Some Issues and Subtleties in Numerical Simulation of X-ray FEL's

W.M. Fawley[a]

[a]Center for Beam Physics, Lawrence Berkeley National Laboratory, Berkeley CA 94720-8211 USA

Part of the overall design effort for x-ray FEL's such as the LCLS and TESLA projects has involved extensive use of particle simulation codes to predict their output performance and underlying sensitivity to various input parameters (*e.g.* electron beam emittance). This paper discusses some of the numerical issues that must be addressed by simulation codes in this regime. We first give a brief overview of the standard approximations and simulation methods adopted by time-dependent (i.e. polychromatic) codes such as GINGER[1], GENESIS[2], and FAST3D[3], including the effects of temporal discretization and the resultant limited spectral bandpass, and then discuss the accuracies and inaccuracies of these codes in predicting incoherent spontaneous emission (*i.e.* the extremely low gain regime).

1. Introduction

Over the past decade there has been increasing interest in basing so-called 4th-generation light sources upon x-ray FEL's. Due both to the present day absence of reasonably reflective optics at x-ray energies (*i.e.* $\geq$ 1 keV) which precludes multipass oscillator configurations and the absence of x-ray laser input sources which could function at repetition rates in the 100-Hz and faster regime, use of self-amplified spontaneous emission (SASE) appears to be the most promising approach to construct an x-ray FEL in the immediate future. In the last year, the U.S. Dept. of Energy has given the initial "green light" to construction of the Linac Coherent Light Source (LCLS)[4] at SLAC and the German government's scientific advisory council has also recently given high marks (and hopefully eventually Euros) to the TESLA x-ray FEL project[5] at DESY.

The cost of each of these facilities will exceed well over $250 million, especially when downstream x-ray optics and experimental beamlines are included. Since the shortest wavelength that an operating SASE FEL has achieved to date is $\sim$ 80 nm[6] or nearly three orders of magnitude longer than the desired $\sim$ 0.1 nm wavelength sought by x-ray FEL proponents, many design choices for the proposed facilities have been based upon numerical simulation studies rather than detailed experiments. Consequently, it is reasonable to question the believability and accuracy of simulation codes such as GINGER[1], GENESIS[2], and FAST3D[3] in modeling x-ray FELs.

In my oral talk given at FEL02 conference, I addressed both the structure of the simulation codes and their relative success in reproducing the experimental results of recent optical/UV SASE FEL experiments at the Argonne APS/LEUTL facility[7], the DESY TTF-FEL facility[6], and the VISA facility[8] at Brookhaven. However, in this paper, rather than give an review of previously published work concerning modeling SASE experiments, I make some simple observations only:

1. The codes have accurately reproduced "core" physical observables which arise from the basic FEL and EM field equations such as exponential gain lengths, saturation lengths, output spectral width, and far field angular mode size.

2. When available from experimental measurements, the codes have also reproduced the effective start-up power and saturation power.

3. For all the above cited experiments, the use of an electron beam compressor has played a very important role. The resultant 6D macroscopic phase space is not at all

a simple convolution of independent Gaussian distributions. Significantly, in each of the three experiments, the FEL emission came overwhelmingly from a high current "spike". For the VISA experiment in particular, it was crucial that a detailed start-to-end simulation be performed in order for the GENESIS code to reproduce the experimental results.

For the rest of this paper, I will concentrate upon the underlying simulation code assumptions, approximations and algorithms. I first discuss the standard assumptions and approximations that nearly all FEL codes make such as the slowly-varying envelope (eikonal) and wiggle-period-averaging approximations. I then discuss the underlying spatial and temporal structure of the codes and how this affects what EM field modes are present. Finally, I discuss the treatment of shot noise and spontaneous emission and how the previous approximations affect the accuracy of these predictions.

2. Standard Assumptions and Approximations in FEL Simulation Codes

2.1. Slowly-Varing Envelope Approximation

Due to Courant condition limitations arising from the large ratio of undulator length $L_w \sim \mathcal{O}(1-100\,\mathrm{m})$ to electromagnetic wavelength $\lambda_s \sim \mathcal{O}(10^{-10}-10^{-6}\,\mathrm{m})$, it is impossible with reasonable computational resources to use a full E&M simulation code to model short wavelength FELs in the laboratory frame. Consequently, nearly all FEL codes make the slowly-varying envelope approximation (SVEA) (also variously referred to as the paraxial and/or eikonal approximation) in which all electromagnetic field and particle microbunching quantities are presumed to be composed of the product of a slowly varying "envelope" quantity times a fast variation at a central frequency and wavenumber, *e.g.* :

$$E(\vec{r},z,t) \equiv \tilde{E}(\vec{r},z,t) \times \exp\, i(k_0 z - \omega_0 t) \quad (1)$$

