Modeling and multidimensional optimization of a tapered free electron laser

Y. Jiao,[1,2,*] J. Wu,[1,†] Y. Cai,[1] A. W. Chao,[1] W. M. Fawley,[1] J. Frisch,[1] Z. Huang,[1] H.-D. Nuhn,[1]
C. Pellegrini,[1,3] and S. Reiche[4]

[1]*SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA*
[2]*Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, China*
[3]*Department of Physics and Astronomy, UCLA, Los Angeles, California 90095-1547, USA*
[4]*Paul Scherrer Institute, Villigen PSI, 5232, Switzerland*
(Received 12 March 2012; published 3 May 2012)

Energy extraction efficiency of a free electron laser (FEL) can be greatly increased using a tapered undulator and self-seeding. However, the extraction rate is limited by various effects that eventually lead to saturation of the peak intensity and power. To better understand these effects, we develop a model extending the Kroll-Morton-Rosenbluth, one-dimensional theory to include the physics of diffraction, optical guiding, and radially resolved particle trapping. The predictions of the model agree well with that of the GENESIS single-frequency numerical simulations. In particular, we discuss the evolution of the electron-radiation interaction along the tapered undulator and show that the decreasing of refractive guiding is the major cause of the efficiency reduction, particle detrapping, and then saturation of the radiation power. With this understanding, we develop a multidimensional optimization scheme based on GENESIS simulations to increase the energy extraction efficiency via an improved taper profile and variation in electron beam radius. We present optimization results for hard x-ray tapered FELs, and the dependence of the maximum extractable radiation power on various parameters of the initial electron beam, radiation field, and the undulator system. We also study the effect of the sideband growth in a tapered FEL. Such growth induces increased particle detrapping and thus decreased refractive guiding that together strongly limit the overall energy extraction efficiency.

I. INTRODUCTION

Recent results on single pulse coherent diffraction imaging of proteins [1] and viruses [2] using an x-ray free electron laser (FEL) show that the resolution can be improved by both increasing the number of the coherent photons and simultaneously reducing the pulse duration to about 10 femtoseconds (fs) or less, thus requiring a peak power of one terawatt (TW) or larger compared to the present values of 20 to 50 GW available at saturation from the self-amplified spontaneous emission (SASE) mode. Theoretical work done at DESY [3] and SLAC [4] shows that one way to increase the peak radiation power of a SASE x-ray FEL to the TW level is to use a tapered undulator, following a concept initially proposed by Kroll, Morton, and Rosenbluth (KMR) [5], together with the self-seeding option [6].

The SLAC work shows the existence of a saturation effect that limits the efficiency of energy transfer from the electrons to the radiation to values below those predicted by the one-dimensional (1D) KMR theory [7,8]. The

*jiaoyi@SLAC.Stanford.EDU; jiaoyi@ihep.ac.cn
†jhwu@SLAC.Stanford.EDU

work is based on numerical simulations, using the three-dimensional (3D), time-dependent codes GENESIS [9] and GINGER [10]. Studies suggest that the saturation effect is due to some combination of diffraction, refraction, radial dependence of the radiation field, and time-dependent, slippage effects.

Better understanding of the limits and capabilities of a self-seeded tapered FEL requires an in-depth study of the 3D effects, in particular diffraction and refraction, absent from the 1D KMR analysis. While GENESIS and GINGER have been benchmarked against many experimental results and give a reliable evaluation of an FEL performance, they require moderately long computing time, limiting the possibility of multidimensional, parametric optimization of FEL performance when 3D and time-dependent effects are important. To this end, we have developed a model of a tapered FEL, using the 1D KMR theory and the optical fiber approximation to describe 3D effects and optical guiding [11–13]. While the model contains some approximations, it allows us to explore the full potential of a high-peak-power FEL not only by tapering the undulator parameters in longitudinal dimension but also by optimizing the transverse effects. The results from the model are compared with those from GENESIS single-frequency numerical simulations to establish the limits of its validity. We then propose a multidimensional optimization scheme, including the change

with z of both the undulator field and the electron beam radial profile. The time-dependent effects not included in the model are also investigated with the GENESIS time-dependent numerical simulations.

KMR use the Hamiltonian approach to derive the equations that describe the electrons' synchrotron oscillations in the bucket associated with the ponderomotive potential in terms of the wiggler magnetic field, wiggler period, and radiation field [5]. The introduction of the synchronous phase Ψ_r is used to formulate both the deceleration rate and the electron trapping fraction F_t. However, the theory is one dimensional, and it assumes constant radiation beam size r_s, electron beam radius r_b, and uniform transverse distribution of the radiation field and electron density. In fact, as shown by the numerical simulations, at the end of the exponential gain regime, the radiation beam in a tapered FEL can expand substantially in the radial direction. Thus, a pure 1D model cannot in general accurately predict the behavior of the electron and radiation beams in the tapered region.

By contrast, the optical guiding approach considers many important transverse effects. As shown in Refs. [11,12], coherent interaction between the radiation and electrons can optically guide and focus the light. Because of its microbunching, the electron beam has an effective complex index of refraction n,

$$n = 1 + \frac{\omega_{p0}^2}{\omega_s^2} \frac{r_{b0}^2}{r_b^2} \frac{a_w}{2|a_s|} [JJ] \left\langle \frac{e^{-i\Psi}}{\gamma} \right\rangle, \tag{1}$$

where ω_p is the electron plasma frequency, $\omega_s = k_s c$ is the radiation frequency, with $\lambda_s = 2\pi/k_s$ the radiation wavelength and c the speed of light, r_b is the electron beam radius, $a_w = |e|B_w/k_w mc^2$ and $a_s = |e|A_s/mc^2$ are the normalized vector potentials of the helical undulator and on-axis radiation field (an additional $2^{1/2}$ factor in denominator for a linearly polarized undulator). Further symbols include e for the elementary charge, mc^2 the rest mass energy of electron, B_w the undulator field amplitude and $\lambda_w = 2\pi/k_w$ the undulator period, $[JJ] = 1$ for helical undulator and $[JJ] = J_0(\kappa) - J_1(\kappa)$ for linearly polarized undulator with $\kappa = a_w^2/2(1 + a_w^2)$, γ the electron's Lorentz factor, and Ψ the electron phase relative to the ponderomotive potential. Quantities with subscript 0 indicate the initial electron and radiation beam parameters. The free space surrounding the electron beam, by contrast, has a refraction index of 1.

After the exponential gain regime in a tapered FEL, the imaginary part of n, Im(n) measuring the gain, is generally close to 0, while the real part, Re(n) describing the refractive guiding, dominates the FEL process. In this region, $|a_s|$ increases much less rapidly with z than is true in the upstream exponential gain regime; there is also an increase in r_s. Moreover, the refractive guiding decreases with increasing $|a_s|$, which in turn leads to slowing down or even stopping of the $|a_s|$ growth [13,14]. This phenomenon can strongly reduce the radiation peak intensity growth rate

in z as compared to what would be predicted by a purely 1D approach that presumes a constant r_s. The optical guiding approach (e.g., [12]) includes the transverse effects self-consistently and thus provides more precise predictions for the behavior of the electron and radiation beams.

In Sec. II A, we include the transverse effects revealed by the optical guiding approach in the framework of the 1D KMR theory and formulate a physical model explicitly with several principal assumptions, such as *ray-equation*, *monoenergy*, *resonant-phase*, and *Gaussian-profile approximations* (see below for specifications). In the presented model, the electron and radiation beam profiles remain approximately Gaussian, and particles travel straight-line orbits at constant radius. Only trapped electrons are presumed to contribute to the microbunching term $\langle \exp(-i\Psi) \rangle$ in Eq. (1) and, furthermore, with a value equal to $\exp[-i\Psi_r(r)]$ at each radial position r. Accordingly as we show in Sec. II A, the term $\langle \exp(-i\Psi) \rangle$ can be calculated directly by averaging the product $F_t(r) \exp[-i\Psi_r(r)]$ over r, where $F_t(r)$ and $\Psi_r(r)$ are the radially resolved trapping fraction and synchronous phase. These assumptions result in a more explicit formulation than the traditional, nonradially resolved optical guiding approach, and simultaneously a far more simple and more efficient computational approach than the brute-force approach of traditional 3D numerical simulation.

