Nuclear Science and Techniques (2026) 37:2
https://doi.org/10.1007/s41365-025-01855-5

Steady-state microbunching based on transverse-longitudinal coupling

Xiu-Jie Deng[1] · Alexander Wu Chao[1] · Wen-Hui Huang[2] · Zi-Zheng Li[2] · Zhi-Long Pan[2] · Chuan-Xiang Tang[2]

Received: 11 October 2024 / Revised: 13 February 2025 / Accepted: 23 February 2025 / Published online: 5 December 2025

Abstract
In this study, three specific scenarios of a novel accelerator light source mechanism called steady-state microbunching (SSMB) were studied: longitudinal weak focusing, longitudinal strong focusing, and generalized longitudinal strong focusing (GLSF). At present, GLSF is the most promising method for realizing high-power short-wavelength coherent radiation with mild requirements on modulation laser power. Its essence is to exploit the ultrasmall natural vertical emittance of an electron beam in a planar storage ring for efficient microbunching formation, like a partial transverse-longitudinal emittance exchange in the optical laser wavelength range. Based on an in-depth investigation of related beam physics, a solution for a GLSF SSMB storage ring that can deliver 1 kW average-power EUV light is presented. The work in this paper, such as the generalized Courant–Snyder formalism, analysis of theoretical minimum emittances, transverse-longitudinal coupling dynamics, and derivation of the bunching factor and modulation strengths for laser-induced microbunching schemes, is expected to be useful not only for the development of SSMB but also for future accelerator light sources in general that demand increasingly precise electron beam phase space manipulations.

Keywords Steady-state microbunching · Extreme ultraviolet · ARPES · Generalized Courant–Snyder formalism · Theoretical minimum emittances · Longitudinal weak focusing · Longitudinal strong focusing · Generalized longitudinal strong focusing · Transverse-longitudinal coupling · Damping wiggler

1 Introduction

Light source accelerator is arguably the most active driving force for accelerator development at the moment. Currently, there are two types of workhorses for these sources, storage ring-based synchrotron radiation sources and linear accelerator (linac)-based free-electron lasers (FELs). They deliver light with high repetition rate and high peak power or brilliance, respectively. We are attempting to develop a new storage ring-based light source mechanism called steady-state microbunching (SSMB) [1–19], which hopefully combines the advantages of these two sources and promises both high repetition and high-power radiation, realizing an accelerator-based fully coherent light source. The schematic layout of the SSMB storage ring and its comparison with the present synchrotron radiation source and FEL are shown in Fig. 1. In a conventional storage ring, the electron bunches are longitudinally focused by one or multiple radio-frequency (RF) cavities, whereas in an SSMB, such bunching systems are replaced by one or several optical laser modulation systems. The wavelength of the laser is six orders of magnitude shorter than that of RF. The bunch length or structure created by the laser is very short; thus, the term microbunching is used. When a beam becomes microbunched, it can radiate coherently and strongly, similar to a laser. However, note that in an SSMB, there is no exponential growth of the radiation power as that in a high-gain FEL [20] or conventional quantum lasers. In this context, the term laser mainly reflects that the radiation is coherent, both transversely and longitudinally. To ensure that the electron

This work was supported by the National Key Research and Development Program of China (No. 2022YFA1603401), National Natural Science Foundation of China (Nos. 12035010 and 12342501), Beijing Outstanding Young Scientist Program (No. JWZQ20240101006), and the Tsinghua University Dushi Program.

✉ Xiu-Jie Deng
dengxiujie@mail.tsinghua.edu.cn

[1] Institute for Advanced Study, Tsinghua University, Beijing 100084, China

[2] Department of Engineering Physics, Tsinghua University, Beijing 100084, China

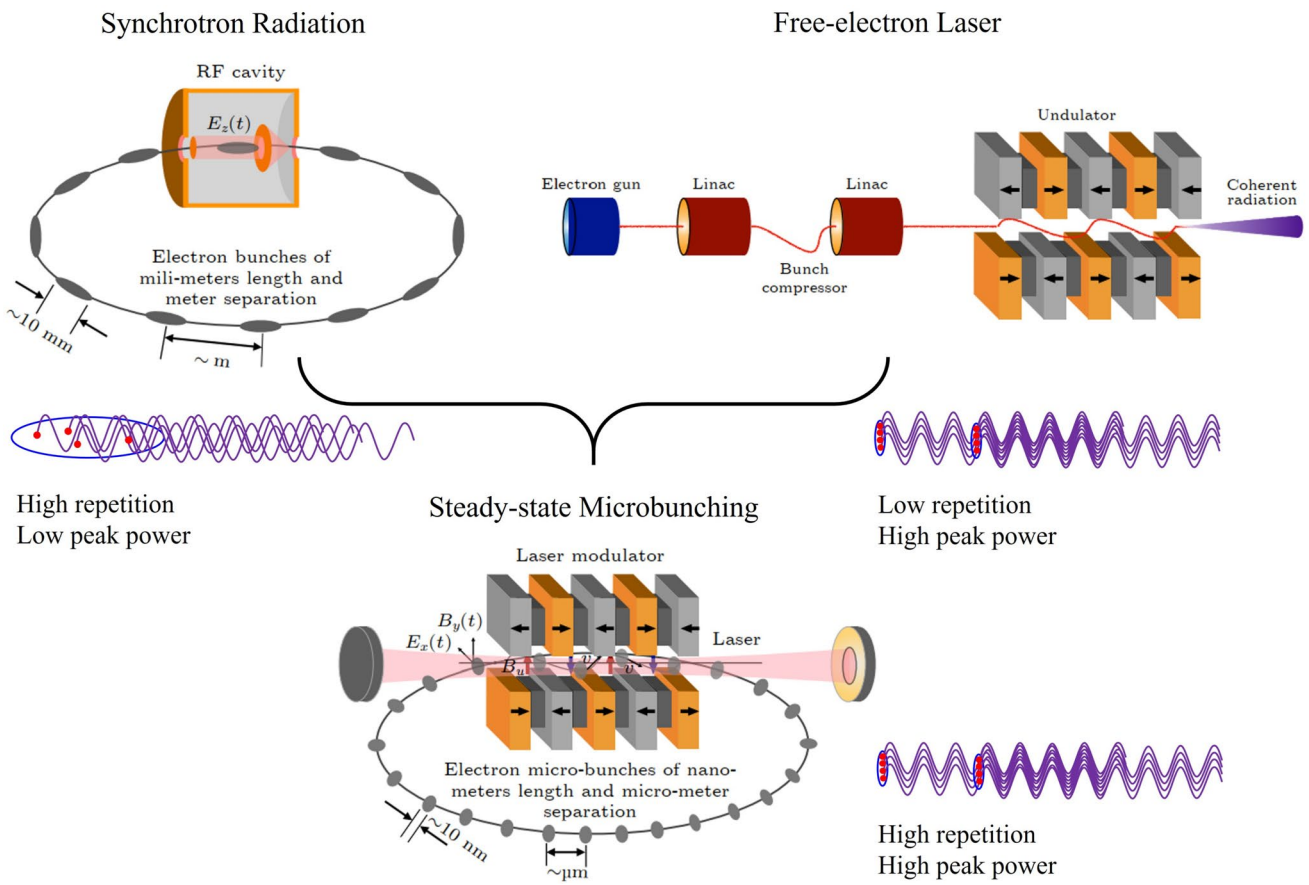

Fig. 1 (Color online) A schematic layout of an SSMB storage ring and comparison to a synchrotron radiation source and free-electron laser. (Figure adapted from Ref. [19])

beam properties can be preserved, the SSMB radiator length is comparatively short, typically at the meter level, and the peak current of the electron beam in the SSMB is also lower than that in a high-gain FEL. Therefore, the radiation back reaction on the electron beam is not violent and can be balanced by the radiation damping in the ring.

Once realized, such an SSMB ring can produce EUV radiation with greatly enhanced power and flux, allowing sub-meV energy resolution in angle-resolved photoemission spectroscopy (ARPES) and providing new opportunities for fundamental physics research, such as revealing key electronic structures in topological materials. A kilowatt (kW)-level EUV source based on this scheme is also promising for EUV lithography for high-volume chip manufacturing. Therefore, the reward for such an SSMB ring is significant. However, there are problems to be investigated and solved on the road of every new concept into reality. To generate coherent EUV radiation in a storage ring, the electron bunch length should reach the nanometer level, which is not a trivial task considering that the typical bunch length in the present electron storage rings is at the millimeter level. This study focused on our efforts to achieve this challenging goal.

The remainder of this paper is organized as follows: In Sect. 2, to build the foundation for the following analysis, we first introduce the generalized Courant–Snyder formalism which applies to a 3D general coupled lattice and present its application in electron storage ring physics. In Sect. 3, based on the formalism, we derive the theoretical minimum longitudinal emittance in an electron storage ring to provide the basis for later investigation, because the SSMB is about obtaining a short bunch length and small longitudinal emittance. Subsequently, in Sect. 4, with a motivation to realize nm-bunch length and high-average-power EUV radiation, we conducted a key analysis of three specific SSMB scenarios, which are, longitudinal weak focusing (LWF), longitudinal strong focusing (LSF), and generalized longitudinal strong focusing (GLSF). A brief summary of the three schemes is provided below. An LWF SSMB ring can be used to generate bunches with a bunch length of a couple of 10 nm; thus, it can be used to generate coherent visible and infrared radiation. From an engineering viewpoint, if we want to push the bunch length to an even shorter value, the required phase slippage factor of the LWF ring will be too small to be realized in practice. For comparison, an LSF SSMB ring can create bunches with a

bunch length at the nanometer level, thus generating coherent EUV radiation. However, the required modulation laser power is at the gigawatt (GW) level, which makes the laser modulator, typically consisting of an optical enhancement cavity with an incident laser and an undulator, used to longitudinally focus the electron beam at the laser wavelength scale, can only work at a low duty cycle pulsed mode, thus limiting the average output EUV radiation power. At present, a GLSF SSMB ring is the most promising among the three for obtaining a nm-bunch length with a mild-modulation laser power, thus allowing a high-average-power radiation output. The basic idea of the GLSF is to exploit the ultrasmall vertical emittance in a planar ring and apply partial transverse-longitudinal emittance for bunch compression with a shallow energy modulation strength, and thus a small modulation laser power. The backbone of such a GLSF ring is the transverse-longitudinal coupling (TLC) dynamics, which were analyzed in depth in this study. Following this analysis, before going into concrete examples, we prove three theorems in Sect. 5 for TLC-based bunch compression or harmonic generation schemes. Next, we discuss the details of the various TLC schemes in Sect. 6 devoted to energy modulation-based schemes, and Sect. 7 is dedicated to angular modulation-based schemes. We derived the bunching factors and required modulation laser powers. The conclusion from the analysis is that energy modulation-based coupling is favored for our application in the GLSF SSMB. Based on the investigations and other critical physical considerations, an example parameter set of a 1 kW average-power EUV light source is presented in Sect. 8. A short summary is given in Sect. 9.

2 Generalized Courant–Snyder formalism

In this section, we introduce a generalized Courant–Snyder formalism for storage ring physics to provide the basis for the following discussions. The particle state vector $\mathbf{X} = \left(x\ x'\ y\ y'\ z\ \delta\right)^{\mathrm{T}}$ is used throughout this paper, with its components meaning the horizontal position, horizontal angle, vertical position, vertical angle, longitudinal position, and $\delta = \Delta E/E_0$ the relative energy deviation of a particle with respect to the reference particle, respectively. E_0 denotes the energy of the reference particle. The superscript "T" denotes the transpose of a vector or matrix.

2.1 Generalized Beta functions in a general coupled Lattice

Inspired by Chao's solution by linear matrix (SLIM) formalism [21], we introduce the definition of the generalized beta functions in a 3D general coupled storage ring lattice as

$$\begin{aligned} &\beta_{ij}^{k} = 2\mathrm{Re}\left(\mathbf{E}_{ki}\mathbf{E}_{kj}^{*}\right), \\ &k = \pm I, II, III, \\ &i,j = 1,2,3,4,5,6, \end{aligned} \tag{1}$$

where * means the complex conjugate, the sub- or superscript k is the eigenmode index, Re() means the real component of a complex number or matrix, $\mathbf{E}_{ki}$ is the i-th component of the vector $\mathbf{E}_k$ which is the eigenvector of the 6×6 symplectic one-turn map $\mathbf{M}$ of storage ring with the eigenvalue $e^{i2\pi\nu_k}$

$$\mathbf{M}\mathbf{E}_k = e^{i2\pi\nu_k}\mathbf{E}_k, \tag{2}$$

satisfying the following normalization condition

$$\mathbf{E}_k^{\dagger}\mathbf{S}\mathbf{E}_k = \begin{cases} i, & k = I, II, III, \\ -i, & k = -I, -II, -III, \end{cases} \tag{3}$$

and

$$\mathbf{E}_k^{\dagger}\mathbf{S}\mathbf{E}_j = 0 \text{ for } k \neq j, \tag{4}$$

where i here means the imaginary unit, † means complex conjugate transpose, and

$$\mathbf{S} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix}. \tag{5}$$

This normalization condition can be preserved around the ring owing to the symplecticity of the transfer matrix. Since the one-turn map is a real symplectic matrix, for a stable motion, we have

$$\nu_{-k} = -\nu_k,\ \mathbf{E}_{-k} = \mathbf{E}_k^{*}, \tag{6}$$

where ν_k is the eigen-tune.

Similar to the real generalized beta function, here we define the imaginary generalized beta functions as

$$\hat{\beta}_{ij}^{k} = 2\mathrm{Im}\left(\mathbf{E}_{ki}\mathbf{E}_{kj}^{*}\right), \tag{7}$$

where Im() denotes the imaginary component of a complex number or matrix. Note that $\hat{\beta}_{ij}^{k}$ is a real value, similar to β_{ij}^{k}. The names 'real' and 'imaginary' here reflect that their definitions take the real and imaginary parts of $\mathbf{E}_{ki}\mathbf{E}_{kj}^{*}$, respectively. We use the symbol ^ on the top of these functions to indicate 'imaginary'. Further, we can define the real and imaginary generalized 6×6 Twiss matrices of a storage ring lattice corresponding to each eigenmode as

$$\left(\mathbf{T}_k\right)_{ij} = \beta_{ij}^k,\ \left(\hat{\mathbf{T}}_k\right)_{ij} = \hat{\beta}_{ij}^k. \tag{8}$$

From their definitions, we have

$$\begin{aligned}&\mathbf{T}_{-k} = \mathbf{T}_k,\ \mathbf{T}_k^* = \mathbf{T}_k,\ \mathbf{T}_k^{\mathrm{T}} = \mathbf{T}_k,\\ &\hat{\mathbf{T}}_{-k} = -\hat{\mathbf{T}}_k,\ \hat{\mathbf{T}}_k^* = \hat{\mathbf{T}}_k,\ \hat{\mathbf{T}}_k^{\mathrm{T}} = -\hat{\mathbf{T}}_k.\end{aligned} \tag{9}$$

Therefore, we know that $\mathbf{T}_k$ is a real symmetric matrix, whereas $\hat{\mathbf{T}}_k$ is a real anti-symmetric matrix. Further, we can prove that

$$\sum_{k=I,II,III} \hat{\mathbf{T}}_k = -\mathbf{S}. \tag{10}$$

The generalized Twiss matrices at different places around the ring are related to each other according to

$$\begin{aligned}&\mathbf{T}_k(s_2) = \mathbf{R}(s_2,s_1)\mathbf{T}_k(s_1)\mathbf{R}^{\mathrm{T}}(s_2,s_1),\\ &\hat{\mathbf{T}}_k(s_2) = \mathbf{R}(s_2,s_1)\hat{\mathbf{T}}_k(s_1)\mathbf{R}^{\mathrm{T}}(s_2,s_1).\end{aligned} \tag{11}$$

where $\mathbf{R}(s_2,s_1)$ is the symplectic transfer matrix of the state vector $\mathbf{X}$ from s_1 to s_2

$$\mathbf{X}(s_2) = \mathbf{R}(s_2,s_1)\mathbf{X}(s_1), \tag{12}$$

and we have

$$\mathbf{T}_k(s+C_0) = \mathbf{T}_k(s),\ \hat{\mathbf{T}}_k(s+C_0) = \hat{\mathbf{T}}_k(s), \tag{13}$$

where C_0 is the ring circumference.

With the help of the generalized Twiss matrices and eigentunes, the one-turn map $\mathbf{M}$ can be parametrized as

$$\begin{aligned}\mathbf{M} &= e^{\mathbf{S}\left(\sum_{k=I,II,III}\mathbf{G}_k\Phi_k\right)}\\ &= \left(\sum_{k=I,II,III}\left[\mathbf{T}_k\sin\Phi_k + \hat{\mathbf{T}}_k\cos\Phi_k\right]\right)\mathbf{S},\end{aligned} \tag{14}$$

where

$$\mathbf{G}_k \equiv \mathbf{S}^{\mathrm{T}}\mathbf{T}_k\mathbf{S}, \tag{15}$$

and $\Phi_k = 2\pi\nu_k$ is the phase advance of the corresponding mode in one turn. For $\mathbf{M}^n$ we only need to replace the above Φ_k with $n\Phi_k$. We can write the matrix terms of the one-turn map more explicitly as

$$\begin{aligned}\mathbf{M}_{ij} = (-1)^j \sum_{k=I,II,III}\Big[&\beta^k_{i(j-(-1)^j)}\sin\Phi_k\\ &+\hat{\beta}^k_{i(j-(-1)^j)}\cos\Phi_k\Big].\end{aligned} \tag{16}$$

Applying $\beta_{ij}^k = \beta_{ji}^k$, $\hat{\beta}_{ij}^k = -\hat{\beta}_{ji}^k$, we have

$$\begin{aligned}\mathbf{M}_{12} &= \sum_{k=I,II,III}\beta_{11}^k\sin\Phi_k,\\ \mathbf{M}_{21} &= -\sum_{k=I,II,III}\beta_{22}^k\sin\Phi_k,\\ \mathbf{M}_{34} &= \sum_{k=I,II,III}\beta_{33}^k\sin\Phi_k,\\ \mathbf{M}_{43} &= -\sum_{k=I,II,III}\beta_{44}^k\sin\Phi_k,\\ \mathbf{M}_{56} &= \sum_{k=I,II,III}\beta_{55}^k\sin\Phi_k,\\ \mathbf{M}_{65} &= -\sum_{k=I,II,III}\beta_{66}^k\sin\Phi_k.\end{aligned} \tag{17}$$

Using the generalized Twiss matrices, the actions or generalized Courant–Snyder invariants of a particle are defined according to

$$J_k \equiv \frac{\mathbf{X}^{\mathrm{T}}\mathbf{G}_k\mathbf{X}}{2}. \tag{18}$$

It is easy to prove that J_k are invariants of a particle when it travels around the ring, from Eqs. (11), (12), and the symplecticity of the transfer matrix

$$\mathbf{R}^{\mathrm{T}}\mathbf{S}\mathbf{R} = \mathbf{S}. \tag{19}$$

For a beam with N_{p} particles, the three beam invariants can then be defined according to

$$\epsilon_k \equiv \langle J_k\rangle = \frac{\sum_{i=1}^{N_{\mathrm{p}}} J_{k,i}}{N_{\mathrm{p}}},\ k = I, II, III, \tag{20}$$

where $J_{k,i}$ denotes the k-th mode invariant of the i-th particle. These invariants are the generalized root-mean-square (RMS) emittances of the beam in the ring.

The beam emittances defined above are based on the eigenmode motion of the particles in the storage ring. Another definition of emittance is based directly on the second moments matrix of a particle beam

$$\mathbf{\Sigma} = \langle\mathbf{X}\mathbf{X}^{\mathrm{T}}\rangle, \tag{21}$$

where $\langle\rangle$ denotes the particle ensemble average. The beam second moment matrix at different places are related according to

$$\mathbf{\Sigma}(s_2) = \mathbf{R}(s_2,s_1)\mathbf{\Sigma}(s_1)\mathbf{R}^{\mathrm{T}}(s_2,s_1). \tag{22}$$

From the symplecticity of $\mathbf{R}$, we can prove that the eigenvalues of $\mathbf{\Sigma}(-i\mathbf{S})$ remain unchanged with beam transport in a linear symplectic lattice. The beam eigen-emittances can thus be defined as the positive eigenvalues of $\mathbf{\Sigma}(-i\mathbf{S})$.

When the particle beam matches the storage ring lattice, which means the beam distribution at a given location repeats itself turn by turn, we have

$$\mathbf{\Sigma}(s+C_0)=\mathbf{M}(s)\mathbf{\Sigma}(s)\mathbf{M}^{\mathrm{T}}(s)=\mathbf{\Sigma}(s). \tag{23}$$

It can be proven that for a matched beam the RMS emittances defined in Eq. (20) are the eigenvalues of $\mathbf{\Sigma}(-i\mathbf{S})$ with the eigenvector $\mathbf{E}_k$, i.e.,

$$\mathbf{\Sigma}(-i\mathbf{S})\mathbf{E}_k=\mathrm{sgn}(k)\epsilon_k\mathbf{E}_k, \tag{24}$$

where we have used $\epsilon_{-k}\equiv\epsilon_k$ and

$$\mathrm{sgn}(k)=\begin{cases}1,\ k=I,II,III,\\ -1,\ k=-I,-II,-III.\end{cases} \tag{25}$$

Using the generalized Twiss and second moments matrices, the eigen-emittances for a matched beam can be calculated as

$$\epsilon_k=\mathrm{Tr}\left(\mathbf{T}_k\mathbf{S}\mathbf{\Sigma}\mathbf{S}^{\mathrm{T}}\right)=\mathrm{Tr}\left(\mathbf{G}_k\mathbf{\Sigma}\right). \tag{26}$$

where Tr() denotes the trace of the matrix.

To ensure that the eigenvectors $\mathbf{E}_k$ are uniquely defined all around the ring once they are determined at a given location, we will let

$$\mathbf{E}_k(s_2)=e^{-i\frac{s_2-s_1}{C_0}\Phi_k}\mathbf{R}(s_2,s_1)\mathbf{E}_k(s_1). \tag{27}$$

Following this definition, we have

$$\mathbf{E}_k(s+C_0)=e^{-i\Phi_k}\mathbf{M}(s)\mathbf{E}_k(s)=\mathbf{E}_k(s). \tag{28}$$

Using the generalized beta function, we can write the eigenvector component in an amplitude-phase form

$$\mathbf{E}_{kj}=\sqrt{\frac{\beta_{jj}^k}{2}}e^{i\phi_j^k}. \tag{29}$$

And according to the definition, we have

$$\beta_{ij}^k=\sqrt{\beta_{ii}^k\beta_{jj}^k}\cos(\phi_i^k-\phi_j^k). \tag{30}$$

Using the generalized Courant–Snyder invariants and beta functions, we can express the phase space coordinate of a particle at s as

$$\mathbf{X}_i(s)=\sum_{k=I,II,III}\sqrt{2J_k\beta_{ii}^k(s)}\cos\left[\psi_i^k(s)\right], \tag{31}$$

with ψ_i^k determined by the initial condition of the particle state. The phase term ψ_i^k at different locations are related according to

$$\psi_i^k(s_2)=\psi_i^k(s_1)+\frac{s_2-s_1}{C_0}\Phi_k+\phi_i^k(s_2)-\phi_i^k(s_1). \tag{32}$$

In particular, after n revolutions in the ring, we have

$$\mathbf{X}_i(s+nC_0)=\sum_{k=I,II,III}\sqrt{2J_k\beta_{ii}^k(s)}\cos\left[\psi_i^k(s)+n\Phi_k\right]. \tag{33}$$

2.2 Perturbations

After considering the parametrization of a general coupled lattice and the prescribed particle motion in it, let us now add perturbations, that from the lattice and also that from the beam. Assume there is a perturbation $\mathbf{K}$ to the one-turn map $\mathbf{M}$, i.e.,

$$\mathbf{M}_{\mathrm{per}}=(\mathbf{I}+\mathbf{K})\mathbf{M}_{\mathrm{unp}}, \tag{34}$$

where $\mathbf{I}$ is the identity matrix, the subscripts 'per' and 'unp' denote 'perturbed' and 'unperturbed', respectively. When the perturbation is small, from canonical perturbation theory, the tune shift of the k-th eigenmode can be calculated as

$$\Delta\nu_k=-\frac{1}{4\pi}\mathrm{Tr}\left[\left(\mathrm{sgn}(k)\mathbf{T}_k+i\hat{\mathbf{T}}_k\right)\mathbf{S}\mathbf{K}\right]. \tag{35}$$

This tune shift formula can be used to calculate the real and imaginary tune shifts due to symplectic (e.g., lattice error) and non-symplectic (e.g., radiation damping) perturbations.

For example, given the radiation damping matrix $\mathbf{D}$ around the ring, the damping rate of each eigenmode per turn is

$$\alpha_k=-\frac{1}{2}\oint\mathrm{Tr}\left(\hat{\mathbf{T}}_k\mathbf{S}\mathbf{D}\right)\mathrm{d}s,\ k=I,II,III, \tag{36}$$

where $\oint$ denotes integration around the ring. Note that the damping rates here are those for the corresponding eigenvectors. The damping rates for particle actions or beam eigen-emittances are a factor of two larger, because they are quadratic with respect to the phase space coordinates.

Apart from the radiation damping, there are also various beam diffusion effects in the ring, such as quantum excitation and intra-beam scattering. Using Eq. (26), once we know the diffusion matrix $\mathbf{N}$ around the ring, the emittance growth per turn due to diffusion is

$$\Delta\epsilon_k=\frac{1}{2}\oint\mathrm{Tr}\left(\mathbf{G}_k\mathbf{N}\right)\mathrm{d}s=-\frac{1}{2}\oint\mathrm{Tr}\left(\mathbf{T}_k\mathbf{S}\mathbf{N}\mathbf{S}\right)\mathrm{d}s. \tag{37}$$

The equilibrium eigen-emittance between a balance of diffusion and damping can be calculated as

$$\begin{aligned}\epsilon_k&=\frac{\Delta\epsilon_k}{2\alpha_k}=\frac{\frac{1}{2}\oint\mathrm{Tr}\left(\mathbf{G}_k\mathbf{N}\right)\mathrm{d}s}{-\oint\mathrm{Tr}\left(\hat{\mathbf{T}}_k\mathbf{S}\mathbf{D}\right)\mathrm{d}s}\\&=\frac{-\frac{1}{2}\sum_{i,j}\oint\beta_{ij}^k(\mathbf{S}\mathbf{N}\mathbf{S})_{ij}\mathrm{d}s}{\sum_{i,j}\oint\hat{\beta}_{ij}^k(\mathbf{S}\mathbf{D})_{ij}\mathrm{d}s}.\end{aligned} \tag{38}$$

It can be proven that the equilibrium beam distribution in the 6D phase space given by such a balance in a linear lattice is

Gaussian [22], with $\langle \mathbf{X} \rangle = \mathbf{0}$. After getting the equilibrium eigen-emittances, the second moments of beam can be written as

$$\Sigma_{ij} = \langle \mathbf{X}_i \mathbf{X}_j \rangle = \sum_{k=I,II,III} \epsilon_k \beta_{ij}^k, \tag{39}$$

or in matrix form as

$$\mathbf{\Sigma} = \sum_{k=I,II,III} \epsilon_k \mathbf{T}_k. \tag{40}$$

2.3 Application to electron storage rings

In an electron storage ring, intrinsic diffusion and damping are both caused by the emission of photons, namely quantum excitation and radiation damping. For quantum excitation, we have all the other components of the diffusion matrix $\mathbf{N}$ zero except that

$$N_{66} = \frac{\left\langle \dot{\mathcal{N}} \frac{u^2}{E_0^2} \right\rangle}{c} = \frac{2C_{\mathrm{L}}\gamma^5}{c|\rho|^3}, \tag{41}$$

where $\dot{\mathcal{N}}$ is the number of photons emitted per unit time, u is the photon energy, $C_{\mathrm{L}} = \frac{55}{48\sqrt{3}} \frac{r_{\mathrm{e}}\hbar}{m_{\mathrm{e}}}$, where r_{e} is the classical electron radius, $\hbar$ is the reduced Planck's constant, m_{e} is the electron mass, γ is the Lorentz factor, c is the speed of light in free space, and ρ is the bending radius. We take the convention that the sign of ρ is positive when the bending is inward and negative when the bending is outward. $\langle \dot{\mathcal{N}} u^2 \rangle$ in the above formula is a result of Campbell's theorem [23, 24].

For the damping effect, we have two sources: dipole magnets and RF cavity. For a horizontal dipole, we have all the other components of damping matrix $\mathbf{D}$ zero except that

$$D_{66} = -\frac{C_\gamma E_0^3}{\pi} \frac{1}{\rho^2}, \; D_{61} = -\frac{C_\gamma E_0^3}{2\pi} \frac{1-2n}{\rho^3}, \tag{42}$$

where $C_\gamma = \frac{4\pi}{3} \frac{r_{\mathrm{e}}}{(m_{\mathrm{e}}c^2)^3} = 8.85 \times 10^{-5} \frac{\mathrm{m}}{\mathrm{GeV}^2}$, $n = -\frac{\rho}{B_y} \frac{\partial B_y}{\partial x}$ is the transverse field gradient index. The physical origin of D_{66} is the fact that a higher energy particle tends to radiate more photon energy in a given magnetic field, while D_{61} is due to the fact that a transverse displacement of the particle will affect its path length in the dipole and when there is a transverse field gradient, the magnetic field strength observed, thus the radiation energy loss. For an RF cavity, we have all the other damping matrix terms of $\mathbf{D}$ zero except that

$$D_{22} = D_{44} = -\frac{eV_{\mathrm{RF}} \sin\phi_{\mathrm{RF}}}{E_0} \delta(s_{\mathrm{RF}}), \tag{43}$$

where e is the elementary charge assumed positive in this study, V_{RF} and ϕ_{RF} are the RF voltage and phase, respectively, and $\delta(s)$ is Dirac's delta function. Here, we assume that the RF cavity is zero length one. The physical origin of these damping terms is that the momentum boost of a particle in the RF cavity is along the longitudinal direction, whereas the transverse momenta of the particle are unchanged. Therefore, a damping effect was observed on the horizontal and vertical angles of the particles. We remind the readers that if we use the horizontal and vertical particle momentum p_x and p_y instead of x' and y', as the phase space coordinates, then the damping occurs only at bending magnets because the RF acceleration does not affect $p_{x,y}$.

In an electron storage ring, RF acceleration compensates for the radiation energy loss of electrons. If there are N cavities in the ring, we have

$$\sum_{i=1}^{N} eV_{\mathrm{RF},i} \sin\phi_{\mathrm{RF},i} = U_0, \tag{44}$$

where

$$U_0 = \frac{C_\gamma E_0^4}{2\pi} I_2 \tag{45}$$

is the radiation energy loss per particle per turn, with

$$I_2 = \oint \frac{1}{\rho^2} \mathrm{d}s. \tag{46}$$

If the ring consists of iso-bending magnets, then $U_0 = C_\gamma \frac{E_0^4}{\rho}$. From Eq. (42), we have

$$\oint D_{66}(s)\mathrm{d}s = -\frac{2U_0}{E_0}. \tag{47}$$

Similarly, based on Eqs. (43) and (44), we have

$$\oint D_{22}(s)\mathrm{d}s = \oint D_{44}(s)\mathrm{d}s = -\frac{U_0}{E_0}. \tag{48}$$

Combining with Eqs. (10) and (36), it is easy to show that for radiation damping, we have

$$\begin{aligned} \sum_{k=I,II,III} \alpha_k &= -\frac{1}{2} \oint \mathrm{Tr}\left[\left(\sum_{k=I,II,III} \hat{\mathbf{T}}_k \right) \mathbf{SD} \right] \mathrm{d}s \\ &= -\frac{1}{2} \oint \mathrm{Tr}(\mathbf{D}) \mathrm{d}s \\ &= -\frac{1}{2} \oint \left(D_{22} + D_{44} + D_{66} \right) \mathrm{d}s \\ &= \frac{2U_0}{E_0}. \end{aligned} \tag{49}$$

This is the well-known Robinson's sum rule [25].

