ANL-AAI-PUB-2007-002
December 2006

A Review of X-ray Free-Electron Laser Theory

Zhirong Huang
Stanford Linear Accelerator Center, Stanford University, Stanford, CA 94309

Kwang-Je Kim
Argonne National Laboratory, 9700 S. Cass Ave, Argonne, IL, 60439

ABSTRACT

High-gain free-electron lasers (FELs) are being developed as extremely bright sources for a nextgeneration x-ray facility. In this paper, we review the basic theory of the start-up, the exponential growth, and the saturation of the high-gain process, emphasizing the self-amplified spontaneous emission. The radiation characteristics of an x-ray FEL, including its transverse coherence, temporal characteristics, and harmonic content, are discussed. FEL performance in the presence of machine errors and undulator wakefields is examined. Various enhancement schemes through seeding and beam manipulations are summarized.

Review of x-ray free-electron laser theory

Zhirong Huang
Stanford Linear Accelerator Center, Stanford, California 94309, USA

Kwang-Je Kim
Argonne National Laboratory, Argonne, Illinois 60439, USA
(Received 25 August 2006; published 12 March 2007)

High-gain free-electron lasers (FELs) are being developed as extremely bright sources for a next-generation x-ray facility. In this paper, we review the basic theory of the start-up, the exponential growth, and the saturation of the high-gain process, emphasizing the self-amplified spontaneous emission. The radiation characteristics of an x-ray FEL, including its transverse coherence, temporal characteristics, and harmonic content, are discussed. FEL performance in the presence of machine errors and undulator wakefields is examined. Various enhancement schemes through seeding and beam manipulations are summarized.

DOI: 10.1103/PhysRevSTAB.10.034801 PACS numbers: 41.60.Cr

I. INTRODUCTION

Free-electron lasers (FELs), invented by John Madey [1] and subsequently demonstrated experimentally by his group at Stanford University in the 1970s [2], hold great promise as tunable, high-power, coherent sources for short-wavelength radiation. To circumvent the need for mirrors or coherent seeds, the initial random field of spontaneous radiation in an undulator may be amplified in the medium of a bright electron beam traveling through a long undulator to intense, quasicoherent radiation [3–5]. In the x-ray wavelength range (from a few nm down to 1 Å or less), a high-gain FEL operated in this self-amplified spontaneous emission (SASE) mode can generate multigigawatt (GW) and femtosecond (fs) coherent x-ray pulses. The extreme high power together with the excellent transverse coherence of such x-ray sources provide about 10 orders of magnitude improvement in peak brightness above that offered by the current synchrotron radiation sources based on electron storage rings, making FELs suitable probes for both the ultrasmall and the ultrafast worlds. Tremendous progress in accelerator and FEL technologies has been made in past years towards realizing such a "fourth-generation" radiation facility, demonstrated by the sequence of recent SASE FEL experiments at visible and ultraviolet wavelengths [6–9]. More recently, the VUV-FEL at DESY, now called FLASH, reported FEL lasing at wavelengths down to 13 nm [10]. Because of these successes, several x-ray FEL projects are either under construction or being proposed (see, e.g., Refs. [11–14]).

This paper reviews the basic theory behind the x-ray FELs and points out possible improvement of these sources. In Sec. II we discuss qualitatively how the coherent radiation is amplified and developed from the initial seed signal or the electron shot noise. It is then followed by mathematical analysis of the FEL equations and their solutions in Sec. III, including diffraction, optical guiding, and effects of beam emittance and energy spread. The main characteristics of x-ray FELs, including coherence properties and harmonic content, are presented in Sec. IV, where the temporal manipulation and seeding schemes are also briefly discussed. In Sec. V we analyze degrading effects of undulator errors and wakefields and study their tolerances or compensations. Several electron beam-manipulation methods to enhance the SASE performance are described in Sec. VI. We conclude the paper with final remarks in Sec. VII.

II. QUALITATIVE DISCUSSIONS

The radiation discussed in this paper is generated in a periodic magnetic device called an "undulator." Consider a planar undulator for which the magnetic field is in the vertical y direction and varies sinusoidally along the z direction:

$$B_y = B_0 \sin(k_u z). \tag{1}$$

Here $k_u = 2\pi/\lambda_u$, λ_u is the undulator period, and B_0 is the magnetic pole field. As shown in Fig. 1, a relativistic electron entering the undulator will wiggle periodically in the horizontal x direction and can spontaneously emit radiation at the resonant wavelength (see, e.g., Ref. [15])

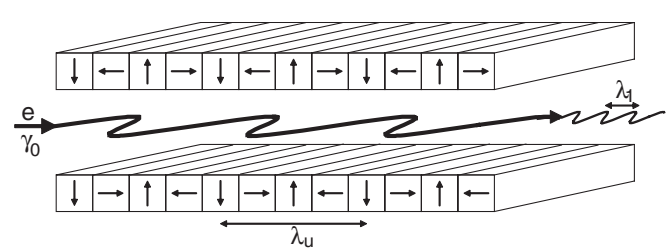


FIG. 1. A wiggling electron in a planar undulator emits undulator radiation.

 © 2007 The American Physical Society

$$\lambda_r = \frac{\lambda_u}{2\gamma_0^2}\left(1 + \frac{K_0^2}{2} + \gamma_0^2\phi^2\right). \tag{2}$$

Here γ_0 is the electron energy in units of the rest energy mc^2,

$$K_0 = \frac{eB_0}{mck_u} = 0.934B_0[\text{Tesla}]\lambda_u[\text{cm}] \tag{3}$$

is the dimensionless undulator strength parameter, and ϕ is the observation angle relative to the undulator z axis. Spontaneous undulator radiation at the resonant wavelength λ_r (and its associated harmonics) is the workhorse of the third-generation synchrotron facilities.

The electromagnetic (EM) wave in the forward direction copropagates with the electron beam and may exchange energy with the electrons. In free space, the interaction cannot be sustained because the EM wave is always faster than the electrons. In the undulator, the copropagating radiation overtakes the electrons in one undulator period by λ_1, where

$$\lambda_1 = \lambda_r(\phi = 0) = \frac{\lambda_u}{2\gamma_0^2}\left(1 + \frac{K_0^2}{2}\right) = \frac{2\pi}{k_1} = \frac{2\pi c}{\omega_1} \tag{4}$$

is the resonant wavelength in the forward direction. In this case, due to the periodicity of the system, the interaction with a plane EM wave carrying the resonant wavelength λ_1 can be sustained as shown in Fig. 2. Depending on the relative phase of the electrons to the plane wave, some electrons gain energy from the radiation while other electrons lose energy to the radiation, hence the energy of a long electron bunch is periodically modulated at λ_1. As faster electrons (with higher energies) catch up with the slower electrons, a periodic density modulation at the radiation wavelength (the so-called "microbunching") begins to develop in the undulator. Under favorable conditions, the microbunched electron beam emits coherent radiation at the expense of the beam kinetic energy; then the EM wave gains net energy and the FEL amplification occurs.

If the total energy gain in the undulator is a small fraction of the EM energy, the FEL is said to operate in the low-gain regime. In this case, an FEL oscillator using an optical cavity together with many electron bunches is necessary to build up the radiation intensity. Such FEL oscillators have been constructed and used successfully in the visible and longer wavelengths (e.g., Ref. [16]). For a sufficiently bright electron beam and a sufficiently long undulator, the collective interaction of the beam-radiation system leads to an exponential growth of the radiation intensity along the undulator distance as illustrated in Fig. 3. Such a high-gain FEL does not require any optical cavity and can amplify either an input seed signal or the spontaneous undulator radiation produced by the electron shot noise. Thus, in the x-ray wavelength range where both mirrors and coherent input sources are difficult to obtain, intense, quasicoherent radiation can be generated by the SASE process. The exponential gain eventually stops as the beam loses enough energy to upset the resonant condition. Both the radiation intensity and the electron beam microbunching reach a maximum saturation level (see Fig. 3).

A fundamental scaling parameter for a high-gain FEL is the dimensionless Pierce parameter ρ defined as [5]

$$\rho = \left[\frac{K_0^2[\text{JJ}]^2}{32}\frac{k_p^2}{k_u^2}\right]^{1/3} = \left[\frac{1}{16}\frac{I_e}{I_A}\frac{K_0^2[\text{JJ}]^2}{\gamma_0^3\sigma_x^2k_u^2}\right]^{1/3}, \tag{5}$$

where the Bessel function factor [JJ] is equal to $[J_0(\xi) - J_1(\xi)]$ with $\xi = K_0^2/(4 + 2K_0^2)$ for a planar undulator trajectory as shown in Fig. 1. A helical undulator containing both the vertical and horizontal magnetic fields can produce a helical undulator trajectory (see, e.g., Ref. [15]). In this case, we should take $[\text{JJ}] = 1$ in Eq. (5). $k_p = \sqrt{2I_e/(\gamma_0^3 I_A \sigma_x^2)}$ is the longitudinal plasma oscillation

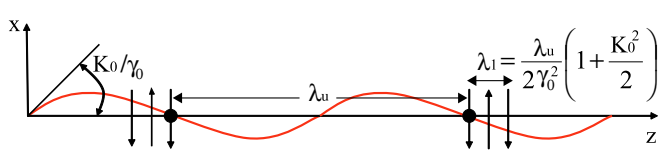


FIG. 2. (Color) After an electron (black dot) travels one undulator period λ_u of the sinusoidal trajectory (in red), a plane wave (represented by alternating vertical arrows) overtakes the electron by one resonant wavelength λ_1. Thus, the undulator radiation carrying this resonant wavelength can exchange energy with the electron over many undulator periods.

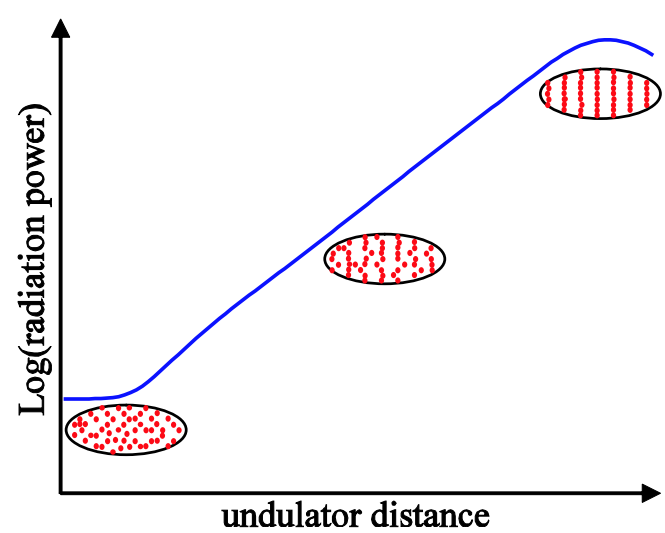


FIG. 3. (Color) Growth of the radiation power and the electron beam microbunching as a function of the undulator distance for a high-gain FEL.

wave number, I_e is the electron peak current, $I_A = ec/r_e \approx 17$ kA is the Alfvén current, $r_e \approx 2.8 \times 10^{-15}$ m is the classical electron radius, and σ_x is the rms transverse size of the electron beam. In terms of this parameter, the one-dimensional (1D) power gain length of a monoenergetic beam is

$$L_{G0} = \frac{\lambda_u}{4\pi\sqrt{3}\rho}. \tag{6}$$

The relative FEL bandwidth at saturation is close to ρ, and the saturation power is about ρ times the electron beam power.

As the electron beam develops a periodic microbunching with the modulation wavelength λ_1, the longitudinal space-charge field between electrons tends to counteract the bunching process if the reduced plasma oscillation wavelength k_p^{-1} is comparable to the FEL gain length as given in Eq. (6). Examination of Eq. (5) for $K_0 \sim 1$ shows that this condition requires that $k_p \to k_u$ and that $\rho \to 1$. In typical short-wavelength FELs using high-energy electron beams, $\rho \sim 10^{-3}$, hence we can neglect beam self-fields at these wavelengths and focus on the beam-radiation interaction.

III. ANALYTICAL TREATMENTS

In this section, we illustrate analytical methods to treat the high-gain FELs. The beam-radiation interaction in the undulator can usually be described by classical physics. Quantum effects are expected to be small and will be discussed at the end of this section.

A. Electron motion in the presence of undulator radiation

Consider an ultrarelativistic electron traversing in the undulator magnetic field described by Eq. (1). In a first approximation, we take the longitudinal velocity $v_z \approx c$ and apply the Lorentz equation in the horizontal direction:

$$\gamma m \frac{dv_x}{dt} = -ev_z B_y \approx -ecB_0 \sin(k_u z). \tag{7}$$

Here γmc^2 is the electron energy that can be slightly different from the reference energy $\gamma_0 mc^2$. Since γ is a constant in a magnetic field, we can integrate Eq. (7) to obtain

$$\begin{aligned} v_x &= \frac{eB_0}{\gamma m k_u}\cos(k_u z) = \frac{K_0 c}{\gamma}\cos(k_u z), \\ v_z &= c\sqrt{1 - \frac{1}{\gamma^2} - \frac{v_x^2}{c^2}} \\ &\approx \underbrace{c\left(1 - \frac{1 + K_0^2/2}{2\gamma^2}\right)}_{\equiv \bar{v}_z} - \frac{K_0^2 c}{4\gamma^2}\cos(2k_u z), \end{aligned} \tag{8}$$

where $\bar{v}_z$ is the average longitudinal velocity over an undulator period. In the presence of a horizontal electric field $E_x = E_0 \cos(k_1 z - \omega_1 t + \psi_0)$, the change of the electron energy is given by

$$\begin{aligned} mc^2 \frac{d\gamma}{dt} &= ev_x E_x \\ &= \frac{eE_0 K_0 c}{2\gamma}\{\cos[(k_1 + k_u)z - \omega_1 t + \psi_0] \\ &\quad + \cos[(k_1 - k_u)z - \omega_1 t + \psi_0]\}. \end{aligned} \tag{9}$$

Here E_0 and ψ_0 are the initial amplitude and phase of the electric field, respectively.

It is convenient to use the distance z from the undulator entrance as the independent variable and change the time variable to a phase variable relative to the EM wave:

$$\theta = (k_1 + k_u)z - \omega_1 \bar{t}, \tag{10}$$

where $\bar{t} = \int dz/\bar{v}_z$ is the electron arrival time averaged over the undulator period at z. The phase change can be calculated as

$$\frac{d\theta}{dz} = k_1 + k_u - \frac{\omega_1}{\bar{v}_z} = k_u - k_1 \frac{1 + K_0^2/2}{2\gamma^2}. \tag{11}$$

Defining a relative energy variable $\eta = (\gamma - \gamma_0)/\gamma_0 \ll 1$ and inserting the resonant condition Eq. (4) into Eq. (11), we obtain

$$\frac{d\theta}{dz} = 2k_u \eta, \tag{12}$$

i.e., the electron's phase relative to the EM wave remains constant if its energy satisfies the resonant condition [i.e., Eq. (4)]. Thus, the first term on the right-hand side of Eq. (9) varies slowly, contributing to the resonant energy exchange, while the second term varies quickly, being oscillatory with the period $2\lambda_u$. Properly taking into account the fact that the electron's longitudinal motion also has an oscillatory part as given in Eq. (8), Eq. (9) after retaining only the slowly varying part becomes

$$\frac{d\eta}{dz} = \frac{eK_0[\mathrm{JJ}]}{2\gamma_0^2 mc^2} E_0 \cos(\theta + \psi_0). \tag{13}$$

Equations (12) and (13), known as the "pendulum equations," describe the motion of electrons under the influence of the "ponderomotive potential" due to the combined undulator and radiation fields [17]. The motion of electrons in the (θ, η) phase space under the influence of the ponderomotive potential is illustrated in Fig. 4. A nearly monoenergetic electron beam with its energy satisfying the resonant condition develops an energy modulation at the resonant wavelength according to Eq. (13). After a certain undulator distance, the energy modulation is turned

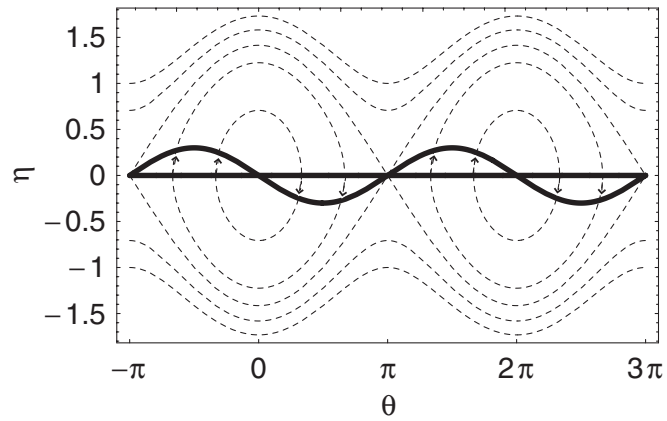


FIG. 4. Electron motion in the longitudinal phase space (θ, η) due to the presence of a resonant EM wave (with an initial phase $\psi_0 = \pi/2$) in the undulator. An initial distribution of the electron beam, shown as a straight line at $\eta = 0$, changes into a distribution on a sinusoidal line, implying that the energy and the density of the electron beam is modulated, i.e., microbunched. The dashed lines are the phase space trajectories.

into a density modulation as the relative longitudinal position of an electron changes by an amount determined by its energy deviation from the resonant energy [see Eq. (12)]. Note that the net energy exchange is still zero, as the number of electrons gaining energy is the same as the number of electrons losing energy. However, such a microbunched beam changes the phase of the EM wave (through the Maxwell equation, see Sec. III C) so that the buckets shift to the left. As a result, the electron beam begins to lose its average energy and the radiation intensity starts to grow exponentially further along the undulator.

Since the FEL interaction is a resonant energy exchange between the electron and the radiation field, the evolution of the electrons' phase may affect the FEL performance critically. For example, a beam with an initial energy spread can cause a phase spread through Eq. (12) that degrades the microbunching process. As we will discuss next, the betatron motion of a finite-emittance beam introduces another mechanism for the phase spread.

B. Electron focusing and emittance effect

A beam with a finite emittance ε_x has an rms angular spread $\sigma_{x'} = \varepsilon_x/\sigma_x$, where σ_x is the rms beam size and will expand its size in free space. Hence, the electron beam in a long undulator channel should be properly focused to keep the beam size nearly constant for the effective FEL interaction. We discuss two types of undulator focusing: "natural" focusing and "strong" focusing, and study their effects on the FEL interaction.

