DISCLAIMER

BNL-228488-2025-TECH

EIC-ADD-TN-130

Electron Ion Collider Strong Hadron Cooling design Summary

E. Wang

July 2025

Electron-Ion Collider

Brookhaven National Laboratory

U.S. Department of Energy

USDOE Office of Science (SC), Nuclear Physics (NP)

DISCLAIMER

Electron Ion Collider Strong Hadron Cooling design Summary

Erdong Wang,* William Bergan, Panagiotis Baxevanis, Michael Blaskiewicz, Alexei Blednykh, Luca Cultrera, Jun Ma, Vadim Ptitsyn, Triveni Rao, Gang Wang, Ferdinand Willeke, and Derong Xu
Brookhaven National Laboratory, Upton, NY 11973,USA

Nick Sereno, Steve Benson, Z. A. Conway, Kirsten Deitrick, Bamunuvita Gamage, Jiquan Guo, Isurumali Neththikumara, Robert Rimmer, Todd Satogata, Sadiq Setiniyaz, and Shaohang Wang
Thomas Jefferson National Accelerator Facility, Newport News, VA, USA

Ji Qiang
Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Alexander Zholents
Argonne National Laboratory, Argonne, Illinois, USA

Gennady Stupakov
SLAC National Accelerator Laboratory, Menlo Park, CA, USA

Ningdong Wang
Cornell University, Ithaca, NY, USA

David Douglas, Joe Conway, Colwyn Gulliford, Christopher Mayes, and Nicholas Taylor
Xelera Research LLC, Ithaca, NY, USA

Robert Apsimon, Graeme Burt, Rancheng Bi, Nattaphong Nuchsirikulaphong, and Sara J. O'Toole
Lancaster University, Lancaster, UK and
Cockcroft Institute, Daresbury Laboratory, Warrington, UK

Chris Hall
RadiaSoft, Boulder, CO,USA

The Electron-Ion Collider (EIC) requires a high-energy cooler to maintain excellent beam quality and achieve high luminosity throughout long collision stores. To meet this requirement, the EIC project has adopted a novel approach known as Coherent Electron Cooling (CeC)—referred to as Strong Hadron Cooling (SHC)—which can provide rapid cooling rates at high energies. The SHC relies on an Energy Recovery Linac (ERL) to provide the intense, high-quality, and low-noise electron beam essential for the cooling process. This paper summarizes the design progress of the Strong Hadron Cooler for the EIC. We discuss key aspects of the project, including cooling physics modeling, luminosity evolution during collisions, the ERL design, risk mitigation strategies, remaining challenges, and ongoing R&D efforts.

I. Introduction

The US Department of Energy (DOE) has selected the Electron-Ion Collider (EIC), a new facility for studying the QCD frontier. Achieving the high average luminosity requires maintaining a small hadron beam emittance during collisions. This is accomplished through Strong Hadron Cooling (SHC) to counteract emittance growth from effects like intra-beam scattering (IBS), thus keeping the average luminosity nearly equal to its peak value of $1.03 \times 10^{34}\,\mathrm{s}^{-1}\,\mathrm{cm}^{-2}$ for e-p collisions at a center-of-mass energy (E_{cm}) of 105 GeV. Beam cooling is essential for substantially reducing the emittance of the proton or ion beams and preserving these parameters throughout the beam store.

Due to high particle densities, conventional stochastic cooling systems for the EIC's proton beams require a 10- to 100-fold increase in bandwidth to provide the cooling rates needed to suppress IBS-induced emittance growth and the corresponding luminosity reduction. This bandwidth is realized by replacing traditional hardware (pick-ups, cables, amplifiers, kickers) with a cold electron beam

* wange@bnl.gov

that functions as both a pickup and a kicker, exploiting electron beam dynamics to achieve the necessary wideband amplification. Other approaches are also under investigation.

Two primary hadron cooling systems are planned for the EIC: a Low-Energy Cooler (LEC) for hadrons at the injection energy of 24.5 GeV, and a high-energy cooler to maintain luminosity at collision energies of 100 GeV and 275 GeV. The latter is the main Strong Hadron Cooling (SHC) system for the project. This report summarizes the design progress of the SHC.

II. EIC high energy cooling requirements

The EIC SHC is essential for the EIC to achieve a luminosity of $1 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The key cooling requirements are:

1. Counteract longitudinal and transverse emittance growth during long stores. The cooling time must be less than or equal to the emittance growth time from all diffusion sources. The dominant source is Intra-Beam Scattering (IBS), with longitudinal and horizontal growth times of 2 h to 3 h. The beam-beam diffusion time is approximately 5 h in the vertical plane.
2. Provide cooling at top energies (275 GeV for protons, as well as 100 GeV and 41 GeV).
3. Ensure the cooling section hardware fits within the available space in the IR-2 tunnel.
4. Accommodate various ion species, including protons, ^{3}He, and heavy ions up to gold.
5. Operate with ion beam currents up to 1 A and bunch repetition rates from 24.6 MHz (at injection) to 98.5 MHz (at collision).
6. Maintain effective horizontal cooling for particles with transverse amplitudes up to 5σ.

III. Various high energy coolers overview

From August to December 2020, an alternative accelerator design review was conducted to select the preferred cooling method for the EIC. The alternatives considered to achieve high luminosity for the EIC were:

1. **Best Luminosity without Strong Hadron Cooling**: This alternative excludes cooling from the project scope. However, it allows for contingency funds to potentially add cooling later if project performance permits.
2. **Strong Hadron Cooling (SHC)**: This alternative assumes Coherent Electron Cooling (CeC) as the primary method, proceeding with cooling R&D and including SHC in the baseline design.
3. **Frequent On-Energy Injection**: This alternative assumes no strong cooling R&D or construction within the project scope, relying instead on frequent on-energy injection.
4. **Combination of SHC and On-Energy Injection**: This alternative includes a lattice compatible with frequent on-energy injection while simultaneously pursuing cooling R&D, keeping both options available until CD-3.

Technical performance, overall risks, and life-cycle costs were evaluated for each solution. Alternative One serves as a fallback for Alternative Two, should the SHC development timeline exceed the current construction schedule, while still permitting a future SHC upgrade. Based on the analysis by the independent review panel, the Office of Nuclear Physics recommended Alternative Two, strong hadron cooling, as the preferred approach for the EIC Project.

During the EIC hadron cooling design phase, additional concepts were proposed. A task force was assembled to evaluate and assess these multiple high-energy cooling proposals, which are discussed within the hadron cooling design framework.

There are two major categories of cooling: stochastic cooling and electron cooling. Stochastic cooling can be subdivided into microwave stochastic cooling (typically operating in the 1 GHz to 10 GHz range), coherent electron cooling, and optical stochastic cooling (operating at optical frequencies, 30 THz to 300 THz).

We summarize these concepts and findings below. Proposals 1–3 are based on stochastic cooling, utilizing electrons or photons for amplification. Proposals 4–8 are variations of electron cooling.

1. Coherent electron cooling by micro-bunching amplifier
2. Coherent electron cooling by plasma cascade amplifier
3. Optical stochastic cooling
4. Ring-based electron cooling using damping wigglers
5. Ring-based electron cooling driven by an induction linac
6. ERL-based cooler with circulator ring
7. A dual-energy storage ring-based electron cooling
8. Single-turn ERL-based cooling using multiple guns

A. Coherent Electron Cooling by Micro-Bunching Amplifier

Coherent electron cooling (CeC), originally proposed by Derbenev and Litvinenko [1, 2], is expected to achieve higher cooling rates than those available from conventional electron cooling techniques. It was selected as the baseline cooling method for the EIC due to its high efficiency in cooling high-energy protons. The CeC mechanism is similar to stochastic cooling but utilizes an electron beam instead of RF signals to increase the amplifier bandwidth. D.Ratner proposed a broadband amplification scheme[3], Microbunched Electron Cooling (MBEC), where amplification is achieved via a sequence of drifts and chicanes. Then G. Stupakov developed detailed theory and discussed the possible solution for cool the proton beam for the EIC [4, 5]. In this scheme, density perturbations in the drift sections execute a quarter-wavelength plasma oscillation. This cooling system is designed to deliver the required performance across the full EIC energy range, with cooling times on the order of one hour. This method is discussed in detail in Section V E.

B. Coherent Electron Cooling by Plasma-Cascade Amplifier

Similar to the above method, CeC based on a Plasma-Cascade Amplifier utilizes the plasma-cascade microbunching instability [6], which arises in a beam propagating along a straight path. This instability is driven by variations in the beam density and the corresponding modulation of the beam's plasma frequency. A key advantage of the PCA concept is that it does not require physical separation of the electron and hadron beams, which simplifies their longitudinal alignment. While PCA experiments are in progress at RHIC, a preliminary evaluation of a PCA for EIC hadron cooling has been performed. Compared to the parameters for the ongoing CeC experiment, the proposed EIC PCA requires a higher beam energy, higher peak current, and lower emittance, all of which are beyond the current state of the art. Furthermore, due to the short bunch length and the limited range of the wake potential, this method can only cool a very narrow range of ions.

C. Optical Stochastic Cooling

Optical stochastic cooling (OSC) substantially extends the frequency range of conventional microwave stochastic cooling from 1 GHz to 10 GHz to the optical regime of 30 THz to 300 THz. This method has been experimentally demonstrated at Fermi Lab Integrable Optics Test Accelerator[7]. In an OSC system, undulators serve as both the pickup and the kicker, with the signal being amplified by an optical amplifier. Following a successful demonstration of OSC at Fermilab, V.A. Lebedev proposed its use for high-energy cooling at the EIC [8]. An essential component of OSC is a chicane, which delays the ion beam to match the travel time of its own radiation through the optical amplifier. This synchronization ensures that each ion interacts with its own amplified radiation signal in the kicker.

Zholents, Rebuffi, and Shi proposed using an electron storage ring to cool high-energy hadron beams in the EIC using traditional electron cooling, replacing previously considered radiation cooling [9–11] with optical stochastic cooling (OSC). They showed in [12] that using extreme ultraviolet (EUV) light from wigglers can boost the cooling system's bandwidth to ∼10 PHz, enabling fast damping without light amplification. This is achieved by applying multiple uncorrelated corrections in a single cooling insertion and combining several insertions into an efficient cooling cascade.

However, employing a large bandwidth increases the probability for electrons to miss interaction in the kicker wiggler with light they radiate in the pickup wiggler due to an incorrect delay obtained on the path from the pickup to the kicker. A follow-up study [13] concluded that forcing each electron to interact with its own radiation signal multiple times in the cooling cascade is incompatible with a requirement for a compact design of the cooling insertion.

D. Ring based electron cooling using damping wigglers

The ring-based cooler approach utilizes conventional electron cooling, where electrons that continuously interact with the hadron beam are themselves cooled via radiation damping in a storage ring wiggler. This concept was proposed for the EIC in 2021 [11] and has since undergone substantial development. The transverse electron emittance is determined by a balance between radiation damping and various heating effects, including quantum excitation in dipoles and wigglers, beam-beam scattering with the ions, and intra-beam scattering (IBS) of the electrons in regions with non-zero dispersion. A preliminary lattice design has been developed that balances the competing requirements of small dispersion to minimize IBS and larger dispersion for chromaticity correction. The expected horizontal and longitudinal cooling times are on the order of 2 h to 3 h. This method was considered as the EIC alternative option and had significantly design progress. The details description can be found in the tech-notes [14].

E. Ring based electron cooling driven by an induction linac

A proposal from Fermilab suggests employing a pulsed, DC-like electron beam (830 ns pulse length) with a current of 50 A to 100 A at energies of 55 MeV to 147 MeV [15]. In this scheme, an induction linac accelerates each pulse for injection into a cooling ring, as depicted in Fig. 15. The beam circulates for approximately 6000 turns while cooling the EIC proton beam before being extracted to a beam dump and replaced with a fresh pulse every 5 ms. This pulsed replacement approach significantly reduces the average power required from the induction linac compared to a continuous source. The use of a coasting (DC) electron beam structure simplifies the overall cooling scheme design. At the maximum energy of 147 MeV, the average power deposited in the beam dump is approximately 2.5 MW. This proposed system is projected to deliver the required cooling performance across the entire EIC energy range, achieving cooling times of 1 h to 2 h.

F. Multiple guns beam merging ERL-based electron cooling

A single-turn Energy Recovery Linac (ERL) is expected to provide the optimal beam quality at the high average currents required for hadron cooling [16]. However, the necessary current is at least six times greater than that achieved by state-of-the-art high-voltage DC (HVDC) electron guns. To overcome the current limitations of a single gun and mitigate the energy-spread increase caused by Coherent Synchrotron Radiation (CSR) in multi-turn ERLs, a beam merging system using multiple guns is proposed. For example, to achieve a target current of 400 mA, one could use four guns producing 100 mA each. Alternatively, employing six guns would lower the requirement per gun to 67 mA, a current that has been experimentally demonstrated.

The concept of merging beams, which relies on a kicker system to interleave bunches onto a common axis, was explored during early EIC R&D. To produce a final beam of 4 nC bunches at 98.5 MHz, each of four guns would generate 4 nC bunches at a quarter of the repetition rate, 24.6 MHz. The 24.6 MHz merging kicker could be realized as either an LC circuit-based capacitor kicker [17] or a cavity-based, ferrite-dominated RF kicker [18]. After merging, the combined beam is accelerated to over 10 MeV for injection into the main ERL. As described in the EIC pre-CDR, this single-turn ERL configuration is expected to preserve beam quality more effectively than multi-turn designs.

A major challenge is the resulting beam power at the dump, estimated at 4 MW, which exceeds the capacity of typical high-power beam dumps. To manage this, the spent beam can be decelerated in an RF structure before reaching the dump. For improved wall-plug efficiency, an RF output coupler can extract power from the decelerating beam. This recovered RF power can then be used to compensate for power losses in the main E-linac cavities, enhancing the overall energy recovery process.

IV. Roadmap to achieve the cooling requirements

Following a 2020 review, Coherent Electron Cooling (CeC) was selected as the primary method for the EIC's high-energy cooler, referred strong hadron cooling (SHC) in the EIC. An EIC hadron cooling roadmap, initiated in 2021, has been updated annually to reflect design, scope, and schedule progress.

During the 2023 design phase, it became clear that a conventional electron cooler was a desirable addition at the injection energy. This cooler would be integrated into the same ERL to achieve the required beam emittances at lower energies (e.g., the 24 GeV injection energy) and provide a potential path for a 41 GeV storage cooler. Consequently, the EIC hadron cooler scope was updated to include both an SHC and a low energy conventional electron cooler (LEC).

The novel physics and technology underlying the SHC present a significant risk of delayed delivery to the EIC project. To mitigate this schedule risk, a pre-conceptual design study was initiated for an alternative cooling method based on incoherent electron cooling in a storage ring [14]. The cooling capability of such a system is, however, limited to high energies, specifically 275 GeV for protons and 100 GeV for gold ions.

In parallel, the CeC-X experiment at RHIC, supported by RHIC operational funds and not the EIC project, aims to demonstrate coherent cooling on a low-energy Au ion beam. Although CeC-X uses a different amplification scheme (plasma-cascade amplification) than the one planned for the EIC, its results could provide valuable design information. However, the outcome of CeC-X is not considered a deciding factor for the viability of the SHC in meeting EIC requirements.

As an additional risk mitigation strategy against the delayed availability of the SHC, the existing RHIC stochastic cooler can be upgraded for EIC use by increasing its amplifier power for heavy ion cooling. Further studies will be assigned to define the capabilities, optimal hardware locations, and cost of this stochastic cooling upgrade.

By 2025, significant technical risks associated with the high-energy cooling system remained unresolved, and luminosity studies indicated that even with the SHC scenario, the projected average luminosity would be less than half of the ultimate requirement. The detailed analysis is described in section V D and section V J. Consequently, due to these unresolved risks and performance limitations, the design efforts for advanced high-energy cooling—including both SHC and the storage ring cooler designs, were terminated within the EIC project scope. The mitigation of remaining risks is now being pursued by resources outside of the main project.

V. Strong Hadron Cooling design

The design study for microbunched electron cooling includes the following activities:

- **Cooling Physics Modeling:** Develop and apply physics models and simulation codes to represent the cooling process accurately.
- **Cooling Section Design:** Engineer the lattice, magnets, and diagnostic systems for the complete cooling section.
- **Luminosity Performance Modeling:** Integrate the cooler model into full collider simulations to project and optimize the machine's luminosity.
- **ERL Design:** Develop the complete design for the ERL, including RF systems, magnets, and beam dynamics.
- **Critical Component R&D:** Conduct targeted research and development for key enabling technologies, such as the high-current electron gun and the amplification system.

The schmatic drawing of the SHC facility as well as hadron beamline is shown in the Fig. 1. It consists of Energy Recovery Linac, electron-hadron overlap section, Injector Cooler(precooler) bypass beamline.

A. 1D, 2D and hybrid model

The simplest way to model the process of microbunched electron cooling is to treat the electrons and hadrons as rigid Gaussian discs of charge, with transverse sizes equal to the relevant beam

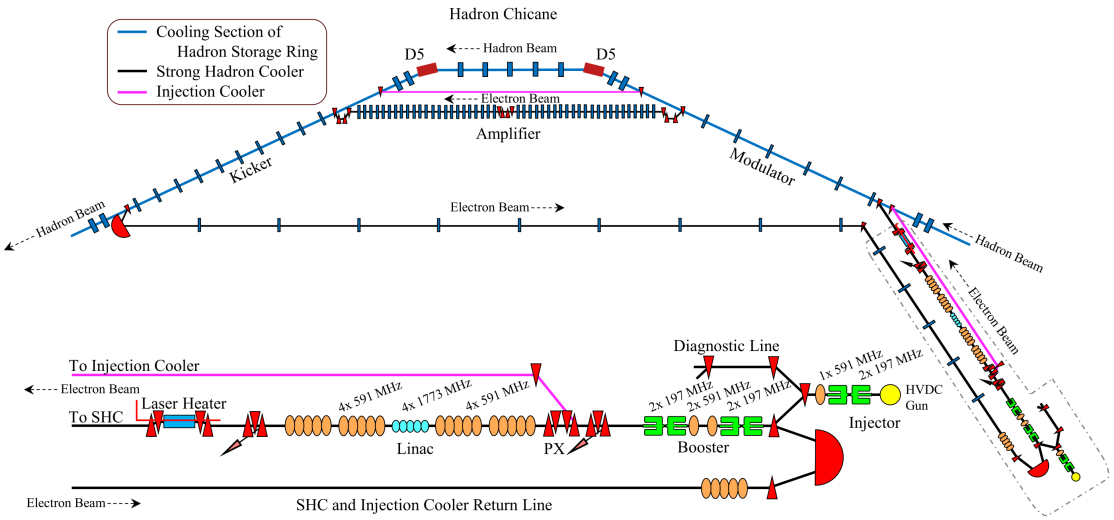


FIG. 1. Representative diagram of the ERL with the cooling section at the 2 o'clock straight. The driver ERL which is housed in its own building is shown on the right of the figure inside the dashed line. The top part of the figure is the cooling section consisting of modulator, amplifier, and kicker. Below the cooling section is a return line that takes the beam back to the booster and main linacs. The bottom part of the figure shows an expanded view of the ERL including the gun, injector linac, booster (BSR) linac, main (LA) linac, PX, P2 and P3 chicanes, (P2/P3 not shown for clarity), laser heater and beam dump chicanes. Shown in the return line is the 591 MHz "chirper" cavity identical to the 591 MHz 5-cell cavities in the LA linac used to control energy spread of the second-pass beam. The pink beamline is a bypass beginning at the PX chicane for the 13 MeV beam used for cooling the injected beam (precooling) in the HSR. The precooling beam bypasses the amplifier section in the cooler but otherwise follows the same beamlines as the modes A and B beams.

sizes. We can then write down expressions for the longitudinal force between a hadron and electron disc, or between two electron discs, as a function of the longitudinal distance between them. From here, an expression for the one-dimensional wake function can be derived, which gives the energy kick a hadron receives in the kicker as a function of its delay in moving from modulator to kicker. Detailed expressions and derivations are provided in [4, 5, 19].