In Sec. II B, we compare the predictions of the model with those of the GENESIS single-frequency simulation for a 120-m, hard x-ray, tapered FEL and obtain good agreement. Based on the model, in Sec. II C we present a general physical picture of the FEL process in a tapered FEL, and divide the tapered undulator region, beginning from the *initial saturation* location to the end of the undulator, into three successive regions according to the behavior of the electron and radiation beams. We show that the unavoidable weakening of refractive guiding causes a decreasing $F_t(z)$, and finally leads to saturation of the radiation power. To improve the overall energy extraction efficiency requires sophisticated control of the decrease of refractive guiding and particle detrapping along the undulator. For this purpose, we propose a multidimensional optimization scheme in Sec. III, and illustrate its applications to hard x-ray tapered FELs. We discuss the difference between the proposed optimization scheme and GINGER code's KMR-based self-design taper algorithm [15]. We also compare the optimization results for a tapered FEL with and without break sections. In addition, we qualitatively summarize the dependence of the available maximum radiation power on various parameters of the initial electron beam, radiation field, and the undulator system, with only the effect of decreasing refractive guiding taken into account. In Sec. IV, we introduce another important phenomenon that leads to premature saturation of the radiation power in a tapered FEL, i.e., the sidebands excited by the SASE components that originate from the shot noise on the

electron beam. We show that sideband growth in a tapered FEL eventually causes rapid particle detrapping and decreased refractive guiding that together significantly reduce the overall energy extraction efficiency. Our conclusions are given in Sec. V.

To aid the reader and to keep the flow of the text clear, we first specify several terminologies and assumptions that appear frequently in this paper.

Initial saturation point.—This location is the end of the exponential gain regime in a single-pass high-gain FEL with a constant-parameter undulator. At this location, $P_{\text{sat}} \approx \rho P_{\text{beam}}$, and the fractional average energy loss and energy spread of the electron beam are of the same order as the FEL parameter ρ.

Tapered FEL.—While empirically the best taper start point is usually slightly before the *initial saturation* location, to simplify the analytical study, it is assumed that the undulator following a self-seeding monochromator is tapered from the *initial saturation* location. Both the undulator field a_w and undulator period λ_w can be varied. However, it is more economical and operationally simpler to vary only a_w than to also change λ_w for an actual undulator. Thus, in this paper, our tapers have a varied a_w but constant λ_w.

Ray-equation approximation.—It is assumed that the electrons' trajectories are straight lines and parallel to the undulator axis; i.e., we neglect radial betatron motion.

Monoenergy approximation.—It is assumed that those electrons trapped in the ponderomotive bucket have a radially independent, resonant energy $\gamma_r\text{mc}^2$ [γ_r is defined by Eq. (10) below]; i.e., we neglect the energy spread of the trapped electrons, which, from a wide range of simulations, is found to remain small and to be of the same order as the FEL parameter ρ.

Resonant-phase approximation.—It is assumed that the trapped electrons are uniformly distributed in ponderomotive phase at each radial position r, and that their contribution to the microbunching term is $\exp[-i\Psi_r(r)]$. By contrast, the contribution of the detrapped particles to the microbunching term is presumed to be zero.

Gaussian-profile approximation.—It is assumed that both the radiation and electron beam transverse profiles approximately follow Gaussian distributions throughout the undulator; i.e., we neglect higher order modes in the radiation field that can develop when r_s starts to significantly exceed r_b.

II. PHYSICAL MODEL OF A TAPERED FEL

A. Formulation

Generally, before and up to the *initial saturation* location, the fundamental transverse Gaussian mode dominates the radiation field. We find this behavior applies also in the tapered region of the undulator [4]. Under the standard eikonal approximation, the normalized vector potential of the radiation field a_s is

$$a_s(r, z) = a_{s0}(z)e^{i\phi(r,z)}e^{(-r^2)/[r_s^2(z)]}, \quad (2)$$

where a_{s0} and ϕ are the slowly varying on-axis amplitude and phase of the radiation field, respectively. Accordingly, the radiation power P (in unit of watts) is given, after integrating over the radial coordinate, by

$$P(z) = \frac{\pi r_s^2(z)a_{s0}^2}{4Z_0}\left(\frac{k_s\text{mc}^2}{e}\right)^2, \quad (3)$$

with $Z_0 \approx 376.7 \ \Omega$ being the impedance of free space. Similarly, we presume the initial electron beam profile obeys Gaussian distribution,

$$f_0(r) = \frac{N_e}{\pi r_{b0}^2}e^{r^2/r_{b0}^2}, \quad (4)$$

where $N_e = \int f_0(r)2\pi rdr$ is the initial electron population.

Both a_{s0} and r_s will vary with z in a tapered FEL [12,13]. From energy conservation, the changes in $a_{s0}(z)$ and $r_s(z)$ in a small longitudinal distance Δz follow

$$a_{s0}^2(z + \Delta z)r_s^2(z + \Delta z) - a_{s0}^2(z)r_s^2(z)$$
$$= \frac{\omega_{p0}^2}{\omega_s^2}r_{b0}^2F_t(z)[\langle\gamma(z)\rangle - \langle\gamma(z + \Delta z)\rangle], \quad (5)$$

where we make the assumption that the variation of F_t is much slower than that of a_{s0}, which is usually valid in the tapered region of the undulator. We take $\Delta r_s(z) = r_s(z + \Delta z) - r_s(z)$, $\Delta a_{s0}(z) = a_{s0}(z + \Delta z) - a_{s0}(z)$ with $|\Delta r_s| \ll r_s$ and $|\Delta a_{s0}| \ll a_{s0}$, and set $\langle\gamma(z)\rangle = \gamma_r(z)$ under the *monoenergy approximation*. Then we can rewrite Eq. (5) as

$$2a_{s0}^2(z_1)r_s(z_1)\Delta r_s(z) + 2r_s^2(z)a_{s0}(z)\Delta a_{s0}(z)$$
$$\approx -\frac{\omega_{p0}^2}{\omega_s^2}r_{b0}^2F_t(z)\gamma_r'(z)\Delta z, \quad (6)$$

where the prime refers to the z derivative. In Eq. (6), the first and second terms on the left-hand side are proportional to the radiation power increment due to radial expansion and growth in on-axis field, respectively. In the following discussion, we denote them as $\Delta E_r(\text{rad})$ and $\Delta E_a(\text{rad})$. Similarly we represent the term on the right-hand side as $\Delta E(e-)$.

Under aforementioned *ray-equation*, *monoenergy*, and *Gaussian-profile approximations*, the evolution of r_s follows the envelope equation [12],

$$r_s'' + K^2r_s = 0, \quad (7)$$

where the optical focusing parameter K^2 can be written in terms of the refractive guiding "fiber parameter" $V^2 = (n^2 - 1)k_s^2r_b^2$ with on-axis $|a_s|$, the average of sine and cosine of Ψ, and $G(z) = (1 - f)/(1 + f)^2$ with the filling factor $f(z) = (r_b/r_s)^2$ [12],

$$K^2 = \frac{4}{k_s^2}\left(-1 + V^2G + \frac{1}{4}V^4G^2\frac{\langle\sin\Psi\rangle^2}{\langle\cos\Psi\rangle^2}\right.$$
$$\left. + \frac{1}{4}k_sr_s^2\frac{d(V^2G/\langle\cos\Psi\rangle)}{dz}\langle\sin\Psi\rangle\right)r_s^{-4}. \quad (8)$$

As mentioned, after the *initial saturation* in a tapered FEL, $\mathrm{Re}(n) \approx 1$ and $\mathrm{Im}(n) \ll 1$. Thus, in this region, V^2 can be written as [11,13]

$$V^2(z) \approx 2[\mathrm{Re}(n)-1]k_s^2r_b^2 \approx \frac{\omega_{p0}^2r_{b0}^2}{c^2}\frac{a_w}{a_{s0}}[JJ]\frac{\langle\cos\Psi\rangle}{\gamma_r}. \quad (9)$$

In Eqs. (8) and (9), $\langle\sin\Psi\rangle$ and $\langle\cos\Psi\rangle$ are approximated by a radial average over only the trapped electrons [see Eqs. (16) and (17)]. In Eq. (9), we also make the *mono-energy approximation* and assume that all the trapped electrons lie exactly at the resonant energy $\gamma_r\mathrm{mc}^2$, with γ_r in the form [5]

$$\gamma_r^2(z) = \frac{k_s}{2k_w(z)}[1 + a_w^2(z)]. \quad (10)$$

In order to close all the above equations, it is necessary to determine both the trapping fraction and the average ponderomotive phase as the electron beam propagates through the tapered region of the undulator. We note that the ponderomotive potential is proportional to $|a_s|$ [5]. With our assumption of Gaussian profile for the radiation field [see Eq. (2)], the ponderomotive potential becomes smaller at larger r, and thus a smaller fraction of electrons will be trapped in the outer edges of the electron beam. As a result, the transverse distribution of the trapped electrons will deviate some from a Gaussian profile.