1. Natural focusing

Equation (1) for the undulator magnetic field is valid only near the $y = 0$ midplane. An exact solution of the Maxwell equation reducing to Eq. (1) for $y = 0$ is [18]

$$\boldsymbol{B} = [0, B_0 \sin(k_u z)\cosh(k_u y), B_0 \cos(k_u z)\sinh(k_u y)]. \tag{14}$$

This magnetic field is a reasonable representation of a planar undulator with wide and flat pole faces.

The Lorentz force on the electron is given by

$$\gamma m\left(\frac{d^2x}{dt^2}, \frac{d^2y}{dt^2}, \frac{d^2z}{dt^2}\right) = e\boldsymbol{v} \times \boldsymbol{B} = e\left(B_z\frac{dy}{dt} - B_y\frac{dz}{dt}, -B_z\frac{dx}{dt}, B_y\frac{dx}{dt}\right). \tag{15}$$

A rigorous derivation of the linear and nonlinear equations of motion in the undulator is presented in Ref. [19]. Here we expand Eq. (15) up to the second order in x and y. The zeroth-order solution is the wiggling motion

$$x = \frac{K_0}{\gamma_0 k_u}\sin(k_u z), \qquad y = 0. \tag{16}$$

The solution up to the second order can be written as

$$x = \frac{K_0}{\gamma_0 k_u}\sin(k_u z)(1 + \chi) + x_{\beta n}, \qquad y = y_{\beta n}, \tag{17}$$

where χ is the correction term to the zeroth-order wiggling motion, and $(x_{\beta n}, y_{\beta n})$ denotes a slow evolution of the trajectory superimposed on the fast wiggling motion. Inserting Eq. (17) into Eq. (15), keeping terms up to the second order, and separating the fast oscillation with the slow motion, we obtain

$$\chi = \frac{k_u^2 y_{\beta n}^2}{2}, \qquad \frac{d^2x_{\beta n}}{dz^2} = 0, \qquad \frac{d^2y_{\beta n}}{dz^2} \approx -\left(\frac{K_0^2 k_u^2}{2\gamma_0^2}\right)y_{\beta n} \equiv -k_{n0}^2 y_{\beta n}, \tag{18}$$

where $k_{n0} = K_0 k_u/(\sqrt{2}\gamma_0) \ll k_u$ is the natural-focusing strength of the undulator, and we have changed the independent variable from t to z for convenience. The vertical motion is indeed focused. The focusing is due to the intrinsic property of the periodic nature of the undulator magnetic field and is referred to as natural focusing. Note that focusing is absent in the x direction because of the uniform extent of the undulator in this direction.

We now calculate the average longitudinal velocity, as it will determine the electron's phase relative to the EM wave through Eq. (11). From Eq. (8), we have

$$\frac{v_z}{c} \approx 1 - \frac{1}{2\gamma^2} - \frac{v_x^2 + v_y^2}{2c^2}. \tag{19}$$

Here

$$v_x \approx c\frac{dx}{dz} = \frac{K_0 c}{\gamma_0}\cos(k_u z)\left(1 + \frac{k_u^2 y_{\beta n}^2}{2}\right),$$
$$v_y \approx c\frac{dy}{dz} = c\frac{dy_{\beta n}}{dz} \equiv c p_y. \tag{20}$$

We can again average Eq. (19) over the fast oscillation and retain terms up to the second order in y to obtain

$$\frac{\bar{v}_z}{c} = 1 - \frac{1 + K_0^2/2}{2\gamma^2} - \frac{k_{n0}^2 y_{\beta n}^2 + p_y^2}{2}. \tag{21}$$

To generate focusing in the x direction, one can shape the undulator pole faces to be parabolic [20]. The natural-focusing strengths in the x and y directions can be shown to satisfy

$$k_{nx}^2 + k_{ny}^2 = k_{n0}^2. \tag{22}$$

Typically, one wants equal focusing in both transverse directions, in which case $k_{nx} = k_{nx} = k_{n0}/\sqrt{2} \equiv k_n$. The average longitudinal velocity is then [20]

$$\frac{\bar{v}_z}{c} = 1 - \frac{1 + K_0^2/2}{2\gamma^2} - \frac{k_n^2 \boldsymbol{x}^2 + \boldsymbol{p}^2}{2}$$
$$= 1 - \frac{1 + K_0^2/2}{2\gamma^2} - \frac{J_x + J_y}{\beta_n}, \tag{23}$$

where $\boldsymbol{x} = (x_{\beta n}, y_{\beta n})$, $\boldsymbol{p} = (dx_{\beta n}/dz, dy_{\beta n}/dz)$, and $\beta_n = 1/k_n = \gamma_0 \lambda_u/(\pi K_0)$ is the natural-focusing beta function in either transverse direction. We have also introduced the transverse actions, $J_x = \beta_n(k_n^2 x^2 + p_x^2)/2$ and $J_y = \beta_n(k_n^2 y^2 + p_y^2)/2$, which are invariants of the transverse motion.

2. Strong focusing

The natural focusing is usually too weak to be effective for the high-energy electrons that drive an x-ray FEL. Thus, quadrupole magnets are inserted in the undulator section breaks to provide the necessary strong focusing, normally in the form of a FODO lattice. A FODO cell (a single lattice period) consists of a focusing (F) and a defocusing (D) quadrupole with drift spaces (O) in between. In Fig. 5 we show two FODO cells with the drift spaces replaced by undulator sections. Following the standard accelerator notation (e.g., Ref. [15]), the electron trajectory can be described by

$$x_\beta(z) = \sqrt{2J_x\beta_x}\cos\Phi_x(z),$$
$$p_{x\beta}(z) \equiv \frac{dx_\beta}{dz} = -\sqrt{\frac{2J_x}{\beta_x}}[\sin\Phi_x(z) + \alpha_x\cos\Phi_x(z)], \tag{24}$$

where β_x is the horizontal beta function, $\alpha_x = -d\beta_x/(2dz)$, and Φ_x is the betatron phase advance. Similar equations describe the vertical betatron motion.

To avoid a large beam size variation in the undulator, the FODO lattice is usually designed to have a small phase advance per cell, i.e., the FODO lattice period L_C is much smaller than the average beta function $\bar{\beta}$ (usually the same in both transverse planes). Such a FODO lattice has the properties

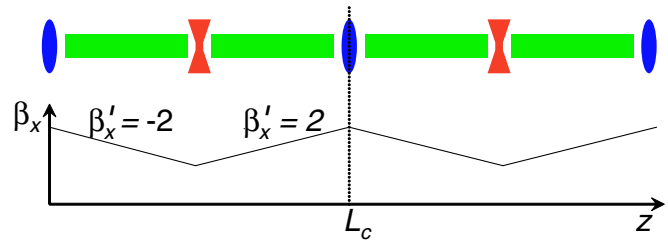


FIG. 5. (Color) Variation of the horizontal beta function along the distance of two FODO cells for a small phase advance per cell. Its derivative $\beta_x' \equiv d\beta_x/dz$ is close to the values ± 2, but the deviation of β_x from the average value $\bar{\beta}$ is relatively small. The FODO cell length L_c is assumed to be much smaller than the average beta function $\bar{\beta}$.

$$\alpha_x = -\frac{1}{2}\frac{d\beta_x}{dz} \approx \pm 1, \qquad \alpha_y = -\frac{1}{2}\frac{d\beta_y}{dz} \approx \mp 1, \tag{25}$$

with the sign alternating per half cell. Figure 5 illustrates the variation of the beta function in a FODO lattice with a small phase advance per cell.

Neglecting the natural focusing, the longitudinal velocity averaged over several undulator periods in a strong focusing undulator is

$$\frac{\bar{v}_z}{c} \approx 1 - \frac{1 + K_0^2/2}{2\gamma^2} - \frac{\boldsymbol{p}_\beta^2}{2}, \tag{26}$$

where

$$\boldsymbol{p}_\beta^2 = p_{x\beta}^2 + p_{y\beta}^2$$
$$= \frac{2J_x}{\beta_x}[1 \pm \sin(2\Phi_x)] + \frac{2J_y}{\beta_y}[1 \mp \sin(2\Phi_y)]. \tag{27}$$

Because the oscillatory terms in Eq. (27) change sign every half FODO lattice period, we can further average the longitudinal velocity over the FODO cell length if it is smaller than the power gain length to obtain [21]

$$\frac{\bar{v}_z}{c} \approx 1 - \frac{1 + K_0^2/2}{2\gamma^2} - \frac{J_x + J_y}{\bar{\beta}}. \tag{28}$$

Note that Eq. (28) has the same form as Eq. (23), with the electron's average longitudinal velocity depending linearly on its transverse actions (constants of the motion in the undulator). This simplifies the theoretical treatment of the emittance effect in an undulator with either natural focusing (used in many low- and medium-energy FELs) or strong focusing (commonly used in x-ray FELs). Although $J_{x,y}$ are constants of motion for each electron, different electrons with different betatron amplitudes have different transverse actions. In fact, the ensemble average of $J_{x,y}$ over all electrons is the rms transverse emittance of

the beam, i.e.,

$$\langle J_{x,y}\rangle = \varepsilon_{x,y}. \tag{29}$$

Because of the finite emittance, electrons with different betatron amplitudes spread out in longitudinal positions (and phases) relative to the EM wave and hence degrade the FEL gain. Inserting Eq. (28) into Eq. (11), we obtain the phase equation in the presence of the betatron motion as

$$\frac{d\theta}{dz} = 2k_u\eta - \frac{k_1(J_x + J_y)}{\bar{\beta}}. \tag{30}$$

C. Paraxial wave equation

In the high-gain regime, the radiation intensity increases rapidly along the length of the undulator. The equations of motion must be solved together with the Maxwell equation for the transverse radiation field:

$$\left[\left(\frac{1}{c}\frac{\partial}{\partial t}\right)^2 - \left(\frac{\partial}{\partial z}\right)^2 - \nabla_\perp^2\right]E_x(\boldsymbol{x}, t; z) = -\frac{1}{\epsilon_0 c^2}\left[\frac{\partial j_x}{\partial t} + c^2\frac{\partial(en_e)}{\partial x}\right], \tag{31}$$

where $\nabla_\perp^2$ is the transverse Laplacian, ϵ_0 is the permittivity of free space, n_e is the electron volume density,

$$j_x = ev_x n_e = eK_0\cos(k_u z)\sum_{j=1}^{N_e}\frac{1}{\gamma_j}\delta[\boldsymbol{x} - \boldsymbol{x}_j(z)]\delta[t - t_j(z)] \tag{32}$$

is the transverse current, and N_e is the total number of electrons in the beam.

As we will discuss later, in order to have efficient FEL interaction, both the e-beam and the radiation cross sections should be reasonably matched with the rms size approximately given by

$$\sigma_x \sim \sigma_r \sim \sqrt{\frac{\lambda_1 2L_{G0}}{4\pi}} \sim \frac{1}{4\pi}\sqrt{\frac{\lambda_1\lambda_u}{\rho}}. \tag{33}$$

For short-wavelength FELs, the rms size of the radiation cross section is usually much smaller than the vacuum chamber radius; hence we may neglect any boundary conditions in solving the Maxwell equation. Furthermore, we have

$$j_x = ev_x n_e \sim \frac{cK}{\gamma_0}en_e, \qquad \frac{\partial}{\partial x} \sim \frac{1}{\sigma_x} \sim 4\pi\sqrt{\frac{\rho}{\lambda_1\lambda_u}}, \qquad \frac{\partial}{\partial t} \sim \omega_1 \sim \frac{2\pi c}{\lambda_1}, \tag{34}$$

then the ratio

$$\frac{ec^2\partial n_e/\partial x}{\partial j_x/\partial t} \sim \sqrt{\rho} \ll 1. \tag{35}$$

Thus, the charge density term in the Maxwell equation (31) may be dropped.

It is convenient to introduce the frequency domain field amplitude E_ν via

$$E_x(\boldsymbol{x}, t; z) = \frac{1}{2}\int d\nu E_\nu(\boldsymbol{x}; z)e^{i\nu k_1(z-ct)} + \text{c.c.}, \tag{36}$$

where c.c. stands for complex conjugate. Although the increase in the radiation amplitude over the entire length of the undulator is large, we will assume that the change is small over a few undulator period. The radiation spectrum will then be narrowly peaked around the resonant frequency. In other words, $\Delta\nu \equiv \nu - 1$ is much smaller compared to unity, and

$$\left|\frac{\partial^2 E_\nu}{\partial z^2}\right| \ll \left|2\nu k_1\frac{\partial E_\nu}{\partial z}\right|. \tag{37}$$

The Maxwell equation now becomes the paraxial wave equation [22]

$$(2i\nu k_1\frac{\partial}{\partial z} + \nabla_\perp^2)E_\nu(\boldsymbol{x}; z) = \frac{1}{\epsilon_0 c^2}\int_{-\infty}^{\infty}\frac{ck_1 dt}{\pi}e^{-i\nu k_1(z-ct)}\frac{\partial}{\partial t}j_x(\boldsymbol{x}, z, t) \approx -\frac{e}{\epsilon_0}K\int_{-\infty}^{\infty}\frac{\nu k_1 dt}{\pi}e^{-i\nu k_1(z-ct)}\frac{i\nu k_1}{\gamma_0}\cos(k_u z)\sum_{j=1}^{N}\delta(\boldsymbol{x} - \boldsymbol{x}_j)\delta(t - t_j). \tag{38}$$

Here we have approximated $\gamma_j = \gamma_0$ in the transverse velocity of the beam (assuming the energy spread is small) and performed integration by parts over the time variable.

Equation (38) can be further simplified as follows. We change the integration variable from t to θ according to Eq. (10) and average the right-hand side properly over the fast wiggling motion to obtain

$$\left(\frac{\partial}{\partial z} + \frac{\nabla_\perp^2}{2i\nu k_1}\right)E_\nu(\boldsymbol{x}; z) = -\frac{eK[\text{JJ}]}{2\epsilon_0\gamma_0}e^{i\Delta\nu k_u z}\int\frac{k_1 d\theta}{2\pi}e^{-i\nu\theta}\sum_{j=1}^{N_e}\delta(\boldsymbol{x} - \boldsymbol{x}_j)\delta(\theta - \theta_j). \tag{39}$$

Here θ_j is the position of the jth electron relative to the bunch center in units of $\lambda_1/(2\pi)$ at the undulator distance z, and the left-hand side ν can be replaced by 1 around the fundamental radiation frequency. Although $\Delta\nu = \nu - 1 \ll 1$, we cannot ignore $\Delta\nu k_u z$ in the exponent of Eq. (39) as $k_u z \gg 1$ for a high-gain undulator. Section IV C gives a derivation of the [JJ] factor including harmonic emissions.

D. Coupled Maxwell-Klimontovich equations

To take into account the discreteness of electrons that initiates the SASE process, we use the Klimontovich distribution function to describe the microscopic electron distribution in the phase space (see, e.g., Ref. [23]):

$$F(\theta, \eta, \boldsymbol{x}, \boldsymbol{p}; z) = \frac{k_1}{n_0}\sum_{j=1}^{N_e}\delta(\theta-\theta_j)\delta(\eta-\eta_j)\delta(\boldsymbol{x}-\boldsymbol{x}_j) \times \delta(\boldsymbol{p}-\boldsymbol{p}_j), \tag{40}$$

where n_0 is the peak electron volume density. The evolution of the Klimontovich distribution function F is governed by the continuity equation

$$\frac{\partial F}{\partial z} + \frac{\partial F}{\partial \theta}\frac{d\theta}{dz} + \frac{\partial F}{\partial \eta}\frac{d\eta}{dz} + \frac{\partial F}{\partial \boldsymbol{x}}\frac{d\boldsymbol{x}}{dz} + \frac{\partial F}{\partial \boldsymbol{p}}\frac{d\boldsymbol{p}}{dz} = 0. \tag{41}$$

Here $\boldsymbol{x}$ and $\boldsymbol{p}$ are the averaged transverse variables that satisfy the smooth focusing approximation, i.e.,

$$\frac{d\boldsymbol{x}}{dz} = \boldsymbol{p}, \qquad \frac{d\boldsymbol{p}}{dz} = -\frac{\boldsymbol{x}}{\bar{\beta}} \equiv -k_\beta \boldsymbol{x}. \tag{42}$$

The averaging is over the (FODO) lattice period for strong focusing. For natural focusing, the averaging is over the undulator period, and k_β should be replaced by k_n.

Using this set of smoothed transverse variables, the phase Eq. (30) for a strong focusing undulator becomes

$$\frac{d\theta}{dz} = 2k_u\eta - \frac{k_1}{2}(k_\beta^2\boldsymbol{x}^2 + \boldsymbol{p}^2). \tag{43}$$

The energy equation can also be rewritten as

$$\frac{d\eta}{dz} = \frac{eK[\text{JJ}]}{4\gamma_0^2 mc^2}\int d\nu e^{i\nu\theta}e^{-i\Delta\nu k_u z}E_\nu(\boldsymbol{x}, z) + \text{c.c.} \tag{44}$$

In terms of the electron distribution function F, the paraxial wave equation (39) becomes

$$\left(\frac{\partial}{\partial z} + \frac{\boldsymbol{\nabla}_\perp^2}{2ik_1}\right)E_\nu(\boldsymbol{x}; z) = -\frac{eK[\text{JJ}]n_0}{2\epsilon_0\gamma_0}e^{i\Delta\nu k_u z}\int\frac{d\theta}{2\pi}e^{-i\nu\theta}\int d^2p\int d\eta F. \tag{45}$$

We note that the fluid limit of the Klimontovich distribution function F satisfies the same continuity equation, often referred to as the Vlasov equation in the literature. Here, we retain the microscopic description of the electrons in order to describe the SASE radiation initiated by the electron shot noise. Equation (41), with F given by Eq. (40), will be referred to as the Klimontovich equation.

