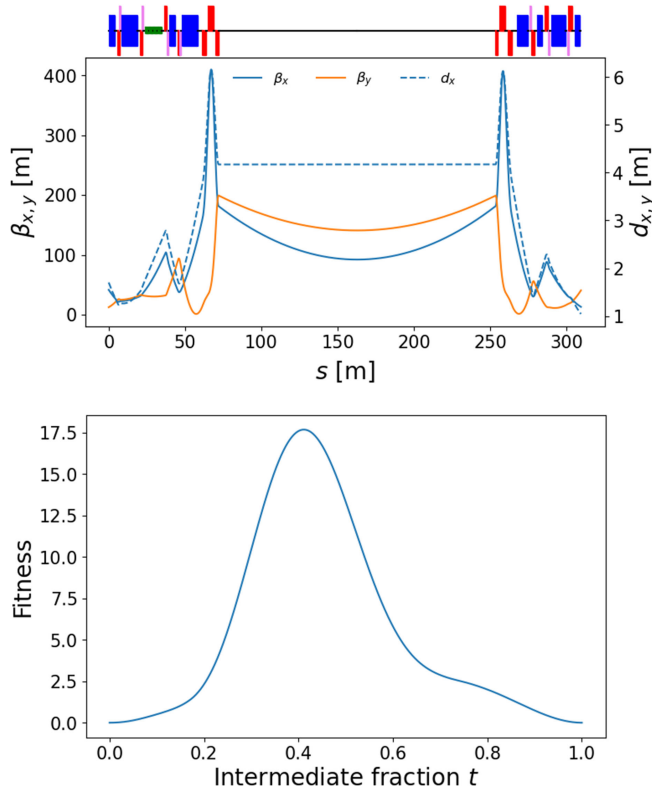


FIG. 5. Matched top-energy optics (110 GeV) with feasible connection to injection solution. Top: matched β -functions and horizontal dispersion at top energy, satisfying the constraints in Eq. (5) and Table II. Bottom: fitness $\|\mathbf{F}[\mathbf{k}(t)]\|^2$ as a function of intermediate fractions during linear interpolation in quadrupole strength, as defined by Eq. (6). The fitness remains low across the ramp, demonstrating a smooth and feasible connection between injection and top-energy optics.

process. Between stones, a cubic spline interpolation is applied to the quadrupole current to ensure smooth variation within hardware limits. The beam energy is incremented in uniform steps of 0.5 GeV. At each energy step, a local optimizer is applied to fine-tune the quadrupole strengths so that the full set of optics constraints of Eq. (5) and Table II is satisfied. The final ramping current profiles for all quadrupoles are illustrated in Fig. 6, demonstrating that the proposed method is practically effective.

The intermediate-penalty method is not restricted to the specific optics matching problem presented here. Its formulation relies only on the existence of two endpoint configurations and the ability to define an interpolation path between them. The penalty can be placed at any intermediate point along this path (not necessarily at the midpoint in this paper), depending on where the transition between constraint sets becomes most difficult. In fact, multiple intermediate penalty points can be introduced if the constraint landscape requires better control along the interpolation path.

The penalty tolerance δ , which controls the softness of the matching requirement at the intermediate point, is problem dependent and should be tuned based on the structure and scale of the constraint residuals. This

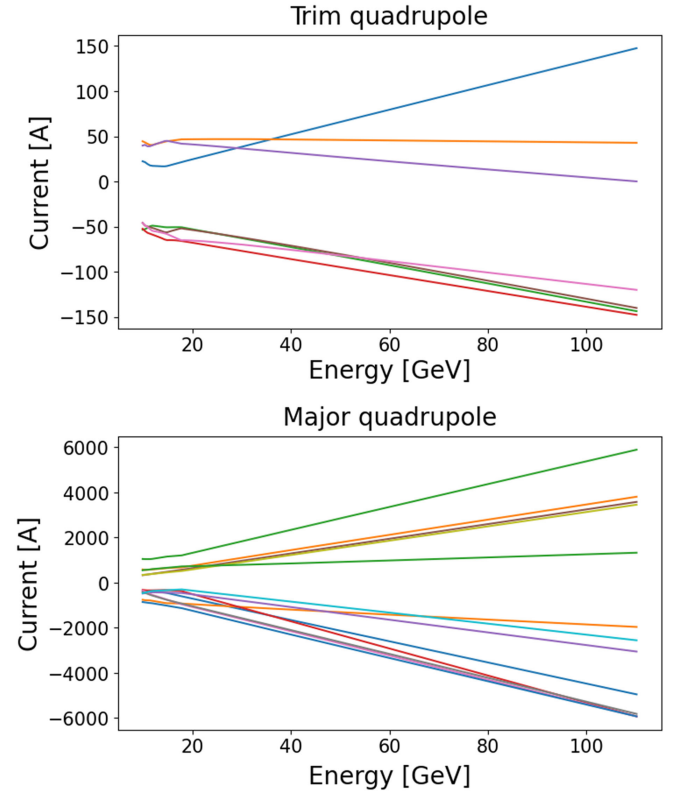


FIG. 6. Quadrupole currents during the energy ramping from injection to top energy, top: the current evolution of the trim quadrupoles, bottom: current of the main quadrupoles. Each curve corresponds to an individual quadrupole in the IR2 region.

flexibility allows the method to be adapted to a wide range of applications.

