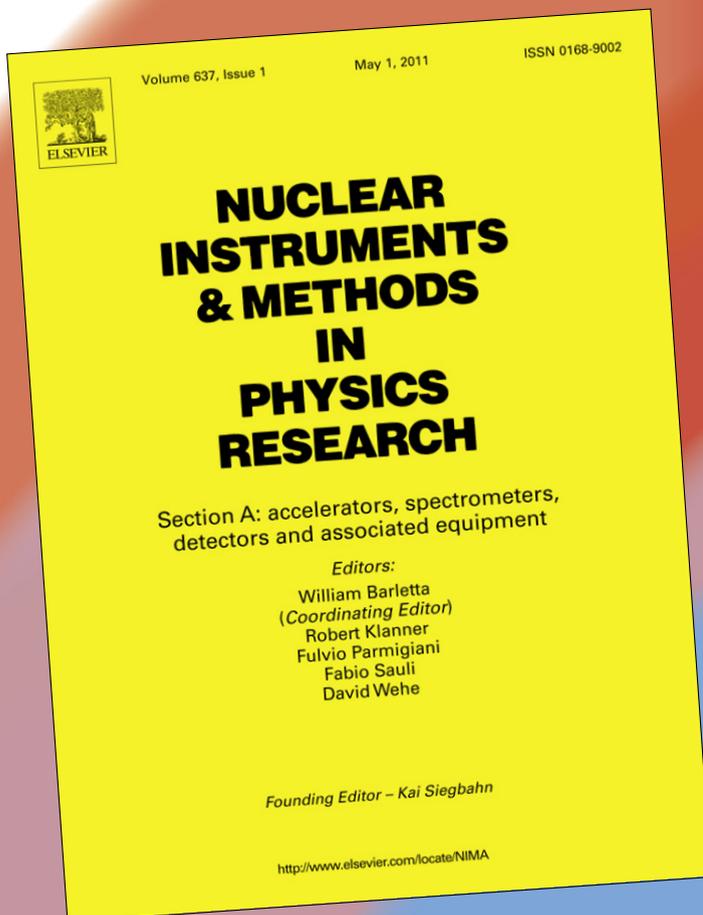


Accelerator Physics for Future Light Sources

A collection of reports from the workshop
sponsored by the
Department of Energy
Office of Basic Energy Sciences
held in Gaithersburg, Maryland
September 15–17, 2009



Co-chairs:
W. A. Barletta
Massachusetts Institute of Technology
J. N. Corlett
Lawrence Berkeley National Laboratory



Contents

Basic Energy Sciences Workshop: Accelerator Physics for Future Light Sources

1. Editorial

2. Free electron lasers: Present status and future challenges

W.A. Barletta, J. Bisognano, J.N. Corlett, P. Emma, Z. Huang, K.-J. Kim, R. Lindberg, J.B. Murphy, G.R. Neil, D.C. Nguyen, C. Pellegrini, R.A. Rimmer, F. Sannibale, G. Stupakov, R.P. Walker and A.A. Zholents
(*Nuclear Instruments and Methods in Physics Research A* 618 (2010) 69–96)

3. X-ray Sources by Energy Recovered Linacs and Their Needed R&D

S. Benson, M. Borland, D.R. Douglas, D. Dowell, C. Hernandez-Garcia, D. Kayran, G.A. Krafft, R. Legg, E. Moog, T. Obina, R. Rimmer and V. Yakimenko
(*Nuclear Instruments and Methods in Physics Research A* 637 (2011) 1–11)

4. The Potential of an Ultimate Storage Ring for Future Light Sources

M. Bei, M. Borland, Y. Cai, P. Elleaume, R. Gerig, K. Harkay, L. Emery, A. Hutton, R. Hettel, R. Nagaoka, D. Robin and C. Steier
(*Nuclear Instruments and Methods in Physics Research A* 622 (2010) 518–535)

5. New source technologies and their impact on future light sources

B.E. Carlsten, E.R. Colby, E.H. Esarey, M. Hogan, F.X. Kärtner, W.S. Graves, W.P. Leemans, T. Rao, J.B. Rosenzweig, C.B. Schroeder, D. Sutter and W.E. White
(*Nuclear Instruments and Methods in Physics Research A* 622 (2010) 657–668)

6. Enabling Instrumentation and Technology for 21st Century Light Sources

J.M. Byrd, T.J. Shea, P. Denes, P. Siddons, D. Attwood, F. Kaertner, L. Moog, Y. Li, A. Sakdinawat and R. Schlueter
(*Nuclear Instruments and Methods in Physics Research A* 623 (2010) 910–920)

7. Cathode R&D for Future Light Sources

D.H. Dowell, I. Bazarov, B. Dunham, K. Harkay, C. Hernandez-Garcia, R. Legg, H. Padmore, T. Rao, J. Smedley and W. Wan
(*Nuclear Instruments and Methods in Physics Research A* 622 (2010) 685–697)

Editorial

Basic energy sciences workshop: Accelerator physics for future light sources executive summary

On September 15–17, 2009, the Office of Basic Energy Sciences sponsored an expert workshop to provide a sound scientific and technical basis that would underpin future investment in accelerator R&D in support of the next major synchrotron radiation research facilities. These future facilities are expected to enable the United States to maintain world-leading capabilities in X-ray science. The workshop considered four machine architectures: free electron lasers (FELs), energy recovery linacs (ERLs) and other recirculating linac designs, so-called “Ultimate” Storage Rings (USRs) that minimize electron beam emittance, and laser-driven X-ray sources. Each architecture had an assigned working group; in addition a fifth group considered the status of cross-cutting technologies and instrumentation.

Participation in the 2(1/2)day workshop was strictly by invitation only and was restricted to approximately 50 participants plus 5 early career, scientific scribes – one for each working group. Attendance was balanced with respect to staff from DOE laboratories and universities and also included several foreign experts from major laboratories. The participants were distributed in five working groups of roughly equal size. The Associate Director for Basic Energy Sciences gave the following charge to the workshop:

- (1) Evaluate the state of readiness of the machine architectures for the next major X-ray science user facility.
- (2) Identify what capabilities could be ready in 5 years? In 10 years?
- (3) Identify the physics and technological challenges of the respective approaches.
- (4) Describe research steps and directions toward a new generation of photon sources in the form of an R&D roadmap.
- (5) No project specific proposals are to be considered.

The response to these charges will be written in the form of a set of five scientific papers to be submitted to a peer-reviewed international journal. The working groups also produced closeout presentations summarized in the following sections.

Free electron lasers

The LCLS is a game changer for pulsed X-ray sources especially for hard X-rays. Building on its great success in achieving an 11-order of magnitude increase in peak brightness over storage ring sources at an X-ray wavelength of 0.15 nm, the group examined in which directions advances could be made in a FEL source operating in the future. The technological springboards of progress are the demonstration at LCLS of sub-micron electron emittance at a relatively high beam charge (~ 100 pC), the efficacy of laser heating to suppress microbunching of the electron beam, demonstration of very low emittance at low (20 pC) e-beam charge, and the excellent agreement of theory and simulation with experimental results. A further important development at the SPARC test facility in Frascati, Italy, is the successful demonstration of the velocity-bunching technique, which may enable operation of an FEL with high quality, low charge bunches that are only a single cooperation length in duration (~ 1 – 2 fs).

While the LCLS produces longitudinally incoherent radiation via SASE, what could follow rapidly is a hard X-ray source with marginally coherent, ultra-short pulses using velocity bunching (~ 1 fs), or an extension of SASE lasing to ~ 20 keV. At LCLS the latter would require a new undulator or high beam energy. Other developments for hard X-ray FELs are common with the needs for soft X-ray facilities and are discussed below.

In the soft X-ray regime, today's status is exemplified by the FLASH facility at DESY, a SASE source at ~ 7 nm, and FERMI at Sincrotrone Trieste, now under construction as the world's first seeded FEL user facility with operation foreseen to ~ 2 nm. Sources similar as these – and extending to 1 nm wavelength – have no counterpart in the US, and could be built today. Seeded operation up to 1 kHz is feasible now, although further R&D on electron sources is needed to extend the lifetime of the photocathode with respect to maintaining good quantum efficiency over multiple weeks of user operation. Within five years more advanced FELs could generate stable, short (from sub-femtosecond to ~ 100 fs), temporally coherent soft X-ray pulses if the system is seeded, with natural synchronization for pump-probe experiments. While room temperature rf-linacs are limited to ~ 1 kHz operation, superconducting linacs could run at much greater repetition rates providing ~ 10 – 100 kHz or greater per FEL, assuming appropriate optical technology development.

On the five-to-ten year horizon one might be able to obtain very narrow bandwidth in picosecond pulses from X-ray FEL oscillators or regenerative amplifiers using multi-layer dielectric mirrors at wavelength down to 20 nm or possibly with Bragg plane reflectors in the hard X-ray region. The group notes that at IR to UV wavelengths oscillators are operational today and could be developed to optimize cost and performance. A principal challenge for shorter X-ray wavelengths is finding ways to produce large mirrors with high reflectivity over broad bandwidth, positional control and stability, adequate damage resistance, and reduced costs.

Hard and soft X-ray sources share many technological challenges. The principle challenges requiring focused R&D include:

- (1) high brightness, high repetition rate injectors with flexible and highly reproducible output currents that incorporate robust, efficient cathodes,
- (2) efficient seeding schemes for high harmonic upshifting that minimize the need for extensive development of laser systems,
- (3) control of e-beam jitter in charge, energy, intensity and trajectory,
- (4) diagnostics for pC electron and fs electron and photon pulses, and
- (5) high quality fast electron beam switchyards for proposed high repetition rate FEL facilities with multiple undulators.

Optimized CW superconducting rf structures and cryostats are needed for high repetition rate facilities. R&D to improve short-period undulators would benefit a future FEL or ERL facility by lowering the requisite beam energy, thereby reducing cost. Development of high power laser systems is needed for photoinjectors and FELs seeding if FEL-based facilities are to reach their full potential. The extent of development in terms of laser power depends largely on the success of developing materials for robust, high quantum efficiency cathodes, and effective seeding techniques such as echo-enabled harmonic generation (EEHG).

Much of the recommended R&D program could and should be demonstrated in realistic operating environments as part of dedicated FEL test beds. The FEL working group recommended using existing test beds in Europe (SPARC) and Japan (SSCS) and participating in machine physics studies of the FERMI FEL at Sincrotrone Trieste. In addition to the international collaborations, the group recommends that DOE proceed with construction of test bed facilities of two types, which could be built now:

- (1) A low repetition rate, room temperature system with E_e 1–2 GeV to reach $E_{ph} \sim 1$ nm, for tests of several seeding schemes and all pulse-by-pulse research issues.
- (2) A high repetition rate system, with $E_e \sim 100$ MeV (set by emittance freezing) to test gun and injector designs with high repetition rate lasers, cathodes, and diagnostics in an integrated system.

ERLs

Energy recovery linacs offer the potential to reach high spectral brightness $\sim 10^{22}$ ph/s/mm²/mrad²/0.1% BW in the hard X-ray range (5–50 keV photons), with high spatial coherence, and control of pulse duration to ~ 1 picosecond. These performance projections are three orders of magnitude beyond the current brightness of the APS, and one order of magnitude shorter in pulse duration. Such performance is predicated primarily on development of: (1) a long-lived, high-brightness and high average current (high repetition rate) electron source informed by relevant cathode material science, (2) development of appropriate drive lasers, (3) multi-pass (two-passes or more) recirculation of 10–100 mA injected beam currents through optimized CW superconducting rf cryomodules and supporting high beam power (e.g. 700MW for 100 mA at 7 GeV), (4) beam halo control, (5) development of high-resolution, multi-physics simulation codes. State-of-the-art technology ERLs have demonstrated energy recovery of up to 1.35MW (JLAB FEL), and two-pass recirculation, however with much lower than desired beam current and with beam emittances one-to-two orders of magnitude greater than required for an ultra-bright hard X-ray source. Extensive R&D is required to bring the technology to the level of maturity appropriate for initiating a construction project. Opportunities do exist for significant advances to be made using existing facilities at DOE institutions operated by SC offices. A dedicated test facility (~ 600 MeV) for injectors and multi-pass recirculation to allow tests of beam physics challenges is part of the R&D roadmap recommended by the working group. With appropriate R&D resources and the experience with a test bed facility, a proposal for a hard X-ray ERL facility could be developed to CD-1 level in ~ 10 years.

Ultimate storage rings

For users, the reliability ($\sim 98\%$), stability, average flux, tuning range, and polarization control of storage rings form the gold standard of light sources. All high repetition rate sources must be compared to a fully optimized storage ring source outfitted with a new generation of detectors that could be developed on the time scale of a major construction project. The working group foresees the possibility of reducing electron beam emittance (and thereby raising average X-ray beam brightness) by a factor of 100 to 1000 over existing storage rings to 10^{22} ph/s/mm²/mrad²/0.1% BW at 1 keV and 10^{23} at 10 keV. The newly approved, 3 GeV MAX-IV ring in Sweden will go a long way toward the best foreseeable electron beam emittance.

Typically a several GeV ring with such bright beams would be quite large (1 to 4 km in circumference) with a commensurately higher price tag. Two architectures have been studied for large rings: (1) a ring using conventional off-axis accumulation from a low duty cycle injector and (2) on-axis injection from a booster and accumulator ring or from a long pulse linac, in which the beam is swapped into the main ring. The choice of architecture is driven by the small dynamic apertures that have been calculated in design studies thus far. A ring with larger dynamic aperture allows for accumulation, while a ring with small dynamic aperture requires on-axis injection. The MAX-IV ring has adopted a 7-bend achromat design with conventional accumulation. The working group judged that all the scenarios considered were at most low-to-medium technical risk, and are readily amenable to risk reduction through an R&D program.

Almost all accelerator physics and technologies required to realize an ultimate storage ring are in hand today. However, there is no complete, integrated design at several GeV that optimizes machine performance. Therefore, the group recommends a focused international design study of recognized experts from the US, Europe, and Japan to define an optimized and realistic zero-order design. The bottom line is that there are no known showstoppers, and the major issue is the cost to construct such a machine.

Other sources

Alternates to conventional rf-accelerator technology are being developed through the application of lasers to produce X-rays by interaction with matter either directly, or by generating extremely high electric fields (> 10 GeV/m) to drive compact electron accelerators. Laser-based sources using High-Harmonic Generation (HHG) in gases offer stand-alone EUV/XUV sources with spatial and temporal coherence, high polarization, ultra-fast pulses of $\sim 10^{10}$ photons per pulse at 10Hz repetition rate and $\sim 10^{10}$ photons per second at 1 MHz, and at wavelengths shorter than 40 nm. These HHG sources have direct applications in ultrafast EUV/XUV spectroscopy and as seed sources for EUV/XUV-soft X-ray FELs. Significant R&D is needed for a well-characterized and reliable seed source for an FEL especially at high repetition rates.

Inverse Compton scattering allows production of hard X-rays over a broad spectrum (~ 1 –100 keV), from a compact accelerator and laser system. R&D is needed in high repetition rate, high power lasers, high-brightness and high power low-energy accelerators, and their

integration, to enable user sources within the next 10 years. Additional R&D could lead to spatially coherent X-ray production using Compton scattering – an approach that is in its infancy.

Lasers can provide high electric fields for accelerating particles in small structures (the optical extension of high-frequency rf acceleration) and to provide very short period undulators for X-ray generation. Laser-driven vacuum structures are at a very early stage of development. They offer potential for accelerating gradients \sim GeV/m or more, and for extremely high peak power X-rays from the short pulses naturally associated with optical wavelength devices. The horizon for implementation is greater than 10 years, with major challenges in sub-fs synchronization, materials damage from both the laser beam and electron beams, charging of structures by beam halo, and diagnostics for attosecond beams. Several key concepts still need proof-of-principle experiments. There may be a specific facility niche (low cost, low yield, small size) that this technology will be able to address better than existing technologies.

Laser-plasma accelerators (LPA) offer extremely high accelerating gradients (10–100 GV/m), intrinsically short electron bunches (1–10 fs), and intrinsic synchronization at the femtosecond time scale, with potential for compact X-ray FELs, directed gamma-rays, and coherent THz production in a fully synchronized hyper-spectral source. Electron beams of up to \sim 10pC charge, 50 fs duration, few % energy spread, 0.1 mrad divergence, and up to \sim 1 GeV energy have been produced by several research groups world wide. THz and incoherent X-ray production has been demonstrated. Soft X-ray production using a FEL and other processes is the topic of multiple experiments now in progress. Demonstration of a LPA-driven FEL at 1 nm and the design of a LPA-driven FEL user facility is expected within \sim 10 years. Substantial R&D is required in lasers, creation and tailoring of plasma channels, injection and acceleration schemes, diagnostics, compact undulators, and 3D simulation codes.

Enabling instrumentation and technologies

This working group identified several critical areas of crosscutting instrumentation and technologies: (1) attosecond instrumentation, (2) photocathodes, (3) insertion devices, (4) lasers, and (5) photon detectors. Development of insertion devices and photon detectors benefits all future and existing light sources, while attosecond instrumentation, high repetition rate lasers, and photocathodes are relevant to at least two facility architectures of the future. Developments are required in X-ray optics to preserve ultra-fast structure in beam splitters, etalons and other optical components, and to avoid damage from high peak power. Moreover, advances in metrology for beam diagnostics will be needed for femtosecond electron and photon pulses and low charge electron beams.

High precision timing techniques are required for synchronization of accelerator and laser systems, and for high-resolution measurement of electron and photon beam pulse duration and timing. The current state of timing and synchronization technology is \sim 10 fs over \sim 100 m distances. R&D is required to extend stability of timing and synchronization systems to 0.1 fs in \sim 5 years (over distance \sim 300 m), and distance scales up to 10 km in \sim 10 years.

R&D in photocathode materials and surface science is essential over the next few years to develop reliable and long-lifetime structures with high quantum efficiency at visible wavelengths, for efficient operation at very high repetition rates (1–1000 MHz) and for understanding of beam dynamics near the cathode surface to produce low intrinsic emittance electron beams.

Short-period undulators, potentially using superconducting technology to avoid moving parts, could extend the photon energy reach, or investment in the accelerator to reach sufficient beam energy to radiate at a given wavelength.