Making use of the fundamental FEL scaling parameter ρ defined in Eq. (5), we introduce the following scaled variables to simplify our equations:

$$\begin{aligned} &\hat{z} = 2\rho k_u z, \quad \hat{\eta} = \frac{\eta}{\rho}, \quad \Delta\hat{\nu} = \frac{\Delta\nu}{2\rho}, \quad \hat{\boldsymbol{x}} = \boldsymbol{x}\sqrt{2k_1k_u\rho}, \\ &\hat{\boldsymbol{p}} = \boldsymbol{p}\sqrt{\frac{k_1}{2k_u\rho}}, \quad a_\nu = \frac{eK[\text{JJ}]}{4\gamma_0^2 mc^2 k_u\rho}e^{-i\Delta\nu k_u z}E_\nu, \quad f = \frac{2k_u\rho^2}{k_1}F. \end{aligned} \tag{46}$$

The coupled Maxwell-Klimontovich equations using these scaled variables are

$$\left(\frac{\partial}{\partial\hat{z}} + i\Delta\hat{\nu} + \frac{\hat{\boldsymbol{\nabla}}_\perp^2}{2i}\right)a_\nu(\hat{\boldsymbol{x}}; \hat{z}) = -\int\frac{2\rho d\theta}{2\pi}e^{-i\nu\theta}\int d^2\hat{p}\int d\hat{\eta}f(\theta, \hat{\eta}, \hat{\boldsymbol{x}}, \hat{\boldsymbol{p}}; \hat{s}), \tag{47}$$

$$\frac{\partial f}{\partial\hat{z}} + \theta'\frac{\partial f}{\partial\theta} + \hat{\boldsymbol{p}}\frac{\partial f}{\partial\hat{\boldsymbol{x}}} - \hat{k}_\beta^2\hat{\boldsymbol{x}}\frac{\partial f}{\partial\hat{\boldsymbol{p}}} + \left(\int d\Delta\hat{\nu}e^{i\nu\theta}a_\nu + \text{c.c.}\right)\frac{\partial f}{\partial\hat{\eta}} = 0, \tag{48}$$

where

$$\theta' = \frac{d\theta}{d\hat{z}} = \hat{\eta} - \frac{\hat{\boldsymbol{p}}^2 + \hat{k}_\beta^2\hat{\boldsymbol{x}}^2}{2}, \tag{49}$$

and $\hat{k}_\beta = k_\beta/(2k_w\rho)$ is the scaled focusing strength.

Throughout this paper, we denote most of the scaled (dimensionless) parameters and variables with caret, except for a few special cases such as ρ, a_ν, and f. A list of symbols and their physical meanings is given in Appendix A.

E. Solution in the exponential growth regime

Equation (48) can be linearized in the small signal regime before saturation when the scaled radiation field

is small, i.e.,

$$a(\theta, \hat{\boldsymbol{x}}; \hat{z}) = \int d\Delta\hat{\nu} a_\nu(\hat{\boldsymbol{x}}; \hat{z}) e^{i\nu\theta} \ll 1. \tag{50}$$

Let us write the distribution function as

$$f = f_0 + f_1, \tag{51}$$

where f_0 is the distribution function averaged over θ, and f_1 contains the shot-noise fluctuation and the FEL-induced modulation. Equation (48) can be rewritten in two parts:

$$\frac{\partial f_0}{\partial \hat{z}} + \hat{\boldsymbol{p}} \frac{\partial f_0}{\partial \hat{\boldsymbol{x}}} - \hat{k}_\beta^2 \hat{\boldsymbol{x}} \frac{\partial f_0}{\partial \hat{\boldsymbol{p}}} + \left\langle \left[\int d\Delta\hat{\nu} e^{i\nu\theta} a_\nu(\hat{\boldsymbol{x}}; \hat{z}) + \text{c.c.} \right] \frac{\partial f_1(\theta, \hat{\eta}, \hat{\boldsymbol{x}}, \hat{\boldsymbol{p}}; \hat{z})}{\partial \hat{\eta}} \right\rangle_\theta = 0, \tag{52}$$

$$\frac{\partial f_1}{\partial \hat{z}} + \theta' \frac{\partial f_1}{\partial \theta} + \hat{\boldsymbol{p}} \frac{\partial f_1}{\partial \hat{\boldsymbol{x}}} - \hat{k}_\beta^2 \hat{\boldsymbol{x}} \frac{\partial f_1}{\partial \hat{\boldsymbol{p}}} + \left[\int d\Delta\hat{\nu} e^{i\nu\theta} a_\nu(\hat{\boldsymbol{x}}; \hat{z}) + \text{c.c.} \right] \frac{\partial f_0(\hat{\eta}, \hat{\boldsymbol{x}}, \hat{\boldsymbol{p}}; \hat{z})}{\partial \hat{\eta}} = 0. \tag{53}$$

We regard both f_1 and the field a defined in Eq. (50) as first-order quantities, hence the last term of Eq. (52) (averaged over θ) is a second-order quantity and will be ignored in the linear theory. We will discuss its effect at the onset of the saturation in the quasilinear theory of Sec. III I.

Using the method of integration along the unperturbed trajectory [23] and following the derivation of Ref. [24], we obtain

$$f_1 = f_1(0) + \int_0^{\hat{z}} ds \int d\Delta\hat{\nu} e^{i\nu\theta^{(0)}} a_\nu(\hat{\boldsymbol{x}}^{(0)}, s) \times \frac{\partial}{\partial \hat{\eta}} f_0(\hat{\eta}, \hat{\boldsymbol{x}}^{(0)}, \hat{\boldsymbol{p}}^{(0)}; s) + \text{c.c.}, \tag{54}$$

where $f_1(0)$ is the initial fluctuation from the smooth distribution at $z = 0$, and the unperturbed trajectory is given by

$$\begin{aligned} \theta^{(0)} &= \theta + \theta'(s - \hat{z}), \\ \hat{\boldsymbol{x}}^{(0)} &= \hat{\boldsymbol{x}} \cos[\hat{k}_\beta(s - \hat{z})] + \frac{\hat{\boldsymbol{p}}}{\hat{k}_\beta} \sin[\hat{k}_\beta(s - \hat{z})], \\ \hat{\boldsymbol{p}}^{(0)} &= -\hat{k}_\beta \hat{\boldsymbol{x}} \sin[k_\beta(s - \hat{z})] + \hat{\boldsymbol{p}} \cos[\hat{k}_\beta(s - \hat{z})]. \end{aligned} \tag{55}$$

One can easily confirm that Eq. (54) satisfies Eq. (53).

We now assume that the electron beam is transversely matched to the undulator channel and is uniform in the longitudinal direction (this can be approximately satisfied by a bunch that is very long compared to the fundamental radiation wavelength λ_1). Then $f_0 = f_0(\hat{\boldsymbol{p}}^2 + \hat{k}_\beta^2 \hat{\boldsymbol{x}}^2, \hat{\eta})$ is a solution of Eq. (52) in the absence of the second-order term. Inserting $f = f_0 + f_1$ with f_1 given by Eq. (54) into Eq. (47), we find that each frequency component of the radiation field is independently amplified and is governed by [24]

$$\left(\frac{\partial}{\partial \hat{z}} + i\Delta\hat{\nu} + \frac{\hat{\boldsymbol{\nabla}}_\perp^2}{2i} \right) a_\nu(\hat{\boldsymbol{x}}; \hat{z}) - \int d^2\hat{p} \int d\hat{\eta} \int_0^{\hat{z}} ds e^{i\theta'(s-\hat{z})} a_\nu(\hat{\boldsymbol{x}}^{(0)}; s) \frac{\partial f_0}{\partial \hat{\eta}} = \int d^2\hat{p} \int d\hat{\eta} f_\nu(\hat{\eta}, \hat{\boldsymbol{x}}, \hat{\boldsymbol{p}}; 0), \tag{56}$$

where

$$\begin{aligned} f_\nu(\hat{\eta}, \hat{\boldsymbol{x}}, \hat{\boldsymbol{p}}; 0) &= \int \frac{2\rho d\theta}{2\pi} e^{-i\nu\theta} f_1(\theta, \hat{\eta}, \hat{\boldsymbol{x}}, \hat{\boldsymbol{p}}; 0) \\ &= \frac{2k_u \rho^3}{\pi n_0} \sum_{j=1}^{N_e} e^{-i\nu\theta_j(0)} \delta[\hat{\eta} - \hat{\eta}_j(0)] \\ &\quad \times \delta[\hat{\boldsymbol{x}} - \hat{\boldsymbol{x}}_j(0)] \delta[\hat{\boldsymbol{p}} - \hat{\boldsymbol{p}}_j(0)] \end{aligned} \tag{57}$$

is the Fourier transformation of the initial fluctuation with $\theta_j(0) = -i\omega_1 t_j(0)$ at $z = 0$.

We will postpone the discussions of shot-noise start-up until Sec. III H and focus on the homogeneous part of Eq. (56). In this case, we seek a solution of the form $A_n(\hat{\boldsymbol{x}}) e^{-i\mu_n \hat{z}}$, where the complex growth rate μ_n and the transverse mode profile $A_n(\hat{\boldsymbol{x}})$ of the nth eigenmode ($n = 0, 1, 2, \ldots$) satisfy

$$\left(-i\mu_n + i\Delta\hat{\nu} + \frac{\hat{\boldsymbol{\nabla}}_\perp^2}{2i} \right) A_n(\hat{\mathbf{x}}) - \int d^2\hat{p} \int d\hat{\eta} \int_{-\infty}^0 d\tau A_n(\hat{\boldsymbol{x}}_+) e^{i(\theta' - \mu_n)\tau} \frac{\partial f_0}{\partial \hat{\eta}} = 0. \tag{58}$$

Here $\hat{\boldsymbol{x}}_+ = \hat{\boldsymbol{x}} \cos(\hat{k}_\beta \tau) + (\hat{\boldsymbol{p}}/\hat{k}_\beta) \sin(\hat{k}_\beta \tau)$. Equation (58) generalizes Moore's guided mode equation [25,26] to include effects of beam energy spread and emittance [27,28]. In general, there are many discrete solutions of Eq. (58), and the radiation field can be written as an expansion of eigenmodes:

$$a_\nu(\hat{\boldsymbol{x}}; \hat{z}) = \sum_{n=0}^{\infty} C_n A_n(\hat{\boldsymbol{x}}) e^{-i\mu_n \hat{z}}, \tag{59}$$

where C_n is the mode expansion coefficient that can be determined by solving the initial value problem. In the high-gain regime, a Gaussian-like fundamental mode (for $n = 0$) with the largest growth rate $\text{Im}\mu_0$ usually dominates over other higher-order modes, i.e.,

$$a_\nu(\hat{\boldsymbol{x}}; \hat{z}) \approx C_0 A_0(\hat{\boldsymbol{x}}) e^{-i\mu_0 \hat{z}} \quad \text{when } \hat{z} = 2\rho k_u z \gg 1. \tag{60}$$

Thus, the transverse profile of the radiation appears to be guided with an exponentially growing amplitude. We will discuss this remarkable feature of a high-gain FEL and its implication to the transverse coherence in Sec. IV A.

Before considering a fully three-dimensional (3D) solution of Eq. (58), it is useful to consider a simpler case where electrons are all parallel with a vanishing angular

spread. We can then set $k_\beta = 0$ and take $f_0 = U(\hat{\boldsymbol{x}})V(\hat{\eta})$, where $U(\hat{\boldsymbol{x}})$ describes the electron beam transverse profile with $U(0) = 1$ and $V(\hat{\eta})$ describes the energy distribution with the normalization $\int d\hat{\eta} V(\hat{\eta}) = 1$. In this case, Eq. (58) becomes

$$\left[\mu_n - \Delta\hat{\nu} + \frac{\hat{\boldsymbol{\nabla}}_\perp^2}{2} - U(\hat{\boldsymbol{x}}) \int d\hat{\eta} \frac{dV/d\hat{\eta}}{\hat{\eta} - \mu_n}\right] A_n(\hat{\boldsymbol{x}}) = 0. \tag{61}$$

This equation has been studied by various authors during the early days when optical guiding was first discovered [25,26,29,30]. In the 1D limit where the electron beam is uniform and infinite in transverse dimensions, we can drop the transverse dependence to obtain

$$\mu - \Delta\hat{\nu} - \int d\hat{\eta} \frac{dV/d\hat{\eta}}{\hat{\eta} - \mu} = 0. \tag{62}$$

Here all transverse modes become degenerate with the same growth rate μ. For a cold beam with a vanishing energy spread [i.e., $V(\hat{\eta}) = \delta(\hat{\eta})$], Eq. (62) reduces to the well-known cubic equation for the complex growth rate [31]

$$\mu^2(\mu - \Delta\hat{\nu}) = 1. \tag{63}$$

At the optimal detuning $\Delta\hat{\nu} = 0$, the growing solution of Eq. (63) is $\mu = (-1 + i\sqrt{3})/2$. Hence, the radiation power builds up as $P \propto \exp(z/L_{G0})$, with $L_{G0} = \lambda_u/(4\pi\sqrt{3}\rho)$ as in Eq. (6).

F. Dispersion relation with four scaled parameters

We will now return to the general 3D case and solve Eq. (58) for a finite-emittance beam with Gaussian transverse and energy distributions:

$$f_0(\hat{\boldsymbol{p}}^2 + \hat{k}_\beta^2\hat{\boldsymbol{x}}^2, \hat{\eta}) = \frac{1}{2\pi\hat{k}_\beta^2\hat{\sigma}_x^2} \exp\left(-\frac{\hat{\boldsymbol{p}}^2 + \hat{k}_\beta^2\hat{\boldsymbol{x}}^2}{2\hat{k}_\beta^2\hat{\sigma}_x^2}\right) \frac{1}{\sqrt{2\pi}\hat{\sigma}_\eta} \times \exp\left(-\frac{\hat{\eta}^2}{2\hat{\sigma}_\eta^2}\right), \tag{64}$$

where

$$\hat{\sigma}_x = \sigma_x\sqrt{2k_1k_u\rho}, \qquad \hat{\sigma}_\eta = \sigma_\eta/\rho, \tag{65}$$

and σ_η is the relative rms energy spread. The electron beam emittance is specified by

$$\varepsilon_x = \varepsilon_y = \varepsilon = \hat{\sigma}_x^2\hat{k}_\beta/k_1. \tag{66}$$

Inserting Eq. (64) into Eq. (58) and performing the $\hat{\eta}$ integral, we obtain a dispersion relation for the fundamental mode as [32,33]

$$\left(\mu_0 - \Delta\hat{\nu} + \frac{\hat{\boldsymbol{\nabla}}_\perp^2}{2}\right)A_0(\hat{\boldsymbol{x}}) - \frac{1}{2\pi\hat{k}_\beta^2\hat{\sigma}_x^2}\int_{-\infty}^{0} \tau d\tau e^{-\hat{\sigma}_\eta^2\tau^2/2 - i\mu_0\tau} \int d^2\hat{p} A_0\left(\hat{\boldsymbol{x}}\cos(\hat{k}_\beta\tau) + \frac{\hat{\boldsymbol{p}}}{\hat{k}_\beta}\sin(\hat{k}_\beta\tau)\right) \times \exp\left[-\frac{\hat{\boldsymbol{p}}^2 + \hat{k}_\beta^2\hat{\boldsymbol{x}}^2}{2}\left(i\tau + \frac{1}{\hat{k}_\beta^2\hat{\sigma}_x^2}\right)\right] = 0. \tag{67}$$

The complex growth rate μ_0 and the fundamental mode $A_0(\hat{\boldsymbol{x}})$ are completely determined by four scaled parameters [32], such as $\hat{\sigma}_x$, $\hat{k}_\beta$, $\hat{\sigma}_\eta$, and $\Delta\hat{\nu}$ used here. They can be cast in different ways [32,33], but they characterize four distinct physical effects to be illustrated below.

(i) $\hat{\sigma}_x$ characterizes the diffraction effect. Let us rewrite

$$\hat{\sigma}_x^2 = \sigma_x^2 2k_1k_u\rho = \frac{2\pi\sigma_x^2}{\lambda_1}\frac{4\pi\rho}{\lambda_u} = \frac{1}{\sqrt{3}}\frac{Z_R}{2L_{G0}}, \tag{68}$$

where $Z_R = 4\pi\sigma_x^2/\lambda_1$ is the Rayleigh length assuming that the rms size of the optical beam is the same as that of the electrons. The transverse area of the radiation will double in one Rayleigh length, reducing the interaction efficiency. A strong gain will mitigate the diffraction. The condition for that is

$$\hat{\sigma}_x \geq 1 \quad \text{or} \quad Z_R \geq 2L_{G0}. \tag{69}$$

(ii) $\Delta\hat{\nu}$ represents the effect of the frequency detuning (the normalized deviation of the radiation wavelength λ from the resonant wavelength λ_1). Since the electron with the energy $\gamma_0 mc^2$ slips exactly λ_1 behind the radiation per undulator period, it will experience a different radiation phase if $\lambda \neq \lambda_1$. In order to not degrade the gain significantly, the phase spread introduced by the frequency detuning over the $\sim 1/\rho$ undulator period should be smaller than unity, or

$$|\Delta\hat{\nu}| = \frac{|\Delta\lambda|}{\lambda_1}\frac{1}{2\rho} = \frac{|\Delta\omega|}{\omega_1}\frac{1}{2\rho} < 1. \tag{70}$$

The length λ_u/ρ is roughly the saturation length [see the discussion following Eq. (92) in Sec. III I]. Thus, we expect that the relative SASE bandwidth at saturation is about ρ, i.e., $\sigma_\nu \sim \rho$.

(iii) $\hat{\sigma}_\eta$ represents the effect of the beam energy spread. The resonant wavelength spread caused by the energy spread over the saturation length must also be less than unity, i.e.,

$$\hat{\sigma}_\eta = \frac{|\Delta\gamma|}{\gamma\rho} \sim \frac{|\Delta\lambda|}{\lambda_1}\frac{1}{\rho} < 1 \quad \text{or} \quad \sigma_\eta < \rho. \tag{71}$$

(iv) $\hat{\sigma}_x \hat{k}_\beta$ represents the effect of the angular spread for a finite-emittance beam. According to Eq. (2), the beam angular spread $\sigma_{x'} = k_\beta \sigma_x = \sigma_x/\bar{\beta}$ inevitably introduces a spread in the resonant wavelength $|\Delta\lambda|/\lambda_1 = \lambda_u \sigma_{x'}^2/(2\lambda_1)$. Using similar arguments, we have

$$(\hat{\sigma}_x \hat{k}_\beta)^2 = \frac{\lambda_u \sigma_{x'}^2}{2\lambda_1}\frac{1}{\rho} = \frac{|\Delta\lambda|}{\lambda_1}\frac{1}{\rho} < 1. \tag{72}$$

Writing $\sigma_{x'}^2 = \varepsilon/\bar{\beta}$, and $L_{G0} = \lambda_u/(4\pi\sqrt{3}\rho) \approx \lambda_u/(8\pi\rho)$, we have the emittance requirement

$$\varepsilon < \frac{\lambda_1}{4\pi}\frac{\bar{\beta}}{L_{G0}}. \tag{73}$$

G. Variational solution of the power gain length

One of the most important FEL design parameters is the power gain length of the fundamental mode given by

$$L_G = \frac{\lambda_u}{8(\mathrm{Im}\mu_0)\pi\rho} \equiv L_{G0}(1+\Lambda), \tag{74}$$

where Λ quantifies the degrading effects discussed in the previous section over the shortest possible gain length $L_{G0} = \lambda_u/(4\sqrt{3}\pi\rho)$ and will be determined explicitly here. Given the four scaled parameters, the complex growth rate of the fundamental mode μ_0 can be obtained by solving Eq. (67) with a variational approximation [30,32,33], with an orthogonal function expansion [34], or with a direct numerical method [33]. The variational method is very fast and accurate in terms of the growth rate and will be illustrated here.