More generally, the approach can be viewed as a structured way to probe the connectivity of feasible regions in under-constrained optimization problems. By augmenting the objective function with a localized penalty at a strategically chosen intermediate point, the optimizer is guided toward solutions that respect both endpoint constraints while maintaining continuity. This makes the intermediate-penalty method applicable to a broader class of lattice design problems or other scenarios where the goal is to interpolate between two known configurations.

IV. CONTINUATION-BASED METHOD

This section presents a method to construct a ramping path that satisfies the significantly different matching constraints imposed by the low-energy cooler (LEC) at injection and the ring cooler (RC) at top energy.

A. Matching for LEC and RC

In this scenario, the LEC and the RC impose different optics requirements at the center of the long drift, as shown in Table II. As a result, the strategy described in Sec. III, which defines a common fitness constraint across the entire ramping process, is no longer applicable.

If only loose constraints are imposed (e.g., $\beta_y < 1000$ m), they fail to sufficiently reduce redundancy, offering little guidance to the optimizer. Conversely, enforcing strict constraints—such as requiring that the Twiss and dispersion functions at the drift center interpolate smoothly between the injection and top-energy optics—can overconstrain the problem and render it intractable, given the bounded quadrupole strength.

Most critically, without clearly defined intermediate constraints, it becomes impossible to determine whether the injection and top-energy optics lie on a connected solution branch in $\mathbf{k}$ space. The previous midpoint-penalty strategy cannot be directly applied in this setting.

Instead of independently matching the beam optics at injection and top energy and constructing the ramping path by interpolating the quadrupole strengths $\mathbf{k}$, we adopt a different strategy. To avoid disconnected solutions in $\mathbf{k}$ space, we begin by matching the top-energy optics, as shown in Fig. 7, and then gradually transition toward satisfying the LEC constraints at injection.

This top-down approach is motivated by two considerations. First, the top-energy requirements listed in Table II are more difficult to satisfy. Second, the allowed quadrupole strength range is wider at lower energy, ensuring that the high-energy solution remains within bounds as the energy decreases. This helps prevent local optimization from failing due to initial guesses violating variable limits.

The injection constraints are categorized into two groups. The first, $\mathbf{F}_{\text{ramp}}$, includes the boundary conditions

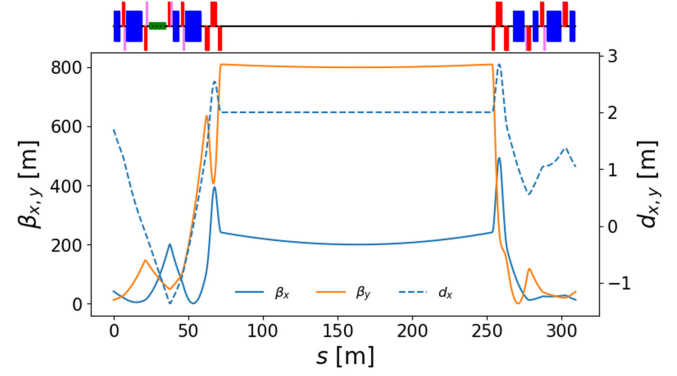


FIG. 7. Matched top-energy (110 GeV) beam optics configuration satisfying quadrupole strength limitations in Table IV and the beam-optics constraints in Table II.

in Table II and must be satisfied throughout the entire ramp. The second, $\mathbf{F}_{\text{inj}}$, includes the aperture constraints in Table III and the low-energy optics targets at the drift center, also defined in Table II, and is enforced only at injection energy.

To implement the gradual transition, we introduce a weighting parameter λ on the injection-specific constraints $\mathbf{F}_{\text{inj}}$. Instead of minimizing the full objective vector $\mathbf{F}$, we minimize the augmented objective vector:

$$\mathbf{F}_\lambda(\mathbf{k}) = \begin{bmatrix} \mathbf{F}_{\text{ramp}} \\ \lambda \cdot \mathbf{F}_{\text{inj}} \end{bmatrix}, \quad (8)$$

where λ is the tunable weight controlling the importance of injection-specific constraints in the optimization.

Here the optimization first tries to find a good solution on the ramp by minimizing $\mathbf{F}_{\text{ramp}}$ because λ is small. Then, as λ is dynamically increased, more weight is given to satisfy the injection constraint $\mathbf{F}_{\text{inj}}$ while still satisfying $\mathbf{F}_{\text{ramp}}$ at the same time. Slowly increasing λ thus shifts the focus of the optimizer to injection, but prevents it to lose sight of the constraints on the ramp.

As shown in Algorithm 1, an adaptive strategy is used to automatically select an appropriate λ that balances feasibility and convergence. Starting from a small λ , we iteratively update it using a randomized scaling factor:

$$\lambda \leftarrow \lambda \times 2^{\text{rand}() - 0.49}, \quad (9)$$

where $\text{rand}()$ generates a random number uniformly in $(0, 1)$. This introduces bidirectional perturbation, with a slight bias toward increasing λ . A fixed λ often causes the algorithm to converge to a local minimum. In contrast, the randomized scaling promotes exploration of the solution space and helps the optimizer escape local minima.