The above formulation applies to a 3D general coupled lattice. For a ring without x-y coupling and when the

RF cavity is placed at dispersion-free location, we can express the normalized eigenvectors using the classical Courant–Snyder functions [26] α, β, γ and dispersion D and dispersion angle D' as

$$\mathbf{E}_I = \frac{1}{\sqrt{2}}\begin{pmatrix} \sqrt{\beta_x} \\ \frac{-\alpha_x+i}{\sqrt{\beta_x}} \\ 0 \\ 0 \\ \frac{-(\alpha_x D_x+\beta_x D'_x)+iD_x}{\sqrt{\beta_x}} \\ 0 \end{pmatrix} e^{i\Psi_I},$$

$$\mathbf{E}_{II} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 0 \\ \sqrt{\beta_y} \\ \frac{-\alpha_y+i}{\sqrt{\beta_y}} \\ \frac{-(\alpha_y D_y+\beta_y D'_y)+iD_y}{\sqrt{\beta_y}} \\ 0 \end{pmatrix} e^{i\Psi_{II}}, \tag{50}$$

$$\mathbf{E}_{III} = \frac{1}{\sqrt{2}}\begin{pmatrix} \frac{-\alpha_z+i}{\sqrt{\beta_z}}D_x \\ \frac{-\alpha_z+i}{\sqrt{\beta_z}}D'_x \\ \frac{-\alpha_z+i}{\sqrt{\beta_z}}D_y \\ \frac{-\alpha_z+i}{\sqrt{\beta_z}}D'_y \\ \sqrt{\beta_z} \\ \frac{-\alpha_z+i}{\sqrt{\beta_z}} \end{pmatrix} e^{i\Psi_{III}},$$

where the subscripts x, y, and z correspond to the horizontal, vertical, and longitudinal dimensions, respectively, and $\Psi_{I,II,III}$ are the phase factors of the eigenvectors. There is flexibility in choosing these phase factors as they do not affect the calculation of our defined Twiss matrices and physical quantities. However, note that once they are set at one location, their values around the ring are determined according to Eq. (27). In this case, the real and imaginary generalized Twiss matrices can be obtained explicitly

$$\mathbf{T}_I = \begin{pmatrix} \beta_x & -\alpha_x & 0\ 0 & -\alpha_x D_x - \beta_x D'_x & 0 \\ -\alpha_x & \gamma_x & 0\ 0 & \gamma_x D_x + \alpha_x D'_x & 0 \\ 0 & 0 & 0\ 0 & 0 & 0 \\ 0 & 0 & 0\ 0 & 0 & 0 \\ -\alpha_x D_x - \beta_x D'_x & \gamma_x D_x + \alpha_x D'_x & 0\ 0 & \mathcal{H}_x & 0 \\ 0 & 0 & 0\ 0 & 0 & 0 \end{pmatrix},$$

$$\hat{\mathbf{T}}_I = \begin{pmatrix} 0 & -1 & 0\ 0 & -D_x & 0 \\ 1 & 0 & 0\ 0 & -D'_x & 0 \\ 0 & 0 & 0\ 0 & 0 & 0 \\ 0 & 0 & 0\ 0 & 0 & 0 \\ D_x & D'_x & 0\ 0 & 0 & 0 \\ 0 & 0 & 0\ 0 & 0 & 0 \end{pmatrix},$$

$$\mathbf{T}_{II} = \begin{pmatrix} 0\ 0 & 0 & 0 & 0 & 0 \\ 0\ 0 & 0 & 0 & 0 & 0 \\ 0\ 0 & \beta_y & -\alpha_y & -\alpha_y D_y - \beta_y D'_y & 0 \\ 0\ 0 & -\alpha_y & \gamma_y & \gamma_y D_y + \alpha_y D'_y & 0 \\ 0\ 0 & -\alpha_y D_y - \beta_y D'_y & \gamma_y D_y + \alpha_y D'_y & \mathcal{H}_y & 0 \\ 0\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\hat{\mathbf{T}}_{II} = \begin{pmatrix} 0\ 0\ 0 & 0 & 0 & 0 & 0 \\ 0\ 0\ 0 & 0 & 0 & 0 & 0 \\ 0\ 0\ 0 & -1 & -D_y & 0 \\ 0\ 0\ 1 & 0 & -D'_y & 0 \\ 0\ 0\ D_y & D'_y & 0 & 0 \\ 0\ 0\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\mathbf{T}_{III} = \begin{pmatrix} \gamma_z D_x^2 & \gamma_z D_x D'_x & \gamma_z D_x D_y & \gamma_z D_x D'_y & -\alpha_z D_x & \gamma_z D_x \\ \gamma_z D_x D'_x & \gamma_z D'^2_x & \gamma_z D'_x D_y & \gamma_z D'_x D'_y & -\alpha_z D'_x & \gamma_z D'_x \\ \gamma_z D_x D_y & \gamma_z D'_x D_y & \gamma_z D_y^2 & \gamma_z D_y D'_y & -\alpha_z D_y & \gamma_z D_y \\ \gamma_z D_x D'_y & \gamma_z D'_x D'_y & \gamma_z D_y D'_y & \gamma_z D'^2_y & -\alpha_z D'_y & \gamma_z D'_y \\ -\alpha_z D_x & -\alpha_z D'_x & -\alpha_z D_y & -\alpha_z D'_y & \beta_z & -\alpha_z \\ \gamma_z D_x & \gamma_z D'_x & \gamma_z D_y & \gamma_z D'_y & -\alpha_z & \gamma_z \end{pmatrix},$$

$$\hat{\mathbf{T}}_{III} = \begin{pmatrix} 0 & 0 & 0 & 0 & D_x & 0 \\ 0 & 0 & 0 & 0 & D'_x & 0 \\ 0 & 0 & 0 & 0 & D_y & 0 \\ 0 & 0 & 0 & 0 & D'_y & 0 \\ -D_x & -D'_x & -D_y & -D'_y & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}. \tag{51}$$

Correspondingly, the generalized Courant–Snyder invariants are given by

$$\begin{aligned} J_I &\equiv \frac{(\mathbf{SX})^T \mathbf{T}_I \mathbf{SX}}{2} = \frac{\left(x - D_x\delta\right)^2 + \left[\alpha_x\left(x - D_x\delta\right) + \beta_x\left(x' - D'_x\delta\right)\right]^2}{2\beta_x}, \\ J_{II} &\equiv \frac{(\mathbf{SX})^T \mathbf{T}_{II} \mathbf{SX}}{2} = \frac{\left(y - D_y\delta\right)^2 + \left[\alpha_y\left(y - D_y\delta\right) + \beta_y\left(y' - D'_y\delta\right)\right]^2}{2\beta_y}, \\ J_{III} &\equiv \frac{(\mathbf{SX})^T \mathbf{T}_{III} \mathbf{SX}}{2} = \frac{\left(z + D'_x x - D_x x' + D'_y y - D_y y'\right)^2 + \left[\alpha_z\left(z + D'_x x - D_x x' + D'_y y - D_y y'\right) + \beta_z\delta\right]^2}{2\beta_z}. \end{aligned} \tag{52}$$

The equilibrium emittances determined by the balance of quantum excitation and radiation damping in an electron storage in this case reduce to the classical Sands radiation integrals formalism found in textbooks [27]

$$\begin{aligned}
\epsilon_x &\equiv \langle J_I \rangle = \frac{C_L \gamma^5}{2c\alpha_I} \oint \frac{\beta_{55}^I}{|\rho|^3} \mathrm{d}s = C_q \frac{\gamma^2}{J_x} \frac{I_{5x}}{I_2}, \\
\epsilon_y &\equiv \langle J_{II} \rangle = \frac{C_L \gamma^5}{2c\alpha_{II}} \oint \frac{\beta_{55}^{II}}{|\rho|^3} \mathrm{d}s = C_q \frac{\gamma^2}{J_y} \frac{I_{5y}}{I_2}, \\
\epsilon_z &\equiv \langle J_{III} \rangle = \frac{C_L \gamma^5}{2c\alpha_{III}} \oint \frac{\beta_{55}^{III}}{|\rho|^3} \mathrm{d}s = C_q \frac{\gamma^2}{J_z} \frac{I_{5z}}{I_2},
\end{aligned} \tag{53}$$

with $C_\mathrm{q} = \frac{55\lambda_\mathrm{e}}{32\sqrt{3}} = 3.8319 \times 10^{-13}\ m$ for electrons, $\lambda_\mathrm{e} = \frac{\lambda_\mathrm{e}}{2\pi} = \frac{hc}{m_\mathrm{e}c^2} = 386\ fm$ is the reduced Compton wavelength of electron, and the radiation integrals are given by

$$\begin{aligned}
&I_2 = \oint \frac{1}{\rho^2} \mathrm{d}s,\ I_{4x} = \oint D_x \left(\frac{1-2n}{\rho^3} \right) \mathrm{d}s, \\
&I_{5x} = \oint \frac{\mathcal{H}_x}{|\rho|^3} \mathrm{d}s,\ I_{5y} = \oint \frac{\mathcal{H}_y}{|\rho|^3} \mathrm{d}s,\ I_{5z} = \oint \frac{\beta_z}{|\rho|^3} \mathrm{d}s,
\end{aligned} \tag{54}$$

where

$$\mathcal{H}_{x,y} = \beta_{55}^{I,II} = \frac{D_{x,y}^2 + \left(\alpha_{x,y} D_{x,y} + \beta_{x,y} D'_{x,y} \right)^2}{\beta_{x,y}} \tag{55}$$

are the horizontal and vertical dispersion invariants, and $\beta_z = \beta_{55}^{III}$ is the longitudinal beta function [8]. The damping rate of three eigenmodes is given by

$$\alpha_I = \frac{U_0}{2E_0} J_x,\ \alpha_{II} = \frac{U_0}{2E_0} J_y,\ \alpha_{III} = \frac{U_0}{2E_0} J_z, \tag{56}$$

where $J_{x,y,z}$ are the damping partition numbers

$$J_x = 1 - \frac{I_{4x}}{I_2},\ J_y = 1,\ J_z = 2 + \frac{I_{4x}}{I_2}. \tag{57}$$

The Robinson's sum rule corresponds then to

$$J_x + J_y + J_z = 4. \tag{58}$$

The horizontal, vertical, and longitudinal radiation damping times are

$$\tau_{x,y,z} = \frac{T_0}{\alpha_{x,y,z}}, \tag{59}$$

where T_0 is the particle revolution period in the ring.

After obtaining the equilibrium emittance, the beam second moments can be obtained by substituting the generalized Twiss matrices from Eq. (51) into Eq. (39). For example,

$$\begin{aligned}
\sigma_x^2 &= \langle x^2 \rangle = \epsilon_x \beta_x + \epsilon_z \gamma_z D_x^2 = \epsilon_x \beta_x + \sigma_\delta^2 D_x^2, \\
\sigma_z^2 &= \langle z^2 \rangle = \epsilon_z \beta_z + \epsilon_x \mathcal{H}_x + \epsilon_y \mathcal{H}_y.
\end{aligned} \tag{60}$$

We remind the readers that the generalized Courant–Snyder formalism presented in this section has been briefly reported in Ref. [28].

3 Theoretical minimum emittances

After introducing the generalized Courant–Snyder formalism, we applied it to analyze the theoretical minimum longitudinal emittance in an electron storage ring. This work serves as the basis for the following investigation of SSMB, since the longitudinal weak focusing and strong focusing SSMB to be introduced soon are about lowering the equilibrium bunch length and longitudinal emittance in an electron storage ring. For completeness, in this section, we also present an analysis of the theoretical minimum transverse emittance, as it can be treated within a single framework.

3.1 Theoretical minimum horizontal emittance

In Eq. (53), we can see that $\mathcal{H}_x$ and β_z, i.e., β_{55}^I and β_{55}^{III} as defined by us, at the bending magnets are of vital importance in determining the horizontal and longitudinal emittance, respectively. Therefore, it is necessary to understand how they evolve inside a bending magnet. The transfer matrix of the particle state vector for a sector bending magnet is

$$\mathbf{B}(\alpha) = \begin{pmatrix}
\cos\alpha & \rho\sin\alpha & 0 & 0 & 0 & \rho(1-\cos\alpha) \\
-\frac{\sin\alpha}{\rho} & \cos\alpha & 0 & 0 & 0 & \sin\alpha \\
0 & 0 & 1 & \rho\alpha & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
-\sin\alpha & -\rho(1-\cos\alpha) & 0 & 0 & 1 & \rho\left(\frac{\alpha}{\gamma^2} - \alpha + \sin\alpha\right) \\
0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}, \tag{61}$$

with ρ and α being the bending radius and angle of the bending magnet, respectively. In a planar uncoupled ring, the normalized eigenvectors of the one-turn map are expressed as Eq. (50). Assuming the Courant–Snyder functions, dispersion and dispersion angle at the dipole center where we set to be $\alpha = 0$ are given by $\alpha_{x0}, \beta_{x0}, \alpha_{z0}, \beta_{z0}, D_{x0}, D'_{x0}$. From Eqs. (1), (61) and (50) we have the evolution of $\mathcal{H}_x$ in the dipole

$$\begin{aligned}\mathcal{H}_x(\alpha) \equiv \beta^I_{55}(\alpha) &= 2|\mathbf{E}_{I5}(\alpha)|^2 = 2|\left(\mathbf{B}(\alpha)\mathbf{E}_I(0)\right)_5|^2 \\ &= \left(\sqrt{\beta_{x0}}(\sin\alpha + D'_{x0}) + \frac{\alpha_{x0}}{\sqrt{\beta_{x0}}}\left[D_{x0} - \rho(1-\cos\alpha)\right]\right)^2 \\ &\quad + \left(\frac{D_{x0} - \rho(1-\cos\alpha)}{\sqrt{\beta_{x0}}}\right)^2.\end{aligned} \tag{62}$$

Similarly, the evolution of β_z in the dipole is given by

$$\begin{aligned}\beta_z(\alpha) \equiv \beta^{III}_{55}(\alpha) &= 2|\mathbf{E}_{III5}(\alpha)|^2 = 2|\left(\mathbf{B}(\alpha)\mathbf{E}_{III}(0)\right)_5|^2 \\ &= \left(\sin\alpha\frac{\alpha_{z0}}{\sqrt{\beta_{z0}}}D_{x0} + \rho(1-\cos\alpha)\frac{\alpha_{z0}}{\sqrt{\beta_{z0}}}D'_{x0} + \sqrt{\beta_{z0}} - \rho(-\alpha+\sin\alpha)\frac{\alpha_{z0}}{\sqrt{\beta_{z0}}}\right)^2 \\ &\quad + \left(-\sin\alpha\frac{1}{\sqrt{\beta_{z0}}}D_{x0} - \rho(1-\cos\alpha)\frac{1}{\sqrt{\beta_{z0}}}D'_{x0} + \rho(-\alpha+\sin\alpha)\frac{1}{\sqrt{\beta_{z0}}}\right)^2.\end{aligned} \tag{63}$$

For simplicity, we have neglected the contribution of $\frac{\rho\alpha}{\gamma^2}$ to R_{56} of dipole in the above calculation of β_z, since we are interested in the relativistic cases. However, we remind the readers that in a quasi-isochronous ring, the contribution of $\frac{C_0}{\gamma^2}$ to the ring R_{56} or phase slippage may not be negligible.

With the evolution of $\mathcal{H}_x$ and β_z known, we can now derive the theoretical minimum emittances. For simplicity, we assume the ring consists of iso-bending magnets, with the bending angle induced by each bending magnet being θ, and the optical functions are identical in each bending magnets. Then from Eq. (53), we have

$$\epsilon_x = C_q\frac{\gamma^2}{J_x}\frac{1}{\rho}f_x(\alpha_{x0}, \beta_{x0}, D_{x0}, D'_{x0}), \tag{64}$$

with

$$f_x(\alpha_{x0}, \beta_{x0}, D_{x0}, D'_{x0}) = \frac{1}{\theta}\int_{-\frac{\theta}{2}}^{\frac{\theta}{2}}\mathcal{H}_x(\alpha)\mathrm{d}\alpha. \tag{65}$$

f_x can be interpreted as the average value of $\mathcal{H}_x$ in the dipoles. The lengthy explicit expression of $f_x(\alpha_{x0}, \beta_{x0}, D_{x0}, D'_{x0})$ is omitted here. It can be obtained straightforwardly by inserting Eq. (62) in Eq. (65). The mathematical problem we are trying to solve is to minimize f_x by adjusting α_{x0}, β_{x0}, D_{x0}, D'_{x0}. From Eq. (65) we have

$$\begin{aligned}\frac{\partial f_x}{\partial\alpha_{x0}} &= \alpha_{x0}\frac{2D_{x0}\left(D_{x0} - 2\rho + 2\rho\frac{\sin\frac{\theta}{2}}{\frac{\theta}{2}}\right) + \rho^2\left(3 - 4\frac{\sin\frac{\theta}{2}}{\frac{\theta}{2}} + \frac{\sin\theta}{\theta}\right)}{\beta_{x0}} \\ &\quad + D'_{x0}2\left(D_{x0} - \rho + \rho\frac{\sin\frac{\theta}{2}}{\frac{\theta}{2}}\right), \\ \frac{\partial f_x}{\partial D'_{x0}} &= \alpha_{x0}2\left(D_{x0} - \rho + \rho\frac{\sin\frac{\theta}{2}}{\frac{\theta}{2}}\right) + D'_{x0}2\beta_{x0}.\end{aligned} \tag{66}$$

We notice that the requirement of $\frac{\partial f_x}{\partial\alpha_{x0}} = 0$ and $\frac{\partial f_x}{\partial D'_{x0}} = 0$ leads to $\alpha_{x0} = 0$ and $D'_{x0} = 0$. Under the above conditions, then

$$\frac{\partial f_x}{\partial D_{x0}} = \frac{2\left(D_{x0} - \rho + \rho\frac{\sin\frac{\theta}{2}}{\frac{\theta}{2}}\right)}{\beta_{x0}}. \tag{67}$$

The requirement of $\frac{\partial f_x}{\partial D_{x0}} = 0$ leads to

$$D_{x0} = \rho\left(1 - \frac{\sin\frac{\theta}{2}}{\frac{\theta}{2}}\right) \approx \frac{\rho\theta^2}{24}. \tag{68}$$

Under the above conditions, then

$$\frac{\partial f_x}{\partial\beta_{x0}} = \frac{\theta(\theta - \sin\theta) - \frac{\rho^2\left(\theta^2 + \theta\sin\theta + 4\cos\theta - 4\right)}{\beta^2_{x0}}}{2\theta^2}. \tag{69}$$

The requirement of $\frac{\partial f_x}{\partial\beta_{x0}} = 0$ leads to

$$\beta_{x0} = \rho\sqrt{\frac{\left(\theta^2 + \theta\sin\theta + 4\cos\theta - 4\right)}{\theta(\theta - \sin\theta)}} \approx \frac{\rho\theta}{2\sqrt{15}}. \tag{70}$$

Summarizing, the extreme value of f_x is realized when

$$\alpha_{x0}=0,\ \beta_{x0}\approx\frac{\rho\theta}{2\sqrt{15}},\ D_{x0}\approx\frac{\rho\theta^2}{24},\ D'_{x0}=0, \tag{71}$$

which means

$$\mathcal{H}_{x0}=\rho\frac{\left(1-\frac{\sin\frac{\theta}{2}}{\frac{\theta}{2}}\right)^2}{\sqrt{\frac{(\theta^2+\theta\sin\theta+4\cos\theta-4)}{\theta(\theta-\sin\theta)}}}\approx\frac{5\rho\theta^3}{96\sqrt{15}}, \tag{72}$$

and

$$\begin{aligned}f_{x,\min}&=\rho\left(1-\frac{\sin\theta}{\theta}\right)\sqrt{\frac{\left(\theta^2+\theta\sin\theta+4\cos\theta-4\right)}{\theta(\theta-\sin\theta)}}\\&\approx\frac{\rho\theta^3}{12\sqrt{15}}.\end{aligned} \tag{73}$$

One can verify that this is the minimum value of f_x. Note that $f_{x,\min}=\frac{8}{5}\mathcal{H}_{x0}$. Under these conditions, we get the minimum horizontal emittance

$$\epsilon_{x,\min}=C_{\mathrm{q}}\frac{\gamma^2}{J_x}\frac{\theta^3}{12\sqrt{15}}. \tag{74}$$

These results are consistent with the classical results of Teng [29]. For practical use, and considering nominally $J_x\approx 1$, the above scaling can be written as

$$\epsilon_{x,\min}[\mathrm{nm}]=31.6E_0^2[\mathrm{GeV}]\theta^3[\mathrm{rad}]. \tag{75}$$

For example, if $E_0=6$ GeV and $\theta=\frac{2\pi}{300}$ rad, we have $\epsilon_{x,\min}=10.4$ pm.

3.2 Theoretical minimum longitudinal emittance

We now analyze the theoretical minimum longitudinal emittance. Similar to the horizontal direction, we have

$$\epsilon_z=C_{\mathrm{q}}\frac{\gamma^2}{J_z}\frac{1}{\rho}f_z(\alpha_{z0},\beta_{z0},D_{x0},D'_{x0}), \tag{76}$$

with

$$f_z(\alpha_{z0},\beta_{x0},D_{x0},D'_{x0})=\frac{1}{\theta}\int_{-\frac{\theta}{2}}^{\frac{\theta}{2}}\beta_z(\alpha)\mathrm{d}\alpha, \tag{77}$$

which can be interpreted as the average value of β_z in dipoles. Then

$$\begin{aligned}\frac{\partial f_z}{\partial\alpha_{z0}}&=\alpha_{z0}\mathcal{G}(\rho,\theta,D_{x0},D'_{x0},\beta_{z0})+D'_{x0}2\rho\left(1-\frac{\sin\frac{\theta}{2}}{\frac{\theta}{2}}\right),\\\frac{\partial f_z}{\partial D'_{x0}}&=\alpha_{z0}2\rho\left(1-\frac{\sin\frac{\theta}{2}}{\frac{\theta}{2}}\right)\\&\quad+D'_{x0}\frac{\left(1+\alpha_{z0}^2\right)\rho^2\left(3-4\frac{\sin\frac{\theta}{2}}{\frac{\theta}{2}}+\frac{\sin\theta}{\theta}\right)}{\beta_{z0}},\end{aligned} \tag{78}$$

where the lengthy explicit expression of $\mathcal{G}(\rho,\theta,D_{x0},D'_{x0},\beta_{z0})$ is omitted here. Similar to the analysis of transverse minimum emittance, we notice again that the requirement of $\frac{\partial f_z}{\partial\alpha_{z0}}=0$ and $\frac{\partial f_z}{\partial D'_{x0}}=0$ leads to $\alpha_{z0}=0$ and $D'_{x0}=0$. Under the above conditions, then

$$\frac{\partial f_z}{\partial D_{x0}}=\frac{D_{x0}\left(1-\frac{\sin\theta}{\theta}\right)-\rho\left(1-\frac{\sin\theta}{\theta}-2\frac{\sin\frac{\theta}{2}}{\frac{\theta}{2}}+2\cos\frac{\theta}{2}\right)}{\beta_{z0}}. \tag{79}$$

The requirement of $\frac{\partial f_z}{\partial D_{x0}}=0$ leads to

$$D_{x0}=\frac{\rho\left(1-\frac{\sin\theta}{\theta}-2\frac{\sin\frac{\theta}{2}}{\frac{\theta}{2}}+2\cos\frac{\theta}{2}\right)}{1-\frac{\sin\theta}{\theta}}\approx-\frac{\rho\theta^2}{40}. \tag{80}$$

Under the above conditions, then

$$\frac{\partial f_z}{\partial\beta_{z0}}=1-\frac{\rho^2\left(\theta^4-12\theta^2-\left(\theta^2-48\right)\theta\sin\theta-12\left(\theta^2-4\right)\cos\theta-48\right)}{12\theta(\theta-\sin\theta)\beta_{z0}^2}. \tag{81}$$

The requirement of $\frac{\partial f_z}{\partial\beta_{z0}}=0$ leads to

$$\beta_{z0}=\rho\sqrt{\frac{\left(\theta^4-12\theta^2-\left(\theta^2-48\right)\theta\sin\theta-12\left(\theta^2-4\right)\cos\theta-48\right)}{12\theta(\theta-\sin\theta)}}\approx\frac{\rho\theta^3}{120\sqrt{7}}. \tag{82}$$

Summarizing, the extreme value of f_z is realized when

$$\alpha_{z0}=0,\ \beta_{z0}\approx\frac{\rho\theta^3}{120\sqrt{7}},\ D_{x0}\approx-\frac{\rho\theta^2}{40},\ D'_{x0}=0, \tag{83}$$

and

$$f_{z,\min}=\rho\sqrt{\frac{\left(\theta^4-12\theta^2-\left(\theta^2-48\right)\theta\sin\theta-12\left(\theta^2-4\right)\cos\theta-48\right)}{3\theta(\theta-\sin\theta)}}\approx\frac{\rho\theta^3}{60\sqrt{7}}. \tag{84}$$

One can verify that this is the minimum value of f_z. Note that $f_{z,\min}=2\beta_{z0}$. Under these conditions, we get the minimum longitudinal emittance

$$\epsilon_{z,\min}=C_{\mathrm{q}}\frac{\gamma^2}{J_z}\frac{\theta^3}{60\sqrt{7}}. \tag{85}$$

The above result was also reported in Ref. [7]. For practical use, and considering nominally $J_z\approx 2$, the above scaling can be written as

$$\epsilon_{z,\min}[\mathrm{nm}]=4.62E_0^2[\mathrm{GeV}]\theta^3[\mathrm{rad}]. \tag{86}$$

For example, if $E_0=600$ MeV and $\theta=\frac{2\pi}{50}$ rad, we have $\epsilon_{z,\min}=3.3$ pm.

Here, we remind the readers that in reality, it may not be easy to reach the optimal conditions Eq. (83) for all dipoles in a ring. This is based on the observation that when we realize Eq. (83) at the dipole center, the dipole as a whole will have a nonzero R_{56} or, more accurately, a nonzero phase slippage. Therefore, to make the longitudinal optics identical in different dipoles, there should be RF or laser modulator kicks between neighboring dipoles; otherwise, the required drift space between each pair of dipoles will be very long to compensate this nonzero phase slippage [7]. It may not be easy to apply too many RF cavities or laser modulators in a ring to manipulate the longitudinal optics, while in the transverse dimension it is straightforward to implement many quadrupoles to manipulate the transverse optics. Instead, we may choose a more practical strategy to realize a small longitudinal emittance, which is to let each half of the bending magnet be isochronous, and the longitudinal optics for each dipole can then be identical. This can be realized by requiring

$$\alpha_{z0}=0,\ \beta_{z0}\approx\frac{\rho\theta^3}{12\sqrt{210}},\ D_{x0}\approx-\frac{\rho\theta^2}{24},\ D'_{x0}=0. \tag{87}$$

In this case, we still have $f_z=2\beta_{z0}$. Under such conditions, the minimum longitudinal emittance is [17]

$$\epsilon_{z,\min,\mathrm{ISO}}=C_{\mathrm{q}}\frac{\gamma^2}{J_z}\frac{\theta^3}{6\sqrt{210}}. \tag{88}$$

The emittance given in Eq. (88) is larger than the real theoretical minimum Eq. (86), but offers a more practical reference for the TMEz. For practical use, and considering nominally $J_z\approx 2$, the above scaling can be written as

$$\epsilon_{z,\min,\mathrm{ISO}}[\mathrm{nm}]=8.44E_0^2[\mathrm{GeV}]\theta^3[\mathrm{rad}]. \tag{89}$$

Based on this emittance and β_{z0}, we can determine the bunch length at the dipole center contributed by the longitudinal emittance. We remind the readers that the bunch length can, in principle, be even smaller than this value, by pushing β_{z0} to an even smaller value, although the longitudinal emittance will actually grow then because β_z will diverge faster when moving away from the dipole center. If the ring works in a longitudinal weak focusing regime to be introduced in next section, there is actually a lower limit of bunch length in this process, which is a factor of $\sqrt{2}$ smaller than the bunch length given by conditions of Eqs. (89) and (87). The energy spread diverges when the bunch length is pushed to this limit [30]. Putting in the numbers, we have this lower limit of bunch length in a longitudinal weak focusing ring [17]

$$\sigma_{z,\min,\mathrm{ISO}}[\mu\mathrm{m}]=4.93\rho^{\frac{1}{2}}[\mathrm{m}]E_0[\mathrm{GeV}]\theta^3[\mathrm{rad}]. \tag{90}$$

For example, if $E_0=600$ MeV, $\theta=\frac{2\pi}{50}$ rad and $\rho=1.5$ m which corresponds to a bending field strength $B_0=1.33$ T, then $\sigma_{z,\min,\mathrm{ISO}}=7.2$ nm.

3.3 Application of transverse gradient bend

The previous analysis assumed that the transverse gradient of the bending field was zero. We now consider the application of transverse gradient bending (TGB) magnets to lower the horizontal and longitudinal emittances. For simplicity, we considered the case of a constant gradient. The transfer matrix of a sector bending magnet with a constant transverse gradient $n=-\frac{\rho}{B_y}\frac{\partial B_y}{\partial x}$ is

$$\mathbf{B}_{\mathrm{TGB}}(\alpha)=\begin{pmatrix} \cos\left(\sqrt{1-n}\alpha\right) & \frac{\rho}{\sqrt{1-n}}\sin\left(\sqrt{1-n}\alpha\right) & 0 & 0 & 0 \\ -\frac{\sqrt{1-n}}{\rho}\sin\left(\sqrt{1-n}\alpha\right) & \cos\left(\sqrt{1-n}\alpha\right) & 0 & 0 & 0 \\ 0 & 0 & \cos\left(\sqrt{n}\alpha\right) & \frac{\rho}{\sqrt{n}}\sin\left(\sqrt{n}\alpha\right) & 0 \\ 0 & 0 & -\frac{\sqrt{n}}{\rho}\sin\left(\sqrt{n}\alpha\right) & \cos\left(\sqrt{n}\alpha\right) & 0 \\ -\frac{1}{\sqrt{1-n}}\sin\left(\sqrt{1-n}\alpha\right) & -\frac{\rho}{1-n}\left(1-\cos\left(\sqrt{1-n}\alpha\right)\right) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & \frac{\rho}{1-n}\left(1-\cos\left(\sqrt{1-n}\alpha\right)\right) \\ 0 & \frac{1}{\sqrt{1-n}}\sin\left(\sqrt{1-n}\alpha\right) \\ 0 & 0 \\ 0 & 0 \\ \frac{\rho\alpha}{\gamma^2}+\frac{\rho}{(1-n)^{\frac{3}{2}}}\left(-\sqrt{1-n}\alpha+\sin\left(\sqrt{1-n}\alpha\right)\right) & 0 \\ 0 & 1 \end{pmatrix}. \tag{91}$$

3.3.1 Horizontal emittance

Following similar steps presented in the above analysis, we find the minimum value of f_x in Eq. (65) is now realized when

$$\begin{aligned} &\alpha_{x0}=0,\ \beta_{x0}\approx\frac{\rho\theta}{2\sqrt{15}}\left[1+\frac{(1-n)\theta^2}{140}\right]+\mathcal{O}(\theta^6),\\ &D_{x0}\approx\frac{\rho\theta^2}{24}\left[1-\frac{(1-n)\theta^2}{48}\right]+\mathcal{O}(\theta^6),\ D'_{x0}=0, \end{aligned} \tag{92}$$

where $\mathcal{O}(\theta^n)$ means terms of order θ^n and higher, which means

$$\mathcal{H}_{x0}=\frac{D_{x0}^2}{\beta_{x0}}\approx\frac{5\rho\theta^3}{96\sqrt{15}}\left[1-\frac{9(1-n)\theta^2}{280}\right]+\mathcal{O}(\theta^7), \tag{93}$$

and

$$f_{x,\min}\approx\frac{\rho\theta^3}{12\sqrt{15}}\left[1-\frac{3(1-n)\theta^2}{70}\right]+\mathcal{O}(\theta^7). \tag{94}$$

So, we have

$$\epsilon_{x,\min,\mathrm{TGB}}=C_{\mathrm{q}}\frac{\gamma^2}{J_x}\frac{\theta^3}{12\sqrt{15}}\left[1-\frac{3(1-n)\theta^2}{70}\right]. \tag{95}$$

Therefore, the impact of the transverse gradient on the theoretical minimum emittance is on a higher order than the bending angle of each magnet. However, we should recognize that n can be quite large in practice. Therefore, its impact may actually be not small. In addition, a transverse gradient bend can also affect the damping partition, the details of which are not discussed here.

3.3.2 Longitudinal emittance

Similarly, using TGB, the minimum value of f_z in Eq. (77) is realized when

$$\begin{aligned} &\alpha_{z0}=0,\ \beta_{z0}\approx\frac{\rho\theta^3}{120\sqrt{7}}\left[1+\frac{(1-n)\theta^2}{90}\right]+\mathcal{O}(\theta^7),\\ &D_{x0}\approx-\frac{\rho\theta^2}{40}\left[1+\frac{19(1-n)\theta^2}{1680}\right]+\mathcal{O}(\theta^6),\ D'_{x0}=0, \end{aligned} \tag{96}$$

and

$$f_{z,\min}\approx\frac{\rho\theta^3}{60\sqrt{7}}\left[1+\frac{(1-n)\theta^2}{90}\right]+\mathcal{O}(\theta^7). \tag{97}$$

Note that $f_{z,\min}=2\beta_{z0}$ still holds here. So, we have

$$\epsilon_{z,\min,\mathrm{TGB}}=C_{\mathrm{q}}\frac{\gamma^2}{J_z}\frac{\theta^3}{60\sqrt{7}}\left[1+\frac{(1-n)\theta^2}{90}\right]. \tag{98}$$

3.4 Application of longitudinal gradient bend

We can also apply longitudinal gradient bends (LGBs) to lower the transverse and longitudinal emittances. For simplicity, we studied the case of an LGB consisting of several sub-dipoles, each with a constant bending radius. Furthermore, we assume that each sub-dipole is a sector dipole. The analysis for the case of rectangular dipoles is similar, as long as the impact of the edge angles on the transfer matrix and damping partition is properly handled. We now investigate the case of sector sub-dipoles. For example, we may choose to let the LGB has a symmetric structure

$$(\rho_2,\theta_2),(\rho_1,\theta_1),(\rho_0,2\theta_0),(\rho_1,\theta_1),(\rho_2,\theta_2). \tag{99}$$

The total bending angle of such a LGB is $\theta_{\mathrm{T}} = 2(\theta_0+\theta_1+\theta_2)$, and the total length is $2(\rho_0\theta_0+\rho_1\theta_1+\rho_2\theta_2)$. Note that $\rho_i\theta_i \geq 0$. We used this structure as an example for the analysis. However, the proposed method also applies to a more general setup.

3.4.1 Horizontal emittance

Now, we calculate the theoretical minimum horizontal emittance by invoking LGBs with each structure given in Eq. (99). Still we assume all the LGB setup in the ring and optical functions in each LGB are identical, then

$$\epsilon_x = C_{\mathrm{q}}\frac{\gamma^2}{J_x}I_{5x}(\alpha_{x0},\beta_{x0},D_{x0},D'_{x0}) \tag{100}$$

with

$$I_{5x}(\alpha_{x0},\beta_{x0},D_{x0},D'_{x0}) = \frac{1}{2\left(\frac{\theta_0}{\rho_0}+\frac{\theta_1}{\rho_1}+\frac{\theta_2}{\rho_2}\right)}\int_{-(\rho_0\theta_0+\rho_1\theta_1+\rho_2\theta_2)}^{(\rho_0\theta_0+\rho_1\theta_1+\rho_2\theta_2)}\frac{\mathcal{H}_x(s)}{|\rho(s)|^3}\mathrm{d}s. \tag{101}$$

Note that the dimension of I_{5x} here is equivalent to $\frac{f_x}{\rho}$ given in the previous sections. The mathematical problem is then to minimize I_{5x}, by adjusting $\alpha_{x0}, \beta_{x0}, D_{x0}, D'_{x0}$. The calculation method of $\mathcal{H}_x$ evolution in an LGB is the same as that given in Eq. (62), but note that here for different parts (subdipole) of the LGBs, we should apply the transfer matrix from the midpoint of the LGB to the corresponding location.

Following similar procedures, we find that to obtain the minimum emittance, we still need $\alpha_{x0} = 0$ and $D'_{x0} = 0$. Then we have $\mathcal{H}_x(-s) = \mathcal{H}_x(s)$, which means

$$I_{5x} = \frac{1}{\left(\frac{\theta_0}{\rho_0}+\frac{\theta_1}{\rho_1}+\frac{\theta_2}{\rho_2}\right)}\left(\int_0^{\rho_0\theta_0}\frac{\mathcal{H}_x(s)}{|\rho_0|^3}\mathrm{d}s+\int_{\rho_0\theta_0}^{\rho_0\theta_0+\rho_1\theta_1}\frac{\mathcal{H}_x(s)}{|\rho_1|^3}\mathrm{d}s+\int_{\rho_0\theta_0+\rho_1\theta_1}^{\rho_0\theta_0+\rho_1\theta_1+\rho_2\theta_2}\frac{\mathcal{H}_x(s)}{|\rho_2|^3}\mathrm{d}s\right). \tag{102}$$

We find a general analytical discussion of the combination of ρ_i and θ_i to be cumbersome. For simplicity and to obtain a concrete feeling, we first consider one specific case: $\rho_i = \rho_0 2^i$. The physical consideration behind this choice is that $\mathcal{H}_x$ is smaller in the central part of the bending magnet and larger at the entrance and exit. Therefore, we make the bending field in the center stronger and smaller at the entrance and exit regions, to minimize the quantum excitation of the horizontal emittance. For example, we may choose

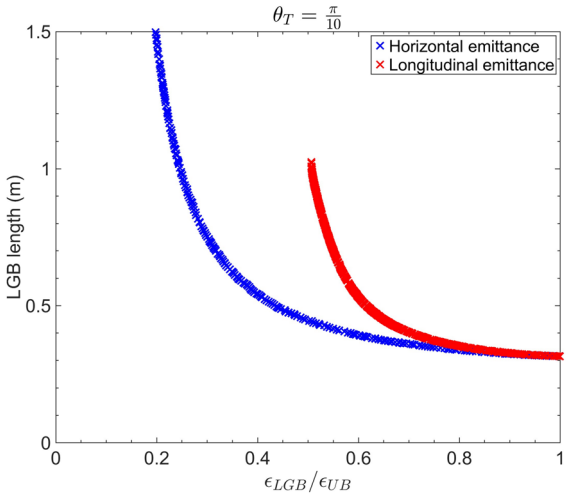

Fig. 2 Application of LGB to minimize horizontal (blue) and longitudinal (red) emittance, respectively. The subscripts "LGB" and "UB" represent the longitudinal gradient bend and uniform bend, respectively

$$\left(4\rho,\frac{\theta}{8}\right),\left(2\rho,\frac{\theta}{8}\right),\left(\rho,\frac{\theta}{4}\right),\left(\rho,\frac{\theta}{4}\right),\left(2\rho,\frac{\theta}{8}\right),\left(4\rho,\frac{\theta}{8}\right). \tag{103}$$

The total length of such an LGB is $2\rho\theta$. The minimum value of I_{5x} and horizontal emittance ϵ_x in this case is realized when

$$\alpha_{x0} = 0,\ \beta_{x0} \approx \frac{\sqrt{\frac{35977}{133755}}\rho\theta}{4},\ D_{x0} \approx \frac{127\rho\theta^2}{7104},\ D'_{x0} = 0, \tag{104}$$

which means

$$\mathcal{H}_{x0} = \frac{D_{x0}^2}{\beta_{x0}} \approx 0.183\times\frac{5}{96\sqrt{15}}\rho\theta^3, \tag{105}$$

and

$$I_{5x,\min} \approx 0.344\times\frac{\theta^3}{12\sqrt{15}}, \tag{106}$$

which compared to Eq. (74) means that the theoretical minimum horizontal emittance can now become approximately one-third of the case with no longitudinal gradient.