For an azimuthally symmetric fundamental mode, we take $A_0(\hat{\boldsymbol{x}}) = A_0(R)$, where $R = |\hat{\boldsymbol{x}}|/\hat{\sigma}_x = |\boldsymbol{x}|/\sigma_x$. Equation (67) can be simplified considerably after some algebra:

$$\left[\mu_0 - \Delta\hat{\nu} + \frac{\hat{\sigma}_x^2}{2}\frac{d}{RdR}\left(R\frac{d}{dR}\right)\right]A_0(R) = \int_0^\infty R'dR' G_0(R,R')A_0(R'), \tag{75}$$

where

$$G_0(R,R') = \int_{-\infty}^0 \frac{\tau d\tau}{\sin^2(\hat{k}_\beta\tau)} \exp\left[-\frac{\hat{\sigma}_\eta^2\tau^2}{2} - i\mu_0\tau - \frac{(R^2+R'^2)(1+i\hat{k}_\beta^2\hat{\sigma}_x^2\tau)}{2\sin^2(\hat{k}_\beta\tau)}\right] I_0\left(\frac{RR'(1+i\hat{k}_\beta^2\hat{\sigma}_x^2\tau)\cos(\hat{k}_\beta\tau)}{\sin^2(\hat{k}_\beta\tau)}\right), \tag{76}$$

and I_0 is the zeroth-order modified Bessel function. A variational functional may be constructed as follows [30,32,33]:

$$\int_0^\infty RdRA_0(R)\left[\mu_0 - \Delta\hat{\nu} + \frac{\hat{\sigma}_x^2}{2R}\frac{d}{dR}\left(R\frac{d}{dR}\right)\right]A_0(R) = \int_0^\infty RdRA_0(R)\int_0^\infty R'dR'G_0(R,R')A_0(R'). \tag{77}$$

The variational principle states that a first-order approximation in $A_0(R)$ yields a stationary solution μ_0 that is accurate to the second order. For a Gaussian-like fundamental mode, we take a trial function $A_0(R) \propto \exp(-wR^2)$ and insert it into Eq. (77) to obtain

$$\frac{\mu_0 - \Delta\hat{\nu}}{4w} - \frac{1}{4\hat{\sigma}_x^2} = \int_{-\infty}^0 \frac{\tau d\tau e^{-\hat{\sigma}_\eta^2\tau^2/2 - i\mu_0\tau}}{(1+i\hat{k}_\beta^2\hat{\sigma}_x^2\tau^2)^2 + 4w(1+i\hat{k}_\beta^2\hat{\sigma}_x^2\tau^2) + 4w^2\sin^2(\hat{k}_\beta^2\tau)}. \tag{78}$$

Differentiating Eq. (78) with respect to w and applying the variational condition $\partial\mu_0/\partial w = 0$, we obtain the second relation between μ_0 and w. Solving these two equations, we can determine μ_0 and hence the power gain length.

Based on the variational solution of the FEL dispersion relation, Ming Xie obtained a very useful fitting formula for the power gain length of the fundamental mode that depends on three scaled parameters [33,35]:

$$\begin{aligned}
\eta_d &= \frac{1}{2\sqrt{3}\hat{\sigma}_x^2} = \frac{L_{G0}}{2k_1\sigma_x^2} \text{ (diffraction parameter)},\\
\eta_\varepsilon &= \frac{2}{\sqrt{3}}\hat{k}_\beta^2\hat{\sigma}_x^2 = k_\beta L_{G0}\frac{\varepsilon}{\lambda_1/(4\pi)} \text{ (angular spread parameter)},\\
\eta_\gamma &= \frac{\hat{\sigma}_\eta}{\sqrt{3}} = 4\pi\frac{L_{G0}}{\lambda_u}\sigma_\eta \text{ (energy spread parameter)},
\end{aligned} \tag{79}$$

while the fourth parameter, the frequency detuning, is optimized to yield the minimum power gain length. The gain length degradation factor Λ defined in Eq. (74) is written as

$$\Lambda = a_1\eta_d^{a_2} + a_3\eta_\varepsilon^{a_4} + a_5\eta_\gamma^{a_6} + a_7\eta_\varepsilon^{a_8}\eta_\gamma^{a_9} + a_{10}\eta_d^{a_{11}}\eta_\gamma^{a_{12}} + a_{13}\eta_d^{a_{14}}\eta_\varepsilon^{a_{15}} + a_{16}\eta_d^{a_{17}}\eta_\varepsilon^{a_{18}}\eta_\gamma^{a_{19}}, \tag{80}$$

where the fitting coefficients are

$$\begin{aligned}
&a_1 = 0.45, \quad a_2 = 0.57, \quad a_3 = 0.55, \quad a_4 = 1.6,\\
&a_5 = 3, \quad a_6 = 2, \quad a_7 = 0.35, \quad a_8 = 2.9,\\
&a_9 = 2.4, \quad a_{10} = 51, \quad a_{11} = 0.95, \quad a_{12} = 3,\\
&a_{13} = 5.4, \quad a_{14} = 0.7, \quad a_{15} = 1.9, \quad a_{16} = 1140,\\
&a_{17} = 2.2, \quad a_{18} = 2.9, \quad a_{19} = 3.2.
\end{aligned} \tag{81}$$

The discrepancy between Xie's fitting formula and numeri-

cal solutions of the FEL eigenmode equation is typically less than 10%. These positive fitting coefficients quantitatively show that all three scaled beam parameters in Eq. (79) should be kept small to avoid a large gain reduction, corresponding to the qualitative beam requirements discussed in Sec. III F.

H. Start-up process

The 3D initial value problem for an FEL starting up from electron shot noise [i.e., Eq. (56)] can be solved using Van Kampen's normal mode expansion [27,36]. Equivalent methods using the bi-orthogonality theorem of a generalized Hamiltonian for the beam-radiation system are given in Refs. [37,38]. For Van Kampen's normal mode expansion, we refer to the detailed derivation in Appendix A of Ref. [39]. In the high-gain limit, we may keep only the fundamental mode and arrive at

$$a_\nu(\hat{\boldsymbol{x}};\hat{z}) = e^{-i\mu_0\hat{z}}A_0(\hat{\boldsymbol{x}})\Bigg[\int d^2\hat{x}' A_0(\hat{\boldsymbol{x}}')a_\nu(\hat{\boldsymbol{x}}';0) + \int d^2\hat{x}' \int d^2\hat{p} \int d\hat{\eta} f_\nu(\hat{\eta},\hat{\boldsymbol{x}}',\hat{\boldsymbol{p}};0) \times \int_{-\infty}^{0} d\tau A_0\left(\hat{\boldsymbol{x}}'\cos\hat{k}_\beta\tau - \frac{\hat{\boldsymbol{p}}}{\hat{k}_\beta}\sin\hat{k}_\beta\tau\right)e^{i(\theta'-\mu_0)\tau}\Bigg]. \tag{82}$$

The first term in the square bracket describes the process of coherent amplification (CA), which starts from a coherent input signal $a_\nu(\hat{\boldsymbol{x}};0)$ at the frequency $\omega = \nu\omega_1$. The second term describes the SASE process that starts from electron shot noise [27,37,40,41]. Although the ensemble average of $f_\nu(\hat{\eta},\hat{\boldsymbol{x}}',\hat{\boldsymbol{p}};0)$ in Eq. (57) is zero, the average radiation intensity is not and can be computed by using the relation [23]

$$\langle f_\nu(\hat{\eta},\hat{\boldsymbol{x}},\hat{\boldsymbol{p}};0)f_\nu(\hat{\eta}',\hat{\boldsymbol{x}}',\hat{\boldsymbol{p}}';0)\rangle = \frac{2k_1^3k_u\rho^3cT}{\pi^2n_0}\delta(\hat{\eta}-\hat{\eta}')\delta(\hat{\boldsymbol{x}}-\hat{\boldsymbol{x}}')\delta(\hat{\boldsymbol{p}}-\hat{\boldsymbol{p}}')f_0, \tag{83}$$

where cT is the bunch length for a flattop current profile.

Including the frequency dependence of the complex growth rate in Eq. (60) and integrating over the transverse coordinates, we obtain the FEL power spectrum in the high-gain regime as

$$\frac{dP}{d\omega} = \frac{\gamma_0mc^3n_0\pi}{2\rho\omega_1^2T}\int dx^2\langle|a_\nu|\rangle^2 = g_A\left(\frac{dP_0}{d\omega} + g_S\frac{\rho\gamma_0mc^2}{2\pi}\right)\exp\left(\frac{z}{L_G} - \frac{\Delta\omega^2}{2\sigma_\omega^2}\right), \tag{84}$$

where $dP_0/d\omega$ is the input power spectrum; $\rho\gamma_0mc^2/(2\pi)$ is the 1D SASE noise power spectrum [27] and can be identified as the spontaneous undulator radiation in the first two power gain lengths [42]; g_A and g_S determine the input coupling to the fundamental mode and the effective start-up noise in units of $\rho\gamma_0mc^2/(2\pi)$, respectively; and σ_ω is the SASE bandwidth. In the 1D, cold beam limit, $g_A = 1/9$, $g_S = 1$, and the rms SASE bandwidth is [40,41]

$$\sigma_\omega = \sqrt{\frac{3\sqrt{3}\rho}{k_uz}}\omega_1. \tag{85}$$

For a more general beam distribution, the SASE bandwidth can be found by solving the dispersion relation [i.e., Eq. (58)] and typically decreases to about $\rho\omega_1$ at the FEL saturation point. g_A and g_S can also be computed for a general beam distribution by solving the 3D initial value problem [38,43]. It is noted that the effective start-up noise increases with larger energy spread and emittance mainly because of the corresponding increase in the gain length. Integrating the SASE term over the frequency, we have the average SASE power as

$$P = g_AP_n\exp\left(\frac{z}{L_G}\right). \tag{86}$$

Here $P_n = g_S\rho\gamma_0mc^2\sigma_\omega/\sqrt{2\pi}$ is the effective noise power for SASE.

As a numerical example, Fig. 6 shows the total radiated energy in the LEUTL FEL [6] at $\lambda_1 = 130$ nm. The agreement of the high-gain behavior between time-dependent SASE simulations using either GINGER [44] or GENESIS [45] and Eq. (86) are very good when the proper input coupling coefficient and effective noise power (i.e., g_A and g_S) are calculated. Note that GENESIS is a 3D code that does not assume azimuthal symmetry in the radiation profile. Hence, it takes into account more higher-order transverse modes than GINGER (with only azimuthally symmetric modes) in the start-up regime and leads to more radiated energy in the early part of the undulator length.

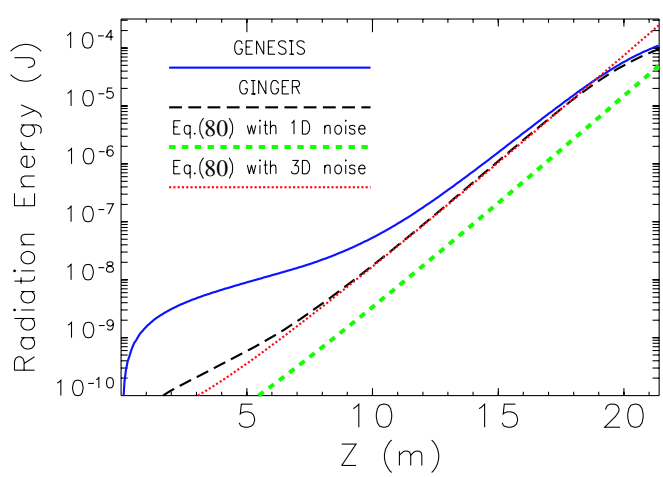


FIG. 6. (Color) GINGER (black) and GENESIS (blue) simulations of the LEUTL FEL energy at 130 nm versus the undulator distance z, as compared from predictions of Eq. (86) with 3D noise (red) and 1D noise (green).

I. Saturation mechanism

The exponential growth cannot continue indefinitely, and the power must saturate at a certain level. This is because the average beam energy must decrease to conserve the total energy. (In the linear theory, the average beam energy is unchanged to the first order in a_ν because the field energy is second order in a_ν.) In addition, the beam energy spread inevitably increases so that the growth rate becomes negligible. The saturation effect is due to nonlinear interactions and can be studied by a quasilinear theory that takes into account the second-order term in Eq. (52) [27,46]. Here we illustrate the basic steps in the 1D case, where we write the distribution function as

$$f = \underbrace{f_0(\hat{\eta};\hat{z})}_{\theta\text{ indep. average distribution}} + \underbrace{\int f_\nu(\hat{\eta};\hat{z})e^{i\nu\theta}d\nu + \text{c.c.}}_{\theta\text{ dep. microbunching}} \tag{87}$$

Unlike in the linear theory discussed in Sec. III E, the smoothed distribution function $f_0(\hat{\eta};\hat{z})$ is also a function of $\hat{z}$. Consider $\nu \sim 1$ for the fundamental frequency (dropping higher harmonics is equivalent to dropping higher than second-order terms; see Sec. IV C), the set of nonlinear equations following Eqs. (47), (52), and (53) are

$$\left(\frac{d}{d\hat{z}} + i\Delta\hat{\nu}\right)a_\nu + \int d\hat{\eta} f_\nu = 0, \tag{88}$$

$$\frac{\partial f_\nu}{\partial \hat{z}} + i\hat{\eta} f_\nu + \frac{\partial f_0}{\partial \hat{\eta}} a_\nu = 0, \tag{89}$$

$$\frac{\partial f_0}{\partial \hat{z}} + \left(\int d\nu a_\nu^* \frac{\partial f_\nu}{\partial \hat{\eta}} + \text{c.c.}\right) = 0. \tag{90}$$

In the small signal regime where $|a_\nu|, |f_\nu| \ll 1$, $\partial f_0/(\partial \hat{z}) = 0$ and we recover the 1D linearized Maxwell-Klimotovich equations. When $|a_\nu|, |f_\nu| \sim 1$, the initial energy distribution changes rapidly with $\hat{z}$. The average energy decreases with an increasing energy spread and the FEL saturation sets in.

This set of nonlinear equations can be solved numerically given the initial conditions. Since the saturation power level is quite insensitive to the start-up process, we start with a small seed signal $a_\nu(0)$ at various initial frequency detunings and find the dependence of the maximum power (at the optimal detuning) as a function of the initial rms energy spread. The results [46] are shown in Fig. 7 and are compared with a simulation fitting formula given by [35]

$$P_{\text{sat}} \approx 1.6\left(\frac{L_{G0}}{L_G}\right)^2 \rho P_{\text{beam}} = \frac{1.6}{(1+\Lambda)^2}\rho P_{\text{beam}}, \tag{91}$$

where Λ is defined in Eq. (74), and $P_{\text{beam}}[\text{GW}] = (\gamma_0 mc^2/e)[\text{GV}]I_e[\text{A}]$ is the total electron beam power.

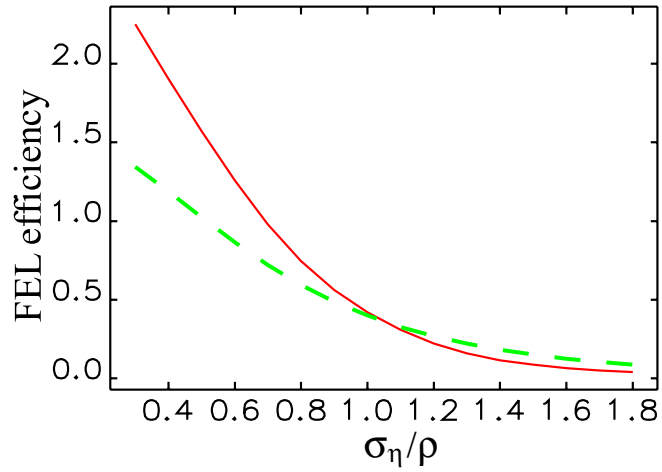


FIG. 7. (Color) Maximum FEL efficiency $P_{\text{sat}}/(\rho P_{\text{beam}})$ versus the initial normalized rms energy spread $\sigma_{\hat{\eta}} = \sigma_\eta/\rho$ from the quasilinear theory [46] (red solid curve) and from a simulation fitting formula [35] (green dashed curve).

To estimate the saturation distance of a SASE FEL, we require that Eq. (86) is equal to Eq. (91). In the 1D case, we obtain

$$\frac{z_{\text{sat}}}{L_G} \approx \ln\frac{20 I_e t_c}{e}, \tag{92}$$

where $t_c = \sqrt{\pi}/\sigma_\omega$ is the coherence time [see Eq. (102) below]. Thus, the saturation distance is a numerical factor times the power gain length. The numerical factor depends logarithmically on the number of electrons within one coherence time (i.e., $N_c = I_e t_c/e$) and typically varies little from 18 to 20. We note that, numerically, $4\pi\sqrt{3} \approx 20$. Therefore the saturation length is simply λ_u/ρ, if $L_G \approx L_{G0}$.

J. Quantum effects

Despite the fact that the first FEL theory is based on a quantum mechanical analysis [1], subsequent analysis shows that the classical theory is adequate in most practical devices. In this section, we discuss both the quantum recoil and the quantum diffusion effects that tend to decrease the FEL interaction efficiency if they are not negligible.

1. Quantum recoil

As an electron emits a photon of energy $\hbar\omega_1$, its energy is reduced due to the quantum recoil. If the fractional energy change is on the order of or larger than the FEL gain bandwidth, the quantum recoil may significantly degrade the FEL gain. For short-wavelength, high-gain FELs using magnetic undulators and high-energy electron beams, the typical FEL gain bandwidth is on the order of 10^{-3}, while the fraction energy change after a photon emission is no more than 10^{-6}; hence, the quantum recoil is negligible. The quantum recoil effect may become an

issue when an extremely bright and low-energy electron beam interacts with an electromagnetic undulator, as the fractional energy change due to an x-ray photon emission may be comparable to or exceed the FEL bandwidth. High-gain FELs in the quantum regime were studied before (see, e.g., Ref. [47–49]) and have been revisited recently in the context of SASE [50,51]. For a 1D, cold electron beam, the classical cubic equation (63) is modified to [48,50,51]

$$\left(\mu^2 - \frac{1}{4\bar{\rho}^2}\right)(\mu - \Delta\hat{\nu}) = 1, \tag{93}$$

where $\bar{\rho} = \rho\gamma mc^2/(\hbar\omega_1)$ can be regarded as the ratio of the classical FEL bandwidth to the fractional energy recoil of an FEL photon. In view of Eq. (62), we see that the effect due to the quantum recoil is equivalent to a flattop energy spread with a width $\hbar\omega_1$ that decreases the FEL growth rate μ.