At each iteration, the Levenberg-Marquardt algorithm is applied within a small neighborhood around the current solution to update the quadrupole strengths $\mathbf{k}$.

ALGORITHM 1. Adaptive λ -weighted optics matching strategy for ramping constraints.

Require: $\lambda_{\max} = 10$
Require: Top-energy solution $\mathbf{k}_{\text{top}}$, range bounds $\mathbf{k}_l, \mathbf{k}_u$
Require: Constraint function $\mathbf{F}_\lambda(\mathbf{k}) = [\mathbf{F}_{\text{ramp}}, \lambda \cdot \mathbf{F}_{\text{inj}}]^\top$
Require: Search radii of quadrupole strength in each step $\Delta \mathbf{k} = 10^{-3} \text{ m}^{-2}$
Require: Tolerances $\text{lo_tol} = 10^{-3}$ (target fit quality) and $\text{hi_tol} = 10^{-2}$ (acceptable maximum during λ search)

- 1: Initialize $\lambda \leftarrow 1, \mathbf{k} \leftarrow \mathbf{k}_{\text{top}}$
- 2: **while** $\lambda < \lambda_{\max}$ **do**
- 3: $\lambda \leftarrow \lambda \times 2^{\text{rand}() - 0.49}$ $\triangleright$ Randomized λ scaling to promote exploration
- 4: Compute local bounds: $\mathbf{k}'_l = \max(\mathbf{k} - \Delta \mathbf{k}, \mathbf{k}_l), \mathbf{k}'_u = \min(\mathbf{k} + \Delta \mathbf{k}, \mathbf{k}_u)$
- 5: Run local optimizer to solve $\min_{\mathbf{k}} \|\mathbf{F}_\lambda(\mathbf{k})\|^2$ within bounds
- 6: Record new solution $\mathbf{k}'$ and objective $J = \|\mathbf{F}_\lambda(\mathbf{k}')\|^2$
- 7: **while** $J > \text{hi_tol}$ **do**
- 8: $\lambda \leftarrow \lambda/2$
- 9: Reoptimize with updated λ
- 10: **end while**
- 11: **if** $J < \text{lo_tol}$ **then**
- 12: $\lambda \leftarrow 2 \cdot \lambda$ $\triangleright$ Bias toward increasing λ if fit is very good
- 13: **end if**
- 14: Update: $\mathbf{k} \leftarrow \mathbf{k}'$
- 15: **end while**

If the resulting fit exceeds a high threshold, λ is halved, and the process repeats until an acceptable λ is found. If the fit falls below a low threshold, λ is doubled to further enforce injection constraints.

The iterative process terminates when λ exceeds a preset upper limit (e.g., 10), indicating that the injection optics are sufficiently well approximated while all ramping constraints are maintained.

The evolution of λ during the optimization process is illustrated in Fig. 8. For the majority of iterations, λ remains

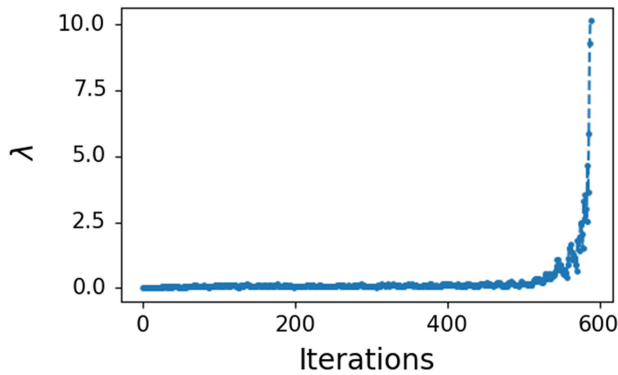


FIG. 8. Evolution of the weight parameter λ during adaptive optimization. The stochastic update allows the optimizer to escape local minima and better balance between the ramping and injection constraints.

small and exhibits oscillatory behavior, allowing for sufficient exploration of the solution space and enabling the optimizer to balance between $\mathbf{F}_{\text{ramp}}$ and $\mathbf{F}_{\text{inj}}$. Toward the final stage of the search, λ rapidly increases over the course of several tens of iterations, reflecting a transition into the convergence phase where the full set of constraints is actively enforced.

By setting $\lambda = 0$, another local optimizer is applied to each solution obtained during the adaptive iteration. This yields the solution set

$$\mathcal{S} = \{\mathbf{k} | \mathbf{F}_{\text{ramp}}(\mathbf{k}) = \mathbf{0}\}, \quad (10)$$

which satisfies all ramping constraints across the process. The top-energy optics shown in Fig. 7 and the final matched injection optics shown in Fig. 9 are both elements of this set. Since our adaptive search algorithm constrains the quadrupole strengths within a small neighborhood during each update, the resulting solutions remain close in parameter space. Therefore, a continuous ramping path