This decomposition is similar to FM radio where the fast term is the carrier wave. By presuming that all longitudinal variations (*e.g.* due to diffraction, gain and refraction) of field envelope quantities occur on length scales much greater than k_0^{-1} and that temporal (non-monochromatic) variations occur on time scales much greater than ω_0^{-1}, the normally hyperbolic field equations can be transformed into parabolic diffusion equations. Numerically, this frees one from the Courant condition which otherwise limits $k_0\Delta z$ and $\omega_o \Delta t$ to $\mathcal{O}(2)$. For x-ray FEL's where gain lengths and Rayleigh lengths typically exceed one meter and where the normalized gain bandpass is generally less than 1 part in 10^3, there is little serious danger for growing modes of violating the SVEA approximation.

2.2. Wiggler-Period Averaging

A commonly-used approximation more singular to the FEL community is so-called wiggle-period-averaging[9]. This is a stronger approximation in z than the eikonal approximation because here one asserts there is little change in field and particle envelope quantities over a z length scale corresponding to one undulator period λ_w, typically 1-10 centimeters, orders of magnitude greater than λ_s for short wavelength FELs. Here, too, the large gain length of x-ray FEL's make this a good approximation.

However, there is at least two possible areaa where wiggler-period averaging poses a *potential* problem. The first is regards to spontaneous emission and is discussed further in §3. The second has to do with high order (*i.e.* $M_\perp > 1$) spatial modes which can have very small Rayleigh ranges compared to the lowest order mode whose transverse size is comparable to the focused beam radius:

$$Z_R(M_\perp) \approx Z_R^0 \,/\, M_\perp^2 \quad (2)$$

Requiring the mode diffraction length to exceed the undulator period is equivalent to

$$M_\perp \leq \left(\frac{4\pi\varepsilon}{\lambda_s}\right)^{1/2} \left(\frac{\beta}{4\lambda_w}\right)^{1/2} \quad (3)$$

For small ratios of focusing β to λ_w and ε to λ_s, as was true for the VISA experiment, even $M_\perp = 1$ barely obeys the above relation indicating wiggler-period-averaging may not be appropriate at all for higher order modes. On the

other hand, higher order spatial modes have relatively little gain and in many cases (such as VISA) the size of the fastest growing radiation mode can be significantly larger than the electron beam (due to strong diffraction effects). Consequently, this problem is perhaps more academic than real. Nonetheless, it indicates that one must exercise some caution when numerically examining the gain/propagation properties of high order modes for visible/IR FEL amplifiers operating with very low emittance electron beams.

There is another point on the subject of wiggler-period averaging that is relevant to this discussion. If one makes a Lorentz transformation to the frame comoving with the average *longitudinal* velocity $c\beta_z$ of the electron beam, the FEL instability becomes identical in nature to a Raman backscatter instability acting in the weak plasma, strong pump regime. More importantly, in this frame, the radiation wavelength and the undulator wavelength become equal (presuming FEL resonance), and the eikonal and wiggler-period-averaging approximations become identical. Radiation modes in this frame with $k_\perp/k_z \sim 1$ may not obey the eikonal approximation; this point is equivalent to one made by Saldin *et al.* [10] who cautioned that modes with opening angles $\geq \gamma_z^{-1}$ might not be properly represented. This problem can also apply to high spatial frequency undulator errors when applying wiggler-period averaging.

2.3. Discrete Temporal and Slippage Resolution

Most FEL simulation codes such as GINGER subdivide field and microbunching quantities onto a temporal grid with a uniform spacing ΔT. The slow variation of envelope quantities such as $\tilde{E}$ in Eq. 1 implies that one may generally use a much coarser temporal scale than λ_0/c without losing significant physics information. The discretization of quantities [*e.g.* $E(t)$] in the temporal domain implies that they are similarly discretized in the frequency domain [*e.g.* $E(\omega)$] with a maximum spectral bandpass $\omega_0 \pm 0.5\Delta\omega_{\max}$ with $\Delta\omega_{\max}$ given by the Nyquist frequency:

$$\Delta\omega_{\max} = 2\pi/\Delta T \tag{4}$$

For x-ray FEL's where the FEL gain parameter[11] $\rho \ll 1$, one typically chooses $c\Delta T/\lambda_0 \leq 32\pi/\rho \sim \mathcal{O}(5-20)$ which ensures that the simulation's total spectral bandpass is four times or greater than the FEL gain bandpass. The minimum spectral resolution scales inversely as the total number N_S of temporal slices:

$$\frac{\Delta\omega_{\min}}{\omega_0} = \frac{\lambda_0}{N_S\, c\Delta T} \tag{5}$$

Normally one needs $N_S \geq 96$ to obtain simultaneously good spectral resolution of the gain bandpass and to ensure that the total spectral range is much larger than the gain bandpass. Since these codes operate in the time domain, spectral information is normally available only after post-processing where one usually applies Fourier transforms to obtain $P(\omega)$.