To determine the radial distribution of the trapped electrons, we follow Ref. [5] to formulate the r-dependent synchronous phase Ψ_r in the form

$$\sin[\Psi_r(r,z)] = \frac{-\gamma_r(z)\gamma_r'(z)}{a_w(z)k_s|a_s(r,z)|}. \quad (11)$$

Adopting the aforementioned *resonant-phase approximation*, a given electron stays trapped if and only if its ponderomotive phase is within the closed orbit portion of the (r, z)-dependent bucket, and will detrap instantly once its phase falls outside this (generally shrinking) border. Thus, the local trapping fraction is given by

$$F_t(r,z) = \frac{\Psi_2(r,z) - \Psi_1(r,z)}{2\pi}, \quad (12)$$

where Ψ_1 and Ψ_2 are the minimum and maximum Ψ of the ponderomotive bucket,

$$\Psi_2(r,z) = \pi - \Psi_r(r,z),$$
$$\cos\Psi_1(r,z) + \Psi_1(r,z)\sin\Psi_r(r,z) \quad (13)$$
$$= \cos\Psi_2(r,z) + \Psi_2(r,z)\sin\Psi_r(r,z).$$

Note that $F_t(r,z) = 0$ when $\Psi_r(r,z) = \pi/2$.

From combining Eqs. (2), (11), and (12), one can see there is a maximum r beyond which there can be no trapped electrons, whose value is given by

$$r_{\max}(z) = r_s(z)\{-\ln[\sin\Psi_r(r=0,z)]\}^{1/2}$$
$$= r_s(z)\left\{-\ln\left[\frac{-\gamma(z)\gamma'(z)}{a_w(z)k_sa_{s0}(z)}\right]\right\}^{1/2}. \quad (14)$$

Locally, the number of the trapped electrons is $F_t(r,z)f_0(r)$, the overall electron trapping fraction $F_t(z)$ is then obtained by averaging $F_t(r,z)$ over all r from 0 to $r_{\max}$,

$$F_t(z) = \frac{1}{N_e}\int_0^{r_{\max}}F_t(r,z)f_0(r)2\pi rdr. \quad (15)$$

Similarly, $\langle\cos\Psi\rangle$ and $\langle\sin\Psi\rangle$ for the trapped electrons are given by

$$\langle\cos\Psi\rangle(z) = \frac{1}{N_e}\int_0^{r_{\max}}\cos\Psi_r(r,z)F_t(r,z)f_0(r)2\pi rdr, \quad (16)$$

$$\langle\sin\Psi\rangle(z) = \frac{1}{N_e}\int_0^{r_{\max}}\sin\Psi_r(r,z)F_t(r,z)f_0(r)2\pi rdr. \quad (17)$$

As mentioned above, we assume the undulator taper begins from the *initial saturation* location. Thus, starting from the electron and radiation beam parameters at *initial saturation*, especially ($a_{s0,\mathrm{sat}}$, $r_{s,\mathrm{sat}}$, $\Psi_{r,\mathrm{sat}}$, $F_{t,\mathrm{sat}}$), one can iterate Eqs. (2) through (17) to evolve these parameters in z along the undulator for a specific taper profile, without requiring the 6D phase space details of the electron and radiation beam distribution from a numerical simulation code.

At *initial saturation*, according to the relation $P_{\mathrm{sat}} \approx \rho P_{\mathrm{beam}}$, the saturation length is estimated by [16]

$$L_{\mathrm{sat}} \approx \frac{\lambda_w}{4\pi\sqrt{3}\rho}\ln\left(\frac{9P_{\mathrm{sat}}}{P_{\mathrm{in}}}\right), \quad (18)$$

with P_{in} being the input radiation power. For the case of self-seeding, this length would be measured from the point where the electron beam rejoins the monochromatized radiation beam in the undulator. The normalized vector potential of the radiation field $a_{s0,\mathrm{sat}}$ in a helical undulator can be evaluated by [14]

$$a_{s0,\mathrm{sat}} \approx \frac{2(1 + a_{w0}^2)\rho^2}{a_{w0}[JJ]}, \quad (19)$$

adopting the assumption of $r_{s,\mathrm{sat}} = r_{b0}$; an additional $2^{1/2}$ factor in the denominator is necessary when considering a linearly polarized undulator. Since the undulator parameters are assumed to be constant before *initial saturation*, we have $\gamma_{r,\mathrm{sat}} = \gamma_0$, $\gamma'_{r,\mathrm{sat}} = 0$, $\Psi_{r,\mathrm{sat}} = 0$, and $F_{t,\mathrm{sat}} = 1$.

Our model has one other *ad hoc* assumption that we use for the undulator region just past *initial saturation* where in actual numerical simulation it is difficult to differentiate
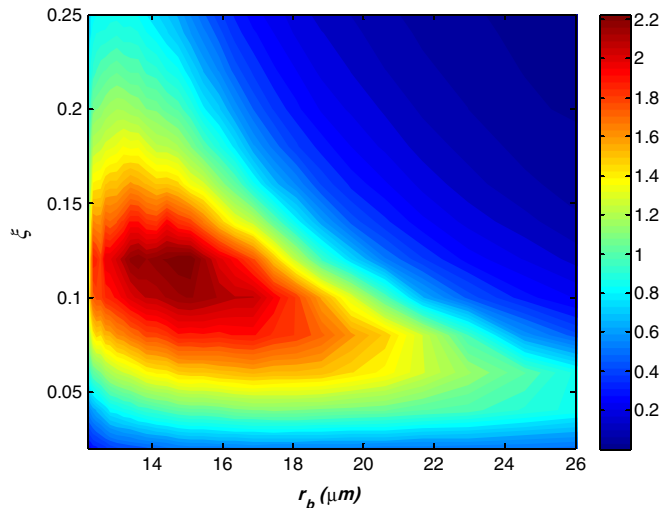
FIG. 1. Contour plot of the output radiation power for an 120-m, hard x-ray, tapered FEL with respect to the taper ratio ξ and electron beam radius r_b, obtained by GENESIS single-frequency simulations. The taper starts at $z_0 = 12$ m. The scale of the color bar is in TW.

TABLE I. Main parameters for a hard x-ray tapered FEL with linearly polarized undulator.

Parameters	Value	Unit
E-beam energy	13.64	GeV
E-beam current	4000	Ampere
Normalized emittances $\varepsilon_{x,n}/\varepsilon_{y,n}$	0.3/0.3	μm rad
Energy spread	1.3	MeV
E-beam pulse length (FWHM)	10	fs
Normalized undulator parameter a_{w0}	2.3832	
Undulator period λ_w	32	mm
Undulator length L_w	120	m
Radiation wavelength λ_s	1.5	Angstrom
Peak radiation input power P_{in}	5	MW

particles which will or will not eventually be trapped in the ponderomotive bucket (e.g., see Fig. 4 below). Namely, for a distance two Rayleigh lengths long where the Rayleigh length is defined by $z_r \approx k_s r_{s,\text{sat}}^2/2 \approx k_s r_b^2/2$, we presume F_t remains exactly equal to one and in effect we delay applying Eqs. (11) through (17). In this region, the electron energy loss and radiation gain are calculated from the taper alone given $F_t = 1$. We justify this assumption only empirically in that it appears to work well for both hard and soft x-ray FEL cases.

It is worth mentioning that the presented model is based on several principal assumptions, such as *ray-equation*, *monoenergy*, *resonant-phase*, and *Gaussian-profile approximations*, rather than being derived strictly from Maxwell equations. It well describes the evolution of the radiation field within r_{max}, which is given by Eq. (14) and more or less $3r_{b0}$, while not considering the radiation propagating outside this region. The principal *ray-equation approximation* implies that the model applies only to the cases where the betatron wavelength is much longer than the synchrotron wavelength of the trapped electrons, an approximation that is well satisfied for the current LCLS-II design [4].

B. Verification of the physical model

To verify the accuracy of the above physical model, we compare its predictions with the results from GENESIS simulations for an 120-m tapered FEL with LCLS-II like radiation seed, electron beam, and undulator parameters [4]. The main parameters are shown in Table I. With a preset quadratic taper started from a somewhat arbitrarily chosen location of $z_0 = 12$ m, we scan the taper ratio $\xi = 1 - a_w(L_w)/a_w(z_0)$ with L_w being the undulator

length and the electron beam radius r_b (constant throughout the undulator) by use of the GENESIS single-frequency simulations, with the results shown in Fig. 1. The minimum scanned r_b is 12 μm, determined by the available maximum field strength of LCLS-II quadrupole magnets. The parameter-scanning results show that the output radiation power depends upon both the taper profile and transverse focusing. Generally, achieving high radiation power requires strong transverse focusing, especially in a case with a relatively large taper ratio.