2. Quantum diffusion

Even when a single photon emission is not capable of changing the electron energy outside the FEL bandwidth, the accumulated effects of spontaneous undulator emission may alter the electron energy significantly over the long undulator line. The classical part of this effect is that the electron changes its energy due to the spontaneous undulator radiation, given by

$$\frac{(\Delta\gamma)_{\rm SR}}{\gamma_0} = -\frac{1}{3} r_e \gamma_0 K_0^2 k_u^2 L_u, \tag{94}$$

where L_u is the total length of the undulator. For the Linac Coherent Light Source (LCLS) at 14 GeV (see Table I), the fractional energy loss over the 100-m undulator is about 0.17%, which causes the electrons' central energy γ_0 to move away from the resonant bandwidth. In view of Eq. (4), the resonant wavelength can be kept constant by tapering the undulator magnetic field strength (i.e., adjusting the K_0 parameter) to compensate for the energy loss.

TABLE I. Main parameters for the LCLS FEL.

Parameter	Symbol	Value
Electron energy	$\gamma_0 mc^2$	13.6 GeV
Bunch duration (FWHM)	T	200 fs
Bunch current (flat part)	I_e	3.4 kA
Transverse norm. emittance	ε_n	1.2 μm
Average beta function	$\bar{\beta}$	25 m
Undulator period	λ_u	0.03 m
Undulator field	B_0	1.25 T
Undulator strength parameter	K_0	3.5
Active undulator length	L_u	110 m
Fundamental wavelength	λ_1	1.5 Å
Pierce parameter	ρ	4.5×10^{-4}
FEL power gain length	L_G	4.3 m

In addition to the average energy loss, the discrete nature of spontaneous photon emissions (over a wide energy spectrum) increases the uncorrelated energy spread of the beam, much like the effect of quantum excitation in an electron storage ring (see, e.g., Refs. [15,52]). The diffusion rate of the energy spread is calculated to be [53]

$$\frac{d\langle(\Delta\gamma)^2\rangle}{dz} = \frac{7}{15} r_e \lambdabar_c \gamma_0^4 K_0^2 k_u^3 F(K_0), \qquad F(K_0) = 1.2K_0 + \frac{1}{1 + 1.33K_0 + 0.40K_0^2}, \tag{95}$$

where $\lambdabar_c = \hbar/mc \approx 3.86 \times 10^{-13}$ is the Compton wavelength. For the LCLS case (see Table I), this quantum diffusion process increases the uncorrelated energy spread in the 100-m undulator to more than 1×10^{-4} even if the initial energy spread is zero. Although this level of energy spread is still acceptable for the LCLS and other similar x-ray FEL projects, the quantum diffusion effect may impose a practical limit on the minimum achievable wavelength for a given transverse emittance and peak current [54].

IV. CHARACTERISTICS OF X-RAY FELS

In the last section, we reviewed the basic theory of a high-gain FEL. In the following three sections, we will apply the theory to discuss in turn, the properties of x-ray FEL output, various effects that can degrade FEL performance, and novel schemes designed to enhance the usefulness of the x-ray FEL. In order to make the length of the paper manageable, we will present most results without much derivation and refer the reader to an extensive list of literature.

In this section, we describe the radiation characteristics of an x-ray FEL including its transverse and temporal coherence, harmonic content, and saturation behaviors.

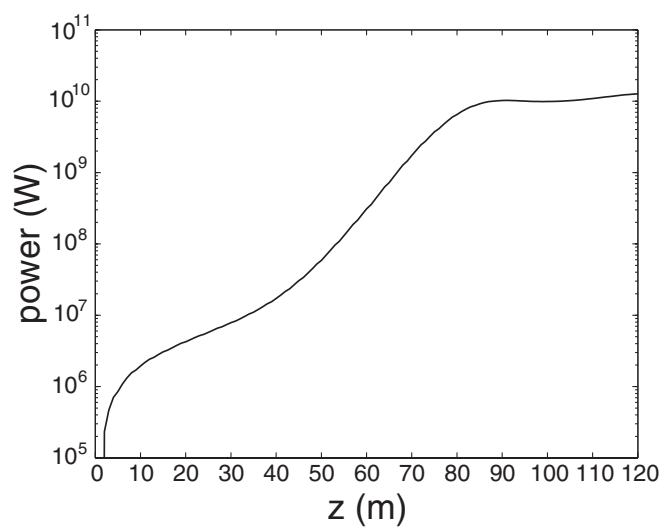


FIG. 8. GENESIS SASE simulation of the LCLS power along the undulator distance.

We will mainly concentrate on the SASE FEL but will also comment on some advanced seeding schemes in Sec. IV B. As a numerical example, we use the LCLS at $\lambda_1 = 1.5$ Å [11] as given in Table I. Figure 8 shows the GENESIS simulation of the radiation growth along the undulator distance obtained with these parameters.

A. Transverse coherence and mode properties

As briefly discussed in Sec. III E, a unique feature of the exponential growth regime is optical guiding [25,29], the phenomena in which the transverse profile of the radiation does not change significantly. This arises because the field amplitude is dominated by the fundamental mode with the largest growth rate as given by Eq. (60). Because of stronger diffraction and less spacial overlap with the electron beam, higher-order modes usually have smaller growth rates and hence are negligible after a few e-folding lengths of the fundamental modes. For an explicit numerical demonstration, we refer the reader to the results of the LCLS higher-order mode calculation reported in Ref. [55].

Because of optical guiding, the SASE FEL can reach almost full transverse coherence before saturation, even when the emittance of the electron beam ε is larger than the diffraction-limited radiation emittance $\varepsilon_{r0} = \lambda_1/(4\pi)$ as in the above LCLS example. Figure 9 shows the GENESIS simulation of the LCLS radiation angular patterns at different z locations. At the initial start-up stage, the large beam emittance excites many transverse modes, and the radiation is dominated by incoherent spontaneous emission with its emittance $\varepsilon_r \approx \varepsilon$ [Fig. 9(a)]. The mode pattern cleans up in the exponential growth regime due to optical guiding [Fig. 9(b)]. Near saturation, the guided fundamental mode dominates the radiation pattern [Fig. 9(c)], and the radiation emittance is almost given by ε_{r0}. In fact, the frequency-dependence of the fundamental mode profile within the finite SASE bandwidth introduces a slight increase of the radiation emittance above the minimum emittance ε_{r0} [56,57]. If we define the transverse mode parameter as

$$M_T^2 = \left(\frac{\varepsilon_r}{\varepsilon_{r0}}\right)^2, \tag{96}$$

then $M_T^2 = (\varepsilon/\varepsilon_{r0})^2 \gg 1$ at the start-up stage and decreases to a value slightly above unity at saturation. Detailed numerical simulations characterizing the dependence of transverse coherence on the ratio of the electron emittance to the radiation emittance is presented in Ref. [58].

In general, the evaluation of the transverse mode size (and angular divergence) requires numerical solutions of the FEL eigenmode equation. Two limiting cases can be discussed here for a "parallel" beam with a vanishing angular spread. For a large electron transverse size or a small diffraction parameter [i.e., $\eta_d \ll 1$ in Eq. (79)], it is plausible (but wrong) to estimate the rms mode size σ_r by $\sigma_D = \sqrt{\lambda_1 2L_{G0}/4\pi}$, which is the size of a coherent optical beam with its Rayleigh length identified as the 1D field amplitude gain length $2L_{G0}$. A correct calculation of the mode size proceeds from Eq. (61) by setting $U(\hat{\boldsymbol{x}}) = \exp(-\hat{\boldsymbol{x}}^2/2\hat{\sigma}_x^2) \approx 1 - \hat{\boldsymbol{x}}^2/2\hat{\sigma}_x^2$ and solving the equation exactly for a Gaussian fundamental mode. In this 1D limit, we find that the rms mode size is the geometric average of σ_D and σ_x [33], i.e.,

$$\sigma_r \approx \sqrt{\sigma_D \sigma_x} \approx \sqrt{\eta_d}\sigma_x. \tag{97}$$

On the other hand, when the radiation mode size is much larger than the electron transverse size due to strong diffraction effect (i.e., $\eta_d \gg 1$), the beam-radiation interaction is no longer characterized by ρ. In this 3D limit, we should redefine ρ by replacing σ_x^2 in Eq. (5) with $\sigma_r^2 = (\lambda_1/4\pi)2L_G^D$, where L_G^D is the power gain length in this diffraction-dominated regime. If we assume L_G^D takes the form of Eq. (6) with the redefined ρ, both L_G^D and σ_r can be determined self-consistently as [32]

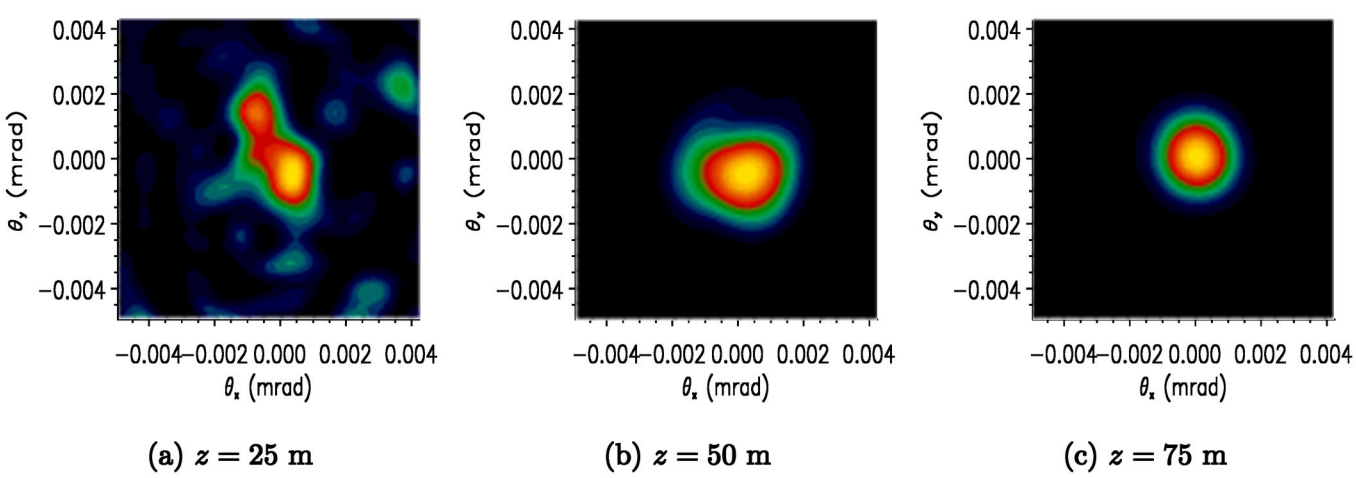


FIG. 9. (Color) Evolution of the LCLS transverse profiles at different z locations (courtesy of Sven Reiche, UCLA).

$$L_G^D \approx \frac{\lambda_u}{4\pi}\sqrt{\frac{\gamma_0 I_A}{I_e}\frac{1+K_0^2/2}{K_0^2[\mathrm{JJ}]^2}}, \qquad \text{and} \qquad \sigma_r \approx \sqrt{\frac{\lambda_1}{4\pi}2L_G^D}. \tag{98}$$

In this case, both the gain length and the mode size are independent of the transverse beam size.

B. Temporal characteristics and manipulations

Because of its noisy start-up, the temporal property of a SASE FEL is that of a chaotic light [59–62]. This may be illustrated with the analytical result in the 1D case. Dropping the transverse dependencies and mode index 0 in Eq. (82) and assuming a cold beam without any initial energy spread, the second term of Eq. (82) is simplified to

$$a_\nu(\hat{z}) = \frac{ie^{-i\mu\hat{z}}}{3\mu}\sum_{j=1}^{N_e} e^{-i\nu\theta_j(0)}, \tag{99}$$

where μ satisfies the 1D FEL cubic equation (63) and may be expanded to the second order in $\Delta\hat{\nu}$ as

$$\mu(\Delta\hat{\nu}) \approx -\frac{1}{2}\left[1-\frac{2}{3}\Delta\hat{\nu}+\frac{\Delta\hat{\nu}^2}{6}\right]+\frac{\sqrt{3}}{2}i\left(1-\frac{\Delta\hat{\nu}^2}{6}\right). \tag{100}$$

Inserting it into Eq. (99) and Fourier transforming a_ν using the relations in Eqs. (36) and (46), we have

$$\begin{aligned} E_x(t;z) &\propto e^{i(k_1 z-\omega_1 t)}\int d\Delta\hat{\nu} e^{-i\mu(\Delta\hat{\nu})\hat{z}}\sum_{j=1}^{N_e}\exp\{-i\theta_j(0)+i2\rho\Delta\hat{\nu}[\theta-\theta_j(0)]\} \\ &\propto \frac{e^{(\sqrt{3}+i)\rho k_u z}}{\sqrt{z}}\sum_{j=1}^{N_e}\exp\left[ik_1 z - i\omega_1[t-t_j(0)]-\frac{3}{4}\left(1+\frac{i}{\sqrt{3}}\right)\sigma_\omega^2\right]\left(t-t_j(0)-\frac{z}{v_g}\right)^2, \end{aligned} \tag{101}$$

where $t_j(0)$ is the random arrival time of the jth electron at the undulator entrance. Note that the group velocity of the wave packet is $v_g = \omega_1/(k_1 + 2k_u/3)$ [59], slower than the speed of light but faster than the electrons that amplify the radiation.

Such a chaotic light can be analyzed by statistical methods (see, e.g., [63]). From the first-order time correlation function, we obtain the coherence time as [60–62]

$$t_c = \frac{\sqrt{\pi}}{\sigma_\omega}. \tag{102}$$

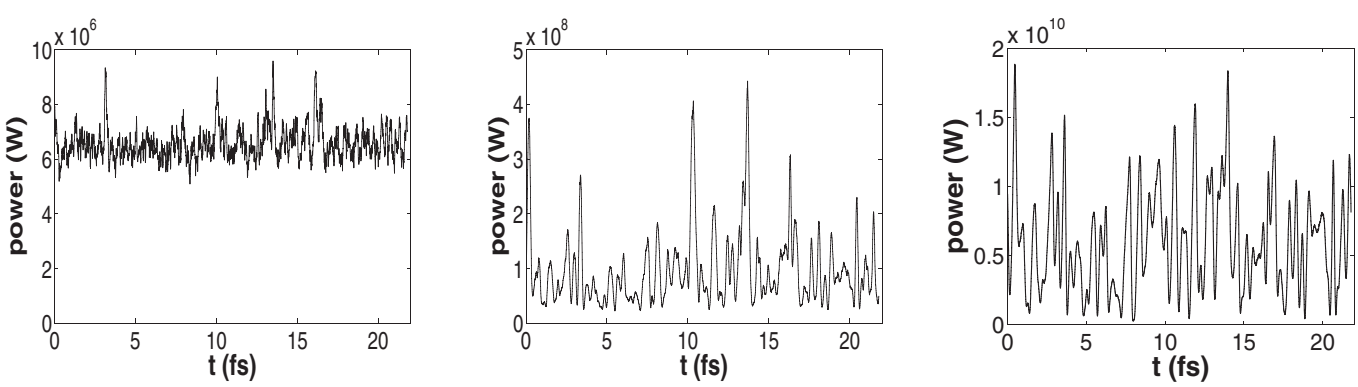
The energy of a SASE pulse W with a flattop duration T fluctuates according to the Gamma probability distribution [60]:

$$p(W) = \frac{M^M}{\Gamma(M)}\frac{W^{M-1}}{\langle W\rangle^M}\exp\left(-M\frac{W}{\langle W\rangle}\right), \tag{103}$$

where $\langle W\rangle$ is the average radiation energy and $\Gamma(M)$ is the Gamma function. The relative rms energy fluctuation σ_W is given by [60–62]

$$M = \frac{1}{\sigma_W^2} = \frac{\langle W\rangle^2}{\langle W^2\rangle - \langle W\rangle^2} = \begin{cases} T/t_c & \text{when } T \gg t_c, \\ 1 & \text{when } T \le t_c. \end{cases} \tag{104}$$

Thus, the M parameter characterizes the degree of freedom or the temporal "mode" of the pulse. For hard x-ray wavelengths, the coherence time t_c determined by Eq. (102) is only a few hundred attoseconds. Since the SASE pulse duration T is on the same order as the typical electron pulse of a few hundred femtoseconds, $M \gg 1$, and the Gamma



(a) z = 25 m (b) z = 50 m (c) z = 75 m

FIG. 10. Temporal structures of 10% of the LCLS pulse at different z locations.

distribution of shot-to-shot pulse energies approaches a Gaussian distribution with a small relative rms fluctuation given by $1/\sqrt{M}$. Figure 10 illustrates the temporal power profile for about 10% of the LCLS x-ray pulse. Note that these intensity spikes are roughly separated by t_c. A statistical analysis shows that the average number of intensity spikes in the time domain is about $0.7M$ [64]. In the frequency domain, the SASE spectral profile is also similarly spiky. The full SASE spectral width is about $2\sqrt{\pi}\sigma_\omega$, consisting of M independent spectral modes. Each mode is characterized by the spectral coherence range $2\pi/T$.

The statistical fluctuation can be generalized to the 3D case by redefining $M = M_L M_T^2$, where $M_L \approx T/t_c$ (for $T \gg t_c$) or 1 (for $T < t_c$) is the longitudinal mode number and M_T^2 is the transverse mode number as defined in Eq. (96). Thus, the instantaneous power (with $M_L = 1$) at the start-up stage (with $M_T^2 \gg 1$) does not fluctuate as much as the exponential growth stage (with $M_T^2 \to 1$), as shown from Figs. 10(a) to 10(b) and 10(b) to 10(c).