Temporal discretization also leads to a discretization with regards to how slippage is applied. In general (ignoring the presence of drift spaces, dispersive sections, *etc.*), a given electron beam slice will "fall back" one radiation slice in a z interval

$$\Delta z_{\text{slip}} \equiv \lambda_w \times \frac{c\Delta T}{\lambda_r} \tag{6}$$

Normally, $\lambda_0 \approx \lambda_r$, the wavelength corresponding to FEL resonance, and one either ignores the γ dependence of the slippage rate and/or presumes in the case of significant energy extraction from the electron beam that the undulator strength is tapered to maintain FEL resonance for the ponderomotively-trapped particles. Due to the wiggler-period-averaging approximation, $\Delta z_{\text{slip}} \geq \lambda_w$ which leads to the maximum spectral bandpass $\Delta\omega_{\max} \leq \omega_0$. Discrete slippage and temporal resolution permit relatively easy parallelization of time-dependent FEL simulation codes because each electron beam slice interacts with one and only one radiation slice over a particular Δz_{slip} interval. This also minimizes "message-passing" between different processors on multiprocessor computers as it is necessary only when slippage is applied, not on each z-step where the field and particle quantities are advanced in z.

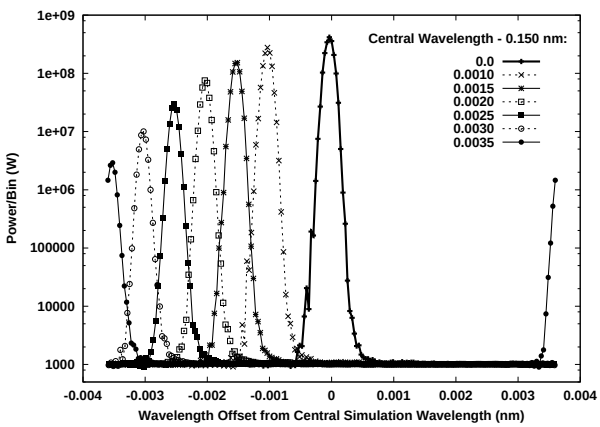

Figure 1. Spectral power versus $\lambda - \lambda_0$ for a series of LCLS-like GINGER simulations for different shifts of λ_0 from the wavelength λ_r corresponding to FEL resonance (0.15000 nm for these runs).

As discussed above, one usually makes the central wavelength λ_0 of the simulation spectral "window" equal to or quite close to the wavelength λ_r corresponding to FEL resonance (and normally peak FEL gain). However, if, for whatever reason (*e.g.* energy chirping, a large temporal variation in mismatch or emittance leading to a variation in β_z), λ_r moves away from λ_0 toward the boundary of the spectral window, there can be a strong numerical (*i.e.* unphysical) suppression of gain. In Fig. 1 we plot $P(\lambda - \lambda_0)$ for a series of LCLS-like time-dependent GINGER simulations in which λ_0 was increasingly moved away from FEL resonance. Each simulation was started with a constant radiation noise of 1000W/bin and run to a point approximately 10 gain lengths in z, well before power saturation. Hence, on this semilog plot, the height of each point is linearly proportional to gain. One sees that as $|\lambda_0 - \lambda_r|$ approaches the boundary of the spectral window, the gain is suppressed by nearly a factor of 2. Moreover, for the largest separation value plotted (0.0035 nm), there is an aliasing effect where gain starts to appear at the other boundary. These results indicates that one must be extremely careful when simulating FEL's with energy-chirped beams, such as may well be true for the LCLS and Tesla X-ray FEL either intentionally or due to time-dependent wakefield losses.

3. Modeling Spontaneous Emission with FEL Simulation Codes

An extremely important issue regarding the believability of x-ray FEL modeling is the accuracy of SASE startup. The codes must both have algorithms which produce the correct amount of electron beam microbunching — at least at those wavelengths which have significant FEL gain — and which correctly treat coupling to radiation emission and eventual exponential gain.

There are a number of published papers, both relatively old (*e.g.* [12,13]) and new (*e.g.* [14–17]), concerning shot noise simulation algorithms for FEL codes. Although the specific details and particular features (*e.g.* ability to handle energy spread, harmonics) may differ, in general these algorithms are very similar and give the proper zero- and first-order distribution properties for a spectral band around the central wavelength of the simulation's spectral window.