Laser technology is rapidly evolving and developments would be beneficial not only for HHG sources in the EUV/XUV range and laserplasma based sources, but also for photocathode systems and seeding for FELs. Major investments are being made overseas with industrial spin-offs and training of next generation scientists and technologists as a result. Substantial developments are needed in high-power lasers up to the kW level (and beyond for acceleration applications), and can be expected in the \sim 10 year timeframe. However, optical technology being developed for commercial and military applications no longer parallels the needs of the DOE Office of Science. Therefore, DOE sponsored laser research is crucial.

Advanced photon detectors could markedly extend the scientific reach of existing storage ring sources and will be essential to the full utilization of future light sources. Relevant R&D includes materials development, microelectronics design and fabrication capabilities, and radiation hardness.

William A. Barletta
Massachusetts Institute of Technology, Cambridge MA

John N. Corlett
Lawrence Berkeley National Laboratory



Free electron lasers: Present status and future challenges

W.A. Barletta^{b,*}, J. Bisognano^{i,*}, J.N. Corlett^e, P. Emma^h, Z. Huang^h, K.-J. Kim^a, R. Lindberg^a, J.B. Murphy^{g,**}, G.R. Neil^j, D.C. Nguyen^f, C. Pellegrini^c, R.A. Rimmer^j, F. Sannibale^e, G. Stupakov^h, R.P. Walker^d, A.A. Zholents^a

^a Argonne National Laboratory, 9700 S. Cass Avenue, Argonne, IL 60439, USA

^b Department of Physics, Massachusetts Institute of Technology, Bldg. 26-563, 77 Massachusetts Avenue, Cambridge, MA 02139-4307, USA

^c Department of Physics and Astronomy, University of California-Los Angeles, 405 Hilgard Avenue, Los Angeles, CA 90095, USA

^d Diamond Light Source Ltd, Diamond House, Harwell Science and Innovation Campus, Didcot, Oxfordshire OX11 0DE, UK

^e Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

^f Los Alamos National Laboratory, P.O. Box 1663, Los Alamos, NM 87545, USA

^g National Synchrotron Light Source, Building 725C, Upton, NY 11973, USA

^h SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025, USA

ⁱ Synchrotron Radiation Center, 3731 Schneider Dr., Stoughton, WI 53589-3097, USA

^j Thomas Jefferson National Accelerator Laboratory, 12000 Jefferson Avenue, Newport News, VA 23606, USA

ARTICLE INFO

Article history:

Received 11 February 2010

Accepted 28 February 2010

Available online 20 March 2010

Keywords:

Free electron laser

FEL

Amplifier

Oscillator

SASE

HGHG

XFEL

ABSTRACT

With the scientific successes of the soft X-ray FLASH facility in Germany and the recent spectacular commissioning of the Linac Coherent Light Source at SLAC, free electron lasers are poised to take center stage as the premier source of tunable, intense, coherent photons of either ultra-short time resolution or ultra-fine spectral resolution, from the far infrared to the hard X-ray regime. This paper examines the state of the art in FEL performance and the underlying enabling technologies. It evaluates the state of readiness of the three basic machine architectures—SASE FELs, seeded FELs, and FEL oscillators—for the major X-ray science user facilities on the 5–10 years time scale and examines the challenges that lie ahead for FELs to achieve their full potential throughout the entire spectral range. In soft and hard X-rays, high longitudinal coherence, in addition to full transverse coherence, will be the key performance upgrade; ideas using laser-based or self-seeding or oscillators can be expected to be qualitatively superior to today's SASE sources. Short pulses, from femtoseconds to attoseconds, can be realistically envisioned. With high repetition rate electron sources coupled to superconducting radiofrequency linear accelerators, unprecedented average beam brightness will be possible and many users would be served simultaneously by a single accelerator complex.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

The recent successful commissioning [1] of the Linac Coherent Light Source (LCLS) [2] hard X-ray (1.5 Å) free electron laser (FEL) amplifier is a culmination of nearly 35 years of continuous advances in electron beam and FEL physics since the first operation of infrared FELs ($\lambda = 10.6 \mu\text{m}$ amplifier and $3.4 \mu\text{m}$ oscillator configuration) in the late 1970s [3,4]. With some further development, free electron lasers are now poised to take center stage as the premier source of tunable, intense, coherent photons of either ultra-short time resolution or ultra-fine spectral resolution, from the far infrared to the hard X-ray regime. FELs can provide an unprecedented enhancement by as much as 10^{11}

in the peak brightness of the photon beams or by 10^6 in average brightness compared to the highly productive storage ring based sources of synchrotron radiation as shown in Fig. 1 [5]. These revolutionary photon probes, with femtosecond time resolution or meV spectral resolution, will foster entirely new avenues of discovery science in such diverse fields as biology, chemistry and material science. In this paper we survey the state of the art in FEL performance and the underlying enabling technologies and also examine the challenges that lie ahead for FELs to achieve their full potential throughout the entire spectral range.

1.1. FEL fundamentals

In a free electron laser the kinetic energy of a relativistic electron beam is transformed into an intense beam of electromagnetic radiation by wiggling the electrons transversely in a periodic magnetic field known as an undulator. Motz [6] and

* Corresponding author.

** Principal corresponding author.

E-mail address: jbm@bnl.gov (J.B. Murphy).

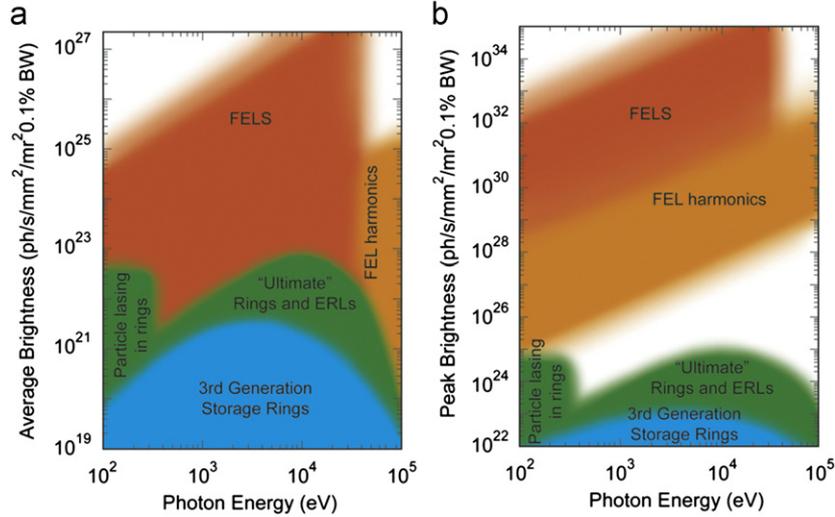


Fig. 1. Average and peak brightness versus photon energy for light sources [5].

Phillips [7] were among the earliest to realize the advantages of the undulator scheme together with relativistic electron beams to break free of the limits of conventional microwave tubes, which required some physical length of the tube to scale proportionally with the output wavelength. The interest in this type of device for short wavelength applications took off with the seminal work of Madey [8,3,4], who coined the name “free electron laser” for such devices [9].

The output wavelength of the FEL radiation at the n th harmonic in a planar undulator is given in terms of the undulator period, λ_u , the relativistic factor, $\gamma \equiv E/mc^2$ and the undulator strength parameter, $K \equiv eB_0\lambda_u/2\pi mc$ by

$$\lambda_n = \frac{\lambda_u}{2n\gamma^2} \left(1 + \frac{K^2}{2} \right), \quad n = 1, 3, 5, \dots \quad (1)$$

where E , e and m are the energy, charge and mass of the electron, respectively, c is the speed of light and B_0 is the peak magnetic field in the undulator with $B(z) = B_0 \cos(2\pi z/\lambda_u)$. The FEL output wavelength scales as $\lambda \propto \lambda_u/2\gamma^2$, and is tunable over a very large range by varying the electron energy, γ , or the magnetic field strength in the undulator by adjusting the gap between the magnetic poles.

FELs can be operated in three different modes: the self-amplified spontaneous emission (SASE), the seeded high-gain amplifier, and the oscillator. In the first two modes, the FELs operate in the high gain regime while the oscillators operate in the low gain regime.

1.1.1. High gain FELs

In both the self-amplified spontaneous emission (SASE) [10,11] mode of FEL operation and the seeded high gain amplifier mode, the electron beam amplifies the initial spontaneous undulator radiation emitted in the early portion of the undulator or the externally injected seed radiation, respectively. As the electrons and the radiation co-propagate through the undulator, the radiation slips through the electron beam in accordance with the resonance condition, so that the radiation field grows exponentially along the undulator, $P \propto \text{Exp}[z/L_g]$, if the electron beam and undulator are of sufficient quality.

The scaling behavior of a high gain amplifier FEL amplifier in the one dimensional, cold beam limit can be well characterized by the so-called one dimensional FEL ρ -parameter or so-called

Pierce parameter [11,12],

$$\rho \equiv \frac{1}{2\gamma} \left[\frac{I_{peak} \lambda_u^2 K^2 [JJ]^2}{I_A 8\pi^2 \varepsilon_x \beta_x} \right]^{1/3} \quad (2)$$

where I_{peak} is the peak current, $I_A \approx 17$ kA is the Alfvén current, ε_x is the electron beam emittance, β_x is the betatron function characterizing the beam envelope and $[JJ]$ is the usual Bessel function coupling factor for a planar undulator. The subscript “x” on ε_x and β_x refers to the horizontal plane; however, in the case of FELs the vertical plane behaves in a similar fashion. The one dimensional power gain length is given by

$$L_g \approx \frac{\lambda_u}{4\pi\sqrt{3}\rho}. \quad (3)$$

Various factors reduce the FEL performance from the ideal, cold beam limit listed above, such as electron beam energy spread, emittance, and magnetic field errors. The gain length in Eq. (3) is evaluated for a beam with no energy spread. Energy spread prevents bunching of all electrons at the same phase, because of the smearing out of the bunching due to variation in longitudinal velocities. The bandwidth in energy, which contributes to the FEL performance, is given by the FEL parameter ρ . To keep the increase in the gain length reasonable, the initial rms energy spread must satisfy [11],

$$\frac{\sigma_\gamma}{\gamma_R} \ll \rho. \quad (4)$$

Following the same argument, the deviation of the mean energy from the resonant energy is limited to

$$\left| \frac{\langle \gamma \rangle - \gamma_R}{\gamma_R} \right| \ll \rho. \quad (5)$$

This condition must be further refined in the case of seeded High Gain Harmonic Generation (HGHG) FELs where, to first approximation, this condition must be applied locally, not to the value averaged over the full bunch length. Thus there is a requirement on the maximum deviation from linearity of the energy distribution along the bunch.

Additional limitations are related to the transverse position and momentum spread in the electron beam, measured by the emittance (ε_x) and by the beam focusing elements along the undulator needed to preserve the electron density and thus the ρ parameter. Electrons oscillate around the undulator axis with a periodicity much larger than the undulator period. However, the betatron oscillations in the focusing system change the transverse

electron momentum, and thus the longitudinal momentum. Hence the spread in the betatron oscillation, which scales with the emittance, has an impact similar to an energy spread [13]. The requirement for the transverse emittance is

$$\varepsilon_x \ll \frac{\beta_x \lambda}{\lambda_u} \rho. \quad (6)$$

Note that the optimum value for β_x is driven by two competing needs: increasing the electron density, which makes the FEL parameter larger (by decreasing β_x), and minimizing the beta function's degrading effect on the synchronism condition (by increasing β_x). A rough estimate for the optimum β_x function is given by

$$\beta_x \approx L_G \approx \frac{\lambda_u}{4\pi\rho}. \quad (7)$$

Inserting into Eq. (6) shows that the beam emittance (ε_x) has to be smaller than the photon beam emittance

$$\varepsilon_x < \frac{\lambda}{4\pi}. \quad (8)$$

This condition can also be seen as a matching of the electron and photon beams' transverse phase space.

Because of diffraction effects, the radiation field escapes from the electron beam in the transverse direction. The field intensity at the location of the electron beam is reduced and the FEL amplification is inhibited. The compensation for field losses due to diffraction is the FEL amplification itself. After the lethargy regime [11] the FEL achieves equilibrium between diffraction and amplification. The transverse radiation profile becomes constant and the amplitude grows exponentially. This "quasi-focusing" of the radiation beam is called "gain guiding" [14], and the profile of the radiation field is an eigenmode of the FEL amplification [15–17].

Similar to the eigenmodes of the free-space propagation of a radiation field (e.g., Hermite–Gauss modes), there is an infinite number of FEL eigenmodes. Each couples differently to the electron beam and, thus, they have different growth rates or gain lengths. The mode with the largest growth rate dominates after several gain lengths and the radiation field becomes transversely coherent. At saturation, gain guiding vanishes and the electron beam radiates into multiple modes. Typically, the fundamental FEL eigenmode is similar to that of free-space propagation and the resulting reduction in transverse coherence at saturation is negligible.

The characteristic measures for diffraction and FEL amplification are the Rayleigh length Z_R and gain length L_G . To calculate the Rayleigh length we approximate the radiation size at its waist by the transverse rms electron beam size (σ_x) as the effective size of the radiation source, giving $Z_R = 4\pi\sigma_x^2/\lambda$ [18]. For $Z_R \ll L_G$, the FEL amplification is diffraction limited with a gain length significantly larger than the 1-D value. In the opposite limit, $Z_R \gg L_G$, the 1-D model remains valid.

At saturation the peak power and bandwidth of the FEL radiation, and the energy spread of the electron beam are given in terms of ρ ,

$$P_{peak}^{sat} \approx \rho E I_{peak} \quad (9)$$

$$\frac{\Delta\omega}{\omega} \approx \rho \quad (10)$$

and

$$\frac{\sigma_E}{E} \approx \rho. \quad (11)$$

SASE output exhibits fluctuations due to the inherent SASE process as well as electron beam fluctuations. The linewidth of the SASE output is typically $\sim 10^{-3}$.

The characteristics of the radiation emitted by the electron beam acting as a gain medium can be improved with respect to producing narrower line width radiation or increasing the temporal (longitudinal) coherence of the radiation or both. The two practical alternatives to improved characteristics are (1) to inject the gain medium (power amplifier) with a co-propagating master oscillator seed signal from a phase coherent source such as a laser, or (2) to have the gain medium (the beam) traverse an optical resonator to form an FEL oscillator. SASE has been important in the development of hard X-ray FELs because amplifiers and oscillators are difficult in this spectral region due to the absence of suitable seed radiation for the former and of suitable high-reflectivity mirrors for the latter.

In the seeded configuration, the temporal coherence is determined by the properties of the master oscillator or seed laser. Seeded amplifiers are combined with various harmonic generation schemes to produce short wavelength coherent radiation. With the advances in FEL and laser harmonic generation techniques, the soft X-ray region can be reached by a seeded harmonic generation amplifier scheme. A further discussion of practical limitations of amplifier performance in VUV and X-ray FELs can be found in Ref. [19].

1.1.2. Low gain FELs: oscillators

Oscillator FELs are well understood, both practically and theoretically. In an oscillator, one of several optical pulses are trapped within an optical cavity, and the FEL gain occurs each time the trapped optical pulse meets and travels together with an electron bunch through an undulator located within the optical cavity. Oscillators are feasible from the IR to hard X-ray regions by employing suitable reflectors for the cavity, which may be multilayer dielectrics, metals at grazing incidence, or Bragg crystals [20] as appropriate for the wavelength region in which the FEL operates.

The threshold condition for an oscillator is

$$(1 + G_0)R > 1 \quad (12)$$

where G_0 is the initial single pass gain and R is the total round-trip reflectivity. If this condition is satisfied then the intensity of the intracavity X-ray pulse increases exponentially with a concomitant reduction in gain. The system will reach saturation when the term $(1 + G)R$ is equal to unity (where G is the intensity-dependent gain).

In the cold beam limit,

$$G_0 = 0.22(2\pi\rho N_u)^3 \quad (13)$$

where N_u is the number of the undulator periods [21]. Although this equation is usually adequate for an IR FEL oscillator, a more refined expression is necessary for hard X-ray FELs since the effects of the electron beam emittance, energy spread and the diffraction are significant [22]. FEL oscillators at soft and hard X-ray wavelengths are discussed in Sections 4.5 and 5.4, respectively.

2. Laser manipulation and seeding

The term laser manipulation (LM) is often used in context of beam physics when laser light is applied to modify the electron beam distribution in phase space. The role of the laser is typically limited to providing an energy modulation of the electrons and that is accomplished by co-propagating the laser beam together with the electrons in the wiggler magnet. This mechanism is also

known as an Inverse Free Electron Laser (IFEL) [23]. For a given laser pulse energy A_L , optimal pulse duration $\sigma_\tau \approx 0.25N\lambda/c$ (where N is a number of wiggler periods and λ is the laser wavelength), and optimal focusing with the Rayleigh length $Z_R = (0.15 - 0.3)L_w$ (where L_w is the wiggler length), the amplitude of energy modulation can be estimated as $\Delta E \approx 2\sqrt{A_L A_s}$ [24], where $A_s \approx 5\pi e^2/\lambda$ is the energy of an electron spontaneous emission into the fundamental undulator mode for a wiggler with a large K and e is the electron charge [25].

One example of LM is “slicing”, currently used for selection of ultra-short X-ray pulses [26–28] in storage ring synchrotron light sources. Another example is the “laser heater” [29–31], where the radial dependence of the laser intensity is used to achieve electron energy modulation dependent on electron transverse coordinate and facilitate mixing of electrons within neighboring slices of the electron beam. A proposal for a current enhanced SASE [32] takes advantage of LM by producing a large energy modulation of electrons during the earlier part of the acceleration at lower beam energy and converting it into density modulation at the undulator entrance, with a significant increase of the peak current using a magnetic chicane. This reduces the FEL gain length and permits synchronizing pump and probe sources, e.g., an optical pump pulse (or any other signal derived from an optical signal) and X-ray probe pulse or *vice versa*. Electron beam conditioning [33] has long been a dream, carrying the promise of relaxing emittance requirements for the electron beam and, thus, improving the FEL performance. It was recently proposed [34] that, perhaps, this problem could also be solved by employing LM.