An improvement to this paradigm is to use a hybrid model which continues to consider the electrons as discs but treats the hadrons as point particles with arbitrary transverse offsets in the modulator and kicker. This requires only a change to the electron-hadron interaction terms, as detailed in [20].

In addition to the coherent kick which a hadron receives due to its own wake, it also receives energy kicks arising from the wakes of neighboring hadrons and electrons. Since these additional kicks are not correlated with the phase-space coordinates of the kicked particle, they act in an incoherent manner to increase the beam emittance in a process termed "diffusion." Diffusion due to noise in the hadron beam is discussed in [4, 5, 19, 21] and an extension of this model to the case of electron noise is provided in [22].

In order to maximize the cooling rate, it is desirable to significantly amplify the imprint of the hadrons on the electron beam, resulting in fractional fluctuations in the electron density that are a significant fraction of one. In this case, the linear theory upon which the previous results are based breaks down. Additional complications arise from the fact that electrons and hadrons with different energy offsets will drift relative to one another in the modulator and kicker and the fact that electrons receive delays in the amplification sections that depend on their transverse actions. In order to handle these effects, we have developed a particle-in-cell (PIC) code which tracks the hadron and electron macroparticles within a narrow slice of the beam through a single pass of the cooling section. Tracking once with initial random noise in both beams gives us a baseline kick-vs-z curve. Tracking again with the same initial noise as before, but with an additional hadron macroparticle at the origin in the modulator, and subtracting off the baseline gives us an effective wake function. Comparing this to the wake functions extracted from theory allows us to derive effective scaling factors by which to reduce the theoretical wakes in order to take account of the nonlinearities. Additional details are provided in [23] and Section III of [22].

B. 3D model

A set of fully three-dimensional theoretical and simulation techniques has also been developed for the study of microbunched electron cooling (MBEC)[24]. Our model incorporates previously neglected features and effects, such as the point charge nature of the electrons, as well as their angular spread and transverse motion (betatron oscillations). A three-dimensional macroparticle simulation algorithm, along with a Vlasov equation-based, frequency-domain theoretical formalism, have been used to track the modulation of the electron beam along the entire cooling lattice, allowing us to perform a comprehensive study of the generalized wakefield of the system (a key figure of merit for the cooler). In general, good agreement has been observed between theory and simulation-based approaches (see Figure 2). Moreover, the results from our 3D toolkit have been compared to their counterparts from simplified but computationally faster approximate models (disk-based or hybrid) used for optimization purposes. For the parameter set considered here (based on the EIC baseline for 100 GeV protons and assuming a cooler electron beam with a 10 A peak current), the 3D calculation essentially confirms the validity of the performance figures from the approximate models. In conclusion, although computationally intensive and relatively slow, the 3D model provides a necessary complement and benchmark for faster but less rigorous techniques, such as the hybrid model, as the sufficient accuracy of the latter cannot be taken for granted.

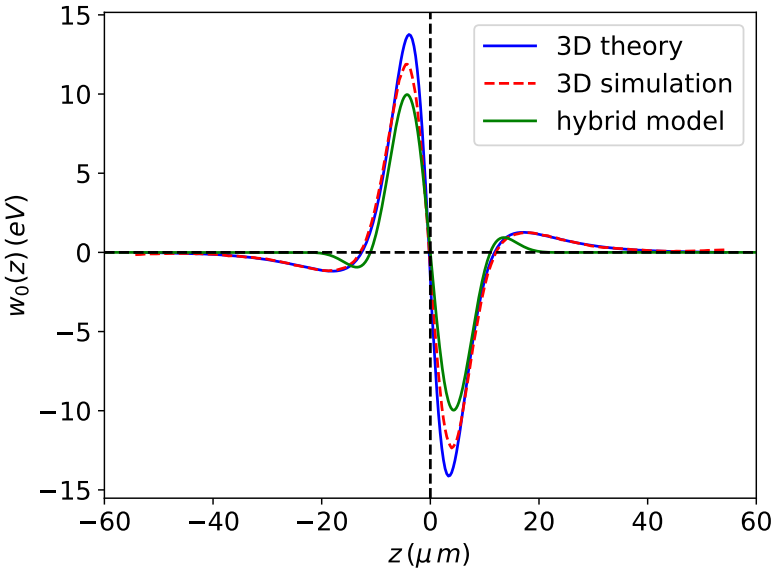


FIG. 2. On-axis wake profiles for the peak electron current. Here, w_0 (which represents the total energy change of the kicker proton) is plotted versus the longitudinal offset z. Data shown are from 3D theory (blue solid lines), 3D simulation (red dashed lines), and the hybrid model (green solid lines). The hybrid model data are courtesy of W. Bergan.

C. Beam requirements and main parameter table

Optimal parameters for cooling 275 and 100 GeV protons were obtained using a multi-objective genetic optimizer [25] followed by manual fine-tuning, and are displayed in Tab. I. Since the cooling force depends on the local electron current, it is desirable to have a flat-top electron current distribution, which we approximate as a 4th-order supergaussian. The "equivalent Gaussian bunch length" quoted in the table corresponds to the length which a Gaussian bunch of the same charge would need to have in order for its central current to match that of the supergaussian bunch. The current distribution of the supergaussian is given by $I(z) = I_0 e^{-(z^2/2\sigma^2)^N}$ where $I_0 = \frac{Qc}{\sqrt{2\pi}\sigma_z}$ is the peak current, N is the order of the supergaussian, $\sigma = \frac{N\sqrt{\pi}}{\Gamma(1/2N)}\sigma_z$, σ_z is the "equivalent Gaussian bunch length" of the electron beam, and Q is the total bunch charge.

TABLE I. Parameters for Longitudinal and Transverse Cooling at Store

Case	100 GeV	275 GeV
Geometry		
Modulator Length (m)	33	33
Kicker Length (m)	33	33
Number of Amplifier Drifts	2	2
Amplifier Drift Lengths (m)	49	49
Proton Parameters		
Protons per Bunch	6.9e10	6.9e10
Fit Proton Bunch Length (cm)	7	6
Proton Fractional Energy Spread	8.5e-4	6.0e-4
Proton Emittance (x/y) (nm)	30 / 2.7	11.3 / 1
Horizontal/Vertical Proton Betas in Modulator (m)	16.6 / 16.4	21.0 / 19.6
Horizontal/Vertical Proton Alphas in Modulator	0 / 0	0 / 0
Horizontal/Vertical Proton Dispersion in Modulator (m)	0.0036 / 0.096	0.0019 / 0.067
Horizontal/Vertical Proton Dispersion Derivative in Modulator	0.030 / -0.0099	0.030 / -0.0049
Horizontal/Vertical Proton Betas in Kicker (m)	16.6 / 16.4	21.0 / 19.6
Horizontal/Vertical Proton Alphas in Kicker	0 / 0	0 / 0
Horizontal/Vertical Proton Dispersion in Kicker (m)	0.0036 / 0.096	0.0019 / 0.067
Horizontal/Vertical Proton Dispersion Derivative in Kicker	-0.030 / 0.0099	-0.030 / 0.0049
Proton Horizontal/Vertical Phase Advance (rad)	3.227 / 4.72	3.162 / 4.44
R56 in Proton Chicane (mm)	4.2	0.95
Electron Parameters		
Electron Bunch Charge (nC)	1	1
Equivalent Gaussian Electron Bunch Length (mm)	12	9.4
RMS Electron Bunch Length (mm)	9	7
Electron Peak Current (A)	~ 10	~ 13
Electron Supergaussian Order	4	4
Electron Fractional Slice Energy Spread	1e-4	5.9e-5
Electron Normalized Emittance (x/y) (mm-mrad)	2.8 / 2.8	2.8 / 2.8
Horizontal/Vertical Electron Betas in Modulator (m)	20 / 20	21.4 / 21.4
Horizontal/Vertical Electron Betas in Kicker (m)	29.7 / 4.1	7.9 / 7.9
Horizontal/ Vertical Electron Betas in Amplifiers (m)	12.0 / 12.0	4.9 / 4.9
R56 in First Electron Chicane (mm)	23.3	12.0
R56 in Second Electron Chicane (mm)	-16.7	-6.7
R56 in Third Electron Chicane (mm)	-18.2	-6.8
Cooling Times		
Horizontal/Vertical/Longitudinal Initial IBS Times (hours)	2.0 / 4.0 / 2.5	2.0 / - / 2.9
Horizontal/Vertical/Longitudinal Beam-Beam Times (hours)	20.0 / 5.0 / -	20.0 / 5.0 / -
Horizontal/Vertical/Longitudinal Initial Cooling Times (hours)	1.9 / 4.7 / 3.3	1.0 / 13.5 / 1.5

D. Luminosity model

In order to understand the long-term evolution of the proton bunch, including microbunched electron cooling as well as growth due to IBS and the beam-beam effect, we have developed a long-term luminosity model, as discussed in [26]. This initializes ten thousand hadron macroparticles and uses a handful of transfer matrices to track them through a simplified hadron storage ring (HSR) consisting of an RF cavity, modulator, kicker, and interaction point (IP). In order to perform the tracking in a reasonable time, we treat each simulated passage through the ring as $N = 10^6$ real turns. Therefore, we scale up the coherent cooling kicks by a factor of N, and incoherent kicks from diffusion, IBS, and beam-beam by a factor of $\sqrt{N}$.

The RF cavity uses both 591 and 197 MHz sine waves to provide an energy kick to each hadron each simulated turn based on its longitudinal offset from the bunch center, enabling synchrotron motion. At this stage, we calculate the IBS rates using Sergei Nagaitsev's rewriting [27] of the Bjorken-Mtingwa formulas [28]. We extract the necessary optics functions around the ring from a realistic lattice, obtain the horizontal and vertical emittances of the bunch from exponential fits of individual hadron actions, and use Gaussian fits of hadron longitudinal positions and energy offsets to get the bunch length and energy spread. The use of fits rather than direct RMS values here de-emphasizes long tails which may develop under the influence of cooling and instead focuses on the parameters at the core of the beam. With the IBS rates known, we apply a Gaussian random kick to each hadron's transverse angles and energy offset with the size chosen to give the correct heating rate. Such a kick is applied once per simulated turn at the RF location, where we assume

zero dispersion. The kick for a particle at coordinates (x, y, z) is scaled by $\sqrt{\rho(x,y,z)/\langle\rho(x,y,z)\rangle}$, where $\rho(x, y, z)$ is the local 3D hadron density and the average is taken over all macroparticles. This reflects the fact that hadrons near the beam core will have more frequent collisions, generating faster IBS growth. Explicitly, the kicks are given by:

$$\begin{aligned} x' &\to x' + g_x\sqrt{\frac{2\lambda_x\epsilon_x\rho(x,y,z)\Delta t}{\beta_x\langle\rho(x,y,z)\rangle}} \\ y' &\to y' + g_y\sqrt{\frac{2\lambda_y\epsilon_y\rho(x,y,z)\Delta t}{\beta_y\langle\rho(x,y,z)\rangle}} \\ \delta &\to \delta + g_\delta\sigma_\delta\sqrt{\frac{2\lambda_z\rho(x,y,z)\Delta t}{\langle\rho(x,y,z)\rangle}} \end{aligned} \tag{1}$$

where $g_{x,y,z}$ are independent Gaussian random numbers with mean 0 and RMS 1, $\lambda_{x,y,z}$ are the IBS rates in the three planes, $\epsilon_{x,y}$ are the transverse RMS emittances, $\beta_{x,y}$ are the transverse beta functions at the kick location, σ_δ is the RMS fractional energy spread, and Δt is the length of the simulated timestep. It can be verified that such kicks will yield the desired emittance growth rates.

The beam-beam effect is assumed to increase the beam emittance with constant growth times of 20 hours horizontally and 5 hours vertically. As in the case of IBS, Gaussian random kicks are applied to each hadron's transverse angles with mean 0 and standard deviation chosen to provide the correct growth rate. The formulas for applying these kicks are nearly identical to the horizontal and vertical IBS case, except that we do not scale the beam-beam kick with local density.

Particles are considered lost if their actions violate the condition $J_x/J_{x_{max}} + J_y/J_{y_{max}} + J_z/J_{z_{max}} < 1$, where the maximum transverse actions are those which would result in the particle's transverse coordinate exceeding 6 times the initial beam size and the maximum longitudinal action is the size of the 591 MHz bucket. Additionally, each simulated turn, each hadron has a probability of being lost due to the Touschek effect equal to $1-\exp[-\Delta t/\tau]$, where Δt is the length of the timestep, τ is the loss time based on the results of [29]:

$$\frac{1}{\tau} = \oint \frac{ds}{C} \frac{Nr_h^2 c\left[1 + \frac{\sqrt{2}}{\pi}\ln\left(\frac{\sigma_{x'}^2+\sigma_{y'}^2}{2\sigma_{x'}\sigma_{y'}}\right) - 0.055\left(\frac{\sigma_{x'}^2-\sigma_{y'}^2}{\sigma_{x'}^2+\sigma_{y'}^2}\right)^2\right]}{4\sqrt{2}\gamma^3\sigma_z\sigma_x\sigma_y\delta_{max}^2\sqrt{\sigma_{x'}^2+\sigma_{y'}^2}} \tag{2}$$

δ_{max} is the fractional energy kick which would cause a zero-action particle to be immediately lost, $r_h \equiv (Ze)^2/(4\pi\epsilon_0 m_h c^2)$ is the classical hadron radius, N is the bunch population, C is the HSR circumference, c is the speed of light, and the various beam sizes and divergences (σ_x, $\sigma_{y'}$, etc.) are computed from the local dispersion, dispersion derivative, and Courant-Snyder parameters.

Cooling is implemented using the hybrid model introduced in [20], and discussed briefly in subsubsection V A of this note. Each hadron receives a coherent kick based on its delay in moving between the modulator and kicker, on its transverse coordinates in each, and on the density of the cooling electrons it overlaps with. The wake function is scaled down based on the nonlinear tracking simulations, as discussed in [23]. Additionally, a Gaussian random energy kick is applied to each hadron to simulate the effect of diffusion, with the size of the kick depending on the hadron's transverse position in the kicker and its longitudinal position within the bunch.

At the IP, we assume that the crab cavities work perfectly, so that the hadron and electron beams collide head-on with negligible transverse angles and longitudinal velocities of c. In this case, we can obtain the instantaneous luminosity from a special case of Eqtn. 4.1 of [30]:

$$\mathcal{L} = 2cf_{rep}\int \rho_+(\vec{x},t)\rho_-(\vec{x},t)d\vec{x}^3dt \tag{3}$$

where c is the speed of light, f_{rep} is the frequency of bunch crossings, and ρ_+ and ρ_- are the densities of the two colliding beams. Numerically, this integral can be evaluated for arbitrary hadron bunch distributions by assuming a perfectly Gaussian electron bunch, tracking each hadron macroparticle as it passes near the interaction point, integrating the electron density it sees (taking into account variations of the electron β^* as it moves away from the IP and the longitudinal motion of the electron bunch), summing the result of this integration over all macroparticles, and multiplying by

the number of real hadrons represented by each macroparticle. The beta functions at the IP are adjusted each turn to prevent the hadron beam size from sinking below its initial value, to prevent the hadron beam divergence from rising above its initial value, and to maintain the electron beam size equal to that of the hadrons.

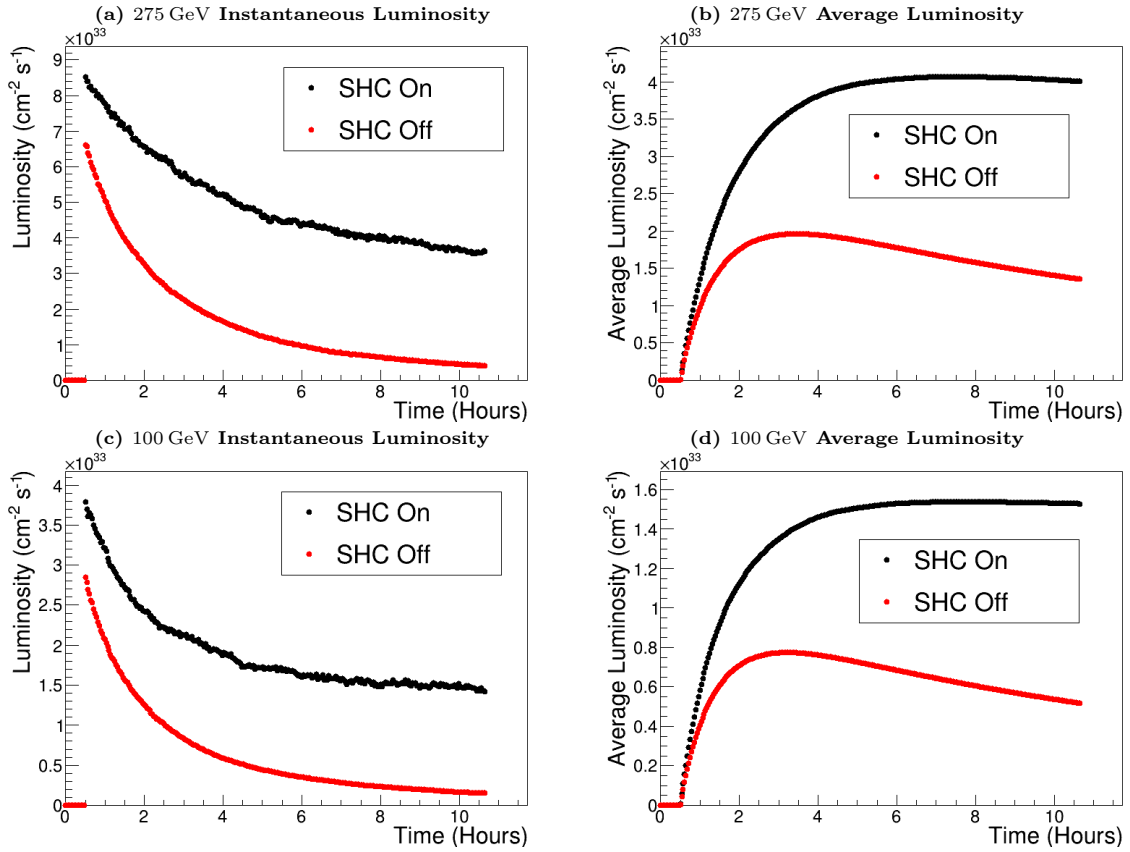


FIG. 3. Evolution of the luminosity over time. (a) and (b) show results for 275 GeV protons, while (c) and (d) show results for 100 GeV protons. (a) and (c) show instantaneous luminosity, while (b) and (d) show average luminosity for a given store length, taking into account the 2 hours needed to ramp down the HSR, refill, do pre-cooling, and ramp to collision energy. Additionally, the luminosity is set to zero during the first half-hour of the store, since this is when we will be filling the electrons and turning on the detector. We see that at both energies the use of SHC based on MBEC reduces the rate of luminosity loss, roughly doubling the achievable average luminosity.

Plots of the instantaneous and average luminosity as a function of time are shown in Fig. 3. We see that microbunched electron cooling reduces the rate with which luminosity decays over time, but does not stop it altogether, leading to a factor of roughly 2 improvement in average luminosity. This cooling method works well at cooling the core of the bunch, but does not do a good job of cooling the tails, since those particles overlap less frequently with the cooling electron bunch and have large actions, leading to large delays relative to the wake wavelength and sampling of the nonlinear region of the wake. This contributes to continued particle loss and emittance growth.