Although the shot noise statistics may be reproduced properly, there are a number of complex issues regarding how fully the overall spontaneous emission is reproduced. Saldin *et al.* in Ref. [10] give an excellent discussion on some of these issues which include:

1. Given the finite spectral bandpass of the simulation, wavelengths below and (especially) above the bandpass will be neither emitted nor represented at all.

2. Most simulation codes employ a transverse grid with a finite number of zones upon which particle microbunching and radiation quantities are decomposed. Hence, only a artificially-limited number of transverse modes can be emitted/represented. This is true both in the near-field *and* far-field. In the case of GINGER, the radiation field is presumed axisymmetric which further limits the number of modes.

As a partial illustration of this issue, a series of GINGER simulations were performed to exam-

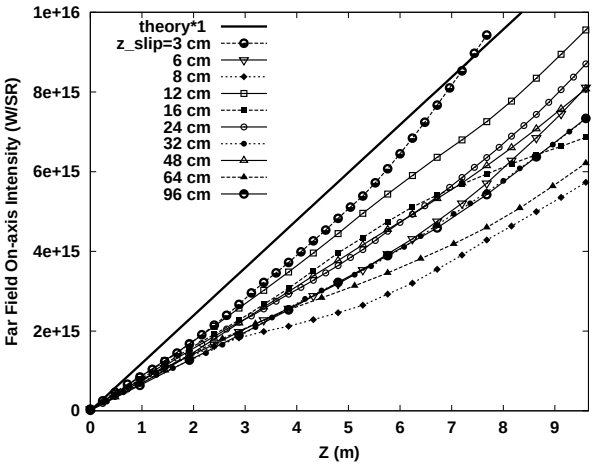

Figure 2. On-axis, far field radiation power for an LCLS-like GINGER simulation showing the effects of changing the overall spectral bandpass via $\Delta z_{\rm slip}$.

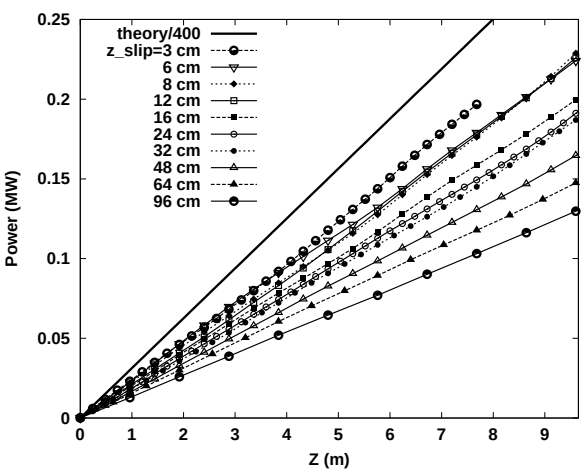

Figure 3. Total spontaneous radiation power for the same example as in Fig. 2.

ine spontaneous emission early in the undulator. In the first case, a extremely low gain $1-\mu$m wavelength FEL with $4\pi\epsilon/\lambda_r \approx 0.25$, the code accurately reproduced the on-axis far field intensity and the total output power contained within the central radiation cone (see Figs. 2 and 3 of Ref. [16]). However, the total output power was low by a factor of five which was attributed to the simulation's finite spectral bandpass.

In a more relevant case employing "standard" LCLS parameters ($I_B = 3400\,$A, $E_B = 14.35\,$GeV, $\varepsilon_N = 1.2\,$mm-mrad, $\lambda_r = 0.15\,$nm, $\lambda_w = 3\,$cm, $K = 3.71$, $4\pi\epsilon/\lambda_r \approx 3.6$), we examined the computed spontaneous emission as a function of spectral bandpass (equivalently $\Delta z_{\rm slip}$). Fig. 2 shows that the computed on-axis far field intensity generally agrees with the analytic value to within $\sim 25\%$ although the agreement tends worsen as the bandpass decreases ($\Delta z_{\rm slip}$ increases). However, the total emitted power (see Fig. 3) disagrees by greater than a factor of 400 from the analytic result and there is a clear trend toward greater disagreement as the simulation's bandpass is reduced (via larger $\Delta z_{\rm slip}$).

The nearly two orders of magnitude increase in the theory/simulation discrepancy ratio between the low emittance[16] and the present LCLS examples is almost certainly due to the importance of higher order modes in the latter. A similar point was previously made in Ref. [10]. In these conference proceedings, Huang and Kim[18] have made an independent study which confirms the important contribution of higher order modes to the total spontaneous power where they find that as the number of numerically-permitted modes increases (*i.e.* going from a 2-D GINGER run to a 3-D GENESIS run), the spontaneous emission grows by more than an order of magnitude. However, the two codes do agree well in the asymptotic exponential gain region where the lowest order, highest gain (and axisymmetric) mode dominates.