The external seeding idea has much in common with LM and had been proposed for FELs much earlier than any of the ideas involving LM [32,35–38]. Seeded FELs employing the process of High Gain Harmonic Generation (HGFG) [38] use a laser to modulate the energy of electrons in the first undulator, then convert it into density modulation using a magnetic chicane, resulting in a relatively large microbunching of electrons at a high harmonic n of a laser frequency, and produce amplified radiation in the downstream undulator tuned on the FEL resonance at a harmonic frequency. This technique has been proposed as a marked improvement to SASE FELs capable of producing laser-like X-rays with a time-bandwidth product approaching the Fourier transform limit. A proof of principle experiment [39] has demonstrated substantially increased spectral brightness of an HGFG FEL compared to a SASE FEL and much better wavelength stability. Typically, the harmonic number is limited to less than five due to debunching caused by the incoherent energy spread of the electrons. Thus, further extension of the HGFG technique consists of adding two or more cascades. In this case, a fresh part of the electron bunch can be used in each cascade [38]. Besides the technical complexity of a multi-staged FEL, the amplitude and phase noise in the initial laser seed will set a limit on how many cascades are practical [40]. Thus, extending the seeding technique into a hard X-ray range seems unlikely and self-seeding [41] could be the only seeding option in this case. Several studies of HGFG FELs with two or more cascades have been carried out [42,43]. The first FEL to be built with more than one cascade will likely be the FERMI@Elettra FEL [42].

The process of High Harmonic Generation (HHG) in gases [44] and solids [45] is also considered as a source for the Fourier transform limited seed signal. In that case the FEL just amplifies the seed if it is already at the requisite frequency [46]. Alternatively one can use it in the same way as the laser seed in a HGFG cascade. In the first case the power of the HHG seed signal must be substantially larger than the shot noise power of the electrons $P_{\text{Noise}} \approx \pi\alpha\omega\hbar I_{\text{peak}}/e$, where α is a fine structure constant, \hbar is Planck’s constant, e is the electron charge, ω is the seed frequency, and I_{peak} is the electron peak current. In the

second case it must be large enough, e.g., often at a few-megawatt level, to be able to induce energy modulation of electrons at least n times larger than the electron energy spread. A combination of the two cases could be a process [47] in which, first, the HHG seed is amplified in a configuration using an optical klystron, and then it is used with fresh electrons in an HGFG process.

Echo-Enabled Harmonic Generation EEHG [48,49] is a different technique that uses two consecutive sections, each of which comprises an energy modulator and a dispersion section. The echo scheme has remarkable up-frequency conversion efficiency and allows for generation of high harmonics with a relatively small energy modulation. In general, the frequencies of the first, ω_1 , and the second, ω_2 , modulators can be different. The beam modulation is observed at the wavelength $2\pi/k_{\text{echo}}$, where $ck_{\text{echo}} = n\omega_1 + m\omega_2$, with n and m integer numbers. The first dispersion section is chosen to be strong enough so that the energy and density modulations induced in the first modulator are macroscopically smeared due to slippage. At the same time, this smearing introduces a complicated fine structure into the phase space of the beam. The echo then occurs as a recoherence effect caused by the mixing of the correlations between the modulation in the second modulator and the structures imprinted onto the phase space by the combined effect of the first modulator and the first dispersion section (Fig. 2). The key advantage of the echo scheme is that the amplitude of high harmonics of the echo is a slow decaying function of the harmonic number. Simulations show that harmonic number $n=50-100$ might be attainable using the EEHG approach.

All seeding options are very sensitive to fluctuations in the electron beam energy and energy chirp when it comes to attaining a narrow bandwidth in the output signal of the order $10^{-4}-10^{-5}$. These fluctuations are caused by (1) phase and amplitude fluctuations in the rf field, (2) jitter in the electron bunch launching time, and (3) fluctuations in the electron bunch charge affecting energy chirp via the wakefields. We also note that the quadratic component in the energy chirp, e.g., d^2E/dt^2 , causes broadening of the bandwidth, and the effect is larger for FELs with

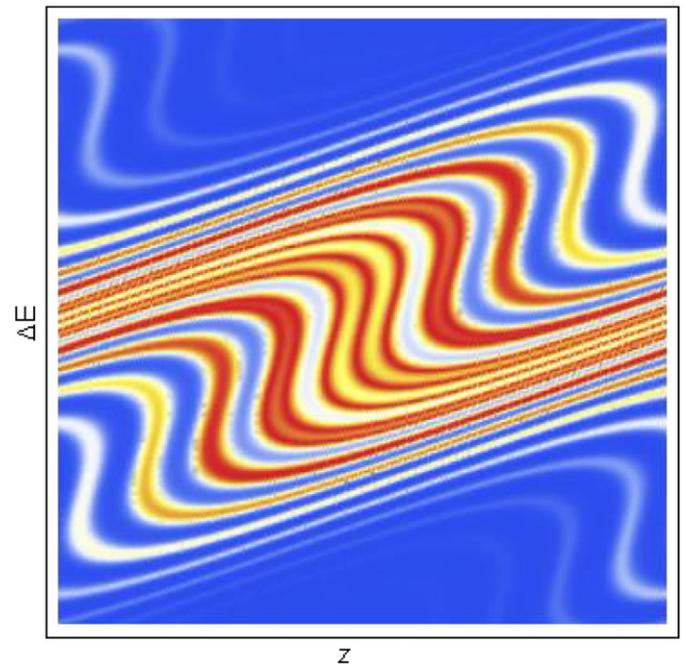


Fig. 2. The longitudinal phase space of the electron bunch showing the microbunching obtained by echo-enabling manipulations.

a larger overall time-off-flight parameter R_{56} . In this regard the microbunching instability (see Section 11) creating small but rapid variations of the electron beam energy along the electron bunch can significantly broaden the bandwidth.

3. IR and UV FELs

3.1. Introduction

The development of future soft and hard X-ray FEL user facilities rests on a long history of accomplishment in the longer wavelength regimes. In particular the experimental confirmation of FEL theory and the benchmarking of codes in the IR, and later into the visible and UV, has permitted a rational design process leading to the recent successes achieved at X-ray wavelengths. The more fundamental reason for the establishment of a large number of FEL facilities around the world is that FELs can yield photon beams with unique characteristics that are not available from conventional sources.

A list published annually by the International FEL Conference [50] shows more than three dozen FELs demonstrated (as of 2009) in the FIR to UV range and a number of proposed devices. Wavelengths of the operating devices ranged from 1.2 to 13.1 nm with applications as varied as their operating range. In the longer wavelength regime, FELs are used in pump-probe studies of molecular dynamics since the frequencies can be chosen to match myriad vibrational modes of organic and other molecules. Such long wavelengths also match well as probes of magnetoresonances, plasmon coupling and electron spin. Refs. [51–53] explore the science case for such facilities.

While the majority of these devices are oscillators based on pulsed, room temperature rf accelerators, a sprinkling of other accelerator types are in the mix, including electrostatic accelerators (1) [54], microtrons (1), and storage rings (7 facilities). Five facilities use superconducting rf acceleration, a technology which is expected to play a major role in future X-ray user facilities. Only five short-wavelength FELs were amplifiers or used SASE because in this wavelength region high-reflectivity mirror technology has not been available to meet the stringent optical tolerances required of the FEL. While it is beyond the scope of this paper to explore the full range of capabilities and applications of FELs in this wavelength region, it is useful to look at a small number of representatives in more detail as a basis for our extrapolation of this technology into the X-ray region.

3.2. Room temperature RF accelerator based oscillator facilities

A large number of pulsed copper rf accelerator based FELs in the world operate as user facilities. A major fraction of these are university facilities due to the ease with which pulsed, copper rf accelerators can be incorporated into an FEL. Many of the facilities have S-band linacs producing 20–30 MeV and utilize rf guns to produce macropulses of 5 micro second at ~ 10 Hz. The beams are sent to short (~ 1.5 m) permanent magnet undulators surrounded by a broad band optical cavity with hole out-coupling. The following two examples illustrate the sort of activities that are underway.

One of the most productive IR user facilities in the world is FELIX at FOM in the Netherlands [55]. At FELIX two lasers on the system have been extensively used for studies in solid state dynamics, atomic clusters, and magnetic materials. FELIX offers lasers from 16–250 μm or 4–30 μm from its two FELs at 10 Hz with 100 mJ per macropulse. The machine is stable enough that

the users typically operate it themselves. The facility has successfully enhanced the utility of its photon beams by introducing innovations such as synchronized pulse slicing of individual micropulses as well as other synchronized lasers. Among the substantial set of publications reviewed on their website [56], a particularly interesting result has been the identification of titanium carbide as a constituent of interstellar plasmas. An upgrade presently in construction will permit high field, intracavity tests at an internal focus to permit the study of a number of non-linear phenomena.

The Budker Institute in Novosibirsk operates a high power FEL in the mid-IR to FIR range for materials research. The Budker system incorporates energy recovery of over 30 mA of average current and recently achieved two-pass acceleration; they are on their way to a five-pass recirculation up to 80 MeV followed by five-pass deceleration [57] with multiple undulator systems. The system has produced over 400 W of average power at 60 μm and now operates down to 15 μm .

3.3. IR, visible and UV FEL amplifiers

FEL amplifiers typically rely on high brightness electron sources together with magnetic bunch compression to generate high current density bunches with peak currents in the range of $I_{\text{peak}} \sim 200\text{--}1000$ A to provide for exponential high gain amplification of either initial shot noise in the electron beam (SASE FEL) or an initial coherent seed laser pulse (seeded FEL). The first high-gain SASE amplifier experiments were the ELF experiments performed on an induction linac at LLNL at Ka-band wavelengths [58]. These were quickly followed by experiments at LANL using an L-band copper integrated injector/linac with a Cs_2Te photocathode to demonstrate a gain of 300 in a 1-m undulator and later, a gain of 3×10^5 in a 2-m undulator by a consortium of UCLA, LANL and RRR [59,60]. However, most of the recent high gain amplifiers (LEUTL, ATF, SDL, etc.) have been based on RF photoinjectors with Cu or Mg cathodes to generate beams in Cu S-Band linacs with rep rates on the order of 2–100 Hz and high gains of $10^5\text{--}10^6$ [61–64]. In fact, the LCLS is a straightforward, albeit extremely challenging, scaling of the successful IR/VUV machines into the X-ray regime [1].

Virtually all FELs produce light with full transverse coherence in a single mode with $M^2 \sim 1$, where M is the number of transverse modes as defined in Ref. [65]. In the IR/VUV regime the use of coherent seed lasers permits the generation of light that is transform limited ($\Delta t_{\text{rms}} \times \Delta \omega_{\text{rms}} \approx 1/2$) or longitudinally coherent. Amplification of a titanium-sapphire seed laser has been demonstrated at the fundamental wavelength of 800 nm, and also at its harmonics of 400 and 266 nm [66]. High harmonic generation (HHG) seeding of an FEL amplifier has recently been demonstrated at 160 nm [67], which will be followed by several shorter wavelength HHG experiments ($\lambda \sim 30$ nm) in the near future. At IR/VUV wavelengths SASE is relegated to the realm of machine tune-up.

FELs based on harmonic generation were first demonstrated in the IR and VUV regimes [68,63,64]. In these devices an input IR seed laser ($\lambda \sim 800$ nm) modulates the energy of the electrons via FEL interaction in an undulator. The energy modulation of the e-beam is converted into a density modulation in a strongly dispersive magnet. The sharp, periodic density modulation yields rich harmonic content that initiates coherent harmonic radiation in a final stage undulator. High gain harmonic generation (HG) has been demonstrated for harmonic number $n=2\text{--}4$ down to $\lambda \sim 200$ nm [64]. Cascaded HG stages have the potential for extending the attractive properties of seeded FELs to the soft X-ray regime [69,70].

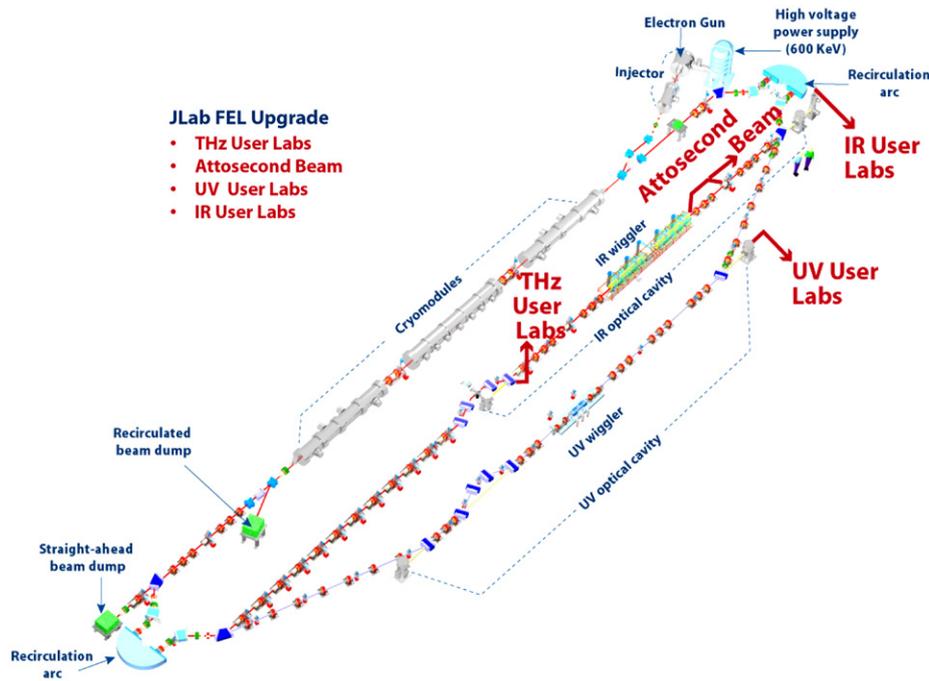


Fig. 3. JLab 10 kW upgrade FEL.

3.4. Superconducting RF (SRF) accelerator IR/UV oscillators

An example of the sort of operation made possible by SRF technology is given by the JLab IR/UV Demo FEL (Fig. 3). This machine has undergone several upgrades and improvements since its initial operation in 1996. Its present configuration has delivered up to 14.3 kW of continuous average power at 1.6 μm and produced many kilowatts at wavelengths ranging out to 10 μm [71]. The electron beam is produced from a DC gun from a GaAs photocathode. Electron bunches of 135 pC are accelerated up to 160 MeV and undergo lasing in a 29 period permanent magnet undulator surrounded by a 32 m optical cavity. Up to 9.3 mA of average current has been accelerated. During lasing the electrons give up about 1.5% of their energy to photons but the full energy spread of the beam increases from 0.3% to nearly 12%. Despite this large energy spread the electrons are sent back through the linac a second time 180° out of rf phase so the beam decelerates back down below its injection energy before being dumped. This machine was the first high power demonstration of such energy recovery technology and has led to the design of (non-lasing) light sources based on energy recovery technology. The system was recently upgraded with an undulator and optical cavity optimized for UV lasing. The beamline has recently transported its first electron beam and the optical system will be installed in January 2010.

The continuous SRF accelerator can produce macropulses of light as short as 10 ms (the cavity filling time) at arbitrary repetition rates up to a continuous train of 130 fs pulses at 4.68–74.85 MHz. The optical pulses of the FEL have a measured bandwidth within 25% of the Fourier transform limit for each wavelength with a mode quality of $M^2 < 1.6$. Significant contributions to the technology of SRF-based light sources include studies of SRF acceleration, beam breakup control, coherent synchrotron generation, high power optics, and rf control. The FEL output can be directed to a number of user labs for studies. A wide array of topics that have or are being addressed include carbon nanotube generation, searches for dark matter (axions), pulsed laser vapor deposition, selective necrosis of fat cells, and

vibrational energy transfer of interstitial hydrogen in semiconductors [72–74].

3.5. Implications for next generation facilities

The broad success of these facilities establishes the experience base to extrapolate FEL operation into the much more demanding X-ray regime. Although these systems are much smaller and simpler than those envisioned for a next generation user facility in the X-ray regime, they have come to grips with key technology challenges and bring the relevant performance near to, if not already meeting, technical requirements of the next generation. As the FEL itself is often the best diagnostic of the accelerator performance, systems with reliable and productive experimental programs point to useful future approaches. Key technologies needing further development include optics, optimized SRF accelerator designs, and the crucial injector technology. While pulsed injectors meet our present requirements, achieving the same performance in a continuous bunch train remains an R&D goal with several groups pursuing possible approaches. Physics issues that must be resolved for success of a next generation facility at EUV to X-ray wavelengths include coherent synchrotron radiation (CSR), longitudinal space charge effects, wakefield and resistive wall effects, halo generation and mitigation, etc. The solid base of operating FELs constitutes a set of test beds available to provide the answers needed.

4. Soft X-ray FELs

4.1. Rationale for VUV-soft X-ray FEL user facilities

4.1.1. Science drivers

The history of accelerator-based light has seen ever increasing demands on the brightness and intensity of the photon sources. The scientific driver is to probe matter with substantially finer length, time, and energy resolution, where physical, chemical, and biological systems can be viewed on their critical temporal, spatial, and energetic scales—femtoseconds, nanometers, and

millivolts. Dynamics will be the theme of the future, complementing the studies of static systems that are the bread and butter of today's light sources.

Desired characteristics of next generation VUV-soft-X-ray light sources include full transverse and temporal coherence, pulse durations extending to femtoseconds and possibly attoseconds, high peak and average flux, rapid tunability of photon energy, and full polarization control. These features will enable a far reaching scientific enterprise. Breakthroughs can be expected, for example, in exploring the dynamics of catalysis and chemical transformation, understanding how correlations of electrons and spins create high T_c superconductors, and elucidating the remarkable functionality of complex biological systems. High brightness and coherence would enable exploitation of new imaging techniques such as coherent diffraction imaging and holography. Flash imaging (where femtosecond snapshots are taken before damage can occur) may enable new levels of understanding of sub-cellular and macro-molecule organization within cells. Short pulses with full temporal coherence would access new regimes of condensed matter, using time-of-flight methods, time resolved imaging, and inelastic X-ray scattering. One could pursue selective fabrication of atomic clusters and other nanostructures. Fundamental atomic/molecular science would be advanced with ultra-fast observations of electronic motions and spectroscopy of rare and exotic species. Ultra-brilliant X-ray sources will enable studies of strong-field and multiphoton physics, nonlinear dynamics, and out-of-equilibrium dynamics and would allow, for the first time, studies in the gas phase. The broad scientific case for fourth generation VUV-soft-X-ray sources has been the topic of numerous scientific workshops and studies. More complete expositions can be found in several recent reports [5,75].