Therefore, the application of the LGB is quite effective in lowering the transverse emittance.

The next question is: what is the optimal combination of $\theta_{0,1,2}$ and $\rho_{0,1,2}$? This question is not straightforward to answer using analytical method. Here, we refer to the numerical method to perform the optimization directly. α_{x0} and D'_{x0} are set to be zero in the optimization. The variables in the numerical optimization are $(\theta_0, \theta_1, \theta_2, \rho_0, \rho_1, \rho_2, \beta_{x0}, D_{x0})$. The two optimization goals are $\frac{\epsilon_{x,\mathrm{LGB}}}{\epsilon_{x,\mathrm{UB}}}$ and the length of the LGB L_{LGB}, where $\epsilon_{x,\mathrm{UB}}$ is the theoretical minimum emittance of applying bending magnet without longitudinal gradient. We require $\rho_i\theta_i > 0$. In the optimization, we kept the total bending angle of the LGB at a constant value.

The optimization result of a specific case where $\theta_{\mathrm{T}} = 2(\theta_0 + \theta_1 + \theta_2) = \frac{\pi}{10}$ is presented in Fig. 2, from which we can see that by applying LGBs, we can lower the horizontal emittance by a factor of five with a reasonable length of the LGB. Note that in this optimization and the one in the following section, we assume that $\rho_i > 0$. However, the formalism also applies to the case with anti-bends.

Table 1 Main characteristics of the LWF, LSF, and GLSF SSMB storage rings

LWF	$\nu_s \ll 1$	$\frac{\beta_{z,\max}}{\beta_{z,\min}} \approx 1$	2D phase space dynamics
LSF	$\nu_s \sim 1$	$\frac{\beta_{z,\max}}{\beta_{z,\min}} \gg 1$	2D phase space dynamics
GLSF	–	$\frac{\sigma_{z,\max}}{\sigma_{z,\min}} \gg 1$	4D or 6D phase space dynamics

3.4.2 Longitudinal emittance

Similarly, for the longitudinal emittance we have

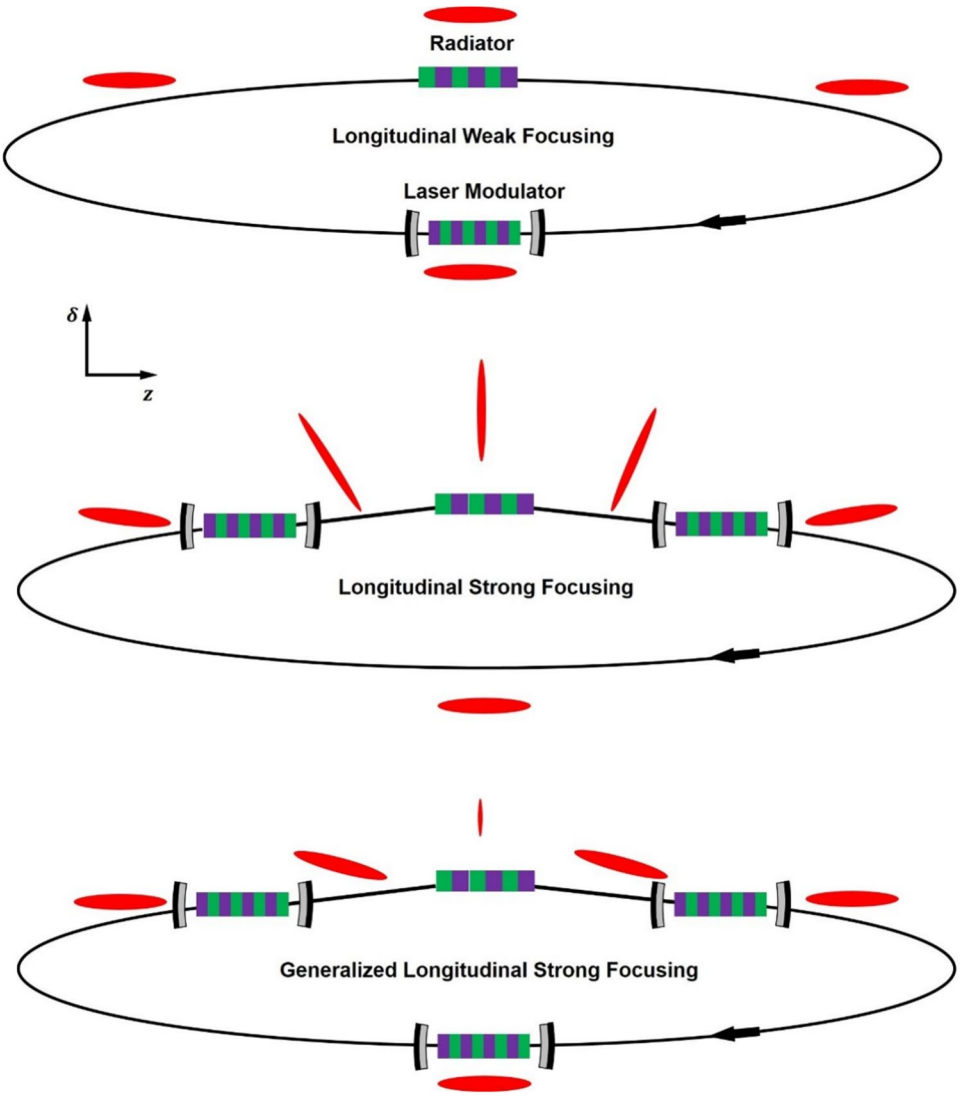

Fig. 3 (Color online) Schematic layout of the longitudinal weak focusing (LWF), longitudinal strong focusing (LSF), and generalized longitudinal strong focusing (GLSF) SSMB storage rings. The red ellipses illustrate the beam distribution in the longitudinal phase space. To realize the same bunch length compression ratio, the required energy chirp strength in the GLSF scheme is much smaller than that in the LSF

$$\epsilon_z = C_q \frac{\gamma^2}{J_z} I_{5z}(\alpha_{z0}, \beta_{z0}, D_{x0}, D'_{x0}) \tag{107}$$

with

$$I_{5z}(\alpha_{z0}, \beta_{z0}, D_{x0}, D'_{x0}) = \frac{1}{2\left(\frac{\theta_0}{\rho_0} + \frac{\theta_1}{\rho_1} + \frac{\theta_2}{\rho_2}\right)} \oint_{-(\rho_0\theta_0+\rho_1\theta_1+\rho_2\theta_2)}^{(\rho_0\theta_0+\rho_1\theta_1+\rho_2\theta_2)} \frac{\beta_z(s)}{|\rho(s)|^3} \mathrm{d}s. \tag{108}$$

For example, if we still choose the setup given in Eq. (103), the minimum value of I_{5z} and longitudinal emittance in this case is realized when

$$\alpha_{z0} = 0,\ \beta_{z0} \approx \frac{\sqrt{\frac{36233641}{62419}}\rho\theta^3}{7680},$$
$$D_{x0} \approx -\frac{169\rho\theta^2}{9640},\ D'_{x0} = 0, \tag{109}$$

and

$$F_{z,\min} \approx 0.838 \times \frac{\theta^3}{60\sqrt{7}}, \tag{110}$$

which compared with Eq. (85) means that the theoretical minimum longitudinal emittance can become slightly smaller than the case of applying a constant bending field.

Similar to what was presented previously about lowering the transverse emittance, we also apply numerical optimization to choose a better combination of ρ_i and θ_i for lowering the longitudinal emittance. Presented in Fig. 2 is the result of one specific case where $\theta_T = 2(\theta_0 + \theta_1 + \theta_2) = \frac{\pi}{10}$, from which we can see that in this case by applying LGBs, in principle we can lower the longitudinal emittance by a factor of two with a reasonable length of the LGB. Therefore, LGB is generally more effective in lowering the horizontal emittance than the longitudinal emittance.

4 Steady-state microbunching storage rings

In this section, based on the theoretical minimum emittance derived in the last section, we conduct a key analysis of three specific SSMB scenarios to realize nanometer bunch length and high-average-power EUV radiation, i.e., longitudinal weak focusing (LWF), longitudinal strong focusing (LSF), and generalized longitudinal strong focusing (GLSF). The analysis aims to answer the question of why GLSF SSMB is the present choice for realizing high-average-power EUV radiation. Before going into the details, here first we use Table 1 and Fig. 3 to briefly summarize the characteristics of these three scenarios. Note that in Fig. 3, the beam distribution in the longitudinal phase space is that of the microbunch, whose length is in the laser wavelength range. We also remind the readers that in the figure, the energy chirp strength in GLSF is much smaller than that of LSF. The physical reason should be clear with the analysis in this section unfolded.

In all the example calculations to be shown in the following part of this paper, we set the electron energy to be $E_0 = 600$ MeV and modulation laser wavelength to be $\lambda_L = 1064$ nm. This beam energy was selected because it is appropriate for EUV generation using an undulator as a radiator. On one hand, it is not too high; otherwise, the laser modulation will become more difficult, which means that more laser power is required to imprint a given modulation strength. However, it is not too low; otherwise, intra-beam scattering (IBS) could become too severe. Actually as we will see in Sect. 8, IBS is a fundamental issue in SSMB storage rings which require at least one of the three eigenemittances to be small. The reason for choosing this laser wavelength is due to the fact that it is the common wavelength range for a high-power optical enhancement cavity, which is used together with an undulator to form the laser modulator of the SSMB.

4.1 Longitudinal weak focusing

We now begin the quantitative analysis. We begin with a longitudinal weak focusing (LWF) SSMB ring. In a LWF ring with a single laser modulator (LM) as shown in Fig. 3, the single-particle longitudinal dynamics turn by turn is modeled as

$$\delta_{n+1} = \delta_n + \frac{h}{k_L}\sin(k_L z_n),$$
$$z_{n+1} = z_n - \eta C_0 \delta_{n+1}, \tag{111}$$

where the subscript $n, n+1$ means the number of revolutions, h is the energy chirp strength around the zero-crossing phase, $k_L = \frac{2\pi}{\lambda_L}$ is the wavenumber of the modulation laser and λ_L is the laser wavelength, C_0 is the ring circumference, and

$$\eta = \frac{\Delta T/T_0}{\Delta E/E_0} = \frac{1}{C_0}\oint\left(\frac{D_x}{\rho} - \frac{1}{\gamma^2}\right)\mathrm{d}s \tag{112}$$

is the phase slippage factor of the ring. Note that in the above model, we have assumed that the radiation energy loss in an SSMB storage ring will be compensated by other systems instead of the laser modulators; thus, microbunching will be formed around the laser zero-crossing phase. The laser can also be used for energy compensation; however, it is not a cost-effective choice and will also limit the output radiation power. Induction linacs or RF cavities can be used to supply the radiation energy loss. Linear stability of Eq. (111) around the zero-crossing phase requires that $0 < h\eta C_0 < 4$. To avoid strong chaotic dynamics which may destroy the

regular longitudinal phase space structure, an empirical criterion is

$$0 < h\eta C_0 \lesssim 0.1. \tag{113}$$

In a LWF ring, when the synchrotron tune $|\nu_s| \approx \sqrt{h\eta C_0}/2\pi \ll 1$, the longitudinal beta function at the center of laser modulator is [17]

$$\beta_{z\mathrm{M}} \approx \sqrt{\frac{\eta C_0}{h}}. \tag{114}$$

Note that in this paper, we will use the subscript 'M' to represent modulator and 'R' to represent radiator. From Eq. (113) we then require

$$|\eta C_0| \lesssim \frac{\beta_{z\mathrm{M}}}{\sqrt{10}}. \tag{115}$$

Now, we can use the previous analysis of the theoretical minimum longitudinal emittance, more specifically, Eq. (90), and the above result to evaluate a LWF ring. If $E_0 = 600$ MeV, the bending radius of dipoles in the ring $\rho_{\mathrm{ring}} = 1.5$ m, which corresponds to a bending field strength $B_{\mathrm{ring}} = 1.33$ T, and if our desired bunch length is $\sigma_z = 50$ nm ($\sigma_z \lesssim \lambda_{\mathrm{L}}/20$ for microbunches to be safely stored in the optical microbuckets for $\lambda_{\mathrm{L}} = 1064$nm), then we may need $\sigma_{z,\mathrm{min,ISO}} \leq 50/\sqrt{2}$ nm to avoid significant energy widening when we reach the desired bunch length. From Eq. (90) we then need $\theta \leq \frac{2\pi}{30}$ rad. Therefore, at least 30 bending magnets are required in the ring. Assuming that the length of each isochronous cell containing a bending magnet is 3 m, the arc section of such a storage ring has a length of approximately 90 m. Considering the straight section for beam injection/extraction, radiation energy loss compensation, and insertion device for radiation generation, the circumference of such a ring could be 100 m to 120 m.

To achieve the desired bunch length, according to Eq. (87) we need $\beta_{z\mathrm{M}} = \beta_{z0} = \rho\theta^3/12\sqrt{210}$. In our present example case, if $\rho = 1.5$ m and $\theta = \frac{2\pi}{30}$, then this value we need is $\beta_{z\mathrm{M}} = 79.2$ μm. Then from Eq. (115) we have

$$|\eta C_0| \lesssim 25\ \mu\mathrm{m}. \tag{116}$$

If $C_0 = 100$ m, then we need a phase slippage factor $|\eta| \lesssim 2.5 \times 10^{-7}$, which is a quite small value. If we want a bunch length even smaller than 50 nm at a beam energy of 600 MeV, then the required phase slippage will be too demanding to be realized using the present technology. More details on the lattice design of an LWF SSMB ring that can store microbunches with a couple of 10 nm-bunch lengths can be found in Ref. [16].

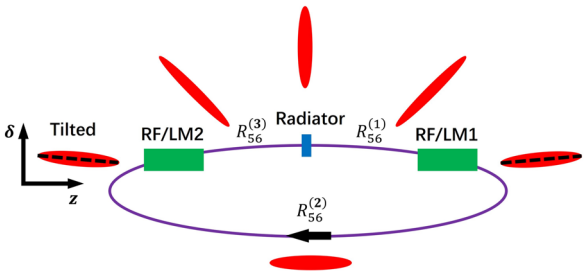

Fig. 4 (Color online) A schematic layout of a storage ring using two RF systems for longitudinal strong focusing and an example beam distribution evolution in the longitudinal phase space. (Figure from Ref. [8])

4.2 Longitudinal strong focusing

After discussing the LWF SSMB ring, we now begin the analysis of the LSF. First, we observe that the above analysis of the LWF SSMB considers the case with only a single LM. When there are multiple LMs, the longitudinal dynamics are similar to those of multiple quadrupoles in the transverse dimension, and the beam dynamics can have more possibilities. For example, a longitudinal strong focusing scheme can be invoked, similar to its transverse counterpart, which is the foundation of modern high-energy accelerators. Here, we use a setup with two LMs for the SSMB as an example to show the scheme of manipulating β_z around the ring using the strong focusing regime. The schematic layout of the ring is illustrated in Fig. 4. The treatment of cases with more LMs is similar.

We divide the ring into five sections from the transfer matrix viewpoint, i.e., three longitudinal drifts (R_{56}) and two LM kicks (h), with the linear transfer matrices of the state vector in the longitudinal dimension $(z, \delta)^{\mathrm{T}}$ given by

$$\begin{aligned}
\mathbf{T}_{\mathrm{D1}} &= \begin{pmatrix} 1 & R_{56}^{(1)} \\ 0 & 1 \end{pmatrix}, \ \mathbf{T}_{\mathrm{LM1}} = \begin{pmatrix} 1 & 0 \\ h_1 & 1 \end{pmatrix}, \\
\mathbf{T}_{\mathrm{D2}} &= \begin{pmatrix} 1 & R_{56}^{(2)} \\ 0 & 1 \end{pmatrix}, \ \mathbf{T}_{\mathrm{LM2}} = \begin{pmatrix} 1 & 0 \\ h_2 & 1 \end{pmatrix}, \\
\mathbf{T}_{\mathrm{D3}} &= \begin{pmatrix} 1 & R_{56}^{(3)} \\ 0 & 1 \end{pmatrix}.
\end{aligned} \tag{117}$$

Then the one-turn map at the radiator center is

$$\mathbf{M}_{\mathrm{R}} = \mathbf{T}_{\mathrm{D3}}\mathbf{T}_{\mathrm{LM2}}\mathbf{T}_{\mathrm{D2}}\mathbf{T}_{\mathrm{LM1}}\mathbf{T}_{\mathrm{D1}}. \tag{118}$$

For the generation of coherent radiation, we usually want the bunch length to reach its minimum at the radiator, and then we need $\alpha_z = 0$ at the radiator.

With the primary goal of presenting the principle, instead of a detailed design, here for simplicity, we focus on one special case: $R_{56}^{(1)} = R_{56}^{(3)}$, $h_1 = h_2 = h$. Denote

$$\zeta_1 \equiv 1 + R_{56}^{(1)} h,\ \zeta_2 \equiv 1 + \frac{R_{56}^{(2)}}{2} h, \tag{119}$$

we then have

$$\mathbf{M}_{\mathrm{R}} = \begin{pmatrix} 2\zeta_1\zeta_2 - 1 & 2\frac{\zeta_1^2\zeta_2 - \zeta_1}{h} \\ 2h\zeta_2 & 2\zeta_1\zeta_2 - 1 \end{pmatrix}. \tag{120}$$

The linear stability requires $|2\zeta_1\zeta_2 - 1| < 1$ which means $0 < \zeta_1\zeta_2 < 1$.

We remind the readers that the analysis above in this subsection about LSF has been presented in Ref. [17] and is presented here again for completeness of the paper. We now attempt to gain further insight. The longitudinal beta function at the radiator is

$$\beta_{z\mathrm{R}} = \frac{2\frac{\zeta_1^2\zeta_2 - \zeta_1}{h}}{\sin\Phi_z} = \frac{1}{|h|}\sqrt{\frac{\zeta_1(1-\zeta_1\zeta_2)}{\zeta_2}}, \tag{121}$$

where $\Phi_z = 2\pi\nu_s$ is the synchrotron phase advance per turn. The longitudinal beta function at the opposite of the radiator from where to LM2 has an $R_{56} = \frac{R_{56}^{(2)}}{2}$, is

$$\beta_{z\mathrm{RO}} = \frac{1}{|h|}\sqrt{\frac{\zeta_2(1-\zeta_1\zeta_2)}{\zeta_1}}, \tag{122}$$

which can be obtained by switching ζ_1 and ζ_2 in the expression for $\beta_{z\mathrm{R}}$ in Eq. (121). So, we have

$$\frac{\beta_{z\mathrm{RO}}}{\beta_{z\mathrm{R}}} = \frac{\zeta_2}{\zeta_1}. \tag{123}$$

The longitudinal beta function at the LM1 and LM2 (here we name them as modulators) is

$$\beta_{z\mathrm{M}} = \beta_{z\mathrm{R}} + \frac{R_{56}^{(1)2}}{\beta_{z\mathrm{R}}}. \tag{124}$$

For efficient bunch compression from the modulator to the radiator, we have $|\zeta_1| \ll 1$. Then

$$\frac{\beta_{z\mathrm{M}}}{\beta_{z\mathrm{R}}} = \frac{\zeta_1 - 2\zeta_1\zeta_2 + \zeta_2}{\zeta_1(1-\zeta_1\zeta_2)} \approx \frac{\zeta_2}{\zeta_1(1-\zeta_1\zeta_2)} = \frac{1}{h^2\beta_{z\mathrm{R}}^2}, \tag{125}$$

from which we have the required energy chirp strength

$$|h| \approx \frac{1}{\sqrt{\beta_{z\mathrm{R}}\beta_{z\mathrm{M}}}}. \tag{126}$$

This relationship is similar to the theorems presented in Sect. 5 about transverse-longitudinal coupling-based bunch compression schemes, which are the backbone of the GLSF scheme. The relation can also be casted as

$$|h| \approx \frac{\sqrt{\epsilon_z/\beta_{z\mathrm{M}}}}{\sqrt{\epsilon_z\beta_{z\mathrm{R}}}} = \frac{\sqrt{\epsilon_z/\beta_{z\mathrm{M}}}}{\sigma_{z\mathrm{R}}}, \tag{127}$$

with $\sigma_{z\mathrm{R}} = \sqrt{\epsilon_z\beta_{z\mathrm{R}}}$ being the bunch length at the radiator.

We now investigate the required energy chirp strength and modulation laser power based on this analysis. We assume that in an LSF ring, the longitudinal emittance is dominated by quantum excitation in the ring dipoles. This assumption will be justified later to determine whether it is the case. Furthermore, we assume that the average longitudinal beta function at the dipoles equals that at the modulator, i.e., $\langle\beta_z\rangle = \beta_{z\mathrm{M}}$. Then the equilibrium longitudinal emittance given by the balance of quantum excitation and radiation damping in a ring consisting of iso-bending magnets has the scaling

$$\epsilon_{z,\mathrm{LSF}} = C_{\mathrm{q}}\frac{\gamma^2}{J_z}\frac{I_{5z}}{I_2} = C_{\mathrm{q}}\frac{\gamma^2}{J_z}\frac{\frac{\langle\beta_z\rangle}{\rho_{\mathrm{ring}}^3}2\pi\rho_{\mathrm{ring}}}{\frac{1}{\rho_{\mathrm{ring}}^2}2\pi\rho_{\mathrm{ring}}} \propto \gamma^2\frac{\beta_{z\mathrm{M}}}{\rho_{\mathrm{ring}}}, \tag{128}$$

which combining with Eq. (127) gives

$$|h| \propto \frac{\gamma}{\sigma_{z\mathrm{R}}\sqrt{\rho_{\mathrm{ring}}}}. \tag{129}$$

Here, $\sigma_{z\mathrm{R}}$ is determined by the desired radiation wavelength. Therefore, given the beam energy and desired bunch length, to lower the required energy chirp strength in the LSF, we should use as large ρ_{ring} as possible, which means as weak a bending magnet as possible. However, the total length of the bending magnets should be within a reasonable range. Also note that when the bending magnets in the ring are very weak, the assumption that the longitudinal emittance in a LSF ring is dominantly from them will fail, since the quantum excitation in the bending magnets will become weaker, while there are other contributions like quantum excitation at the laser modulators.

As mentioned in Sect. 1, a short bunch can generate coherent radiation. The parameter used to quantify the capability of beam for coherent radiation generation is called bunching factor, and in the 1D case is defined as

$$b(\omega) = \int_{-\infty}^{\infty}\psi(z)e^{-i\frac{\omega}{c}z}\mathrm{d}z, \tag{130}$$

where ω is the radiation frequency, $\psi(z)$ is the longitudinal charge density distribution satisfying the normalization $\int_{-\infty}^{\infty}\psi(z)\mathrm{d}z = 1$. We will present a more in-depth discussion of the bunching factor in Sect. 6.1. The coherent radiation power of a beam with N_{p} particles at frequency ω is related to the radiation of a single particle according to

$$P_{\mathrm{beam}}(\omega) = N_{\mathrm{p}}^2 b^2(\omega) P_{\mathrm{single}}(\omega). \tag{131}$$

For a Gaussian bunch with an RMS bunch length of σ_z, we have $b(\omega) = \exp\left[-\left(\frac{\omega}{c}\sigma_z\right)^2/2\right]$. For significant 13.5 nm-wavelength coherent EUV radiation generation, we may need $\sigma_{z\mathrm{R}} \lesssim 4$ nm which corresponds to $b_{13.5\,\mathrm{nm}} \gtrsim 0.18$. To increase the radiation power, we may need a radiator, which is assumed to be an undulator, with a large period number N_u (for example, $N_\mathrm{u} \approx 300$). To avoid significant bunch lengthening from the energy spread and the undulator $R_{56} = 2N_\mathrm{u}\lambda_\mathrm{R}$, we then need $N_\mathrm{u}\lambda_\mathrm{R}\sigma_\delta \lesssim \sigma_{z\mathrm{R}}$, which then requires the energy spread at the radiator $\sigma_{\delta\mathrm{R}} \lesssim 1\times10^{-3}$. So, we need $\epsilon_z = \sigma_{z\mathrm{R}}\sigma_{\delta\mathrm{R}} \lesssim 4$ pm. Since in a LSF ring, as explained we cannot make the optimal conditions Eq. (83) be satisfied in all the bending magnets, a reasonable argument is that the real longitudinal emittance should be at least a factor of two larger than the true theoretical minimum, which then requires

$$\epsilon_{z,\min} \lesssim 2\ \mathrm{pm}. \tag{132}$$

Then according to Eq. (86), if $E_0 = 600$ MeV, we need $\theta \lesssim \frac{2\pi}{59}$, which means 59 bending magnets are needed. Assuming that the length of each isochronous cell containing a bending magnet is 3 m, the arc section of such a ring has a length of ~177 m. Considering the straight section for beam injection/extraction, radiation energy loss compensation, and longitudinal strong focusing section, the circumference of such a ring is ~200 m.

If $\epsilon_z = 4$ pm, to get $\sigma_{z\mathrm{R}} \lesssim 4$ nm, we then need

$$\beta_{z\mathrm{R}} = \frac{\sigma_{z\mathrm{R}}^2}{\epsilon_z} \lesssim 4\ \mu\mathrm{m}. \tag{133}$$

From Eq. (126), given $\beta_{z\mathrm{R}}$, we should apply as large $\beta_{z\mathrm{M}}$ as possible to decrease the energy chirp strength h. Because $\beta_{z0} \propto \rho$ with θ given, we will choose a reasonably large bending radius for the dipoles in the ring. If $\rho = 10$ m, which corresponds to $B_0 = 0.2$ T and a total bending magnet length of 62.8 m, then the longitudinal beta function at the dipole center required to reach the practical theoretical minimum longitudinal emittance (see Eqs. (89) and (87)) is $\beta_{z0} = \rho\theta^3/12\sqrt{210} = 69.5\ \mu$m. Then we may let

$$\beta_{z\mathrm{M}} \approx 2\beta_{z0} \approx 139\ \mu\mathrm{m}. \tag{134}$$

Then from Eqs. (126), (133) and (134), the energy chirp strength required in such a LSF SSMB ring is

$$|h| \approx \frac{1}{\sqrt{\beta_{z\mathrm{R}}\beta_{z\mathrm{M}}}} \gtrsim 4.24\times10^4\ \mathrm{m}^{-1}. \tag{135}$$

According to Eq. (221) to be presented later about the laser modulator induced energy chirp strength, if $E_0 = 600$ MeV, and for a modulator undulator with period $\lambda_\mathrm{uM} = 8$ cm ($B_{0\mathrm{M}} = 1.13$ T), length $L_\mathrm{uM} = 1.6$ m, and for a laser with wavelength $\lambda_\mathrm{L} = 1064$ m, to introduce the required energy chirp strength, we need a laser power $P_\mathrm{L} \approx 1$ GW. This is a large value, three orders of magnitude higher than the average stored laser power reachable in an optical enhancement cavity at the moment [31], which is at the level of one megawatt (MW). This causes the optical enhancement cavity to only work in a low duty cycle pulsed mode, thus limiting the filling factor of the microbunched electron beam in the ring and the average output EUV power.

We remind the readers that there is a subtle point in an LSF SSMB ring if we consider the nonlinear sinusoidal modulation waveform, since the dynamical system is then strongly chaotic and requires careful analysis to ensure a sufficiently large stable region for particle motion in the longitudinal phase space. More details on this respect can be found in Ref. [18].

For the completeness of discussion, let us evaluate the contribution of modulator undulators to the longitudinal emittance in our above example case, since there is also quantum excitation at the modulators. The quantum excitation contributions of two modulators to ϵ_z in a LSF ring are

$$\begin{aligned}\Delta\epsilon_{z\mathrm{M}} &= C_\mathrm{q}\frac{\gamma^2}{J_z}\frac{\Delta I_{5z\mathrm{M}}}{I_2}\\ &= C_\mathrm{q}\frac{\gamma^2}{J_z}\frac{1}{I_2}\times 2\int_{-\frac{L_{u\mathrm{M}}}{2}}^{\frac{L_{u\mathrm{M}}}{2}}\frac{\beta_{z\mathrm{M}}}{|\rho(s)|^3}\mathrm{d}s\\ &= C_\mathrm{q}\frac{\gamma^2}{J_z}\frac{1}{I_2}\times 2\frac{\beta_{z\mathrm{M}}}{\rho_{0\mathrm{M}}^3}\frac{4}{3\pi}L_\mathrm{uM},\end{aligned} \tag{136}$$

where $\Delta I_{5z\mathrm{M}}$ is the contribution of the two modulators to I_{5z} defined in Eq. (54), $\rho_{0\mathrm{M}}$ is the bending radius at the peak magnetic field of the modulator. Putting in the numbers, and taking the approximation $I_2 \approx \frac{2\pi}{\rho_\mathrm{ring}}$ which means the radiation loss is mainly from dipoles in the ring, $J_z \approx 2$, we have

$$\Delta\epsilon_{z\mathrm{M}}[\mathrm{nm}] = 8.9B_\mathrm{ring}^{-1}[\mathrm{T}]B_{0\mathrm{M}}^3[\mathrm{T}]\beta_{z\mathrm{M}}[\mathrm{m}]L_\mathrm{uM}[\mathrm{m}]. \tag{137}$$

In our example case, $B_\mathrm{ring} = 0.2$ T, $B_{0\mathrm{M}} = 1.13$ T, $\beta_{z\mathrm{M}} = 139\ \mu$m, $L_\mathrm{uM} = 1.6$ m, we have

$$\Delta\epsilon_{z\mathrm{M}} = 14.3\ \mathrm{pm}, \tag{138}$$

which is even larger than the desired 4 pm longitudinal emittance and is therefore unacceptable.

The above evaluation of the longitudinal emittance contribution indicates that a weaker or shorter modulator should be used. Since our desired longitudinal emittance is $\epsilon_z \lesssim 4$ pm, we need to control the contribution from the two modulators to be $\Delta\epsilon_{z\mathrm{M}} \lesssim 1$ pm, since the ring dipoles will also contribute longitudinal emittance with a theoretical minimum of approximately 2 pm. For example, we may choose to weaken the modulator field by a factor of more than two. If $\lambda_\mathrm{uM} = 0.15$ m and $B_{0\mathrm{M}} = 0.435$ T, $L_\mathrm{uM} = 1.5$

m, then the contribution of two modulators to longitudinal emittance is

$$\Delta\epsilon_{z\mathrm{M}} = 0.76\ \mathrm{pm}, \tag{139}$$

which should be acceptable for a target total longitudinal emittance of 4 pm. However, to introduce the desired energy chirp strength, we now need $P_{\mathrm{L}} \approx 2$ GW.

Note that in this updated parameter choice, there is still one issue that needs to be addressed. In the evaluation of the quantum excitation contribution of modulators to the longitudinal emittance, we implicitly assumed that the longitudinal beta function did not change inside it. This is not strictly true. The undulator itself has an $R_{56} = 2N_{\mathrm{u}}\lambda_0$, with λ_0 being the fundamental resonance wavelength of the undulator, which in our case is the modulation laser wavelength. The criterion for whether the thin-lens approximation applies is to evaluate whether or not $|hR_{56}| \ll 1$, where R_{56} is that of the undulator. Here in this updated example, we have $hR_{56} = h2N_{\mathrm{u}}\lambda_{\mathrm{L}} = 0.9$ for the modulator, which means that the thin-lens kick approximation does not apply here. Therefore, more accurately we should use the thick-lens map of the modulator [17] to calculate the evolution of longitudinal beta function in the modulator, and then evaluate the contribution to longitudinal emittance.

With all the subtle points carefully handled, the LSF, as analyzed above, can realize the desired nm-bunch length and thus generate coherent EUV radiation. The main issue with such an EUV source is that the required modulation laser power (GW level) is too high, and the optical enhancement cavity can only work in low-duty-cycle pulsed mode, thus limiting the average EUV output power.

4.3 Generalized longitudinal strong focusing

The previous analysis of the LWF and LSF led us to consider the generalized longitudinal strong focusing (GLSF) scheme [9]. The basic idea of the GLSF is to take advantage of the ultrasmall natural vertical emittance in a planar electron storage ring. More specifically, we will apply a partial transverse-longitudinal emittance exchange in the optical laser wavelength range to achieve efficient microbunching generation. As shown in Fig. 3, the schematic setup of a GLSF ring is very similar to that of an LSF ring. However, as stressed before, the energy chirp strength required in GLSF is much smaller than that in the LSF scheme, which means that the required modulation laser power can also be smaller. A sharp reader may also notice that in Fig. 3, the longitudinal phase space area of the beam is not conserved in the bunch compression or harmonic generation section of a GLSF ring. Fundamental physical laws, such as Liouville's theorem, cannot be violated in a symplectic system. The reason for this apparent "contradiction" is that GLSF invokes 4D or 6D phase space dynamics, as summarized in Table 1, and what is conserved are the eigen-emittances, instead of the projected emittances. It should be noted that in the plot, the phase space rotation direction in the GLSF scheme is reversed after the radiator compared to that before the radiator, whereas this is not the case in the LSF scheme. In other words, in the GLSF, we choose to make the upstream and downstream modulations cancel each other. In this sense, this setup is a special case of the reversible seeding scheme of the SSMB [32]. The reason for this is that we want to make the system transverse-longitudinal coupled only in a limited local region in the ring, the so-called GLSF section, such that we can maintain $\mathcal{H}_y = 0$ being at the majority places of the ring to minimize the quantum excitation and IBS contribution to vertical emittance, thus keeping the small vertical emittance of a planar uncoupled ring. Furthermore, the cancellation of the nonlinear sinusoidal modulation waveforms makes the nonlinear dynamics of the ring easier to handle. To ensure that the modulations perfectly cancel, the lattice between the upstream and downstream modulators must be an isochronous achromat. This reversible seeding setup makes the decoupling of the system straightforward. All we need to do is make the GLSF section an achromat, as the section from the upstream modulator to the downstream modulator is transparent to the longitudinal dynamics. Another advantage of this reversible seeding setup is that it makes the bunch length at the modulator more flexible. It can be a short microbunned beam, as shown in Fig. 3. It can also be a conventional RF-bunched or coasting beam. A coasting beam in this context means the beam is not pre-microbunched, and its length is much longer than the modulation laser wavelength, or even longer than that of an RF bunch. In fact, in our present design, which will be presented in Sect. 8, we use an RF-bunched beam in the ring. Therefore, the third laser modulator of such a GLSF ring is an RF system. Having explained the reason why we choose this reversible seeding setup for the GLSF, we remind the readers that this is not the only possible way to realize the GLSF scheme [9]. For example, a symmetric lattice setup with respect to the radiator is also possible, although the nonlinear dynamics may be challenging.