From these GENESIS simulations, we obtain a maximum power of 2.2 TW for an optimal case in which $\xi = 0.12$ and $r_b = 15$ μm; for convenience, we name it as "Case A." Figure 2 presents the Case A evolution of the radiation power within $r = 3r_{b0}$, radius r_s and on-axis $|a_s|$ by fitting a Gaussian to the radiation field data, and the electron trapping fraction F_t in solid lines.

We note that the rms relative energy spread σ_γ/γ of the trapped electrons remains close to $\rho = 0.0011$, a value much less than the fractional energy reduction of the trapped electrons over the undulator $\Delta\gamma/\gamma_0 \approx 0.1$. We also note that the betatron oscillation period, $L_\beta \approx \pi\gamma r_b^2/\varepsilon_{x,n} \approx 60$ m, is 5 times greater than the synchrotron oscillation period [5] $L_s \approx \lambda_w[(1 + a_w^2)/|a_{s0}|/a_w]^{1/2}/2 \approx 12$ m. As a result, the *monoenergy* and *ray-equation approximations* used in the model are approximately satisfied for Case A.

To begin our physical model calculations for Case A, we use Eqs. (18) and (19) to obtain $L_{\text{sat}} = 16.1$ m and $a_{s0,\text{sat}} \approx 6.4 \times 10^{-6}$, which agree fairly well with the GENESIS simulation results, $L_{\text{sat}} \approx 20$ m and $a_{s0,\text{sat}} \approx 5.1 \times 10^{-6}$. We then advance Eqs. (2) through (17) adopting a z-step size of $\lambda_w = 3.2$ cm (while keeping constant $\Psi_r = 0$ and $F_t = 1$ in the first two Rayleigh lengths after *initial saturation*) and generate the evolution of (P, a_s, r_s, F_t) along the undulator. The model's predictions are plotted as dashed lines in Fig. 2. Considering the assumptions used in the formulation, the agreement between the model and the GENESIS single-frequency simulation is reasonably good.
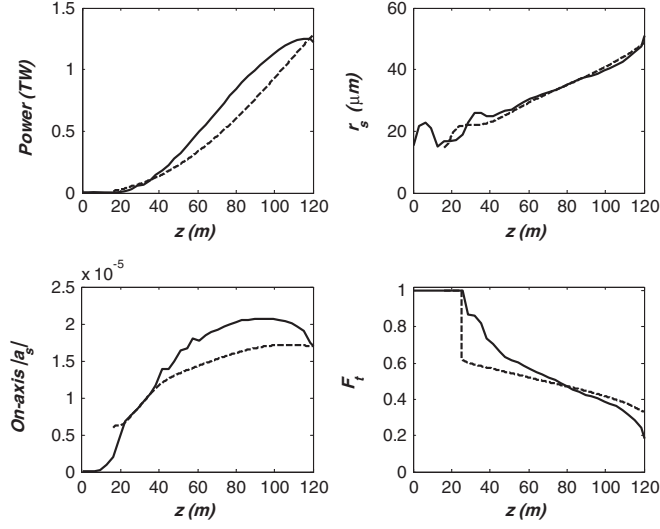
FIG. 2.   Evolution with z of the radiation power within $r = 3r_{b0}$, radius r_s and on-axis $|a_s|$ (obtained by fitting a Gaussian to the radiation field data), and radially averaged trapping fraction F_t for Case A, obtained by GENESIS single-frequency simulation (solid lines) and the physical model (dashed lines).

The success of the model is also reflected in the transverse distribution of the trapped electron number and radiation field, as shown in Fig. 3. Most importantly, the figure also shows that the variation with radius of the trapped electron number and their average ponderomotive phase $\langle \Psi \rangle = \langle \theta + \phi \rangle$ (θ is the electron phase relative to a plane wave) obtained from GENESIS shows the same behavior as that revealed by our model, i.e., electrons at larger r have greater Ψ_r and detrap more rapidly.
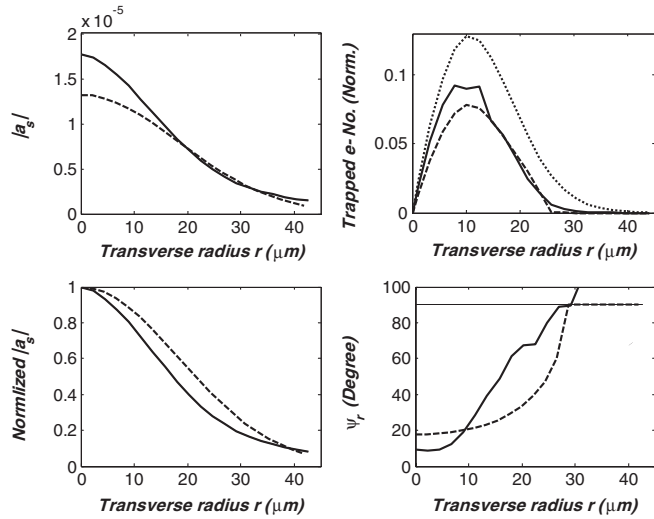

FIG. 3.   Transverse distribution of the radiation field $|a_s|$, normalized trapped electron number per radial bin, and synchronous phase Ψ_r at $z = 50$ m for Case A, obtained by GENESIS single-frequency simulation (solid lines) and the physical model (dashed lines). The dotted line in the top right plot indicates the initial normalized electron number (without detrapping).

C. Saturation mechanism in a tapered FEL

Based on the verification of the presented physical model, we give a more general description for the evolution of the electron-radiation interaction in a tapered FEL, and show that the decreasing of refractive guiding and trapping fraction is the major cause of the saturation of the radiation intensity and power in a tapered FEL. According to the behavior of the electron and radiation beams, we divide the tapered undulator region, from the *initial saturation* location to the end, into three successive regions.

1. Particle trapping development region

With our aforementioned assumptions, this region starts from the *initial saturation* location and lasts about two Rayleigh lengths.

At the beginning of this region, because we have assumed that the upstream gain guiding leads to $r_s \approx r_b$, the factor G in Eq. (8) is close to 0. Although the refractive guiding "fiber parameter" V^2 is considerably larger than 1 (e.g., ~ 10 for Case A), with $G \approx 0$ the -1 term dominates the expression for K^2 [see Eq. (8)], leading to strong diffraction of the radiation and an increase in r_s. From Eq. (6), the increase in r_s corresponds to a large $\Delta E_r(\text{rad})$, and therefore a large portion of the energy extracted from the electron beam contributes to the radiation expansion in the radial direction. However, as r_s increases, G also increases. By the end of this region, G has grown large enough that the V^2G term in Eq. (8) becomes comparable to or even larger than 1, and the optical focusing effects now become strong. At this point the r_s growth becomes much slower and as the radiation power increases, a_{s0} will also begin to increase strongly.

From Eqs. (10) and (11), we have

$$a_{s0}(z)\sin[\Psi_r(r=0,z)] = \frac{|a'_w|}{2k_w}. \qquad (20)$$

To vary a_w smoothly, one usually gradually increases $|a'_w|$ from 0 after the taper start point. Corresponding to the increased $|a'_w|$ and more or less constant a_{s0}, the ponderomotive bucket evolves with the on-axis Ψ_r increasing from 0 to a positive value and with F_t decreasing markedly. Accordingly, the electrons gradually bunch themselves to match the changed ponderomotive bucket, during which evident detrapping occurs. In order to avoid large on-axis Ψ_r and significant detrapping, it usually calls for a gentle taper with small $|a'_w|$ in this region.