4.1.2. Free electron lasers

Free electron lasers are uniquely suited to achieving the performance goals discussed above for the fourth generation VUV-soft-X-ray facilities because of the inherent transverse coherence of the FEL process. Moreover, at VUV to soft-X-ray wavelengths, laser seeding and harmonic up-conversion offer, in contrast to SASE, substantially improved temporal coherence (possibly approaching a few times the Fourier limit) plus good pulse-to-pulse repeatability and amplitude stability. Such FELs have the potential to deliver radiation with spectral characteristics mimicking those of the seed lasers, albeit at shorter wavelength. It appears possible to achieve very energetic (mJ) and short (~ 50 fs) output pulses, or else ultra-short (femtosecond to attosecond) output pulses, with bandwidth near the limit posed by the uncertainty principle. The seed lasers can be integrated with the low-level rf (LLRF) of the accelerator to enable overall synchronization with conventional pulsed pump or probe laser sources. For ultra-short pulses, the coherent radiation signals produced by the electron bunches at wavelengths in the visible region can also be used for synchronization. Laser seeding can also reduce the length of undulator to achieve saturation, and provide frequency stability that is a small fraction of the FEL bandwidth.

CW operation with a superconducting RF linear accelerator driver allows the high average beam power and high repetition rates demanded by many applications and the need to serve many

users simultaneously [76–78]. Room temperature linacs can be considered as FEL drivers at lower duty factors. They are less expensive, but can only operate at ~ 100 Hz repetition rates with high accelerating gradients. By lowering the gradients to ~ 5 MV/m, 10 kHz rates may be possible and would be attractive for a subset of proposed FEL-based research. However, such FEL-optimized designs require further study to establish practicability and costs. High-gain oscillator configurations can also be considered, as described in Sections 4.5 and 5.4.

4.2. Where we are today

4.2.1. Flash

FLASH at DESY (Hamburg, Germany) [79] is presently the world leading FEL source in the VUV and soft X-ray wavelength range. Its configuration and operational characteristics provide both a proof of principle for the viability of FELs as research tools and a starting point to develop the next generation of sources that are designed from their inception as user facilities. FLASH originates from the TESLA Test Facility, a test bed for linear collider development. The accelerator uses superconducting RF technology (a 1300 MHz, 1 GeV linear accelerator), but is operated in a long-pulse mode (≤ 0.8 ms and ≤ 800 pulses at 5 Hz) rather than CW. In the late 1990s a single-pass SASE free electron laser was added. FLASH has operated in SASE mode to produce transversely coherent radiation in 10–50 fs pulses in the wavelength range of 6.5–47 nm. The peak brilliance is 10^{29} – 10^{30} ph/s/mm²/mrad²/0.1%bw and peak power is 1–5 GW with a spectral width of $\sim 1\%$. The FLASH layout, shown in Fig. 4, includes a pulsed photoinjector and bunch compression chicanes.

Since FLASH has a fixed gap undulator, a change of the photon wavelength requires a change of the electron beam energy, which has proven time intensive.

The capability for a seeded FEL, sFLASH, is being installed. It consists of a seed laser system, an undulator section of 10 m, and a photon beam line to transport the FEL radiation to an experimental hutch located outside the FLASH tunnel. Also, beam energy is being increased to 1.2 GeV, reducing the shortest wavelength to below 5 nm. Since summer 2005, almost 100 proposals for user experiments have been accepted, and experiments including diffraction imaging, condensed matter, plasma and cluster physics, femtochemistry, atomic physics, and molecular biology have been successfully carried out.

4.2.2. FERMI@Elettra

The FERMI@Elettra FEL complex in Trieste [70] will be the first user facility to be based on seeded harmonic cascade FELs. The seeded FELs provide high peak power pulses of controlled duration, allowing tailored output for time domain explorations with pulses of 100 fs or less, and high resolution with output bandwidths of the order of 5–10 meV. A high brightness RF photocathode gun will supply single beam pulses to an existing 1.5 GeV room temperature linac at repetition rates up to 50 Hz. FERMI will provide tunable output over a range from ~ 100 to ~ 4 nm. Its Advanced Planar Polarized Light Emitted (APPLE) undulators will allow control of the photon polarization. Initially, two FEL configurations are planned: a single, harmonic generation

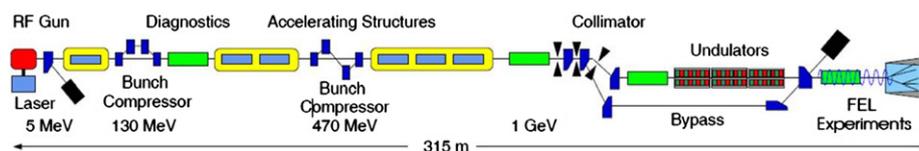


Fig. 4. FLASH facility.

stage to operate over ~ 100 to ~ 20 nm, and a two-stage cascade operating from ~ 25 to ~ 4 nm or shorter wavelength, each with spatial and partial temporal coherent output and peak power in the GW range. FERMI user operations are scheduled to begin in January 2011.

4.3. Facility architectures for the future

4.3.1. Technical specifications

Both FLASH and FERMI represent the first generation of coherent light sources, based on reconfiguration of existing facilities to operate FELs. To move forward to a user facility dedicated both in design and operation requires careful consideration of performance goals such as full coherence and time structure together with more practical issues such as serving multiple users and cost control. The typical photon range for a VUV-Soft X-ray facility would arguably be between 50 eV and 2 keV in the first harmonic, with user desires for “one-stop” research possibly requiring lower photon energies. A list of desirable features would include:

- small line width;
- temporal coherence;
- good amplitude stability;
- short pulses—femtoseconds, possibly shorter;
- high repetition rate;
- regularly spaced pulses;
- variable polarization, both linear and circular;
- multiple, simultaneous users;
- multicolors;
- tunability in real time;
- synchronization for pump/probe;
- parasitic terahertz radiation.

Overall, these user requirements translate into accelerator beam specifications that are well matched to the capabilities of CW superconducting linear accelerators. While subpicosecond electron pulses are beyond the reach of storage rings, they are routinely achievable by linear accelerators. With their weaker wakefields due to relatively larger apertures, L-band SRF linacs can maintain low energy spread at high bunch charge more easily than room temperature linacs. Thus the superb emittances generated by state of the art electron guns can be preserved even at relatively high charge. CW operation allows high average beam power and repetition rates of tens of kilohertz to megahertz required for some applications or for multiple simultaneous users. However, substantial value engineering of superconducting technology will be required to keep capital and operating costs per experiment reasonable. Recirculation has been considered to reduce costs, but beam quality preservation may be marginal in the recirculation arcs because of coherent synchrotron radiation (CSR) [80–82] and wakefield degradation.

A complementary approach aimed at a smaller scale, for sources that could be sited at more locations than just one or two major national/international centers, is based on X-band (8–12 GHz) accelerator technology and is described in more detail in Section 9. Due to the relatively small size of the structure (compared to S-band room temperature linacs at ~ 3 GHz or L-band SRF linacs at ~ 1.5 GHz), kilohertz repetition rates may be realizable. When combined with low charge (few 10 pC), ultra-low emittance sources (~ 0.1 mm mrad) recently demonstrated on the LCLS, such FELs could address an important subset of experiments using ultra-short X-ray pulses. They could also serve as test beds for more expansive concepts. This approach requires

further work to demonstrate performance limits, flexibility, and cost effectiveness.

4.3.2. Design of a CW FEL facility

A defining characteristic of next-generation VUV-Soft X-ray FEL light sources is fully coherent radiation, both transversely and temporally, together with short pulses. For soft X-rays several seeding approaches are available. As developed at BNL [83], seeding with a conventional laser starts the FEL process by modulating the energy of the electron beam on passage through an undulator tuned to the laser wavelength. After passing through a buncher, higher harmonics in the current density are generated, which are amplified and radiate in a subsequent undulator tuned to a harmonic. The resulting radiation has substantially improved temporal coherence, mimicking the coherence of the conventional laser seed. Furthermore, this process can be cascaded to offer several stages of up-conversion in photon energy.

Since these original experiments, the process of high harmonic generation (HHG) [84] using IR laser ionization of a noble gas can produce sufficient power to seed at much shorter wavelengths, ~ 30 nm and potentially lower. Seeding has the added advantage shortening the final amplifier undulator with respect to the SASE case. More flexible X-ray pulse intensity control might be possible, since the FEL need not be driven into saturation to generate reproducible radiation.

Some designs employ a “fresh bunch” technique [85], in which substantial beam degradation develops as the FEL reaches saturation in a cascade stage. To continue cascading, the timing of the longer electron bunch is slipped relative to the photon pulse to offer previously unperturbed electrons for continued FEL amplification. Since the overall peak current must be high (~ 1 kA), nanocoulomb bunch charge is required. An alternative approach utilizing a modified modulator/radiator scheme allows a cascaded FEL to be obtained without the need for the “fresh bunch” technique, thus simplifying the layout and allowing operation with shorter and lower charge electron bunches [77,78]. Low bunch charge could enable more simultaneous users to be supported for a given average current.

Another seeding method recently proposed, echo enhancement [48], utilizes two stages of laser seeding to filament the electron bunch longitudinally. This process, now being investigated experimentally, might generate very high harmonic content enabling production of keV photons with few (if any) cascades.

One potential layout of an SRF based FEL source is shown in Fig. 5. Principal components include a high repetition rate photocathode injector (~ 100 pC charge, ≤ 1 mm mrad normalized emittance), a superconducting L-band linac (~ 2.5 GeV for 1 keV photons in first harmonic, assuming conservative undulator parameters), bunch compressors to generate high peak current for FEL gain, and spreaders to simultaneously deliver beams to multiple FELs. The undulators would offer a variable gap or variable field to allow real time tuning without changing linac energy. With reasonable assumptions on beam current, laser and electron gun performance, and experimental configurations, a dozen FELs could be supported. At 1 keV photon energy, average brightness would exceed 10^{25} ph/s/mm²/mrad²/0.1%bw and peak brightness would exceed 10^{32} ph/s/mm²/mrad²/0.1%bw.

Pulse repetition rates as high as a megahertz appear ultimately feasible, if the electron gun can support them. Kilohertz rates represent present technology. High repetition rates depend on the success of R&D programs on CW electron guns and on high power seed and photocathode lasers. Both low frequency room temperature and high and low frequency SRF guns are under consideration.

An alternate approach to achieving high average brightness is to utilize an oscillator. The possible layout of such a system is

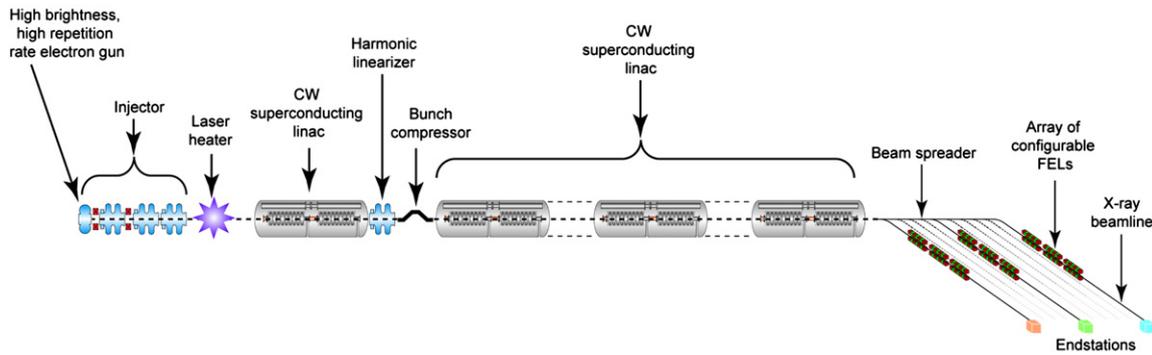


Fig. 5. Possible VUV-soft-X-ray user facility configuration.

similar to Fig. 5 with the addition of an optical cavity around each undulator. The oscillator approach would differ from most FEL oscillators operating in the IR-visible regime in that long undulators could provide gains of 1000 or more, thereby loosening the tight mirror reflectivity, optical figure and alignment tolerances in FEL oscillator systems. As oscillators typically produce nearly Fourier limited bandwidths as determined by the optical cavity configuration, the output spectrum would be determined by the electron bunch length. This feature may limit oscillator pulses to 50 fs to 1 ps range in contrast to the very short pulses that are possible with an amplifier; the associated advantage is that the bandwidth of an oscillator can be very small, of order 10^{-6} . Another limitation is that the optical pulse repetition period cannot be made substantially longer than the round trip time of the optical cavity, typically a microsecond. Therefore, very fast switching of the relatively low power (as compared with high gain amplifiers) output X-rays would be required.

The maximum beam currents envisioned for these light sources, 1 mA, are well within the range of current technology. RF phase and amplitude stability are tight (typical numbers are 0.02% and 0.03°), but seem feasible in light of the recent results obtained at Cornell on their ERL injector, with control of $<0.01^\circ$ in phase and 2×10^{-5} in amplitude [86]. For designs without recirculation, high order mode (HOM) damping is much less stringent than in energy recovery linacs (ERLs), which must accelerate substantially higher average currents. Although present design concepts assume existing L-band structures, the required bunch repetition rates (megahertz) would permit other frequency choices. A full study of the optimal frequency for cost and beam dynamics considerations is highly desirable.

4.4. Technical readiness and principal R&D challenges

The successful operation of FLASH has initiated the era of VUV-soft-X-ray FEL enabled science. Several options could be deployed immediately. For example, the European XFEL [87] project will generate multi-keV, hard X-rays using a higher energy extrapolation of FLASH in a pulsed SASE mode. One could construct dedicated room temperature C or S-band linacs duplicating the approach of FERMI@Elettra project, but at higher energies. A direct extension of current CW SRF linac and electron gun technology would allow construction of a kilohertz repetition rate, keV-photon-energy SASE FEL facility. The predictions of well-benchmarked computer simulations, and the present state of the art in conventional laser technology, support the immediate design of a seeded FEL operating at kilohertz repetition rates. The FERMI@Elettra project will confirm at least one seeding approach by early 2011.

A VUV-soft-X-ray FEL facility providing X-ray pulses to multiple users with full transverse coherence, substantially improved temporal coherence (possibly approaching a few times the Fourier limit), 100 kHz or higher repetition rate, and femtosecond pulses appears well within reach given a dedicated R&D program. A summary of several of the key R&D challenges is as follows:

- (1) The greatest challenge for FEL-related R&D is expanding the technology base to a level consistent with megahertz rate, multi-user facilities. With respect to beam physics, the most critical subsystems are the photoinjector and seed lasers necessary to take full advantage of the underlying potential of the CW superconducting linac. Presently, high brightness gun operation above the kilohertz level is not in hand. L-band and VHF SRF guns, VHF room temperature guns, and DC guns offer different strengths and weaknesses, and all appear viable with further development. Substantial cathode development is essential to greatly extend cathode lifetime and quantum efficiency. Also critical is the development of diagnostics for low-emittance/low charge electron beams that are capable of measuring the beam phase-space distribution of femtosecond to attosecond pulses of electrons and photons.
- (2) Alternative seeding and harmonic up-conversion processes (HGHC, echo, self-seeding, etc.) must be evaluated experimentally including determining the required tolerances to achieve the highest temporal coherence. Technical issues include control of the 6-dimensional phase space distribution of the electron bunch, the effects of non-linear energy chirp, propagation of laser phase noise in harmonic up-conversion, and jitter control. Oscillator designs in the VUV and soft X-ray region will also require extensive modeling and proof-of-principle experiments including the development of optical cavities for operation in the 50–500 eV region.
- (3) Laser systems for seeding and to drive photo-emission must operate at rates commensurate with the electron guns; approaches of interest include high harmonic generation (for seeding) and cryo-cooled crystals. Seed pulses must be fully characterized to fully uncover their effect on the FEL performance. Improved FEL simulations codes will be needed to evaluate the practical advantages and disadvantages of the many seeding scenarios.
- (4) Bunch compression, microbunching, and other collective effects, which can strongly influence the final X-ray pulse characteristics, must be managed to maximize performance and reduce costs. To furnish beam to multiple FEL lines for simultaneous operation, beam switchyard hardware for repetition rates exceeding ≥ 100 kHz needs demonstration. Studies of recirculation, advanced undulators and/or smaller gaps, and advanced harmonic schemes may allow lower linac

energies and lower costs. Optical elements for X-ray pulse shaping, transport and focusing, without spoiling the characteristic FEL properties, also require development.

- (5) A CW SRF linac represents the mostly costly technical systems of a high rep-rate facility. Existing pulsed SRF designs optimized for linear colliders are not cost-optimum for CW FEL applications. Optimized SRF structures for FELs that minimize operating and capital costs for CW operation require development. High order mode (HOM) damping requirements should recirculation prove viable must be determined. See Section 6 for further details.
- (6) The beam emittance obtained at LCLS shows that the undulator period is now a limiting factor in X-ray FEL design. Undulators with shorter period and sufficiently high K can be used to allow lower beam energy for fixed photon energy. The R&D on undulators would include both technological approaches such as in-vacuum, superconducting, and cryo-cooled devices together with understanding of the beam physics limitations on undulator gap. Novel approaches, such as microwave undulators, should also be studied.

4.5. VUV-soft X-ray oscillators

4.5.1. Rationale for VUV/soft X-ray FEL oscillators

Although the concept of a VUV/soft X-ray oscillator is not new [88–91], advances in both accelerator and optical technology have made feasible previously inaccessible approaches. To date storage ring FELs have been limited to wavelengths just extending into the VUV region, with 176 nm having been achieved at Elettra [92]. Due to the long pulses and low peak currents possible in storage rings, this architecture is unlikely to achieve the high brightness desired for a next generation source.

It is worth examining the potential advantages of such systems as compared to amplifiers. At visible and near UV wavelengths, where good mirrors are available, oscillators are the dominant approach. Oscillators eliminate the need for a seed laser with its attendant timing issues and complexities of continuous tuning. The oscillator output naturally grows to saturation with a near Fourier transform limited pulse producing an optical pulse that naturally extracts energy from a large fraction of the full electron pulse and (in the case of low gain systems) is relatively insensitive to the energy spread of the electron beam.