After this general introduction of the GLSF scheme, we now appreciate in a more physical way why GLSF could be favored compared to LSF, in lowering the required modulation laser power. The key is that the LSF has a contribution of ϵ_z from both the LSF section and the ring dipoles, while the GLSF has only a contribution of ϵ_y from the GLSF section, because $\mathcal{H}_y$ outside the GLSF section is zero, as explained above. This is the key physical argument from the single-particle dynamics perspective of why the GLSF may require a smaller energy modulation than the LSF to realize the same desired bunch length at the radiator. If the longitudinal emittance in the LSF is only from

the quantum excitation of LSF modulators, and the vertical emittance in the GLSF is only from the quantum excitation of GLSF modulators, then GLSF and LSF are equivalent in essence (only a factor of two difference in damping rates) concerning the requirement on energy modulation strength from the single-particle dynamics perspective.

Now, we explain the above argument more clearly using formulas. As we will show in the following section, more specifically Eq. (150), in GLSF at best case we have

$$|h| = \frac{1}{\sqrt{\mathcal{H}_{y\mathrm{R}}\mathcal{H}_{y\mathrm{M}}}} = \frac{\sqrt{\epsilon_y/\mathcal{H}_{y\mathrm{M}}}}{\sqrt{\epsilon_y\mathcal{H}_{y\mathrm{R}}}} = \frac{\sqrt{\epsilon_y/\mathcal{H}_{y\mathrm{M}}}}{\sigma_{z\mathrm{R}}}, \tag{140}$$

where $\mathcal{H}_y = \beta_{55}^{II}$ is defined in Sect. 2 and quantifies the contribution of the vertical emittance to the bunch length. Note that we have used $\sigma_{z\mathrm{R}} = \sqrt{\epsilon_y\mathcal{H}_{y\mathrm{R}}}$, which means that the bunch length at the radiator in the GLSF scheme is solely determined by the beam vertical emittance. One can appreciate the similarity between the above formula and Eq. (127) for the LSF case. Therefore, GLSF will be advantageous to LSF in lowering the required energy modulation strength if

$$\frac{\epsilon_{y,\mathrm{GLSF}}}{\mathcal{H}_{y\mathrm{M,GLSF}}} < \frac{\epsilon_{z,\mathrm{LSF}}}{\beta_{z\mathrm{M,LSF}}}. \tag{141}$$

We now compare the two schemes in a more quantitative manner. We assume that the two schemes operate at the same beam energy. As we will show in Sect. 8.3, in a GLSF SSMB ring and if we consider only single-particle dynamics, the dominant contribution of vertical emittance is from the quantum excitation of two modulators in the GLSF section, and we have

$$\epsilon_{y,\mathrm{GLSF}} \approx C_\mathrm{q}\frac{\gamma^2}{J_y}\frac{1}{I_2}\times 2\frac{\mathcal{H}_{y\mathrm{M,GLSF}}}{\rho_{0\mathrm{M}}^3}\frac{4}{3\pi}L_{u\mathrm{M,GLSF}}. \tag{142}$$

In the LSF, we assume that the longitudinal emittance is mainly from the quantum excitation in dipoles of the ring, and the average longitudinal beta function around the ring dipoles is the same as that at the modulators $\langle\beta_z\rangle \approx \beta_{z\mathrm{M}}$.

Taking the approximation $I_2 \approx \frac{2\pi}{\rho_\mathrm{ring}}$ which means in both rings the radiation loss is mainly from the dipoles in the ring, $J_z \approx 2$, $J_x \approx 1$, and combining with Eqs. (128), (141) then corresponds to

$$4\rho_{\mathrm{ring,GLSF}}\frac{\frac{4}{3\pi}L_{u\mathrm{M,GLSF}}}{\rho_{0\mathrm{M,GLSF}}^3} < \rho_{\mathrm{ring,LSF}}\frac{2\pi\rho_{\mathrm{ring,LSF}}}{\rho_{\mathrm{ring,LSF}}^3}. \tag{143}$$

The above condition should be straightforward to fulfill in practice. For example, if the bending magnet strengths are the same in both schemes, i.e., if $\rho_{\mathrm{ring,GLSF}} = \rho_{\mathrm{ring,LSF}} = \rho_\mathrm{ring}$, the above relation corresponds to

$$L_{u\mathrm{M,GLSF}} < \frac{3\pi^2}{8}\frac{\rho_{0\mathrm{M,GLSF}}^3}{\rho_\mathrm{ring}^2}, \tag{144}$$

which is easily satisfied in practice. Therefore, GLSF can be favored compared to LSF in lowering the required modulation laser power.

4.4 Short summary

From the analysis in this section, our tentative conclusion is that an LWF SSMB ring can be used to generate bunches with a couple of 10 nm-bunch lengths, thus generating coherent visible and infrared radiation. If we want to push the bunch length to an even shorter range, the required phase slippage factor of the ring will be too small from an engineering perspective. An LSF SSMB ring can create bunches with a bunch length at the nanometer level, thus generating coherent EUV radiation. However, the required modulation laser power is at the GW level, and the optical enhancement cavity can only work at a low duty cycle pulsed mode, thus limiting the average output EUV radiation power. At present, a GLSF SSMB ring is the most promising among these three schemes to realize nm-bunch length with a smaller modulation laser power compared to LSF SSMB, thus allowing a higher average power for EUV radiation generation.

5 Transverse-longitudinal coupling for Bunch compression and harmonic generation

In the following sections, we will discuss the GLSF scheme in detail; specifically, we will systematically investigate the backbone of a GLSF SSMB storage ring, the transverse-longitudinal phase space coupling dynamics. As the first step, in this section, we present three theorems or inequalities that dictate such TLC-based bunch compression or harmonic generation schemes. If the initial bunch is longer than the modulation RF or laser wavelength, then the compression of the bunch or microbunch can be viewed as a harmonic

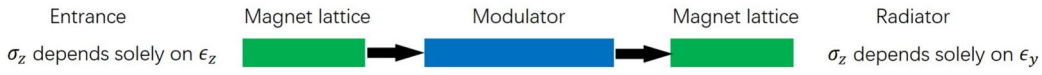

Fig. 5 (Color online) Schematic layout of applying TLC dynamics for bunch compression. (Figure adapted from Ref. [33])

generation scheme. Therefore, in this study, we considered bunch compression and harmonic generation to be the same thing in essence. We remind the readers that the theorems presented here are generalizations of those presented in Refs. [17, 33] from the 4D phase space to the 6D phase space. These formal mathematical relations will be useful in our later detailed study of a GLSF SSMB light source.

5.1 Problem definition

First, let us define the problem that we are trying to solve. We assume ϵ_y is the small eigen-emittance we want to exploit. The case of using ϵ_x is similar. The schematic layout of the TLC-based bunch compression section is shown in Fig. 5. Suppose the beam at the entrance of the bunch compression section is *x*-*y*-*z* decoupled, with its second moments matrix given by

$$\boldsymbol{\Sigma}_i = \langle \mathbf{X}\mathbf{X}^{\mathrm{T}} \rangle_i = \begin{pmatrix} \epsilon_x\beta_{xi} & -\epsilon_x\alpha_{xi} & 0 & 0 & 0 & 0 \\ -\epsilon_x\alpha_{xi} & \epsilon_x\gamma_{xi} & 0 & 0 & 0 & 0 \\ 0 & 0 & \epsilon_y\beta_{yi} & -\epsilon_y\alpha_{yi} & 0 & 0 \\ 0 & 0 & -\epsilon_y\alpha_{yi} & \epsilon_y\gamma_{yi} & 0 & 0 \\ 0 & 0 & 0 & 0 & \epsilon_z\beta_{zi} & -\epsilon_z\alpha_{zi} \\ 0 & 0 & 0 & 0 & -\epsilon_z\alpha_{zi} & \epsilon_z\gamma_{zi} \end{pmatrix}, \tag{145}$$

where α, β, and γ are the Courant–Snyder functions, the subscript i indicates the initial state, and ϵ_x, ϵ_y, and ϵ_z are the eigen-emittances of the beam corresponding to the horizontal, vertical, and longitudinal modes, respectively. The eigen-emittances are beam invariants with respect to linear symplectic transport. For the application of TLC for bunch compression, it means that the final bunch length at the exit or radiator $\sigma_{z\mathrm{R}}$ depends only on the vertical emittance ϵ_y, and neither on the horizontal one ϵ_x and nor on the longitudinal one ϵ_z.

We divide such a bunch compression section into three parts, with their symplectic transfer matrices given by

$$\mathbf{M}_1 = \begin{pmatrix} r_{11} & r_{12} & r_{13} & r_{14} & 0 & r_{16} \\ r_{21} & r_{22} & r_{23} & r_{24} & 0 & r_{26} \\ r_{31} & r_{32} & r_{33} & r_{34} & 0 & r_{36} \\ r_{41} & r_{42} & r_{43} & r_{44} & 0 & r_{46} \\ r_{51} & r_{52} & r_{53} & r_{54} & 1 & r_{56} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\mathbf{M}_2 = \text{modulation kick map}, \tag{146}$$

$$\mathbf{M}_3 = \begin{pmatrix} R_{11} & R_{12} & R_{13} & R_{14} & 0 & R_{16} \\ R_{21} & R_{22} & R_{23} & R_{24} & 0 & R_{26} \\ R_{31} & R_{32} & R_{33} & R_{34} & 0 & R_{36} \\ R_{41} & R_{42} & R_{43} & R_{44} & 0 & R_{46} \\ R_{51} & R_{52} & R_{53} & R_{54} & 1 & R_{56} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

with $\mathbf{M}_1$ representing "from entrance to modulator", $\mathbf{M}_2$ representing "modulation kick" and $\mathbf{M}_3$ representing "modulator to radiator". Note that $\mathbf{M}_1$ and $\mathbf{M}_3$ are in their general thick-lens form and do not necessarily need to be *x*-*y* decoupled. The transfer matrix from the entrance to the radiator is then

$$\mathbf{O} = \mathbf{M}_3\mathbf{M}_2\mathbf{M}_1. \tag{147}$$

From the problem definition, for $\sigma_{z\mathrm{R}}$ to be independent of ϵ_x and ϵ_z, we need

$$O_{51} = 0,\ O_{52} = 0,\ O_{55} = 0,\ O_{56} = 0. \tag{148}$$

5.2 Theorems and proof

5.2.1 Theorems

Given the above problem definition, and assuming that the modulation kick map $\mathbf{M}_2$ is a thin-lens one, we have three theorems that dictate the relation between the modulator kick strength and the optical functions at the modulator and radiator, respectively.

Theorem 1: If

$$\mathbf{M}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & h & 1 \end{pmatrix}, \tag{149}$$

which corresponds to the case of a normal RF or a TEM_{00} mode laser modulator, then

$$h^2\mathcal{H}_{y\mathrm{M}}\mathcal{H}_{y\mathrm{R}} \geq 1, \tag{150}$$

where the subscripts "M" and "R" represent modulator and radiator, respectively.

Theorem 2: If

$$\mathbf{M}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & g & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & g & 0 & 0 & 1 \end{pmatrix}, \tag{151}$$

which corresponds to the case of a transverse deflecting (in *y*-dimension) RF or a TEM_{01} mode laser modulator or other schemes for angular modulation, then

$$g^2\beta_{y\mathrm{M}}\mathcal{H}_{y\mathrm{R}} \geq 1. \tag{152}$$

Theorem 3: If

$$\mathbf{M}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & q & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -q & 0 & 1 \end{pmatrix}, \tag{153}$$

then

$$q^2 \gamma_{y\mathrm{M}} \mathcal{H}_{y\mathrm{R}} \geq 1. \tag{154}$$

There is no commonly used single element realizing the kick map in Eq. (153) directly. Instead, it takes the combination of several elements to realize such a map, making its application less straightforward compared to the cases correspond to Theorem 1 and 2.

5.2.2 Proof

Here, we present the details of the proof of Theorem 1. The proof of the other two is similar. From the problem definition, for $\sigma_{z\mathrm{R}}$ to be independent of ϵ_x and ϵ_z, we need

$$\begin{aligned} O_{51} &= r_{11}R_{51} + r_{21}R_{52} + r_{31}R_{53} + r_{41}R_{54} + r_{51}(hR_{56} + 1) = 0, \\ O_{52} &= r_{12}R_{51} + r_{22}R_{52} + r_{32}R_{53} + r_{42}R_{54} + r_{52}(hR_{56} + 1) = 0, \\ O_{55} &= hR_{56} + 1 = 0, \\ O_{56} &= r_{16}R_{51} + r_{26}R_{52} + r_{36}R_{53} + r_{46}R_{54} + r_{56}(hR_{56} + 1) + R_{56} = 0. \end{aligned} \tag{155}$$

Under the above conditions, we have

$$\mathbf{O} = \begin{pmatrix} \mathbf{A} & \mathbf{B} & \mathbf{C} \\ \mathbf{D} & \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} & \mathbf{I} \end{pmatrix}, \tag{156}$$

with $\mathbf{A} \sim \mathbf{I}$ being 2×2 submatrices of $\mathbf{T}$, and

$$\begin{aligned} \mathbf{G} &= \begin{pmatrix} 0 & 0 \\ r_{51}h & r_{52}h \end{pmatrix}, \\ \mathbf{H} &= \begin{pmatrix} r_{13}R_{51} + r_{23}R_{52} + r_{33}R_{53} + r_{43}R_{54} & r_{14}R_{51} + r_{24}R_{52} + r_{34}R_{53} + r_{44}R_{54} \\ r_{53}h & r_{54}h \end{pmatrix}, \\ \mathbf{I} &= \begin{pmatrix} 0 & 0 \\ h & r_{56}h + 1 \end{pmatrix}. \end{aligned} \tag{157}$$

Note that $\mathbf{I}$ in this subsection does not refer to the identity matrix. The bunch length squared at the modulator and the radiator is obtained as that given in Eq. (60) with the evolution of generalized beta functions properly handled according to Eq. (11) and the results are

$$\begin{aligned} \sigma_{z\mathrm{M}}^2 &= \epsilon_x \frac{\left(\beta_{xi} r_{51} - \alpha_{xi} r_{52}\right)^2 + r_{52}^2}{\beta_{xi}} \\ &\quad + \epsilon_y \frac{\left(\beta_{yi} r_{53} - \alpha_{yi} r_{54}\right)^2 + r_{54}^2}{\beta_{yi}} \\ &\quad + \epsilon_z \left(\beta_{zi} - 2\alpha_{zi} r_{56} + \gamma_{zi} r_{56}^2\right) \\ &= \epsilon_x \mathcal{H}_{x\mathrm{M}} + \epsilon_y \mathcal{H}_{y\mathrm{M}} + \epsilon_z \beta_{z\mathrm{M}}, \\ \sigma_{z\mathrm{R}}^2 &= \epsilon_y \frac{\left(\beta_{yi} O_{53} - \alpha_{yi} O_{54}\right)^2 + O_{54}^2}{\beta_{yi}} = \epsilon_y \mathcal{H}_{y\mathrm{R}}. \end{aligned} \tag{158}$$

According to the Cauchy–Schwarz inequality, we have

$$\begin{aligned} h^2 \mathcal{H}_{y\mathrm{M}} \mathcal{H}_{y\mathrm{R}} &= h^2 \frac{\left[\left(\beta_{yi} r_{53} - \alpha_{yi} r_{54}\right)^2 + r_{54}^2\right]}{\beta_{yi}} \frac{\left[\left(\beta_{yi} O_{53} - \alpha_{yi} O_{54}\right)^2 + O_{54}^2\right]}{\beta_{yi}} \\ &\geq \frac{h^2}{\beta_{yi}^2} \left[-\left(\beta_{yi} r_{53} - \alpha_{yi} r_{54}\right) O_{54} + r_{54}\left(\beta_{yi} O_{53} - \alpha_{yi} O_{54}\right)\right]^2 \\ &= \left(O_{53} r_{54} h - O_{54} r_{53} h\right)^2 = \left(O_{53} O_{64} - O_{54} O_{63}\right)^2 = |\mathrm{Det}(\mathbf{H})|^2, \end{aligned} \tag{159}$$

where Det() denotes the determinant of the matrix. The equality holds when $\frac{-(\beta_{yi} r_{53} - \alpha_{yi} r_{54})}{O_{54}} = \frac{r_{54}}{(\beta_{yi} O_{53} - \alpha_{yi} O_{54})}$. The symplecticity of $\mathbf{O}$ requires that $\mathbf{OSO}^{\mathrm{T}} = \mathbf{S}$, where $\mathbf{S} = \begin{pmatrix} \mathbf{J} & 0 & 0 \\ 0 & \mathbf{J} & 0 \\ 0 & 0 & \mathbf{J} \end{pmatrix}$ and $\mathbf{J} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, so we have

$$\begin{pmatrix} \mathbf{AJA}^{\mathrm{T}} + \mathbf{BJB}^{\mathrm{T}} + \mathbf{CJC}^{\mathrm{T}} & \mathbf{AJD}^{\mathrm{T}} + \mathbf{BJE}^{\mathrm{T}} + \mathbf{CJF}^{\mathrm{T}} & \mathbf{AJG}^{\mathrm{T}} + \mathbf{BJH}^{\mathrm{T}} + \mathbf{CJI}^{\mathrm{T}} \\ \mathbf{DJA}^{\mathrm{T}} + \mathbf{EJB}^{\mathrm{T}} + \mathbf{FJC}^{\mathrm{T}} & \mathbf{DJD}^{\mathrm{T}} + \mathbf{EJE}^{\mathrm{T}} + \mathbf{FJF}^{\mathrm{T}} & \mathbf{DJG}^{\mathrm{T}} + \mathbf{EJH}^{\mathrm{T}} + \mathbf{FJI}^{\mathrm{T}} \\ \mathbf{GJA}^{\mathrm{T}} + \mathbf{HJB}^{\mathrm{T}} + \mathbf{IJC}^{\mathrm{T}} & \mathbf{GJD}^{\mathrm{T}} + \mathbf{HJE}^{\mathrm{T}} + \mathbf{IJF}^{\mathrm{T}} & \mathbf{GJG}^{\mathrm{T}} + \mathbf{HJH}^{\mathrm{T}} + \mathbf{IJI}^{\mathrm{T}} \end{pmatrix} = \mathbf{S}. \tag{160}$$

According to Eq. (157), we have

$$\mathbf{GJG}^{\mathrm{T}} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \ \mathbf{IJI}^{\mathrm{T}} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \tag{161}$$

Therefore,

$$\mathbf{HJH}^{\mathrm{T}} = \mathbf{J}, \tag{162}$$

which means $\mathbf{H}$ is also a symplectic matrix. So, we have $\mathrm{Det}(\mathbf{H}) = 1$. Thus, the theorem is proven.

5.3 Dragt's minimum emittance theorem

Theorem 1 in Eq. (150) can also be expressed as

$$|h| \geq \frac{\epsilon_y}{\sqrt{\epsilon_y \mathcal{H}_{y\mathrm{M}}} \sqrt{\epsilon_y \mathcal{H}_{y\mathrm{R}}}} = \frac{\epsilon_y}{\sigma_{zy\mathrm{M}} \sigma_{z\mathrm{R}}}. \tag{163}$$

Note that in the above formula, $\sigma_{zy\mathrm{M}}$ is the bunch length at the modulator contributed by the vertical emittance ϵ_y. Therefore, given a fixed ϵ_y and desired $\sigma_{z\mathrm{R}}$, a smaller h,

i.e., a smaller RF acceleration gradient or modulation laser power ($P_{\rm L} \propto |h|^2$), means a larger $\mathcal{H}_{y\rm M}$, thus a larger $\sigma_{zy\rm M}$, is needed. As $|h|\sigma_{z\rm M}$ quantifies the energy spread introduced by the modulation kick, we thus also have

$$\sigma_{z\rm R}\sigma_{\delta\rm R} \geq \epsilon_y. \tag{164}$$

Similarly for Theorem 2 and 3, we have

$$|g| \geq \frac{\epsilon_y}{\sigma_{y\beta\rm M}\sigma_{z\rm R}}, \tag{165}$$

and

$$|q| \geq \frac{\epsilon_y}{\sigma_{y'\beta\rm M}\sigma_{z\rm R}}, \tag{166}$$

respectively, and Eq. (164). Note that in the above formulas, the vertical beam size or divergence at the modulator contains only the vertical betatron part, i.e., that from the vertical emittance ϵ_y.

Equation (164) is actually a manifestation of the classical uncertainty principle [34], which states that

$$\begin{aligned}
\Sigma_{11}\Sigma_{22} &\geq \epsilon_{\rm min}^2,\\
\Sigma_{33}\Sigma_{44} &\geq \epsilon_{\rm min}^2,\\
\Sigma_{55}\Sigma_{66} &\geq \epsilon_{\rm min}^2,
\end{aligned} \tag{167}$$

where $\epsilon_{\rm min}$ is the minimum of the three eigen-emittances $\epsilon_{I,II,III}$. In our bunch compression case, we assume that ϵ_y is smaller than ϵ_z.

Actually there is a stronger inequality compared to the classical uncertainty principle, i.e., the minimum emittance theorem [34], which states that the projected emittance cannot be smaller than the minimum one among the three eigen-emittances,

$$\begin{aligned}
\epsilon_{x,\rm pro}^2 &= \Sigma_{11}\Sigma_{22} - \Sigma_{12}^2 \geq \epsilon_{\rm min}^2,\\
\epsilon_{y,\rm pro}^2 &= \Sigma_{33}\Sigma_{44} - \Sigma_{34}^2 \geq \epsilon_{\rm min}^2,\\
\epsilon_{z,\rm pro}^2 &= \Sigma_{55}\Sigma_{66} - \Sigma_{56}^2 \geq \epsilon_{\rm min}^2.
\end{aligned} \tag{168}$$

5.4 Theorems cast in another form

As another way to appreciate the result, here we cast the theorems in a form using the generalized beta functions, as introduced in Sect. 2. According to definition, we have

$$\beta_y \equiv \beta_{33}^{II},\ \gamma_y \equiv \beta_{44}^{II},\ \mathcal{H}_y \equiv \beta_{55}^{II}. \tag{169}$$

Theorem 1: If $\mathbf{M}_2$ is as shown in Eq. (149), then

$$M_{2,65}^2(\rm Mod)\beta_{55}^{II}(\rm Mod)\beta_{55}^{II}(\rm Rad) \geq 1, \tag{170}$$

where $M_{2,65}$ is the $_{65}$ matrix term of $\mathbf{M}_2$, i.e., h. The superscript 2 of $M_{2,65}^2$ means the square of $M_{2,65}$. For better visualization, in this subsection, we use brackets to denote the location, with Ent, Mod, and Rad meaning entrance, modulator, and radiator, respectively.

Theorem 2: If $\mathbf{M}_2$ is as shown in Eq. (151), then

$$M_{2,63}^2(\rm Mod)\beta_{33}^{II}(\rm Mod)\beta_{55}^{II}(\rm Rad) \geq 1. \tag{171}$$

Theorem 3: If $\mathbf{M}_2$ is as shown in Eq. (153), then

$$M_{2,64}^2(\rm Mod)\beta_{44}^{II}(\rm Mod)\beta_{55}^{II}(\rm Rad) \geq 1. \tag{172}$$

At the entrance, the generalized Twiss matrix corresponding to eigenmode I is

$$\mathbf{T}_I(\rm Ent) = \begin{pmatrix}
\beta_{xi} & -\alpha_{xi} & 0 & 0 & 0 & 0\\
-\alpha_{xi} & \gamma_{xi} & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}, \tag{173}$$

and similar expressions for $\mathbf{T}_{II,III}(\rm Ent)$, with x replaced by y, z and the location of the 2×2 matrix shifted in the diagonal direction. Then

$$\beta_{33}^{II}(\rm Mod) = \frac{\left(\beta_{yi}r_{33} - \alpha_{yi}r_{34}\right)^2 + r_{34}^2}{\beta_{yi}}, \tag{174}$$

$$\beta_{44}^{II}(\rm Mod) = \frac{\left(\beta_{yi}r_{43} - \alpha_{yi}r_{44}\right)^2 + r_{44}^2}{\beta_{yi}}, \tag{175}$$

$$\beta_{55}^{II}(\rm Mod) = \frac{\left(\beta_{yi}r_{53} - \alpha_{yi}r_{54}\right)^2 + r_{54}^2}{\beta_{yi}}, \tag{176}$$

$$\beta_{55}^{I}(\rm Rad) = \frac{\left(\beta_{xi}O_{51} - \alpha_{xi}O_{52}\right)^2 + O_{52}^2}{\beta_{xi}}, \tag{177}$$

$$\beta_{55}^{II}(\rm Rad) = \frac{\left(\beta_{yi}O_{53} - \alpha_{yi}O_{54}\right)^2 + O_{54}^2}{\beta_{yi}}, \tag{178}$$

$$\beta_{55}^{III}(\mathrm{Rad}) = \frac{\left(\beta_{zi}O_{55} - \alpha_{zi}O_{56}\right)^2 + O_{56}^2}{\beta_{zi}}. \tag{179}$$

For $\sigma_{z\mathrm{R}}$ to be independent of ϵ_x and ϵ_z, we need $\beta_{55}^{I}(\mathrm{Rad}) = 0$ and $\beta_{55}^{III}(\mathrm{Rad}) = 0$, which then lead to Eq. (148). The following proof procedures are the same as those shown in the Sect. 5.2.2.

6 Energy modulation-based coupling schemes

After introducing the three formal theorems, we conduct a more detailed analysis of the TLC-based bunch compression or microbunching generation schemes. We grouped these schemes into two categories, i.e., energy modulation-based and angular modulation-based schemes. They correspond to the cases of Theorem 1 and 2 presented in the previous section. In this section, we focus on energy modulation-based schemes, and the next section is dedicated to angular modulation-based schemes. The physical realization corresponding to the case of Theorem 3 is not as straightforward as the cases of Theorem 1 and 2, and we do not expand its discussion in this paper.

6.1 Spectral form function and bunching factor

6.1.1 General formula

For coherent radiation generation, a parameter of vital importance is the bunching factor of the electron beam. First, we derive the bunching factor of the energy modulation-based TLC microbunching schemes. The mathematical model was formulated as follows:

6D particle state vector:

$$\mathbf{X} \equiv \left(x \;\; x' \;\; y \;\; y' \;\; z \;\; \delta \right)^{\mathrm{T}}. \tag{180}$$

6D spectral vector:

$$\mathbf{K} \equiv \left(k_x \;\; k_{x'} \;\; k_y \;\; k_{y'} \;\; k_z \;\; k_\delta \right). \tag{181}$$

Normalized particle density distribution in phase space $\psi(\mathbf{X})$:

$$\int \psi(\mathbf{X})\mathrm{d}\mathbf{X} = 1, \;\; \psi(\mathbf{X}) \geq 0, \tag{182}$$

where $\int \mathrm{d}\mathbf{X}$ means $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{d}x\mathrm{d}x'\mathrm{d}y\mathrm{d}y'\mathrm{d}z\mathrm{d}\delta$. Here, we introduce the spectral form function (SFF) of the beam as:

$$\mathcal{F}(\mathbf{K}) \equiv \int \psi(\mathbf{X})e^{-i\mathbf{K}\mathbf{X}}\mathrm{d}\mathbf{X}. \tag{183}$$

$\psi(\mathbf{X})$ and $\mathcal{F}(\mathbf{K})$ then forms a Fourier transform pair

$$\psi(\mathbf{X}) = \frac{1}{(2\pi)^6}\int \mathcal{F}(\mathbf{K})e^{i\mathbf{K}\mathbf{X}}\mathrm{d}\mathbf{K}, \tag{184}$$

where $\int \mathrm{d}\mathbf{K}$ means $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{d}k_x\mathrm{d}k_{x'}\mathrm{d}k_y\mathrm{d}k_{y'}\mathrm{d}k_z\mathrm{d}k_\delta$. This spectral form function is another complete description of the beam distribution and can offer complementary insights for beam dynamics studies. More details on this aspect will be reported in future studies.

The classical 1D bunching factor or form factor used in literature is a specific point in our defined SFF, i.e., with $\mathbf{K} = (0, 0, 0, 0, k_z, 0)$,

$$b(k_z) = \mathcal{F}(0, 0, 0, 0, k_z, 0) = \int_{-\infty}^{\infty}\psi(z)e^{-ik_z z}dz, \tag{185}$$

where $\psi(z) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\psi(\mathbf{X})\mathrm{d}x\mathrm{d}x'\mathrm{d}y\mathrm{d}y'\mathrm{d}\delta$ is the normalized longitudinal distribution of particles. In this paper, we use $\mathcal{F}$ to denote the 6D spectral form function and b the classical 1D bunching factor.

Now, we derive the spectral form function and bunching factor for single-stage energy modulation-based microbunching schemes. A lumped description of the laser-induced energy modulation can be written as:

$$\delta = \delta + A\sin(k_{\mathrm{L}}z), \tag{186}$$

where $k_{\mathrm{L}} = \frac{2\pi}{\lambda_{\mathrm{L}}}$ is the laser wavenumber, and A is the modulation strength. After the modulation, the particle state vector evolves according to:

$$\mathbf{X}_{\mathrm{f}} = \mathbf{R}\mathbf{X}_i, \tag{187}$$

where $\mathbf{R}$ is the linear 6×6 symplectic transfer matrix of the magnet lattice, which can be a single-pass one, such as a linear accelerator, or a multi-pass one, such as a storage ring. In this study, we only considered the case in which the magnet lattice is linear. Denote:

$$\begin{aligned} &\mathbf{A} \equiv (0, 0, 0, 0, 0, A)^{\mathrm{T}}, \\ &\mathbf{U}_p \equiv \left(0 \;\; 0 \;\; 0 \;\; 0 \;\; pk_{\mathrm{L}} \;\; 0 \right), \\ &\mathbf{M}_p \equiv \mathbf{K}\mathbf{R} - \mathbf{U}_p, \end{aligned} \tag{188}$$

with p being an integer. Then the final SFF is

$$\begin{aligned}\mathcal{F}(\mathbf{K}) &= \int \psi_{\mathrm{f}}(\mathbf{X})e^{-i\mathbf{K}\mathbf{X}}\mathrm{d}\mathbf{X}\\ &= \int \psi_{m+}(\mathbf{X})e^{-i\mathbf{K}\mathbf{R}\mathbf{X}}\mathrm{d}\mathbf{X}\\ &= \int \psi_0(\mathbf{X})e^{-i(\mathbf{K}\mathbf{R}\mathbf{X}+\mathbf{K}\mathbf{R}\mathbf{A}\sin(k_{\mathrm{L}}z))}\mathrm{d}\mathbf{X}\\ &= \sum_{p=-\infty}^{\infty} J_p(-\mathbf{K}\mathbf{R}\mathbf{A})\int \psi_0(\mathbf{X})e^{-i(\mathbf{K}\mathbf{R}\mathbf{X}-pk_{\mathrm{L}}z)}\mathrm{d}\mathbf{X}\\ &= \sum_{p=-\infty}^{\infty} J_p(-\mathbf{K}\mathbf{R}\mathbf{A})\int \psi_0(\mathbf{X})e^{-i\mathbf{M}_p\mathbf{X}}\mathrm{d}\mathbf{X}\\ &= \sum_{p=-\infty}^{\infty} J_p(-\mathbf{K}\mathbf{R}\mathbf{A})\mathcal{F}_0(\mathbf{M}_p),\end{aligned} \tag{189}$$

where $\psi_0(\mathbf{X})$, $\psi_{m+}(\mathbf{X})$, and $\psi_{\mathrm{f}}(\mathbf{X})$ are the beam distributions at the beginning, immediately after the energy modulation, and at the final point, respectively. J_p is the p-th order Bessel function of the first type. Jacobi–Anger identity $e^{ix\sin y} = \sum_{n=-\infty}^{\infty} e^{iny}J_n[x]$ has been used in the above derivation. Note that we have also used the fact that for a symplectic matrix $\mathbf{R}$ we have $\mathrm{Det}(\mathbf{R}) = 1$. $\mathcal{F}_0(\mathbf{K})$ is the initial SFF.

The above formula is general and applies to an arbitrary initial beam distribution. If the initial beam distribution is Gaussian in the 6D phase space,

$$\psi_0(\mathbf{X}) = \frac{1}{(2\pi)^3\sqrt{\mathrm{Det}(\boldsymbol{\Sigma}_\mathbf{0})}}\exp\left(-\frac{1}{2}\mathbf{X}^{\mathrm{T}}\boldsymbol{\Sigma}_\mathbf{0}^{-1}\mathbf{X}\right) \tag{190}$$

with $\boldsymbol{\Sigma}_0$ being the initial second moments of the beam, the initial SFF is then

$$\mathcal{F}_0(\mathbf{K}) = \exp\left(-\frac{\mathbf{K}\boldsymbol{\Sigma}_0\mathbf{K}^{\mathrm{T}}}{2}\right). \tag{191}$$

The final SFF is then

$$\mathcal{F}(\mathbf{K}) = \sum_{p=-\infty}^{\infty} J_p(-\mathbf{K}\mathbf{R}\mathbf{A})\exp\left(-\frac{\mathbf{M}_p\boldsymbol{\Sigma}_0\mathbf{M}_p^{\mathrm{T}}}{2}\right), \tag{192}$$

6.1.2 HGHG

In many applications, a classical 1D bunching factor suffices. With $\mathbf{K} = \begin{pmatrix}0&0&0&0&k_z&0\end{pmatrix}$, then

$$\begin{aligned}&\mathbf{K}\mathbf{R}\mathbf{A} = k_zR_{56}A,\\ &\mathbf{M}_p = k_z\begin{pmatrix}R_{51}&R_{52}&R_{53}&R_{54}&R_{55}-\frac{pk_{\mathrm{L}}}{k_z}&R_{56}\end{pmatrix}.\end{aligned} \tag{193}$$

If further $R_{51} = 0$, $R_{52} = 0$, $R_{53} = 0$, $R_{54} = 0$ and $R_{55} = 1$, and the initial beam is transverse-longitudinal decoupled and upright in the longitudinal phase space, which corresponds to the case of high-gain harmonic generation (HGHG) [35], then

$$\begin{aligned}b(k_z) = &\sum_{p=-\infty}^{\infty} J_p\left(-k_zR_{56}A\right)\\ &\exp\left(-\frac{k_z^2}{2}\left[\left(1-\frac{pk_{\mathrm{L}}}{k_z}\right)^2\sigma_{z0}^2+\left(R_{56}\sigma_{\delta0}\right)^2\right]\right),\end{aligned} \tag{194}$$

where σ_{z0} and $\sigma_{\delta0}$ are the initial RMS bunch length and energy spread, respectively. If the initial bunch length is much longer than the laser wavelength, i.e., $k_{\mathrm{L}}\sigma_{z0} \gg 1$, the above exponential terms will be nonzero only when $k_z = pk_{\mathrm{L}}$, which means that there is only bunching at the laser harmonics. In this case, we have the bunching factor at the n-th (n being integer) laser harmonic

$$\begin{aligned}b_n &= b(k_z = nk_{\mathrm{L}})\\ &= J_n(-nk_{\mathrm{L}}R_{56}A)\exp\left[-\frac{\left(nk_{\mathrm{L}}R_{56}\sigma_{\delta0}\right)^2}{2}\right].\end{aligned} \tag{195}$$

For $n > 4$, the maximal value of the Bessel function J_n is about $0.67/n^{1/3}$ and is achieved when its argument is equal to $n + 0.81n^{1/3}$. For large n, this argument corresponds to $k_{\mathrm{L}}R_{56}A \sim 1$. Then, to ensure that the exponential term in Eq. (195) not too small, we need $A \sim n\sigma_{\delta0}$. Therefore, if we want to realize n-th harmonic bunching in HGHG, we need an energy modulation strength a factor of n larger than the initial energy spread.