Consequently, the author presently believes that for high gain, high emittance FEL's such as the LCLS, in order to accurately reproduce quantitatively the spontaneous emission spectrum near the fundamental resonant wavelength,

one will have to use a full 3D code, use a value of $\Delta z_{\rm slip}$ not much larger than λ_w to get a large spectral bandpass, *and* resolve the electron beam with a large number of grid zones [*e.g.* $\geq \mathcal{O}(100)$]. These requirements must be fulfilled also if one wishes to determine in detail how transverse coherence quantitatively scales with z (although one may be able to use fewer grid zones since only a few low order modes will dominate after the first few gain lengths). Fortunately, however, if one is most interested in the total saturated power and/or the on-axis far field emission, it does not appear necessary to use either such a wide spectral bandpass nor a full 3-D code (although for other reasons such as undulator errors the latter may be required).

4. Summary

In this paper I have given a brief introduction to some of the approximations and algorithms adopted by FEL simulation codes such as GINGER, GENESIS, and FAST3D which are presently used to model SASE-based x-ray FEL's. In general, these approximations (*e.g.* wiggler-period-averaging, SVEA) are even safer at x-ray wavelengths than in the visible or infrared where they have been successfully used in FEL simulation for many years. There are some possible pitfalls that require caution when modeling particular scenarios (*e.g.* energy-chirped electron beams; very low emittance, visible/IR FELs). Furthermore, one must be cognizant of which portion of the spontaneous emission spectrum is and is not properly represented in FEL codes when modeling SASE startup.

Acknowledgements

I am pleased to acknowledge useful discussions with Z. Huang, V. Litvinenko, S. Reiche, and M. Yurkov. This work was supported at LBNL by the U.S. Department of Energy under Contract No. DE-AC03-76SF0098. Computational resources were provided in part by the National Energy Research Supercomputer Center (NERSC).

REFERENCES

1. W.M. Fawley, LBNL Report No. LBNL-49625 (2002).
2. S. Reiche, Nucl. Inst. Meth. **A429**, 243 (1999).
3. E.L. Saldin, E.A. Schneidmiller, M.V. Yurkov, Nucl. Inst. Meth. **A429**, 233 (1999).
4. "Linac Coherent Light Source (LCLS) Design Study Report", SLAC-R-521, UC-414 (1998).
5. "Conceptual Design of a 500 GeV e+e- Linear Collider with Intgegrated X-ray Laser Facility, DESY 97-048 (1997).
6. V. Ayvazyan *et al.* , Phys. Rev. Letters, **88**, 104802 (2002).
7. S.V. Milton *et al.* , Science, **292**, 2037 (2001).
8. A. Murokh *et al.* , "Properties of an Ultra-Short Gain Length, Saturated, Self-Amplified Spontaneous Emission FEL", submitted to Phys. Rev, 2002; see also A. Tremaine *et al.* , Phys. Rev. Lett., **88**, 204801 (2002).
9. N. Kroll, P. Morton, M. Rosenbluth, IEEE J. Quantum Elec., **QE-17**, 1436 (1981).
10. E.L. Saldin, E.A. Scneidmiller, M.V. Yurkov, Opt. Commun., **186**, 185 (2000).
11. R. Bonifacio, C. Pellegrini, and L. Narducci, Opt. Commun., **50**, 373 (1984).
12. S. Benson and J.M.J. Madey, Nucl. Inst. Meth. Phys. Res. **A237**, 55 (1985).
13. C. Penman and B.W.J. McNeil, Opt. Commun. **90**, 82 (1992).
14. S. Reiche *et al.* , in *Towards X-Ray Free Electron Lasers* (Garda Lake, Italy; 1997), AIP Conf. Proc. 413 (ed. R. Bonifacio and W.A. Barletta), 29 (1997).
15. E.L. Saldin, E.A. Scneidmiller, M.V. Yurkov, Opt. Commun., **148**, 383 (1998).
16. W.M. Fawley, Phys. Rev. ST-AB, **5**, 070701 (2002).
17. V. Litvinenko, "Macroparticle FEL Model with Self-Consistent Spontaneous Radiation", Poster WE-P-45, these proceedings.
18. Z. Huang and K.-J. Kim, "Spontaneous and Amplified Radiation at the Initial Stage of a SASE FEL", Poster MO-P-15, these proceedings.