1. Shorter x-ray pulses

The temporal property of the SASE pulse can be tailored to a given application by suitable manipulations. One interesting direction is to generate much shorter x-ray pulses than the typical electron pulse of 100-fs in duration. Several methods have been proposed (see Ref. [65] for a recent review), including that based on x-ray pulse compression [66] and slicing [67] of a frequency-chirped SASE. The frequency-chirped SASE can be generated by an energy-chirped electron beam through the FEL resonant condition of Eq. (4). In this case, it can be shown [68] that the coherence time is independent of the frequency chirp $u = \Delta\omega/\Delta t$ as long as the frequency span within a temporal spike is smaller than its bandwidth (i.e., when $|u| \ll \sigma_\omega^2$), while the spectral coherence range increases according to $\sqrt{\pi}|u|/\sigma_\omega$. A narrow-bandwidth monochromator may be used to slice a much shorter section of the chirped x-ray pulse. If σ_m is the rms bandwidth of the monochromator, the sliced rms x-ray pulse duration is [68]

$$\sigma_t = \sqrt{\frac{\sigma_\omega^2 + \sigma_m^2}{u^2} + \frac{1}{4\sigma_m^2}}. \tag{105}$$

The minimum pulse duration for an optimized σ_m is $(\sigma_t)_{\min} \approx \sigma_\omega/|u|$. For $|u| \sim \sigma_\omega^2$, $\sigma_t \sim t_c$, and a single temporal spike of a few hundred attoseconds may be selected.

The energy chirp produced by the rf accelerator over the entire electron bunch is typically much smaller than that required to select a single SASE spike, and the sliced x-ray pulse duration is typically on the order of 10 fs [68]. A sufficiently large energy modulation over a small fraction of the bunch may be produced when a high-power, fs optical laser resonantly interacts with the bunch in a short undulator [69]. The local energy chirp can be sufficiently large (with $|u| \sim \sigma_\omega^2$) but can act like an effective energy spread to degrade the FEL gain. Nevertheless, it is pointed out in Ref. [70] that the FEL gain degradation due to a linear energy chirp can be perfectly compensated for by a proper taper of the undulator parameter. Thus, a tapered undulator can automatically "select" a small fraction of an energy-modulated bunch that has the right chirp with a pulse duration of about 200 attoseconds [70].

2. Coherence enhancement through seeding

Although the poor temporal coherence of a SASE pulse can be improved by a narrow-bandwidth monochromator, the radiation energy will at least be reduced by the ratio of the SASE bandwidth to the monochromator bandwidth. In addition, the statistical fluctuation of the filtered radiation will increase up to 100% as the number of spectral modes is reduced down to unity. In order to provide fully coherent x-ray FEL pulses, the intrinsic noise of the SASE radiation must be overcome with some form of seeding.

(i) HGHG: Since a proper coherent seed does not exist at x-ray wavelengths, a high-gain harmonic generation (HGHG) FEL relies on a coherent seed at subharmonic wavelengths. In this scheme [71], a small energy modulation is imposed on the electron beam by interaction with a seed laser in a short undulator (the modulator). The energy modulation is converted to a coherent spatial density modulation as the electron beam traverses a dispersive section. A second undulator (the radiator), tuned to a higher harmonic of the seed frequency, causes the microbunched electron beam to emit coherent radiation at that harmonic frequency. This shorter-wavelength radiation may then be used as the coherent seed to the next stage HGHG. In this cascaded harmonic conversion process, the ratio of electron shot noise to the laser signal is amplified by at least the square of the harmonic order and may limit its final wavelength reach to the soft x-ray region [72]. Single-stage HGHGs at infrared and ultraviolet wavelengths have been demonstrated at Brookhaven National Laboratory [73,74]. Cascaded HGHG FELs are currently under design studies as soft x-ray sources [75,76].

(ii) Self-seeding FEL: A self-seeding scheme [77,78] to improve the temporal coherence consists of two undulators (of the same undulator period and strength) and an x-ray monochromator located between them. The first undulator operates in the exponential gain regime of a SASE FEL. After the exit of the first undulator, the electron beam is guided through a dispersive bypass that smears out the microbunching induced in the first undulator. The SASE output enters the monochromator, which selects a narrow band of radiation. At the entrance of the second undulator the monochromatic x-ray beam is combined with the electron beam and is amplified up to the saturation level. Since the SASE power over a narrow bandwidth at the exit of the first undulator fluctuates 100% for a single mode, the length of the second undulator is chosen to exceed the

saturation length sufficiently to suppress fluctuation of the final output power level. Thus, this approach requires an undulator system almost twice as long as a single-stage SASE FEL.

(iii) Regenerative amplifier FEL: Another self-seeding scheme, a regenerative amplifier FEL (RAFEL), has been demonstrated in the infrared wavelength region [79] and proposed for VUV FELs [80,81]. Recently, a hard x-ray RAFEL was also proposed and studied [82]. In this scheme, SASE radiation from the leading electron bunch in a bunch train is spectrally filtered by the Bragg crystal reflectors and is brought back to the beginning of the undulator to interact with the second bunch. This process continues bunch-to-bunch, yielding an exponentially growing laser field in the x-ray cavity. The FEL interaction with these short bunches regeneratively amplifies the radiation intensity and broadens its spectrum. The downstream crystal transmits the part of the radiation spectrum outside its bandwidth and feeds back the filtered radiation to continue the amplification process. This approach uses a significantly shorter undulator but requires a bunch train that is uniform in space and energy.

C. Nonlinear harmonic generation

The ability to generate coherent harmonic radiation is an important aspect of an x-ray FEL. In a planar undulator, the electron trajectory is not a pure sinusoid due to the fact that the longitudinal velocity oscillates at one-half of the undulator period [see Eq. (8)]. This fact leads to the odd harmonic emission along the undulator axis. More specifically, when we change the dependent coordinate from t to θ in the paraxial wave Eq. (38), we should use the exact arrival time of the electron $t = \bar{t} + (K_0^2/(8ck_u\gamma^2)\sin(2k_u z)$ and the relation

$$\begin{aligned}\theta(z) &= (k_u + k_1)z - ck_1\bar{t} \\ &= (k_u + k_1)z - ck_1 t + \xi\sin(2k_u z),\end{aligned} \tag{106}$$

where $\bar{t}$ is the undulator-period-averaged arrival time and $\xi = K_0^2/(4 + 2K_0^2)$ as defined previously in Sec. II. Because the right-hand side of Eq. (38) under the time integral is periodic in z due to the fast wiggling motion, we average Eq. (38) over the undulator period λ_u with the help of the Bessel function expansion

$$e^{i\nu\xi\sin(2k_u z)} = \sum_{p=-\infty}^{+\infty} J_p(\nu\xi)e^{i2pk_u z}. \tag{107}$$

This undulator-period averaging is nonzero only when ν is close to an odd integer $h = 2p \pm 1 = \ldots, -3, -1, 1, 3, \ldots$. Thus, the harmonic field amplitude $E_\nu(\boldsymbol{x}; z)$ at $\nu \sim h$ is given by

$$\left(\frac{\partial}{\partial z} + \frac{\nabla_\perp^2}{2ihk_1}\right)E_\nu(\boldsymbol{x}; z) = \frac{eK_h}{2\epsilon_0\gamma_0}e^{i\Delta\nu_h k_u z}\int\frac{k_1 d\theta}{2\pi}e^{-i\nu\theta} \times \sum_{j=1}^{N}\delta(\boldsymbol{x} - \boldsymbol{x}_j)\delta(\theta - \theta_j). \tag{108}$$

Here, the effective coupling strength of the hth harmonic is

$$K_h = K_0(-1)^{(h-1)/2}[J_{(h-1)/2}(h\xi) - J_{(h+1)/2}(h\xi)]. \tag{109}$$

In the previous notation, we had $K_1 = K_0[\mathrm{JJ}]$. Thus, in the forward z direction, the electric field consists of a series of nearly monochromatic waves around the harmonic frequencies hck_1 [22], with the frequency detuning $\Delta\nu_h = \nu - h \ll 1$.

The FEL interaction introduces both energy and density modulations of the electron beam with the period λ_1. Close to saturation, strong bunching at the fundamental frequency ω_1 produces rich harmonic bunching and significant harmonic radiation in a planar undulator [83,84]. This nonlinear harmonic bunching process is qualitatively illustrated in Fig. 11. Taking into account electron energy spread and emittance, as well as the radiation diffraction and guiding, a 3D analysis of nonlinear harmonic generation [39] shows that the gain length, and transverse and temporal properties of the first few harmonics are eventually governed by those of the fundamental after a certain stage of exponential growth. For instance, driven by the third power of the radiation mode at the fundamental wavelength, the third nonlinear harmonic radiation grows 3 times faster than the fundamental with a coherent transverse mode and a more spiky temporal structure. As a numerical example, the third harmonic power P_3 (at 0.5 Å) of the LCLS radiation before saturation is analytically estimated to be [39]

$$\frac{P_3}{\rho P_{\text{beam}}} = \left(\frac{P_1}{\rho P_{\text{beam}}}\right)^3\begin{cases}\times 0.018 & \text{in the seeded mode} \\ \times 0.11 & \text{in the SASE mode,}\end{cases} \tag{110}$$

where P_1 is the fundamental radiation power. Because of the nonlinear statistics, the third-harmonic power in the SASE mode is higher by a factor of 6 than that in the seeded case. Figure 12 shows both the LCLS fundamental and the third-harmonic power of the seeded case obtained

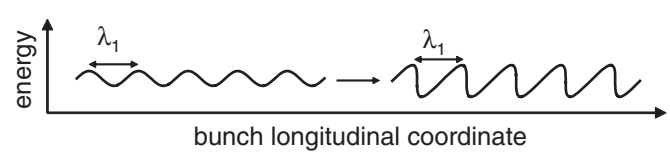


FIG. 11. Longitudinal phase space picture of the electron bunch. In the small signal regime, both the energy and the density modulations are sinusoidal at the fundamental wavelength λ_1 (left). Near saturation, the nonlinear modulation at the fundamental induces strong harmonic bunching in the beam current (right).

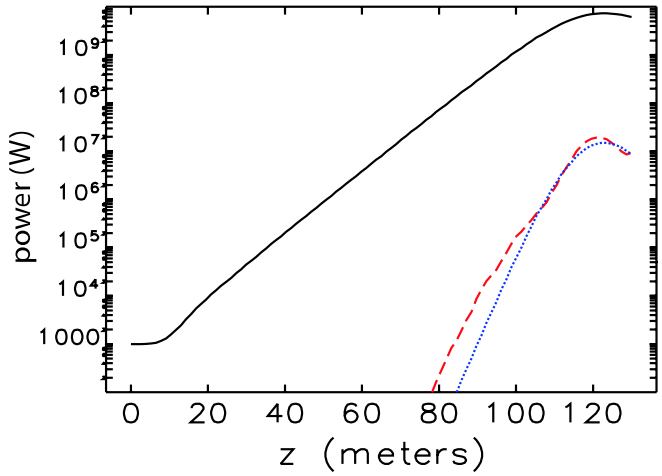


FIG. 12. (Color) Fundamental (black solid curve) and third-harmonic (red dashed curve) powers in GINGER seeded simulation for the LCLS parameters at the optimal frequency detuning. The analytical estimate according to Eq. (110) is shown as the blue dotted curve.

from GINGER simulation at the optimal frequency detuning. Simulation for the SASE mode shows more complicated third-harmonic evolution due to the shot-noise background of the higher harmonics and the saturation effect. The analytical estimate for the third-harmonic power is valid only for a short distance just before saturation. GINGER SASE simulation shows that the third-harmonic power at saturation can reach almost 1% of its fundamental power at 1.5 Å. A 1D SASE simulation study [85] shows that the maximum third-harmonic power at saturation (for a cold beam) is about 2% of the fundamental level. The coherence time at saturation falls inversely proportional to the harmonic number, while the relative spectral bandwidth is independent of the harmonic number.

In general, the third nonlinear harmonic radiation is the most significant harmonic component and can naturally extend the wavelength reach of the x-ray FEL by a factor of 3. The naturally synchronized fundamental and third-harmonic radiation open up possibilities for two-color pumb-probe experiments [86]. The pronounced temporal spikes of the nonlinear harmonic radiation may allow selection of a short temporal pulse with high intensity [87].

Even harmonic radiation exists at an angle away from the undulator z axis. Although the microbunched electron beam at saturation contains more second-harmonic bunching than the third-harmonic bunching, the coupling strengths to even harmonic radiation are usually much weaker for x-ray FELs employing high-energy electrons [88,89]. For instance, the second-harmonic radiation for the LCLS FEL is negligible. Nevertheless, the second-harmonic radiation may still be significant for long-wavelength FELs using relatively low-energy electron beams as experimentally observed in Refs. [7,90]. For an x-ray FEL such as the LCLS, an "afterburner" undulator with its fundamental wavelength tuned to the second harmonic of the main undulator may be used parasitically to extract coherent second-harmonic radiation (at 0.75 Å) at a power level higher than the third-harmonic radiation that accompanies the fundamental in the main undulator [91].

D. Saturation behavior

The radiation characteristics after saturation are more complex, especially for SASE FELs. Linear and quasi-linear theories do not apply, and simulation codes are required to accurately predict the saturation behavior. The FEL bandwidth starts to increase due to the appearance of sidebands associated with synchrotron oscillations of electrons trapped in the ponderomotive potential [92]. In general both the transverse and the temporal coherence decrease with the undulator distance in the saturation regime. Although the fluctuation of the total radiated energy is reduced after saturation, the fluctuation of a single frequency mode filtered by a monochromator is still 100% just as in the exponential growth regime [60]. An analytical model that reproduces such a statistical fluctuation in the early saturation regime was recently developed in Ref. [93].

V. UNDULATOR ERRORS AND WAKEFIELDS

The design of a typical x-ray FEL calls for a small-gap undulator system about 100 m in length, consisting of many undulator sections with beam focusing/steering/diagnostic stations between the sections. Errors in undulator magnetic field and electron beam steering can degrade the FEL performance. In addition, wakefields induced by a high-current beam in the small-gap vacuum chamber can also interfere with the FEL gain process. In this section, we illustrate how FEL theory may be applied to study these effects.

A. Undulator errors

We will assume that each undulator segment is shimmed to have vanishing first and second magnetic field integrals (no net steering errors) and focus on the variations of the undulator parameter K due to magnetic field errors or transverse misalignments among segments. Using the 1D FEL equations, Yu *et al.* [94] studied the effect of undulator errors on FEL performance. When the undulator strength parameter has an error $\Delta K = K - K_0$, Eq. (11) can be written as

$$\frac{d\theta}{dz} = k_u - k_1 \frac{1 + (K_0 + \Delta K)^2/2}{2\gamma^2} \approx 2k_u\eta - k_u \frac{K_0 \Delta K(z)}{1 + K_0^2/2}. \tag{111}$$

Here, the first term describes the ideal motion, and the second term is the amount of the phase kick due to small

changes in K. As a concrete model, we take

$$\Delta K(z) = \Delta K_n \quad \text{for } (n-1)L_c < z < nL_c \quad (n = 1, 2, 3, \ldots), \tag{112}$$

where ΔK_n is a random quantity with the ensemble average $\langle \Delta K_n \rangle = 0$. We have introduced a magnetic correlation length $L_c = N_c \lambda_u$, which is assumed to be much shorter than the approximate field amplitude gain length $2L_G \approx \lambda_u/(4\pi\rho)$. Then the net phase shift per gain length is

$$\Delta\theta = \sum_{n=1}^{2L_G/L_c} N_c \frac{2\pi K_0 \Delta K_n}{1 + K_0^2/2}. \tag{113}$$

For $2L_G/L_c \gg 1$, $\Delta\theta$ has a zero mean and a variance

$$\overline{(\Delta\theta)^2} = \frac{L_g}{L_c}\left(N_c \frac{2\pi K_0^2}{1 + K_0^2/2} \frac{\sigma_K}{K_0}\right)^2 = \frac{\pi N_c K_0^4}{(1 + K_0^2/2)^2} \frac{(\sigma_K/K_0)^2}{\rho} \approx 4\pi N_c \frac{(\sigma_K/K_0)^2}{\rho}, \tag{114}$$

where σ_K is the rms value of ΔK_n. A perturbation analysis yields the radiation power as [94]

$$P \approx P_0 \exp\left[-\frac{z}{L_G}\frac{\overline{(\Delta\theta)^2}}{9}\right], \tag{115}$$

where P_0 is the power along the undulator without any error.

For a negligible power degradation near the SASE saturation at $z \approx 20L_G$, the mean square of the ponderomotive phase shift per gain length is $\overline{\Delta\theta^2} \ll 1$. For errors associated with magnetic pole field B_0 that may occur every undulator period, $N_c \sim 1$, the condition becomes [94]

$$\frac{\sigma_B}{B_0} < \sqrt{\frac{\rho}{4\pi}}. \tag{116}$$

Hence, the pole field error tolerance is quite relaxed because it scales as $\sqrt{\rho}$ instead of ρ. On the other hand, if the length of the undulator segment is a significant fraction of $2L_G$ as in the LCLS case, the error in the average undulator parameter K per segment is now correlated over $N_c \to (4\pi\rho)^{-1}$. Although the perturbation analysis is not strictly valid in this case, Eq. (114) suggests that the error tolerance for K is

$$\frac{\sigma_K}{K_0} < \rho. \tag{117}$$

The LCLS has the FEL parameter $\rho \approx 4.5 \times 10^{-4}$ and 33 undulator segments (each 3.4 m in length) [11]. Figure 13 shows that the GENESIS SASE simulation results for the LCLS undulator segment K errors is in qualitative agreement with the requirement of Eq. (117).

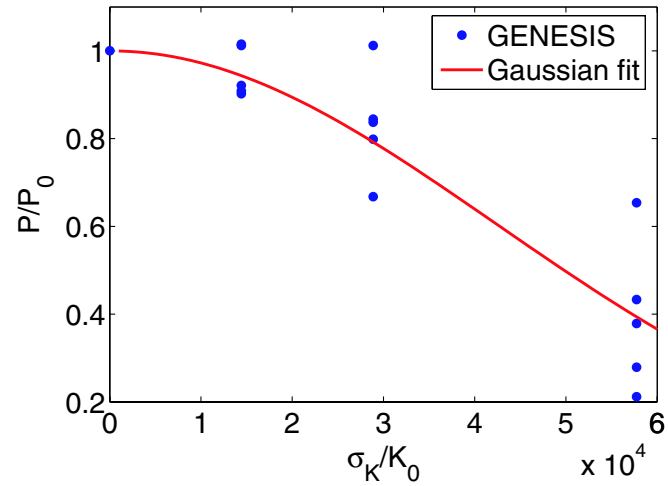


FIG. 13. (Color) Power degradation factor P/P_0 at FEL saturation versus σ_K/K_0 in the LCLS 33 undulator segments. Here, σ_K is the rms value of a uniform segment K error distribution. Five random error distributions are used for a given σ_K. The rms width of the Gaussian fit is 4.2×10^{-4}.