Taking Case A as an example, we show in Fig. 4 the evolution of the electron beam distribution in the first two Rayleigh lengths after *initial saturation*. At the end of this region, the trapped electrons have a visible reduction in energy and are also well separated in energy from the detrapped ones; the measured F_t decreases by about 17%.
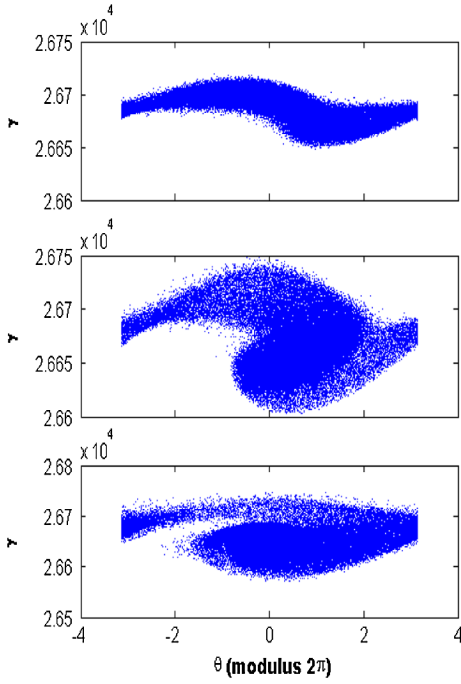
FIG. 4. The electron beam distribution in the phase space of (θ, γ) at the *initial saturation* location (top), and a distance of $z_r = 4.7$ m (middle) and $2z_r = 9.4$ m (bottom) after *initial saturation* for Case A, obtained by GENESIS single-frequency simulation.

2. Radiation intensity growth region

This region follows the *particle trapping development region* and ends where a_{s0} reaches a maximum value. Since G has been close to 1, the $(-1 + V^2 G) \approx (-1 + V^2)$ term in the expression for K^2 dominates provided that the on-axis Ψ_r is small enough, e.g. $\Psi_r(r = 0) < 0.4$.

At the beginning of this region, V^2 is well above 1. Thus, the optical focusing is relatively strong, leading to slower increase in r_s compared to the upstream region. According to Eq. (5), a greater proportion of the energy extracted from the electron beam will contribute to the growth in a_{s0}, resulting in a higher a_{s0} growth rate. Meanwhile, from Eq. (9), the increase in a_{s0} causes a reduction in V^2, and hence a decrease in K^2, i.e., reduced optical focusing. As a result, with the increasing a_{s0}, r_s now increases more rapidly, and the a_{s0} growth gradually slows down, in general eventually reaching an asymptotic value [14].

At the location where $\Delta E_r(\text{rad}) = \Delta E(e-)$, a_{s0} reaches its maximum and decreases thereafter. We denote this location as $z = L_{a,\max}$. A somewhat empirical criterion for the a_{s0} saturation is $V^2 \approx 1$ [11,14] with $K^2 \approx 0$, which results in the maximum a_{s0} in the form

$$a_{s0,\max} \approx \frac{\omega_{p0}^2 r_{b0}^2}{c^2} \frac{a_{w0}[JJ]}{\gamma_0} \langle \cos\Psi \rangle, \qquad (21)$$

where we have used the relation $a_w(L_{a,\max})/\gamma_r(L_{a,\max}) \approx a_{w0}/\gamma_0$ in the case of $a_w^2 \gg 1$. Note that the a_{s0} at the
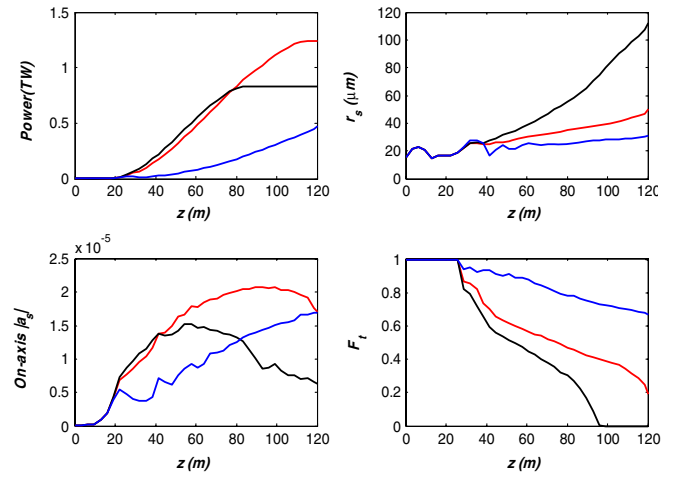

FIG. 5. Evolution with z of the power (within $3r_b$), on-axis $|a_s|$ and r_s (obtained by fitting a Gaussian to the radiation field data), and F_t for the demonstrative tapered FEL in Sec. II B with $r_b = 15$ μm and with quadratic tapers of taper ratio $\xi = 0.02$ (blue lines), 0.12 (red lines), and 0.16 (black lines). The results are obtained by GENESIS single-frequency simulations.

beginning of this region is close to $a_{s0,\text{sat}}$. Thus, the length of this *radiation intensity growth region* can be estimated by $L_{\text{rigr}} \approx (a_{s0,\max} - a_{s0,\text{sat}})/\langle a'_{s0} \rangle$, with $\langle a'_{s0} \rangle$ being the average a_{s0} growth rate in this region.

From Eqs. (19) and (21), the values of $a_{s0,\text{sat}}$ and $a_{s0,\max}$ are mainly determined by the parameters of the initial electron beam, radiation field, and the undulator system. In addition, $a_{s0,\max}$ is also related to the taper profile by the F_t and $\langle \cos\Psi \rangle$ at $z = L_{a,\max}$. Generally speaking, a gentle taper with small $|a'_w|$ will lead to a large F_t and $\langle \cos\Psi \rangle$ and thus a large value of $a_{s0,\max}$ at $z = L_{a,\max}$. However, such a taper will also lead to a small γ'_r and $\langle a'_{s0} \rangle$ and, hence, a long L_{rigr}. Figure 5 shows the evolutions of the power, a_{s0}, r_s, and F_t for the demonstrative tapered FEL in Sec. II B with $r_b = 15$ μm and with three quadratic tapers of different taper ratio. One sees that, when tapering a finite-length undulator with too small an $|a'_w|$ (the $\xi = 0.02$ case), a_{s0} does not reach the predicted $a_{s0,\max}$ even at the end of the undulator. On the other hand, if the undulator taper employs too large a value for $|a'_w|$ (the $\xi = 0.016$ case), a_{s0} reaches a small value of $a_{s0,\max}$ within a short undulator length and then begins to decrease. This reduction also causes rapid detrapping. The radiation power reaches saturation when total detrapping occurs. Both cases produce low radiation power. As a result, a moderate taper is required to achieve both a relatively large $a_{s0,\max}$ and a high radiation power (the $\xi = 0.12$ case).

3. Radiation power growth region

In a tapered FEL with a long enough undulator that a_{s0} reaches its maximum before the exit of the undulator, i.e., $L_w > L_{a,\max}$, there will be a region following the *radiation*

intensity growth region, where the on-axis intensity does not increase but the total radiation power still grows; we name this region the *radiation power growth region*. Note that in the case with $L_w \lesssim L_{a,\max}$, there will not be such a region.

In this power growth region, the refractive guiding is rather weak with $V^2 < 1$. The -1 term in the expression for K^2 dominates again. Radiation diffraction is now important again, leading to a more rapid increase in r_s than is true in the upstream region. This leads to $\Delta E_r(\mathrm{rad})$ larger than $\Delta E(e-)$ and to a decreasing a_{s0} (in some cases, the a_{s0} decrease rate can be small). In spite of the reduction in a_{s0}, the radiation power will continue to increase as long as there exist electrons trapped in the ponderomotive bucket. However, associated with the weakened refractive guiding, the energy extraction becomes less and less efficient, independent of the actual taper profile. In the following, we will show that, in an idealized tapered FEL with arbitrarily long undulator and with three kinds of taper profiles, the radiation power will reach saturation due to either the combination of a decreasing deceleration rate together with detrapping (Cases 1 and 2) or total detrapping (Case 3).

Case 1.—For the typical taper obtained using the KMR-based self-design taper algorithm in which a design particle at a preset radius remains at a constant Ψ_r, F_t remains relatively constant along the undulator. In this case, a_{s0} decreases slowly with z in this *radiation power growth region*. From Eqs. (10) and (20), this design style comes with a slowly decreasing $|\gamma_r'|$ and $|a_w'|$ with z; it also implies that one is not necessarily concerned with minimizing the total undulator length. However, there is always a small amount of detrapping (especially when betatron motion effects are included) so that eventually the product of $F_t|\gamma_r'|$ in the expression for $\Delta E(e-)$ becomes very small and the output power essentially saturates.

Case 2.—For a taper with smaller $|a_w'|$ compared to that of a constant-Ψ_r taper, the overall trapping fraction is generally higher at any given z but the energy extraction rate is slower. Thus, this in general is not an attractive option. Detrapping will still occur to some extent so that one expects an eventual saturation in radiation power, albeit at a much greater undulator length.