In principle, a soft X-ray oscillator based on a CW or long-macropulsed superconducting linac could provide the following desirable attributes in a user facility: small line width; temporal and transverse coherence; good amplitude stability; pulse lengths of 50 to perhaps 1000 fs; very high repetition rate (> MHz); regularly spaced pulses; variable polarization; both linear and circular; and continuous tunability in real time. These factors combine to produce a brighter, more coherent optical beam from a given electron beam, and specifically, given peak current. Against these advantages are balanced (1) the difficulty of supplying a suitable optical cavity with acceptable mirror reflectivities and (2) the necessity of a high repetition rate of order MHz, because the interpulse spacing cannot be made less than a small multiple of the optical cavity round trip period. Furthermore, the optical pulse length, which will be of order of the electron pulse length, is probably > 50 fs for the charge and energy spread required to achieve sufficient gain.

4.5.2. Concept studies

Limited by the properties of available electron injectors, early VUV-Soft X-ray oscillator designs resembled existing IR FELs in which the achievable small signal gain was limited to 10. Their performance was very sensitive to the mirror figure, alignment,

and detuning of the optical cavity. Goldstein et al. [90] studied a ring resonator design operating at a 50 nm with grazing incidence, beam expansion elements and multifaceted aluminum mirrors. They considered outcoupling by means of both a scraper in the resonator backleg and a hole in the first mirror. At $\gamma = 511.5$, 150 A peak current, and 40 mm mrad normalized emittance, an 8 m undulator with 16 mm period yielded a small signal gain prediction of 5.2 at 50 nm. A key aspect of the design was the grazer followed by six faceted end mirrors on the optical cavity with glancing angles of 75° . Five of the mirrors were flat and one was an off-axis paraboloid. The 29.5 m optical cavity was relatively sensitive to tilt showing power drops of 50% for $1 \mu\text{R}$ tilt. Of greater concern was spherical aberration, as an aberration of only $\lambda/5$ reduced the power by 50% and decreased the output beam quality. This surface figure tolerance is severe for an X-ray mirror due to manufacturing limits as well as possible thermal distortions, even taking advantage of the reflectivity enhancement (98.2% for the Si hyperboloid and 92% for the Al flats) and $1/\cos(\theta)$ figure improvement at small angles.

This work was later extended to modeling output wavelengths as short as 4 nm and included the effects of thermal distortion [91] for beams of 10 mm-mrad normalized emittance, a value considered at the time to be “the best conceivable beam quality from a photocathode/linac system”. All coefficients in the Zernike expansion of the mirror distortion had to be smaller than 0.5λ not to extinguish the laser, a result later experimentally verified in the infrared region [93].

The next conceptual advance came with the theoretical and simulation studies of a high gain, high extraction efficiency oscillator [94]—the Regenerative Amplifier Free Electron Laser or RAFEL. Although the simulation and later demonstration [95] were at $16.3 \mu\text{m}$ wavelength from 17 MeV electrons, this approach scales to the X-ray region. The RAFEL relies on very high single pass gain and feed back of a small fraction of the output pulse. The optical feedback allows the radiation intensity to build to saturation after only a few passes and provides significant energy extraction. The experiment used 4.5 nC of charge to produce 270 A of peak current with 7 mm mrad emittance. A 100 period undulator of 20 mm wavelength generated measured gain lengths of only 15 cm, more than sufficient to drive the system to saturation, since net gain was 330 as deduced by a separate SASE experiment. Saturation was observed in a few μs and a sizeable detuning length of 1 mm was observed. The optical cavity consisted of an annular mirror for outcoupling while the return beam was collimated by a pair of spherical and cylindrical mirrors approximating a 90° paraboloid. Of great importance to scaling to the VUV/soft X-ray region was the relative insensitivity to alignment and pulse overlap. This feature is expected if one considers the feedback pulse as simply a low level seed for a high gain amplifier; overlap, mode quality, and alignment tolerances for high gain amplifiers are significantly relaxed over typical low gain oscillator systems.

A regenerative FEL amplifier design for the TESLA test facility at DESY [96] was simulated [97] to operate at beam energies from 180 to 325 MeV with 1 nC bunch charge producing 500 A peak current of 2 mm mrad normalized emittance. The linac could produce $700 \mu\text{s}$ macropulses at 10 Hz. The undulator was a 27.3 mm period 13.5 m long system consisting of three modules. A long optical cavity of 66.4 m employed only a spherical mirror at one end with a hole outcoupling 10% of the pulse and a spherical grating assembly at the other end. A grating was mounted in the Littrow configuration provided bandwidth control while stretching the pulse length by $4 \times$. SiC was chosen as the mirror material in normal incidence because reflectivities of greater than 40% could be obtained in the 10–20 eV photon range. In addition, the thermal properties of this material are excellent, such that

modeling determined that surface distortions caused by the average absorbed power would be negligible. Simulations indicate that an alignment accuracy of $2\ \mu\text{rad}$ is required for reliable operation, a number that has been achieved in several FEL systems. The same simulations predicted fully coherent output pulses with a narrow bandwidth determined by the pulse length at nearly constant saturated power with little dependence on the seed level.

Thomson et al. [98,99] show that the optical properties of the RAFEL differ from a low gain oscillator because the modes experience strong gain guiding in the FEL. The primary function of the cavity is to return a small part of the optical field to seed the next electron pulse. Even 10^{-5} of the output is sufficient to stabilize the output and deliver significantly better temporal coherence than a SASE system. When optimized, the time bandwidth product is approximately double a Fourier transform limited pulse. Extensive studies showed pulses within 36% of a transform limited bandwidth, five times better than the equivalent SASE result. The authors also mention the possibility of combining harmonic operation for up-conversion to shorter wavelengths.

4.5.3. Facility architectures for the future

Many user applications that demand high average brightness drive facility architectures toward CW SRF based FELs. The potential ability of a soft X-ray oscillator to operate at multi-MHz repetition rates is a tantalizing idea. As in other CW systems, a number of technical hurdles must be overcome before such a design could come to fruition. In particular, a high brightness CW injector is key as it is for high repetition rate amplifiers. The output pulse train from the gun must be produced at the cavity round trip frequency with the following caveat. As described in Siegman [100], it is possible to operate an optical cavity slightly off axis such that the mode traces a Lissajous figure on the mirror surface. If the outcoupling is by means of a hole at one point on this curve, then the output pulses will be provided at the cavity period divided by a small integer. Provided the mirror losses are small enough so that, after all the bounces, sufficient seed power still exists to drive the FEL to saturation, the electron beam pulse need only be provided at this lower repetition rate.

A FEL oscillator would have a generic layout similar to that of the SRF-based FEL amplifier shown in Fig. 5 with the addition of an optical system around each undulator. Several oscillators could be supported if the accelerator can operate at a sufficiently high repetition rate (and thus high average current). Principal components are a high repetition rate photocathode injector (a few hundred pC charge, $\leq 1\ \text{mm mrad}$ normalized emittance), a superconducting L-band linac ($\sim 2.5\ \text{GeV}$ for 1 keV photons in the first harmonic, assuming conservative undulator parameters), bunch compressors to generate high peak current for FEL gain, and spreaders to simultaneously deliver beams to multiple oscillator FELs. Freed from the need for a seed pulse, variable gap or variable field undulators would offer continuous, real time tuning at constant linac energy. FEL oscillators would provide an average brightness exceeding $10^{25}\ \text{ph/s/mm}^2/\text{mrad}^2/0.1\ \text{bw}$ and peak brightness exceeding $10^{32}\ \text{ph/s/mm}^2/\text{mrad}^2/0.1\ \text{bw}$.

The FEL oscillator will also produce copious harmonics of the fundamental. Typically, the power at the H^{th} harmonic of an oscillator is of order 10^{-H} of the fundamental. Thus the several harmonics would also provide higher average brightness than achievable with any existing source, as they are outcoupled through a hole in the first mirror along with the fundamental beam.

The beams envisioned for these light sources, of the order of 1 mA at 500–1000 MeV, are well within the range of current

technology though normalized emittances at $\sim 200\ \text{pC}$ need improvement to be substantially less than $1\ \text{mm mrad}$. Since a high repetition rate is essential for this approach, the linac system must support such relatively high average currents. RF phase and amplitude stability are tight but feasible (typically between 0.03° and 0.02%). A conceptual design, optimized for both cost and beam dynamics, has not yet been performed.

4.5.4. Readiness and principal challenges

A VUV/soft-X-ray FEL oscillator facility providing multiple users with fully coherent, MHz or higher repetition rate, 100 fs pulses appears challenging, but ultimately within reach. Beyond the accelerator requirements shared by all X-ray FELs, the key to the success of this approach is the availability of excellent, high reflectivity mirror technology in the desired wavelength band. Reasonable reflectivities can be achieved in glancing incidence for aluminum mirrors down to 4 nm. In such a system multiple glancing bounces in the resonator are required. In normal incidence mirror, losses are higher and more concern exists about thermal loading effects.

The development of high reflectivity mirror coatings for the desired wavelength bands of operation is a priority, especially 50–600 eV where seed laser technology will be limited. Multilayer mirrors are designed so that scattering from the change in refractive index at the interfaces between layers interferes constructively. Alternating layers of two materials, a high atomic number followed by a low atomic number for wavelengths below 100 nm continues to be an area of active development [101], although such mirrors do not have wide tunability. At the soft X-ray wavelength of 2.734 nm (in the so-called water window), titanium oxide and aluminum oxide have achieved over 30% near normal incidence reflectance on a silicon substrate [102]. Atomic layer deposition has been used to deposit $\text{Al}_2\text{O}_3/\text{W}$ multilayer coatings for X-ray optics on bulk substrates for high reflectivity down to 0.15 nm [103]. Multilayer coating methods for X-ray reflectivity enhancement of polysilicon micro-mirrors have even achieved high reflectivity at 1.54 nm wavelength [104]. In the EUV range of 12–25 nm, Mo/Si mirrors [105,106] at normal incidence are now available commercially with reflectivities of 68% at 13.5 nm [107].

This approach is close enough to fruition to warrant a sub-scale test of the full concept; that is, build a SRF driven FEL oscillator in the 20–100 nm range to experimentally determine the tolerances and limits of the approach including extensive tests of X-ray optics. Electron beam energies of a few hundred MeV would be required with CW currents of several hundred microamperes.

5. Hard and ultra-hard X-ray FELs

Hard X-Ray FELs (XFELs) are defined as those producing photon energies between 5 and 50 keV ($0.25\ \text{\AA} < \lambda < 2.5\ \text{\AA}$) whereas ultra-hard XFELs are defined as those producing photon energies greater than 50 keV. At present, no sources of coherent seed radiation exist in this photon energy range (the LCLS third harmonic is already at 25 keV). Moreover, the limits of harmonic upshifting with combinations of HHG lasers and either high gain harmonic generation cascades or echo enhanced harmonic generation are not well understood. Therefore, a hard/ultra-hard X-ray FEL can only use a SASE, self-seeded, oscillator, or regenerative amplifier configuration. The LCLS, the world's first hard X-ray FEL, has demonstrated SASE lasing and saturation at $1.5\ \text{\AA}$ [1]. An oscillator XFEL using Bragg crystals as X-ray optical cavity mirrors is discussed in detail in Section 5.4 of this paper. A regenerative amplifier has been demonstrated at infrared

wavelength [95] and has been proposed for an XFEL [108]. Optical feedback on the order of 10^{-5} is adequate for the hard X-ray FEL to saturate in a few passes. A regenerative amplifier approach can potentially shorten the undulator length needed for saturation and improve the XFEL temporal coherence. The remainder of this section focuses on the SASE design.

Most hard/ultra-hard X-ray FELs are designed as SASE amplifiers which amplify spontaneous noise by 10^7 and reach saturation in a single pass through the undulator. A practical power gain length for the hard/ultra-hard X-ray range is 3–5 m, requiring therefore an undulator length of the order of 60–100 m. The number of X-ray photons per pulse generated at saturation depends on the photon energy and the FEL pulse energy, which is the product of the electron beam energy, bunch charge and FEL parameter ρ . For a typical XFEL operating with nominal 1 nC, 10 GeV and $\rho \approx 10^{-4}$ to 10^{-3} , the pulse energy is about 1 mJ. The number of X-ray photons can be estimated from the photon energy to be 10^{11} – 10^{12} per pulse depending on pulse length and whether the radiation is produced at the fundamental or at a harmonic.

Sample heating is likely to limit hard/ultra-hard XFELs to pulsed, low-duty factor operation. X-ray pulses can be produced at tens to hundreds of pulses per second in the case of normal-conducting copper linac, or in a macropulse consisting of many micropulses, as in the case of an SRF linac. The macropulses can be repeated to increase the average photon flux. To provide hard X-ray photons to multiple simultaneous users, either multiple undulators or X-ray beam splitters (when they are developed) can deliver photon beams to the users.

5.1. Science case examples for hard X-ray SASE FELs

Of the many science drivers studied for the LCLS and the XFEL at DESY, two examples may illustrate the unique capabilities (high intensity and ultra-short pulse) of the hard XFEL. High-energy X-ray beams can be focused to a pencil-thin line in solids to create a uniformly heated volume of warm dense matter (WDM) [109]. WDM is the non-equilibrium state of matter that occupies an intermediate region of temperature and density ($T=0.1$ – 10 eV, $\delta=1$ – 10 g/cc). Room-temperature, uncompressed matter occupies the region ($T=0.01$ eV, $\delta=1$ g/cc) whereas hot dense matter, such as exists in stellar cores and in directly driven inertial fusion plasmas, occupies the region ($T=100$ eV, $\delta=0.01$ – 100 g/cc). WDM occurs in the cores of large planets, in X-ray driven inertial confinement implosions and in systems that begin as solids and end as a plasma. To understand WDM, one needs accurate determination of its equation of state (EOS) in a regime where neither condensed matter physics ($T=0$) nor plasma physics ($T=\infty$) is valid. The fundamental XFEL pulses can heat a solid target rapidly and uniformly to create WDM, and another XFEL pulse or its harmonic can probe the WDM to determine its EOS [110].

A second example is the use of the high-energy (50–100 keV) photons to characterize materials under the intense pressure generated by irradiation from ns-long, high-energy laser pulses [111]. The sub-ps X-ray pulses can probe deeply into the high density material under this shock-free compression with exquisite spatial and temporal resolution. These features are valuable to detect high-pressure-induced phase transitions, to study the dynamics of melting and refreezing, etc.

5.2. Status of hard X-ray SASE FEL facilities

Since October 1, 2009, the LCLS is operating routinely, serving users 120 h per week, with the other 48 h reserved for machine

development and maintenance. Beam availability, with respect to schedule run time, has already achieved 92% (98% if user-requested tuning time is included). In addition, the FEL power stability has been as good as 3% rms and up to 10%, depending on the user's choice of photon energy (780–2000 eV) and pulse width (60–300 fs FWHM). FEL saturation has also been achieved for photon energies from 780 (16 Å) to 9000 eV (1.4 Å), with only soft X-rays delivered to users, as the hard X-ray line is still under construction. In low charge mode (20 pC), X-ray pulses are likely as short as 10 fs, although these values are inferred rather than directly measured.

One of the most remarkable aspects of the LCLS start-up was the immediate demonstration of FEL gain on the very first attempts in April 2009. FEL lasing was observed immediately after 10 undulator segments were inserted, and within 4 days the SASE FEL was fully saturated at the shortest wavelength of 1.5 Å, well within the design undulator length. This encouraging success story demonstrates the real practicality and the great potential of future FELs, which are now well grounded as stable and reliable light sources.

In addition, new extreme pulse compression schemes with reduced bunch charge have been conceived and now tested at LCLS demonstrating stable, saturated operation with high-power X-ray pulse durations of < 10 fs. This new ultra-short pulse mode is now being delivered to a host of very excited users.

Two other hard X-ray FELs are now under construction, one at DESY in Germany and one at Spring-8 in Japan. The characteristics of this facilities are compared with those of LCLS in Table 1.

5.3. Challenges and technologies for future hard/ultra-hard X-ray FELs

The SASE X-ray beams have full transverse coherence, but due to the inherent nature of the SASE process, the pulses only have partial longitudinal coherence, i.e., each X-ray pulse consists of a number of spikes that are coherent over the spike duration, also known as the cooperation length (approximately the slippage length of the radiation). The corresponding Fourier transform has many spectral features, i.e., chaotic and broad spectral bandwidth. Improving the temporal coherence of the SASE XFEL will enable applications that require phase information or narrow linewidth. Of the several possible approaches to enhance the longitudinal coherence, none has been experimentally demonstrated. In the self-seeding approach, SASE coherence is enhanced by dividing

Table 1
Hard X-Ray FEL facilities in operation or under construction.

	LCLS	EU XFEL	SPRING-8
λ (Å)	1.5	1	1
First Light	2009	2014	2011
E (GeV)	13.6	20	8
Linac Type	NCRF	SRF	NCRF
Freq (GHz)	2.856	1.3	5.712
Length (km)	1	3.4	0.75
Gun Type	NCRF	SRF	Pulsed HV
Freq (GHz)	2.856	1.3	DC
Cathode	Cu	CS ₂	CeB ₆
Type	Photo-Injector	Photo-Injector	Thermionic
Q_{bun} (nC)	0.25	1	0.3
I_{peak} (kA)	3	5	3
ε_n (μm)	0.5	1.4	0.8
τ_{FWHM} (fs)	70		
Undulator	Planar	Planar	Planar
L_u (m)	132		100
$N_{ph}/Pulse$	2×10^{12}	1.2×10^{12}	2×10^{11}
f (Hz)	120	30,000	60

the undulator into two sections. The first undulator produces partially coherent FEL light that is monochromatized with a diffraction grating. The monochromatic light is injected into the second undulator where it seeds the amplification to generate a longitudinally coherent X-ray beam [41]. Alternatively, the electrons can be pre-bunched at a sub-harmonic wavelength, as from echo-enhanced harmonic bunching [48], to produce periodic density modulations at that wavelength. The electron beam pre-bunched at a longer wavelength is then injected into the undulator and radiates coherently at the n th harmonic where n is a large number. The coherent power at the n th harmonic power depends on the pre-bunching factor as given by

$$P_{\text{prebunch}} \approx |b_n|^2 \rho \left(\frac{I_{\text{peak}} E_b}{e} \right) \quad (14)$$

where $|b_n|^2$ is the bunching factor at the n th harmonic, I_{peak} is the peak current and E_b is the electron beam energy. With sufficient density modulation the coherent power can exceed the start-up noise and serve as the seed source for the amplification process.