Further development of this code is ongoing, both to better understand the luminosity evolution without SHC and to incorporate stochastic cooling for gold ions.

E. Cooling section design

The cooling section for microbunched electron cooling (MBEC) consists of three parts, with a diagram shown in Fig. 1.

The first section is a straight "modulator," within which the circulating hadron beam co-propagates with an electron beam of the same relativistic gamma. This contains quadrupoles to maintain the electron beam size, but appears as a drift to the high-energy hadrons. It is at this stage that the hadrons provide energy kicks to the electrons.

After the modulator, the hadron and electron beams are separated. The electrons travel through an amplification section, which consists of three chicanes and two straight sections. The chicanes translate energy offsets of individual electrons into longitudinal position shifts, while within the straight sections the electrons provide energy kicks to one another, introducing additional correlated energy shifts. The repetition of these steps induces a microbunching instability to amplify the initial energy perturbations in the electron beam and turn them into density fluctuations.

During this time, the hadrons pass through their own chicane, so that their travel time is dependent on their energy offsets and transverse phase-space coordinates.

The final section is the "kicker." This looks physically similar to the modulator, and once again the two species co-propagate. At this point, the density-modulated electron beam provides energy kicks to the hadrons. By making the proper choice of the beam parameters and optics functions, one can make it so that these energy kicks serve to reduce the actions of the hadrons, reducing the emittance. The cooling section includes a 33 m modulator, a 100 m amplification section, and a 33 m kicker section.The FODO cells are designed to get smaller gamma to achieve a smaller relativistic gamma to reduce the delay caused by betatron oscillations.. The drift sections of the Modulator/amplifier/kicker all use FODO cell.

The amplification section consists of three chicanes to turn the energy modulation into a density modulation. The electron bunch longitudinal space charge will increase the beam energy spread and lengthen the bunch length when through the cooling section if their R_{56} is not zero. However, using the regular four dipoles chicane in the amplifier with the drift space will give negative R_{56} and cause the microbunch slippage. This slippage of the modulated micro-bunches will misalign with the same hadrons at the kicker section. One of the methods to solve this problem is reversing sign of R_{56} in one of chicanes in the amplification section and achieve a the total $R_{56} = 0$ [31]. To avoid anti-cooling, the chicane's $R_{56_1} \cdot R_{56_2} \cdot R_{56_3} > 0$, because the hadron chicane provides a positive R_{56}. Here the positive R_{56} means the high momentum particles move backward in the beam frame. We have designed chicanes with embedded quadrupoles. To generate positive R_{56}, the dispersion crosses the zero between the 1st and the 2nd dipoles, resulting in a shorter path length for the lower energy electrons. Figure 4 shows the dispersion and lattice layout of the chicane. It can tune the R_{56} in large range without change the path length. We also evaluated the CSR enhanced microbunch through the chicane follow the method discussed in Ref [32]. The enhancement factor for microbunching due to CSR is less than one, as shown in Fig. 4. To achieve the sufficient stability of the longitudinal alignment, the chicane dipole field stabilization must be better than $5e-6$, which would cause a shift about 200 nm, CSR caused longitudinal shift 140 nm and longitudinal space charge caused shift 56 nm. These noise requirements are very challenging.

The cooling section design are incorporated in the ERL beamline lattice see Sec. V H 3.

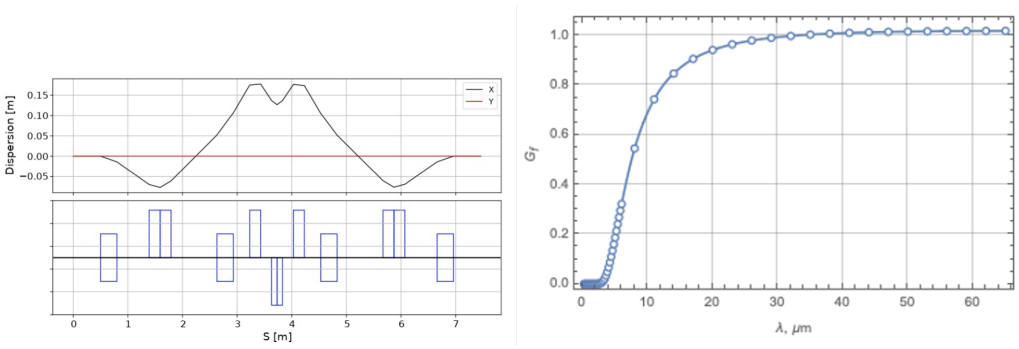


FIG. 4. R_{56} tune-able chicanes at amplification section and amplification due to CSR

It is also necessary for the energy spread in the electron beam to be properly controlled. From [4, 5], we see that an electron chicane of strength R_{56} multiplies the MBEC impedance by a factor of $\exp[-(kR_{56}\sigma_e)^2/2]$, where k is the wavenumber and σ_e is the electron energy spread. If σ_e is much larger than the design, it will wash out the wake features necessary for cooling, while a too-small σ_e will introduce high-frequency noise into the electron beam, saturating the amplifier.

Another potential issue is the higher-order path length delays to the electrons. As discussed in [23], an electron with transverse actions J_x and J_y traveling between points s_1 and s_2 receives phase-averaged delays approximately equal to

$$\Delta z \approx -\int_{s_1}^{s_2} \frac{J_x\gamma_x(s) + J_y\gamma_y(s)}{2} ds. \tag{4}$$

We have found that, despite their short lengths, the large Courant-Snyder gammas in the electron chicanes result in those elements providing extra delays to the electrons on the order of the wake wavelength, significantly smearing out the wake function. This will need to be addressed in any future design.

All these effects of cooling wake smearing have been included in the cooling model.

F. Wiggler-based electron beam microbunching amplifier

The wiggler-based amplifier of the microbunching instability has been studied theoretically and through simulations using the OPAL-FEL code [33]. This study concluded that a large gain can be obtained using a relatively compact system. Benchmarking effects of electron collective interactions in a wiggler with OPAL-FEL is described in [34]. The Argonne Wakefield Accelerator facility (AWA) [35, 36] was used for this experiment. Subsequent theoretical and simulation studies using large wiggler parameter $K = 7.5$ [37] do not show improvement compared to [33], where $K = 1.5$ was used.

G. Hadron beamline design

The Insertion Region 2 (IR2) of the Hadron Storage Ring was selected to host the cooling systems, which include a Pre-Cooler operating at injection energy (24 GeV) and a SHC system operating at top energies (100 GeV and 275 GeV). A major design challenge arises from the geometric constraints imposed by both systems: the SHC requires a chicane between the modulator and the kicker, while the Pre-Cooler demands a sufficiently long drift space of the modulator and kicker.

A proof-of-principle layout and optics had been integrated into the lattice HSR-220512a [38], delivering good SHC control parameter ranges at 100 GeV and 275 GeV as shown in Fig. 5. Peak magnet strengths were consistent with re-purposed RHIC arc quadrupoles and RHIC insertion region dipoles.

An updated geometry was explored to provide longer drift space for pre-cooling [39]. The original symmetrical layout of IR2 was transformed into an asymmetrical one to align with both the colliding and non-colliding IR8. The beam optics was rematched to meet the SHC requirements. The hadron lattice have met both 275 GeV and 100 GeV Horizontal and longitudinal cooling. Minor adjustments to the optics can rebalance H/L cooling rate. A vertical cooling lattice can be achieved by using skew quadrupole between dipole in the middle section.

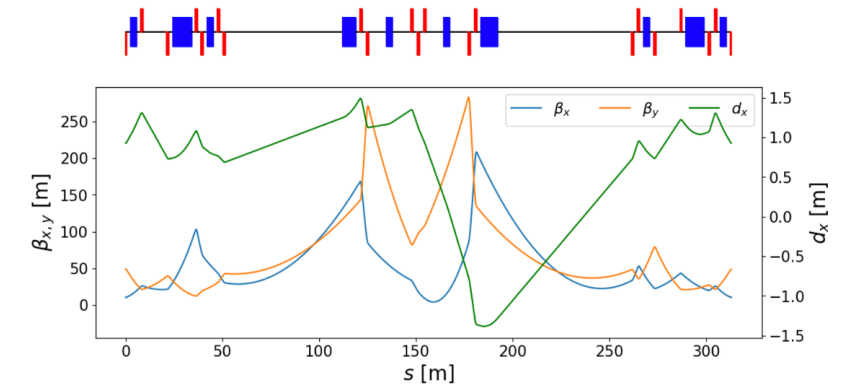


FIG. 5. Hadron lattice including the matching section, electron-hadron copropagate drift section and bypass section

H. Energy Recovery Linac design

A design effort was undertaken at JLab,BNL, and Xelera to specify parameters of an Energy Recovery Linac (ERL) located at the 2 o'clock straight of the EIC that can achieve the SHC requirements for the EIC. Multiple papers discussed the SHC ERL design in various workshops[40–43]. The primary requirements are high average current up to 100 mA used for pre-cooling the

injected HSR beam at 24.5 GeV, and cooling at the hadron beam collision energy of 275 and 100 GeV. The injection and collision beam energies imply required electron beam energies from the ERL of 13, 150 and 55 MeV. The electron beam energies at collision are called modes A (150 MeV) and B (55 MeV) respectively. Figure 1 shows a representative diagram of the major ERL and cooling systems. In the following sections we describe the various design aspects and challenges required to complete a full design of an ERL to meet the requirements. The main ERL design challenges are the relatively low beam energy and high average current which make the machine susceptible to various beam instabilities and losses. Operability also is a challenge in that the machine will have to switch between injection cooling and cooling at collision energy. In the following sections we describe the completed ERL design and plans and important aspects of future R&D required to realize an ERL based SHC system.

1. ERL injector

We use a 400 kV to 550 kV DC gun with $K(Na)_2CsSb$ photocathode as the electron source to produce a 1 nC bunch charge with the repetition rate of 98.5 MHz and less than 2 mm mrad of normalized emittance. 100 ps long beer-can distribution beam is generated from the HVDC gun. A single cryomodule consisting of two of 197 MHz quarter-wave resonator and a single cell 591 MHz cavity as shown in Fig. 6 is used for accelerating beam to 5.6 MeV. We assume each 197 MHz cavity's gap voltage is 2.9 MV. The beam is nearly on crest of 197 MHz cavities differing from crest by $< 0.2°$ to compensate for the energy chirp caused by longitudinal phase space effects. A 3rd harmonic cavity(591 MHz) is placed after the 197 MHz cavity to linearize the bunch longitudinal phase space. At the end of the injector, the RMS normalized emittance is 1.57 mm mrad with an RMS bunch length of 17 mm. At the exit of the ERL merger, the RMS normalized emittance is 1.88 mm mrad.

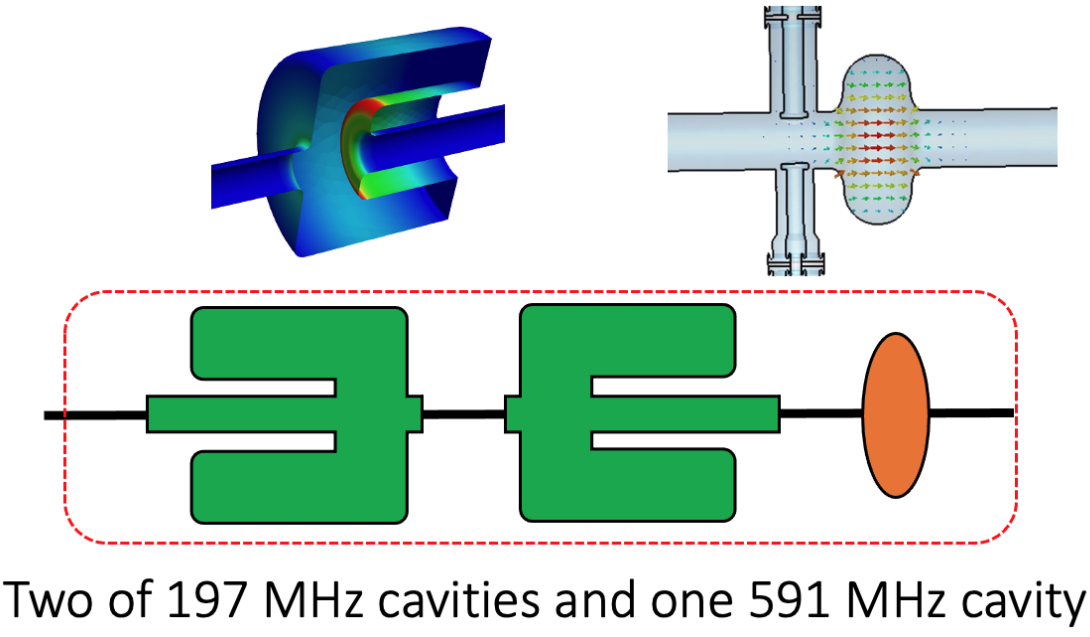


FIG. 6. Two of 197 MHz cavities and one cell 591 MHz cavity in a single cryomodule.

Because SHC will amplify the density-modulated micro-bunch to produce a wakefield strong enough to kick the hadron particles in momentum space. The amplification process depends on the current. For uniform amplification of the microbunch, a constant current along the bunch is required ("beer-can" distribution [23]). Reference [44] described a method to get a smooth current and energy spread profile along the bunch. The method uses micro-Gaussian pulse stacking to generate the desired initial current profile of the electron bunch and thereby generate an initial longitudinal distribution. For practical and stable operation, we use eight micro-pulses stacked together and independently control the amplitude of each pulse. Adjacent pulses are spaced 2 σ apart, and each micropulse with the same polarization direction is spaced at least 4 σ apart to

prevent interference.

We attempt to make the electron distribution a Super-Gaussian distribution at the end of Linac with the functional form in Eq. (5). The optimizer adjusts the micropulse amplitude to maximize the value of p at the end of the linac.

$$f(z) = Ae^{-\left(\frac{z^2}{2\sigma^2}\right)^p} \tag{5}$$

$$A = \frac{p}{\sqrt{2}\sigma\Gamma(\frac{1}{2p})} \tag{6}$$

In the injector, the bunch is always at the RF crest to achieve the maximum accelerating voltage, and the RF voltage is also set to the maximum. The booster cavities between the merger and chicane were configured to generate a beam chirp. All focusing elements, such as solenoid strength, are optimized to minimize transverse emittance after optimizing the longitudinal distribution. After optimization, the initial optimized laser longitudinal distribution was obtained as shown in Figure 7.

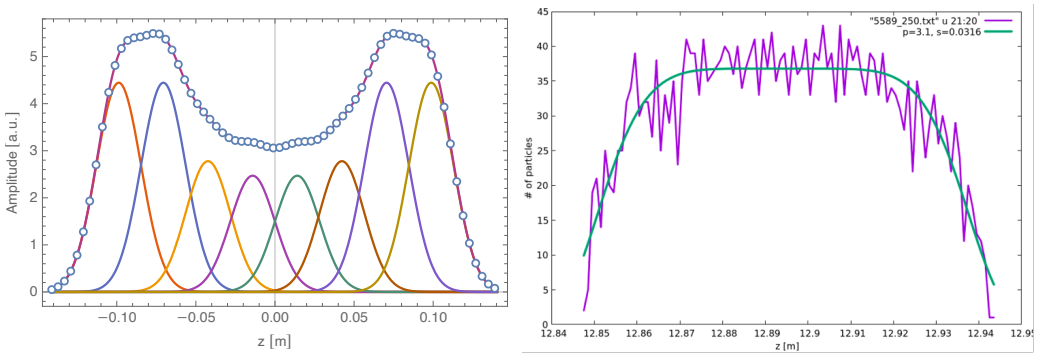


FIG. 7. The laser initial distribution (left) and the beam distribution (right) at the end of injector.

The ERL merger consists of an achromatic arrangement of two dipoles and two solenoids and brings the beam into the Linac section. We use Chevron dipoles that have focusing in both directions. The two solenoids are tuned to keep dispersion zero after the merger. To merge high-energy electrons with energies of 149.77 MeV, 54.46 MeV, or 22.33 MeV, we place a three-dipole chicane before the last merger dipole at the return beamline. To mitigate the nonlinearity of chicane-induced distribution changes, we utilize an upstream third harmonic cavity to produce correlated energy spread. We find that the ratio $V_3/V_1 \approx 1/6$. With this set up, at the end of linac, we get a SuperGaussian distribution with p of 2.61.

An alternative study using space charge tracking in code Bmad and genetic optimization was carried out[45]. Various initial distributions were generated at the cathode and were propagated through the full SHC ERL lattice to the end of the main linac while 3-D space-charge forces remained active up to 13 MeV. In the baseline beer-can initial distribution, we confirmed that the non-linear space charge forces destroyed the desired flat current and uniform slice properties well before the cooling section. To achieve desired beam quality, the study parameterised the cathode laser's longitudinal profile with eight spline-control points, plus the bunch RMS length and an overall phase, giving ten free variables. After tracking each candidate distribution, three slice-based figures of merit—current uniformity, slice energy-spread deviation, and slice transverse-emittance deviation—were computed inside one RMS bunch length. A parallelised CNSGA-II algorithm explored this multidimensional space with a population of 64 over 500 generations, building a Pareto front that exposes the trade-offs among the three objectives.

The optimal solution corresponds to an asymmetric two-peak laser pulse. When tracked, it delivers 10 A peak current and the target slice energy spread while meeting the 1 nC charge requirement. However, space-charge-induced timing shifts still imposed an excessive peak-to-peak energy variation. Eight 591 MHz fundamental and four 1773 MHz third-harmonic cavities—voltages and phases in the main linac were retuned, which reduced the peak-to-peak energy spread dramatically without spoiling current flatness or slice energy spread.

The optimised setting therefore meets two of the three SHC electron-beam specifications. The transverse emittance grows significantly in the injector and bunch-compressor regions and can be mitigated by re-matching solenoids and quadrupoles, which is achieved in the lattice design described in the next section. The initial and final distribution of the optimal solution is shown in Figure 8.

Both methods confirmed an uniform current distribution can be achieved. The hardware specification preliminary studied and shows achievable as well.

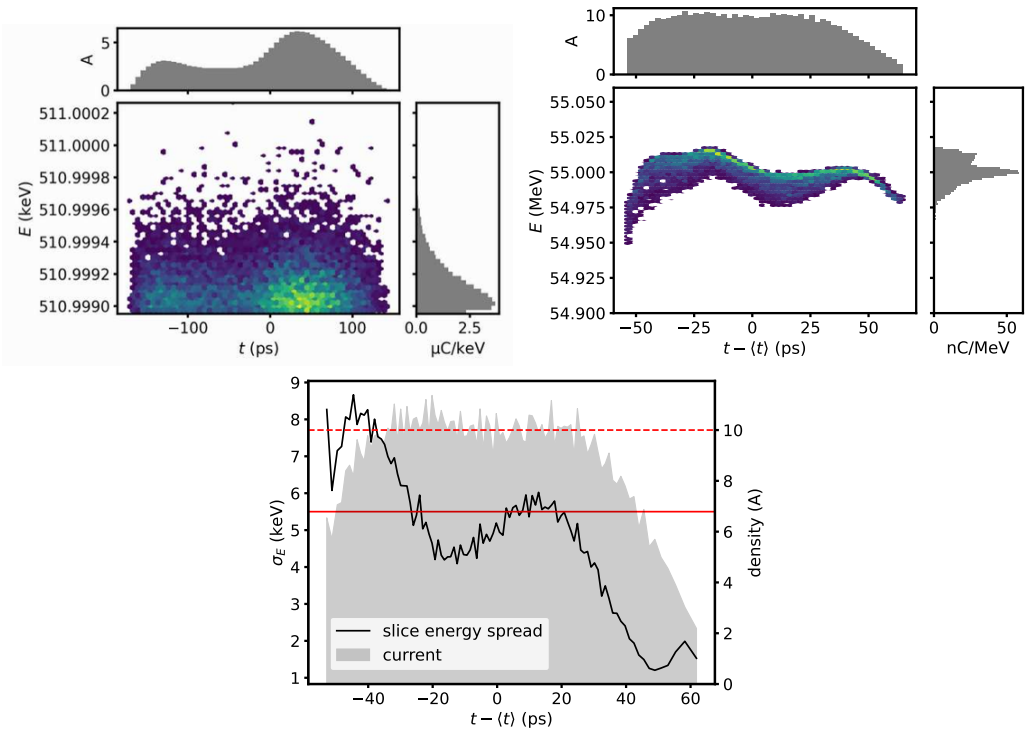


FIG. 8. (top left) Initial longitudinal distribution that resulted in the best cooling distribution at the end of the linac (top right) Final distribution longitudinal phase space after tuning linac cavities. It now has a small peak-to-peak energy spread. (bottom) The slice energy spread and current profile of the final distribution. Red lines indicate the target values for slice transverse emittance (solid) and peak current (dashed). The current is very uniform around the center at the target current 10 A.