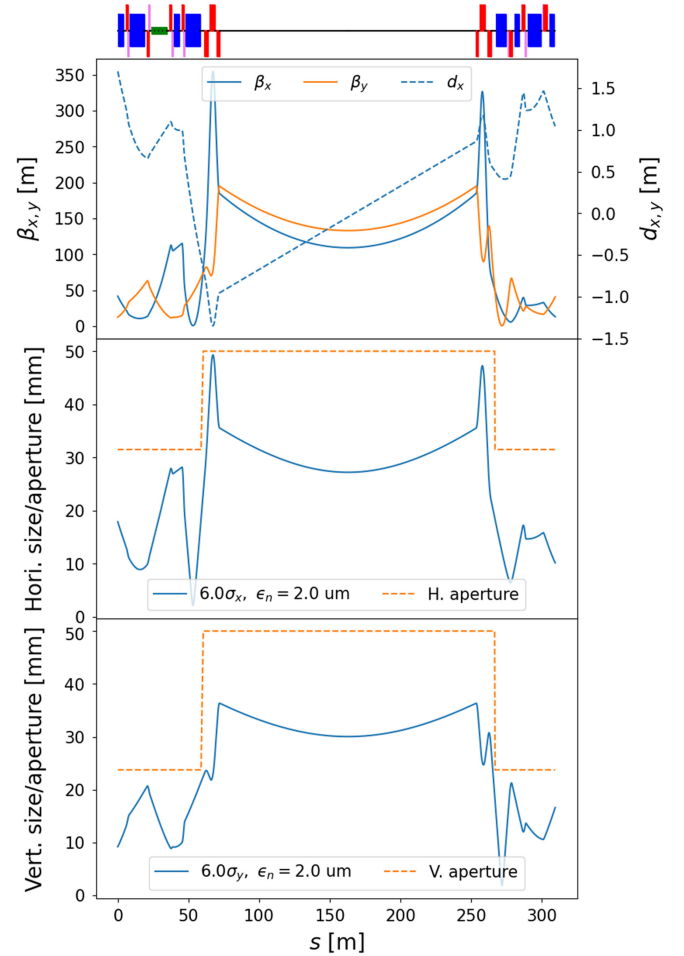


FIG. 9. Injection beam optics ramped down from top-energy configuration in Fig. 7. Top: β -functions and horizontal dispersion; middle: horizontal beam size and aperture limit; bottom: vertical beam size and aperture limit.

ALGORITHM 2. Constructing a smooth ramping path via Dijkstra and local refinement.

Require: Candidate solutions $\mathcal{S}$

Require: Distance threshold $\Delta = 0.1 \text{ m}^{-2}$

- 1: Build a graph G with nodes $\mathbf{k}_i \in \mathcal{S}$, connect edge if $\|\mathbf{k}_i - \mathbf{k}_j\| < \Delta$
 - 2: Use Dijkstra's algorithm to find shortest path from top energy to injection optics
 - 3: Let $\{\mathbf{k}_{s_1}, \dots, \mathbf{k}_{s_M}\}$ be the selected path
 - 4: Construct cubic spline interpolator over $\{\mathbf{k}_{s_j}\}$ to get $\mathbf{k}(E)$
 - 5: **for** each energy step $E \in [E_{\text{inj}}, E_{\text{top}}]$ **do**
 - 6: Use interpolated $\mathbf{k}(E)$ as initial guess
 - 7: Apply Levenberg-Marquardt optimizer to minimize $\|\mathbf{F}_{\text{ramp}}(\mathbf{k}(E))\|$
 - 8: Save refined $\mathbf{k}(E)$
 - 9: **end for**
 - 10: **Output:** Smooth and feasible quadrupole ramp $\{\mathbf{k}(E)\}$
-

connecting the injection and top-energy configurations can be constructed within $\mathcal{S}$.

Algorithm 2 outlines the procedure for constructing a smooth ramping path from the solution set $\mathcal{S}$. A graph is built by connecting nearby points in $\mathcal{S}$, and Dijkstra's algorithm [16] is applied to identify the shortest path from the top-energy optics to the injection optics. This path defines a sequence of intermediate optics configurations, which serve as anchor points for cubic spline interpolation. The interpolated quadrupole strengths are then used as initial guesses at each discrete energy step. This ensures that the resulting ramping path remains both smooth and physically feasible across the entire energy range. The corresponding quadrupole current profiles during the ramping process are shown in Fig. 10, illustrating the continuous evolution of both trim and main quadrupoles.

B. Relation to homotopy method

The method used here can be interpreted as a continuation-based approach inspired by homotopy method [17].

In traditional homotopy method, a nonlinear system is solved by constructing a continuous deformation (homotopy) from a simpler problem with a known solution. One then traces the solution path as a parameter evolves from 0 to 1. The key idea is to preserve continuity and feasibility along the solution path.

An introduction to the homotopy method can be found in [18]. The key concepts are briefly summarized here for readers' convenience.