6.1.3 TLC-based microbunching

Now, let us consider the case of nonzero $R_{51,52,53,54}$, which corresponds to transverse-longitudinal coupling (TLC)-based microbunching [33, 36–40]. Below we use y-z coupling as an example for the analysis and ignore the x-dimension. The analysis for x-z coupling is similar. If $\mathbf{K} \equiv \begin{pmatrix}0&0&0&0&k_z&0\end{pmatrix}$, $R_{51} = 0$, $R_{52} = 0$, $R_{55} = 1$, $R_{66} = 1$, then

$$\begin{aligned}&\mathbf{K}\mathbf{R}\mathbf{A} = k_zR_{56}A,\\ &\mathbf{M}_p = k_z\begin{pmatrix}0&0&R_{53}&R_{54}&1-\frac{pk_{\mathrm{L}}}{k_z}&R_{56}\end{pmatrix}.\end{aligned} \tag{196}$$

Using the real and imaginary generalized beta functions and Twiss matrices introduced in Sect. 2, the bunching factor at the n-th laser harmonic can be expressed as

$$b_n = \sum_{p=-\infty}^{\infty} J_p\left(-nk_{\mathrm{L}}R_{56}A\right)\exp\left[-\frac{\left(\epsilon_{II}\mathbf{M}_p\mathbf{T}_{II}\mathbf{M}_p^{\mathrm{T}}+\epsilon_{III}\mathbf{M}_p\mathbf{T}_{III}\mathbf{M}_p^{\mathrm{T}}\right)}{2}\right], \tag{197}$$

where $\mathbf{T}_{II,III}$ are the real generalized Twiss matrices immediately before the modulation. We then require

$$\mathbf{M}_p\mathbf{T}_{III}\mathbf{M}_p^{\mathrm{T}}\Big|_{p=n} = (nk_{\mathrm{L}})^2\left(\mathbf{R}\mathbf{T}_{III}\mathbf{R}^{\mathrm{T}} - 2\mathbf{R}\mathbf{T}_{III} + \mathbf{T}_{III}\right)_{55} = 0. \tag{198}$$

The physical reason for this requirement is that for $p = n$, we want the longitudinal emittance not to contribute to the bunching factor. By doing this, the bunching factor is mainly determined by the vertical emittance, which is assumed to be small. Therefore, we can realize high-harmonic bunching with a shallow energy modulation strength A. The above relation can be written more specifically using the generalized beta functions and the matrix terms of $\mathbf{R}$ as

$$\begin{aligned} &R_{53}^2\beta_{33}^{III} + R_{54}^2\beta_{44}^{III} + R_{56}^2\beta_{66}^{III} + 2R_{53}R_{54}\beta_{34}^{III} \\ &\quad + 2R_{53}R_{56}\beta_{36}^{III} + 2R_{54}R_{56}\beta_{46}^{III} = 0. \end{aligned} \tag{199}$$

If the generalized Twiss matrices of the eigenmode *II* and *III* right before the modulation are given by Eq. (51), and further assuming that at the modulation point we have $D_x = 0$ and $D_x' = 0$, then Eq. (199) can be casted into

$$R_{53}D_y + R_{54}D_y' + R_{56} = 0. \tag{200}$$

This relation means that the final coordinate z does not depend on the initial energy deviation δ in the linear approximation. Under the above condition, we have

$$\mathbf{M}_p\mathbf{T}_{III}\mathbf{M}_p^{\mathrm{T}} = (nk_{\mathrm{L}})^2\left(1 - \frac{p}{n}\right)^2\beta_{z\mathrm{M}}, \tag{201}$$

and

$$\begin{aligned} \mathbf{M}_p\mathbf{T}_{II}\mathbf{M}_p^{\mathrm{T}} = (nk_{\mathrm{L}})^2\Big[&\mathcal{H}_{y\mathrm{R}} + \left(1 - \frac{p}{n}\right)^2\mathcal{H}_{y\mathrm{M}} \\ &+2\left(1 - \frac{p}{n}\right)\left(\gamma_y R_{54}D_y - \alpha_y R_{53}D_y + \alpha_y R_{54}D_y' - \beta_y R_{53}D_y'\right)\Big]. \end{aligned} \tag{202}$$

Then the bunching factor at the n-th laser harmonic at the final point which in our context means the radiator is

$$\begin{aligned} b_n = &\sum_{p=-\infty}^{\infty} J_p\left(-nk_{\mathrm{L}}R_{56}A\right)\exp\left[-\frac{(nk_{\mathrm{L}})^2}{2}\epsilon_y\mathcal{H}_{y\mathrm{R}}\right] \\ &\exp\left[-\frac{k_{\mathrm{L}}^2}{2}(n-p)^2\left(\epsilon_y\mathcal{H}_{y\mathrm{M}} + \epsilon_z\beta_{z\mathrm{M}}\right)\right] \\ &\exp\left[-\frac{k_{\mathrm{L}}^2}{2}2n(n-p)\epsilon_y(\gamma_y R_{54}D_y - \alpha_y R_{53}D_y + \alpha_y R_{54}D_y' - \beta_y R_{53}D_y')\right]. \end{aligned} \tag{203}$$

Here we remind the readers that there is an extra factor in Eq. (203) compared to the result in our previous publications in Refs. [17, 33], i.e.,

$$\begin{aligned} b_n = &\sum_{p=-\infty}^{\infty} J_p\left(-nk_{\mathrm{L}}R_{56}A\right)\exp\left[-\frac{(nk_{\mathrm{L}})^2}{2}\epsilon_y\mathcal{H}_{y\mathrm{R}}\right] \\ &\exp\left[-\frac{k_{\mathrm{L}}^2}{2}(n-p)^2\left(\epsilon_y\mathcal{H}_{y\mathrm{M}} + \epsilon_z\beta_{z\mathrm{M}}\right)\right]. \end{aligned} \tag{204}$$

We conclude that Eq. (203) here is rigorously more accurate.

If the modulation waveform is linear, according to Eq. (158), the RMS bunch length at the modulator and radiator is given by

$$\begin{aligned} \sigma_{z\mathrm{M}}^2 &= \epsilon_z\beta_{z\mathrm{M}} + \epsilon_y\mathcal{H}_{y\mathrm{M}}, \\ \sigma_{z\mathrm{R}}^2 &= \epsilon_y\mathcal{H}_{y\mathrm{R}}. \end{aligned} \tag{205}$$

So in this paper, we call $\sigma_{z\mathrm{R}} = \sqrt{\epsilon_y\mathcal{H}_{y\mathrm{R}}}$ the linear bunch length at the radiator. In the previous section, we proved that there is a fundamental inequality dictating the energy chirp strength h and $\mathcal{H}_y$ at the modulator and radiator, respectively, i.e., $h^2\mathcal{H}_{y\mathrm{M}}\mathcal{H}_{y\mathrm{R}} \geq 1$. Given the vertical emittance and desired $\sigma_{z\mathrm{R}}$, to lower the energy chirp strength, we need to lengthen the bunch at the modulator. If the initial bunch is shorter than the modulation laser wavelength, considering that the actual laser modulation waveform is sinusoidal, then according to Eq. (203), bunch lengthening at the modulator means a bunching factor drop at the radiator, as can be seen from the second exponential term. For more discussions on this point, the readers are referred to Refs. [17, 33].

When $k_{\mathrm{L}}^2\left(\epsilon_y\mathcal{H}_{y\mathrm{M}} + \epsilon_z\beta_{z\mathrm{M}}\right) \gg 1$ which means that the bunch length at the modulation point is much longer than the laser wavelength, there is only one non-vanishing term in the above summation, i.e., the term with $p = n$. Then

$$b_n = J_n\left(-nk_{\mathrm{L}}R_{56}A\right)\exp\left[-\frac{(nk_{\mathrm{L}})^2}{2}\epsilon_y\mathcal{H}_{y\mathrm{R}}\right]. \tag{206}$$

Here, we make a short comment that our derivation of the bunching factor, and also most of that found in the literature, neglects the collective interactions between the electron beam and the co-propagating electromagnetic field. Such collective interactions may disturb the modulation performance. Although not much work on this subject, the interested readers may refer to a recent relevant work [41].

6.2 Modulation strength

After deriving the bunching factor, we derived the formula for the modulation strength, given the laser, electron, and undulator parameters. This is necessary for quantitative analysis and comparison.

6.2.1 A normally incident laser

The most common method of imprinting energy modulation on an electron beam at the laser wavelength is to use a TEM_{00} mode laser that resonates with the electrons in an undulator. In this study, we used a planar undulator as the modulator. A helical undulator can also be applied for energy modulation; however, because we want to preserve the ultrasmall vertical emittance, we need to avoid x-y coupling as much as possible. Thus, a planar undulator might be preferred. The electromagnetic field of a TEM_{00} mode Gaussian laser polarized in the horizontal plane is [36]

$$\begin{pmatrix} E_x \\ E_y \\ E_z \\ cB_x \\ cB_y \\ cB_z \end{pmatrix} = E_{x0} e^{ik_\mathrm{L} s - i\omega_\mathrm{L} t + i\phi_0} \left(-iZ_\mathrm{R} Q\right) \exp\left[i\frac{k_\mathrm{L} Q}{2}\left(x^2+y^2\right)\right] \begin{pmatrix} 1 \\ 0 \\ -Qx \\ -Q^2 xy \\ Q^2x^2 - \frac{iQ}{k_\mathrm{L}} + 1 \\ -Qy \end{pmatrix}, \tag{207}$$

where $E_{x,y,z}$ and $B_{x,y,z}$ are the horizontal, vertical and longitudinal electric and magnetic field, respectively, s is the longitudinal global path length variable, t is the time variable, c is the speed of light in free space, $\omega_L = k_\mathrm{L} c$, $Z_\mathrm{R} = \pi w_0^2/\lambda_\mathrm{L}$ is the Rayleigh length, w_0 the beam waist radius, and

$$Q = \frac{i}{Z_\mathrm{R}\left(1 + i\frac{s}{Z_\mathrm{R}}\right)} \tag{208}$$

with i here being the imaginary unit. We remind the readers that it has been implicitly assumed that all the fields will take the real part of their complex expressions.

The relationship between the peak electric field E_{x0} and the laser peak power P_L is given by

$$P_\mathrm{L} = \frac{E_{x0}^2 Z_\mathrm{R} \lambda_\mathrm{L}}{4Z_0}, \tag{209}$$

where $Z_0 = 376.73\ \Omega$ is the impedance of free space. The prescribed wiggling motion of electron in a planar undulator is

$$x(s) = \frac{K}{\gamma k_\mathrm{u}} \sin(k_\mathrm{u} s), \tag{210}$$

with γ being the Lorentz factor,

$$K = \frac{eB_0 \lambda_\mathrm{u}}{2\pi m_\mathrm{e} c} = 0.934 \cdot B_0[\mathrm{T}] \cdot \lambda_\mathrm{u}[\mathrm{cm}] \tag{211}$$

being the dimensionless undulator parameter, where B_0 is the peak magnetic field, λ_u is the undulator period, and $k_\mathrm{u} = 2\pi/\lambda_\mathrm{u}$ being the undulator wavenumber. The resonant condition of laser–electron interaction inside a planar undulator is

$$\lambda_\mathrm{L} = \frac{1 + \frac{K^2}{2}}{2\gamma^2} \lambda_\mathrm{u}. \tag{212}$$

From the prescribed motion, we can calculate the electron horizontal and longitudinal velocity

$$\begin{aligned} v_x(s) &\approx \frac{\beta c K}{\gamma} \cos(k_\mathrm{u} s), \\ v_z(s) &= \sqrt{v^2 - v_x^2} \approx \bar{v}_z - \frac{cK^2}{4\gamma^2} \cos(2k_\mathrm{u} s), \end{aligned} \tag{213}$$

with $\beta = \sqrt{1 - \frac{1}{\gamma^2}}$, and

$$\bar{v}_z = c\left(1 - \frac{1 + K^2/2}{2\gamma^2}\right) \tag{214}$$

the average longitudinal velocity of electrons in the undulator. Therefore, we have the longitudinal path length of electron as a function of time given by

$$s(t) \approx \bar{v}_z t - \frac{K^2}{8\gamma^2 k_\mathrm{u}} \sin(2k_\mathrm{u} \bar{v}_z t). \tag{215}$$

Then

$$k_\mathrm{L} s - \omega_\mathrm{L} t \approx -k_\mathrm{u} s - \chi \sin(2k_\mathrm{u} s), \tag{216}$$

where

$$\chi = \frac{K^2 k_\mathrm{L}}{8\gamma^2 k_\mathrm{u}} = \frac{K^2}{4 + 2K^2}. \tag{217}$$

Note that because the longitudinal coordinate of the electron will affect the observed laser phase, we need to calculate its precision to the order of $\frac{1}{\gamma^2}$, while for the horizontal coordinate x, we only need to calculate it to the order of $\frac{1}{\gamma}$. In the following, we will adopt the approximation $\beta \approx 1$ because we are interested in relativistic cases.

Given the electron prescribed motion and laser electric field, the laser and electron exchange energy according to

$$\frac{\mathrm{d}\mathcal{W}}{\mathrm{d}t} = -\left(ev_x E_x + ev_z E_z\right), \tag{218}$$

where $\mathcal{W}$ is the work done on the electron by the laser, or the energy transfer from the laser to the electron. Note that in this manuscript e represents the elementary charge and is

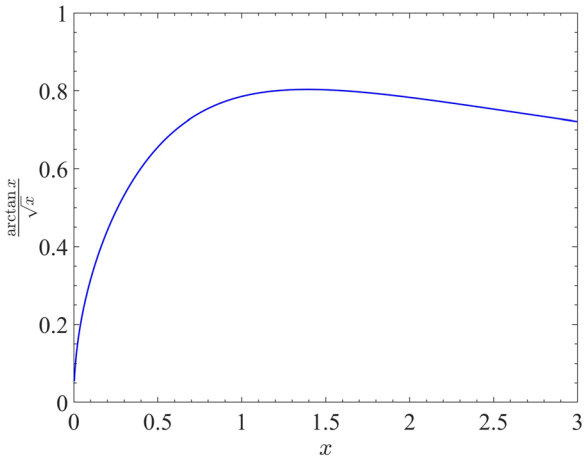

Fig. 6 $f(x) = \frac{\tan^{-1}(x)}{\sqrt{x}}$ vs. x. This curve describes the laser-induced energy modulation strength as a function of the laser Rayleigh length, with the laser power and modulator length given

assumed to be positive. Assuming that the laser beam waist is in the middle of the undulator, whose length is L_u, and when the electron transverse coordinates are much smaller than the laser beam waist $x, y \ll w_0$, which is true in most of the cases under consideration, we drop the factor $\exp\left[i\frac{k_L Q}{2}(x^2+y^2)\right]$ in the laser electric field. Furthermore, when the transverse displacement of the electron is much smaller than the Rayleigh length $x \ll Z_R$ which is also usually the case, we can also drop the contribution from E_z on the energy modulation. Assuming the relative phase of laser field to the electron horizontal velocity v_x at the undulator center is ϕ_0, the integrated modulation voltage induced by the laser on the electron beam in a planar undulator is then

$$\begin{aligned} V_L &\approx \mathrm{Re}\left[\int_{-\frac{L_u}{2}}^{\frac{L_u}{2}} v_x E_x \frac{\mathrm{d}s}{c}\right] \\ &\approx E_{x0}\frac{K}{\gamma}\mathrm{Re}\left[e^{i\phi_0}\int_{-\frac{L_u}{2}}^{\frac{L_u}{2}} \frac{1}{1+i\frac{s}{Z_R}} \sum_{n=-\infty}^{\infty} J_n(-\chi) e^{in2k_u s}\frac{1+e^{-i2k_u s}}{2}\mathrm{d}s\right], \end{aligned} \tag{219}$$

where Re() denotes the real component of the complex number. When $L_u \gg \lambda_u$, which means the undulator period number $N_u \gg 1$, and when $Z_R \gg \lambda_u$, in the above integration, only the term with $n = 0$ and $n = 1$ will give a notable non-vanishing value. Denote $[JJ] \equiv J_0(\chi) - J_1(\chi)$, we then have

$$\begin{aligned} V_L &= E_{x0}\frac{K}{\gamma}\frac{[JJ]}{2}\mathrm{Re}\left[e^{i\phi_0}\int_{-\frac{L_u}{2}}^{\frac{L_u}{2}} \frac{1-i\frac{s}{Z_R}}{1+\left(\frac{s}{Z_R}\right)^2}\mathrm{d}s\right] \\ &= E_{x0}\frac{K}{\gamma}[JJ] Z_R \tan^{-1}\left(\frac{L_u}{2Z_R}\right)\cos\phi_0. \end{aligned} \tag{220}$$

We want the energy modulation strength to be as large as possible; therefore, we choose $\phi_0 = 0$. Putting in the expression of peak electric field from Eq. (209), the linear energy chirp strength around the zero-crossing phase is therefore

$$h = \frac{eV_L}{E_0}k_L = \frac{ek_L K[JJ]}{\gamma^2 m_e c^2}\sqrt{\frac{2P_L Z_0}{\lambda_L}}\frac{\tan^{-1}\left(\frac{L_u}{2Z_R}\right)}{\sqrt{\frac{L_u}{2Z_R}}}\sqrt{L_u}. \tag{221}$$

Once the modulator length is given, we can optimize the laser Rayleigh length to maximize energy modulation. Figure 6 is a plot of $f(x) = \frac{\tan^{-1}(x)}{\sqrt{x}}$ as a function of x. The maximum value of $f(x)$ is 0.8034 and is realized when $x = 1.392$. So, when $Z_R = \frac{L_u}{2\times 1.392} = 0.359L_u$, the energy modulation reaches the maximum value. Note that when Z_R is within a small range close to the optimal value, the impact of Rayleigh length on the energy modulation strength is not very sensitive. Therefore, for easy of remembering, the optimal condition can be expressed as

$$Z_R \approx \frac{L_u}{3}. \tag{222}$$

From Eq. (221), we can see that under the optimal condition, we have $h \propto \sqrt{P_L L_u}$. Using Eq. (221), we can also perform some example calculations to obtain a more concrete feeling. If $E_0 = 600$ MeV, $\lambda_L = 1064$ nm, $\lambda_u = 8$ cm ($B_0 = 1.13$ T), $L_u = 0.8$ m ($N_u = 10$), $Z_R = 0.359L_u$, then the induced energy chirp strength with $P_L = 1$ MW is $h = 955$ m^{-1}.

6.2.2 Dual-tilted laser for energy modulation

Now with a hope to increase the energy modulation strength with a given laser power, we can use a configuration of crossing two lasers for energy modulation. The basic idea is that if two crossing lasers can double the energy modulation strength of a single laser, then the effect is similar to that induced by a single laser with a laser power four times larger. Our calculation shows that indeed dual-tilted laser (DTL) can induce a larger energy modulation compared to that of a single laser, but the issue is that the required crossing angle (less than 2 mrad) is too small from an engineering viewpoint. We remind the readers that we can boost the bunching factor and thus the radiation power by adding one or more high-harmonic modulation in addition to the fundamental frequency modulation, as discussed in Ref. [33]. Here in this paper we focus on the case of using a fundamental frequency modulation only.

The analysis is presented in the following sections. First, we consider the case of two lasers crossing in the y-z plane. The laser field of an oblique TEM_{00} laser is given

by first replacing the physical coordinate with the rotated coordinates

$$x \to x,\ y \to y_1 = y\cos\theta + s\sin\theta,\ s \to s_1 = -y\sin\theta + s\cos\theta,\ t \to t, \tag{223}$$

with θ in this section and below being the tilted angle of the incident laser instead of the dipole bending angle in Sect. 3. We hope their difference is clear from context. The resulting field expression according to Eq. (207) is

$$\begin{pmatrix} E_x \\ E_y \\ E_z \\ cB_x \\ cB_y \\ cB_z \end{pmatrix}_{\text{rot}} = E_{x0} e^{ik_{\text{L}} s_1 - i\omega_{\text{L}} t} \frac{\exp\left[-\frac{k_{\text{L}}(x^2+y_1^2)}{2(Z_{\text{R}}+is_1)}\right]}{1+i\frac{s_1}{Z_{\text{R}}}} \begin{pmatrix} 1 \\ 0 \\ -i\frac{x}{Z_{\text{R}}+is_1} \\ \frac{xy_1}{(Z_{\text{R}}+is_1)^2} \\ -\frac{x^2}{(Z_{\text{R}}+is_1)^2} - \frac{1}{k_{\text{L}}(Z_{\text{R}}+is_1)} + 1 \\ -i\frac{y_1}{Z_{\text{R}}+is_1} \end{pmatrix}. \tag{224}$$

However, in the above expression, x, y, s of the electromagnetic fields are defined according to the oblique laser propagation direction. To get the expression back in the original coordinate system, i.e., undulator axis as the z axis, we need to rotate the laser field as,

$$E_x \to E_{x\text{rot}},\ E_y \to E_y = E_{y\text{rot}}\cos\theta - E_{z\text{rot}}\sin\theta,\ E_z \to E_z = E_{y\text{rot}}\sin\theta + E_{z\text{rot}}\cos\theta,\ t \to t, \tag{225}$$

and for the electric field the result is

$$\begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix}_{\text{unrot}} = E_{x0} e^{ik_{\text{L}} s_1 - i\omega_{\text{L}} t} \frac{\exp\left[-\frac{k_{\text{L}}(x^2+y_1^2)}{2(Z_{\text{R}}+is_1)}\right]}{1+i\frac{s_1}{Z_{\text{R}}}} \begin{pmatrix} 1 \\ -i\frac{x}{Z_{\text{R}}+is_1}\sin\theta \\ -i\frac{x}{Z_{\text{R}}+is_1}\cos\theta \end{pmatrix}. \tag{226}$$

Assume that the two crossing lasers are in phase and have the same amplitude. In addition, we assume that $\theta_2 = -\theta_1 = -\theta$, and that the two lasers have the same Rayleigh length. Then the superimposed field is

$$\begin{aligned} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix}_{\text{unrot}} &= E_{x0} e^{ik_{\text{L}}(-y\sin\theta+s\cos\theta) - i\omega_{\text{L}} t} \frac{\exp\left[-\frac{k_{\text{L}}(x^2+(y\cos\theta+s\sin\theta)^2)}{2(Z_{\text{R}}+i(-y\sin\theta+s\cos\theta))}\right]}{1+i\frac{(-y\sin\theta+s\cos\theta)}{Z_{\text{R}}}} \begin{pmatrix} 1 \\ -i\frac{x}{Z_{\text{R}}+i(-y\sin\theta+s\cos\theta)}\sin\theta \\ -i\frac{x}{Z_{\text{R}}+i(-y\sin\theta+s\cos\theta)}\cos\theta \end{pmatrix} \\ &+ E_{x0} e^{ik_{\text{L}}(y\sin\theta+s\cos\theta) - i\omega_{\text{L}} t} \frac{\exp\left[-\frac{k_{\text{L}}(x^2+(y\cos\theta-s\sin\theta)^2)}{2(Z_{\text{R}}+i(y\sin\theta+s\cos\theta))}\right]}{1+i\frac{(y\sin\theta+s\cos\theta)}{Z_{\text{R}}}} \begin{pmatrix} 1 \\ i\frac{x}{Z_{\text{R}}+i(y\sin\theta+s\cos\theta)}\sin\theta \\ -i\frac{x}{Z_{\text{R}}+i(y\sin\theta+s\cos\theta)}\cos\theta \end{pmatrix}. \end{aligned} \tag{227}$$

As before, we focus on the impact of E_x on the laser-electron interaction and ignore the contribution from E_z. When θ is very small, the superimposed E_x can be approximated as

$$\begin{aligned} E_x &= 2E_{x0} \frac{\exp\left[ik_{\text{L}} s\cos\theta - i\omega t - \frac{k_{\text{L}}(x^2+s^2\sin^2\theta)}{2(Z_{\text{R}}+is\cos\theta)}\right]}{1+\frac{is\cos\theta}{Z_{\text{R}}}} \\ &\approx 2E_{x0} e^{ik_{\text{L}} s\cos\theta - i\omega_{\text{L}} t} \frac{\exp\left[-\frac{k_{\text{L}} s^2\theta^2}{2(Z_{\text{R}}+is)}\right]}{1+i\frac{s}{Z_{\text{R}}}}. \end{aligned} \tag{228}$$

Note that in the final approximated expression, we have retained $s\cos\theta$ in the laser phase term. This is because the laser phase is of key importance in laser–electron interactions, and the accuracy requirement is high. In addition, we have also kept the $s^2\theta^2$ in the intensity decay term, because $L_{\text{u}}\theta$ may not be small compared to the laser beam wait radius w_0. The expression for E_x in the case of crossing in x–z plane is similar. Therefore, the difference in crossing in x–z plane and y–z plane is not significant in inducing energy modulation.

For effective laser–electron interaction, the off-axis resonance condition now is

$$c\frac{\lambda_{\text{u}}}{\bar{v}_z} - \lambda_{\text{u}}\cos\theta = \lambda_{\text{L}}, \tag{229}$$

or

$$\begin{aligned} \lambda_{\text{L}} &= \left(\frac{1}{1-\frac{1+\frac{K^2}{2}}{2\gamma^2}} - \cos\theta\right)\lambda_{\text{u}} \\ &\approx \frac{1+\frac{K^2}{2}+\gamma^2\theta^2}{2\gamma^2}\lambda_{\text{u}}. \end{aligned} \tag{230}$$

Then

$$k_{\text{L}} s - \omega_{\text{L}} t \approx -k_{\text{u}} s - \chi\sin(2k_{\text{u}} s). \tag{231}$$

Note that χ now depends on θ, more specifically,

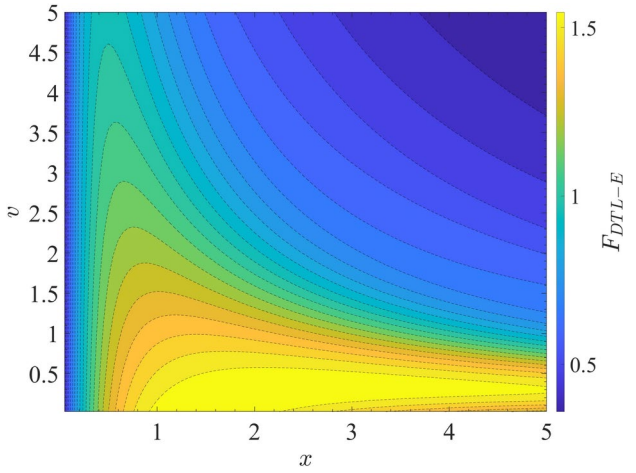

Fig. 7 (Color online) Contour plot of $F_{\text{DTL-E}}(x, v)$ given by Eq. (236)

$$\chi = \frac{K^2 k_\text{L}}{8\gamma^2 k_\text{u}} \approx \frac{K^2}{4 + 2K^2 + 4\gamma^2\theta^2}. \tag{232}$$

Assume that the laser beam waists are in the middle of the undulator, whose length is L_u, and denote

$$v \equiv \frac{k_\text{L} Z_\text{R} \theta^2}{2} = \left(\frac{Z_\text{R}\theta}{w_0}\right)^2, \tag{233}$$

the integrated modulation voltage induced by the DTL in a planar undulator is then

$$\begin{aligned} V_\text{L} &\approx \text{Re}\left[\int_{-\frac{L_\text{u}}{2}}^{\frac{L_\text{u}}{2}} v_x E_x \frac{\text{d}s}{c}\right] \\ &\approx E_{x0}\frac{K[JJ]}{\gamma} Z_\text{R} \int_{-\frac{L_\text{u}}{2Z_\text{R}}}^{\frac{L_\text{u}}{2Z_\text{R}}} \frac{\exp\left(-v\frac{u^2}{1+u^2}\right)}{1+u^2} \\ &\quad \left[\cos\left(v\frac{u^3}{1+u^2}\right) + u\sin\left(v\frac{u^3}{1+u^2}\right)\right]\text{d}u\cos\phi_0. \end{aligned} \tag{234}$$

To maximize the energy modulation, we chose $\phi_0 = 0$. Putting in the expression of E_{x0} from Eq. (209), the linear energy chirp strength around the zero-crossing phase is therefore

$$h = \frac{eV_\text{L}}{E_0} k_\text{L} = \frac{e k_\text{L} K[JJ]}{\gamma^2 m_\text{e} c^2}\sqrt{\frac{2P_\text{L} Z_0}{\lambda_\text{L}}} F_{\text{DTL-E}}\left(\frac{L_\text{u}}{2Z_\text{R}}, \frac{k_\text{L} Z_\text{R}\theta^2}{2}\right)\sqrt{L_\text{u}}, \tag{235}$$

where

$$\begin{aligned} F_{\text{DTL-E}}(x, v) = &\frac{1}{\sqrt{x}}\int_{-x}^{x} \frac{\exp\left(-v\frac{u^2}{1+u^2}\right)}{1+u^2} \\ &\left[\cos\left(v\frac{u^3}{1+u^2}\right) + u\sin\left(v\frac{u^3}{1+u^2}\right)\right]\text{d}u. \end{aligned} \tag{236}$$

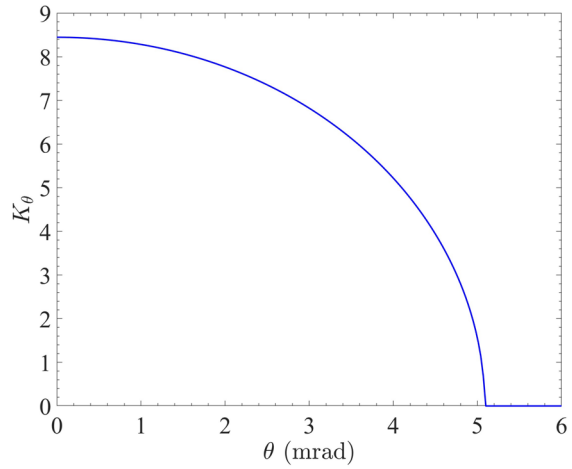

Fig. 8 Undulator parameter K_θ vs. θ with λ_u fixed. Parameters used: $E_0 = 600$ MeV, $\lambda_\text{L} = 1064$ nm, $\lambda_\text{u} = 8$ cm

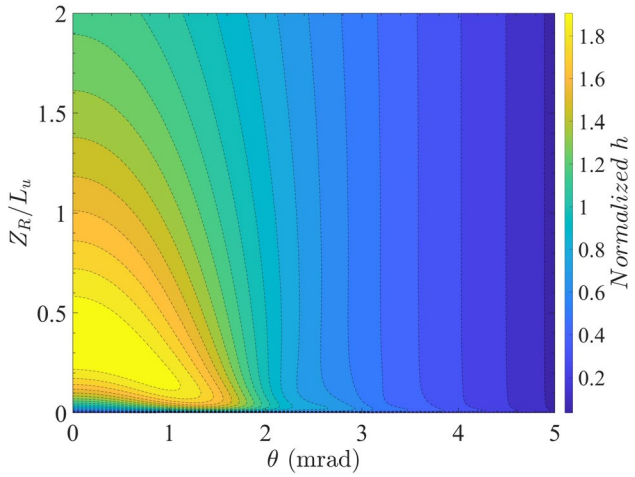

Fig. 9 (Color online) Energy chirp strength h normalized by the largest energy chirp induced by a single normally incident laser vs. θ and Z_R/L_u. Keep λ_u fixed when changing θ. Parameters used: $E_0 = 600$ MeV, $\lambda_\text{L} = 1064$ nm, $\lambda_\text{u} = 0.08$ m, $N_\text{u} = 10$, $L_\text{u} = 0.8$ m

A flat contour plot of $F_{\text{DTL-E}}(x, v)$ is shown in Fig. 7.

Now, we can use the derived formula to calculate the energy chirp strength induced by the DTL. First we consider the case of keeping λ_u fixed when changing θ, then the off-axis resonant condition leads to the undulator parameter as a function of θ given by

$$K_\theta = \sqrt{2}\sqrt{2\gamma^2\left(1 - \frac{1}{\frac{\lambda_\text{L}}{\lambda_\text{u}} + \cos\theta}\right) - 1}. \tag{237}$$

With the increase of θ, K_θ decreases. Note that in this case, $K_\theta \propto B_{0\theta}$, so the magnetic field strength $B_{0\theta}$ also decreases

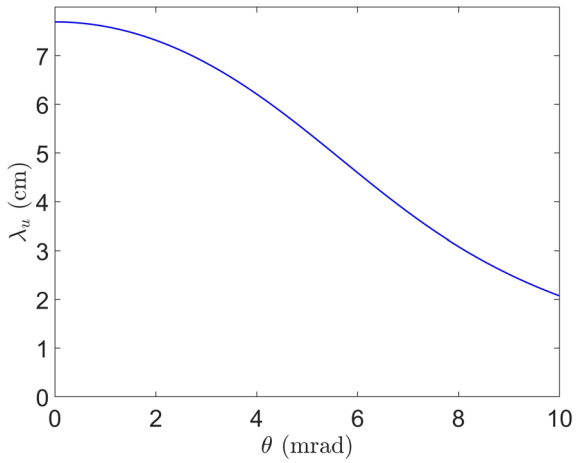

Fig. 10 λ_u vs. θ, with B_0 fixed. Parameters used: $E_0 = 600$ MeV, $\lambda_L = 1064$ nm, $B_0 = 1.2$ T

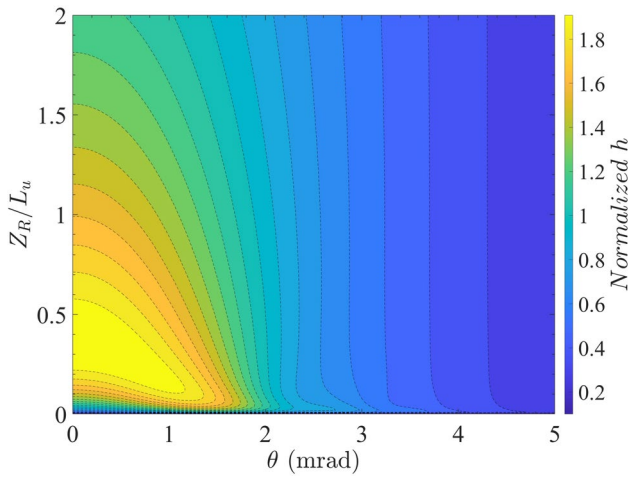

Fig. 11 (Color online) Energy chirp strength h normalized by the largest energy chirp induced by a single normally incident laser vs. θ and Z_R/L_u. Keep B_0 fixed when changing θ. Parameters used: $E_0 = 600$ MeV, $\lambda_L = 1064$ nm, $B_0 = 1.2$ T, $L_u = 0.8$ m

with the increase of θ. An example calculation of K_θ versus θ is shown in Fig. 8. The corresponding contour plot of the energy chirp strength normalized by the largest energy chirp induced by a single normally incident laser, versus θ and Z_R/L_u, is given in Fig. 9.