2. *Laser heater for a uniform energy spread*

The injector design addressed the current uniform issue. However, simulations of the injector and linac revealed that the initial slice energy spread of the electron beam is non-uniform at the end of the linac and too low for efficient cooling, which can cause saturation in the amplifier. To mitigate this, a laser heater section is proposed to increase the uncorrelated energy spread and make it more uniform across the bunch. The laser heater operates through a resonant interaction between the electron beam and a laser pulse within an undulator, which imparts an energy modulation that is then converted into a controlled, uniform energy spread [46]. A series of 3D simulations were performed using the IMPACT-Z code.The study modeled a 150 MeV, 1 nC electron beam from GPT simulation, passing through a laser heater system with parameters similar to those of LCLS-II [46]. The simulations used a direct model of the electron-laser interaction with realistic Gaussian laser fields and undulator magnetic fields. The laser heater parameters are listed in the Table II

The simulation results successfully demonstrated the effectiveness of the laser heater. The final uncorrelated relative energy spread was increased to the target level of a few 10^{-5}, with a much flatter profile across the bunch compared to the initial distribution. This was achieved with no degradation of the transverse beam emittance. The results from the direct, detailed simulations also showed good agreement with predictions from a simplified analytical model. The required laser parameters, including an average power of approximately 1.1 kW, are considered feasible. Therefore a laser heater is a viable solution for generating the electron beam with the uniform energy spread necessary for the EIC SHC. The Laser heater may generates long wavelength (1 μm) noise, which will not be amplified the MBEC and the impact on the cooling is negligible.

TABLE II. Simulation Parameters for Laser and Undulator

Parameter	**Value**
Laser Parameters	
Laser wavelength	1.03 µm
Laser power	100 kW
Laser field amplitude	14.0 MV m^{-1}
Laser waist size	0.5 mm
Undulator Parameters	
Undulator length	1.1 m
Undulator period	0.11 m
Undulator B field on axis	0.11 T
Undulator parameter (K)	1.1

3. ERL lattice

In beamline order during energy recovery, the sections are the injector, the merger, the booster, the bunch compressor and high energy bypass lines (PX), the main linac, the beam dump, the laser heater, transport to the cooling section, the cooling section (modulator, amplifier, and kicker), the first turnaround, the return line, and the second turnaround – after which, the beam begins energy recovery. Alternatively, the beam can be steered through the merger into the diagnostic line.

The injector consists of a photocathode DC gun and a cryomodule containing three superconducting radiofrequency (SRF) cavities – two 197 MHz quarter-wave resonators (QWRs) and a 591 MHz single-cell cavity. The beam is accelerated to roughly 6 MeV and the accelerating cavities are run off-crest so that the bunch is chirped; the single cell cavity is at the third harmonic of the 197 MHz frequency and is used to linearize the longitudinal phase space.

The booster consists of two cryomodules, oriented so that the cavity order is two 197 MHz QWRs, two 591 MHz, and two 197 MHz QWRs. The beam energy is roughly 13 MeV, the bunch is significantly chirped, and the single cell cavities are again used to linearize the longitudinal phase space; however, the longitudinal phase space is evaluated after the bunch compressor in the following section.

Next, the beam enters the PX section, which consists of the P1, P2, and P3 lines. Each line is energy-specific and during the acceleration of both configurations, the 13 MeV beam goes down the P1 line, which compresses the bunch; bunch compression is tunable due to the quadrupole magnets included in the compressor.

Following the PX section is the main linac, consisting of eight 591 MHz five-cell SRF cavities and four 1773 MHz five-cell SRF cavities, with the 1773 MHz cavities placed in the center and four 591 MHz cavities on either side. The beam is accelerated to the top energy of either 150 or 55 MeV and is no longer chirped. Similar to the injector and booster, the longitudinal phase space is linearized with the 1773 MHz cavities.

The dump chicane and the laser heater chicane follow the main linac, after which the beam is transported to the cooling section, which consists of the modulator, amplifier, and kicker. After the cooling section, the beam separates from the HSR before entering a Bates bend and going down the return line.

At the end of the return line is a 591 MHz five-cell SRF cavity, which is run at the zero crossing and chirps the beam, minimizing the energy spread of the beam at the dump. Without this chirp cavity, the only way to minimize the energy spread is to have the beam energy at the dump to be significantly different than the 6 MeV injection energy.

After the return line is a second Bates bend – following this, the high energy line merges with the injection line, and the chirped beam begins deceleration and energy recovery. The beam is decelerated during the second pass through the booster, then enters the PX section. For the second pass through PX, the beam enters either the P2 or P3 beamline, corresponding to the 55 and 150 MeV modes, respectively. During the second pass through the linac, the beam decelerates to the injection energy and is transported to the beam dump. Closed optics exist for both operational modes, shown in Fig. 9.

a. SPACE CHARGE EFFECTS IN THE BEAMLINE

Given that the electron beam has a relatively low energy until the main linac, the space charge effect plays a dominant role in influencing beam quality, and a rigorous treatment is essential for minimizing emittance and energy spread. To account for these effects, the beamline was divided

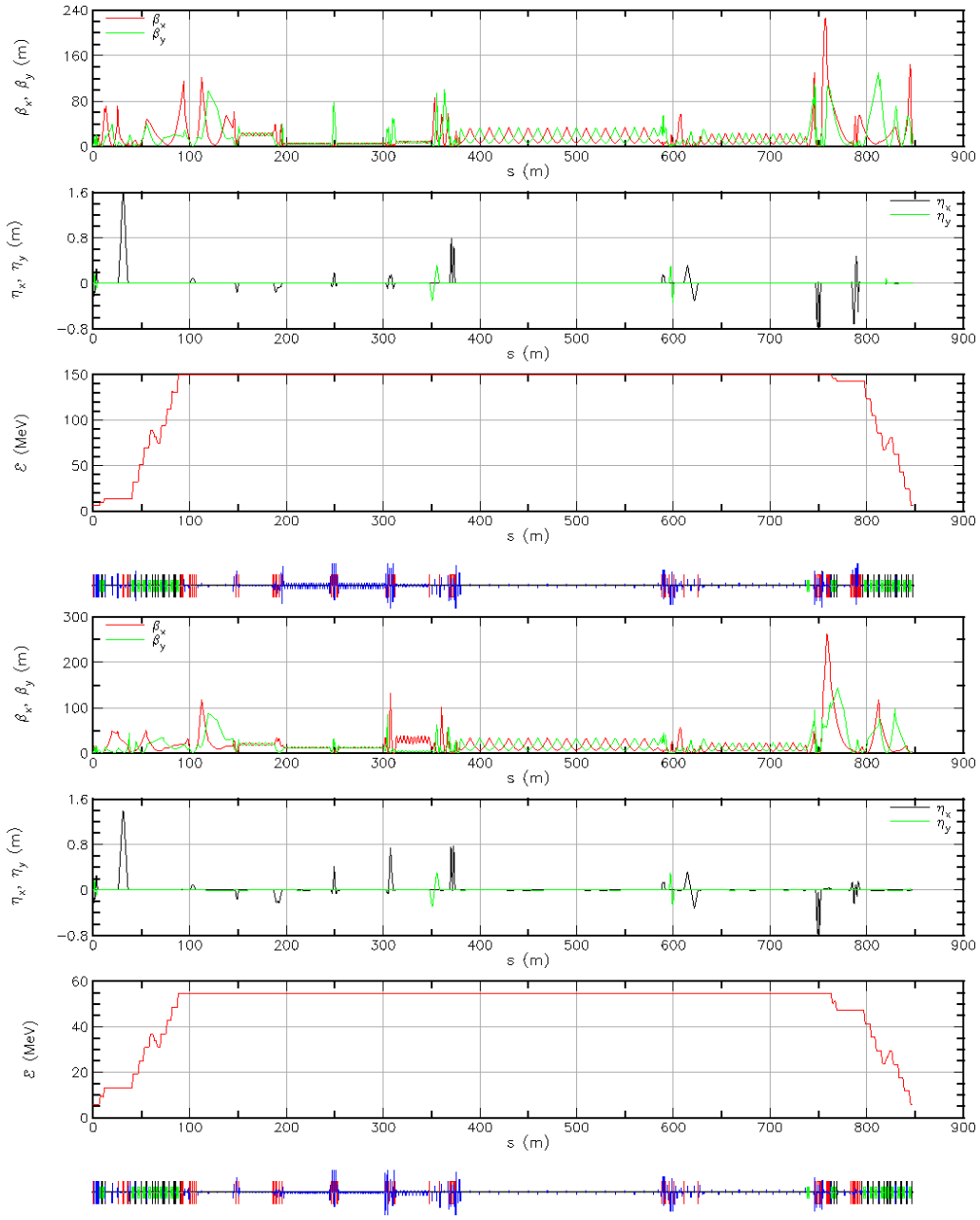


FIG. 9. The closed design optics for the 150 MeV (top) and 55 MeV (bottom) modes of the SHC-ERL. Shown are the beta functions in meters (top), dispersion in meters (middle), and the beam energy in MeV (bottom), all as a function of s in meters. Below the three plots is the component layout.

into two distinct regions: a space charge-dominated regime and a space charge-free regime. An interface point was defined between these regions, ensuring a proper Twiss match there as to maintain required beam quality across both regimes.

For the injector and merger, we utilized the python package Xopt with the continuous NSGA-II optimizer. The primary objectives of the injector optimization are to minimize the beam's emittance and energy spread at the injector exit. The longitudinal and transverse dynamics are largely independent, therefore we decoupled the longitudinal and transverse optimization. The longitudinal optimization tunes three cavity voltages and phases to minimize the bunch energy spread while maintaining 6 MeV beam energy. The transverse optimization uses two solenoids to minimize the horizontal and vertical emittance. The merger section transfers the electron beam from the injector into the pre-cooler linac, ensuring minimal emittance growth and maintaining zero dispersion at the merger exit. The merger optimization uses 5 solenoids to control beam focusing and dispersion. Chromatic aberration effects caused by the inherent large energy spread

of the bunch eventually result emittance growth in the downstream booster and PX sections. This effect was mitigated by introducing an additional objective in to the optimizer. The optimizer was modified to minimize the variation in slice Twiss functions alongside the original objectives, ensuring better phase space alignment and reducing overall emittance growth.

However, the initially chosen space charge interface was shifted further downstream after it was observed that nonlinear effects in the PX chicane were significant. There was evidence suggesting a presence of emittance exchange mechanism within the PX chicane, as the vertical emittance increased at the exit of PX. We investigated potential coupling in the upstream line, examining the injector solenoids and RF setting in the booster cavity in the injector and taken these observations into account for the optimization criteria. The optimized element parameters for the space charge dominated regime of the lattice are explained in a technote [47].

Although the R56 values were determined without accounting for higher-order momentum-dependent terms, these effects nonetheless contributed to an increase in the peak current for Mode A. This led to a deviation from the desired flat-top particle distribution, resulting in a skewed profile, with a peak current of approximately 20 A, for mode A. However, this skewness could be mitigated by optimizing the booster RF parameters [44]. The booster RF parameters were optimized to lower the peak bunch current, and restore a symmetric distribution.

The transition point between the space charge–dominated and non–space charge–dominated regions is then shifted to the end of the first cryomodule in the main linac LA). The normalized emittances at this match point, needed to be closer to the required value for effective cooling. The achieved normalized emittances for both modes are as follows;

For Mode A

- $\epsilon_{nx} = 2.7259 \times 10^{-6}\,\mathrm{m\,rad}$
- $\epsilon_{ny} = 2.9023 \times 10^{-6}\,\mathrm{m\,rad}$

For Mode B

- $\epsilon_{nx} = 2.8567 \times 10^{-6}\,\mathrm{m\,rad}$
- $\epsilon_{ny} = 2.8204 \times 10^{-6}\,\mathrm{m\,rad}$

Figure 10 illustrates the beam sizes and Twiss variation of through the space charge dominated region.

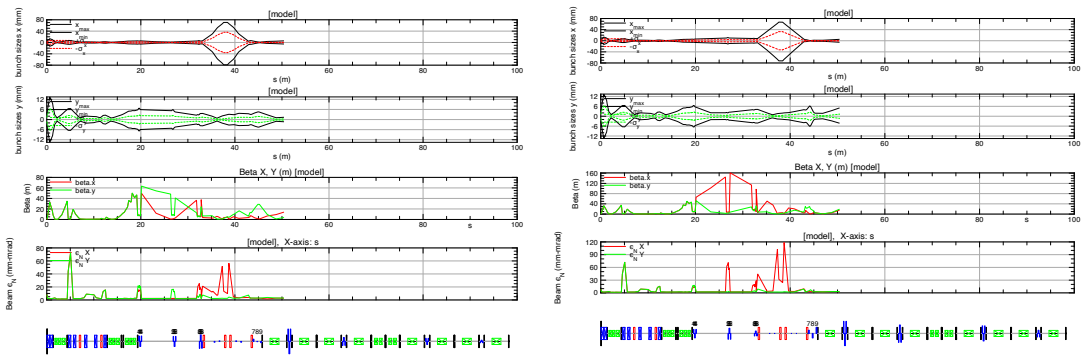


FIG. 10. Beam parameter variation along the space charge dominated region for mode A & mode B lattices

b. TIME OF FLIGHT CONSIDERATIONS

Typically, ERLs have a single linac and consequently, only a single time of flight concern – the time of flight between the linac exit and the linac entrance. However, this machine has two accelerating sections, the booster and the linac, which means that there are two time of flight concerns. The first concern is the time of flight between the booster exit and the booster entrance; the second is the time of flight between the linac exit and the linac entrance.

The booster time of flight uses path length changes in the two Bates bends for flexibility, similar to the Jefferson Lab Free Electron Laser ERL drivers; correctors at the entrance and exit of the Bates bend allow for the beam to enter and exit on-axis, while the beam orbit through the magnet is off-axis, changing the path length through the Bates bend [48]. At the 197 MHz fundamental frequency of the booster and a 2.5 cm maximum orbit excursion at the center of the bend, this translates to a range of ±11.7° per Bates bend. However, the length of this machine introduces a

complication – for a fixed path length on the order of 800 m, the time of flight for the two energies differs by roughly 8° at the booster frequency, but the time of flight between the booster exit and entrance is required to be the same for both modes. The second Bates bend was placed so that the relevant time of flight was roughly 4° from the requirement for both energies and by design, both energies travel through the second Bates bend off-axis to achieve the requisite time of flight.

The linac time of flight is controlled by the path length through the appropriate high energy PX line. Both P2 and P3 have moving stages in order to physically move two dipoles and change the on-axis path length through the lines, similar to the moving stages in the CBETA splitters [49, 50].

c. PX SECTION

The floor plan of the PX section, with the separate lines labeled, is shown in Fig. 11. The design of this section allows for the SHC-ERL to switch operational modes without requiring a change to the physical layout of the machine. While a simpler design would have a single high energy line which bypasses the bunch compressor of P1, such a design required either very large beam pipes or placing the first high energy dipole so close to the common dipole that the risk of field leaking into the P1 line was considered non-negligible.

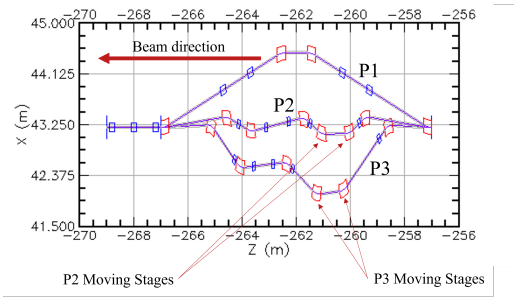


FIG. 11. The floor layout of the PX section; the three separate lines are P1, P2, and P3, top to bottom (as labeled), which correspond to the beam energies of 13, 48, and 143 MeV, respectively. The "common" dipoles are the right most and left most dipoles, which is a common magnet for all three beam energies. In both the P2 and P3 lines, the second and third dipoles are on moving stages which can be remotely controlled to change the physical path length of the on-axis orbit.

Each line in PX corresponds to a specific beam energy; P1, P2, and P3 are designed for beam energies of 13, 48, and 143 MeV, respectively. P1 compresses the accelerating beam for both modes, while P2 and P3 transport the decelerating beams for the top energies of 55 and 150 MeV, respectively. While all lines are achromats, none is isochronous; though it would be ideal for R_{56} to be zero in P2 and P3, it is more critical to close dispersion, dispersion prime, and control the transverse optics through these lines. Given that the decelerating beam is chirped, the non-zero R_{56} of these lines does mean that some bunch stretching occurs. If necessary, the R_{56} of the second Bates bend can be tuned to compress the bunch, so that the bunch length during the second pass of the linac is sufficiently short.

In order for the time of flight requirement between the linac exit and entrance to be correct, the booster time of flight must be correct. This is driven by the limited range of the moving stages, i.e., the limited time of flight flexibility in the P2 and P3 lines, as well as the inherent geometry of the PX section; if the decelerating beam enters either high energy line at an energy significantly different than design, the beam will be lost on the beam pipe wall before reaching the first high energy dipole. While the 150 and 55 MeV configurations have very similar linac time of flight requirements, the P3 line has an extra wavelength of path length in order to remove geometry conflicts with the P2 line. As the two lines have different geometries, the time of flight flexibility is different. The P2 line has a range of +20°/-10°, while the P3 line has a range of ±55°, both with respect to the 591 MHz fundamental frequency of the linac. The physical motion of the moving stages corresponding to these ranges is shown in Fig. 12.

The design upstream of the main linac, particularly the booster and PX sections, is motivated by the need to produce a supergaussian longitudinal distribution, with small slice energy spread and transverse emittance. At present, we achieve this by producing a very long bunch at the gun, which controls both the slice energy spread and transverse emittance, until the beam is at a sufficiently high energy that when the bunch is compressed to the requisite bunch length, these parameters are relatively unperturbed through the compression process.