The FEL resonance condition dictates that to produce wavelengths shorter than 1.24 \AA using a practical value for the permanent-magnet undulator period ($\sim 1.5 \text{ cm}$), one would need an electron beam energy between 5 and 35 GeV. The other requirement was given in Eq. (8), that the electron beam's emittance observed in the laboratory, its normalized rms emittance ($\epsilon_{x,N}$) divided by γ , must be less than the photon beam's transverse phase space area. This can be written as

$$\frac{\epsilon_{x,N}}{\gamma} < \frac{\lambda}{4\pi} \quad (15)$$

which sets the minimum electron beam energy, or γ , for a given wavelength and normalized rms emittance.

Future hard/ultra-hard XFELs can benefit from ultra-low normalized emittance ($< 0.1 \mu\text{m}$) at very low bunch charge, where the normalized emittance is dominated by the thermal emittance. A photoinjector operating at low bunch charge, small photoemission radius using photocathodes with small mismatch between the cathode band-gap and the drive laser photon energy will be the key to achieving this ultra-low emittance. These are discussed in details in Section 6 of this paper.

The typical electron beam for the LCLS 1.5 \AA XFEL delivers a normalized rms emittance of $0.5 \mu\text{m}$ at bunch charge of 0.25 nC . As one pushes toward shorter wavelengths (smaller photon emittance), one has to increase the electron beam energy to reduce the beam's geometric emittance. Unfortunately with increasing electron beam energy, energy diffusion due to random (quantum) fluctuations in the emission of spontaneous synchrotron radiation becomes important. The relative energy spread induced by this process scales inversely with wavelength as given by

$$\left\langle \left(\frac{\delta\gamma}{\gamma} \right)^2 \right\rangle = \frac{1.26410^{-19} \text{ cm}^2 \gamma^2 K^3 N_u}{L^2 (1 + 1.33K + 0.4K^2)} \quad (16)$$

where $\delta\gamma/\gamma$ is the relative energy spread induced by quantum fluctuations, K is the rms undulator dimensionless parameter, and N_u is the number of undulator periods. With sufficiently small FEL wavelength and large gamma, at some point along the length of the undulator this energy diffusion will exceed the energy spread induced by the FEL interaction, i.e., the FEL bucket; and the performance of the SASE FEL will degrade or cease altogether. For a 50 keV XFEL using a 35 GeV electron beam, quantum fluctuations are expected to induce an energy spread comparable to the FEL interaction in 120 m of undulator length, the same length needed to saturate with an emittance of $0.2 \mu\text{m}$.

5.4. Hard X-ray FEL oscillators and cavity configurations

Huang and Ruth [108] recently proposed a hard X-ray regenerative amplifier system using 3 Bragg crystals to form a ring resonator. Initially the 1.55 \AA (8 keV) SASE radiation is spectrally filtered by the crystals. Amplification of the FEL broadens the optical spectrum beyond the acceptance of the crystal planes resulting in transmission of some of the photon pulse through one $100 \mu\text{m}$ thick crystal for out-coupling. The net three crystal reflectivity within a 4×10^{-6} bandwidth was 91%. A calculation based on LCLS parameters (9.9 GeV , 3 kA peak current from 300 pC bunches, and 1 mmrad normalized emittance) indicated power gains of 39 after a 20 m undulator. Two to three orders of magnitude higher brightness would result as compared to SASE operation. Such an approach would only work at the hard end of this spectral range (multi-keV) since high reflectivity from crystal diffraction planes is an essential element of the design.

A hard X-ray FEL oscillator (XFEL) appears to be feasible with ultra-low emittance, low charge ($20\text{--}50 \text{ pC}$) electron bunches with a repetition rate of about 1 MHz , an undulator of length $20\text{--}60 \text{ m}$, and optical cavity consisting of near perfect diamond crystals and high-reflectivity, grazing incidence, curved mirrors [112,113]. The output characteristics of an XFEL would be extraordinary and complementary to a SASE FEL. Although its pulse intensity would be lower by about a thousand times, the pulse length would be two to three orders of magnitudes longer, the bandwidth about four orders of magnitudes times narrower, and the repetition rate two to four orders of magnitudes times higher than a SASE amplifier. The relatively low power minimizes the thermal impulse loading of the crystals in the X-ray cavity and avoids damage of the experimental sample. An XFEL can thus revolutionize the field of inelastic scattering, Mössbauer spectroscopy, bulk-sensitive Fermi surface study, X-ray imaging with near atomic resolution, X-ray photon correlation spectroscopy, etc.

Oscillator FELs have struggled to extend their operating range to X-ray wavelengths, because of (1) the shortage of mirrors with sufficiently high reflectivity consistent with the limited single-pass gain of the system, and (2) the sensitivity of many mirrors to radiation damage.

The concept for a hard X-ray FEL oscillator that employs crystals as high-reflectivity Bragg mirrors was first proposed about 25 years ago [20], at the same time as the X-ray SASE was

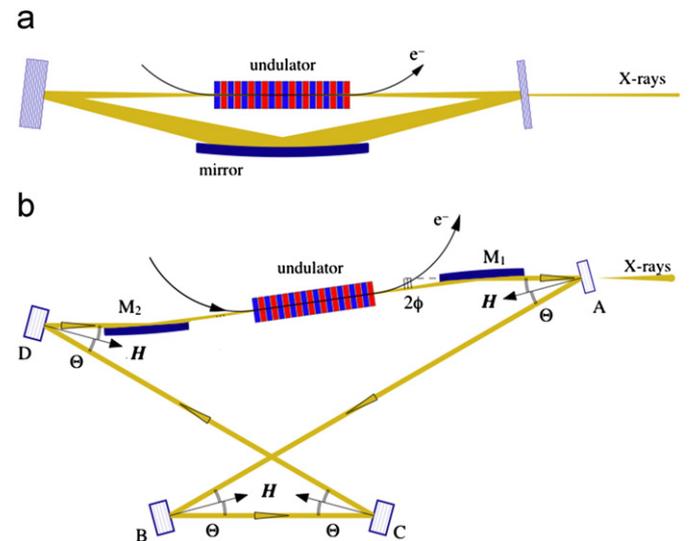


Fig. 6. Schematics of an XFEL optical cavity, (a) a basic configuration with two crystals which is not tunable and (b) a tunable configuration.

proposed. However, the concept did not receive much attention until a recent study showed that an XFEL would be made possible by using low-intensity, ultra-low-emittance electron bunches [112].

Several schemes have been proposed to improve the temporal coherence of high-gain X-ray FELs, such as self-seeding [114], electron beam out-coupling [115], and the regenerative amplifier [108]. Compared to these schemes, the XFEL has the advantage of providing higher coherence.

The basic configuration of an XFEL is shown in Fig. 6(a). An X-ray pulse is stored in an optical cavity consisting of two crystal reflectors and a grazing-incidence mirror for focusing. One of the crystals is designed to be thin so as to couple a fraction of the intra-cavity power to the output power. Each time a pulse arrives at the undulator entrance it meets an electron bunch, and they travel together through the undulator. During this process the X-ray pulse is amplified via the FEL interaction. Although the single-pass gain is small, the pulse intensity increases steadily and exponentially as the pass number increases, as long as the FEL gain can overcome the loss in the optical cavity. Eventually, the gain decreases due to nonlinear effects, and the FEL reaches a steady state in which the gain balances the loss.

Curved mirrors are necessary in an optical cavity to control the transverse mode profile to assure optimum gain. For tuning the photon energy, the Bragg angle must be changed together with the angle of incidence on the curved mirror. The tuning range in this scheme is thus severely limited, since the angle of grazing incidence should be kept small, i.e., less than 1 mrad, for high reflectivity. A broad range of tuning can be achieved with the four-cavity scheme shown in Fig. 6(b), in which the Bragg angle can be changed without changing the angle of grazing incidence on the mirrors or the round-trip path length, by a coordinated translation and rotation of the four crystals [116,117]. An important additional merit of the four-crystal scheme is that several different Bragg planes from the same crystal can cover all spectral regions of interest. The choice material is diamond—a crystal with the highest reflectivity and with excellent thermo-mechanical properties.

5.4.1. Unique characteristics and applications

The output characteristics of an XFEL are extraordinary and complementary to a SASE FEL. In a SASE FEL the bandwidth is set by the rho parameter. In contrast, the bandwidth of the oscillator is monochromatized by the design of the optical cavity. In principle, a bandwidth of a few meV (corresponding to a relative bandwidth of about 10^{-7} at hard X-ray wavelengths) is achievable. Such small bandwidths can be four orders of magnitude smaller than possible with simple SASE FELs. The micro-pulse duration (~ 1 ps) will be 10–100 times greater, while the pulse intensity will be ~ 1000 times smaller. Lower single pulse intensity may be advantage in experiments requiring the sample to survive very large numbers of X-ray pulses or requiring the suppression of nonlinear effects.

The repetition rate of XFEL pulses is constant at ~ 1 MHz—two to four orders of magnitude higher than that of pulsed SASE systems. This repetition rate, while not as flexible as that with SASE or seeded amplifiers, is ideal for many time-resolved techniques in which the maximum desirable rate is set by the relaxation time of the processes under investigation. Consequently the time-averaged brightness of an XFEL may be up to five orders of magnitude higher than that of pulsed SASE FELs. Additionally, FEL oscillators are predicted to generate fully phase-coherent radiation.

Using X-ray pulses from an XFEL, the techniques developed at third-generation synchrotron radiation facilities—particularly those requiring high coherence and high spectral purity—could be

dramatically enhanced. The APS, one of the brightest sources in the hard X-ray range, produces 10^9 photons per second with a meV bandwidth, which will be increased a million-fold by an XFEL. Inelastic X-ray scattering (IXS) [113,118] and nuclear resonance scattering (NRS) techniques [119,120], which are currently limited by the available flux in the desired bandwidth, can be revolutionized by an XFEL. The intensity of an XFEL will also enable hard X-ray photoemission spectroscopy (HAXPES) for time-resolved study of Fermi surfaces in bulk material [121,122]. Hard X-ray imaging will be feasible with nm resolution using multilayer Laue lenses that require an X-ray bandwidth of less than 10^{-5} [123].

5.4.2. Issues and R&D items

Two technologies of critical importance for realizing an XFEL are: an ultra-low-emittance injector producing low intensity bunches at a constant repetition rate, and a stable, low-loss X-ray cavity consisting of Bragg crystals and focusing mirrors. Substantial research efforts in each of these areas are required to demonstrate the practicality of FEL oscillators at hard X-ray wavelengths.

5.4.3. XFEL modeling and parameter optimization

Analytical calculations using supermode theory [124,125,112] describe the initial seeding of the FEL from spontaneous noise [126], and shows that the electron beam length should be longer than the inverse bandwidth of the Bragg crystal to have sufficient gain per pass to reach saturation. Full oscillator simulations including the complex Bragg mirror have been performed [127] with the simulation code GINGER, demonstrating several possible configurations for generating 5–20 keV X-rays using an electron beam with 10–20 A of current in a 1 ps, 0.2×10^{-6} mm mrad emittance bunch at 7 GeV. To reach saturation ~ 50 MW requires ~ 3000 periods of undulator if the losses are taken to be 15%. While the simulations of the basic oscillator configuration are rather mature, some additional work is needed to include the transverse acceptance of the Bragg crystals and to model more accurately the four-mirror tunable cavity.

Recent work has shown [128] that compressing a similar low-charge (25 pC) beam to 0.1 ps and 100 A significantly relaxes the beam and cavity requirements: with a 0.3×10^{-6} mm mrad emittance beam and 50% losses, one can reach saturation using < 1000 undulator periods. This configuration may be more stable and more tolerant of errors in the positions of optical elements.

The intensity of the intracavity harmonics is low, about five orders of magnitude smaller than the fundamental for the cases studied so far. This result is due to (1) the low value of the deflection parameter K , (2) the negligible reflectivity of the harmonics by the thin crystal, and (3) the relatively large energy spread with respect to the FEL bandwidth (proportional to the inverse of the undulator periods). The relative intensity of the harmonics transmitted through the thin crystal, however, is higher, only three to four orders of magnitude smaller than fundamental, due to the higher transmission efficiency of the harmonics.

The narrow bandwidth of the Bragg mirrors precludes directly producing fs X-ray pulses using an XFEL. However, it may be advantageous to use this concept to seed a very short electron beam [128]; an electron beam so prepared could then be extracted from the oscillator region and passed through a short (a couple of gain lengths) undulator section to produce fully coherent, fs X-ray pulses. The fs pulses produced in this scheme should be a stable in magnitude, in contrast to the ultra-fast SASE pulses proposed recently [129,130]. Extracting and transporting a high-energy beam from the oscillator while maintaining the

Å-scale microbunching is extremely challenging and will require a careful electron beam and X-ray optics design.

The normalized emittance of electron bunches for XFEL operation is ~ 0.1 mmrad, much smaller than the nominal value typically quoted for high-gain X-ray FELs. As the peak current is $100\times$ smaller, an XFEL injector would not suffer space charge degradation of the initial thermal emittance. For an XFEL injector, a promising approach starts with a continuous, low-current beam from a thermionic cathode that minimizes space-charge effects and the need for their compensation. A group at RIKEN, SPring-8 (Japan) has constructed a low-emittance injector of this type for their test high-gain FEL employing a pulsed DC voltage at a repetition rate of 60 Hz [131,132].

5.4.4. X-Ray optics

The reflectivity of both the diamond crystals and the grazing incidence, curved mirrors should be very high. At the Bragg peak, the reflectivity of perfect diamond crystals is predicted to be 98–99% in the hard X-ray region between 5–30 keV [133]. Since only an area less than 0.5 mm in diameter is required, it should be possible to select out a small, high-quality single-crystal region from bulk crystals. Still to be demonstrated is the closeness of the reflectivity of the commercial diamond crystals to the predicted value.

For high reflectivity, say greater than 95%, the grazing angle of incidence on the curved mirror should be well below the critical angle, making the mirror large and bulky. The mirror profile should be ellipsoidal for ideal focusing. The specifications on mirror roughness figure error are very tight, 10 nm in height rms error (assuming 1 mrad angle of incidence) and $0.1\ \mu\text{rad}$ rms figure error, but may be feasible with a new method that combines electrolytic in-process-dressing grinding and elastic emission machining [134]. Such mirrors are highly desirable for the general community of synchrotron radiation and X-ray FELs.

Although only a small fraction of the X-ray energy is absorbed in the diamond crystal, the deposited heat will increase the temperature causing the crystal to expand, which can result in a change in the Bragg energy of peak reflection. Related are two heat-load issues: (1) if the heat from one pulse has not sufficiently dissipated the subsequent pulse will see a distorted lattice and different reflective properties of the crystal (the “inter-pulse problem”), and (2) the tail of a single pulse may see an expanded crystal due to the heating from the pulse front (the “intra-pulse problem”). The inter-pulse problem may be solvable by cooling the crystal to a low temperature, say 50 K, at which the heat diffusivity of diamond is large. In addition, the heat expansion coefficient for diamond at temperatures lower than 100 K was measured to be very small [135]. The intra-pulse problem is not fully understood. However, some simple models indicate that crystal expansion time is much longer than 1 ps, in which case the crystal properties remain constant within the duration of a single X-ray pulse.

The requirements to stabilize the crystals in the cavity are very stringent. The angular tolerance determined from the stability of the transverse mode is 10 nrad while the positional tolerance determined from the detuning curve is $3\ \mu\text{m}$ [136]. To achieve these tolerances, the null-detection feedback technique employed at LIGO may be the most promising, because it allows stabilization of several optical axes with a single detector. A preliminary feedback experiment at the APS beam line indicates that an angular stability of 50 nrad within 1 Hz fluctuation bandwidth can be achieved for a single axis [137]. Further experiments for multi-axis feedback with broader bandwidth are under preparation.

5.4.5. Main accelerator

To provide the highest average brightness, an XFEL requires a superconducting linear accelerator operating in a CW mode.

While this technology may be regarded as mature, extensive cost engineering is highly desirable. A straight linear superconducting accelerator may be employed for single-pass acceleration to the final energy of 7–10 GeV and would also be the most versatile in providing different bunch profiles. Substantial savings in the linac length and thus cost may be achieved by one or more recirculation paths, with or without an energy recovery option [138].

An XFEL can also be operated using a pulsed superconducting linac such as the 15 GeV linac for the European XFEL [87]. The macropulses in these linacs are 1–2 ms long, accommodating 1000–2000 micropulses at 1 MHz repetition rate, which is sufficient to drive an XFEL to saturation level if the electron bunch is compressed to 100 fs, as discussed above.

5.4.6. Outlook

X-ray FEL oscillators provide an exciting extension of hard X-ray capability into the realm of ultra-high spectral brightness and high average power. The prerequisite technologies—an ultra-low emittance, low-intensity, 1-MHz repetition rate electron injector and X-ray optics technology producing high-reflectivity Bragg crystals and grazing incidence curved mirrors with accurate and stabler positioning—are challenging, but appear to be within the limits of the current technology. Five years of an intense and successful R&D program advancing each of the critical technologies could make possible a detailed conceptual design of an XFEL.

6. Ultra-short pulses

A definition of the ultra-short X-ray pulses in this paper is the pulse duration of the order of few femtoseconds and shorter. A rather natural way to obtain such pulses is to use ultra-short electron bunches. In fact, pulses of just a few femtoseconds in duration with a $1.5\ \text{\AA}$ carrier wavelength have recently been obtained at the LCLS using electron bunches of a few femtoseconds in duration containing only a charge of 20 pC [139]. In this experiment the bunch length is inferred rather than measured directly. The full exploration of producing and manipulating ultra-short bunches will require a new generation of electron beam and photon diagnostics capable of such fine temporal resolution.

The next goal is to obtain even shorter pulses with pulse duration comparable to a cooperation length (e.g., the length of temporal coherence) in the FEL, thus, obtaining single-spike, nearly Fourier transform limited X-ray pulses. In the case of LCLS this would be X-ray pulses of the order of a few hundred

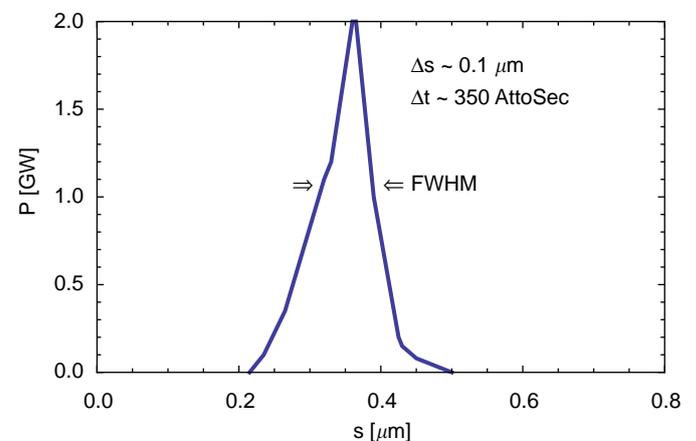


Fig. 7. A nearly Fourier transform limited X-ray pulse at the LCLS simulated using the electron bunch with 1 pC charge compressed to the rms bunch length of 0.2 fs.

attoseconds. It was proposed that such short electron bunches can be obtained by further reducing the electron bunch charge to approximately 1 pC [140]. Remarkably, bunches with lesser charge are also expected to have a higher brightness and, therefore, are expected to saturate the FEL faster. Fig. 7 demonstrates a simulation result for LCLS.