Case 3.—A taper with larger $|a_w'|$ (e.g., a parabolic taper with z) and thus larger on-axis Ψ_r compared to a constant-Ψ_r taper produces a larger $\Delta E(e-)$ and relatively slower decrease in a_{s0} within a short distance from the beginning of this region. However, this is accompanied by a more rapid detrapping in this power growth region. For a monotonically increasing $|a_w'|$, $\Psi_r(r = 0)$ eventually increases to $\pi/2$, total detrapping takes place, and the radiation power reaches saturation in the smallest undulator length among three cases. However, for the optimization of the radiation power at a given z (i.e., for a fixed-length undulator), such a taper may often be more aggressive than

a constant-Ψ_r taper, especially in the latter part of the undulator.

Since the optical focusing and, hence, the evolution of the radiation beam are related to both the longitudinal dynamics (through $|\gamma_r'|$ and Ψ_r) and the transverse dimension parameters (through G), it is worthwhile to consider if allowing a controlled variation in the electron beam radius can be helpful in improving the overall energy extraction efficiency. We turn to this question in the next section.

III. MULTIDIMENSIONAL OPTIMIZATION OF A TAPERED FEL

A. Multidimensional optimization scheme

For a physically realistic scenario, we need to optimize the FEL radiation power within a finite undulator length. From the analytical studies in Sec. II, optical guiding physics limits the on-axis radiation intensity and the overall energy extraction efficiency. Generally, to maximize the overall energy extraction efficiency usually requires a moderate taper within a long enough undulator length, let us say $L_w \gtrsim L_{\mathrm{sat}} + 4z_r$, which results in a relatively large $a_{s0,\max}$ before and close to the exit of the undulator. In addition, it is interesting to see if it is possible to enhance the refractive guiding in the tapered region of the undulator by optimizing the electron beam radius. Our physical model can be used to investigate the possible consequences.

Let us consider a gradually decreased r_b in the latter part of a tapered undulator. Compared to the case with the same taper profile but a constant r_b, the factor G and, hence, the optical focusing parameter K^2 will be slightly larger, leading to increased optical focusing and to smaller r_s as well as smaller $\Delta E_r(\mathrm{rad})$. For a given taper profile, Eq. (6) predicts a larger a_{s0}, resulting in smaller $\Psi_r(r = 0)$ and higher $F_t(r = 0)$ [see Eqs. (11) and (12)]. On the other hand, a smaller r_s leads to a smaller $r_{\max}$ [see Eq. (14)], causing a higher $\Psi_r(r)$ and smaller $F_t(r)$ for electrons at large r. Thus, the detrapping of the electrons at large r will be more rapid, while the electrons around the axis will detrap less rapidly. The stronger on-axis optical guiding on one hand will tend to lead to a more rapidly growing a_{s0}; on the other hand, a smaller r_s will increase diffractive effects. If one squeezes r_b to too small a value such that the diffraction effect dominates, the energy extracted from the electron beam will contribute to rapid radial expansion of the radiation rather than growth in a_{s0}. Thus, one expects there is an optimal value for a decreased r_b. We note that, since G has been close to 1 in the latter part of the tapered undulator (e.g., $r_s \sim 2$–$3r_b$ for Case A), the expected change in G and overall optical guiding due to r_b variation is rather small. Thus, we expect only a relatively small improvement from an r_b variation compared with what is attainable from optimizing the taper profile in z at fixed r_b.

To illustrate the optimization scheme specifically, we formulate the taper profile as [17]

$$a_w(z) = a_w(z_0) \times [1 - c \times (z - z_0)^d], \quad (22)$$

where z_0 indicates the taper start point, d is the taper profile order, and c is the scale coefficient which is related to the taper ratio ξ by $c = \xi/(L_w - z_0)^d$.

From Eq. (20), we can obtain

$$a_{s0}(z) \sin[\Psi_r(r = 0, z)] = \frac{a_w(z_0)}{2k_w} \frac{d(z - z_0)^{d-1}}{(L_w - z_0)^d} \xi. \quad (23)$$

Achieving an increasing $a_{s0}(z)$ requires d greater than 1. Empirically we find that it is best to start the taper slightly before *initial saturation* and use a moderate taper profile order with $d \approx 2$. The taper ratio ξ is closely related to the energy reduction of the trapped electrons and, hence, to the gain of the radiation field. The optimal ξ for the maximum radiation power varies with the undulator length and various initial electron and radiation beam parameters.

Our particular formulation of a variation in electron beam radius is as follows. We introduce a three-segment r_b variation by linearly changing the quadrupole field strength with z,

$$K_q(z) = \begin{cases} K_q(z_1), & \text{with } 0 < z \le z_1 \\ K_q(z_1) \times [1 - f \times (z - z_1)], & \text{with } z_1 < z \le z_2, \\ K_q(z_2) \times [1 - g \times (z - z_2)], & \text{with } z_2 < z \le L_w, \end{cases}$$
$$(24)$$

where $K_q(z)$ is the quadrupole field strength; z_1 indicates the first K_q-variation start point, which is usually around

the *initial saturation* location; z_2 indicates the second K_q-variation start point, which is usually around the location $z = L_{a,\max}$; f and g are related to $K_q(z_1)$, $K_q(z_2)$, and $K_q(L_w)$ by $K_q(z_2) = K_q(z_1) \times [1 - f \times (z_2 - z_1)]$ and $K_q(L_w) = K_q(z_2) \times [1 - g \times (L_w - z_2)]$; f can be either positive or negative, while g is usually negative. In the thin-lens approximation, the average beta function $\beta_{\text{aver}} \propto 1/K_q(z)$ for periodic transverse focusing lattice cells. According to the relation $r_b = (2\varepsilon_n \beta_{\text{aver}}/\gamma)^{1/2}$, a linearly decreasing K_q approximately corresponds to r_b increasing with square root of z, and vice versa.

B. Applications of the multidimensional optimization

With the above specification for functional dependences of the a_w taper and transverse focusing, we obtain the maximum radiation power for a tapered FEL with specific electron beam, radiation seed, and undulator properties by performing multidimensional scans with GENESIS single-frequency simulations over the following eight parameters: z_0, d, ξ, z_1, z_2, $K_q(z_1)$, $K_q(z_2)$, and $K_q(L_w)$. In the following, we present the optimization results for hard x-ray tapered FELs with and without break sections.

1. A 200-m, hard x-ray, tapered FEL without break sections

We first examine the case of a hard x-ray, tapered FEL with the same radiation seed, electron beam, and undulator parameters as those listed in Table I except that the undulator length is now 200 m. For a situation with fixed electron beam radius, we found maximum output with
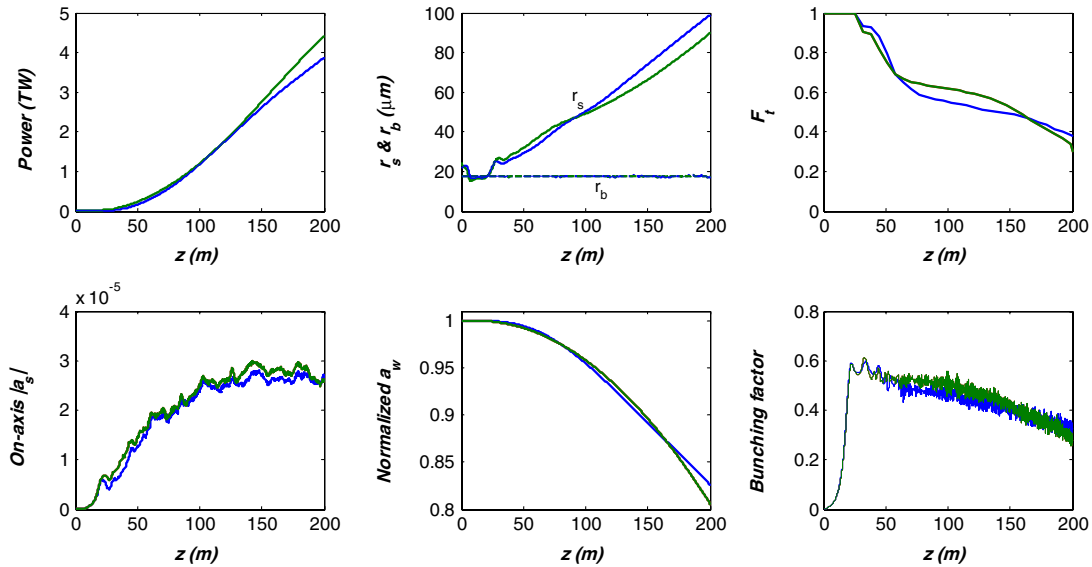

FIG. 6.   GENESIS single-frequency simulations for the evolution of the radiation power, on-axis $|a_s|$, r_s, F_t, and bunching factor in a 200-m, hard x-ray, tapered FEL with constant r_b, and with taper profiles obtained from GINGER code's KMR-based self-design taper algorithm (blue lines) and multidimensional optimization (green lines).

an optimal taper profile with $z_0 = 11.5$ m, $d = 2.04$, $\xi = 0.19$, and optimal r_b of 17.5 μm. With these parameters, Fig. 6 (green lines) shows the evolution of radiation power, on-axis $|a_s|$, r_s, F_t, and bunching factor. The variation of the bunching factor is similar to that of F_t in the deep-tapered region of the undulator. For comparison, we also present a GENESIS run that employs an a_w taper profile obtained with GINGER code's KMR-based self-design taper algorithm [15] with $\Psi_r = 0.4$ at $r \approx r_b$ (blue lines).