The previous calculation assumes λ_u remains unchanged when we adjust the incident angle θ. This will result in a limited region of θ to fulfill the resonant condition, as shown in Fig. 8. We now conduct the calculation by assuming that the peak magnetic field B_0 remains unchanged when we adjust θ. Putting the expression of undulator parameter Eq. (211) in the off-axis resonant condition Eq. (230), we have

$$\frac{1}{2}\left(\frac{eB_0}{2\pi m_e c}\right)^2 \lambda_u^3 + (1+\gamma^2\theta^2)\lambda_u - 2\gamma^2\lambda_L = 0, \tag{238}$$

from which we get

$$\lambda_u = \frac{2\pi^{2/3}\sqrt[3]{\mathcal{D}}}{3^{2/3}B_0^2 e^2} - \frac{4\pi^{4/3}c^2\left(\gamma^2\theta^2+1\right)m_e^2}{\sqrt[3]{3}\sqrt[3]{\mathcal{D}}}, \tag{239}$$

with

$$\mathcal{D} = \sqrt{3B_0^6c^4e^6m_e^4\left(27B_0^2\gamma^4e^2\lambda_L^2 + 8\pi^2c^2\left(\gamma^2\theta^2+1\right)^3m_e^2\right)} + 9B_0^4c^2\gamma^2e^4m_e^2\lambda_L. \tag{240}$$

For example, if $\lambda_L = 1064$ nm and $B_0 = 1.2$ T, then λ_u as a function of θ for the case of $E_0 = 600$ MeV is shown in Fig. 10. Note that in this case, $K_\theta \propto \lambda_u$ which depends on θ, so the undulator parameter also decreases with the increase of θ. However, the decrease was not as rapid as that presented in Fig. 8. The corresponding contour plot of the energy chirp strength induced by the DTL, normalized by the largest energy chirp induced by a single normally incident laser, vs. θ and Z_R/L_u, is shown in Fig. 11.

As shown in the calculation results in Figs. 9 and 11, DTL indeed can induce a larger energy modulation compared to a single normally incident laser. However, the required crossing angle (less than 2 mrad) is too small for engineering. Therefore, the usual setup of a single normally incident TEM_{00} mode laser remains the preferred choice in practical applications.

6.3 Realization examples

After the derivation of the bunching factor and laser-induced modulation strengths, finally in this section we provide some realization examples of microbunching schemes that belong to what we have analyzed. FEL seeding technique like phase-merging enhanced harmonic generation (PEHG) [37, 38] and angular dispersion-induced microbunching (ADM) [39] can be viewed as specific examples of our general definition of TLC-based microbunching schemes in Theorem 1. A more detailed discussion in this respect has been presented in Ref. [33].

Here, we briefly comment on the relation between FEL seeding techniques, such as HGHG, PEHG, and ADM, and the storage ring schemes, such as LSF and GLSF, discussed in this paper. One is single pass, and the other is multipass. One invokes matrix multiplication or nonlinear transfer map once, and the other invokes eigenanalysis or normal form analysis of the one-turn map. They share the bunch compression or harmonic generation mechanism. The relationship between HGHG and LSF was similar to that between PEHG/ADM and GLSF.

We also remind the readers that the GLSF SSMB scheme analyzed in this paper bears similarity to the approach

discussed in Ref. [40], in the transverse-longitudinal-coupling-based microbunching and modulation-demodulation processes. The difference is that here we aim for a true steady state in the context of storage ring dynamics and the electron beam pass the radiator turn by turn, while in Ref. [40] the radiator is placed at a bypass line and the ring is used to prepare the electron beam which is sent to the bypass line every multiple revolutions in the ring.

Note that we must add the second equation above to make the modulation symplectic. Such an angular modulation process also satisfies the Panofsky–Wenzel theorem [42]

$$\frac{\partial \Delta y'}{\partial z} = \frac{\partial \Delta \delta}{\partial y}. \tag{242}$$

Following the derivations in the section of energy modulation-based schemes, the final spectral form function in this case is

$$\begin{aligned}
\mathcal{F}(\mathbf{K}) &= \int \psi_{\mathrm{f}}(\mathbf{X}) e^{-i\mathbf{K}\mathbf{X}} \mathrm{d}\mathbf{X} \\
&= \int \psi_{m+}(\mathbf{X}) e^{-i\mathbf{K}\mathbf{R}\mathbf{X}} \mathrm{d}\mathbf{X} \\
&= \int \psi_0(\mathbf{X}) e^{-i\left(\mathbf{K}\mathbf{R}\mathbf{X} + \mathbf{K}\mathbf{R}_{i4} A \sin(k_{\mathrm{L}} z) + \mathbf{K}\mathbf{R}_{i6} A k_{\mathrm{L}} y \cos(k_{\mathrm{L}} z)\right)} \mathrm{d}\mathbf{X} \\
&= \sum_{p_1=-\infty}^{\infty} J_{p_1}\left(-\mathbf{K}\mathbf{R}_{i4} A\right) \int \psi_0(\mathbf{X}) e^{-i\left(\mathbf{K}\mathbf{R}\mathbf{X} - p_1 k_{\mathrm{L}} z + \mathbf{K}\mathbf{R}_{i6} A k_{\mathrm{L}} y \cos(k_{\mathrm{L}} z)\right)} \mathrm{d}\mathbf{X} \\
&= \sum_{p_1=-\infty}^{\infty} J_{p_1}\left(-\mathbf{K}\mathbf{R}_{i4} A\right) \int \psi_0(\mathbf{X}) \sum_{p_2=-\infty}^{\infty} i^{p_2} J_{p_2}\left(-\mathbf{K}\mathbf{R}_{i6} A k_{\mathrm{L}} y\right) e^{-i\left[\mathbf{K}\mathbf{R}\mathbf{X} - (p_1+p_2) k_{\mathrm{L}} z\right]} \mathrm{d}\mathbf{X}.
\end{aligned} \tag{243}$$

7 Angular modulation-based coupling schemes

7.1 Bunching factor

After investigating the energy modulation-based TLC microbunching schemes, in this section, we discuss angular modulation-based schemes [42–48]. The problem definition is similar to that in energy modulation-based schemes, with the exception of replacing energy modulation with angular modulation. We use y' modulation as an example because we will take advantage of the ultrasmall vertical emittance in a planar ring, as explained before. The lumped laser-induced angular modulation is modeled as:

$$\begin{aligned}
y' &= y' + A \sin(k_{\mathrm{L}} z), \\
\delta &= \delta + A k_{\mathrm{L}} y \cos(k_{\mathrm{L}} z).
\end{aligned} \tag{241}$$

The integration of the above results is given by the generalized hypergeometric function. If $\mathbf{K} = \begin{pmatrix} 0 & 0 & 0 & 0 & k_z & 0 \end{pmatrix}$ and $R_{56} = 0$, then $\mathbf{K}\mathbf{R}_{i6} = 0$, and the initial beam distribution is Gaussian as given in Eq. (190), then the 1D bunching factor is

$$\begin{aligned}
b(k_z) &= \sum_{p=-\infty}^{\infty} J_p\left(-\mathbf{K}\mathbf{R}_{i4} A\right) \int \psi_0(\mathbf{X}) e^{-i\left(\mathbf{K}\mathbf{R}\mathbf{X} - p k_{\mathrm{L}} z\right)} \mathrm{d}\mathbf{X} \\
&= \sum_{p=-\infty}^{\infty} J_p\left(-k_z R_{54} A\right) \exp\left(-\frac{\mathbf{M}_p \boldsymbol{\Sigma}_0 \mathbf{M}_p^{\mathrm{T}}}{2}\right),
\end{aligned} \tag{244}$$

with $\mathbf{M}_p$ given by Eq. (188).

To appreciate the physical principle, we use a specific case as an example instead of a general mathematical analysis. If $R_{51} = 0$, $R_{52} = 0$, $R_{55} = 1$, $R_{56} = 0$, and the initial beam is transverse-longitudinal-decoupled and has an upright distribution in the longitudinal phase space, then

$$b(k_z) = \sum_{p=-\infty}^{\infty} J_p\left(-k_z R_{54} A\right) \exp\left(-\frac{k_z^2}{2}\left[\epsilon_y\left(R_{53}^2 \beta_y - 2R_{53} R_{54} \alpha_y + R_{54}^2 \gamma_y\right) + \left(1 - \frac{p k_{\mathrm{L}}}{k_z}\right)^2 \sigma_{z0}^2\right]\right), \tag{245}$$

with $\alpha_y, \beta_y, \gamma_y$ being the Courant–Snyder functions before the modulation, and $\sigma_{\delta 0}$ being the initial RMS energy spread. If the initial bunch length is much longer than the laser wavelength, the above exponential terms will be nonzero only when $1 - \frac{p k_{\mathrm{L}}}{k_z} = 0$, which means that there is only bunching

at the laser harmonics. In this case, we have the bunching factor at the n-th laser harmonic to be

$$b_n = b(k_z = nk_{\mathrm{L}}) = J_n(-nk_{\mathrm{L}}R_{54}A)\exp\left[-\frac{(nk_{\mathrm{L}})^2}{2}\epsilon_y\mathcal{H}_{y\mathrm{R}}\right], \tag{246}$$

where $\mathcal{H}_{y\mathrm{R}} = R_{53}^2\beta_y - 2R_{53}R_{54}\alpha_y + R_{54}^2\gamma_y$. One can appreciate the similarity of the above result with the bunching factor of the energy modulation-based TLC microbunching schemes, i.e., Eq. (206). The R_{54} here plays the role of R_{56} in the previous equation. If further $R_{53} = 0$, then Eq. (246) reduces to

$$b_n = J_n(-nk_{\mathrm{L}}R_{54}A)\exp\left[-\frac{(nk_{\mathrm{L}}R_{54}\sigma_{y'})^2}{2}\right], \tag{247}$$

where $\sigma_{y'}$ is the initial RMS beam angular divergence. Therefore, R_{54} and $\sigma_{y'}$ in this scheme play the role of R_{56} and σ_δ in HGHG, as shown in Eq. (195). In a planar uncoupled ring, the natural vertical emittance is quite small, as is $\sigma_{y'}$. Therefore, using this scheme, we can realize high-harmonic bunching in a storage ring to generate ultrashort soft X-ray pulse.

As can be seen from our analysis, both the energy modulation-based and angular modulation-based TLC microbunching schemes share the same spirit, i.e., to take advantage of the small transverse emittance, the vertical emittance in our case, to generate microbunching with a shallow modulation strength. These TLC-based microbunching schemes can be viewed as partial transverse-longitudinal emittance exchanges in the optical laser wavelength range. They do not necessarily need to be complete emittance exchanges because for microbunching, the most important coordinate is z, and δ is relatively less important. As we will show soon, although the spirit is the same, given the same level of modulation laser power, the physical realization of energy modulation-based TLC microbunching schemes is more effective for our SSMB application compared to the angular modulation-based schemes.

7.2 Modulation strength

7.2.1 TEM01 mode laser-induced angular modulation

After deriving the bunching factor, we derived the laser-induced angular modulation strength for quantitative evaluation. We begin with the usual angular modulation proposal by applying a TEM_{01} mode laser in an undulator [43]. The electric field of a Hermite-Gaussian TEM_{01} mode laser polarized in the horizontal plane is [36]

$$\begin{pmatrix} E_x \\ E_y \\ E_z \\ cB_x \\ cB_y \\ cB_z \end{pmatrix} = E_{x0}e^{ik_{\mathrm{L}}s - i\omega_{\mathrm{L}}t}\left(-iZ_{\mathrm{R}}Q\right)^2 \exp\left[i\frac{k_{\mathrm{L}}Q}{2}\left(x^2+y^2\right)\right]\left(\frac{2\sqrt{2}}{w_0}\right)\begin{pmatrix} y \\ 0 \\ -Qxy \\ i\frac{Qx}{k_{\mathrm{L}}} - Q^2xy^2 \\ \left(Q^2x^2 - \frac{iQ}{k_{\mathrm{L}}} + 1\right)y \\ \frac{i}{k_{\mathrm{L}}} - Qy^2 \end{pmatrix}. \tag{248}$$

The relation between E_{x0} and the laser peak power for a TEM_{01} mode laser is given by

$$P_{\mathrm{L}} = \frac{E_{x0}^2 Z_{\mathrm{R}}\lambda_{\mathrm{L}}}{2Z_0}. \tag{249}$$

Note that there is a factor of two difference in the above laser power formula compared to the case of a TEM_{00} mode laser. The electron wiggles in a horizontal planar undulator according to Eq. (210), and the laser-electron exchanges energy according to Eq. (218). Making the same assumption and following similar procedures as before, we can get the integrated modulation voltage induced by a TEM_{01} mode laser in the undulator

$$\begin{aligned} V_{\mathrm{L}} &= \mathrm{Re}\left[E_{x0}\frac{K}{\gamma}e^{i\phi_0}\frac{[JJ]}{2}\frac{2\sqrt{2}}{w_0}y\int_{-\frac{L_{\mathrm{u}}}{2}}^{\frac{L_{\mathrm{u}}}{2}}\frac{1}{\left(1+i\frac{s}{Z_{\mathrm{R}}}\right)^2}\mathrm{d}s\right] \\ &= E_{x0}\frac{K[JJ]}{\gamma}\frac{2\sqrt{2}}{w_0}yZ_{\mathrm{R}}\frac{\frac{L_{\mathrm{u}}}{2Z_{\mathrm{R}}}}{1+\left(\frac{L_{\mathrm{u}}}{2Z_{\mathrm{R}}}\right)^2}\cos\phi_0. \end{aligned} \tag{250}$$

We will choose $\phi_0 = 0$ to maximize the modulation voltage. Putting in the expression of E_{x0} from Eq. (249) and $w_0 = \sqrt{\frac{Z_{\mathrm{R}}\lambda_{\mathrm{L}}}{\pi}}$, we have

$$V_{\mathrm{L}} = \frac{4K[JJ]}{\gamma}\frac{\sqrt{\pi P_{\mathrm{L}}Z_0}}{\lambda_{\mathrm{L}}}\frac{\frac{L_{\mathrm{u}}}{2Z_{\mathrm{R}}}}{1+\left(\frac{L_{\mathrm{u}}}{2Z_{\mathrm{R}}}\right)^2}y. \tag{251}$$

The induced energy modulation strength with respect to y around the zero-crossing phase is then

$$g = \frac{\partial\left(\frac{eV_{\mathrm{L}}}{E_0}\right)}{\partial y} = \frac{2ek_{\mathrm{L}}K[JJ]}{\gamma^2 m_{\mathrm{e}}c^2}\sqrt{\frac{P_{\mathrm{L}}Z_0}{\pi}}\frac{\frac{L_{\mathrm{u}}}{2Z_{\mathrm{R}}}}{1+\left(\frac{L_{\mathrm{u}}}{2Z_{\mathrm{R}}}\right)^2}. \tag{252}$$

The symplecticity of the dynamical system requires that this formula also provides the linear angular chirp strength around the zero-crossing phase. It is interesting to note that, given the laser power, the modulation kick strength depends on the ratio between Z_R and L_u, instead of their absolute values.

One may wonder that when $\frac{L_u}{2Z_R}$ is fixed, the induced angular chirp strength is independent of the modulator length L_u. As a comparison, in a TEM_{00} mode laser modulator, as given in Eq. (221), we have the energy modulation strength proportional to $\sqrt{L_u}$. Mathematically, this is because in the expression of the TEM_{01} mode laser field, there is a term $\left(-iZ_R Q\right)^2$, while in the TEM_{00} mode laser, this term is $\left(-iZ_R Q\right)$. Physically, this means that the on-axis power density of a TEM_{01} mode laser decays faster than that of a TEM_{00} mode laser when we move away from the laser waist. This may not be surprising if we consider that the intensity peaks of a TEM_{01} mode laser are not on-axis. But we recognize there is a lower limit of the modulator length, below which the laser waist size is too small and our approximation of ignoring particle transverse coordinate and wiggling motion on the modulation breaks down.

According to Eq. (252), the maximal modulation is realized when $Z_R = \frac{L_u}{2}$ and the value is

$$g_{\max} = \frac{\partial\left(\frac{eV_L}{E_0}\right)}{\partial y} = \frac{ek_L K[JJ]}{\gamma^2 m_e c^2}\sqrt{\frac{P_L Z_0}{\pi}}. \tag{253}$$

For example, if $E_0 = 600$ MeV, $\lambda_L = 1064$ nm, $\lambda_u = 8$ cm ($B_0 = 1.13$ T), $L_u = 0.8$ m ($N_u = 10$), $Z_R = \frac{L_u}{2}$, then for $P_L = 1$ MW the induced angular chirp strength is $g = 0.55\ \mathrm{m}^{-1}$. For comparison, the energy chirp strength induced by a 1 MW TEM_{00} laser modulator, as evaluated previously, is $h = 955\ \mathrm{m}^{-1}$. So, generally, a TEM_{01} mode laser is ineffective for imprinting angular modulation. As will be shown in the next section, even a dual-tilted laser setup is still not sufficiently effective for our application.

7.2.2 Dual-titled-laser-induced angular modulation

Another way to imprint angular modulation on the electron beam is to use a titled incident TEM_{00} mode laser to modulate the beam in an undulator. To further lower the required laser power, a dual-tilted laser (DTL) with a crossing configuration can be applied [44, 45]. In this study, we focused on an angular modulation scheme based on a DTL setup. Note that if we want to use a DTL for energy modulation, the two lasers should be in phase to ensure that the two laser-induced energy modulations add. For angular modulation, they should be π-phase shifted with respect to each other. This is because, for angular modulation, the particle on the reference orbit should receive a zero-energy kick. Only when the particle transverse coordinate is nonzero will it receive an energy kick. So, the energy modulations induced by the two lasers should cancel on axis.

To induce vertical angular modulation, we let the two lasers cross in y-z plane and be polarized in the horizontal plane. The laser field of a normal incident TEM_{00} laser is given by Eq. (207). Assuming that the two lasers are π -phase-shifted with respect to each other and have the same amplitude. In addition, we assume that $\theta_2 = -\theta_1 = -\theta$, and the two lasers have the same Rayleigh length. Then the superimposed field is

$$\begin{aligned}\begin{pmatrix}E_x\\E_y\\E_z\end{pmatrix}_{\text{unrot}} &= E_{x0}e^{ik_L(-y\sin\theta+s\cos\theta)-i\omega_L t}\frac{\exp\left[-\frac{k_L(x^2+(y\cos\theta+s\sin\theta)^2)}{2(Z_R+i(-y\sin\theta+s\cos\theta))}\right]}{1+i\frac{(-y\sin\theta+s\cos\theta)}{Z_R}}\begin{pmatrix}1\\-i\frac{x}{Z_R+i(-y\sin\theta+s\cos\theta)}\sin\theta\\-i\frac{x}{Z_R+i(-y\sin\theta+s\cos\theta)}\cos\theta\end{pmatrix}\\ &- E_{x0}e^{ik_L(y\sin\theta+s\cos\theta)-i\omega_L t}\frac{\exp\left[-\frac{k_L(x^2+(y\cos\theta-s\sin\theta)^2)}{2(Z_R+i(y\sin\theta+s\cos\theta))}\right]}{1+i\frac{(y\sin\theta+s\cos\theta)}{Z_R}}\begin{pmatrix}1\\i\frac{x}{Z_R+i(y\sin\theta+s\cos\theta)}\sin\theta\\-i\frac{x}{Z_R+i(y\sin\theta+s\cos\theta)}\cos\theta\end{pmatrix}.\end{aligned} \tag{254}$$

As before, we focus on the impact of E_x on the laser–electron interaction and ignore the contribution from E_z. Note that if $y = 0$, then $E_x = 0$. We want to know $\frac{\partial E_x}{\partial y}$ when y is close to zero. For this purpose, we do Taylor expansion of the above horizontal electric field with respect to θ when θ is small,

$$E_x = E_{x0}\frac{Z_R\left(2isk_L Z_R + 2k_L Z_R^2 + x^2k_L + y^2k_L - 2Z_R - 2is\right)\exp\left[-\frac{k_L(x^2+y^2)}{2(Z_R+iz)} + ik_L s - i\omega_L t\right]}{\left(s - iZ_R\right)^3}\theta y. \tag{255}$$

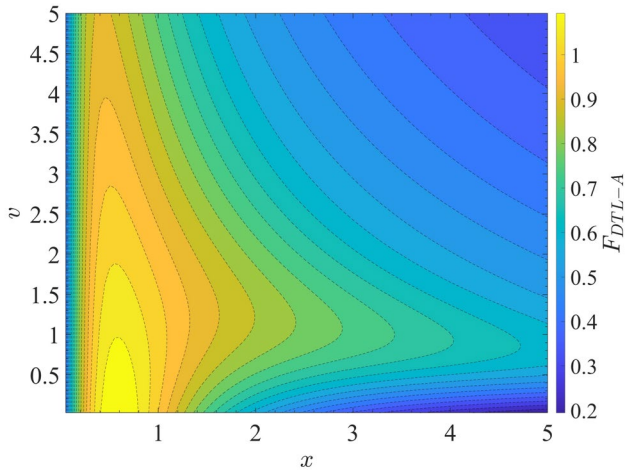

Fig. 12 (Color online) Contour plot of $F_{\text{DTL-A}}(x, v)$ given by Eq. (260)

When $x, y \ll Z_{\text{R}}$, $\lambda_{\text{L}} \ll Z_{\text{R}}$, we have

$$E_x \approx -i2k_{\text{L}}E_{x0}\frac{\exp\left[-\frac{(x^2+y^2)}{w_0^2\left(1+i\frac{s}{Z_{\text{R}}}\right)}+ik_{\text{L}}s-i\omega_{\text{L}}t\right]}{\left(1+i\frac{s}{Z_{\text{R}}}\right)^2}\theta y. \tag{256}$$

However, note that when $L_{\text{u}}\theta$ is comparable to w_0, the term $s\sin\theta$ should be kept in the exponential term. In addition, for the laser phase, we should use a more accurate $s\cos\theta$. These arguments have been explained before when we analyzed the application of a DTL for energy modulation. So, the more correct approximate expression of E_x is

$$E_x \approx -i2k_{\text{L}}E_{x0}\frac{\exp\left(-\frac{s^2\theta^2}{w_0^2\left(1+i\frac{s}{Z_{\text{R}}}\right)}+ik_{\text{L}}s\cos\theta-i\omega_{\text{L}}t\right)}{\left(1+i\frac{s}{Z_{\text{R}}}\right)^2}\theta y. \tag{257}$$

Again taking the same notation as given in Eq. (233), the integrated modulation voltage induced by a DTL (two lasers π-phase-shifted with respect to each other) in a planar horizontal undulator can be calculated to be

$$\begin{aligned}V_{\text{L}} &\approx \text{Re}\left[\int_{-\frac{L_{\text{u}}}{2}}^{\frac{L_{\text{u}}}{2}} v_x E_x\frac{\text{d}s}{c}\right]\\ &\approx k_{\text{L}}E_{x0}\frac{K[JJ]}{\gamma}\theta y Z_{\text{R}}\\ &\left[\int_{-\frac{L_{\text{u}}}{2Z_{\text{R}}}}^{\frac{L_{\text{u}}}{2Z_{\text{R}}}}\frac{\exp\left(-v\frac{u^2}{(1+u^2)}\right)}{(1+u^2)^2}\left[(1-u^2)\cos\left(v\frac{u^3}{(1+u^2)}\right)+2u\sin\left(v\frac{u^3}{(1+u^2)}\right)\right]\text{d}u\right]\sin\phi_0.\end{aligned} \tag{258}$$

We choose $\phi_0 = \frac{\pi}{2}$ to maximize V_{L}.

Therefore, we get the maximum linear angular chirp strength as

$$g = \frac{\partial\left(\frac{eV_{\text{L}}}{E_0}\right)}{\partial y} = \frac{ek_{\text{L}}K[JJ]}{\gamma^2 m_{\text{e}}c^2}\sqrt{\frac{2P_{\text{L}}Z_0}{\lambda_{\text{L}}}}F_{\text{DTL-A}}\left(\frac{L_{\text{u}}}{2Z_{\text{R}}},\frac{k_{\text{L}}Z_{\text{R}}\theta^2}{2}\right)\sqrt{L_{\text{u}}}\theta, \tag{259}$$

where

$$F_{\text{DTL-A}}(x,v) = \frac{1}{\sqrt{x}}\int_{-x}^{x}\frac{\exp\left(-v\frac{u^2}{(1+u^2)}\right)}{(1+u^2)^2}\left[(1-u^2)\cos\left(v\frac{u^3}{(1+u^2)}\right)+2u\sin\left(v\frac{u^3}{(1+u^2)}\right)\right]\text{d}u. \tag{260}$$

A contour plot of $F_{\text{DTL-A}}(x, v)$ is presented in Fig. 12.

Then the angular chirp strength introduced by a DTL compared to the energy chirp strength introduced by a single TEM_{00} laser modulator, i.e., that given in Eq. (221), with the same laser parameters and undulator length can be expressed as

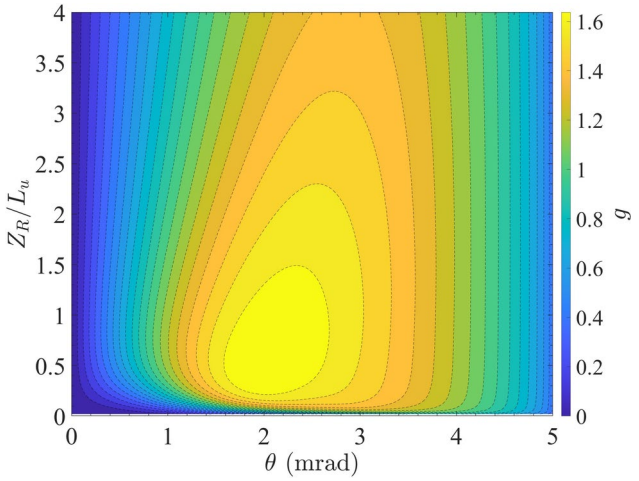

Fig. 13 (Color online) Angular chirp strength g vs. θ and Z_R/L_u, keeping λ_u fixed when changing θ. Parameters used: $E_0 = 600$ MeV, $\lambda_L = 1064$ nm, $P_L = 1$ MW, $\lambda_u = 0.08$ m , $N_u = 10$, $L_u = 0.8$ m

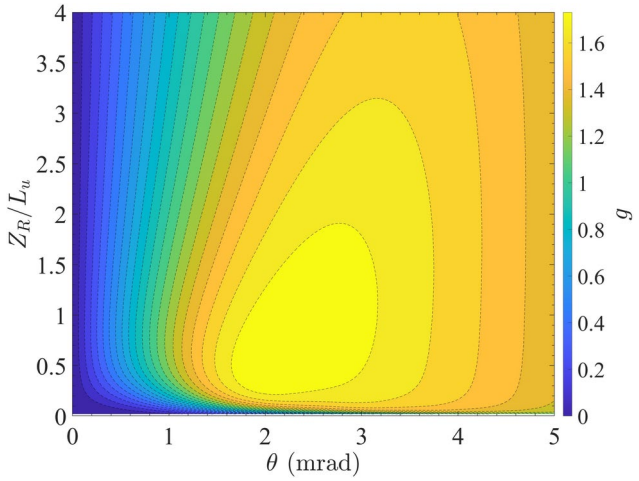

Fig. 14 (Color online) Angular chirp strength g vs. θ and Z_R/L_u, keeping B_0 fixed when changing θ. Parameters used: $E_0 = 600$ MeV, $\lambda_L = 1064$ nm, $P_L = 1$ MW, $B_0 = 1.2$ T, $L_u = 0.8$ m

$$\frac{g}{h} = \xi\theta, \tag{261}$$

with

$$\xi = \frac{F_{\text{DTL-A}}\left(\frac{L_u}{2Z_R}, \frac{k_L Z_R \theta^2}{2}\right)}{\frac{\tan^{-1}\left(\frac{L_u}{2Z_R}\right)}{\sqrt{\frac{L_u}{2Z_R}}}} \frac{K[JJ]}{K[JJ]|_{\theta=0}}. \tag{262}$$

Note that $K[JJ]$ in the numerator is a function of θ according to the off-axis resonance condition given in Eq. (230). Since $\xi \sim 1$, and θ are usually in the mrad level, given the same laser power, the DTL-induced angular chirp strength will be much smaller than the energy modulation strength induced by a TEM_{00} mode laser. This observation was supported by a more quantitative calculation of the angular chirp strength induced by a DTL, as shown in Figs. 13 and 14. As before, we considered the case of keeping λ_u or B_0 unchanged when we change the crossing angle θ. In both cases, the maximal angular chirp strength induced with $P_L = 1$ MW was approximately $g \approx 1.7\ \text{m}^{-1}$. From the evaluation in Sect. 6.2.1, at the same power level, a TEM_{00} mode laser modulation can induce an energy chirp strength of $h \sim 955\ \text{m}^{-1}$. This is because θ is only 1 to 2 mrad.

Therefore, we can see that the DTL-induced angular chirp strength, although a factor of three larger than that induced by a single normally incident TEM_{01} mode laser with the same laser power, is still generally quite small. There are two reasons why a DTL is not effective in imprinting angular modulation:

- The crossing angle between the laser and the electron propagating directions results in that they have a rather limited effective interaction region. For example, if the crossing angle is $\theta = 5$ mrad and the undulator length is $L_u = 0.8$ m. Then, the center of the electron beam and the center of the laser beam at the undulator entrance and exit depart from each other with a distance of $\frac{L_u}{2}\theta = 2$ mm, which is a large value compared to the laser beam waist $w_0 = \sqrt{\frac{Z_R \lambda_L}{\pi}} \approx \sqrt{\frac{L_u \lambda_L}{2\pi}} = 368$ μm and results in a very weak laser electric field felt by the electron there.
- The decrease of undulator parameter K with the increase of θ to meet the off-axis resonance condition, as can be seen in Figs. 8 and 10.

Because the angular chirp strength is small, according to Theorem 2, the required vertical beta function at the modulator β_{yM} will be large. For example, if $\epsilon_y = 4$ pm, $\sigma_{zR} = \sqrt{\epsilon_y \mathcal{H}_{yR}} = 2$ nm, and $g = 2\ \text{m}^{-1}$, then we need $\beta_{yM} = 2.5 \times 10^5$ m, which is too large to be used in practice in a storage ring. To lower β_{yM}, a higher modulation laser power is required. This is why we tend to use energy modulation-based coupling schemes for bunch compression or microbunching generation in the GLSF SSMB. Our analysis and conclusion here is consistent with what reported in Ref. [49] when comparing energy and angular modulation microbunching schemes for laser plasma accelerator-based light source.

Generally, when we compare the energy modulation-based and DTL-induced angular modulation-based bunch compression schemes, from the Theorems 1 and 2, i.e., Eqs. (150) and (152), for the same modulation laser

wavelength λ_L and power P_L, vertical emittance ϵ_y and target linear bunch length at the radiator $\sigma_{zR} = \sqrt{\epsilon_y \mathcal{H}_{yR}}$, we have

$$\beta_{yM}(\text{Angular modulation}) \sim \frac{\mathcal{H}_{yM}(\text{Energy modulation})}{\theta^2}. \tag{263}$$

7.3 Realization examples

Similar to the section on energy modulation-based schemes, we introduce some realization examples of angular modulation-based microbunching. To the best of our knowledge, the first proposal of applying an angular modulated beam for harmonic generation was from Ref. [43]. Later, an emittance exchange-based harmonic generation scheme was proposed in Ref. [46]. These two schemes apply a TEM_{01} mode laser to induce angular modulation. Following these developments, there have been proposals to realize angular modulation using TEM_{00} mode lasers, with Ref. [47] using an off-resonance laser, and Ref. [48] using a tilted incident laser. Later, a dual-tilted-laser (DTL) modulation scheme was applied in emittance exchange in the optical laser wavelength range [44]. Recently, the DTL scheme was proposed to compress the bunch length in SSMB and lower the requirement on the laser power by a factor of four compared to a single-tilted laser scheme [45]. Note that for these angular modulation-based harmonic or bunch compression schemes, we have the inequality given by Eq. (152), i.e., our Theorem 2.

8 1 kW GLSF SSMB-EUV source

Our goal in this study as stated is to find a solution for a high-power EUV source based on SSMB, using parameters within the reach of present technology. According to our analysis, generalized longitudinal strong focusing (GLSF) is the most promising scenario compared to the longitudinal weak focusing and longitudinal strong focusing. The key to a GLSF SSMB ring is the precise transverse-longitudinal coupling dynamics to utilize the ultrasmall natural vertical emittance in a planar electron storage ring for efficient microbunching formation. For our purpose, we find that energy modulation-based coupling schemes are preferred over angular modulation-based coupling schemes, in lowering the required modulation laser power. Therefore, we will use a TEM_{00} laser-induced energy modulation-based coupling scheme in a GLSF SSMB storage ring. In this section, we first present a solution for 1 kW average-power EUV source based on a GLSF SSMB. More detailed analyses to support our solution were then developed.

8.1 A solution of 1 kW EUV source

Based on what we have studied in the previous sections and the various important physical effects to be discussed in this section, we present an example parameter set of a 1 kW average-power EUV light source based on GLSF SSMB, as shown in Table 2. All parameter lists should be feasible from an engineering perspective. Such a table summarizes the investigations presented in this study.

8.2 Some basic considerations

Now, we present the detailed considerations and calculations to support our solution. First, we discuss some basic considerations on the choice of parameters. As explained at the beginning of Sect. 4, we used a beam energy of $E_0 = 600$ MeV and a modulation laser wavelength of $\lambda_L = 1064$ nm. A small vertical emittance is crucial in the GLSF scheme. We assume that the vertical emittance used to accomplish our goal stated above is $\epsilon_y = 40$ pm. It should be noted that the vertical emittance used is not extremely small. This conservative choice is mainly a reflection of the consideration for intra-beam scattering (IBS) to be introduced soon. To realize significant coherent EUV generation, and considering that we may generate microbunching based on a coasting beam or RF-bunched beam in the ring, which is much longer than the modulation laser wavelength, we may need to compress the linear bunch length $\sigma_{zR} = \sqrt{\epsilon_y \mathcal{H}_{yR}}$ to be as short as 2 nm. The bunching factor at 13.5 nm according to Eq. (206) is then 0.0675.