Other ideas for obtaining the ultra-short X-ray pulses are based on selection of a small portion of the electron bunch to produce a dominant radiation. It was proposed in Ref. [141] to use a slotted foil that would spoil the emittance of the entire electron bunch except for a small fraction that goes through a slot in the foil. Due to the sensitivity of the FEL process to beam emittance, only this fraction lases and produces the ultra-short X-ray pulse.

While both methods described above can be implemented at the LCLS immediately, their utilization in pump–probe experiments would suffer from a jitter in the electron bunch arrival time in the undulator. In the case of the short electron bunch, however, it might be possible to use coherent transition radiation from the bunch in the micron wavelength region as a timing signal. Recent advances in a generation of intense few-cycle laser pulses with a carrier-envelope phase stabilization (see for example Ref. [142]) inspired a set of ideas for a generation of sub-femtosecond X-ray pulses [143–149] free from the problem of arrival time jitter. In all these proposals a few-cycle laser pulse (or a combination of two of these pulses) co-propagates with the electron bunch in a one period wiggler magnet that is located right upstream of the FEL, inducing an energy modulation of the electrons that is much larger than the electron energy spread (Fig. 8).

Then, specific means [143–149] are used to force only a fraction of electrons that participated in the interaction with the laser, typically electrons located close to the central peak of energy modulation or within the region of a maximum gradient of the energy modulation, to produce a powerful X-ray pulse that dominates the FEL output signal. One can actually take one ultra-short laser pulse, split it into two identical pulses, and modulate two sections of the electron bunch with a well controlled time delay between them to produce two nearly Fourier transform limited, ultra-short X-ray spikes in the FEL output. In fact, a combination of the recent ideas of echo-enabling microbunching [150] and current enhanced SASE [151] with ultra-short laser pulse technology allows generating even two-color spikes [152].

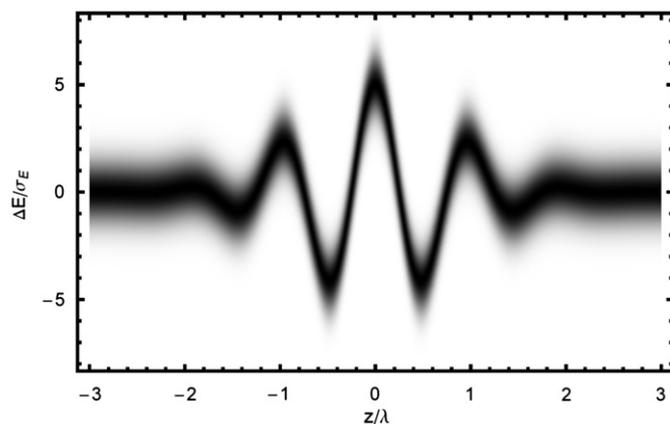


Fig. 8. Density plot showing energy modulation of electrons (in units of the electron energy spread) as a function of the position along the electron bunch normalized on the laser wavelength. The laser pulse width is 5 fs and the actual length of the electron bunch significantly extends beyond the plot boundaries.

7. Injectors

In the previous sections, numerous schemes and different configurations of FELs have been described. Although the final performance and the differences between the various cases can be quite significant, in all of them the required beam quality is ultimately set at the injector and in particular at the electron gun. The electron beam brightness generated at the electron source can be at best preserved in the downstream accelerator but never improved.

This section lists and describes the typical requirements for FEL injectors, introduces the available gun technologies, and analyzes how each approach matches the various requirements and what challenges still must be addressed. The critical issue of the proper choice of the cathode/laser system in photoemission guns (the dominant scheme used nowadays in high quality electron guns) is here only briefly addressed, as it is extensively treated elsewhere [153].

Some of the requirements for FEL injectors are directly defined by the physics of the lasing process; others are set by the requirements of the users' experiments; and still others, in a user facility, have to do with reliability, ease of operation, and cost.

7.1. General features

The performance of a FEL is ultimately determined by the availability of electron beams with high six-dimensional brightness. In most configurations the beam dynamics downstream of the injector can be effectively decoupled between the transverse and longitudinal planes. In that approximation the critical beam parameters at the undulator are the charge per bunch, the electron beam transverse geometric emittance and the longitudinal emittance. For lasing at X-ray wavelengths, the geometric emittance required for matching the small photon emittance is given by $\lambda/4\pi$, where λ is the photon wavelength. As geometric emittance is proportional to the ratio between the normalized emittance and the beam energy, a smaller normalized emittance allows for lower beam energies, thereby lowering the cost and complexity of the FEL (provided that undulators with the required period and K parameter are technically feasible).

The normalized emittance is an invariant quantity that it is ultimately defined at the electron injector. For this reason a large part of the effort of gun design is devoted to generating schemes capable of small normalized emittances. Space charge effects in the injector, if not properly controlled, can quickly and irreversibly increase the normalized emittance. Minimizing such effects requires controlling the transverse and longitudinal distribution of the electron beam, plus using techniques such as emittance compensation [154,155] that require magnetic fields of significant intensity in the gun region. With proper space charge control, the minimum normalized emittance that can be obtained is defined by the characteristics of the cathode and the electron extraction mechanism. In what follows, we will refer to this minimal quantity as the thermal emittance following the common use in the community (although such a term is strictly correct only when referred to thermionic cathodes). An intense worldwide R&D activity is dedicated to finding laser/cathode systems that minimize thermal emittance and generate the electron beam with the desired characteristics.

The longitudinal emittance and the charge per bunch define two other important quantities of the lasing process, energy spread and peak current. Small beam energy spread is required in FELs to permit the proper microbunching in the undulators, and is especially critical to lasing at high undulator harmonics. The path length difference induced by the different energy of the electrons

transiting in the undulator over one gain length must be much smaller than the wavelength of the radiation. Peak currents up to several kA are required for high-gain amplifiers, and are typically obtained by a first compression of the beam in the injector followed by further compression in one or more dedicated compressor chicane downstream in the linac. Short bunches at the gun are helpful because they minimize the subsequent compression required downstream. In short bunches, however, space charge effects increase and therefore a tradeoff must be adopted.

As noted in the previous section, requirements for a hard X-ray FEL oscillator can differ from those for high-gain amplifiers. The lower gain requirements allow for smaller peak currents, and hence for smaller bunch charges; however, emittance requirements are more stringent. Indeed, a X-ray free-electron laser oscillator for high temporal coherence [156] requires an injector providing low-intensity ($Q < 50$ pC), ultra-low normalized emittance ($\sim 0.1 \mu\text{m}$) bunches at a constant repetition rate of about 1 MHz. A photoinjector optimized for this performance should include a low thermal emittance photocathode (Cs:GaAs for example), or as proposed in an alternative injector scheme, a small-diameter thermionic cathode followed by acceleration and beam manipulation sections to compress the bunch to the required length [157,130,158]. An injector with these characteristics could also be used for ultra-fast, SASE X-ray FELs, producing few-fs pulses [159,160], and it can be an option for the high-coherence-mode operation of energy recovery linacs (ERL).

Light sources must be optimized to satisfy the experimental needs of users. The user community defines the principal characteristics of the photon beam required for experiments and so indirectly sets many of the requirements on the electron beam. Typical examples of user defined parameters are: (1) radiation wavelength and thus electron energy and undulator period and field, (2) average brightness and flux and thus repetition rate and average current, (3) photon pulse length and thus lasing mechanism schemes in FELs, and (4) radiation field intensity and thus peak beam current, total charge and so on. These requirements on the accelerator strongly impact those for the electron gun. Therefore, the proper combination gun/accelerator must be found to match the desired light source characteristics.

Like any other light source, FEL-based sources must deliver light to experiments in a reliable, continual, well-defined manner. Electron guns should be designed accordingly; systems with less complexity and greater ease of operation have a better chance of higher reliability. In many experiments, stability of the beam characteristics is important; hence, electron guns must be designed to minimize their contribution to the overall parameter fluctuation.

Table 2 summarizes the main requirements for an “ideal” injector for soft/hard X-ray FELs, the application that presents the

greatest challenges for injectors. The repetition rate dramatically impacts the choice of the accelerator technology and, thus, the complexity and cost of the FEL. FELs operating at up to ~ 1 kHz are compatible with presently available lasers and low quantum efficiency (QE) cathodes ($\sim 10^{-5}$ – 10^{-4}). Megahertz repetition rates will require photoinjectors with high QE cathodes ($\sim 10^{-2}$), preferably operating with visible drive lasers.

High quantum efficiency cathodes are typically based on semiconductor materials that present a higher sensitivity to damage due to ion back-bombardment and to surface contamination compared to lower quantum efficiency metal cathodes. Among the better understood semiconductor cathodes we can mention Cs:GaAs, which for example has been used extensively in DC guns at the Stanford Linear Collider and at the IR FEL at JLab, and Cs_2Te , which is reliably operating in the RF gun at the FLASH FEL in Hamburg for example. Lifetimes compatible with the operation of a user facility have been routinely demonstrated for both of these cathodes when operated at the proper ultra-high vacuum conditions: $\sim 10^{-11}$ Torr for Cs:GaAs, and $\sim 10^{-9}$ Torr for Cs_2Te . An intense R&D activity is also currently addressed to the development and characterization of other semiconductor materials potentially suitable for photocathode applications. In particular, multi-alkali antimonides present high quantum efficiency and can photo-emit in the visible, significantly relaxing laser requirements. Extensive information on these and other photocathodes can be found elsewhere [153]. In general, for the case of semiconductor cathodes an appropriate mechanical design using vacuum load-lock mechanisms is also required to allow for convenient reconditioning (and occasional replacement) of the cathode.

FEL operation in a low charge mode (tens of pC) permits sub-ps bunches at the injector while still controlling space charge effects. The reduced charge permits a smaller beam diameter at the cathode and consequently yields smaller thermal emittances. Higher charges require longer and larger diameter bunches at the gun; operation with ~ 1 nC bunches and normalized emittances smaller than $\sim 1 \mu\text{m}$ are typically necessary. Extensive simulations [161] show that obtaining such emittance values with nC charges requires an energy at the gun exit exceeding ~ 0.5 MeV. Higher electric fields at the cathode increase the space charge limit, allowing extraction of higher charges, and minimizing of space charge degradation of emittance.

Beam dynamics simulations and measurements [162] show that non-Gaussian charge distributions can significantly improve the emittance performance of an electron gun. The requirement of controlling the bunch distribution pushes most gun designers to laser/photocathode based systems, in which proper shaping of the laser pulse can control the electron bunch distribution. The emittance compensation technique [154,155] that is fundamental in obtaining low emittance values requires the presence of magnetic field in the cathode/gun region.

In practice, not all the parameters of Table 2 can be simultaneously satisfied by any available injector technology. In what follows the characteristics of these different technologies are presented and compared with the FEL requirements, pointing out present limitations and challenges.

7.2. Technology options

DC guns with thermionic cathodes have been and still are the electron sources for the majority of accelerators. DC gun schemes using photocathodes and generating quality low-emittance beams have also been developed and have shown reliable operation in a number of facilities. Among the advantages of DC technology are the potential of operating at arbitrarily high

Table 2
Main requirements for X-ray FEL electron guns.

Parameter	Value or comment
Repetition rate	Few Hz to hundreds of MHz
Charge per bunch	Tens of pC to \sim nC
Normalized emittance	~ 0.1 (low charge) to $\sim 1 \mu\text{m}$
Energy at gun exit	Greater than ~ 0.5 MeV
E field at the cathode	Greater than ~ 10 MV/m
Bunch spatial distribution	Controllable
rms bunch length	Tens of fs to tens of ps
B field compatible	Emittance compensation
Vacuum pressure	10^{-9} down to 10^{-11} Torr
Vacuum load-lock compat.	“Easy” cathode replacement
High reliability	For user facility operation

repetition rates, of being compatible with the presence of magnetic fields in the cathode area, and of demonstrated capability to operate at the pressures demanded by semiconductor photocathodes. Indeed, at present, injectors based on DC guns are the only ones that successfully operate with GaAs photocathodes, a semiconductor capable of generating polarized electron beams with very small thermal emittances. The main technical challenge with DC guns is achieving a beam energy at the gun exit greater than ~ 300 keV. Reliable operation at 350 keV has already been consistently obtained [163], but operation at higher energies has been limited by field emission and voltage breakdown at the ceramic insulator. A significant R&D activity is presently pursuing higher voltage operation by improving the technology of the ceramic insulators and by developing schemes using the inverted insulator geometry [164–166].

Relatively recently a conceptually new electron gun has been developed and successfully tested, the pulsed DC gun [157]. In this scheme, a modulator generates a high voltage pulse with a several μ s flat top that is applied to a DC-like gun structure. Qualitatively this approach is similar to that used in early induction linac injectors. From the beam dynamics point of view, the system behaves as a conventional DC gun, but the pulsed time structure overcomes the voltage limitations experienced by DC systems. The gun uses a thermionic cathode to exploit the relatively small thermal emittance of these cathodes and has demonstrated 500 keV beams with the charge and emittance required by X-ray FELs. The technical challenge for this scheme consists in pushing the pulsed modulator technology from the present few hundreds of Hz repetition rates towards higher values. MHz repetition rates may be beyond the range of such a technology.

Electron guns based on superconducting rf (SRF) accelerator cavities are a promising technology that would allow for CW operation with accelerating fields significantly higher than in DC guns and, when coupled with a photocathode/laser system, would be capable of generating FEL quality electron beams at extremely high repetition rates. Many groups worldwide are trying to develop such sources, and first experimental results are beginning to arrive [167–170]. Superconducting schemes have an inherent ability to achieve extremely low vacuum pressures, because of the effective cryo-pumping performed by the superconducting walls. Metallic (niobium) photocathodes operating at cryogenic temperatures had already been demonstrated, but showed very low quantum efficiency. Several groups are developing and studying schemes that allow coupling higher quantum efficiency, high temperature cathodes with superconductive structures. Cavity contamination by pollutants from the photocathodes is an issue under investigation, and a group has opted for a hybrid DC-SRF scheme that separates the cathode, in a DC gun structure at high temperature, from the downstream SRF cells used for accelerating the beam to higher energies [171]. The successful operation of cesium telluride (Cs_2Te) cathodes in a SRF gun [172] is promising. Because of magnetic field exclusion by the superconducting walls (Meissner effect), magnetic fields in the cathode region cannot be easily applied. Alternative schemes in which the required magnetic field configuration is generated by exciting a specific high order mode in the accelerating structure have been proposed and are under study [173]. Although not all of the required performances for a high brightness injector have been simultaneously proved yet in a SRF gun, the progress and the recent results indicate that the full potential of such a scheme could be achieved in the near future.

A number of normal-conducting guns operating in the L (~ 1 –2 GHz) and S (~ 2 –4 GHz) band have been developed and successfully operated in photoinjector schemes [174]. A recent quite remarkable example is the LCLS gun at SLAC that is reliably

delivering beams with the brightness required by the FEL lasing in either the low or high charge operation modes [175,176]. Normal conducting, high frequency guns can be considered a reliable and mature technology that shows several important features. It provides very high gradients (up to ~ 150 MV/m) that allow extraction of high charge bunches with relatively short bunch lengths at the cathode. It is compatible with a large number of cathode types and with the presence of magnetic fields in the cathode area. As the high RF frequency implies relatively small resonant structures, the power density on the cavity walls becomes very high when the cavity is operated at high gradients, thereby limiting the maximum achievable repetition rate to a few 100 Hz to ~ 10 kHz [177], depending on the gun RF frequency, gradient and overall design. As the small RF structures do not allow for large pumping apertures, the overall vacuum performance can be limited.

User demand for operating at high repetition rates exceeding ~ 10 kHz has pushed some gun designers to investigate schemes using normal conducting RF structures with relatively low frequency, ranging from ~ 100 to 700 MHz. At such lower frequencies, the resonant structures become larger, and the power density on the walls decreases. Below a certain frequency the RF cavities can start running in continuous wave (CW) mode. However, lower frequencies also imply lower accelerating fields than their normal-conducting higher frequency counterparts, although still substantially higher than DC fields. Like all normal conducting RF guns, these schemes are compatible with the presence of magnetic fields in the cathode region. Several guns of this type have been proposed [178–181], and a room-temperature gun operating at 700 MHz and using a very sophisticated cooling system capable of dissipating the extremely high heat load has been designed and successfully conditioned at full RF power [182]. If the frequency is pushed low enough (less than ~ 400 MHz), then the power density on the walls becomes sufficiently small to be handled by conventional cooling techniques. Such schemes based on mature RF common in storage ring RF systems offer a remarkable simplicity and reliability. Additionally, because of the longer wavelength, remarkably large apertures can be opened in the cavity walls, allowing the high vacuum conductance necessary for the low pressure needed for the operation of high quantum efficiency semiconductor cathodes. A gun using a cavity resonating at ~ 200 MHz is presently under construction at LBNL [183].

Table 3 shows the present best beam performance obtained by the different gun technologies.

In addition to the electron gun, an injector must include other systems to obtain the beam characteristics required by FELs. Magnetic elements such as solenoids in synergy with accelerating sections (RF booster) perform and consolidate emittance compensation. Compression schemes including RF prebunchers and/or bunchers or de-phasing of the RF field in the booster cavities (velocity bunching) [184] are used in injectors for shortening the bunches and increasing the peak current, thereby relaxing the compression requirements in the downstream linac. Collimators and control of the beam distribution tails, especially at high repetition rates, must be used in injectors to control beam losses and minimize radiation issues. Last but not least, a complete set of beam and laser diagnostics must be distributed along the injector to monitor and tune the injector parameters. This layout is nontrivial because of (1) the limited space available between components, (2) the very different beam characteristics along the injector, and (3) the presence of space charge forces. In practical terms, selecting the technology for the electron gun is only the beginning; a complete injector system must be designed and constructed around it to exploit its full potential and to overcome its limits.