Let $\mathbf{F}(\mathbf{x})$ be a complicated, nonlinear, high-dimensional system of full rank for which the solutions of $\mathbf{F}(\mathbf{x}) = \mathbf{0}$ are desired. Suppose there exists another system $\mathbf{G}(\mathbf{x})$ whose solutions are known and which possesses at least as many solutions as $\mathbf{F}$. The homotopy method constructs a continuous deformation between the two systems via a

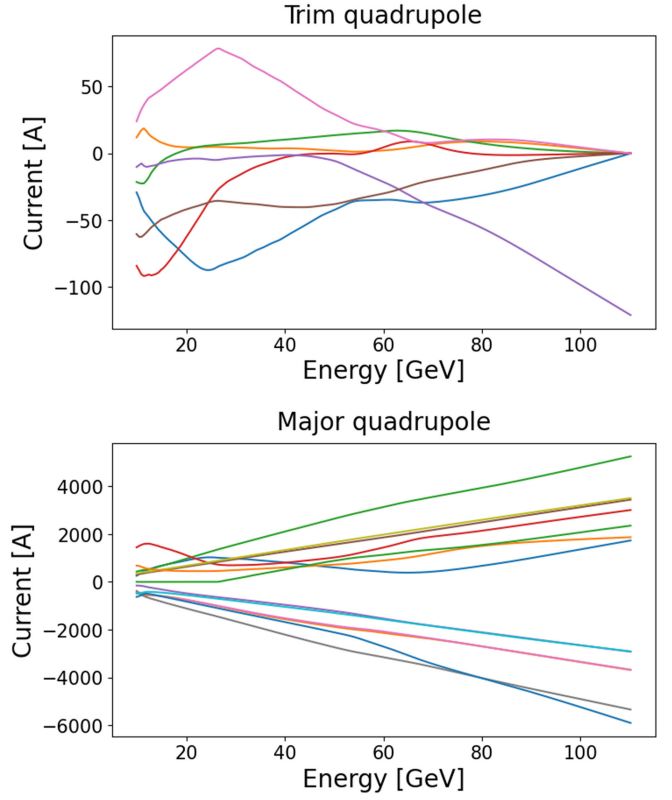


FIG. 10. Quadrupole currents during the energy ramping from injection configuration in Fig. 9 to top-energy configuration in Fig. 7, top: the current evolution of the trim quadrupoles, bottom: current of the main quadrupoles. Each curve corresponds to an individual quadrupole in the IR2 region.

homotopy function $\mathbf{H}(\mathbf{x}, t)$, such that $\mathbf{H}(\mathbf{x}, 1) = \mathbf{F}(\mathbf{x})$ and $\mathbf{H}(\mathbf{x}, 0) = \mathbf{G}(\mathbf{x})$.

Differentiating $\mathbf{H}(\mathbf{x}, t)$ with respect to t yields

$$\frac{d\mathbf{x}}{dt} = -(\nabla_{\mathbf{x}}\mathbf{H})^{-1} \cdot \frac{\partial \mathbf{H}}{\partial t}, \quad (11)$$

where $\nabla_{\mathbf{x}}\mathbf{H}$ denotes the Jacobian of $\mathbf{H}$ with respect to $\mathbf{x}$, assumed to be square and invertible along the path.

Equation (11) defines ordinary differential equations (ODEs) for the evolution of $\mathbf{x}(t)$. By any numerical method solving ODEs from $t = 0$ to $t = 1$, the solution $\mathbf{x}(0)$ of $\mathbf{G}(\mathbf{x}) = \mathbf{0}$ is continuously transformed into a solution $\mathbf{x}(1)$ satisfying $\mathbf{F}(\mathbf{x}) = \mathbf{0}$.

Amstutz *et al.* applied the homotopy method to compute the required quadrupole corrections that ensure continuous optics compensation in a LINAC beam line [19]. However, directly applying the homotopy method to general lattice design problems poses several challenges.

First, most lattice design codes operate as black-box tools, and the analytical derivatives or Jacobian matrix required in Eq. (11) are not available. Although this is not a fundamental limitation—since automatic differentiation

tools are increasingly developed in the accelerator community, such as JuTrack, developed by Wan *et al.* [20]—the tool used for the HSR-IR2 lattice design does not currently support such capabilities.

Second, many of the constraints in the HSR-IR2 design are soft constraints, such as controlling beam sizes at injection. These are implemented via one-sided penalty functions, as shown in Eq. (3). Consequently, the associated derivatives are not smooth, especially near constraint boundaries, which violates assumptions commonly used in the homotopy framework.

Third, the problem studied here is under-constrained: the number of available knobs exceeds the number of hard constraints. As a result, the Jacobian matrix in Eq. (11) is not square and hence not directly invertible, posing difficulties for standard homotopy continuation methods.

Nevertheless, the under-constrained nature of the system offers a distinct advantage: it increases the likelihood of the existence of a smooth ramping path. As discussed in [19], when the number of knobs equals the number of constraints, and the Jacobian remains nonsingular (i.e., $\det \nabla_x \mathbf{H} \neq 0$), a continuous solution path exists locally up to a maximum parameter $t = t_{\max}$. Beyond this point, the determinant vanishes, and the continuation path breaks down.

In contrast, for under-constrained problems, additional degrees of freedom allow one to select one subset of the knobs to preserve the nonsingularity condition along the path. This ensures the existence of continuous ramping path between the starting and target configurations, though such a path is not unique.

In the proposed method, as outlined in Algorithm 1, the parameter λ plays a role analogous to the homotopy parameter t in classical continuation methods. As λ gradually increases from a small value to a large one, the objective vector transitions from encoding top-energy constraints to injection constraints. Correspondingly, the solution evolves from the known top-energy configuration to the unknown injection-compatible configuration that we aim to construct.

Unlike traditional homotopy approaches, which solve the path following ODEs defined by the homotopy function, our method applies a local optimizer at each discrete step in λ . This avoids the need for Jacobian computation and ODE integration, making the method applicable even with black-box lattice codes.