The inclusion of a bunch compressor within the accelerating section of an ERL is not a typical configuration, and its inclusion makes the design more complex. As a design alternative, the injector energy could be increased to 13 MeV and the bunch compressed before the merger. This

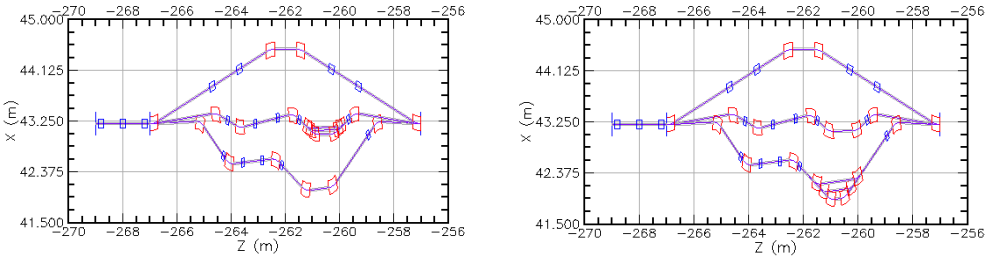


FIG. 12. The floor layout of PX section demonstrating the full range of motion of the moving stages for lines P2 (left) and P3 (right).

would have the benefit of returning to a more simple ERL layout, with a single time of flight requirement and no need for higher-energy bypass lines to transport the decelerating beam around the compressor chicane. However, this design comes with other drawbacks – lower energy efficiency and higher shielding requirements at both the dump and the diagnostics line. Another alternative is to compress the 6 MeV bunch in the injector – however, this approach does not produce the required beam parameters.

4. High fidelity simulation using real number of electrons

Maintaining low initial electron beam noise is critical for effective SHC. Shot noise within the electron beam can be amplified by microbunching instabilities, particularly at short wavelengths. This amplified noise can obscure the intended hadron-induced modulation signal, thereby degrading the cooling rate. To study this, we performed high-fidelity 3D macroparticle tracking simulations of the shot noise evolution using the IMPACT code suite. These simulations, which used the real number of electrons, followed the accelerator layout shown in Fig. 1 from the source to the end of the linac.

The results show a long-wavelength relative current modulation ($\lambda \approx 280\,\mu$m) at the linac exit. This wavelength is well outside the amplifier's operational frequency range and thus will not be amplified. After filtering out this modulation, the remaining relative RMS current fluctuation is approximately 7.5×10^{-4}. This level is consistent with the expected noise from a smooth beam distribution, suggesting that the initial high-frequency shot noise is not significantly amplified through the accelerator. Detailed results are presented in [51].

To evaluate the impact of various collective effects (space charge, coherent synchrotron radiation, wakefields, and intrabeam scattering), we also performed high-fidelity simulations of intentionally modulated electron beams at 55 MeV and 150 MeV. These simulations tracked the beams through the two-stage microbunching amplifier [52]. For an initial modulation, we achieved an amplification factor of 18 at 55 MeV and 11 at 150 MeV. Space charge effects were the dominant factor, leading to greater amplification at the lower beam energy. The contributions from CSR and resistive wall wakefields were found to be weak, while the effect of intrabeam scattering was negligible.

The final modulation amplitude can be optimized by adjusting the bending angle of the chicane located between the two amplification stages. For instance, a 50% increase in this chicane's bending angle was shown to double the resulting current modulation amplitude.

5. Lattice error preliminary study

The ERL lattice error were preliminary studied and discussed in the Ref. [53–55]. The study focused mainly on the 180° Bates bend dipole misalignment and have investigated sources of betatron mismatch when misalignment and excitation errors are present. The modulator and kicker sections of the ERL have been identified as the areas most sensitive to these quad errors.This has two major implications: first, any corrections will likely be non-local to the errors themselves, and second, some level of mismatch will probably be present throughout the ERL.The sensitivity of the cooling process to these beam transport errors (including missteering and focusing errors) is not included in this study. Therefore, a crucial next step is to evaluate how the cooling rate is affected by betatron mismatch in the electron beam. This will define the required tolerance and clarify how precisely the beam transport must be controlled.

The current operational assumption is that it will be possible to validate the lattice using differential orbit measurements—similar to methods developed at CEBAF—and then match the beam properties to the validated transport. However, if this non-local correction scheme proves insufficient, profile diagnostics and correction processes may be required within the modulator and kicker sections themselves. This presents a significant challenge, given the presence of the hadron beam in these areas. Detailed study should be continued to evaluating the error impact on the cooling performance.

6. Beam Halo and Mitigation Strategies

One of the critical challenge in high-intensity beam accelerators is the precise control and transport of the beam, particularly in mitigating the formation of a low-density particle cloud outside the core beam, commonly referred to as the "beam halo." Beam halo formation poses significant risks, including unwanted interactions that can damage lattice components and compromise system performance. This challenge is particularly acute in the SHC-ERL injector, where the beam's long bunch length and strong space-charge forces exacerbate halo development. Mitigating these effects necessitates effective halo suppression strategies and advanced collimation techniques.

To address these challenges, simulation studies utilizing the Bmad library were conducted to model the initial beam distribution at the cathode, incorporating transverse and longitudinal halo components. These simulations informed the design of practical collimation strategies, including transverse halo collimation methods and a longitudinal collimation approach leveraging a high-dispersion section. The proposed solutions aim to minimize halo-induced effects, protect accelerator components, and enhance overall beam quality in the SHC Energy Recovery Linac (SHC-ERL). Additionally, the study evaluated the potential of thin halo spoilers as an alternative to traditional collimators.

a. Halo distribution at cathode

The study of halo collimation begins with a halo distribution initialized at the cathode. The particle distribution at the cathode is generated using *Distgen*, a library developed to generate particle distribution, incorporating the parameters of the core beam of the SHC- ERL. According to the literature, the halo current is typically less than 1.0% of the core beam, corresponding to a halo current of less than 1 pC. The efficiency of the collimator can be assessed using multiple halo beams, as detailed in the following paragraph.

The halo beam consists of both transverse and longitudinal halo particles and is created by superimposing multiple distributions. While the core beam remained unchanged with a charge of 1 nC, the parameters for the halo particles are varied to generate multiple distinct distributions. The transverse space of this beam can include several types of distributions. Halo particles extending beyond the truncation radius of the core beam may follow uniform, radial-Gaussian, or hollow distributions as illustrated in the figure 13.

These distributions correspond to errors in the cathode spot size and surface imperfections in practical scenarios. The longitudinal halo typically originates mostly due to longer cathode response time, resulting in a right-skewed time distribution (t) that satisfies the temporal condition of $\sigma_{t-\mathrm{halo}} > \sigma_{t-\mathrm{core}}$, where $\sigma_{t-\mathrm{halo}}$ and $\sigma_{t-\mathrm{core}}$ are the RMS bunch length of the halo and core, respectively.

Halo formation along the injector-merger beamline using a truncated radial-Gaussian cathode distribution is observed. Figure 14 illustrates the beam phase space at the end of injector merger, where the core and halo particles are marked with two distinctive colors. As illustrated in this figure, the halo particles disperse more in both transverse and longitudinal directions. However, in longitudinal phase space, the tail exhibits a curvature distinct from the core beam. After some studies, we arrived at a plausible explanation for this longitudinal behavior: the effective space charge fields are stronger in regions where the core and halo overlap as the particle density is higher; The fields get weaker in areas with the low density halo tail; Also, the core is more uniform than the halo and particles near beam core experience a uniform RF curvature in the RF cavities and hence have a smaller curvature than the halo-tail.

b. Halo collimation simulations

The injector beamline, shown in Figure 15, includes three collimator locations marked in red. The first two collimators target transverse halo reduction, while the third, located in a high-dispersion region, is dedicated to longitudinal halo collimation. The simulated collimation settings, collimator geometries, and efficiencies are summarized in Table III.

In the first configuration, simulations of transverse halo collimation demonstrated that positioning collimators upstream of the cavities significantly reduced beam halo power by an order

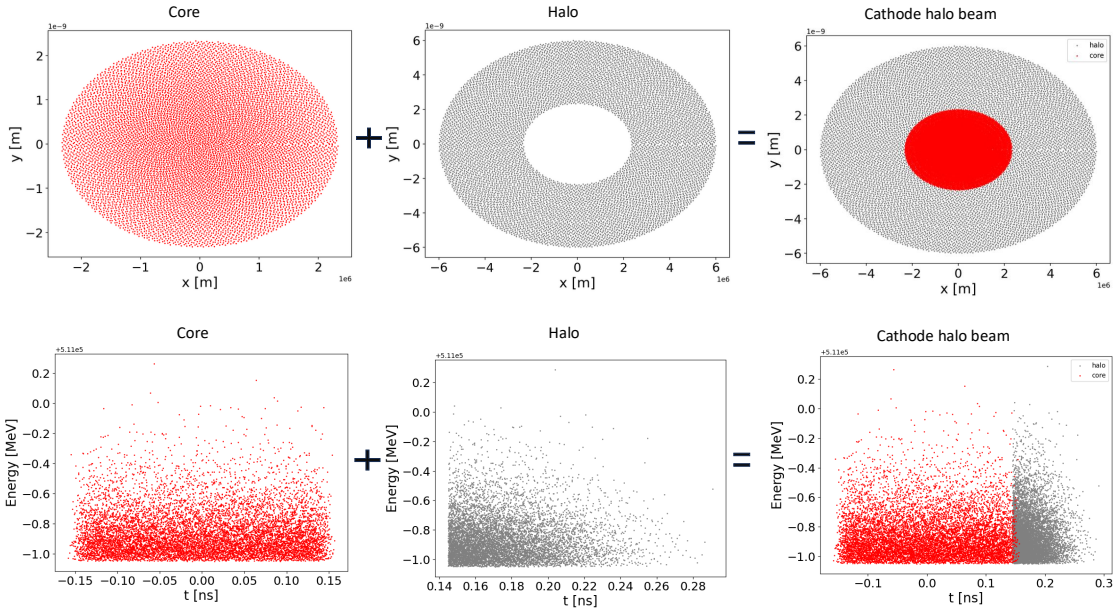


FIG. 13. Generation of transverse (top row) and longitudinal (bottom row) cathode halo distribution as a superposition of core and halo distributions.

.

of magnitude. This improvement is attributed to the lower beam energy in this region, which enhances collimation effectiveness. The setup achieved an impressive efficiency of 98%, though the limited spatial availability in this area necessitates meticulous beamline optimization to integrate the collimators effectively.

The second configuration, also targeting transverse halo collimation, involved placing a collimator downstream of the cavities. This arrangement achieved similarly high efficiency, exceeding 97%, but resulted in increased power deposition on the collimator due to the higher beam energy at this location.

When beam tracking was performed with space charge for the same initial distribution, significant changes in the phase advance and overlapping of the cores and halos were observed, which indicates that the proposed collimation scheme requires further improvements. The new collimator locations are given in the Figure. 16.

For longitudinal halo collimation, a dual-collimator setup was investigated in the third configuration. However, simulations revealed a low collimation efficiency of only 8.5%, primarily due to insufficient dispersion in the beamline. This result highlights the need for further refinements to the injector beamline design to enhance dispersion and improve longitudinal halo collimation effectiveness.

TABLE III. Collimation simulation results.

collimation setting	**halo type**	**space** charge	**collimator geometry**	**collimation efficiency**
collimator 1	transverse	off	R_1 = 2.3 mm & L_1= 10 cm	98%
collimator 2	transverse	off	R_2 = 5.4 mm & L_2= 5 mm	97%
collimator 1B & **2B**	transverse	on	R_1 = 6.5 mm & L_1= 10 cm R_2 = 1.2 mm & L_2 = 10 cm	92%
collimator 1B, 2B & **3B**	transverse	on	R_1 = 6.5 mm & L_1= 10 cm R_2 = 1.2 mm & L_2 = 10 cm R_3 = 1.0 mm & L_3 = 10 cm	96%
collimator 2 & **3**	longitudinal	on	R_2 = 5.5 mm & L_2 = 5 mm R_3 = 5.5 mm & L_3 = 5 mm	8.5%

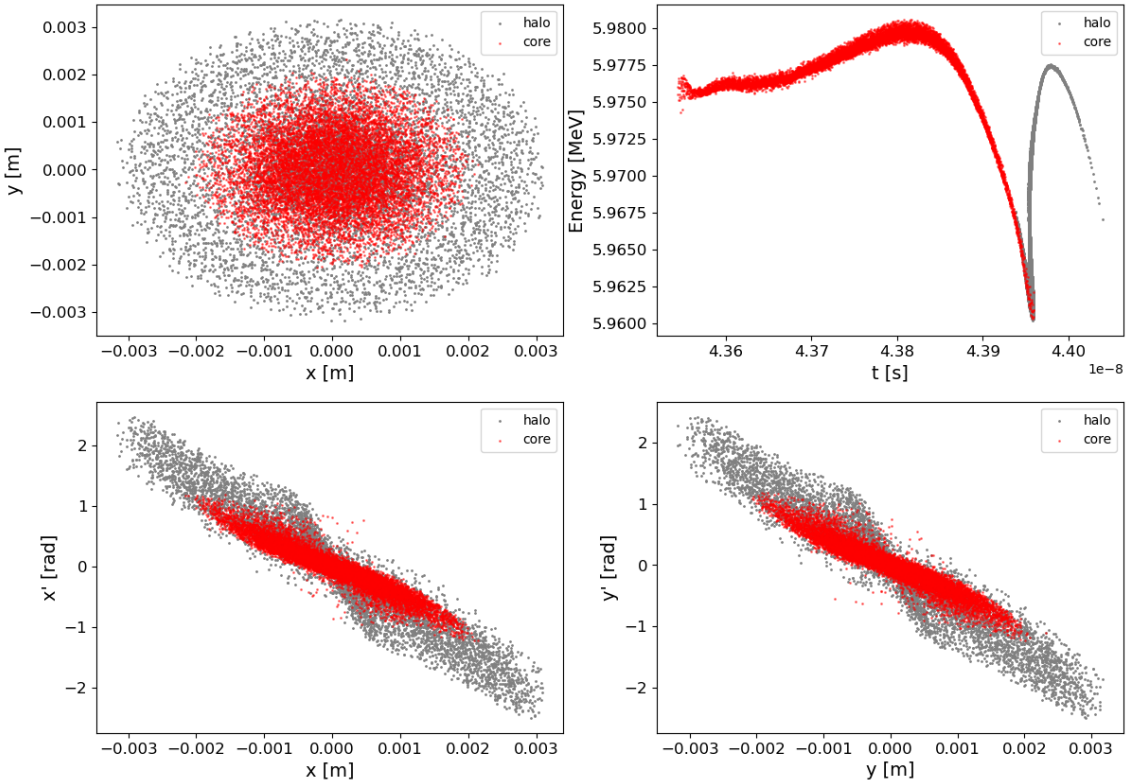


FIG. 14. Halo formation with space charge fields enabled, using a core beam with a charge of 1 nC and halo distribution is of charge $\cong$ 0.5 pC. In the visualization, core particles are represented by yellow dots, while halo particles are depicted as red dots.

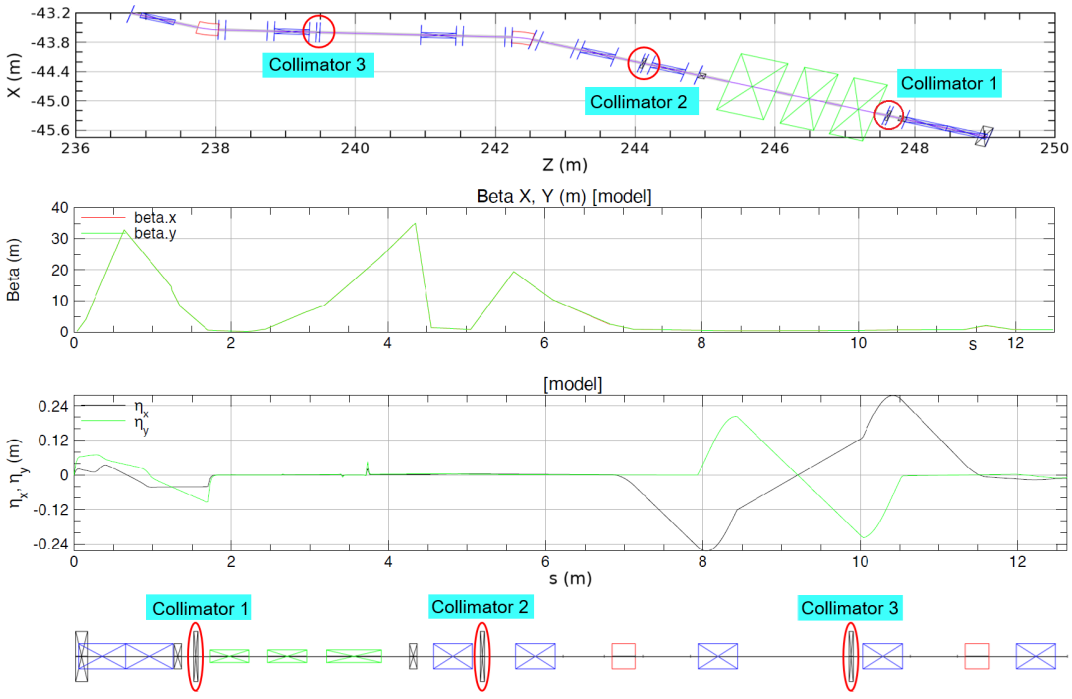


FIG. 15. Injector beamline.

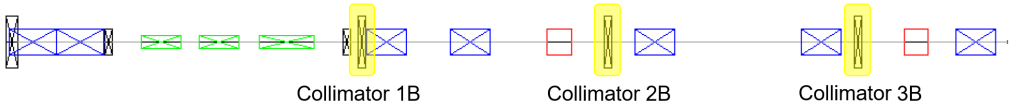


FIG. 16. Placement of the collimators when space charge tracking turned on. The letter “B" is added to differentiate from the locations in Figure 15. The collimator locations are marked with highlighted regions.

7. *Halo Spoiler*

This section investigates the potential use of thin halo spoilers to increase the divergence of halo particles, causing them to be lost on the downstream beam pipe, rather than relying on traditional collimators that completely stop the beam. Unlike conventional collimators, halo spoilers act as emittance spoilers, designed to scatter the halo beam without fully stopping it. This approach significantly reduces the power deposited in the spoiler, potentially eliminating the need for active cooling systems.

The scattering of electrons in the spoiler was simulated using G4Beamline [56], which is based on Geant4 [57]. The beam was transported from the electron gun to the first collimator, located between two solenoids, as modeled in Bmad. The beam distribution was recorded at the entrance of the collimator and subsequently converted into a G4Beamline-compatible format. In G4Beamline virtual detectors were placed before and after the collimator to record the incoming and outgoing electron distributions.

Simulations were conducted using aluminum and silicon spoilers with thicknesses of 0.1 mm and 1 mm for a 6 MeV electron beam. The horizontal beam phase space at the front and back of the spoiler is presented in Figure 17. While the beam size remained largely unchanged, the beam divergence increased by two orders of magnitude after passing through the spoiler. As can be seen, significant degradation of the beam emittance was achieved using spoilers. As shown in sub-figure (f), thicker spoilers result in larger scattered beam divergence. However, there is no significant difference in the scattered beam divergence between aluminum and silicon spoilers of the same thickness. Due to its dielectric properties, a silicon spoiler is preferable, as it generates much lower beam impedance compared to aluminum, making it a more suitable choice for beamline applications.