B. Beam trajectory errors

The effects of nonstraight beam trajectory may be illustrated with a heuristic 3D model when a microbunched beam is kicked by a single error dipole field (e.g., a misaligned quadrupole) [95]. While the direction of the beam trajectory changes after the kick by a deflecting angle ϕ, the wavefront orientation normal to the microbunching plane does not. This discrepancy results in two mechanisms for gain degradation: a decrease in coherent radiation power and an increased smearing of microbunching due to the intrinsic angular spread. Both mechanisms are characterized by a critical angle [95]

$$\phi_c = \sqrt{\frac{\lambda_1}{L_G}}, \tag{118}$$

and the power gain length after the kick becomes approximately $L_G/(1 - \phi^2/\phi_c^2)$. In the LCLS case, $\phi_c \approx 6$ μrad at $\lambda_1 = 1.5$ Å for $L_G \approx 4$ m.

For random trajectory errors that are periodically corrected by steering elements at beam position monitor locations between the undulator sections, a statistical analysis based on the previous phase error model is given in Ref. [94]. When the separation of the corrector stations L_s is smaller than the gain length, the radiation power for an rms trajectory deviation $x_{\rm rms}$ is

$$P \approx P_0 \exp\left[-\left(\frac{x_{\rm rms}}{x_{\rm tol}}\right)^{1/4}\right], \quad x_{\rm tol} = 0.266\left(\frac{L_s}{L_G}\right)^{3/4}\left(\frac{L_G}{z}\right)^{1/4}\sqrt{\lambda_1 L_G}. \tag{119}$$

For the LCLS, we can take $L_s = 3.4$ m, $L_G = 4$ m, and $z/L_G \approx 20$ for the saturation undulator distance, then

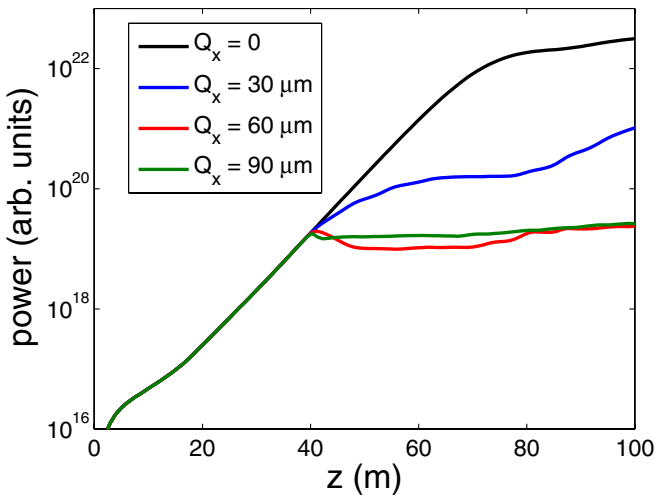


FIG. 14. (Color) GENESIS simulation of the LCLS far-field power for various quadrupole offsets Q_x at $z = 40$ m.

$x_{\rm tol} \approx 3$ μm, and the rms trajectory angle should be controlled to within 1 μrad in order to guarantee a small power degradation.

Since a large trajectory distortion can destroy the FEL interaction, kicking the beam at selected undulator locations may facilitate the z-dependent FEL power measurements using a single diagnostic station at the end of the undulator beam line. This technique is especially useful when intraundulator FEL diagnostic stations are difficult to install. Let us illustrate the trajectory distortion method for the LCLS. For a quadrupole with a focal length $f_Q = 10$ m, a small horizontal offset $Q_x = 60$ μm corresponds to a kick angle $\phi = Q_x/f_Q = \phi_c$. Figure 14 shows that a quadrupole at $z = 40$ m with a horizontal offset $Q_x \geq 60$ μm (i.e., a kick angle $\phi \geq \phi_c$) inhibits further growth of the FEL fundamental mode, producing an approximately constant on-axis radiation intensity, which may be detected by a far-field x-ray diagnostic station after the undulator. Similar conclusions hold at other undulator locations in the exponential growth regime.

C. Wakefield effects

A high-current electron bunch induces a short-range wakefield that changes the beam properties in the long undulator vacuum chamber. For the LCLS, the dominant (longitudinal) wakefield is caused by the resistive wall of the vacuum pipe [96] and creates an energy variation along the undulator distance as well as along the bunch position. Since the x-ray coherence time t_c is much shorter than the x-ray pulse duration T, the wakefield-induced energy variation in an electron temporal slice of length ct_c (known as an FEL slice) is usually negligible for a typical wakefield that does not vary rapidly inside the bunch. Thus, the main effect of the undulator wakefield in an FEL slice is to change the slice central energy and consequently shift its resonant wavelength along the undulator distance. If we still use the initial resonant wavelength to define the ponderomotive phase as was done in Sec. III A, the phase Eq. (11) can now be rewritten as

$$\frac{d\theta}{dz} = 2k_u \frac{[\gamma(z) - \gamma_0]}{\gamma_0} = 2k_u \Bigg[\underbrace{\frac{\gamma(z) - \gamma_c(z)}{\gamma_0}}_{\equiv \eta} + \underbrace{\frac{\gamma_c(z) - \gamma_0}{\gamma_0}}_{\equiv \zeta} \Bigg], \tag{120}$$

where $\eta(z)$ is now the energy deviation from the slice central energy $\gamma_c mc^2$ and is still governed by Eq. (13) due to the FEL interaction, and $\zeta(z)$ is the wakefield-induced energy change relative to the initial energy $\gamma_0 mc^2$ for a particular slice of the bunch. Thus, the wakefield effect for this slice is equivalent to an undulator taper that also changes the resonant wavelength. The last point can be seen by comparing Eqs. (120) and (111). They are equivalent when

$$\zeta(z) = -\frac{K_0 \Delta K(z)}{(2 + K_0^2)} \approx -\frac{\Delta K(z)}{K_0}, \qquad \text{for } K_0^2 \gg 2. \tag{121}$$

Note that ζ in Eq. (120) is not a randomly fluctuating quantity as was the case for undulator errors. Instead, ζ is a linear function of z for wakefield-induced energy change or an equivalent linear taper of the undulator parameter. Unlike the uniform energy loss due to the spontaneous undulator radiation, the wakefield-induced energy change varies from slice to slice along the bunch coordinate, hence a unique undulator taper cannot perfectly compensate the energy change for all bunch slices.

In general, $\zeta(z)$ is not small but can be considered as slowly varying if the fractional energy change per field gain length is less than ρ. In the small signal regime before saturation, the WKB approximation can be used to solve the FEL equations and to obtain the SASE power as [97]

$$P(z) \approx P_m(z) \exp\Bigg[-\frac{1}{2} \Bigg(\frac{\zeta(z) - \zeta_m(z)}{\sqrt{3}\sigma_\omega(z)/\omega_1} \Bigg)^2 \Bigg], \tag{122}$$

where P_m is the maximum power at the optimal energy change $\zeta_m > 0$ or an equivalent undulator taper, and $P_m > P_0$ with P_0 being the radiation power when $\zeta(z) = 0$. Thus, a small energy gain is actually beneficial to the SASE output power. For the LCLS, simulations show that a fractional energy increase of 2ρ over the saturation distance $z_{\rm sat} \approx 90$ m improves the saturation power by about a factor of 2 as compared to the nominal saturation power without any external energy change or taper [e.g., that given by Eq. (91) or FEL simulations without any wakefield and taper]. Because the LCLS bandwidth $\sigma_\omega(z_{\rm sat})/\omega_1$ is close to ρ, Eq. (122) indicates that the SASE power has a FWHM in $\zeta \approx 4\rho$ at saturation.

For a given wake energy variation as a function of the bunch coordinate, Eq. (122) can be used to estimate the

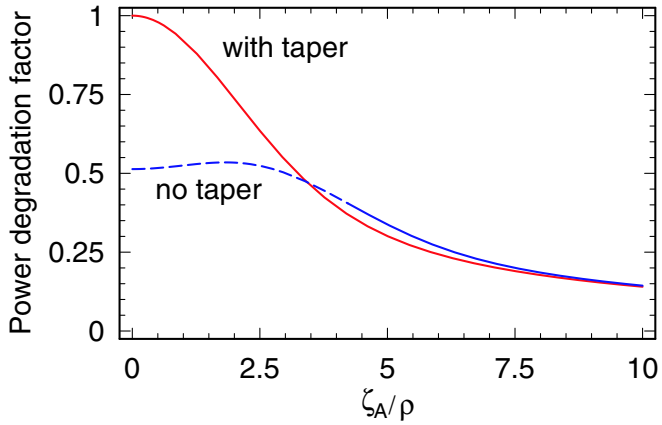


FIG. 15. (Color) Power degradation factor averaged over the core part of the bunch (with about 30 μm in length) versus the sinusoidal wake oscillation amplitude ζ_A/ρ at the LCLS saturation ($z = 90$ m) for a prescribed tapered undulator (red solid curve) and without any taper (blue dashed curve) (from Ref. [97]).

FEL power along the bunch position and to find the average SASE power over the bunch. As a numerical example, Ref. [97] studied the case for a sinusoidal energy oscillation that resembles the resistive wall wakefield in the core part of the 1-nC LCLS bunch [96]. Figure 15 shows the average power degradation factor (with respect to the maximum power P_m) as a function of the fractional energy oscillation amplitude ζ_A without and with a linear taper that yields $\zeta_m = 2\rho$ at $z_{\rm sat}$. For a round 5-mm-diameter vacuum pipe, $\zeta_A \approx 6\rho$ for Cu and 3ρ for Al at $z_{\rm sat} = 90$ m. The average power in this part of the bunch is then about 50% (25%) of P_m for the Al (Cu) vacuum pipe and insensitive to the undulator taper for large energy oscillation amplitudes as shown in Fig. 15. In order to reduce the wakefield effects in the undulator as well as in the accelerator, a 200-pC bunch configuration was recently proposed for the LCLS [98]. In addition to compensating the average wake energy loss, the start-to-end LCLS simulations [99] show that an additional undulator taper of about 2ρ improves the saturation power by about a factor of 2, making its radiation energy comparable to the 1-nC case that suffers stronger wakefield effects.

VI. GAIN ENHANCEMENT METHODS

As discussed in Sec. III F, a key requirement in realizing x-ray FELs is high-quality electron beams. Since the SASE coherence time is relatively short as compared to the electron bunch length, the slice beam qualities (i.e., the local emittance and energy spread on the scale of the coherence time) are more relevant than the global ones. When high-energy electrons are employed to drive the x-ray FELs, the relative slice energy spread is usually too small to affect the SASE gain process. However, the FEL performance depends critically on the transverse brightness of the beam defined as

$$B_\perp = \frac{I_e}{4\pi^2\varepsilon_n^2}. \tag{123}$$

Here $\varepsilon_n = \gamma_0\varepsilon$ is the transverse normalized emittance. In the x-ray wavelength with negligible energy spread, the power gain length at the optimized beta function may be estimated as [100]

$$L_G = 1.2\left(\frac{I_A}{I}\right)^{1/2}\frac{\varepsilon_n^{5/6}\lambda_u^{5/6}}{\lambda_r^{2/3}}\frac{(1+K_0^2/2)^{1/3}}{K_0[\mathrm{JJ}]} \propto B_\perp^{-1/2}\varepsilon_n^{-1/6}. \tag{124}$$

The state-of-the-art photocathode rf gun is expected to produce beams with a normalized emittance of about 1 μm at about 1 nC charge. With an optimized bunch compression configuration, the bunch can be compressed to a peak electron current of 3 to 4 kA. These expectations are reflected in the LCLS design parameters listed in Table I. Hence, the FEL power gain length predicted from Eq. (80) or (124) is about 4 to 5 m, and the saturation length is about 80 to 100 m.

We will now discuss some advanced beam-manipulation methods that may overcome these apparent beam-quality limitations in order to enhance the FEL performance.

A. Beam conditioning

It was realized in Ref. [101] that the angular spread of a finite-emittance beam in the undulator can be compensated for if each electron's energy deviation is made to be proportional to the square of its betatron amplitude. This can be understood by expanding the undulator-period-averaged longitudinal velocity [i.e., Eq. (23) or (28)] for small $\Delta\gamma = \gamma - \gamma_0$. The average longitudinal velocity will not depend on the transverse actions $J_{x,y}$ when

$$\frac{(1+K_0^2/2)}{\gamma_0^2}\frac{\Delta\gamma}{\gamma_0} = \frac{(J_x+J_y)}{\bar\beta}. \tag{125}$$

Averaging Eq. (125) over the beam (i.e., $\langle J_{x,y}\rangle = \varepsilon = \varepsilon_n/\gamma_0$) and using the FEL resonant condition, the required correlated energy spread to compensate for the emittance effect or to "condition" the beam is

$$\langle\Delta\gamma\rangle = \frac{\lambda_u}{\lambda_1}\frac{\varepsilon_n}{\bar\beta}. \tag{126}$$

Thus, the FEL gain degradation due to the angular spread of a finite-emittance beam can be eliminated by conditioning the electron beam prior to the undulator entrance according to Eq. (125). For such a conditioned beam in a natural-focusing undulator, we can take $\hat{k}_\beta\hat{\sigma}_x = 0$ in Eq. (78) and $\eta_\varepsilon = 0$ in Eq. (79), and the power gain length from Eq. (80) can be much reduced. For a strong-focusing

undulator employing FODO cells discussed in Sec. III B, however, the oscillatory effect of the angular spread as shown in Eq. (27) remains even though the average effect of the angular spread is eliminated by conditioning. In this case, we can take $\eta_\varepsilon = 0$ in Eq. (80) only when [102]

$$\varepsilon_n < \gamma_0 \frac{\lambda_1 \bar{\beta}}{\pi L_c}. \tag{127}$$

This is usually a more tolerable emittance requirement than an unconditioned beam [see Eq. (73)].

The nonlinear correlation of Eq. (125) requires special nonlinear accelerator components that pose technical challenges. The discussion of various proposed methods to condition the electron beams is beyond the scope of this paper. Reference [103] contains a recent review on the subject.

B. Current-enhanced SASE

High electron peak current is an essential requirement for efficient FEL interaction (through the FEL scaling parameter ρ). In a typical x-ray FEL accelerator system, bunch-compressor chicanes are designed to increase the beam current to a few-kA level. Further compression is increasingly difficult due to short-bunch collective effects in the accelerator. Recently, Zholents proposed introduction of a GW-level optical laser beam to induce large energy modulation in the electron beam in a special wiggler placed in the accelerator [104]. This energy modulation can then be converted to a large-density modulation at the optical wavelength by a weak chicane prior to the FEL undulator. The local peak current of the modulated beam can be tens of kA without strong emittance-deteriorating effects due to the small amount of charge concentrated in high-current regions. Such a current-enhanced beam may enable the x-ray SASE FEL to saturate in a shorter undulator distance or to decrease the x-ray wavelength for a fixed undulator length. Reference [105] details a possible implementation of this scheme in the LCLS.

C. Optical klystron enhancement

An optical klystron FEL [106] uses dispersive sections (magnetic chicanes) between undulators to speed up the FEL microbunching process and has been successfully implemented in many FEL oscillator facilities. A prerequisite for the effectiveness of the optical klystron is the small relative energy spread. For a high-gain FEL, the requirement is [107]

$$\sigma_\eta \ll \rho. \tag{128}$$

Motivated by the very small uncorrelated energy spread (a few keV) of the electron beam that has been measured in a photocathode rf gun [108], the optical klystron enhancement to SASE FELs was recently studied in Ref. [109]. The optimal momentum compaction R_{56} of the chicane can be determined as

$$R_{56} \approx \frac{1}{k_1 \sigma_\eta}. \tag{129}$$

The additional slippage of the electron beam to the radiation introduced by the chicane is

$$\frac{R_{56}}{2} \approx \frac{\lambda_1}{4\pi\sigma_\eta} \gg \frac{\lambda_1}{4\pi\rho} \sim t_c, \tag{130}$$

where t_c is the SASE coherence time defined in Eq. (102). This large slippage is beneficial to a SASE optical klystron device because the microbunched electron beam does not have to match the radiation phase when the dispersively enhanced microbunching does not overwhelm the radiation from the earlier undulator [109]. Thus, the output power will not be sensitive to a small variation of the chicane R_{56} (at the Angstrom level) or a small energy jitter (at the 10^{-4} level), in contrast to an early simulation study with a seeded FEL [110]. The simulated SASE performance with the addition of four optical klystrons located at the undulator breaks in the LCLS shows significant improvement if the slice energy spread at the undulator entrance can be controlled to 5×10^{-5} [109]. In addition, FEL saturation at shorter x-ray wavelengths (around 1.0 Å) within the LCLS undulator length becomes possible.

D. Emittance exchange

The very small uncorrelated energy spread of the beam from an rf gun also opens up the possibility of exchanging the small longitudinal emittance with a large transverse emittance, and hence increasing the transverse brightness of the beam for an x-ray FEL. First, a special, "flat-beam" rf gun [111] can be used to produce beams with a large ratio of transverse emittances. Then a transverse-to-longitudinal emittance-exchange beam line can be used to switch the larger transverse emittance with the smaller longitudinal one. An approximate optics for emittance exchange consisting of a dipole mode cavity in the middle of two doglegs of opposite kicks (i.e., a magnetic chicane) was discussed in Ref. [112]. The scheme is adequate when the emittance ratio is not too large. An exact emittance-exchange optics found recently by Kim is similar to the one in Ref. [112], but with the second dogleg in the same direction as the first one [113] (see also Ref. [103]). Such a transverse-to-longitudinal emittance-exchange optics is capable of handling very large-emittance ratios. Together with a short-pulse, flat-beam rf gun, this scheme may produce beams with normalized transverse emittances on the order of 0.1 μm and a compressed current on the order of kA, which may be used to drive a sub-Angstrom x-ray FEL.

VII. CONCLUSIONS

We have reviewed the theory of the high-gain FELs, especially the SASE FELs, in both ideal and more realistic accelerator environments. These analytical results are useful in providing physical pictures, benchmarking simulation codes, and guiding the FEL designs and experiments that are currently building toward x-ray lasers. We also highlighted several research directions toward shortening the x-ray pulse lengths, increasing the temporal coherence of the source, and enhancing the FEL performance. We hope that the formulas and ideas summarized here will stimulate further progress in realizing x-ray FELs, improving their performance, and reducing their size and cost.

ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy Contracts No. DE-AC02-76SF00515 and No. DE-AC02-06CH11357.