To mitigate the risk of the local optimizer becoming trapped in poor local minima, the parameter λ is updated stochastically, with a slight bias toward increasing values, as shown by Eq. (9). This introduces a level of robustness and improves the success rate in challenging under-constrained problems like HSR-IR2 design.

V. SUMMARY AND DISCUSSION

This study presents a complete lattice design and optics ramping framework for the IR2 region in the HSR of the

EIC. The key physical motivation is to provide a long drift section for co-propagating electrons and hadrons, which is essential for effective electron cooling. A systematic permutation of RHIC magnet cryomodules is compared to maximize the usable drift length.

We show that the optics design problem is fundamentally under-determined: the number of quadrupole knobs exceeds the number of matching constraints. The separately matched injection and top-energy optics usually belong to disconnected branches in the parameter space. This poses a significant challenge when constructing a continuous and hardware-feasible ramping path.

When top-energy optics is not heavily constrained, we incorporate a penalty term at intermediate states to ensure continuity and feasibility between injection and top-energy optics. When constraints of top-energy cooling are active, we introduce an adaptive optimization scheme that tunes a weighting parameter to gradually enforce low-energy requirements while remaining within hardware limits. A shortest-path search in the quadrupole parameter space is applied to extract a smooth configuration sequence, and local optimizers are used to finalize the ramping path.

The final result satisfies all optics and hardware constraints, while preserving smooth evolution of quadrupole strengths. These two methods are general and applicable to other ramping problems where the solution space is under-constrained and the top-energy solution lies far from the low-energy configuration in variable space. Testing the two methods on other under-constrained design problems may further demonstrate their generality and robustness, which could be explored in future work.

The intermediate-penalty method is effective for simulation-based feasibility screening, while the continuation-based strategy is suited for practical applications. We will also explore the feasibility of adapting the second method for integration with online lattice tuning in realistic ramping scenarios in the future.

ACKNOWLEDGMENTS

The author would like to thank J. S. Berg, G. Robert-Demolaize, and V. Ptitsyn for insightful discussions; J. S. Berg, C. Liu, and A. Fedotov for providing key constraints that define the matching process; and S. G. Peggs, C. Liu, and Y. Li for carefully reviewing the manuscript and offering valuable suggestions. This work was supported by Brookhaven Science Associates, LLC under Contract No. DE-SC0012704 with the U.S. Department of Energy, and by a Department of Energy Early Career Award.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available. The data are available from the authors upon reasonable request.

APPENDIX A: LATTICE LAYOUT AND GEOMETRY REQUIREMENTS

The placement of the Siberian snake in IR2 strongly constrains the magnet layout. The snake and its adjacent quadrupoles occupy about 20 m. To maintain 180° separation from the IR8 snake, IR2 must reserve alignment slots for both colliding and noncolliding IR8 configurations [21,22]. This reduces the maximum available drift to approximately 160 m. Therefore, this study considers only the noncolliding IR8 configuration, consistent with the current EIC baseline. A future detector at IR8 would likely require reconfiguring the IR2 layout.

The HSR-IR2 layout must provide sufficient space for low-energy electron cooling. In this scheme, hadrons co-propagate with electrons at the same average velocity, and Coulomb interactions reduce the hadron emittance [23]. The cooling rate scales with the effective overlap length between the two beams. To satisfy this requirement, a systematic layout search is performed over permutations of dipoles, quadrupoles, the Siberian snake, and drift elements.

The IR2 region includes eight dipole magnets labeled by the RHIC naming convention: D9, D8, D6, D6, D5O, D5I, D8, and D9. In each candidate layout, two dipoles (D9 and D8) are placed before the Siberian snake and the remaining six after it. Accounting for the two identical D6 magnets, the number of unique dipole arrangements is $2! \times \frac{6!}{2} = 720$. Seven insertion points between dipoles are available for placing quadrupoles, the snake, and the long drift. There are two possibilities to place the long drift: (i) (Q, snake, long, Q, Q, Q, Q) and (ii) (Q, snake, Q, long, Q, Q, Q).

Here Q denotes a standard RHIC CQS or CQT cryomodule, which contains a standard quadrupole.

The Siberian snake occupies a reserved length of 12.5 m. One quadrupole is placed on each side of the snake to control the horizontal and vertical beam sizes. The long drift section includes two large-aperture triplets at each end. To fully control the six transverse optical parameters at the center of the long drift—namely, β_x , α_x , β_y , α_y , η_x , and η'_x —a minimum of six quadrupoles is required on each side of the drift. Both candidate layouts satisfy this: one places six quadrupoles left of the long drift and seven on the right; the other reverses this arrangement.

Although permutations of quadrupole types could yield more layouts, they are not considered here, as they cause only minor geometric variations.

A short adjustable drift is added to ensure geometric closure. It can be placed in eight locations: before the first or after the last dipole (two options), or adjacent to one of the quadrupoles or the snake (six options). The total number of configurations is

$$720 \times 2 \times 8 = 11,520. \quad (\text{A1})$$

The lengths of the short and long drifts are adjusted to meet geometric boundary conditions at YO1 and YI2, the

TABLE I. Geometry boundary conditions at YO1 and YI2, which mark the upstream and downstream interfaces to the arc, respectively. Although the hadron beam circulates counterclockwise, the lattice is designed in the clockwise direction. The symbols x , z , and θ follow the MAD-X SURVEY convention [24].