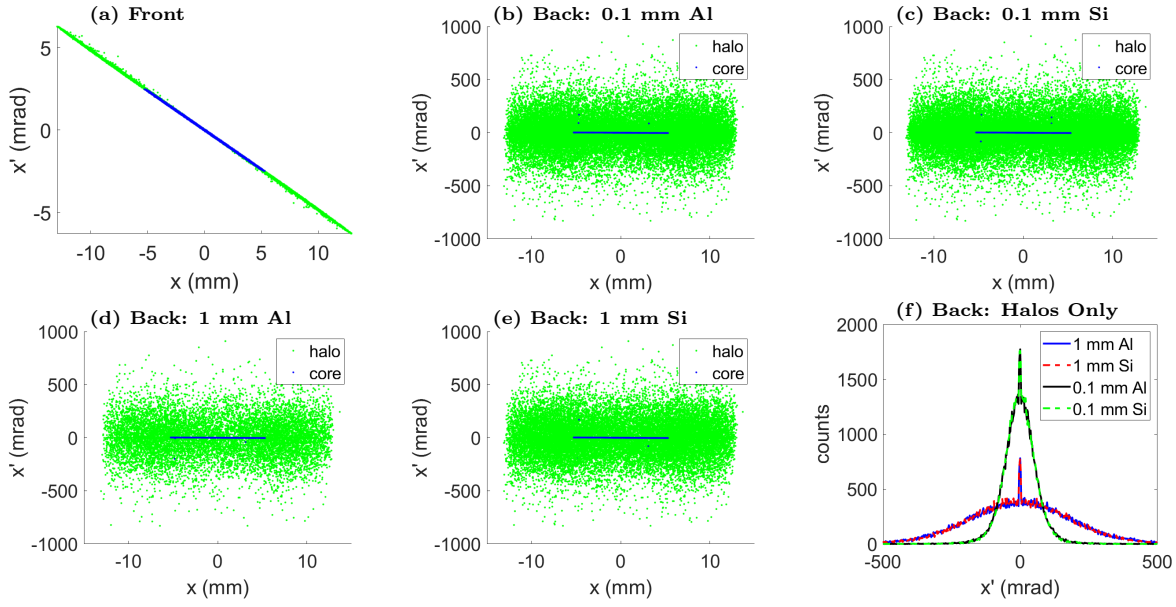


FIG. 17. G4Beamline simulation results: Horizontal phase space at the front (a) and back of the spoiler for various configurations: (b) 0.1 mm aluminum, (c) 0.1 mm silicon, (d) 1 mm aluminum, (e) 1 mm silicon, and (f) scattered halo beam only.

Assuming a halo-to-core ratio of 1:1000 and a core beam current of 100 mA, the total power of the halo was calculated to be 600 W. A regular beam stopper collimator would absorb most of this power. Table IV summarizes the power deposited in the spoilers and the downstream beam pipe.

The power deposited in the spoilers are 1 to 2 orders of magnitude smaller than beam stopper collimator. Notably, the 0.1 mm thick targets resulted in only 4 W of power deposition, which is an order of magnitude less than the 1 mm thick targets. This makes thin targets favorable, as 4 W is sufficiently low to eliminate the need for additional cooling, simplifying design requirements.

TABLE IV. Power deposited in spoilers and beam pipe.

Component	1 mm Al	1 mm Si	0.1 mm Al	0.1 mm Si
Power in spoiler (W)	44.5	39.5	4.2	3.7
Power in 1.5 m SST pipe (W)	457	460	388	384

The results show that the power deposited in the spoilers is 1 to 2 orders of magnitude lower than in a typical beam stopper collimator. Notably, the use of 0.1 mm-thick targets resulted in power deposition as low as 4 W - an order of magnitude less than the 1 mm-thick targets. This substantial reduction makes thin targets particularly advantageous, as the 4 W power level is sufficiently low to negate the need for additional cooling systems, thereby simplifying the overall design and operational requirements.

The beam power deposited in the spoilers and downstream beam pipe was estimated for various configurations. A 1.5 m long, 1.7 mm thick stainless steel (SST) beam pipe was placed immediately after the collimator. The power deposition along the pipe is shown Figure 18. A power deposition of 433 W was estimated to cause a temperature rise of approximately 157 K in the SST pipe with convective cooling. To mitigate this, a water-cooled pipe is needed to manage the heat load effectively.

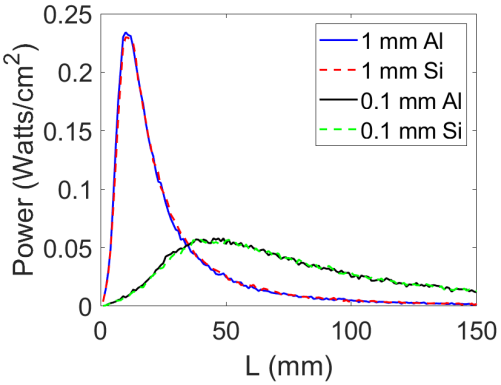


FIG. 18. Power deposition along the beam pipe. A virtual detector (purple) is positioned at the end of the pipe.

8. *Impedance Budget*

The SHC-ERL is designed to deliver a bunch charge of 1 nC with an average current of 100 mA. A key challenge in such high-current accelerators is managing beam impedance and wakefields, which can induce beam instabilities, energy spread, and power losses. This subsection summarizes the longitudinal impedance budget of the SHC-ERL, focusing on key components such as Beam Position Monitors (BPMs), gate valves, bellows, ion pump slots, and pipe connections. The impedance contributions are estimated using a combination of simulations (CST and ECHO3D [58, 59]) and analytical approaches, with benchmarking to ensure accuracy.

a. Wakefield simulation benchmarking

Wakefield simulations were benchmarked against analytical models for a Gaussian bunch with an RMS bunch length of 3 mm in a 10 cm stainless steel beam pipe. The analytical loss factor, as derived in Ref. [60], is given by:

$$k_{\text{loss}} = \frac{Z_0}{8\pi^2 bc}\Gamma\left(\frac{3}{4}\right)\left(\frac{2}{\sigma_c\mu_0}\right)^{1/2}\sigma_t^{-3/2}, \tag{7}$$

where $Z_0 \approx 377\,\Omega$ is the vacuum impedance, c is the speed of light, b is the pipe radius, σ_c is the pipe's electrical conductivity, μ_0 is the permeability, and σ_t is the bunch length. The results,

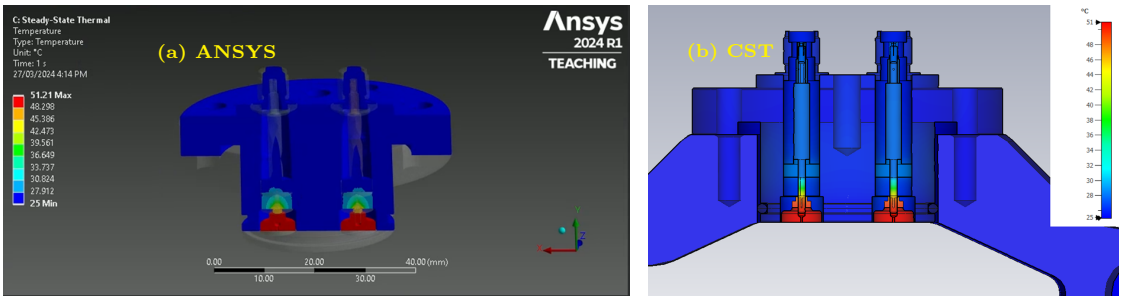


FIG. 19. Thermal simulation results with (a) ANSYS and (b) CST.

shown in Table V, indicate that CST and ECHO3D simulations yield loss factors within 1.1% and 3.3% of the analytical value of 1.80 V/nC, confirming the reliability of the simulation setups.

TABLE V. Loss factor estimated by simulations and analytical formula.

Method	k_{loss} (V/nC)	$k_{\text{loss}}/k_{\text{loss, analytical}}$
Analytical	1.80	0
ECHO3D	1.74	3.3%
CST	1.82	1.1%

b. Thermal simulation benchmarking

Thermal simulations were conducted to assess the temperature rise in components due to wakefield-induced power deposition. Using CST and ANSYS, the NSLS-II dual-button BPM was simulated with a 500 mA beam current, applying 1 W to each button and sleeve. The ambient temperature was set to 25°C, with boundary conditions including radiation (emissivity = 0.6), air convection (23.15 W/m^2/K), and water cooling (10000 W/m^2/K). The results, shown in Fig. 19, indicate a temperature increase to 51°C for both tools, consistent with private communications with the authors of Ref. [61] when identical settings were applied.

c. BPM impedance

The SHC-ERL employs four-button BPMs from BNL, scaled for 73 mm and 121 mm beam pipes, as shown in Fig. 20. With 257 BPMs in the electron beamline and 185 in the hadron beamline, each with 3 mm molybdenum buttons, the impedance was evaluated for a super-Gaussian beam with a 7 mm RMS bunch length (Mode A). Mesh convergence studies, shown in Fig. 21, confirmed convergence at 40 cells per wavelength. The loss factors, listed in Table VI, are 7.67×10^{-4} V/pC for 73 mm BPMs and 4.55×10^{-4} V/pC for 121 mm BPMs, contributing 0.1971 V/pC and 0.0842 V/pC to the total budget, respectively.

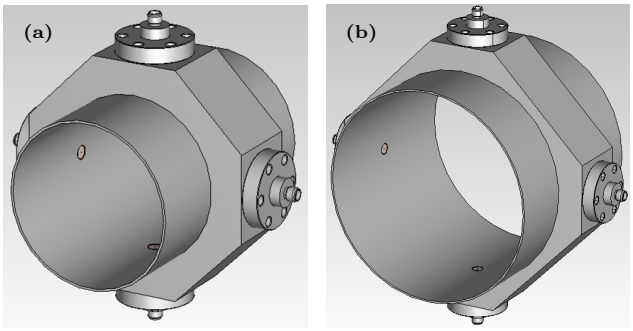


FIG. 20. Baseline BPM for SHC-ERL scaled for (a) 73 and (b) 121 mm beam pipes.

TABLE VI. BPM loss factors.

Object	Number of Components	k_{loss} (V/pC)	Σk_{loss} (V/pC)
BPM 73 mm	257	7.67×10^{-4}	0.1971
BPM 121 mm	185	4.55×10^{-4}	0.0842

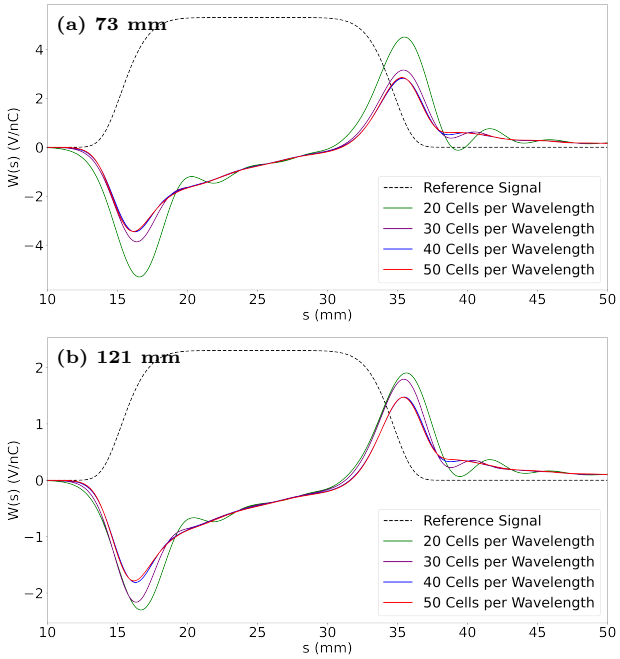


FIG. 21. Wakefield convergence studies of BPMs.

TABLE VII. Loss factor of bellows.

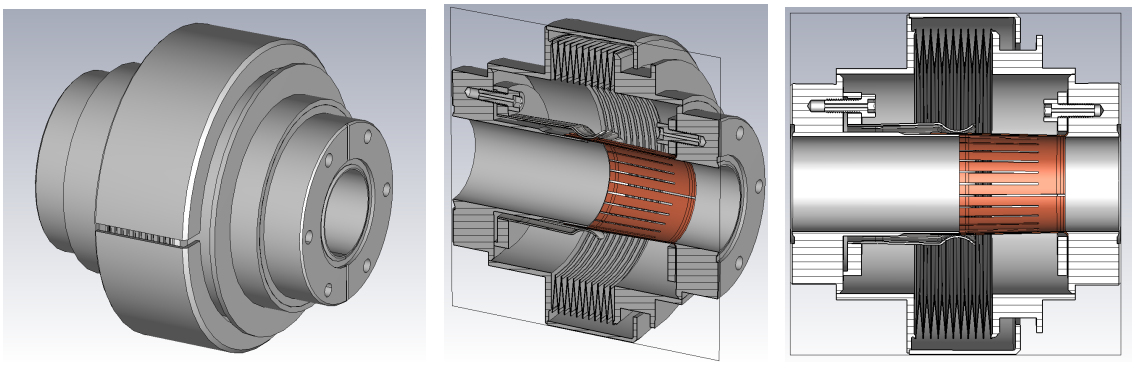
	$k_{\text{loss, 73 mm}}$ (V/nC)	$k_{\text{loss, 121 mm}}$ (V/nC)	$k_{\text{loss, 73 mm}}$ (V/nC)	$k_{\text{loss, 121 mm}}$ (V/nC)
Number of fingers	72	120	26	26
Scaled finger gap width	N/A	N/A	12.57	25.49
Non-scaled finger gap width	12.42	24.99	11.98	12.0

d. Gate valves

MDC Precision gate valves (DN63CF for 73 mm and DN160CF for 121 mm pipes) are used in simulation. Simulations indicate that their impedance contributions are manageable, though specific loss factors are under further optimization to ensure minimal impact on beam stability.

e. Bellows

Bellows with copper RF fingers, shown in Fig. 22, accommodate thermal expansion while minimizing wakefields. Initial simulations with scaled finger gaps yielded loss factors of 12.57 V/nC for 73 mm and 25.49 V/nC for 121 mm bellows, as shown in Fig. 23 and Table VII. The higher loss factor for the 121 mm bellows is due to larger finger gaps, which allow more field leakage. Simulations with non-scaled (equal) finger gaps reduced the 121 mm loss factor to approximately 12.0 V/nC, aligning with expectations for larger pipes.



(a) (b) (c)

FIG. 22. Bellow and RF fingers.

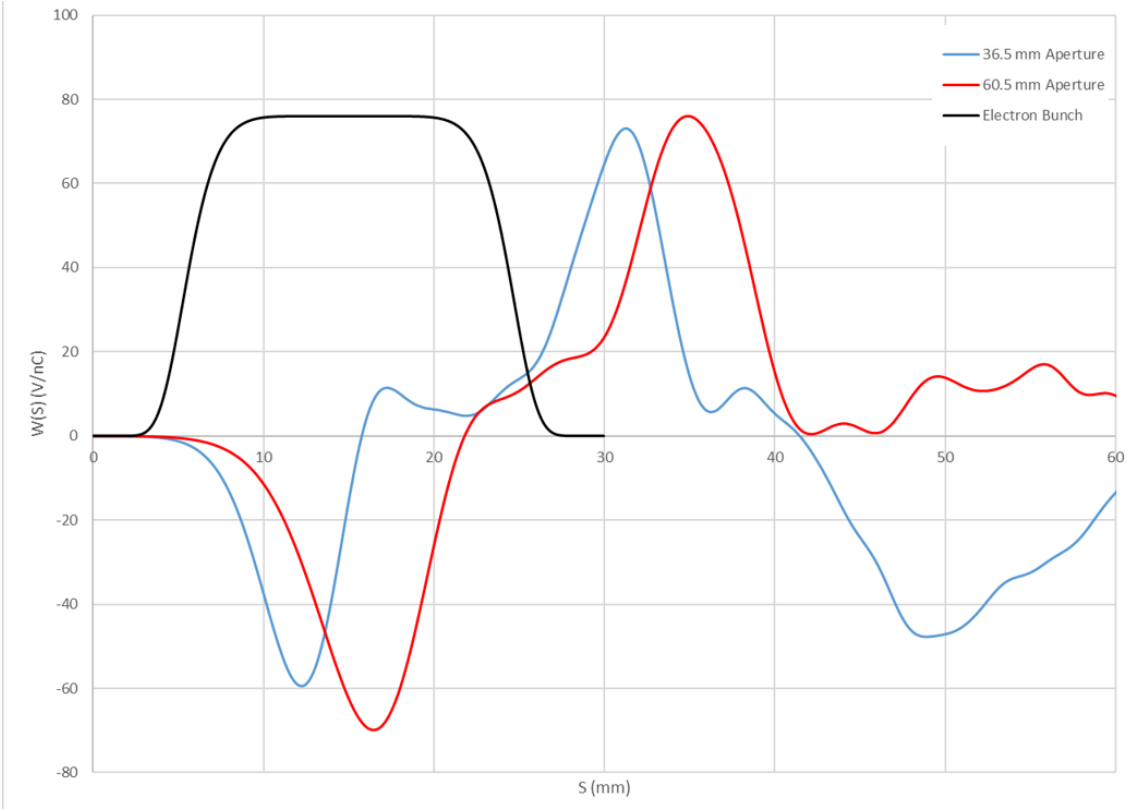
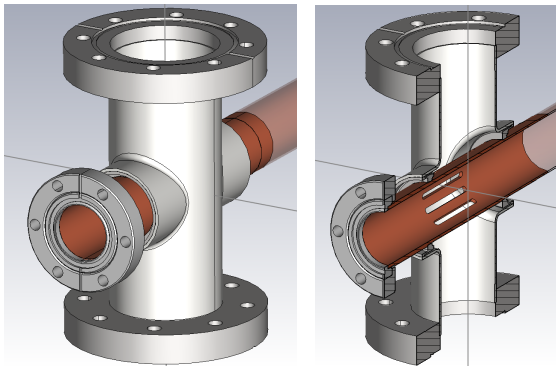


FIG. 23. Bellow wakefields with scaled gaps.

f. Ion pump slots

Ion pump slots use RF shields, as shown in Fig. 24, to mitigate high-frequency wakefields. Preliminary simulations suggest low impedance contributions, with ongoing studies to finalize loss factors.



(a) (b)

FIG. 24. RF shield for the ion pumps.

g. Pipe connections

Pipe connections use ConFlat flanges with copper gaskets, as shown in Fig. 25. Non-flush gaskets with larger inner diameters (39.1 mm, 38.1 mm, 37.1 mm) yield loss factors of 6.04 V/nC, 3.75 V/nC, and 1.72 V/nC for 73 mm pipes, respectively, as shown in Table VIII. Flush gaskets (36.5 mm) are expected to further reduce impedance. Wakefields and loss factors as functions of extension pipe length are shown in Figs. 26 and 27, indicating that shorter extensions minimize impedance.

h. Summary of impedance budget

The longitudinal impedance budget is dominated by bellows and pipe connections due to their higher loss factors. Ongoing optimizations, including refined RF finger designs for bellows and

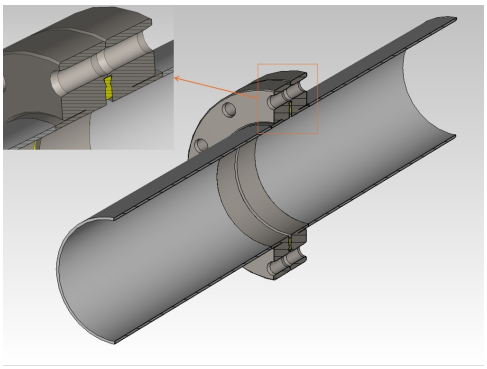


FIG. 25. Pipe connection with two flanges and a gasket.

TABLE VIII. Loss factor at pipe connection.

Gasket type	Gasket radius (mm)	$k_{\text{loss, 73 mm}}$ (V/nC)
Non-flushed	39.1	6.04
Non-flushed	38.1	3.75
Non-flushed	37.1	1.72
Flushed gasket	36.5	–

flush gaskets for pipe connections, aim to reduce these contributions.