APPENDIX A: LIST OF SYMBOLS

Symbol	Physical meaning
a_ν	Scaled electric field
A_n	n^{th}-order transverse mode profile ($n = 0, 1, 2, \ldots$)
$\alpha_{x,y}$	Twiss parameter α
B_0	Undulator peak magnetic field on axis
$\bar{\beta}$	Average transverse beta function
β_n	Natural beta function in a parabolic-pole-faced undulator
$\beta_{x,y}$	Transverse beta function
c	Speed of light in vacuum
$\Delta\nu$	Relative frequency detuning
e	Electron charge
E_0	Transverse electric field amplitude
E_ν	Fourier component of the transverse electric field
E_x	Transverse electric field
η	Relative energy deviation
η_d	Diffraction parameter
η_ϵ	Angular spread parameter
η_γ	Energy spread parameter
ϵ_0	Vacuum permittivity
ε or $\varepsilon_{x,y}$	Transverse emittance of the electron beam
ε_n	Normalized transverse emittance of the electron beam
ε_{r0}	Diffraction-limited radiation emittance
ε_r	Radiation emittance
f or F	Electron phase space distribution function
$f_{0,1}$	Zeroth-order (first-order) phase space distribution function
f_ν	Fourier component of the distribution function
γ_0	Electron reference energy (in units of mc^2)
γ	Electron energy (in units of mc^2)
h	Harmonic order
$\hbar$	Planck constant
I_0	Zeroth-order modified Bessel function
I_e	Electron bunch peak current
I_A	Alfvén current
j_x	Transverse current
$J_{x,y}$	Transverse action
J_n	Bessel function of order n ($n = 0, 1, 2, \ldots$)
k_1	Fundamental undulator radiation wave number
k_β	Average betatron focusing wave number
k_n	Parabolic-pole-faced undulator natural-focusing wave number
k_{n0}	Planar undulator vertical natural-focusing wave number
k_p	Longitudinal plasma oscillation wave number
k_u	Undulator wave number
K_0	Nominal undulator strength parameter
K_h	Effective coupling strength of the hth harmonic radiation
L_{G0}	1D FEL power gain length of a monoenergetic beam
L_G	3D FEL power gain length
L_s	Undulator section length
L_u	Active undulator length
λ_1	Fundamental FEL wavelength
λ_r	Undulator resonant wavelength at an arbitrary angle
λ_u	Undulator period
Λ	FEL gain length degradation factor
m	Electron rest mass
M	Total number of independent mode in a radiation pulse
$M_{T,L}$	Transverse or longitudinal mode number
μ_n	Scaled growth rate of the nth transverse mode ($n = 0, 1, \ldots$)
n_0	Peak electron volume density
n_e	Electron volume density function
N_e	Total number of electrons in a bunch
ω_1	Fundamental undulator radiation frequency
ν	Ratio of the radiation frequency to the fundamental frequency ω_1
$\mathbf{p}_\beta$	Divergence angle vector of transverse betatron motion
$\mathbf{p}$	Divergence angle vector due to natural focusing or smoothed betatron focusing
P	Radiation power
P_{beam}	Electron beam power
P_n	Effective SASE start-up noise power
P_{sat}	FEL saturation power
P_0	Fundamental radiation power without error or taper
$P_{1,3}$	Radiation power at the fundamental (third-harmonic) frequency
P_m	Fundamental radiation power with an optimal taper
ϕ	Angle of the electron trajectory relative to the radiation propagation
ψ_0	Initial phase of the radiation wave
r_e	Classical electron radius
ρ	FEL Pierce parameter
σ_η	rms relative energy spread of the electron beam
σ_r	rms transverse radiation size
σ_ω	rms SASE bandwidth

Symbol	Physical meaning
σ_W	rms radiation energy fluctuation
σ_x	rms transverse size of the electron beam
$\sigma_{x'}$	rms transverse divergence of the electron beam
t	Electron's arrival time at the undulator location z
$\bar{t}$	Electron's arrival time averaged over an undulator period
T	Flattop electron bunch duration
t_c	Radiation temporal coherence time
θ	Electron's phase relative to the radiation wave
u	Radiation frequency chirp (frequency change per unit time)
v_g	SASE radiation group velocity
$v_{x,y}$	Electron's transverse velocity
v_z	Electron's longitudinal velocity
$\bar{v}_z$	Electron's average longitudinal velocity in a planar undulator
W	Total radiation energy
x	Electron's horizontal position
$\mathbf{x}$	Two-component vector representing smoothed transverse betatron motion
y	Electron's vertical position
z	Distance from the undulator beginning
z_{sat}	FEL saturation distance
ζ	Wakefield-induced fractional energy change
Z_R	Rayleigh length of the radiation

[1] J. Madey, J. Appl. Phys. **42**, 1906 (1971).
[2] D. A. G. Deacon, L. R. Elias, J. M. J. Madey, G. J. Ramian, H. A. Schwettman, and T. I. Smith, Phys. Rev. Lett. **38**, 892 (1977).
[3] A. Kondratenko and E. Saldin, Part. Accel. **10**, 207 (1980).
[4] Y. Derbenev, A. Kondratenko, and E. Saldin, Nucl. Instrum. Methods Phys. Res., Sect. A **193**, 415 (1982).
[5] R. Bonifacio, C. Pellegrini, and L. Narducci, Opt. Commun. **50**, 373 (1984).
[6] S. Milton *et al.*, Science **292**, 2037 (2001).
[7] A. Tremaine *et al.*, Phys. Rev. Lett. **88**, 204801 (2002).
[8] V. Ayvazyan *et al.*, Phys. Rev. Lett. **88**, 104802 (2002).
[9] V. Ayvazyan *et al.*, Eur. Phys. J. D **37**, 297 (2006).
[10] J. Rossbach, in Proceedings of the 2006 European Particle Accelerator Conference, 2006, p. 34.
[11] Linac Coherent Light Source Conceptual Design Report, SLAC-R-593, SLAC, 2002.
[12] TESLA Technical Design Report, TESLA FEL 2002-09, DESY, 2002.
[13] SPring-8 Compact SASE Source Conceptual Design Report, http://www-xfel.spring8.or.jp (2005).
[14] I. S. Ko, in *Proceedings of the 2005 Free Electron Laser Conference* (Stanford, CA, USA, 2005), p. 216.
[15] H. Wiedemann, *Particle Accelerator Physics I and II* (Springer-Verlag, Berlin, 1999).
[16] C. Brau, *Free Electron Lasers* (Academic Press, New York, 1990).
[17] W. Colson, Phys. Lett. **64**, 190 (1977).
[18] K. Halbach (unpublished); see also Ref. [15], Vol. II, p. 60.
[19] L. Smith, LBL-21391, LBNL, 1986.
[20] E. Scharlemann, J. Appl. Phys. **58**, 2154 (1985).
[21] S. Reiche, Nucl. Instrum. Methods Phys. Res., Sect. A **445**, 90 (2000).
[22] W. Colson, IEEE J. Quantum Electron. **17**, 1417 (1981).
[23] S. Ichimaru, *Basic Principles of Plasma Physics* (W.A. Benjamin, Inc., London, 1973).
[24] K.-J. Kim, Nucl. Instrum. Methods Phys. Res., Sect. A **318**, 489 (1992).
[25] G. Moore, Opt. Commun. **52**, 46 (1984).
[26] G. Moore, Nucl. Instrum. Methods Phys. Res., Sect. A **239**, 19 (1985).
[27] K.-J. Kim, Phys. Rev. Lett. **57**, 1871 (1986).
[28] L.-H. Yu and S. Krinsky, Phys. Lett. **129**, 463 (1988).
[29] E. Scharlemann, A. Sessler, and J. Wurtele, Phys. Rev. Lett. **54**, 1925 (1985).
[30] M. Xie and D. Deacon, Nucl. Instrum. Methods Phys. Res., Sect. A **250**, 426 (1986).
[31] N. Kroll and W. McMullin, Phys. Rev. A **17**, 300 (1978).
[32] L.-H. Yu, S. Krinsky, and R. Gluckstern, Phys. Rev. Lett. **64**, 3011 (1990).
[33] M. Xie, Nucl. Instrum. Methods Phys. Res., Sect. A **445**, 59 (2000).
[34] Y. Chin, K.-J. Kim, and M. Xie, Phys. Rev. A **46**, 6662 (1992).
[35] M. Xie, in *Proceedings of the 1995 Particle Accelerator Conference* (IEEE, Piscataway, NJ, 1995), p. 183.
[36] N. V. Kampen, Physica (Amsterdam) **21**, 949 (1955).
[37] S. Krinsky and L.-H. Yu, Phys. Rev. A **35**, 3406 (1987).
[38] M. Xie, Nucl. Instrum. Methods Phys. Res., Sect. A **475**, 51 (2001).
[39] Z. Huang and K.-J. Kim, Phys. Rev. E **62**, 7295 (2000).
[40] K.-J. Kim, Nucl. Instrum. Methods Phys. Res., Sect. A **250**, 396 (1986).
[41] J.-M. Wang and L.-H. Yu, Nucl. Instrum. Methods Phys. Res., Sect. A **250**, 484 (1986).
[42] L.-H. Yu and S. Krinsky, Nucl. Instrum. Methods Phys. Res., Sect. A **285**, 119 (1989).
[43] Z. Huang and K.-J. Kim, Nucl. Instrum. Methods Phys. Res., Sect. A **475**, 59 (2001).
[44] W. Fawley, Report No. LBNL-49625, LBL, 2002.
[45] S. Reiche, Nucl. Instrum. Methods Phys. Res., Sect. A **429**, 243 (1999).
[46] N. A. Vinokurov, Z. Huang, O. A. Shevchenko, and K.-J. Kim, Nucl. Instrum. Methods Phys. Res., Sect. A **475**, 74 (2001).
[47] R. Bonifacio and F. Casagrande, Opt. Commun. **50**, 251 (1984).
[48] J. Gea-Banacloche, Phys. Rev. A **31**, 1607 (1985).
[49] C. Schroeder, C. Pelligrini, and P. Chen, Phys. Rev. E **64**, 056502 (2001).
[50] R. Bonifacio, N. Piovella, and G. Robb, Nucl. Instrum. Methods Phys. Res., Sect. A **543**, 645 (2005).
[51] R. Bonifacio, N. Piovella, G. Robb, and A. Schiavi, Phys. Rev. ST Accel. Beams **9**, 090701 (2006).
[52] M. Sands, SLAC-R-121, SLAC, 1970.
[53] E. Saldin, E. Schneidmiller, and M. Yurkov, Nucl. Instrum. Methods Phys. Res., Sect. A **381**, 545 (1996).

[54] J. Rossbach, E. Saldin, E. Schneidmiller, and M. Yurkov, Nucl. Instrum. Methods Phys. Res., Sect. A **374**, 401 (1996).
[55] M. Xie, Nucl. Instrum. Methods Phys. Res., Sect. A **445**, 67 (2000).
[56] E. Saldin, E. Schneidmiller, and M. Yurkov, Opt. Commun. **186**, 185 (2000).
[57] K.-J. Kim and Z. Huang, in *Proceedings of the ICFA Beam Dynamics Workshop on the Physics of and Science with the X-ray Free-Electron Laser*, edited by I. L. S. Chattopadhyay, M. Cornacchia, and C. Pellegrini, AIP Conference Proceedings No. 581 (AIP, New York, 2001), p. 185.
[58] E. Saldin, E. Schneidmiller, and M. Yurkov, DESY 06-137, DESY, 2006.
[59] R. Bonifacio, L. D. Salvo, P. Pierini, N. Piovella, and C. Pellegrini, Phys. Rev. Lett. **73**, 70 (1994).
[60] E. Saldin, E. Schneidmiller, and M. Yurkov, Opt. Commun. **148**, 383 (1998).
[61] L.-H. Yu and S. Krinsky, Nucl. Instrum. Methods Phys. Res., Sect. A **407**, 261 (1998).
[62] K.-J. Kim, in *Towards X-ray Free Electron Lasers*, edited by R. Bonifacio and W. Barletta, AIP Conference Proceedings No. 413 (AIP, New York, 1997), p. 3.
[63] J. Goodman, *Statistical Optics* (John Wiley & Sons, Inc., New York, 2000).
[64] S. Krinsky and R. Gluckstern, Phys. Rev. ST Accel. Beams **6**, 050701 (2003).
[65] P. Emma, in *Proceedings of the 2004 European Particle Accelerator Conference* (Lucerne, Switzerland, 2004), p. 225.
[66] C. Pellegrini, Nucl. Instrum. Methods Phys. Res., Sect. A **445**, 124 (2000).
[67] C. Schroeder, C. Pelligrini, S. Reiche, J. Arthor, and P. Emma, J. Opt. Soc. Am. B **19**, 1782 (2002).
[68] S. Krinsky and Z. Huang, Phys. Rev. ST Accel. Beams **6**, 050702 (2003).
[69] A. Zholents and W. Fawley, Phys. Rev. Lett. **92**, 224801 (2004).
[70] E. Saldin, E. Schneidmiller, and M. Yurkov, Phys. Rev. ST Accel. Beams **9**, 050702 (2006).
[71] L.-H. Yu, Phys. Rev. A **44**, 5178 (1991).
[72] E. Saldin, E. Schneidmiller, and M. Yurkov, Opt. Commun. **202**, 169 (2002).
[73] L.-H. Yu *et al.*, Science **289**, 932 (2000).
[74] L.-H. Yu *et al.*, Phys. Rev. Lett. **91**, 074801 (2003).
[75] D. Kramer, in *Proceedings of the 2004 Free Electron Laser Conference* (Trieste, Italy, 2004), p. 108.
[76] C. J. Bocchetta *et al.*, in Ref. [14], p. 682.
[77] J. Feldhaus, E. Saldin, E. Schneidmiller, and M. Yurkov, Opt. Commun. **140**, 341 (1997).
[78] E. Saldin, E. Schneidmiller, Y. Shvydko, and M. Yurkov, Nucl. Instrum. Methods Phys. Res., Sect. A **475**, 357 (2001).
[79] D. Nguyen, R. Sheffield, C. Fortgang, J. Goldstein, J. Kinross-Wright, and N. Ebrahim, Nucl. Instrum. Methods Phys. Res., Sect. A **429**, 125 (1999).
[80] B. Faatz, J. Feldhaus, J. Krzywinski, E. Saldin, E. Schneidmiller, and M. Yurkov, Nucl. Instrum. Methods Phys. Res., Sect. A **429**, 424 (1999).
[81] N. Thompson, M. Poole, and B. McNeil, in Ref. [14], p. 79.
[82] Z. Huang and R. Ruth, Phys. Rev. Lett. **96**, 144801 (2006).
[83] R. Bonifacio, L. D. Salvo, and P. Pierini, Nucl. Instrum. Methods Phys. Res., Sect. A **293**, 627 (1990).
[84] H. Freund, S. Biedron, and S. Milton, IEEE J. Quantum Electron. **36**, 275 (2000).
[85] E. Saldin, E. Schneidmiller, and M. Yurkov, Phys. Rev. ST Accel. Beams **9**, 030702 (2006).
[86] J. Feldhaus, T. Moller, E. Saldin, E. Schneidmiller, and M. Yurkov, Nucl. Instrum. Methods Phys. Res., Sect. A **507**, 435 (2003).
[87] W. Brefeld, B. Faatz, J. Feldhaus, M. Korfer, T. Moller, J. Pflueger, E. Saldin, E. Schneidmiller, J. Krzywinski, and M. Yurkov, DESY 02-038, DESY, 2002.
[88] Z. Huang and K.-J. Kim, Nucl. Instrum. Methods Phys. Res., Sect. A **475**, 112 (2001).
[89] G. Geloni, E. Saldin, E. Schneidmiller, and M. Yurkov, DESY 05-137, DESY, 2005.
[90] S. Biedron *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **483**, 94 (2002).
[91] Z. Huang and S. Reiche, in Ref. [75], p. 201.
[92] N. Kroll, P. Morton, and M. Rosenbluth, IEEE J. Quantum Electron. **17**, 1436 (1981).
[93] S. Krinsky, Phys. Rev. E **69**, 066503 (2004).
[94] L.-H. Yu, S. Krinsky, R. Gluckstern, and J. van Zeijts, Phys. Rev. A **45**, 1163 (1992).
[95] T. Tanaka, H. Kitamura, and T. Shintake, Nucl. Instrum. Methods Phys. Res., Sect. A **528**, 172 (2004).
[96] K. Bane and G. Stupakov, SLAC-PUB-10707, SLAC, 2004.
[97] Z. Huang and G. Stupakov, Phys. Rev. ST Accel. Beams **8**, 040702 (2005).
[98] P. Emma, Z. Huang, C. Limborg, J. Wu, W. Fawley, M. Zolotorev, and S. Reiche, in *Proceedings of the 2005 Particle Accelerator Conference* (IEEE, Piscataway, NJ, 2005), p. 344.
[99] W. Fawley, K. Bane, P. Emma, Z. Huang, H.-D. Nuhn, S. Reiche, and G. Stupakov, in Ref. [14], p. 510.
[100] E. Saldin, E. Schneidmiller, and M. Yurkov, Opt. Commun. **235**, 415 (2004).
[101] A. Sessler, D. Whittum, and L.-H. Yu, Phys. Rev. Lett. **68**, 309 (1992).
[102] Z. Huang, G. Stupakov, and S. Reiche, to be published in *Proceedings of the ICFA Beam Dynamics Workshop onthe Physics and Applications of High Brightness Electron Beams* (World Scientific, Erice, Italy, 2005).
[103] K.-J. Kim and A. Sessler, in *Beam Cooling And Related Topics: International Workshop on Beam Cooling and Related Topics-COOL05*, edited by S. Nagaitsev and R. Pasquinelli, AIP Conference Proceedings No. 821 (AIP, New York, 2005), p. 115.
[104] A. Zholents, Phys. Rev. ST Accel. Beams **8**, 040701 (2005).
[105] A. Zholents, W. Fawley, P. Emma, Z. Huang, G. Stupakov, and S. Reiche, in Ref. [75], p. 582.
[106] N. Vinokurov and A. Skrinsky, INP 77-59, Novosibirsk, 1977.
[107] R. Bonifacio, R. Corsini, and P. Pierini, Phys. Rev. A **45**, 4091 (1992).
[108] M. Huneing and H. Schlarb, in *Proceedings of the 2003*

Particle Accelerator Conference (Portland, USA, 2003), p. 2074.

[109] Y. Ding, P. Emma, Z. Huang, and V. Kumar, Phys. Rev. ST Accel. Beams **9**, 070702 (2006).

[110] G. Neil and H. Freund, Nucl. Instrum. Methods Phys. Res., Sect. A **475**, 381 (2001).

[111] R. Brinkmann, Y. Derbenev, and K. Flottmann, Phys. Rev. ST Accel. Beams **4**, 053501 (2001).

[112] M. Cornacchia and P. Emma, Phys. Rev. ST Accel. Beams **5**, 084001 (2002).

[113] P. Emma, Z. Huang, K.-J. Kim, and P. Piot, Phys. Rev. ST Accel. Beams **9**, 100702 (2006).