Compared to the proposed multidimensional optimization scheme, the constant-Ψ_r self-design algorithm results in slightly more rapid taper in the *radiation intensity growth region*, a bit smaller value of $a_{s0,\max}$, and then a less rapid taper in the *radiation power growth region*. Taken as a whole, the self-design algorithm generates a taper with a bit smaller taper ratio ξ (0.17 vs 0.19), a somewhat greater trapping fraction, but less radiation power (3.9 vs 4.4 TW) at the end of the undulator.

In Fig. 7 we compare the multidimensional optimization results with an optimal r_b variation (red lines) and with a constant r_b of 17.5 μm (green lines, same results as plotted in Fig. 6). The optimal r_b variation results in a higher F_t, larger on-axis $|a_s|$, and smaller r_s at the exit of the undulator, and eventually increases the predicted output radiation power by a factor of 11% (4.9 vs 4.4 TW).

2. A 200-m, hard x-ray, tapered FEL with break sections

Currently, most FEL facilities [18–21] in operation or under construction with total undulator lengths up to a few hundred meters are composed of individual undulator segments separated by break sections used for transverse

focusing, orbit correction, and diagnostics. As an example, the LCLS-II upgrade project [4,21] adopts a 1-m break section after each 3.4-m undulator segment, to accommodate quadrupoles, diagnostics, and phase shifters.

To investigate the impact of the nonzero-length break sections upon the radiation output, we perform multidimensional optimization for a 200-m tapered FEL with break sections of 1 m per 4.4 m, and with the other parameters the same as those listed in Table I. The optimization results are shown in Fig. 8 as black solid lines and compared with that for a 200-m tapered FEL without break sections (red lines, same results as plotted in Fig. 7). In the case with break sections, the maximum radiation power is 2.65 TW, only 54% of that achieved in the case without break sections. The reduction in the predicted available power is partially due to the 23% decrement of the magnetic length, which leads to a smaller value for the optimal taper ratio (0.12 vs 0.19). Another important cause is the vacuum diffraction of radiation in break sections, which causes a further increase in r_s and decrease in a_{s0}. These changes cascade into a further increase of on-axis Ψ_r and decrease in F_t, and hence even less efficiency in energy extraction compared to the case without break sections.

3. Dependence of the available maximum radiation power on various parameters

From a wide range of multidimensional optimizations based upon GENESIS single-frequency simulations, we can qualitatively summarize the dependence of the available maximum radiation power in a tapered, x-ray FEL on various parameters of the initial electron beam, radiation field, and the undulator system as follows.
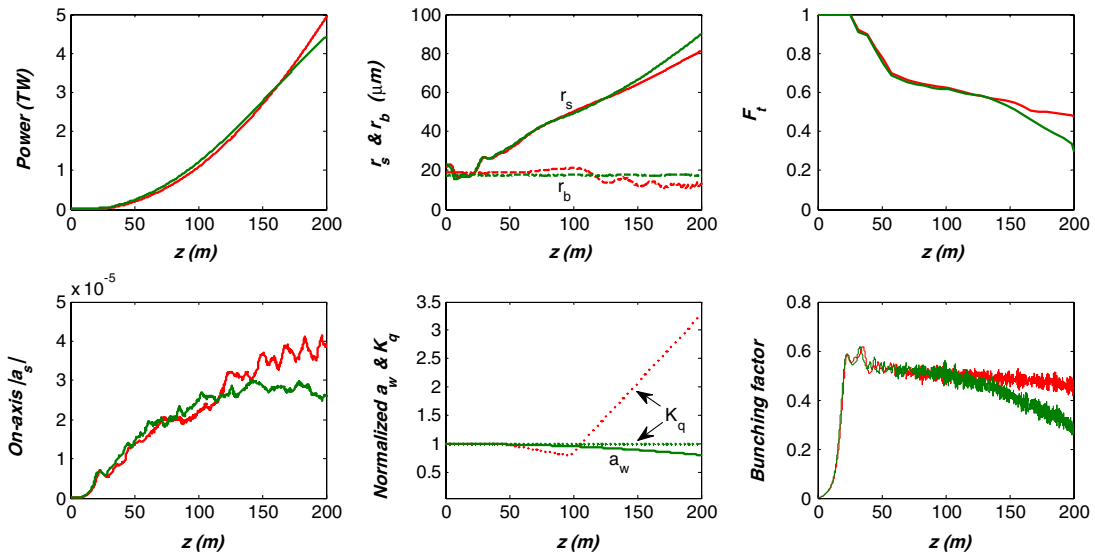

FIG. 7.   GENESIS single-frequency simulations for the evolution of the radiation power, on-axis $|a_s|$, r_s, F_t, and bunching factor in a 200 m, hard x-ray, tapered FEL for the cases of constant r_b (green lines) and varied r_b (red lines). The taper profile is obtained from multidimensional optimization.
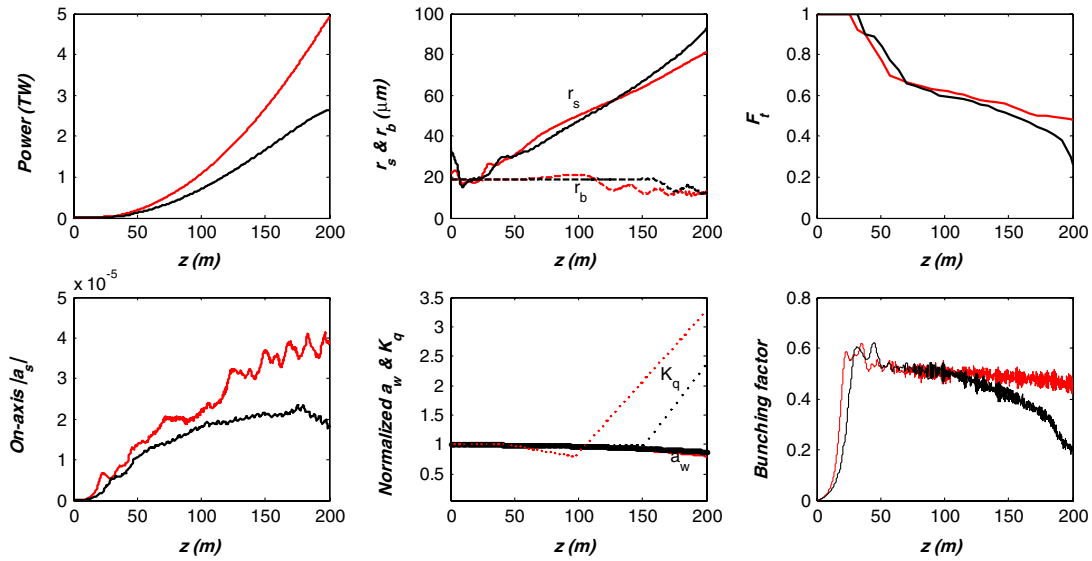
FIG. 8.   GENESIS single-frequency simulations for the evolution of the radiation power, on-axis $|a_s|$, r_s, F_t, and bunching factor in a 200 m, hard x-ray, tapered FEL with (black lines) and without (red lines) break sections. The taper profiles and r_b variations are obtained from multidimensional optimization.

a. Initial electron beam parameters

The maximum radiation power varies approximately quadratically with the initial beam current, and linearly with the initial beam energy. In addition, the smaller the emittance and energy spread of the electron beam, the larger the maximum extractable radiation power.

b. Input radiation field parameters

The impact of the initial radiation field parameters, such as the radiation input power and the initial spot size, upon the maximum extractable radiation power is relatively small, especially when the seed power exceeds a specific value (e.g., 1 MW for the parameters of Table I).

c. Undulator system parameters

For the same values of radiation seed and electron beam parameters, a helical undulator produces higher radiation power than the linearly polarized undulator, by a factor of about $1/[JJ]^2$; an undulator with shorter focusing/ diagnostic break sections produces higher radiation power; a longer undulator can be optimized for higher maximum radiation power and performs best with a taper with smaller $|a'_w|$ relative to what is true for shorter undulators.