Parameter	YO1	YI2
$x(\text{m})$	427.705516	159.486753
$z(\text{m})$	431.526825	585.348499
$\theta(\text{rad})$	-0.951763070	-1.142632032

upstream and downstream interfaces to the arc. These conditions, based on the MAD-X SURVEY convention, are listed in Table I.

All 11,520 configurations are evaluated. A valid layout requires positive drift lengths. Among feasible solutions, 777 provide a long drift exceeding 170 m. The configuration with the longest usable drift is shown in Fig. 2.

APPENDIX B: BEAM-OPTICS CONSTRAINTS

The Twiss functions must match the upstream and downstream arc optics throughout ramping. The required boundary conditions are listed in Table II. At the center of the long drift, the beam optics are optimized for electron cooling. At injection, the beta functions should lie within $100 \sim 200$ m, with near-zero dispersion, consistent with the standalone LEC design derived from LEReC experience at RHIC [6]. At top energy, two scenarios are considered. Without high-energy cooling, the optics at the drift center remain unconstrained. If the ring cooler is implemented, the optics must meet stricter requirements, including large vertical beta function $\beta_y^* > 800$ m and large horizontal dispersion $\eta_x^* \sim 2$ m. These target values are also summarized in Table II.

The physical aperture is limited by the RHIC vacuum chambers, upgraded with copper-coated beam screens to reduce resistive wall heating and suppress electron cloud effects [25]. A 6σ beam envelope is used as the design limit. This is challenging at injection due to large beta functions

TABLE II. Twiss parameters at the upstream (YO1), downstream (YI2), and center of the long drift. The values at YO1 and YI2 match arc optics across all energies; the values at the center of the long drift indicate optimal ranges for electron cooling at injection energy (LEC center), or at top energy (RC center).

Parameter	YO1	YI2	LEC center	RC center
β_x (m)	41.66	13.18	[100, 200]	200
α_x	1.97	0.74	0	0
η_x (m)	1.70	1.04	[-1, 1]	~ 2
η'_x	-0.0825	-0.044	[-0.02, 0.02]	0
β_y (m)	12.53	40.70	[100, 200]	> 800
α_y	-0.71	-1.97	0	0

TABLE III. Aperture constraints at injection. The 6σ beam envelope must remain within the physical aperture.

Parameter	Value
Species	Golden ion beam
Beam energy, E	10 GeV
Normalized emittance, ϵ_n	2.0 μm
Relative momentum spread, σ_δ	6×10^{-4}
Aperture at triplets	$6\sigma_x, 6\sigma_y \leq 50 \text{ mm}$
Aperture at other magnets	$6\sigma_x \leq 31.5 \text{ mm}$ $6\sigma_y \leq 23.75 \text{ mm}$

TABLE IV. Maximum quadrupole strength at top energy. The normalized focusing strength is defined as $k = \frac{1}{B\rho} \frac{\partial B_y}{\partial x}$.

Quadrupole type	Max current (A)	k [m^{-2}]
Main quadrupole	6000	0.098
Trim quadrupole	150	0.046

and the large emittance of low-energy beams. Large-aperture triplets are placed at both ends of the long drift to reduce beam size and control beta functions. The gold ion beam defines the most stringent aperture constraints and is used as the reference case. Table III summarizes the relevant beam parameters and aperture limits in this paper.

As shown in Fig. 2, the IR2 lattice uses two types of quadrupoles: standard RHIC quadrupoles (“main”) and individually powered trim quadrupoles. Each has a distinct transfer function due to differences in magnetic design and power supply configuration [11]. Table IV lists the quadrupole current following the RHIC operational practice and the corresponding normalized focusing strengths at top energy (110 GeV gold beam).

[1] M. S. Camillocci, S. Redaelli, R. Tomás, and J. Wenninger, Combined ramp and squeeze to 6.5 TeV in the LHC, in *Proceedings of the International Particle Accelerator Conference (IPAC-2016), Busan, Korea* (JACoW, Geneva, Switzerland, 2016), Vol. 7, pp. 1509–1512, [10.18429/JACoW-IPAC2016-TUPMW031](https://doi.org/10.18429/JACoW-IPAC2016-TUPMW031).

[2] V. Morozov *et al.*, Correction of the detector solenoid effect in the hadron storage ring of the electron-ion collider, in *Proceedings of the IPAC-2024, 15th International Particle Accelerator Conference No. 15* (JACoW, Geneva, Switzerland, 2024), pp. 156–159.

[3] F. Willeke and J. Beebe-Wang, Electron ion collider conceptual design report 2021, Technical Report, Brookhaven National Laboratory (BNL), Upton, NY; Thomas Jefferson National Accelerator Facility (TJNAF), Newport News, VA, Report No. BNL-221006–2021-FORE, 2021, [10.2172/1765663](https://doi.org/10.2172/1765663).

[4] S. R. Mane, Y. M. Shatunov, and K. Yokoya, Siberian snakes in high-energy accelerators, *J. Phys. G* **31**, R151 (2005).

[5] This paper follows the same naming convention for other regions, such as IR8.