With $\epsilon_y = 40$ pm, to get the desired linear bunch length 2 nm at the radiator, we need $\mathcal{H}_{yR} = 0.1$ μm. If β_y at the radiator is around 1 m, then the required control precision of dispersion and dispersion angle at the radiator is at the level of 0.3 mm and 0.3 mrad. Such precise control of $\mathcal{H}_{yR}$ is challenging but realizable using present technology. We remind the readers that the dispersion function is not well defined when the system is transverse-longitudinal coupled, and $\mathcal{H}_y$ here should be replaced by the defined generalized beta function β^{II}_{55}. However, here we still use the classical definition of $\mathcal{H}_y$ in obtaining the numbers for the above D_x and D'_x control precision to give the readers a more concrete feeling. Following a similar line of thought, we also need a precision control of $\mathcal{H}_x$ or β^I_{55} at the radiator, since the horizontal emittance is even larger than the vertical one. Generally, we require precision control of both $D_{x,y}$ and $D'_{x,y}$. Besides, we also need to ensure the coupling of horizontal emittance ϵ_x to the vertical plane to be less than 1% since our applied $\epsilon_x = 2$ nm is about two orders of magnitude larger than $\epsilon_y = 40$ pm.

Another important parameter is the beam current, including the average and peak currents. First, we observe that

Table 2 An example parameter set of a GLSF SSMB ring for 1 kW average-power EUV radiation generation

Parameter	Value	Description
E_0	600 MeV	Beam energy
C_0	~ 200 m	Circumference
η	$\sim 5 \times 10^{-3}$	Phase slippage factor
I_{P}	40 A	Peak current
f_{e}	0.5%	Electron beam filling factor
I_{A}	200 mA	Average current
B_{ring}	1.33 T	Bending magnet field in the ring
ρ_{ring}	1.5 m	Bending radius in the ring
$U_{0\mathrm{dipoles}}$	7.7 keV	Radiation loss per particle per turn from ring dipoles
$B_{0\mathrm{w}}$	6 T	Bending field of damping wiggler
L_{w}	40 m	Total length of N_{wc} identical damping wigglers
N_{wc}	20	Number of identical damping wigglers
λ_{w}	< 0.168 m	Wiggler period length
$U_{0\mathrm{w}}$	328 keV	Radiation loss per particle per turn from damping wiggler
P_{Rbeam}	68.3 kW	Total radiation loss power of the electron beam
P_{RF}	$100 \sim 200$ kW	Total power consumption of the RF system
$\sigma_{\delta 0}$	4.2×10^{-4}	Natural energy spread (without damping wiggler)
$\sigma_{\delta \mathrm{w}}$	8.2×10^{-4}	Natural energy spread (with damping wiggler)
$\sigma_{\delta \mathrm{IBS}}$	8.5×10^{-4}	Energy spread with IBS (with damping wiggler)
ϵ_x	2 nm	Horizontal emittance
ϵ_y	40 pm	Vertical emittance
$\tau_{y\mathrm{RD}}$	2.38 ms	Vertical radiation damping time with damping wiggler
$\tau_{y\mathrm{IBS}}$	7.11 ms	Vertical IBS diffusion time
λ_L	1064 nm	Modulation laser wavelength
$\sigma_{z\mathrm{R}}$	2 nm	Linear bunch length at the radiator
$\mathcal{H}_{y\mathrm{R}}$	0.1 μm	$\mathcal{H}_y$ at the radiator
$\mathcal{H}_{y\mathrm{M}}$	0.056 m	$\mathcal{H}_y$ at the modulator
$A = \frac{eV_L}{E_0}$	2.3×10^{-3}	Modulator induced energy modulation strength
h	1.33×10^4 m^{-1}	Modulator induced linear energy chirp strength
$\lambda_{u\mathrm{M}}$	0.1 m	Modulator undulator period
$B_{0\mathrm{M}}$	0.806 T	Modulator peak magnetic flux density
K_{uM}	7.53	K of modulator undulator
N_{uM}	15	N_{u} of modulator undulator
L_{uM}	1.5 m	Modulator length
$\Delta\epsilon_{y\mathrm{M}}$	13.4 pm	Modulators' quantum excitation to vertical emittance
Z_{R}	$\sim \frac{L_{\mathrm{uM}}}{3}$	Laser Rayleigh length
P_{LP}	130 MW	Peak modulation laser power
f_{L}	0.5%	Laser beam filling factor
P_{LA}	651 kW	Average modulation laser power
$\lambda_{\mathrm{R}} = \frac{\lambda_L}{79}$	13.5 nm	Radiation wavelength
b_{79}	0.0675	Bunching factor
$\sigma_{\perp \mathrm{R}}$	20 μm	Effective transverse electron beam size at the radiator
$\lambda_{u\mathrm{R}}$	1.8 cm	Radiator undulator period
$B_{0\mathrm{R}}$	0.867 T	Radiator peak magnetic flux density
N_{uR}	79×4	Number of undulator periods
L_{uR}	5.69 m	Radiator length
P_{RP}	224 kW	Peak radiation power
P_{RA}	1.12 kW	Average radiation power

given the same average beam current, the average radiation power will be higher with a decreasing beam filling factor f_e. This is because the peak power of the coherent radiation is proportional to the peak current squared $P_{RP} \propto I_P^2 \propto I_A^2 f_e^{-2}$, where I_P and I_A are the peak and average currents, respectively, and f_e is the electron beam filling factor. The average radiation power is then $P_{RA} = P_{RP} f_e \propto I_A^2 f_e^{-1}$. Therefore, we tend to choose a high average current and a small filling factor as much as possible, as long as collective effects, such as IBS and coherent synchrotron radiation, do not degrade the beam properties. We applied a 40 A peak current and a 200 mA average current in our solution. Note that in the calculation of the filling factor, for simplicity, we assumed that the beam current was similar to square waves. The real beam distribution is more like a smooth curve, for example, a Gaussian profile if there is only a single RF cavity in the ring. However, this observation only results in some numerical factor adjustment and does not affect the core part of our analysis. So, in Table 2, we take this simplification of a square wave current distribution.

8.3 Quantum excitation contribution to vertical emittance

After the general considerations, let us now take a closer look at the critical parameter ϵ_y. The first contribution of the vertical emittance is the quantum excitation in the GLSF section itself, since $\mathcal{H}_y$ there is nonzero. More clearly, according to our Theorem 1, i.e., Eq. (150), the modulator should be placed at a dispersive location if energy modulation-based coupling is used for bunch compression or harmonic generation. Therefore, the quantum excitation of all bend-like elements, such as dipoles, modulators, and radiators in the GLSF section, will contribute to the vertical emittance.

Like the calculation in Eq. (136), the quantum excitation contribution of a radiator to ϵ_y is

$$\Delta\epsilon_{yR} = C_q \frac{\gamma^2}{J_y} \frac{1}{I_2} \times \frac{\mathcal{H}_{yR}}{\rho_{0R}^3} \frac{4}{3\pi} L_{uR}, \tag{264}$$

where L_{uR} is the radiator undulator length, ρ_{0R} is the bending radius corresponding to the peak magnetic field of the radiator B_{0R}. Assuming that the radiation energy loss in the GLSF section is much less than that induced by the bending magnets in the ring, then we have $I_2 \approx \frac{2\pi}{\rho_{ring}}$. Taking the approximation $J_y \approx 1$, in an easy-to-use form we have

$$\Delta\epsilon_{yR}[\text{nm}] = 8.9 B_{ring}^{-1}[\text{T}] B_{0R}^3[\text{T}] \mathcal{H}_{yR}[\text{m}] L_{uR}[\text{m}]. \tag{265}$$

For the parameters listed in Table 2, $B_{ring} = 1.33$ T, $\lambda_{uR} = 1.8$ cm ($B_{0R} = 0.867$ T), $\mathcal{H}_{yR} = 0.1$ μm, $L_{uR} = 5.69$ m ($N_{uR} = 4 \times 79$), then $\Delta\epsilon_{yR} = 2.5$ fm, which is much less than ϵ_y. Generally, because $\mathcal{H}_y$ at the radiator is quite small, the contribution of the radiator to the vertical emittance is not dominant, compared to that from modulators and dipoles to be introduced.

Similar to the analysis for radiator, the contribution of two GLSF modulators to ϵ_y is

$$\Delta\epsilon_{yM}[\text{nm}] = 17.8 B_{ring}^{-1}[\text{T}] B_{0M}^3[\text{T}] \mathcal{H}_{yM}[\text{m}] L_{uM}[\text{m}], \tag{266}$$

where B_{0M} and L_{uM} are the peak magnetic field and length of the modulator undulators, respectively. For the parameters listed in Table 2, $B_{ring} = 1.33$ T, $B_{0M} = 0.806$ T, $\mathcal{H}_{yM} = 0.056$ m, $L_{uM} = 1.5$ m, then $\Delta\epsilon_{yM} = 592$ pm, which is a quite large value compared to that contributed from the radiator. This is mainly due to the fact that $\mathcal{H}_{yM} \gg \mathcal{H}_{yR}$. As will be introduced soon, we can apply a horizontal planar damping wiggler to control this contribution by increasing the radiation damping rate.

In the above analysis, we have assumed $\mathcal{H}_y$ is a constant value throughout an undulator. This is not strictly true, since intrinsic dispersion is generated inside an undulator. We can refer to the transfer matrix of an undulator or laser modulator to study the evolution of $\mathcal{H}_y$ in an undulator radiator and modulator to obtain a more accurate evaluation of their quantum excitation contribution, as shown in Sect. 8.4.2. Our calculation shows that this will result in a significant difference in the radiator's contribution to ϵ_y, but not much difference for the modulators. However, because the radiator's contribution to ϵ_y is negligible compared to that of the modulators, we still use the simplified formula in Eq. (266) in our following discussion.

Vertical bending magnets are also present in the GLSF section for optical manipulation to fulfill the bunch compression or harmonic generation condition. However, in principle, we can use weak dipoles to minimize their quantum excitation contribution to vertical emittance and satisfy the symplectic optics requirement simultaneously. The total length of these dipoles should not be excessively long. Therefore, in our present evaluation, we will assume that the quantum excitation contribution from the two modulators is the dominant source of ϵ_y if we consider single-particle dynamics alone.

8.4 Application of damping wigglers

8.4.1 To speed up damping

It is desired that $\Delta\epsilon_{yM}$ is only a small portion of our desired ϵ_y, since then it provides room for other contributions of vertical emittance, such as IBS and that from x–y coupling. In principle, we can also use a weaker modulator field to weaken the quantum excitation, but this means that the laser-electron interaction will be less efficient, and a larger laser power is needed if we want to imprint the same energy modulation strength. Instead, we chose to increase the radiation damping rate per turn. To speed up damping, which is

helpful in controlling the vertical emittance growth from both the quantum excitation and IBS, we may invoke one or multiple damping wigglers. Horizontal planar wigglers can be placed at dispersion-free locations. Thus, the damping wiggler will contribute only damping and no excitation to the vertical emittance. Assuming that

$$U_{0\mathrm{w}} = R_{\mathrm{w}} U_{0\mathrm{dipoles}}, \tag{267}$$

with $U_{0\mathrm{dipoles}} = C_\gamma E_0^4/\rho_{\mathrm{ring}}$ the radiation energy loss per particle per turn from the bending magnets in the ring, $U_{0\mathrm{w}}$ the radiation loss from all the damping wigglers in the ring, and

$$R_{\mathrm{w}} = \frac{U_{0\mathrm{w}}}{U_{0\mathrm{dipoles}}} = \frac{1}{2}\left(\frac{B_{0\mathrm{w}}}{B_{\mathrm{ring}}}\right)^2 \frac{L_{\mathrm{w}}}{2\pi\rho_{\mathrm{ring}}}, \tag{268}$$

where $B_{0\mathrm{w}}$ and L_{w} are the peak magnetic field and total length of the damping wigglers, respectively. Since the damping rate is proportional to the radiation energy loss per turn, then the damping constant will be a factor of R_{w} larger by applying the damping wiggler

$$\alpha_V = (1 + R_{\mathrm{w}})\alpha_{\mathrm{V0}}, \tag{269}$$

with α_{V0} being the natural vertical damping rate without the damping wigglers. Then, the above evaluated emittance growth from the radiator, modulators, and dipoles will become $\frac{1}{1+R_{\mathrm{w}}}$ of the original value. For example for the modulators, we have

$$\Delta\epsilon_{y\mathrm{M}}[\mathrm{nm}] = \frac{17.8 B_{\mathrm{ring}}^{-1}[\mathrm{T}] B_{0\mathrm{M}}^3[\mathrm{T}] \mathcal{H}_{y\mathrm{M}}[\mathrm{m}] L_{\mathrm{uM}}[\mathrm{m}]}{1 + R_{\mathrm{w}}}. \tag{270}$$

Putting the above relation in another way

$$L_{\mathrm{uM}}[\mathrm{m}] \approx 56.2\big(1 + R_{\mathrm{w}}\big) \frac{B_{\mathrm{ring}}[\mathrm{T}]\Delta\epsilon_{y\mathrm{M}}[\mathrm{pm}]}{\mathcal{H}_{y\mathrm{M}}[\mu\mathrm{m}] B_{0\mathrm{M}}^3[\mathrm{T}]}. \tag{271}$$

We will use a rather strong superconducting damping wiggler or damping wigglers to speed up the damping to fight against IBS diffusion and quantum excitation to maintain a small vertical emittance. For the parameters listed in Table 2, $B_{0\mathrm{ring}} = 1.33$ T, $U_{0\mathrm{dipoles}} = 7.7$ keV, $B_{0\mathrm{w}} = 6$ T and $L_{\mathrm{w}} = 40$ m, then $U_{0\mathrm{w}} = 328$ keV and $R_{\mathrm{w}} = 42.9$. By applying such strong damping wigglers, we now have $\Delta\epsilon_{y\mathrm{M}} = 13.4$ pm, which is a factor of three smaller than the desired $\epsilon_y = 40$ pm and should be acceptable. Assuming the circumference of the ring is $C_0 = 200$ m, the longitudinal and vertical radiation damping times are correspondingly

$$\tau_{\delta\mathrm{RD}} = 1.19\ \mathrm{ms},\ \tau_{y\mathrm{RD}} = 2.38\ \mathrm{ms}. \tag{272}$$

We will compare this radiation damping speed with that of the IBS diffusion later.

8.4.2 Impact of damping wigglers on energy spread, horizontal emittance and phase slippage

Our primary goal of applying a damping wiggler is to speed up radiation damping, but the damping wiggler also contributes to quantum excitation, thus affecting the energy spread and horizontal emittance. We first investigated the energy spread. Considering both the ring dipoles and damping wiggler, the new equilibrium energy spread is

$$\sigma_{\delta\mathrm{w}} = \sigma_{\delta 0}\sqrt{\frac{1 + \frac{4}{3\pi}\left(\frac{B_{0\mathrm{w}}}{B_{\mathrm{ring}}}\right)^3 \frac{L_{\mathrm{w}}}{2\pi\rho_{\mathrm{ring}}}}{1 + \frac{1}{2}\left(\frac{B_{0\mathrm{w}}}{B_{\mathrm{ring}}}\right)^2 \frac{L_{\mathrm{w}}}{2\pi\rho_{\mathrm{ring}}}}}, \tag{273}$$

where

$$\sigma_{\delta 0} = \sqrt{\frac{C_{\mathrm{q}}}{J_z}\frac{\gamma^2}{\rho_{\mathrm{ring}}}} \tag{274}$$

is the natural energy spread in the absence of a damping wiggler. Nominally, we have a longitudinal damping partition $J_z \approx 2$. So

$$\sigma_{\delta 0} \approx 4.69 \times 10^{-4} B_{\mathrm{ring}}^{\frac{1}{2}}[\mathrm{T}] E_0^{\frac{1}{2}}[\mathrm{GeV}]. \tag{275}$$

For example, if $B_{\mathrm{ring}} = 1.33$ T, and $E_0 = 600$ MeV, then $\sigma_{\delta 0} = 4.2 \times 10^{-4}$. When $\frac{1}{2}\left(\frac{B_{0\mathrm{w}}}{B_{\mathrm{ring}}}\right)^2 \frac{L_{\mathrm{w}}}{2\pi\rho_{\mathrm{ring}}} \gg 1$ and $\frac{4}{3\pi}\left(\frac{B_{0\mathrm{w}}}{B_{\mathrm{ring}}}\right)^3 \frac{L_{\mathrm{w}}}{2\pi\rho_{\mathrm{ring}}} \gg 1$, which means the energy spread is dominant by the damping wigglers, we have

$$\begin{aligned}\sigma_{\delta\mathrm{w}} &\approx \sigma_{\delta 0}\sqrt{\frac{8}{3\pi}\frac{B_{0\mathrm{w}}}{B_{\mathrm{ring}}}} \\ &\approx 4.32 \times 10^{-4} B_{0\mathrm{w}}^{\frac{1}{2}}[\mathrm{T}] E_0^{\frac{1}{2}}[\mathrm{GeV}].\end{aligned} \tag{276}$$

Therefore, given the beam energy, the new equilibrium energy spread depends solely on the peak magnetic field of the damping wiggler $B_{0\mathrm{w}}$. For example, if $E_0 = 600$ MeV and $B_{0\mathrm{w}} = 6$ T, then the asymptotic energy spread is $\sigma_{\delta\mathrm{w}} = 8.2 \times 10^{-4}$. This energy spread can affect the coherent EUV radiation power in a long radiator, as will be studied in Sect. 8.9.

Now, let us examine the impact on the horizontal emittance. We have said that we can place the damping wigglers at horizontally dispersion-free locations to minimize their quantum excitation on ϵ_x. However, there will be some intrinsic horizontal dispersion and dispersion angle, and thus $\mathcal{H}_x$ generated inside the wigglers. The strong field strength raises the concern that the quantum excitation of the damping wigglers may result in horizontal emittance growth.

Here, we present a quantitative evaluation of this. Part of the content in this section can also be found in Ref. [50].

Typically, the central part of the wiggler has a sinusoidal field strength pattern along the longitudinal axis. We set the origin of the global path length coordinate $s = 0$ to be the location of the peak magnetic field closest to the wiggler center. Then the vertical magnetic field of a horizontal planar wiggler is

$$B_y = B_{0\mathrm{w}} \cosh(k_x x) \cosh(k_y y) \cos(k_\mathrm{w} s), \tag{277}$$

with $B_{0\mathrm{w}}$ the peak magnetic field and $k_\mathrm{w} = 2\pi/\lambda_\mathrm{w}$ the wavenumber of wiggler, and $k_x^2 + k_y^2 = k_\mathrm{w}^2$. The linear transfer matrix of $\mathbf{X}$ from $s = 0$ to $s \in \left[-\frac{L_\mathrm{w}}{2}, \frac{L_\mathrm{w}}{2}\right]$ with L_w the wiggler length is then [51]

$$\mathbf{W}(s|0) = \begin{pmatrix} 1 & s & 0 & 0 & 0 & -\frac{K}{\gamma k_\mathrm{w}}[1-\cos(k_\mathrm{w} s)] \\ 0 & 1 & 0 & 0 & 0 & -\frac{K}{\gamma}\sin(k_\mathrm{w} s) \\ 0 & 0 & \cos(k_y s) & \frac{\sin(k_y s)}{k_y} & 0 & 0 \\ 0 & 0 & -k_y \sin(k_y s) & \cos(k_y s) & 0 & 0 \\ W_{51} & W_{52} & 0 & 0 & 1 & W_{56} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \tag{278}$$

where $W_{51} = -W_{26}$, $W_{52} = -sW_{26} + W_{16}$ and

$$W_{56} = \frac{2\lambda_0}{\lambda_\mathrm{w}} s + \frac{K^2}{\gamma^2}\left[\frac{\sin(2k_\mathrm{w} s) - 4\sin(k_\mathrm{w} s)}{4k_\mathrm{w}}\right], \tag{279}$$

with $\lambda_0 = \frac{1+K^2/2}{2\gamma^2}\lambda_\mathrm{w}$ being the fundamental on-axis resonant wavelength.

In a planar uncoupled ring, the normalized eigenvector of the storage ring one-turn map corresponding to the horizontal eigenmode at $s = 0$ can be expressed as the first vector in Eq. (50). We assume α_{x0}, β_{x0} are the Courant–Snyder functions and D_{x0}, D'_{x0} are the dispersion and dispersion angle corresponding to the horizontal plane at $s = 0$, respectively. Then the horizontal chromatic function $\mathcal{H}_{x0}$ at $s = 0$ is

$$\begin{aligned} \mathcal{H}_{x0} &\equiv \beta_{55}^I(0) = 2|E_{I5}(0)|^2 \\ &= \frac{D_{x0}^2 + \left(\alpha_{x0} D_{x0} + \beta_{x0} D'_{x0}\right)^2}{\beta_{x0}}. \end{aligned} \tag{280}$$

The evolution of $\mathcal{H}_x$ from $s = 0$ to $s \in \left[-\frac{L_\mathrm{w}}{2}, \frac{L_\mathrm{w}}{2}\right]$ is

$$\begin{aligned} \mathcal{H}_x(s) &\equiv \beta_{55}^I(s) = 2|E_{I5}(s)|^2 = 2|\left(\mathbf{W}(s|0)\mathbf{E}_I(0)\right)_5|^2 \\ &= \frac{(D_{x0} + W_{16} - W_{26}s)^2 + \left[\alpha_{x0}(D_{x0} + W_{16} - W_{26}s) + \beta_{x0}(D'_{x0} + W_{26})\right]^2}{\beta_{x0}}, \end{aligned} \tag{281}$$

where W_{ij} denotes the i-th row and j-th column matrix term of $\mathbf{W}(s|0)$. Putting in the explicit expression of the wiggler matrix terms, we have

$$\begin{aligned} \mathcal{H}_x(s) = \frac{1}{\rho_\mathrm{w}^2 k_\mathrm{w}^4 \beta_{x0}} \big\{ &[D_{x0}\rho_\mathrm{w} k_\mathrm{w}^2 + \sin\left(k_\mathrm{w} s\right) k_\mathrm{w} s + \cos(k_\mathrm{w} s) - 1]^2 \\ &+ [\beta_{x0} k_\mathrm{w}\left(\rho_\mathrm{w} k_\mathrm{w} D'_{x0} - \sin\left(k_\mathrm{w} s\right)\right) + \alpha_{x0}\left(D_{x0}\rho_\mathrm{w} k_\mathrm{w}^2 + \sin\left(k_\mathrm{w} s\right) k_\mathrm{w} s + \cos(k_\mathrm{w} s) - 1\right)]^2 \big\}. \end{aligned} \tag{282}$$

We assume that the quantum excitation contribution from the entrance and exit regions of the wiggler, where the field strength deviates from the ideal sinusoidal pattern, is much smaller than that of the central sinusoidal field region. Then the quantum excitation of a wiggler to the horizontal beam emittance can be evaluated by the integral

$$I_{5\mathrm{w}} = \int_{-\frac{L_\mathrm{w}}{2}}^{\frac{L_\mathrm{w}}{2}} \frac{\mathcal{H}_x(s)}{|\rho(s)|^3} \mathrm{d}s = \frac{1}{\rho_\mathrm{w}^3} \int_{-\frac{L_\mathrm{w}}{2}}^{\frac{L_\mathrm{w}}{2}} \mathcal{H}_x(s) |\cos(k_\mathrm{w} s)|^3 \mathrm{d}s, \tag{283}$$

where $\rho_\mathrm{w} = \frac{\gamma m_\mathrm{e} \beta c}{e B_{0\mathrm{w}}}$ corresponds to the bending radius at the location of peak magnetic field $B_{0\mathrm{w}}$, with $\beta = \sqrt{1 - \frac{1}{\gamma^2}}$. Putting Eq. (282) in, we have

$$I_{5\mathrm{w}} = \frac{4}{15\pi} \frac{L_\mathrm{w}}{\rho_\mathrm{w}^5 k_\mathrm{w}^2} \left[\beta_{x0} + \gamma_{x0} L_\mathrm{w}^2 \mathcal{R} + 5\rho_\mathrm{w}^2 k_\mathrm{w}^2 \mathcal{H}_{x0} - \left(10 + \frac{15\pi}{8}\right) \frac{\rho_\mathrm{w}\left(D_{x0} + \alpha_{x0} D_{x0} + \beta_{x0} D'_{x0}\right)}{\beta_{x0}}\right], \tag{284}$$

where $\gamma_{x0} = \frac{1+\alpha_{x0}^2}{\beta_{x0}}$ and

$$\mathcal{R} = \frac{15}{32\pi^2} \frac{1}{N_\mathrm{w}^3} \int_{-N_\mathrm{w}\pi}^{N_\mathrm{w}\pi} [\sin(x)x + \cos(x) - 1]^2 |\cos(x)|^3 \mathrm{d}x, \tag{285}$$

with $N_\mathrm{w} = L_\mathrm{w}/\lambda_\mathrm{w}$ being the number of wiggler period, which is assumed to be an integer. Equation (284) is the exact formula for the wiggler contribution to the radiation integral I_{5x} in an electron storage ring.

Given a specific N_u, $\mathcal{R}$ can be obtained straightforwardly by integration in Eq. (285). When $N_\mathrm{w} \gg 1$, we have $\mathcal{R} \approx \frac{1}{12}$ and

$$I_{5\mathrm{w}} \approx \frac{4}{15\pi}\frac{L_\mathrm{w}}{\rho_\mathrm{w}^5 k_\mathrm{w}^2}\left[\langle\beta_x\rangle_\mathrm{w} + 5\rho_\mathrm{w}^2 k_\mathrm{w}^2 \mathcal{H}_{x0} - \left(10 + \frac{15\pi}{8}\right)\frac{\rho_\mathrm{w}\left(D_{x0} + \alpha_{x0}D_{x0} + \beta_{x0}D'_{x0}\right)}{\beta_{x0}}\right], \tag{286}$$

where $\langle\beta_x\rangle_\mathrm{w}$ is the average value of β_x along the wiggler

$$\langle\beta_x\rangle_\mathrm{w} = \frac{1}{L_\mathrm{w}}\int_{-\frac{L_\mathrm{w}}{2}}^{\frac{L_\mathrm{w}}{2}} \beta_x(s)\mathrm{d}s = \beta_{x0} + \frac{\gamma_{x0}L_\mathrm{w}^2}{12}. \tag{287}$$

Denote $\chi_{x0} = \mathrm{Arg}\left(\frac{E_{I5}(0)}{E_{I1}(0)}\right)$, where $E_{I5}(0)$ and $E_{I1}(0)$ represent the fifth and first term of the first eigenvector in Eq. (50) and Arg() means the angle of a complex number, then Eq. (286) can be written as

$$I_{5\mathrm{w}} \approx \frac{4}{15\pi}\frac{L_\mathrm{w}}{\rho_\mathrm{w}^5 k_\mathrm{w}^2}\left[\langle\beta_x\rangle_\mathrm{w} + 5\rho_\mathrm{w}^2 k_\mathrm{w}^2 \mathcal{H}_{x0} - \left(10 + \frac{15\pi}{8}\right)\rho_\mathrm{w}\sqrt{\frac{2\mathcal{H}_{x0}}{\beta_{x0}}}\sin\left(\chi_{x0} - \frac{\pi}{4}\right)\right]. \tag{288}$$

The first term in the above bracket corresponds to the approximate formula found in the literature

$$I_{5\mathrm{w,intrinsic}} \approx \frac{4}{15\pi}\frac{L_\mathrm{w}\langle\beta_x\rangle_\mathrm{w}}{\rho_\mathrm{w}^5 k_\mathrm{w}^2}. \tag{289}$$

It can be viewed as the intrinsic contribution of a wiggler to the radiation integral I_{5x}, since there will be intrinsic dispersion and dispersion angle generated inside the wiggler even if $D_{x0} = 0$ and $D'_{x0} = 0$ as can be seen from the matrix terms W_{16} and W_{26} of the wiggler. The second and third terms in the bracket arise from a nonzero $\mathcal{H}_{x0}$. When D'_{x0} or $\frac{D_{x0}}{\beta_{x0}}$ is of the order $\frac{K}{\gamma} = \frac{1}{\rho_\mathrm{w} k_\mathrm{w}}$, which can easily be the case in a real lattice, the contribution from this nonzero $\mathcal{H}_{x0}$ in the bracket could be comparable or even larger than the first term and cannot be neglected. The more accurate formula derived here should then be invoked to calculate the wiggler quantum excitation of the beam emittance. When the third term is much smaller than the second term, i.e., roughly when $\mathcal{H}_{x0}\beta_{x0} \gg \left(\frac{K}{\gamma}\lambda_\mathrm{w}\right)^2$, Eq. (288) can be further approximated as

$$I_{5\mathrm{w}} \approx \frac{4}{15\pi}\frac{L_\mathrm{w}\langle\beta_x\rangle_\mathrm{w}}{\rho_\mathrm{w}^5 k_\mathrm{w}^2} + \frac{4}{3\pi}\frac{L_\mathrm{w}\mathcal{H}_{x0}}{\rho_\mathrm{w}^3}, \tag{290}$$

where the first term accounts for the intrinsic contribution and the second term accounts for the nonzero $\mathcal{H}_{x0}$.

From the above analysis, we can see that the minimum $I_{5\mathrm{w}}$ is realized when

$$\alpha_{x0} = 0,\ \beta_{x0} = \frac{L_\mathrm{w}}{2\sqrt{3}},\ D_{x0} = 0,\ D'_{x0} = 0,\ \left(\mathcal{H}_{x0} = 0\right), \tag{291}$$

and the minimal value is

$$I_{5\mathrm{w,min}} \approx \frac{4}{15\sqrt{3}\pi}\frac{L_\mathrm{w}^2}{\rho_\mathrm{w}^5 k_\mathrm{w}^2}. \tag{292}$$

Now we can evaluate the impact of the damping wigglers on the equilibrium horizontal emittance to ensure that the desired $\epsilon_x = 2$ nm can be realized. From Eq. (53), the equilibrium emittance with the damping wigglers is given by

$$\epsilon_{x\mathrm{w}} = C_\mathrm{q}\frac{\gamma^2}{J_x}\frac{I_{50} + I_{5\mathrm{w}}}{I_{20} + I_{2\mathrm{w}}}, \tag{293}$$

where I_{20} and I_{50} are the radiation integrals of the ring without damping wigglers. Note that the natural emittance without the damping wigglers is given by

$$\epsilon_{x0} = C_\mathrm{q}\frac{\gamma^2}{J_x}\frac{I_{50}}{I_{20}}. \tag{294}$$

If we want $\epsilon_{x\mathrm{w}} \le \epsilon_{x0}$, then we need $C_\mathrm{q}\frac{\gamma^2}{J_x}\frac{I_{5\mathrm{w}}}{I_{2\mathrm{w}}} \le \epsilon_{x0}$. Using $I_{2\mathrm{w}} \approx \frac{L_\mathrm{w}}{2\rho_\mathrm{w}^2}$ and Eq. (292), we then have

$$\lambda_\mathrm{w} \le 2\pi\sqrt{\frac{15\sqrt{3}\pi J_x \epsilon_{x0}\rho_\mathrm{w}^3}{8C_\mathrm{q}\gamma^2 L_\mathrm{w}}}. \tag{295}$$

According to the above scaling, a longer wiggler length requires a shorter wiggler period to control the quantum excitation contribution to the horizontal emittance. This is because the average values of β_x and $\mathcal{H}_x$ in the wiggler are linearly proportional to L_w. However, note that the above analysis assumes that there is only a single wiggler. If we split the wiggler into N_wc identical cells with a fixed total

length, we can make the contribution of the wigglers to I_{5w} and thus quantum excitation to horizontal emittance becomes a factor of N_{wc} smaller, while the contribution to I_{2w} and thus the effect on radiation damping is unchanged. Then, the tolerance of λ_w can be a factor of $\sqrt{N_{wc}}$ larger. Putting in the numerical numbers, we have a more useful scaling

$$\lambda_w[\mathrm{m}] \le 3.19\sqrt{\frac{N_{wc}J_x E_0[\mathrm{GeV}]\epsilon_{x0}[\mathrm{nm}]}{B_{0w}^3[\mathrm{T}]L_w[\mathrm{m}]}}. \tag{296}$$

Nominally $J_x \approx 1$. For our example parameters given in Table 2, $E_0 = 600$ MeV, $\epsilon_{x0} = 2$ nm, $B_{0w} = 6$ T, $L_w = 40$ m, $N_{wc} = 20$ which means each small wiggler has a length of 2 m, then to realize $\epsilon_{xw} \le \epsilon_{x0}$ we need

$$\lambda_w \le 0.168 \text{ m}. \tag{297}$$

A wiggler with a period length of 10 cm to 15 cm and a peak field of 6 T is feasible using superconducting magnet technology. We recognize the impact of such a long (total length) damping wiggler on single-particle nonlinear dynamics and collective instabilities require further in-depth study, especially considering that the gap of such a wiggler is small to realize such a strong field strength. Another practical reason for splitting the wigglers into shorter sections is to avoid the synchrotron radiation generated heating the magnet poles. We also recognize that it may take some effort in the lattice design to achieve the optimal conditions given by Eq. (291).

In addition to quantum excitation, the damping wiggler also contributes to phase slippage. If $B_w = 6$ T, $\lambda_w = 0.1$ m, $L_w = 40$ m which means a total wiggler period number $N_w = 400$, then the wiggler undulator parameter is $K_w = 0.934 \times 6 \times 10 = 56.04$, then the fundamental resonance wavelength of the wiggler is $\lambda_{rw} = \frac{1+K_w^2/2}{2\gamma^2}\lambda_w = 57$ μm. The R_{56} of the entire damping wiggler is twice the fundamental frequency radiation slippage length $R_{w,56} = 2N_w\lambda_{rw} = 45.6$ mm [17]. The total R_{56} of the ring used in Table 2 is approximately 1 m; then the contribution of the damping wigglers to R_{56} is acceptable.

8.5 Intra-beam scattering

We mentioned that our conservative choice of $\epsilon_y = 40$ pm is mainly due to the consideration of IBS. This can be understood with more quantitative calculations. We will see that IBS is the most fundamental obstacle in obtaining ultrasmall vertical emittance in GLSF SSMB. This is partially because our choice of beam energy was not too high. In addition, to realize high EUV power, we need a high peak current, which means a high charge density in the phase space.