I. Major components

1. *Longitudinal alignment measurement*

One of the most significant challenges for an MBEC cooler is the longitudinal alignment of the hadron and electron beams. During its propagation through the modulator, a given hadron will provide energy kicks to the neighboring electrons. The two beams then separate, propagate independently for roughly 100 m, and recombine in the kicker. In order for the cooling to work, each hadron must see the same electrons in the kicker as it did in the modulator, with the length scale for "same" set by the wake wavelength, on the order of 1-10 μm. Since there is no way that we could dead-reckon the necessary few parts-per-billion relative path-length difference, we need some method to determine if the beams are properly aligned or not. While measuring the cooling itself is one such option, the cooling timescale, and therefore the timescale of such a diagnostic, is on the order of hours, making startup painful and any sort of feedback prohibitive. We have therefore proposed the use of systematic changes in the radiative power emitted by protons in a downstream

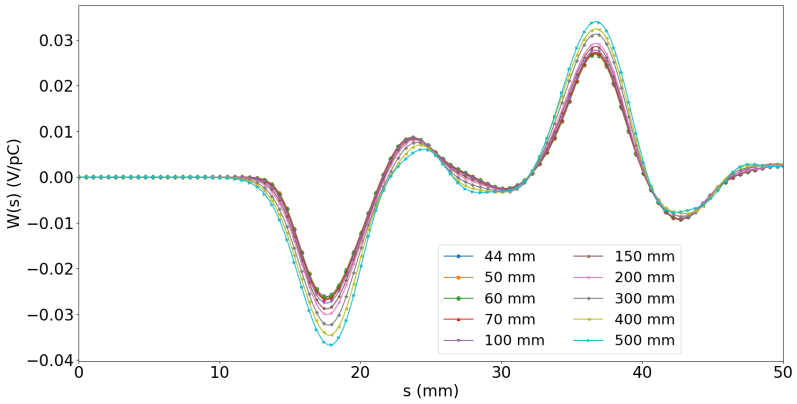


FIG. 26. Wakefield of extension pipe lengths.

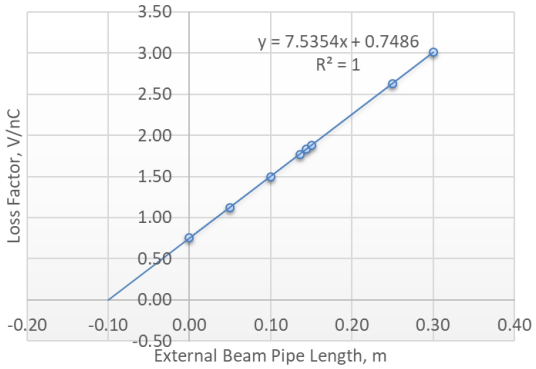


FIG. 27. Loss factor as function of extension pipe length.

dipole at wavelengths near the wake wavelength as a potential means of detecting alignment on the scale of seconds [62]. This would require a detector capable of seeing parts-per-thousand changes in nW-levels of radiation at wavelengths of a few to a few tens of microns [23]. Detectors with such sensitivity at these wavelengths are hard to come by and would require significant R&D efforts. Additional care must be taken to ensure that the longitudinal structure of the hadron bunch is preserved on this submicron-scale between the kicker and the dipole where the radiation would be produced.

2. Velocity alignment instruments

Effective cooling will depend on the accuracy of velocity matching between the electron and proton beam. Another process, the rate of ion-electron recombination, is also maximized when the velocities are matched, the its detect time is much faster. Detecting and maximizing recombination signals by tuning electron beam energy will be helpful in finding the narrow velocity matching window conducive to cooling. When a proton picks up an electron with a similar velocity, it is converted into H atom with nearly the same momentum. The detection of the recombined ions can be done by driving H atom into the beam pipe wall, creating showers of secondary particles which then can be detected outside the cryostat by using beam loss monitor. The SHC recombination monitor evaluation is discussed in the reference [63]. The average electron-ions recombination rate given by:

$$R = \gamma^{-2} N_p N_{\text{bunch}} \frac{\sigma_{es}}{\sigma_{ps}} n_e \frac{L_{\text{cool}}}{L_{\text{ring}}} \alpha_r \tag{8}$$

where,

$$\begin{aligned} \alpha_r = {} & Ac \frac{h\nu_0}{m_e \beta_x \beta_y} \sqrt{\frac{2h\nu_0}{m_e c^2}} \\ & \int_0^\infty \frac{1}{\sqrt{y + \frac{m_e v_{z0}^2}{2h\nu_0}}} \left[-\frac{1}{2} \ln\left(y + \frac{m_e v_{z0}^2}{2h\nu_0} \right) + \gamma_1 + \gamma_2 \left(y + \frac{m_e v_{z0}^2}{2h\nu_0} \right)^{1/3} \right] \\ & \exp\left(-\frac{h\nu_0}{kT_1} y \right) I_0 \left(\frac{h\nu_0}{kT_1} \eta \cdot y \right) dy \end{aligned} \tag{9}$$

including particle numbers (N_p), bunch number (N_{bunch}), electron density (n_e), and the geometry of the cooling section (L_{cool}) relative to the ring circumference (L_{ring}),the α_r is the recombination rate coefficient of an ion bunch moving together with an electron bunch. The analytical expression for α_r was applied to the EIC SHC configuration, which assumes 1160 proton bunches interacting with the electron beam over a 35-meter cooling section. Fig. 28 shows the recombination rate as the function of velocity different.

This rate is on the same order of magnitude as the ∼14,000 s^{-1} expected for the CeC-X experiment. Experimental testing is planned in the RHIC CeC-X experiment.

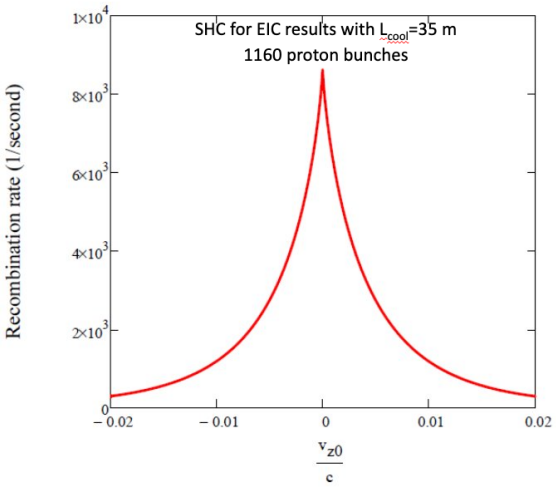


FIG. 28. The recombination rate of the electron and ion in the one drift section

3. *RF cryomodules*

a. SHC ERL RF cryomodules considerations: The current baseline ERL design employs BNL-type RF cavities, where each cryomodule houses a single cavity. While this approach ensures stability, it imposes significant spatial and cost constraints. We studied the feasibility of integrating multiple PERLE-type cavities, which allow for 4 cavities per cryomodule, thereby reducing the footprint and cost [64].

A key challenge in the adoption of multi-cavity cryomodules is the risk of Beam Breakup (BBU) instabilities, which arise from interactions between the beam and higher-order modes (HOMs) trapped in the cavities. To evaluate the impact of HOMs, a frequency-scaling approach was employed, assuming constant quality factors and shunt impedances while scaling HOM frequencies linearly. The study then compared the BBU threshold currents of BNL and PERLE cavities using two independent simulation tools: Bmad BBU and a MATLAB-based BBU tracking code.

The simulation results indicate that BNL-type cavities exhibit significantly higher BBU threshold currents than PERLE-type cavities, suggesting greater beam stability. The dominant HOMs contributing to BBU instabilities originate from the 1773 MHz cavities, making them the most critical elements for stability considerations. The findings suggest that while full adoption of PERLE cavities may not be viable due to their lower threshold currents. A hybrid approach—using BNL-type cavities for higher frequencies and PERLE-type cavities for lower frequencies—could provide a practical solution.

The study concludes that while PERLE-type cavities offer clear spatial and cost advantages, their susceptibility to BBU instabilities makes them less suitable for high-current ERL applications without additional HOM damping strategies. Future work will explore methods such as BBU feedback systems and emittance exchange techniques to improve stability while maintaining compact cryomodule designs.

J. R&D needs and in progress

1. *R&D for High current electron source*

SHC requires a high-brightness electron source capable of delivering a 100 mA average current. While radio-frequency (RF) guns are generally well-suited for high-brightness beams, developing a superconducting RF (SRF) gun that can provide 50 mA to 100 mA is not considered feasible within the EIC project timeline [65].

Consequently, the selected approach is to develop a high-voltage DC (HVDC) gun. This gun will generate a long, "pencil-shaped" electron bunch that is subsequently compressed to the desired length before injection into the main linac. To this end, and leveraging expertise from previous polarized source development and LEReC gun operations, a dedicated high current HVDC gun for hadron cooling is under development at Stony Brook University.

TABLE IX. Comparison of SHC electron source requirements and the state of the art value

	SHC	The state of the art value
Average current, mA	90-100	65
Bunch charge, nC	1.2	2
Lifetime,days	> 3	< 1

The HVDC gun R&D is guided by the following requirements:

- Development of an electron source capable of delivering a 98 mA beam with a single operational lifetime charge of 60 000 C.
- Operation at voltages exceeding 500 kV with an electric field gradient on the cathode surface of more than 5 MV m^{-1}.

The novel features of this R&D effort include:

- Development of an efficient cathode cooling mechanism through the high voltage feedthru to maintain the cathode temperature below 50 C under high laser power
- Development of an inverted feedthrough rated for above 500 kV to achieve to achieve good beam quality at nC range bunch charge
- Using epitaxially grown large crystal K2CsSb photocathode with CSi single crystal substrate to increase lifetime
- Gun attached Resistor network with Pan-cake connector to fully eliminate the 1 MV cable stored energy

The HVPS is a two-stage Cockcroft-Walton voltage multiplier, housed in a vessel that can be pressurized with either N_2 or SF_6 gas. It is designed to deliver an average current up to 120mA at 550 kV.

An external power inverter operating at a nominal 20 kHz drives the system. This relatively low operating frequency allows the inverter and multiplier tank to be located remotely from the gun, connected via a flexible high-voltage cable. The resulting voltage drop and ripple are on the order of 4×10^{-4} Two risk mitigation strategies have been developed to address potential issues with cathode survivability under high laser power:

1. **Alternative Cathode:** In the event the primary K2CsSb cathode cannot withstand the required laser power, a more robust NaKSb cathode will be used.
2. **Alternative Wavelength:** Switch to a 350 nm laser. This is expected to increase the Quantum Efficiency (QE) by a factor of 3 to 6, allowing for a corresponding reduction in laser power and significantly extending the cathode's operational lifetime.

The detailed gun design, test plan and progress can be found in [66]

2. R&D for high current ERL

The SHC cooling ERL would be the first SRF linac to routinely operate at the 100 mA current level. It has very tight beam specifications and must be extremely reliable. The specifications are given in Table X.

These specifications are evolving but the average current of 98.5 mA is a minimum requirement and some headroom is desirable. The transverse emittance is also rather small for this much charge. In addition to these specifications, the noise on the beam in the 100 THz region must be no more than 1.5 times the Poisson limit for the 275 GeV case and 3 times the Poisson limit for the 100 GeV case.

The highest charge recirculated in an ERL is 1 nC at the BINP FEL in Novosibirsk. The highest charge in an SRF ERL is 270 pC. The highest current in an ERL is 30 mA at BINP and the highest in an SRF ERL is 9 mA at Jefferson Lab. These were both FEL drivers so their design stressed high peak current, short pulses, and moderate transverse emittance. The longitudinal emittance of 250-350 keV-psec is comparable to the best achieved for such high charges.

TABLE X. ERL specifications for 100 GeV and 275 GeV proton cooling.

Parameter	100 GeV	275 GeV	Unit
Energy	55	150	MeV
Norm. Emit. (x/y)	2.8 / 2.8	2.8 / 2.8	mm-mrad
Repetition rate	98.5	98.5	MHz
Bunch Charge	≥ 1	≥ 1	nC
Average current	≥ 98.5	≥ 98.5	mA
Peak Current	8.5	17	A
Bunch Length[a]	14	7	mm rms
Fractional Energy Spread	10^{-4}	10^{-4}	—
H/V electron β_0 in Modulator	29.7 / 29.7	40 / 25	m
H/V electron β_0 in Kicker	10 / 10	4 / 4	m
Modulator Length	≥ 39	≥ 39	m
Kicker Length	≥ 39	≥ 39	m

[a] Gaussian bunches assumed.

There are many potential challenges with the required ERL design. The average current, high brightness with minimal halo, energy spread and stability, all at high charge, are unprecedented. R&D on the following topics is needed.

1. **Halo** – Halo is always a problem with any high current ERL. One can collimate no more than one part in 10^{-4} using a shielded collimator at a few strategically located positions. If collimation beyond this level is required, a significantly larger number of collimators would be necessary. Can one run as much as 100 mA with 1 nC bunches while keeping the halo from the injector and all other sources (Touschek, residual gas scattering, non-linear focusing,etc), below the 10^{-4} level?

2. **Ion trapping** – Ion trapping will probably be a problem. It is important to know the level of ions that might be trapped in this design. We also have to determine which is the best way to eliminate the ions. One would want to test out a few different solutions to reduce ion trapping.

3. **Turn-on transients** – It must be possible to turn the beam on from zero up to the 100 mA level with no losses or instabilities along the way. One cannot just turn the beam on instantly but should ramp the beam up adiabatically to its final value. It may cause losses or faults.

4. **Beam Break-Up** – HOM loading and Beam Break Up (BBU) can also lead to a limit on the average current. This can sometimes be tuned away by altering the phase advance of a phase trombone to change the net M_{12} or M_{34} between a cavity and itself (i.e., on accelerating and decelerating passes) but this only works if you are close to threshold. The best way to raise the threshold for BBU is to increase the HOM damping.

5. **Beam stability** – The RF control circuits see a very large reactive load on the cavities. The phase loop gain must be very high to keep the cavity stable. Can one keep the phase sufficiently stable to achieve 10^{-4} overall energy stability?

6. **Electron Beam Transport** – Errors in the transport may accumulate over tens of meters of transport and lead to strong aberrations or other non-linearities. It is possible to simulate these using ELEGANT or BMAD but it is most reassuring to see if you can manage these in a real machine.

7. **Control of microbunching** – With such a small slice energy spread, microbunching gain will tend to be very high. But one must keep the noise low so it is very important to keep the microbunching gain around unity before the cooling section.

8. **Commissioning approach** – Appropriate diagnostics must be available that allow both the setup and the stabilization of the beam to the required precision.

9. **Machine stability** – The stability of the machine must meet the stringent requirements of the SHC. This will probably require feedback on the beam position and energy. Can the required specifications be met?

10. **Longitudinal gymnastics** – The existing design assumes that the bunches can be stretched and compressed by approximately a factor of two. Can this be done while maintaining the beam specifications? A big advantage of the cooler ERL is that the energy spread is very small so chromatics should not be a problem. This is no longer true when one has to stretch and compress the bunch.

Possible Test Facilities

Some of the issues listed above can be addressed in an injector test stand that produces 100 mA at about 500 keV kinetic energy. Some may be addressed in a facility like LEReC, which can run up to 50 mA but does not energy recover. A few issues can be verified at high charge but moderate average current. Many of the most challenging problems can only be addressed by a working high-current ERL. Some may only be addressed in the final machine. Whatever the average current, one would like to run the tests at full charge, so, for example, if one can only run 20 mA one would want to run 20 MHz repetition rate at 1 nC. On the other hand, the ion trapping threshold should be lower for a higher repetition rate at lower charge so that option would be nice to have as well. If the test facility being used does not use 591 MHz RF it might be good to scale the charge by the inverse of the frequency so that the peak current is the same and the phase spread in degree is the same.

High current ERL and major components test could be carried out in the following test facilities.

1. **Local Gun test facility** – Currently a new high average current R&D gun is being built at SUNY, cathode cooling and the capability of high charge and long lifetime. To run at 100 mA this facility would need some expensive upgrades. This facility might test out lifetimes and beam quality in an electron source with a cooled cathode and can test out new type of cathodes as well. The facility could also test out the production of different laser pulse shapes that will be required for the EIC ERL. To test out operation at 100 mA, it will need a dump and electron beam source that can handle 100 mA of current. The dump must be sufficiently separated from the gun so that there is very little backflow of gas from the dump. A high power beam dump and HVPS have been constructed and are under testing now.

2. **LEReC** – This facility, though it does not energy recover, is capable of high current and might be useful in testing some of the ideas we have for transport and preservation of beam quality, including some tests of ion trapping.

3. **CBETA** – This facility was designed for 40 mA and four passes and, in principle, could reach 100 mA in single-pass mode. The injector has shown performance up to 65 mA. The single-pass ERL operation attained 60 μA but this was limited by a couple of factors that have since been fixed. The machine is not currently running and would require a moderately large amount of work to get it back up and running. The cost to run the machine is about \$2.5M per year.

4. **CERL at KEK** – This facility might be able to run at 10 mA and test out some of the issues above. It can run at low charge and high repetition rate leading to a smaller threshold for ion trapping and has well-instrumented collimators.

5. **BerlinPRO** – This facility was designed for 100 mA and the beamline and vault are built to handle this current. Unfortunately they lost their cryomodule to BESSY and have had many problems with their SRF gun. One could think of shipping them the CBETA cryomodule and RF power supplies as a possible way to get up to 100 mA. The ERL Roadmap committee wants to get this machine up and running at high current as a test facility for PEARL.

6. **New facility at BNL** – Once the ERL tunnel is built, one could install the ERL but with the first 180° bend before a temporary shielding wall near the end of the tunnel. The diagnostic line would also have to be moved upstream so that the beam quality can be verified. This machine could be run in parallel with HSR commissioning. This would allow continuous running of the ERL and testing out many features of it before cooling is required. The biggest drawback of this facility is the long wait before obtaining any results. Initial tests could occur at lower energy and current to allow for a staged purchase of RF for the machine. One could also run with only 4 cavities in the linac and still get to 55 MeV.

3. IR detector for longitudinal alignment

A major risk is the development of a detector capable of seeing weak proton radiation in order to ensure proper electron/proton longitudinal alignment, as discussed in subsubsection V I 1. Alternatively, finding some other method to quickly and reliably determine the longitudinal alignment of those two beams will eliminate this risk. Two principal detection schemes are evaluated: direct detection using semiconductor-based sensors and indirect detection via nonlinear wavelength conversion. Direct Detection

For the 275 GeV case, the optimal 5 µm to 8 µm range falls within the operating capabilities of commercially available HgCdTe photodetectors. These detectors offer high detectivity and can be operated with liquid nitrogen cooling (77 K), which is standard for laboratory applications. Environmental background from blackbody radiation of the vacuum pipe is a potential concern, but its contribution can be effectively removed using phase-locked measurements with a lock-in amplifier synchronized to the signal modulation frequency.

For the 100 GeV case, the 30 µm requirement is far more challenging. This wavelength falls into the long-wave infrared (LWIR) or far-infrared (FIR) region. The most suitable detectors are Si:As (Silicon doped with Arsenic) Blocked-Impurity-Band (BIB) detectors. While these devices offer the necessary sensitivity, they are not widely available commercially, as their primary applications are in defense and space-based astronomy (e.g., the James Webb Space Telescope). Furthermore, they require cryogenic cooling to liquid helium temperatures (at 4 K), which adds significant complexity and cost to the system.

An alternative to the challenges of direct LWIR detection is to use a nonlinear optical process to convert the signal to a more convenient wavelength. Three-wave mixing, also known as optical parametric amplification (OPA), is a promising technique. In this scheme, the weak, incoherent IR signal (ω_1) is mixed with a strong, coherent pump laser beam (ω_2) in a nonlinear crystal. This process generates an idler beam ($\omega_3 = \omega_1 - \omega_2$) at a shorter wavelength, typically in the near-infrared or visible spectrum, where highly efficient and low-noise silicon-based detectors are available.