Note that these relations are obtained with optimization based only on time-steady, single-frequency, 3D simulations. We will show in the next section that, when one includes time-dependent, multifrequency effects such as sideband growth, the radiation power will reach saturation in a shorter undulator length and at a much smaller value.
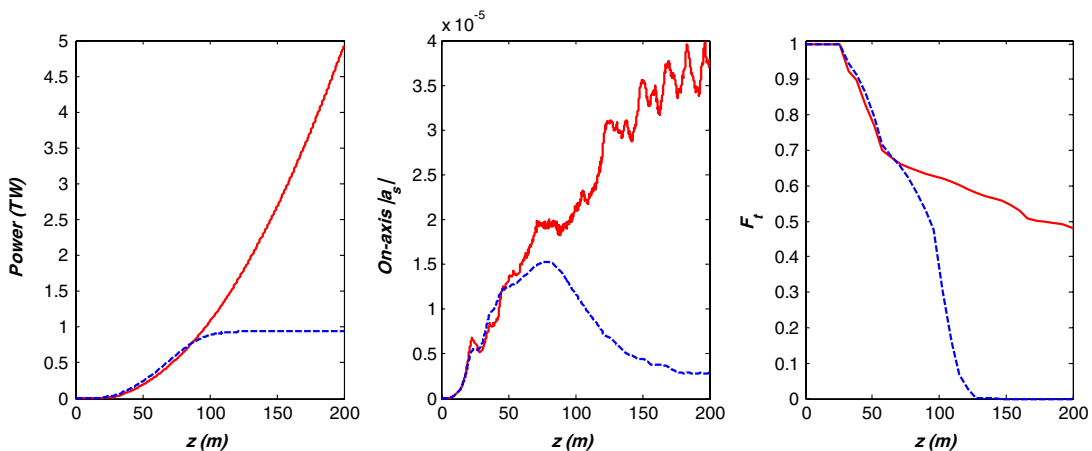

FIG. 9.   Evolution with z of the average radiation power, on-axis $|a_s|$ and F_t in a 200-m, hard x-ray, tapered FEL without break sections, as predicted by GENESIS single-frequency (red solid lines) and time-dependent simulations (blue dashed lines).

IV. IMPACT OF SIDEBAND GROWTH ON THE OPTIMIZATION OF A TAPERED FEL

In this section, we discuss the impact of time-dependent, multifrequency effects on the optimization of a tapered FEL. Sideband growth [5,22] generally occurs after *initial saturation* and can be excited by the SASE components that grow in the exponential gain regime and originate from the shot noise on the electron beam. From the viewpoint of energy conservation, the sideband power accumulates along the undulator associated with the slowly varying $|a_s|$ in the tapered region, generally reducing the energy gain of the primary wave and causing a more rapid detrapping due to the resultant decrease of longitudinal coherence, particularly with respect to the eikonal phase. Eventually, these sideband-induced effects lead to saturation of the radiation power at a much earlier point in z than would be true in their absence.

To illustrate the impact of the sidebands upon the radiation output, we perform GENESIS time-dependent simulation for a 200-m, hard x-ray, tapered FEL without break sections. Figure 9 compares the results from time-dependent simulation (blue dashed lines) with that from a single-frequency run (red solid lines); the taper, electron beam, and radiation seed parameters are the same as used in Fig. 7. It is clear that time-dependent effects lead to significant detrapping by around 100 m, and the average radiation power and on-axis $|a_s|$ reach saturation much earlier in z and, most importantly, at much smaller values than is true for the single-frequency run. The output spectra (not shown here) display well-developed upper and lower sideband peaks offset from the central resonant wavelength.

The simulation results shown in Fig. 9 indicate that time-dependent, sideband effects, introduce important limitations. A systematic theoretical and simulation analysis of these effects is under way and results will be presented in a forthcoming paper.

V. CONCLUSION

In this paper, we have built an explicit physical model of a tapered FEL by extending the 1D KMR theory of Ref. [5] to include certain 3D effects, e.g., refraction, diffraction, and the radial dependence of both the radiation field and the electron trapping fraction. While the model contains some approximations, it has good success in predicting the behavior with z of the electron and radiation beams, for the case of a hard x-ray with LCLS-II like parameters. Based on the model, we give a general description of the evolution of the electron-radiation interaction beginning from the *initial saturation* location to the end of a tapered undulator, and illustrate that the decreasing of refractive guiding is the major cause of the reduction of energy extraction efficiency, particle detrapping, and finally saturation of the radiation power. This framework provides guidance for maximizing the overall energy extraction efficiency in a tapered FEL.

We then propose and apply multidimensional optimization via GENESIS single-frequency simulations to find a taper profile that maximizes the output radiation power. We also find that a reasonable variation of the transverse focusing and thus electron beam cross section can, to some extent, further enhance energy extraction efficiency. When we extend our studies to include time-dependent, multifrequency phenomena, we find that sideband growth in a tapered x-ray FEL leads to strong detrapping effects and, contrary to single-frequency results, to saturation of the radiation power in a shorter undulator length and at a much smaller value. The physical model and analysis in this paper can provide a good foundation for future studies.

[1] H. N. Chapman *et al.*, Nature (London) **470**, 73 (2011).
[2] N. M. Seibert *et al.*, Nature (London) **470**, 78 (2011).
[3] G. Geloni, V. Kocharyan, and E. Saldin, DESY Report No. 10-108, 2010.
[4] W. M. Fawley, J. Frisch, Z. Huang, Y. Jiao, H.-D. Nuhn, C. Pellegrini, S. Reiche, and J. Wu, in Proceedings of the 33rd International Free Electron Laser Conference, Shanghai, China, 2011, TUOA4.
[5] N. M. Kroll, P. L. Morton, and M. N. Rosenbluth, IEEE J. Quantum Electron. **17**, 1436 (1981).
[6] G. Geloni, V. Kocharyan, and E. Saldin, DESY Report No. 11-049, 2011.
[7] L.-H. Yu, Phys. Rev. A **44**, 5178 (1991).
[8] K.-J. Kim, Report No. LBL-31969, 1992.
[9] S. Reiche, Nucl. Instrum. Methods Phys. Res., Sect. A **429**, 243 (1999).
[10] W. M. Fawley, Reports No. LBID-2141, CBP Technical Note-104, UC-414, 1995.
[11] E. T. Scharlemann, A. M. Sessler, and J. S. Wurtele, Phys. Rev. Lett. **54**, 1925 (1985).
[12] P. Sprangle, A. Ting, and C. M. Tang, Phys. Rev. Lett. **59**, 202 (1987).
[13] B. Hafizi, A. Ting, P. Sprangle, and C. M. Tang, Phys. Rev. Lett. **64**, 180 (1990).
[14] W. M. Fawley, Nucl. Instrum. Methods Phys. Res., Sect. A **375**, 550 (1996).
[15] W. M. Fawley, Z. Huang, K.-J. Kim, and N. A. Vinokurov, Nucl. Instrum. Methods Phys. Res., Sect. A **483**, 537 (2002).
[16] M. Xie, in *Proceedings of the Particle Accelerator Conference, Dallas, TX, 1995* (IEEE, New York, 1995), p. 183.
[17] Except the presented model for the taper profile, one can use a polynomial taper profile to high orders.

[18] J. Arthur *et al.*, LCLS Design Report No. SLAC-R-593, 2002.

[19] SCSS XFEL Conceptual Design Report, 2005.

[20] R. Abela *et al.*, Report No. DESY-2006-097, 2006.

[21] LCLS-II, https://slacportal.slac.stanford.edu/sites/lcls_public/lcls_ii/Pages/default.aspx.

[22] N. M. Kroll, P. L. Morton, and M. N. Rosenbluth, in *Free-Electron Generators of Coherent Radiation*, edited by S. F. Jacobs, H. S. Pilloff, M. Sargent, M. O. Scully, and R. Spitzer, Physics of Quantum Electronics (Addison-Wesley, Reading, MA, 1980), Vol. 7, p. 147.