We use Bane's high-energy approximation [52] to calculate the IBS diffusion rate

$$\begin{aligned}
&\frac{1}{T_\delta} \approx \frac{r_e^2 cNL_c}{16\gamma^3\epsilon_x^{\frac{3}{4}}\epsilon_y^{\frac{3}{4}}\sigma_z\sigma_\delta^3}\left\langle \sigma_H g_{\mathrm{Bane}}\left(\frac{a}{b}\right)(\beta_x\beta_y)^{-\frac{1}{4}}\right\rangle,\\
&\frac{1}{T_{x,y}} = \frac{\sigma_\delta^2\langle\mathcal{H}_{x,y}\rangle}{\epsilon_{x,y}}\frac{1}{T_\delta},\\
&\frac{1}{\sigma_H^2} = \frac{1}{\sigma_\delta^2} + \frac{\mathcal{H}_x}{\epsilon_x} + \frac{\mathcal{H}_y}{\epsilon_y},\\
&g_{\mathrm{Bane}}(\alpha) = \frac{2\sqrt{\alpha}}{\pi}\int_0^\infty \frac{du}{\sqrt{1+u^2}\sqrt{\alpha^2+u^2}},\\
&a = \frac{\sigma_H}{\gamma}\sqrt{\frac{\beta_x}{\epsilon_x}},\ b = \frac{\sigma_H}{\gamma}\sqrt{\frac{\beta_y}{\epsilon_y}},
\end{aligned} \tag{298}$$

where $r_e = 2.818 \times 10^{-15}$ m is the classical electron radius and $L_c = \ln\left(\frac{b_{\max}}{b_{\min}}\right)$ is the Coulomb Log factor, with $b_{\max}$ and $b_{\min}$ being the maximum and minimal impact factor for the scattering process, respectively. Typically L_c is in the range of 10 to 20. N is the number of electrons in the bunch. For a coasting beam, we need to replace $\sigma_z \to L/(2\sqrt{\pi})$ where L is the bunch length. Note that $\frac{eN}{L/c} = I_P$ according to our definition, where I_P is the peak current.

Now, we input some example numbers to estimate the IBS diffusion rate in a GLSF EUV SSMB ring:

$$\begin{aligned}
&E_0 = 600 \text{ MeV},\ I_P = 40 \text{ A},\ \sigma_\delta = 8.5\times10^{-4},\\
&\epsilon_x = 2 \text{ nm},\ \epsilon_y = 40 \text{ pm}, \langle\sigma_H\rangle = 4\times10^{-5},\\
&\frac{a}{b} = \frac{1}{10},\ g_{\mathrm{Bane}}\left(\frac{1}{10}\right) = 0.744,\ \langle(\beta_x\beta_y)^{-\frac{1}{4}}\rangle = 0.32,\\
&\langle\mathcal{H}_y\rangle = \frac{2\times1.6\times0.056 \text{ m}}{200} = 0.9 \text{ mm},\ L_c = 10.
\end{aligned} \tag{299}$$

Note that in Table 2 we have $\mathcal{H}_y = 0.056$ m at the two modulators, whose length are both 1.6 m. In evaluating $\langle\mathcal{H}_y\rangle$ we only considered the contribution from the two modulators, where $\mathcal{H}_y$ reaches its maximum value. This is a simplification but should provide a correct order of magnitude estimation. Putting in the example numbers, we then have

$$\tau_{\delta\mathrm{IBS}} = 113 \text{ ms},\ \tau_{y\mathrm{IBS}} = 7.11 \text{ ms}. \tag{300}$$

Compared with the radiation damping times given in Eq. (272), we can see that even for the vertical dimension,

the IBS diffusion is more than three times slower than the radiation damping. Therefore, the IBS diffusion can now be controlled by the strong damping induced by damping wigglers. This calculation also justifies the necessity or benefit of applying the damping wigglers.

8.6 Microwave instability

Now, we want to evaluate whether the peak current of 40 A we applied in the above example is feasible. One of the main limitations of the peak current is the microwave instability induced by coherent synchrotron radiation (CSR). According to Ref. [53], the CSR-induced microwave instability threshold is

$$(S_{\rm CSR})_{\rm th} = 0.5 + 0.12\Pi, \tag{301}$$

with

$$S_{\rm CSR} = \frac{I\rho^{1/3}}{\sigma_{z0}^{4/3}},\ I = \frac{r_{\rm e}N_{\rm b}}{2\pi\nu_{s0}\gamma\sigma_{\delta 0}},\ \Pi = \frac{\sigma_{z0}\rho^{1/2}}{g^{3/2}}, \tag{302}$$

and $2g$ is the separation between the two plates. So, the threshold peak current is

$$I_{\rm th,peak} = \frac{eN_{\rm b}}{\sqrt{2\pi}\sigma_{z0}/c} = \frac{1}{2\sqrt{2\pi}} I_{\rm Alf}\gamma\left(1 + 0.24\frac{\sigma_{\delta 0}|R_{56}|^{1/2}\rho^{1/2}}{|h_{\rm RF}|^{1/2}g^{3/2}}\right) \frac{\sigma_{\delta 0}^{\frac{4}{3}}|R_{56}|^{\frac{2}{3}}|h_{\rm RF}|^{\frac{1}{3}}}{\rho^{1/3}}, \tag{303}$$

with $I_{\rm Alf} = \frac{ec}{r_{\rm e}} = 17$ kA being the Alfven current. The $R_{56} = -\eta C_0$ is that of the entire ring. $h_{\rm RF} = \frac{eV_{\rm RF}\cos\phi_{\rm s}}{E_0}k_{\rm RF}$ is the linear energy chirp strength around the synchronous RF phase $\phi_{\rm s}$. Putting in some typical parameters for the EUV SSMB: $E_0 = 600$ MeV, $B_0 = 1.33$ T, $\rho = 1.5$ m, $|R_{56}| = 1$ m, $\sigma_{\delta 0} = 8.5\times10^{-4}$, $g = 4$ cm, $h_{\rm RF} = 0.01$ m^{-1}, then

$$I_{\rm th,peak} = 60.3\times(1+0.31)\ {\rm A} = 79\ {\rm A}. \tag{304}$$

Therefore, our application of a peak current of 40 A should be safe from microwave instability.

The astute reader may notice that one of the main reasons we have a large threshold current here is the large phase slippage or R_{56} we applied for the ring. To avoid confusion, first we need to clarify that in this example of a GLSF SSMB-EUV source, the electron bunch in the ring can be a coasting beam or an RF-bunched beam, and microbunching appears only at the radiator, owing to the phase space manipulation of the electron beam in the GLSF section. In our setup, we used an RF-bunched beam in the ring. Therefore, the phase slippage factor of the ring does not need to be small, whereas this is required in an LWF SSMB ring. Then, the question becomes whether the required large phase slippage is feasible and what beam dynamics effects it may have. As a reference, the Metrology Light Source storage ring [54] in standard user mode has an $|R_{56}| \approx 1.6$ m, which means a phase slippage factor of 3.3×10^{-2} given a circumference of 48 m. Therefore, we believe that our application of $|R_{56}| = 1$ m is realizable. However, we recognize that such a large R_{56} requires a large horizontal dispersion D_x at the dipoles, since $-R_{56} = \eta C_0 \approx 2\pi\langle D_x\rangle_\rho$ where $\langle\rangle_\rho$ means average around the dipoles in the ring. A large D_x may result in a large $\mathcal{H}_x$, and the quantum excitation of dipoles to the horizontal emittance ϵ_x should be carefully evaluated. A large ϵ_x will degrade the coherent EUV radiation [17, 55]. What we need is $|R_{56}| \sim 1$ m, and at the same time $\epsilon_x \lesssim 2$ nm, and some optimization may be required to realize such a goal.

Now, we check if the required RF system in the above example calculation is feasible. The longitudinal beta function at the RF cavity is $\beta_z \approx \sqrt{\eta C_0/h_{\rm RF}} = 10$ m, and the RMS bunch length is $\sigma_z = \sigma_{\delta 0}\beta_z = 8.5$ mm. To obtain a beam filling factor of 0.5%, we require an RF wavelength of approximately 1.7 m, which corresponds to an RF frequency of 176.5 MHz. Then, the required energy chirp strength of $h_{\rm RF} = 0.01$ m^{-1} means that the required RF voltage is 1.62 MV. Such an RF voltage should be achievable in this frequency range. Multiple cavities can be invoked if it is too demanding for a single cavity to reach the desired voltage.

We remind the readers that in this evaluation, we assumed that the beam was RF-bunched. In our evaluation of the IBS, we assumed that the beam was a coasting beam. These evaluations mainly serve as an order of magnitude estimation that supports the general feasibility of our parameter choice. A more detailed analysis of the collective effects will be necessary in the future development of such a GLSF SSMB light source.

8.7 Energy compensation system

The large radiation loss of the electron beam in the ring, particularly that induced by the strong damping wigglers, must be compensated. Here, we present some preliminary analysis on the requirements of the energy compensation system of such a GLSF SSMB-EUV light source. From Table 2, the total radiation loss per particle per turn is $U_0 = U_{\rm 0dipoles} + U_{\rm 0w} + U_{\rm 0R} = 341.3$ keV, where the three terms on the right hand side represent the radiation loss in dipoles, damping wigglers and radiator, respectively. Therefore, the synchronous acceleration phase corresponds to an acceleration voltage of $V_{\rm acc} = 341.3$ kV. Such a high voltage is not easy to realize using an induction linac, considering

that the required repetition rate is at the MHz level. Therefore, we use conventional RF cavities to supply radiation loss. Following the discussion in the last section, and assuming that we have used an RF frequency of 166.6 MHz (RF wavelength 1.8 m), we need an RF voltage of 1.72 MV to realize an energy chirp strength of $h_{\rm RF} = 0.01\ {\rm m}^{-1}$. Under this parameter set, and consider $V_{\rm RF} \gg V_{\rm acc}$, the RF bucket half-height is [17]

$$\hat{\delta}_{\frac{1}{2}} = \frac{2}{\beta_z k_{\rm RF}} = \frac{\lambda_{\rm RF}}{\pi}\sqrt{\frac{h}{\eta C_0}} = 5.73\times 10^{-2}, \tag{305}$$

which is $67.4\sigma_{\delta 0}$ with the energy spread $\sigma_{\delta 0} = 8.5\times 10^{-4}$. Therefore, the bucket half-height should be sufficiently large to ensure the beam quantum and Touschek lifetime.

To relieve the burden on the RF cavities, we may use three RF cavities to achieve the total RF voltage, with each cavity having a voltage of 573 kV. To minimize the power dissipated on the cavity, we need a large shunt impedance of each cavity which here we assume to be $R_{\rm s} = 20\ {\rm M\Omega}$, then the total power dissipated on the three cavity walls is

$$P_{\rm diss} = \frac{1}{3}\frac{V_{\rm RF}^2}{R_{\rm s}} = 49.2\ {\rm kW}. \tag{306}$$

The power delivered to the 200 mA average current beam is

$$P_{\rm b} = I_{\rm A}\frac{U_0}{e} = 68.3\ {\rm kW}. \tag{307}$$

Therefore, the power dissipated on the cavity walls was at the same level as that delivered to the beam. We recognize that a large shunt impedance may require effort to realize in practice. If the shunt impedance is lower than the assumed value, the power consumption on the wall will be correspondingly higher. Superconducting RF cavities can be used to reduce the power dissipation on the wall.

Generally, the total power consumption of the RF system of an SSMB ring is at 100–200 kW level. Together with the power consumption of other systems, such as electromagnets, superconducting damping wigglers, and vacuum and water cooling systems, the overall power consumption of such an SSMB storage ring is at the level of several hundred kWs. The output EUV power per radiator is approximately 1 kW. In principle, an SSMB ring can accommodate multiple GLSF insertions and, therefore, multiple radiators; however, we consider the case of only one radiator in the ring. Therefore, for such a GLSF SSMB storage ring, it takes a couple of 100 kWs of electricity to generate 1 kW of EUV light. Such a large power consumption may raise the question of the advantage of such an SSMB-EUV source compared to the superconducting RF-based high repetition rate FEL-EUV source, particularly the energy recovery linac-based FEL-EUV source. For comparison, according to Ref. [56], it takes 7 MW overall power consumption to generate 10 kW EUV light in an ERL-FEL EUV source, which means 700 kW electricity power per 1 kW EUV light. Therefore, the overall power efficiency from electricity to EUV light for the SSMB-EUV and ERL-FEL EUV source is comparable. However, we remind readers that these numbers are only rough estimates, and more in-depth studies are required to reach a concrete conclusion. Another side comment is that the radiation emitted by the damping wigglers may also be useful.

8.8 Modulation laser power

We now evaluate the modulation laser power required. Given the laser wavelength, modulator undulator parameters, and the required energy chirp strength, we can use Eq. (221) to calculate the required laser power

$$P_{\rm L} = \frac{\lambda_{\rm L}}{4Z_0 Z_{\rm R}}\left(h\frac{1}{\frac{eK[JJ]}{\gamma^2 m_{\rm e}c^2}\tan^{-1}\left(\frac{L_{\rm u}}{2Z_{\rm R}}\right)k_{\rm L}}\right)^2. \tag{308}$$

Using Theorem 1, i.e., Eq. (150), and let the equality in the relation holds, we have

$$P_{\rm L} \approx \frac{1}{1+R_{\rm w}}\frac{\epsilon_{\rm y}}{\Delta\epsilon_{\rm yM}}\frac{1}{(K[JJ])^2}\frac{\lambda_L^3}{3\pi^3 Z_0}\frac{55}{48\sqrt{3}}\frac{\alpha_F c^2\lambda_{\rm e}^2\gamma^7 B_{\rm 0M}^3}{C_\gamma E_0^3 B_{\rm ring}}\frac{1}{\sigma_{z\rm R}^2}\frac{\frac{L_{\rm u}}{2Z_{\rm R}}}{\left[\tan^{-1}\left(\frac{L_{\rm u}}{2Z_{\rm R}}\right)\right]^2}, \tag{309}$$

where $\alpha_{\rm F} = \frac{1}{137}$ is the fine structure constant and $R_{\rm w}$ is given by Eq. (268). In the above derivation, we used the electron momentum $P_0 = \gamma m_{\rm e} c$ and the approximation $E_0 \approx P_0 c$. We now attempt to derive more useful scaling laws to offer guidance in our parameter choice for a GLSF SSMB storage ring. As shown in Fig. 6, to maximize the energy modulation, we need $\frac{Z_{\rm R}}{L_{\rm u}} = 0.359 \approx \frac{1}{3}$. When $K > \sqrt{2}$, we approximate the resonance condition as

$$\lambda_{\rm L} = \frac{1+\frac{K^2}{2}}{2\gamma^2}\lambda_{\rm u} \approx \frac{K^2}{4\gamma^2}\lambda_{\rm u}, \tag{310}$$

and $[JJ] \approx 0.7$. Then we have

$$\begin{aligned} P_{\rm L} &\propto \frac{1}{1+R_{\rm w}}\frac{\epsilon_{\rm y}}{\Delta\epsilon_{\rm yM}}\frac{\lambda_{\rm L}^3}{K^2}\frac{\gamma^4 B_{\rm 0M}^3}{B_{\rm ring}}\frac{1}{\sigma_{z\rm R}^2} \\ &\propto \frac{1}{1+R_{\rm w}}\frac{\epsilon_{\rm y}}{\Delta\epsilon_{\rm yM}}\frac{\lambda_{\rm L}^{\frac{7}{3}}\gamma^{\frac{8}{3}}B_{\rm 0M}^{\frac{7}{3}}}{B_{\rm ring}}\frac{1}{\sigma_{z\rm R}^2}. \end{aligned} \tag{311}$$

Putting in the numbers for the constants, we obtain the quantitative expressions of the above scalings for practical use

$$P_{\mathrm{L}}[\mathrm{kW}] \approx 5.7 \frac{1}{1+R_{\mathrm{w}}} \frac{\epsilon_y}{\Delta\epsilon_{y\mathrm{M}}} \frac{\lambda_{\mathrm{L}}^{\frac{7}{3}}[\mathrm{nm}] E_0^{\frac{8}{3}}[\mathrm{GeV}] B_{0\mathrm{M}}^{\frac{7}{3}}[\mathrm{T}]}{\sigma_{z\mathrm{R}}^2[\mathrm{nm}] B_{\mathrm{ring}}[\mathrm{T}]}. \tag{312}$$

The above scaling laws are accurate when $K > \sqrt{2}$, and it should be noted that the calculated power refers to the peak power of the laser. For completeness, here we give also the modulator length scaling

$$L_{\mathrm{uM}}[\mathrm{m}] \approx 56.2\left(1+R_{\mathrm{w}}\right) \frac{B_{\mathrm{ring}}[\mathrm{T}] \Delta\epsilon_{y\mathrm{M}}[\mathrm{pm}]}{\mathcal{H}_{y\mathrm{M}}[\mu\mathrm{m}] B_{0\mathrm{M}}^3[\mathrm{T}]}. \tag{313}$$

Therefore, to lower the required modulation laser power, we can apply a large R_w, which indicates strong damping wigglers. A low beam energy and short laser wavelength are also preferred. However, their choices should consider more factors, such as IBS and engineering experience of the optical enhancement cavity, as explained in Sect. 4. A strong bending magnetic field in the ring is also desired. A weaker modulator field strength is favored to lower the required laser power. However, note that the required modulator length may be longer if we keep the quantum excitation contribution of modulators to vertical emittance $\Delta\epsilon_{y\mathrm{M}}$ unchanged in this process.

8.9 Radiation power

After microbunching is formed, now comes the radiation generation. We will use a planar undulator as a radiator. Coherent undulator radiation power at the odd-H-th harmonic from a transversely round electron beam is [17, 55]

$$P_{H,\mathrm{peak}}[\mathrm{kW}] = 1.183 N_{\mathrm{u}} H \chi [JJ]_H^2 F F_\perp(S) |b_{z,H}|^2 I_{\mathrm{P}}^2[\mathrm{A}], \tag{314}$$

where N_{u} is the number of undulator periods,

$$[JJ]_H^2 = \left[J_{\frac{H-1}{2}}(H\chi) - J_{\frac{H+1}{2}}(H\chi) \right]^2, \tag{315}$$

with $\chi = \frac{K^2}{4+2K^2}$, and the transverse form factor is

$$F F_\perp(S) = \frac{2}{\pi} \left[\tan^{-1}\left(\frac{1}{2S}\right) + S \ln\left(\frac{(2S)^2}{(2S)^2+1} \right) \right], \tag{316}$$

where $S = \frac{\sigma_\perp^2 \frac{\omega}{c}}{L_{\mathrm{u}}}$, $\sigma_\perp$ is the RMS transverse electron beam size, $b_{z,H}$ is the bunching factor at the H-th harmonic determined by the longitudinal current distribution, and I_{P} is the peak current.

The above formula was derived by assuming that the longitudinal and transverse distributions of the electron beam do not change significantly in the radiator. The energy spread of the electron beam can lead to a current distribution change inside the undulator, considering that the undulator has an $R_{56} = 2N_{\mathrm{u}}\lambda_0$, where λ_0 is the fundamental on-axis resonant wavelength of the undulator. If we consider the impact of the energy spread on coherent radiation, and assuming that the microbunching length reach its minimum at the radiator center, there will be a correction or reduction factor multiplied to the radiation power given by Eq. (314)

$$\mathcal{C} = \frac{\sqrt{\pi}}{2} \frac{\mathrm{erf}\left(\frac{\omega}{c} \sigma_\delta N_{\mathrm{u}} \lambda_0 \right)}{\frac{\omega}{c} \sigma_\delta N_{\mathrm{u}} \lambda_0}. \tag{317}$$

Here, we remind the readers that in our case of the GLSF SSMB, the beam is not transversely round at the radiator. Therefore, we used an effective round beam size to evaluate the radiation power using the above formula. This effective transverse beam size $\sigma_\perp$ is between σ_x and σ_y. For a more accurate calculation of the radiation power, we can use the real beam distribution and invoke a numerical method [17].

With all the beam physics issues properly handled, we have finally obtained a solution for a 1 kW EUV source based on GLSF SSMB, as shown in Table 2.

9 Summary

This paper discusses our efforts to obtain a solution for 1 kW EUV light source based on SSMB. Here, we provide a brief summary of this endeavor. We start by presenting the generalized Courant–Snyder formalism to build the theoretical framework for the following investigations. Based on the formalism, we conducted a theoretical minimum emittance analysis in an electron storage ring, from which we determined that to obtain a small longitudinal emittance, we need to decrease the bending angle of each bending magnet, which means increasing the number of bending magnets in the ring. In principle, we can obtain a small longitudinal emittance and a short bunch length along this line. However, there is practical limitation. To obtain a short bunch length, we need to increase the bending magnet number and lower the phase slippage factor of the ring. Using the present realizable phase slippage, which at minimum is on the order of 1×10^{-6}, a bunch length of a couple of 10 nm is the lower limit if we apply the longitudinal weak focusing regime. To compress the bunch length further, a longitudinal strong focusing regime can be invoked, similar to its transverse counterpart in the final focus of a collider, to compress the longitudinal beta function, thus the bunch length at the radiator significantly. This scheme can realize a bunch length of the nanometer level, thus allowing coherent EUV radiation generation. However, because the compression of the

longitudinal beta function requires a strong energy chirp of the electron beam, which is similar to a strong quadrupole focusing strength in the transverse dimension, the modulation laser power required is at the GW level, making the optical enhancement cavity of the SSMB only work in a low duty cycle pulsed mode, and thus limits the filling factor of the microbunched beam in the ring, and thus the average output EUV power. This leads us to the generalized longitudinal strong focusing (GLSF) regime, which is the focus of this study. The basic idea of the GLSF is to exploit the ultrasmall natural vertical emittance in a planar electron storage ring and apply partial transverse-longitudinal emittance exchange to compress the bunch length or generate high-harmonic bunching with a shallow energy modulation strength, thus lowering the requirement for the modulation laser power. The backbone of such a scheme is transverse-longitudinal phase space coupling. To find a solution based on the GLSF scheme, we first conduct a formal mathematical analysis of transverse-longitudinal coupling (TLC)-based bunch compression and harmonic generation schemes and prove three related theorems that are useful in the later choice of parameters and evaluation of laser power. We then discuss the details of different specific coupling schemes, grouping them into two categories: energy modulation-based coupling schemes and angular modulation-based coupling schemes. We derived the formulas for the bunching factor and laser-induced modulation strength in each case and used them for quantitative calculations and comparisons. Our conclusion from these analyses is that the commonly used TEM_{00} mode laser-induced energy modulation-based schemes are favored for our application in SSMB, as its requirement on the modulation laser power is lower than that in the angular modulation-based schemes. There are also various other important physical issues to be considered in finding a solution, such as the quantum excitation contribution of the GLSF section to the vertical emittance, the application of damping wigglers to speed up damping and its impact on the transverse emittance, intra-beam scattering and coherent synchrotron radiation-induced microwave instability, and the energy compensation system. The motivation of these studies is to make our choice of parameters as self-consistent as possible from a beam physics perspective. Finally, based on all the analyses and calculations, we present an example parameter set of a GLSF SSMB light source that can deliver 1 kW average-power EUV radiation. This 1 kW EUV solution as shown in Table 2 summarizes the investigations presented in this study. Our study provides a basis for the future development of SSMB.

Acknowledgements We thank Chao Feng, Ji Li, and Yujie Lu for helpful discussions on storage ring-based coherent light source, tilted laser modulation schemes, and the issue of quantum excitation contribution of damping wiggler to the horizontal beam emittance. We appreciate the helpful discussions with our colleagues in Tsinghua and other institutes.

Author Contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Xiu-Jie Deng. The first draft of the manuscript was written by Xiu-Jie Deng, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. D.F. Ratner, A.W. Chao, Steady-state microbunching in a storage ring for generating coherent radiation. Phys. Rev. Lett. **105**, 154801 (2010). https://doi.org/10.1103/PhysRevLett.105.154801
2. Y. Jiao, D.F. Ratner, A.W. Chao, Terahertz coherent radiation from steady-state microbunching in storage rings with X-band radio-frequency system. Phys. Rev. ST Accel. Beams **14**, 110702 (2011). https://doi.org/10.1103/PhysRevSTAB.14.110702
3. A. Chao, E. Granados, X. Huang, et al. High power radiation sources using the steady-state microbunching mechanism. The 7th International Particle Accelerator Conference, Busan. pp. 1048 (2016)
4. C. Tang, X. Deng, A. Chao, et al. An overview of the progress on SSMB. The 60th ICFA advanced beam dynamics workshop on future light sources, Shanghai. p. 166 (2018). https://doi.org/10.18429/JACoW-FLS2018-THP2WB02
5. X.J. Deng, A.W. Chao, W.H. Huang et al., Single-particle dynamics of microbunching. Phys. Rev. Accel. Beams **23**, 044002 (2020). https://doi.org/10.1103/PhysRevAccelBeams.23.044002
6. X.J. Deng, R. Klein, A.W. Chao et al., Widening and distortion of the particle energy distribution by chromaticity in quasi-isochronous rings. Phys. Rev. Accel. Beams **23**, 044001 (2020). https://doi.org/10.1103/PhysRevAccelBeams.23.044001
7. Y. Zhang, X.J. Deng, Z.L. Pan et al., Ultralow longitudinal emittance storage rings. Phys. Rev. Accel. Beams **24**, 090701 (2021). https://doi.org/10.1103/PhysRevAccelBeams.24.090701
8. X.J. Deng, A.W. Chao, W.H. Huang et al., Courant-Snyder formalism of longitudinal dynamics. Phys. Rev. Accel. Beams **24**, 094001 (2021). https://doi.org/10.1103/PhysRevAccelBeams.24.094001
9. Z. Li, X. Deng, Z. Pan et al., Generalized longitudinal strong focusing in a steady-state microbunching storage ring. Phys. Rev. Accel. Beams **26**, 110701 (2023). https://doi.org/10.1103/PhysRevAccelBeams.26.110701

10. C.-Y. Tsai, A.W. Chao, Y. Jiao et al., Coherent-radiation-induced longitudinal single-pass beam breakup instability of a steady-state microbunch train in an undulator. Phys. Rev. Accel. Beams **24**, 114401 (2021). https://doi.org/10.1103/PhysRevAccelBeams.24.114401
11. C.-Y. Tsai, Theoretical formulation of multiturn collective dynamics in a laser cavity modulator with comparison to Robinson and high-gain free-electron laser instability. Phys. Rev. Accel. Beams **25**, 064401 (2022). https://doi.org/10.1103/PhysRevAccelBeams.25.064401
12. C.-Y. Tsai, Longitudinal single-bunch instabilities driven by coherent undulator radiation in the cavity modulator of a steady-state microbunching storage ring. Nucl. Instrum. Methods Phys. Res. A **1042**, 167454 (2022). https://doi.org/10.1016/j.nima.2022.167454
13. Z. Zhao, N. Wang, H. Xu et al., Calculations of space-charge tune shifts in storage rings with extremely short bunches and small bunch spacing. Phys. Rev. Accel. Beams **27**, 094201 (2024). https://doi.org/10.1103/PhysRevAccelBeams.27.094201
14. X. Deng, A. Chao, J. Feikes et al., Experimental demonstration of the mechanism of steady-state microbunching. Nature **590**, 576 (2021). https://doi.org/10.1038/s41586-021-03203-0
15. A. Kruschinski, X. Deng, J. Feikes et al., Confirming the theoretical foundation of steady-state microbunching. Commun. Phys. **7**, 160 (2024). https://doi.org/10.1038/s42005-024-01657-y
16. Z. Pan, Research on optimization and design of advanced laser-driving storage ring. PhD thesis, Tsinghua University, Beijing (2020)
17. X. Deng, *Theoretical and experimental studies on steady-state microbunching*. (Springer Nature, 2024). https://doi.org/10.1007/978-981-99-5800-9
18. Y. Zhang, Research on longitudinal strong focusing SSMB ring. PhD Thesis, Tsinghua University, Beijing (2022)
19. C.X. Tang, X.J. Deng, Steady-state micro-bunching accelerator light source. Acta Phys. Sin. **71**, 152901 (2022). https://doi.org/10.7498/aps.71.20220486
20. R. Bonifacio, C. Pellegrini, L.M. Narducci, Collective instabilities and high-gain regime in a free electron laser. Optics Commun. **50**, 373 (1984). https://doi.org/10.1016/0030-4018(84)90105-6
21. A.W. Chao, Evaluation of beam distribution parameters in an electron storage ring. J. Appl. Phys. **50**, 595 (1979). https://doi.org/10.1063/1.326070
22. M.C. Wang, G.E. Uhlenbeck, On the theory of the Brownian motion II. Rev. Mod. Phys. **17**, 323 (1945). https://doi.org/10.1103/RevModPhys.17.32
23. N. Campbell, The study of discontinuous phenomena. Proc. Cambridge Philos. Soc. **15**, 117 (1909)
24. M. Sands, Synchrotron oscillations induced by radiation fluctuations. Phys. Rev. **97**, 470 (1955). https://doi.org/10.1103/PhysRev.97.470
25. K.W. Robinson, Radiation effects in circular electron accelerators. Phys. Rev. **111**, 373 (1958). https://doi.org/10.1103/PhysRev.111.373
26. E.D. Courant, H.S. Snyder, Theory of the alternating-gradient synchrotron. Ann. Phys. (N.Y.) **3**, 1 (1958). https://doi.org/10.1006/aphy.2000.6012
27. M. Sands, The physics of electron storage rings: An introduction. SLAC report SLAC-121 (1970)
28. X. Deng, Optical stochastic cooling in a general coupled lattice. in Proceedings of the 67th ICFA Adv. Beam Dyn. Workshop Future Light Sources, Luzern, Switzerland (JACoW, Geneva, 2023), p. 143. https://doi.org/10.18429/JACoW-FLS2023-TU4P30
29. L.C. Teng, Minimum emittance lattice for synchrotron radiation storage rings. Argonne National Laboratory Technical Note. No. ANL-LS-17. (1985)
30. X.J. Deng, A.W. Chao, J. Feikes et al., Breakdown of classical bunch length and energy spread formula in a quasi-isochronous electron storage ring. Phys. Rev. Accel. Beams **26**, 054001 (2023). https://doi.org/10.1103/PhysRevAccelBeams.26.054001
31. X.-Y. Lu, R. Chiche, K. Dupraz et al., 710 kW stable average power in a 45,000 finesse two-mirror optical cavity. Opt. Lett. **49**, 6884 (2024). https://doi.org/10.1364/OL.543388
32. D. Ratner, A. Chao, Reversible seeding in storage rings. in Proc. of the 33th International Free-electron Laser Conference. Shanghai, p. 57 (2011)
33. X.J. Deng, W.H. Huang, Z.Z. Li et al., Harmonic generation and bunch compression based on transverse-longitudinal coupling. Nucl. Instrum. Methods Phys. Res. A **1019**, 165859 (2021). https://doi.org/10.1016/j.nima.2021.165859
34. Dragt, Alex J. Lie methods for nonlinear dynamics with applications to accelerator physics University of Maryland (2020). https://www.physics.umd.edu/dsat/dsatliemethods.html
35. L.H. Yu, Generation of intense uv radiation by subharmonically seeded single-pass free-electron lasers. Phys. Rev. A **44**, 5178 (1991). https://doi.org/10.1103/PhysRevA.44.5178
36. A.W. Chao, Focused laser. Unpublished Note (2023)
37. H. Deng, C. Feng, Using off-resonance laser modulation for beam-energy-spread cooling in generation of short-wavelength radiation. Phys. Rev. Lett. **111**, 084801 (2013). https://doi.org/10.1103/PhysRevLett.111.084801
38. C. Feng, H. Deng, D. Wang et al., Phase-merging enhanced harmonic generation free-electron laser. New J. Phys. **16**, 043021 (2014). https://doi.org/10.1088/1367-2630/16/4/043021
39. C. Feng, Z. Zhao, A storage ring based free-electron laser for generating ultrashort coherent EUV and X-ray radiation. Sci. Rep. **7**, 4724 (2017). https://doi.org/10.1038/s41598-017-04962-5
40. B. Jiang, C. Feng, C. Li et al., A synchrotron-based kilowatt-level radiation source for EUV lithography. Sci. Rep. **12**, 3325 (2022). https://doi.org/10.1038/s41598-022-07323-z
41. C.-Y. Tsai, Analytical study of higher harmonic bunching and matrix formalism in linear high-gain free-electron laser model. Nucl. Instrum. Methods Phys. Res. A **1048**, 167974 (2023). https://doi.org/10.1016/j.nima.2022.167974
42. A.W. Chao, *Physics of collective beam instabilities in high energy accelerators* (Wiley, D, 1993)
43. D. Xiang, W. Wan, Generating ultrashort coherent soft X-ray radiation in storage rings using angular-modulated electron beams. Phys. Rev. Lett. **104**, 084803 (2010). https://doi.org/10.1103/PhysRevLett.104.084803
44. X. Wang, C. Feng, C. Yang et al., Transverse-to-longitudinal emittance-exchange in optical wavelength. New J. Phys. **22**, 063034 (2020). https://doi.org/10.1088/1367-2630/ab8e5d
45. Y. Lu, X. Wang, X. Deng et al., Methods for enhancing the steady-state microbunching in storage rings. Results Phys **40**, 105849 (2022). https://doi.org/10.1016/j.rinp.2022.105849
46. B. Jiang, J.G. Power, R. Lindberg et al., Emittance-exchange-based high harmonic generation scheme for a short-wavelength free electron laser. Phys. Rev. Lett. **106**, 114801 (2011). https://doi.org/10.1103/PhysRevLett.106.114801
47. K. Feng, C. Yu, J. Liu et al., Coherent X-ray source generation with off-resonance laser modulation. Opt. Express **26**, 19067 (2018). https://doi.org/10.1364/OE.26.019067
48. X. Wang, C. Feng, C.-Y. Tsai et al., Obliquely incident laser and electron beam interaction in an undulator. Phys. Rev. Accel. Beams **22**, 070701 (2019). https://doi.org/10.1103/PhysRevAccelBeams.22.070701
49. W. Liu, C. Feng, Y. Jiao et al., A coherent harmonic generation method for producing femtosecond coherent radiation in a laser plasma accelerator based light source. J. Synchrotron Rad. **28**, 669 (2021). https://doi.org/10.1107/S1600577521002745

50. X. Deng, J. Li, Y. Lu, Contribution of Wiggler to Radiation Integral in an Electron Storage Ring. arXiv:2410.12863 (2024). https://doi.org/10.48550/arXiv.2410.12863
51. J. Zhao, Undulator in SLIM. Unpublished Note. (Oct. 2023)
52. K. Bane, A simplified model of intrabeam scattering, in Proceedings of the 8th European Particle Accelerator Conference, Paris, France, 2002 (EPS-IGA and CERN, Geneva, 2002), p. 1443
53. K.F. Bane, Y. Cai, G. Stupakov, Threshold studies of the microwave instability in electron storage rings. Phys. Rev. ST Accel. Beams **13**, 104402 (2010). https://doi.org/10.1103/PhysRevSTAB.13.104402
54. J. Feikes, M. von Hartrott, M. Ries et al., Metrology Light Source: The first electron storage ring optimized for generating coherent THz radiation. Phys. Rev. ST Accel. Beams **14**, 030705 (2011). https://doi.org/10.1103/PhysRevSTAB.14.030705
55. X.J. Deng, Y. Zhang, Z.L. Pan et al., Average and statistical properties of coherent radiation from steady-state microbunching. J. Synchrotron Radiat. **30**, 35 (2023). https://doi.org/10.1107/S1600577522009973
56. N. Nakamura, R. Kato, H. Sakai et al., High-power EUV free-electron laser for future lithography. Jpn. J. Appl. Phys. **62**, SG0809 (2023). https://doi.org/10.35848/1347-4065/acc18c