This approach requires a nonlinear crystal with high transparency at all three wavelengths and a large nonlinear coefficient. For the 3-6 µm range, materials like Periodically Poled Lithium Niobate (PPLN) are suitable. For the more challenging 20-30 µm range, crystals such as $AgGaSe_2$ (AGSe) or $LiGaTe_2$ may be viable. While OPA is typically used with coherent sources, studies have shown its applicability to incoherent or broadband radiation, suggesting its feasibility for this application.

Based on these findings, a two-pronged research and development plan is proposed:

Develop a proof-of-principle direct detection system for the 5 µm to 8 µm range. This would involve procuring a commercial HgCdTe detector and a lock-in amplifier to build a prototype and validate the performance predicted by simulations.

Investigate and prototype a wavelength conversion system. This involves researching and procuring suitable nonlinear crystals for both the mid-IR and far-IR ranges. A proof-of-concept experiment should be conducted, initially converting a 3 µm to 6 µm source and subsequently extending the technique to the more challenging 20 µm to 30 µm range required for the 100 GeV case.

VI. Summary and Outlook

From 2020 to 2024, the EIC strong hadron cooling collaboration made significant progress in designing a high-energy cooler based on Coherent Electron Cooling (CeC). The design effort produced a comprehensive, which is essential for maintaining and improve high luminosity at the EIC. The key achievement includes:

- **Advanced Modeling and Simulation**: Sophisticated 1D, 2D, and full 3D models were developed to understand the cooling physics, including non-linear effects and diffusion. A long-term luminosity evolution code was also created, which incorporated effects like IBS and beam-beam interactions.

- **Comprehensive ERL Design**: A preliminary design for the ERL was established, specifying the injector, main linac, complex transport lines needed to deliver the high-quality electron beam, cooling section and return beamline. This included solutions such as a laser heater,uniform current generation, bunch compressive using splitter, wakes evaluation. High-fidelity tracking with the real number of electrons was used to verify amplifier performance and show that initial shot noise is not significantly amplified.

- **Critical R&D Initiatives**: R&D began on the most challenging components. This includes the development of a high-current, high-brightness HVDC electron gun and novel detectors for the sub-micrometer scale longitudinal alignment between the electron and hadron beams.

Meanwhile, we reviewed various hadron cooling options and planned for ERL testing and implantation. While the design studies produced a conceptual framework, we also realized the significant technical risks and performance limitations such as even with a fully functional SHC, the average luminosity would still fall short of the ultimate EIC requirement. Furthermore, critical challenges of beam stability, the high-current ERL operation, cooling measurement instruments could not be fully retired within the project's construction timescale. A successful proof-of-principle CeC experiment, such as the ongoing CeC-X effort at RHIC, will be necessary before committing to a large-scale cooling facility. Successful outcomes in these foundational R&D areas could pave the way for a future proposal to implement SHC as an upgrade to the EIC, unlocking its full luminosity potential.

---

[1] Y. S. Derbenev, On possibilities of fast cooling of heavy particle beams, AIP Conference Proceedings **253**, 103 (1992), https://pubs.aip.org/aip/acp/article-pdf/253/1/103/12003721/103_1_online.pdf.

[2] V. N. Litvinenko and Y. S. Derbenev, Coherent electron cooling, Phys. Rev. Lett. **102**, 114801 (2009).

[3] D. Ratner, Microbunched electron cooling for high-energy hadron beams, Phys. Rev. Lett. **111**, 084802 (2013).

[4] G. Stupakov, Cooling rate for microbunched electron cooling without amplification, Phys. Rev. Accel. Beams **21**, 114402 (2018).

[5] G. Stupakov and P. Baxevanis, Microbunched electron cooling with amplification cascades, Phys. Rev. Accel. Beams **22**, 034401 (2019).

[6] V. N. Litvinenko, Y. Jing, D. Kayran, P. Inacker, J. Ma, T. Miller, I. Petrushina, I. Pinayev, K. Shih, G. Wang, and Y. H. Wu, Plasma-cascade instability, Phys. Rev. Accel. Beams **24**, 014402 (2021).

[7] J. Jarvis, V. Lebedev, A. Romanov, D. Broemmelsiek, K. Carlson, S. Chattopadhyay, A. Dick, D. Edstrom, I. Lobach, S. Nagaitsev, H. Piekarz, P. Piot, J. Ruan, J. Santucci, G. Stancari, and A. Valishev, Experimental demonstration of optical stochastic cooling, Nature 2022 608:7922 **608**, 287 (2022).

[8] V. Lebedev, High energy cooling, Journal of Instrumentation **19** (06), P06015.

[9] D. Cline, A. Garren, H. Herr, F. E. Mills, C. Rubbia, A. Ruggiero, and D. Young, "High energy electron cooling to improve the luminosity and lifetime in colliding beam machines", in *IEEE Transaction on Nuclear Science*, Vol. NS-26 (1979).

[10] M. Gentner, R. Brinkmann, Y. Derbenev, D. Husmann, and C. Steier, "On the possibilities of electron cooling for HERA", Nucl. Instrum. Methods in Phys. Research A **424**, 277 (1999).

[11] 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).

[12] A. Zholents, L. Rebuffi, and X. Shi, Stochastic cooling of electrons and positrons with euv light, Phys. Rev. Accel. Beams **24**, 022803 (2021).

[13] A. Zholents and C. Hall, Isochronous Lattice Design Study for Cooling Insertion in the Electron Storage Ring, APS Light Source Note LS-376 (2024).

[14] S. Seletskiy, J. Kewisch, Y. Jing, A. Fedotov, D. Kayran, J. Unger, and G. Hoffstaetter, *Ring Electron Cooler for Electron Ion Collider*, Tech. Rep. (Brookhaven National Laboratory and Cornell University, 2025) eIC Tech-Note, to be submitted.

[15] V. Lebedev, S. Nagaitsev, A. Burov, V. Yakovlev, I. Gonin, I. Terechkine, A. Saini, and N. Solyak, Conceptual design report: a ring-based electron cooling system for the eic, Journal of Instrumentation **16** (01), T01003–T01003.

[16] I. Bazarov, Overview of energy recovery linacs, in *Proceedings of the 2005 Particle Accelerator Conference* (2005) pp. 382–386.

[17] E. Wang, High current polarized electron source for future erhic, AIP Conference Proceedings **1970**, 050008 (2018), https://pubs.aip.org/aip/acp/article-pdf/doi/10.1063/1.5040227/13296584/050008_1_online.pdf.

[18] I. S. K. Gardner, Ferrite dominated cavities, in *CERN - Rutherford Accelerator School: RF Engineering for Particle Accelerators* (1991) pp. 349–374.

[19] P. Baxevanis and G. Stupakov, Transverse dynamics considerations for microbunched electron cooling, Phys. Rev. Accel. Beams **22**, 081003 (2019).

[20] W. F. Bergan and G. Stupakov, Effects of transverse dependence of kicks in simulations of microbunched electron cooling, in *Proceedings of NAPAC2022* (JACoW Publishing, 2022).

[21] P. Baxevanis and G. Stupakov, Hadron beam evolution in microbunched electron cooling, Phys. Rev. Accel. Beams **23**, 111001 (2020).

[22] W. F. Bergan, Electron diffusion in microbunched electron cooling, Phys. Rev. Accel. Beams **27**, 084402 (2024).

[23] W. F. Bergan *et al.*, Coherent electron cooling physics for the eic, in *Proceedings of IPAC2024* (JACoW Publishing, 2024).
[24] P. Baxevanis, Three-dimensional analysis of microbunched electron cooling, Phys. Rev. Accel. Beams **28**, 031003 (2025).
[25] E. Zitzler, M. Laumanns, and L. Thiele, Spea2: improving the strength pareto evolutionary algorithm for multiobjective optimization, in *Proceedings of EUROGEN2001* (2001) p. 95–100.
[26] A. Blednykh *et al.*, *Advancements in Accelerator Physics and System Developments for the EIC*, Tech Note BNL-227548-2025-TECH (Brookhaven National Laboratory, 2025).
[27] S. Nagaitsev, Intrabeam scattering formulas for fast numerical evaluation, Phys. Rev. ST Accel. Beams **8**, 064403 (2005).
[28] J. D. Bjorken and S. K. Mtingwa, Intrabeam scattering, Part. Accel. **13**, 115 (1983).
[29] V. A. Lebedev, J. S. Hangst, N. Madsen, and A. Labrador, Single and multiple intrabeam scattering in a laser cooled beam, Nucl. Instrum. Methods Phys. Res., Sect. A **391**, 176 (1997).
[30] M. Furman and M. Zisman, Luminosity, in *Handbook of Accelerator Physics and Engineering*, edited by A. W. Chao, K. H. Mess, M. Tigner, and F. Zimmermann (World Scientific, 2013) 2nd ed., Chap. 4.1, pp. 311 – 318.
[31] E. Wang and M. Blaskiewicz, *Longitudinal space charge kick in Coherent Electron Cooling*, Tech. Rep. (Brookhaven National Lab. (BNL), Upton, NY (United States), 2020).
[32] S. Heifets, G. Stupakov, and S. Krinsky, Coherent synchrotron radiation instability in a bunch compressor, Phys. Rev. ST Accel. Beams **5**, 064401 (2002).
[33] A. Zholents and C. Hall, Wiggler-based electron beam microbunching amplifier study, Phys. Rev. Accel. Beams **27**, 081601 (2024).
[34] A. Albà, J. Seok, A. Adelmann, S. Doran, G. Ha, S. Lee, Y. Piao, J. Power, M. Qian, E. Wisniewski, J. Xu, and A. Zholents, Benchmarking collective effects of electron interactions in a wiggler with OPAL-FEL, Computer Physics Communications **280**, 108475 (2022).
[35] M. Conde, S. Antipov, D. Doran, W. Gai, Q. Gao, G. Ha, C. Jing, W. Liu, N. Neveu, J. Power, *et al.*, Research program and recent results at the argonne wakefield accelerator facility (AWA), Proc. IPAC'17 , 2885 (2017).
[36] G. Ha, J. G. Power, M. Conde, D. S. Doran, and W. Gai, Limiting effects in double EEX beamline, in *Journal of Physics: Conference Series*, Vol. 874 (IOP Publishing, 2017) p. 012061.
[37] A. Zholents and C. Hall, Wiggler-based microbunching amplifier in the case of a large wiggler parameter K, APS Light Source Note LS-377 (2024).
[38] S. Peggs, S. Benson, W. Bergan, D. Bruno, Y. Gao, D. Holmes, R. Lambiase, C. Liu, H. L. III, G. Mahler, T. Michalski, F. Micolon, V. Ptitsyn, G. Robert-Demolaize, R. Than, J. Tuozzolo, E. Wang, D. Weiss, and D. Xu, Optics for Strong Hadron Cooling in EIC HSR-IR2, in *Proc. IPAC'22*, International Particle Accelerator Conference No. 13 (JACoW Publishing, Geneva, Switzerland, 2022) pp. 1920–1923.
[39] D. X. et al., Beam optics update for eic hsr-ir2, in *Proc. IPAC'23*, IPAC'23 - 14th International Particle Accelerator Conference No. 14 (JACoW Publishing, Geneva, Switzerland, 2023) pp. 116–119.
[40] E. Wang, W. F. Bergan, F. J. Willeke, S. V. Benson, K. E. Deitrick, D. Douglas, C. M. Gulliford, C. E. Mayes, N. F. Taylor, and J. Qiang, Electron ion collider strong hadron cooling injector and erl (Linac, 2022).
[41] C. Gulliford, J. Conway, D. Douglas, B. Dunham, R. Eichhorn, V. Kostroun, C. Mayes, K. Smolenski, N. Taylor, E. Wang, *et al.*, Design and optimization of an erl for cooling eic hadron beams (Thomas Jefferson National Accelerator Facility (TJNAF), Newport News, VA (United States), 2023).
[42] K. Deitrick, I. Neththikumara, S. Setiniyaz, S. Benson, T. Satogata, A. Fedotov, D. Xu, D. Kayran, E. Wang, W. Bergan, *et al.*, Development of an erl for coherent electron cooling at the electron-ion collider (Thomas Jefferson National Accelerator Facility (TJNAF), Newport News, VA (United States); Brookhaven National Laboratory (BNL), Upton, NY (United States); Argonne National Laboratory (ANL), Argonne, IL (United States); SLAC National Accelerator Laboratory (SLAC), Menlo Park, CA (United States), 2024).
[43] K. Deitrick, T. Satogata, S. Setiniyaz, I. Neththikumara, J. Guo, S. Benson, C. Mayes, C. Gulliford, N. Taylor, and N. Sereno, A preliminary feasibility study on multi-cavity cryomodule integration for the electron ion collider energy recover linac cooler (Thomas Jefferson National Accelerator Facility (TJNAF), Newport News, VA (United States); SLAC National Accelerator Laboratory (SLAC), Menlo Park, CA (United States); Argonne National Laboratory (ANL), Argonne, IL (United States), 2024).
[44] E. Wang, J. Qiang, S. Benson, and W. Bergan, Generating Super-Gaussian distribution and uniform sliced energy spread bunch for EIC strong hadron cooling, JACoW **IPAC2024**, MOPC23 (2024).
[45] N. Wang, C. Mayes, C. Gulliford, D. Sagan, G. Hoffstaetter, E. Wang, W. Bergan, I. Neththikumara, K. Deitrick, S. Benson, *et al.*, Optimization of cooling distribution of the eic shc cooler erl (SLAC National Accelerator Laboratory (SLAC), Menlo Park, CA (United States); Brookhaven National Laboratory (BNL), Upton, NY (United States); Thomas Jefferson National Accelerator Facility (TJNAF), Newport News, VA (United States); Argonne National Lab., Idaho Falls, ID (United States), 2024).
[46] Z. Huang, A. Brachmann, F.-J. Decker, Y. Ding, D. Dowell, P. Emma, J. Frisch, S. Gilevich, G. Hays, P. Hering, R. Iverson, H. Loos, A. Miahnahri, H.-D. Nuhn, D. Ratner, G. Stupakov, J. Turner,

J. Welch, W. White, J. Wu, and D. Xiang, Measurements of the linac coherent light source laser heater and its impact on the x-ray free-electron laser performance, Phys. Rev. ST Accel. Beams **13**, 020703 (2010).

[47] I. Neththikumara, *Space Charge Studies on the ERL design for Strong Hadron Cooling*, Tech. Rep. (EIC-SHC-TN-25-004, 2025).

[48] D. Douglas, *Some Features of the FEL Upgrade* Π-*Bends*, Tech. Rep. (JLAB-TN-01-024, 2021).

[49] G. H. Hoffstaetter, D. Trbojevic, C. Mayes, N. Banerjee, J. Barley, I. Bazarov, A. Bartnik, J. S. Berg, S. Brooks, D. Burke, J. Crittenden, L. Cultrera, J. Dobbins, D. Douglas, B. Dunham, R. Eichhorn, S. Full, F. Furuta, C. Franck, R. Gallagher, M. Ge, C. Gulliford, B. Heltsley, D. Jusic, R. Kaplan, V. Kostroun, Y. Li, M. Liepe, C. Liu, W. Lou, G. Mahler, F. Meot, R. Michnoff, M. Minty, R. Patterson, S. Peggs, V. Ptitsyn, P. Quigley, T. Roser, D. Sabol, D. Sagan, J. Sears, C. Shore, E. Smith, K. Smolenski, P. Thieberger, S. Trabocchi, J. Tuozzolo, N. Tsoupas, V. Veshcherevich, D. Widger, G. Wang, F. Willeke, and W. Xu, Cbeta design report, cornell-bnl erl test accelerator (2017), arXiv:1706.04245 [physics.acc-ph].

[50] A. Bartnik, N. Banerjee, D. Burke, J. Crittenden, K. Deitrick, J. Dobbins, C. Gulliford, G. H. Hoffstaetter, Y. Li, W. Lou, P. Quigley, D. Sagan, K. Smolenski, J. S. Berg, S. Brooks, R. Hulsart, G. Mahler, F. Meot, R. Michnoff, S. Peggs, T. Roser, D. Trbojevic, N. Tsoupas, and T. Miyajima, Cbeta: First multipass superconducting linear accelerator with energy recovery, Phys. Rev. Lett. **125**, 044803 (2020).

[51] J. Qiang and E. Wang, Simulation of shot noise effects in the eic strong hadron cooling accelerator using real number of electrons, Journal of Physics: Conference Series **2687**, 062020 (2024).

[52] J. Qiang, W. Bergan, and E. Wang, Simulation of electron beam transport through the coherent electron cooling amplification section using real number of electrons, in *Proc. 16th Int. Particle Accelerator Conf. (IPAC'25)* (Taipei, Taiwan, 2025) p. WEPS015.

[53] D. Douglas, *Path Length Tolerance driven constraint on transverse residual orbit in amplifier transport*, Tech. Rep. (SHC notes, 2022) unpublished.

[54] D. Douglas, *Updates to simulations of orbit and excitation errors in cooler ERL recirculation transport*, Tech. Rep. (SHC notes, 2022) unpublished.

[55] D. Douglas, *Cooling channel error dispersion and compaction sensitivity to beamline and injection*, Tech. Rep. (SHC notes, 2023) unpublished.

[56] G4beamline release 3.06, Available online: `http://www.muonsinternal.com/muons3/G4beamline` (2019), accessed on 09 Apr 2021.

[57] S. Agostinelli, Geant4—A Simulation Toolkit, Nucl. Instruments Methods Phys. Res. A **506**, 250 (2003).

[58] Dassault Systèmes, Cst studio suite: Cst microwave studio (2023), accessed: 2025-01-01.

[59] I. Zagorodnov, Echo: Applications, Presentation slides, The Second Topical Workshop on Instabilities, Impedance, and Collective Effects, Abingdon, Oxfordshire, 8 February 2016 (2016), accessed: 2025-01-02.

[60] N. Nakamura, Resistive-wall impedance effects for the new kek light source, Journal of Physics: Conference Series **874**, 012069 (2017).

[61] B. Kosciuk, A. Blednykh, D. Padrazo, and O. Singh, Overview of diagnostics and instrumentation for national synchrotron light source ii, Diamond Light Source Proceedings **1**, e18 (2010).

[62] W. F. Bergan, M. Blaskiewicz, and G. Stupakov, Schottky signal modification as a diagnostic tool for coherent electron cooling, Phys. Rev. Accel. Beams **25**, 094401 (2022).

[63] G. Wang, *Estimates of the Recombination Rate for the Strong Hadron Cooling System in the EIC*, Tech. Rep. (Brookhaven National Laboratory (BNL), Upton, NY (United States), 2023).

[64] S. Setiniyaz, S. Benson, I. Neththikumara, K. Deitrick, C. Gulliford, J. Guo, C. Mayes, T. Satogata, N. Sereno, and N. Taylor, A preliminary feasibility study on multi-cavity cryomodule integration for the electron-ion collider energy recovery linac cooler, in *Proceedings of the 15th International Particle Accelerator Conference (IPAC'24)* (JACoW Publishing, 2024).

[65] E. Wang, V. N. Litvinenko, I. Pinayev, M. Gaowei, J. Skaritka, S. Belomestnykh, I. Ben-Zvi, J. C. Brutus, Y. Jing, J. Biswas, J. Ma, G. Narayan, I. Petrushina, O. Rahman, T. Xin, T. Rao, F. Severino, K. Shih, K. Smith, G. Wang, and Y. Wu, Long lifetime of bialkali photocathodes operating in high gradient superconducting radio frequency gun, Scientific Reports **11**, 4477 (2021).

[66] E. Wang, High average current HVDC electron gun for EIC hadron cooling , JACoW **ERL 2024**, WEO12 (2024).