[6] A. Fedotov, Z. Altinbas, S. Belomestnykh, I. Ben-Zvi, M. Blaskiewicz, M. Brennan, D. Bruno, C. Brutus, M. Costanzo, A. Drees *et al.*, Experimental demonstration of hadron beam cooling using radio-frequency accelerated electron bunches, *Phys. Rev. Lett.* **124**, 084801 (2020).

[7] W. Bergan, P. Baxevanis, M. Blaskiewicz, E. Wang, and G. Stupakov, Design of an MBEC cooler for the EIC, in *Proceedings of the IPAC-2021* (JACoW, Geneva, Switzerland, 2021), p. 1819.

[8] H. Zhao, J. Kewisch, M. Blaskiewicz, and A. Fedotov, Ring-based electron cooler for high energy beam cooling, *Phys. Rev. Accel. Beams* **24**, 043501 (2021).

[9] W. F. Bergan *et al.*, Coherent electron cooling physics for the EIC, in *Proceedings of the IPAC-2024, 15th International Particle Accelerator Conference No. 15* (JACoW, Geneva, Switzerland, 2024), pp. 2937–2942.

[10] Y. Luo, D. Xu, M. Blaskiewicz, and C. Montag, Experimental demonstration of a large transverse emittance ratio 11:1 in the relativistic heavy ion collider for the electron-ion collider, *Phys. Rev. Lett.* **132**, 205001 (2024).

[11] Brookhaven National Laboratory, *RHIC Configuration Manual* (Collider-Accelerator Department, Upton, 2006).

[12] M. Gomez, LeastSquaresOptim.jl, <https://github.com/matthieugomez/LeastSquaresOptim.jl> (2018).

[13] R. Feldt and A. Stukalov, BlackBoxOptim.jl, <https://github.com/robertfeldt/BlackBoxOptim.jl> (2025).

[14] J. Kewisch, J. van Zeijts, S. Peggs, and T. Satogata, Ramp management in RHIC, in *Proceedings of the 1999 Particle Accelerator Conference (Cat. No. 99CH36366)* (IEEE, New York, 1999), Vol. 2, pp. 705–707.

[15] S. Redaelli, M. Lamont, J. Wenninger, G. Muller, X. Buffat, M. Camillocci, and R. Tomas, Operation of the betatron squeeze at the LHC, in *Proceedings of the International Particle Accelerator Conference (IPAC-2013), Shanghai, China* (JACoW, Geneva, Switzerland, 2013), pp. 1430–1432.

[16] E. W. Dijkstra, A note on two problems in connexion with graphs, in *Edsger Wybe Dijkstra: His Life, Work, and Legacy* (Association for Computing Machinery, New York, NY, 2022), pp. 287–290, [10.1145/3544585.3544600](https://doi.org/10.1145/3544585.3544600).

[17] S.-J. Liao, The proposed homotopy analysis technique for the solution of nonlinear problems, Ph.D. thesis, Shanghai Jiao Tong University Shanghai, 1992.

[18] E. L. Allgower and K. Georg, *Introduction to Numerical Continuation Methods* (SIAM, Philadelphia, 2003).

[19] P. Amstutz, T. Plath, S. Ackermann, J. Bödwadt, C. Lechner, and M. Vogt, Confining continuous manipulations of accelerator beam-line optics, *Phys. Rev. Accel. Beams* **20**, 042802 (2017).

[20] J. Wan, H. Alamprese, C. Ratcliff, J. Qiang, and Y. Hao, JuTrack: A Julia package for auto-differentiable accelerator modeling and particle tracking, *Comput. Phys. Commun.* **309**, 109497 (2025).

[21] J. S. Berg *et al.*, Lattice design for the hadron storage ring of the electron-ion collider, in *Proceedings of the IPAC-2023, 14th International Particle Accelerator Conference No. 14* (JACoW, Geneva, Switzerland, 2023), pp. 903–905.

- [22] C. Liu *et al.*, Reconfiguration of RHIC straight sections for the EIC, in *Proceedings of the IPAC-2022*, International Particle Accelerator Conference No. 13 (JACoW, Geneva, Switzerland, 2022), pp. 1916–1919.
- [23] S. Nagaitsev, D. Broemmelsiek, A. Burov, K. Carlson, C. Gattuso, M. Hu, T. Kroc, L. Prost, S. Pruss, M. Sutherland, C. W. Schmidt, A. Shemyakin, V. Tupikov, A. Warner, G. Kazakevich, and S. Seletskiy, Experimental demonstration of relativistic electron cooling, *Phys. Rev. Lett.* **96**, 044801 (2006).
- [24] L. Deniau, H. Grote, G. Roy, and F. Schmidt, *The MAD-X Program User's Reference Manual* (JACoW, Geneva, 2023).
- [25] S. Verdú-Andrés *et al.*, A beam screen to prepare the RHIC vacuum chamber for EIC hadron beams: Conceptual design and requirements, in *Proceedings of the IPAC-2021*, International Particle Accelerator Conference No. 12 (JACoW, Geneva, Switzerland, 2021), pp. 2066–2069, [10.18429/JACoW-IPAC2021-TUPAB260](https://doi.org/10.18429/JACoW-IPAC2021-TUPAB260).