Table 3
Beam performance of existing electron guns.

Gun	Technology	Repetition rate (Hz)	Field (MV/m)	Energy (MeV)	ϵ_n (μm)	Charge (pC)
LCLS	NC RF 3 GHz	120	140	6	0.5(0.14)	250(20)
PITZ	NC RF 1.3 GHz	10, 800 μs pulse 1 MHz structure	60	> 5	1.3	1000
JLab	DC	75×10^6	6	0.35	3	140
SCSS	Pulsed DC	60	~ 10	0.5	0.6	300
Rossendorf	SRF	125×10^6	5 (damaged cavity)	~ 1	3	80

Clearly, no one injector technology has fully demonstrated the capability of operating with every FEL scheme; many fundamental challenges still need to be addressed. A focused R&D program is necessary to solve such open issues. Highly desirable are injector test facilities in which the beam is accelerated up to energies that freeze the emittance (few hundred MeV) and where emittance compensation, emittance exchange, beam manipulation, compression schemes and beam diagnostics can be efficiently tested. In particular, because the repetition rate has a large impact on the accelerator technology to be used, two different injector test facilities would be useful, one for high (\sim MHz) and another for low repetition rates. The high repetition rate test facility would allow demonstrating the still unproven capability of generating FEL quality beams at high repetition rate, to develop and test high repetition rate beam diagnostics, and to perform R&D on high quantum efficiency cathodes fundamental to operate the high repetition rate FELs. The low repetition rate facility would allow investigating and improving the beam dynamics at high accelerating gradients and performing research on low QE (metal) cathodes targeting lower thermal emittances.

7.3. Injectors for an XFEL

One concept [159,160] for an ultra-low emittance injector capable of producing electron bunches at MHz repetition rates begins with a thermionic cathode embedded in a low frequency CW, 100 MHz rf cavity similar to one under construction at LBNL [177]. Electrons from most of the rf cycles are deflected out to a dump, the remaining few MHz portion of the beam being transmitted through a slit that selects out 0.5 ns pieces in each rf cycle via energy filtering. The beam is then subjected to various manipulations to produce the desired characteristics. Preliminary simulations show that the concept is promising. If successful, the injector would also be useful for ultra-fast SASE with a MHz repetition rate.

8. RF acceleration and power

While the injector technology will have a major impact on the performance (and limits) of any new linac-based light source, the capital and operating cost will be largely determined by the approaches chosen for rf acceleration and power. The two major technical approaches for rf acceleration of high energy electrons, copper accelerators and superconducting accelerators, have rather different capabilities and technical issues. Each is treated separately below, followed by a discussion and of other technical issues that can have a major impact on the performance and cost of next generation light sources.

8.1. Copper accelerator technology

Linacs based on room temperature copper accelerating structures have been used for discovery science for decades. The

technology and beam dynamics are well understood at frequencies from the FM broadcast band and extending to X-band, about 12 GHz. Reliable power sources to power the linac exist at many frequencies, including the L, S, C and X bands. In most cases, they are commercially available.

The main demonstrated advantages are: (1) large accelerating field, up to 35 MV/m at C-band [185] and 100 MV/m at X-band [186], and (2) well-developed technology, with a cost/GeV smaller than \$15M. Balancing these significant advantages is the limited repetition rate, generally less than 1 kHz, and greater potential for wakefield effects due to smaller beam apertures. The maturity of this approach and the sizable established infrastructure that allowed the development of the LCLS X-ray FEL on such a short time scale and within a relatively modest budget. The success of that program and the ease with which the experimental user program is getting underway is a testament to the years of experience on such machines.

These systems are ideal for smaller facilities and exploratory developments, but are limited in delivering the higher average brightness possible in a large scale user facility, with beams switching at high rates between many undulators. Despite the maturity of the technology, work is still in progress to optimize such systems for light source applications. For example, research is being actively done to increase the accelerating field for X-band structures to 100–150 MV/m, mostly for the TeV class lepton linear collider but with application to more compact hard X-ray sources [187].

Another line of development, of particular interest for light sources, is increasing the repetition rate increase to 1–10 kHz, albeit at lower accelerating fields. For light sources the beam energy is limited to less than 2.5 GeV for soft X-rays and less than 30 GeV for hard X-ray systems. Therefore, the linac design can be optimized for maximum repetition rate instead of maximum accelerating field; by reducing the peak accelerating field and increasing the structure frequency, the 1–10 kHz levels can be reached with limited input average power. An additional advantage of this approach is that the modulator-klystron system powering the linac is optimized for low peak power, and thus can be cheaper and more reliable.

Light sources based on room temperature linacs, with a repetition rate of 1 kHz or higher, offer an interesting alternative to those based on the more expensive superconducting linac alternative. This is particularly important for light sources optimized for one particular area of science, like ultra-short X-ray pulses. Because of the much smaller cost, it is possible to imagine a number of these facilities, each optimized for one particular application.

8.2. SRF accelerator technology

For next generation linac based light sources it will be highly desirable to have bunch rates in excess of 100 kHz, possibly higher than 1 MHz with fast switching between undulator lines, to provide high average brightness X-rays to as many users as possible. Combined with the need for high average accelerating

gradient to keep the machines to manageable size, such rates drive the requirement for CW superconducting RF (CW SRF) technology. Fortunately large-scale installations such as CEBAF [188] have shown that this technology is mature enough to make such facilities feasible. Although much more energy efficient at low to moderate beam loading than normal-conducting RF, the operating costs of high gradient CW SRF linacs are nevertheless sizable. Capital costs of installed SRF as well as large-scale cryogenic capability are cost drivers for potential next generation FEL projects. Average currents of the order of 1 mA imply substantial installed RF power capacity (e.g., a 1 GeV linac with 1 mA beam current requires more than 1 MW of RF power). RF power can be supplied by a large number of relatively small sources, giving maximum flexibility of operation, or fewer higher power sources with more sophisticated control and distribution may possibly reduce costs as discussed below, provided the beam emittance can be preserved. Although SRF requirements for FELs and energy recovery linacs (ERLs) are similar in many ways, optimization may be significantly different. FELs typically require high peak and modest average currents, placing a premium on wakefield control and emittance preservation. This choice may favor cavities with larger apertures and highly symmetric structures to minimize transverse kicks (see Fig. 9), for example from HOM or power couplers; however the relaxed HOM damping requirements may allow longer structures (typically up to 9 cells per cavity are being considered [189]), yielding higher packing factor. The machine will have extremely tight stability requirements for energy and phase jitter [190], and require very low trip rates for user operations.

For CW SRF the most significant factor in operating cost is the efficiency of the accelerating system. For a given cavity geometry efficiency translates to the highest possible quality factor Q_0 . The achievable Q_0 depends on many factors, including material properties, processing history and surface morphology, as well as on operating frequency and temperature. The well-known BCS theory [191] describes the ideal variation of superconducting cavity surface resistance with temperature and frequency and implies that the optimum operating point would be at the lowest practical temperature and frequency. The actual surface resistance of presently produced niobium cavities deviates from this ideal behavior and asymptotes to a residual surface resistance significantly higher than BCS theory predicts [192]. The large variation in practically achieved surface resistance in cavities is as yet poorly understood. The typical average value is high enough above BCS that the theoretical gains from lower temperature and frequency are not realized in practice. Consequently, most present or proposed CW machines remain at relatively high frequency (1.3–1.5 GHz). The origins of anomalous losses at typical operating gradients are the subject of ongoing investigation; any

advance in this area will pay large dividends in terms of usable gradient and overall facility costs. Recent excellent results with electro-polished cavities [193] suggest that higher operating Q 's with fields in the range of 20–25 MV/m may be reliably attainable. If these results prove to be typical, they may shift the cost-optimal operating point to higher gradient. Field emission can seriously degrade the Q_0 and limit the usable gradient, if cavities are imperfectly processed or mishandled after cleaning. Although great strides have been made in combating field emission, including recent tests to over 35 MV/m with no detectable X-rays [194], elaborate procedures will be necessary to ensure this performance can be achieved reliably for a large ensemble of cavities. Care must also be taken in the design of magnetic shielding in the cryomodule to achieve the full potential of the cavities. In practice the ideal shielding configuration is often compromised by the many penetrations necessary for tuners, couplers, etc.

The choice of frequency may be influenced by other factors besides SRF operating efficiency. Lower frequency cavities may accommodate higher charge per bunch, but have longer RF buckets; depending on the detailed user beam requirements this feature may or may not be advantageous. Final choice of operating frequency, structure type and cryogenic temperature should be the result of a complex optimization and may be quite unique to a specific facility. A number of light-source optimized cavity designs are under development worldwide [195]; they vary considerably in frequency, cell shape, number of cells, HOM damper type and power coupler configuration. An important milestone in cavity development will be beam tests of these designs. The best verification of HOM damping, microphonics, power coupler, etc. is with beam in a real machine or test facility. Several such facilities exist or are proposed globally [196,197].

For cavities with such a high natural Q factor a certain amount of RF power overhead is required to ensure stability of the cavity gradient when the structure undergoes small tuning excursions due to external disturbances. This microphonic effect can be measured and to some degree mitigated by careful design, stiffening of the structure, and good isolation or active feedback, but nevertheless places a practical limit on the maximum external Q that can be operated stably. For linac based FELs, microphonic detuning is typically not a problem in practice as the optimum external Q needed to compensate for beam loading is much lower than this limit, and there is a modest RF overhead required to ensure amplitude regulation at the maximum frequency excursions.

Two recent parametric studies have been performed for proposed light source facilities; the Cornell ERL [198] and the UK NLS FEL [199]. The optimal configurations for each are quite sensitive to detailed assumptions in the models; however, they

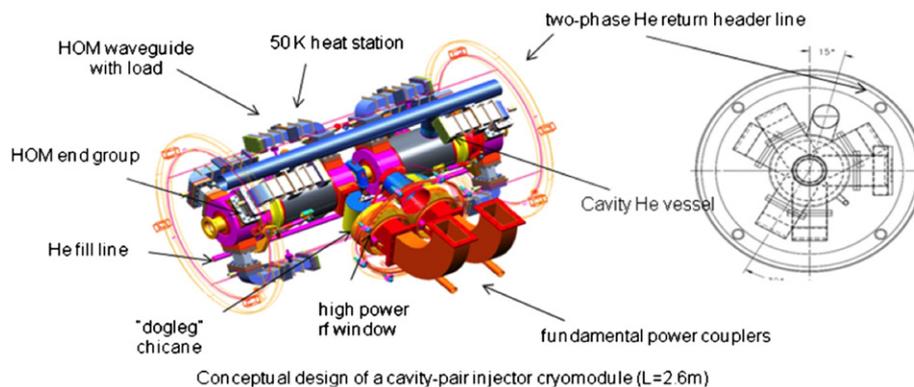


Fig. 9. A highly symmetric SRF cavity. The trifold symmetry of the HOM damping minimizes beam kicks while the large apertures minimize wakefield generation.

share several common features. Both find a broad cost minimum (capital+10 years operating costs), between about 15–25 MV/m; both chose gradients in the lower half of this range to be conservative. Both chose 1.3 GHz as the preferred operating frequency. The optimum operating temperature could be as low as 1.8 K depending on the assumed residual resistance. As might be expected, the ERL study favors shorter (7-cell), strongly HOM damped cavities, while the FEL study uses LLC-like 9-cell cavities. Some variables were not included in these studies such as the variation of RF power costs, residual resistance and optimum operating temperature with frequency, and the relationship between end-use optical output and electron beam properties as influenced by all of the above. A comprehensive evaluation including such variables should be undertaken for any new major facility.

The biggest cost leverage would be obtained by improvements in Q_0 at the operating gradient. Research is ongoing in this area, including studies of the high-field Q drop [200] and origins of residual resistance. Improvements in field emission free cavity processing would provide more confidence in higher operating gradients and could also allow cavity shapes to be contemplated that have lower operating losses but higher surface electric fields. Improvements in HOM damping, packing factor in the cryomodule, static losses and construction costs would all be worthwhile. Success in this regard might ultimately allow solid-state amplifiers to be used with attendant simplifications and reliability.

8.3. Supporting accelerator technology

In addition to the main areas of research on optimizing accelerator structures for light source application there are additional areas of accelerator R&D that need resolution to minimize the cost and maximize performance of a next generation light source. We will concentrate on technology development for the SRF system since a major user facility will require the high repetition rates provided by SRF and the copper technology is relatively mature. However, R&D to develop high repetition rate room temperature copper linacs should be part of any program of research on accelerators for soft and hard X-ray FELs.

The key cost driver of a facility is the linac itself with the associated conventional facility, cooling, and rf power. A sizeable reduction amounting to a significant fraction of the total facility cost (perhaps \sim \$1B for a multipurpose, next generation facility with many undulators and photon beam lines) could be achieved if cost reductions in each of these technologies prove viable. One idea under consideration for the linac is the use of recirculation, sending the electron beam through the same structures multiple times. While recirculation requires the addition of a magnetic transport lattice for each pass, such lattices cost substantially less than superconducting linac structures. Thus if the linac could be one-half to one-third the single pass length, a savings of several hundred million dollars could be realized.

Recirculation would also reduced costs substantially for conventional facilities, rf power, and cooling systems. This technique was successfully employed in the CEBAF accelerator at Jefferson Lab, with the beam passing through the linacs five times before being directed to the nuclear physics target areas. The application to light sources is more difficult, particularly for soft or hard X-ray FELs, because the electron bunch charges are much higher and the demand for ultimate electron beam brightness at the wiggler is greater. Bending electron beams leads not only to emittance dilution but also a number of potentially serious beam heating effects such as coherent synchrotron emission, wakefields and longitudinal space charge effects. These are discussed in more detail in Section 11. Presently it is unknown

whether successful approaches can be found to bend electron beams of say, 200 pC multiple times for recirculation without significantly damaging the brightness; however, the prospect of substantial cost savings is sufficient to warrant experimental studies.

A second area of potentially beneficial R&D is the use of a single rf source to drive multiple cavities. While copper cavity linacs use this approach, it is typically avoided in SRF machines because the cavities are subject to microphonic noise, which leads to high demands on rf control. Frequency shifts of 10 Hz would not be uncommon in L-band cavities with $Q_0 > 10^9$. This shift increases the required control power substantially and, since klystrons typically use full electric power whether or not they are operating at fully saturated output, electricity costs also increase. Such demand could be reduced if multiple SRF cavities were driven from one rf source. Typically the cost of an rf system is very nonlinear with power. There is a significant buy-in cost regardless of power that increases slowly as the average power is raised. Thus driving 5 cavities from one rf source might only cost half what five separate systems would cost and use one-third the electricity. Since accelerating 1 mA of current by 20 MV in an srf cavity requires only 20 kW of power (plus perhaps 10 kW of control overhead) having five cavities driven from one 150 kW tube is a practical approach. Studies are required to determine whether the rf control to 0.01° of rf phase can be achieved while providing power from one source to multiple cavities.

Another major cost driver for SRF machines is the helium refrigeration cost. The cost of a 5000 W cryoplant at 2 K is now in the range of \$20M, and it uses on the order of 5 MW of electricity continuously. Significant progress has been made in the last 5 years in improving refrigeration efficiency and reducing capital cost but benefits could accrue from further research. As a single pass 2.5 GeV SRF linac to drive a 1 keV FEL might require as many as three of these cryogenic refrigerators, improvements will have high leverage. Longer term, the development of $\beta = 1$ low frequency (< 500 MHz) cavities may permit operation of such an SRF system at 4 K instead of 2 K which is presently needed. Although there is little experience at the present time with such structures, this approach would effectively double the refrigeration efficiency and halve the capital cost of the helium refrigeration plants. A modest R&D program could effectively begin to determine realistic goals and technical issues in such an approach.

9. Test beds

To assess the integrated performance of new elements of technology or physics models requires testbed facilities. Testbeds that are in operation or under construction in the United States, Europe and Japan are listed in Table 4. The two active testbeds in

Table 4
FEL testbed facilities: status.

Name	E (MeV)	Mode	Status
C-AD ERL	20	SASE	C
BINP ERL	40	Oscillator	O
JLAB	160	Oscillator	O
SDUV	300	HGHG	C
NSLS SDL	300	HGHG	O
SPARC	300	HHG	O
SCSS Proto	300	HHG	O
DESY FLASH	1000	SASE	O
FERMI	1500	HGHG	C

C, construction; O, operation.



Fig. 10. NLSL Source Development Laboratory Testbed.

the USA are the Source Development Laboratory (SDL) FEL [66] at BNL (see Fig. 10) and the IR/VUV FEL at JLAB [71] (see Fig. 3).

While much can be learned through collaborative efforts using these existing machines, two new integrated FEL testbeds in the USA would be highly desirable at the earliest possible time to maximize the cost effectiveness and scientific utility of future user facilities. As both the FLASH and FERMI facilities must support full user science programs, a new testbed capable of radiating in the VUV-soft-X-ray regime, with radiation pulses as short as 1 fs would allow extensive testing and comparison of many seeding schemes as well as the most challenging timing and synchronization techniques. Evaluation of the performance of cathodes, laser systems, emittance compensation, beam manipulation, and diagnostics at MHz rates requires a testbed operating at an energy at which the emittance is frozen ($E \sim 100$ MeV). If that energy is raised to ~ 500 MeV, the limits and utility of recirculation of nC charge bunches can be quantified.

10. Synchronization and timing

Pump–probe experiments require synchronizing the FEL X-ray pulse (the probe) with an external laser. In most cases the time duration of the pump and probe pulses is rather short, from a few to tens of femtoseconds, and the two pulses must be synchronized to within a small fraction of their duration. The problem is clearly easier in the case of pulses with time duration of 100–1000 fs, and we will limit our remarks to the case of the shorter pulses.

The synchronization of the FEL components, from the injector to the linac to other systems, done through the low level RF system and the photocathode laser system, has been developed to a high level of performance. Here we discuss only the ideas that are being considered to synchronize the pulses using the interaction of the electron pulses with external laser signals, and the coherent radiation at long wavelength generated by the electron bunch.

In addition to the temporal jitter due to RF phase fluctuations and to other system fluctuations, the electron beam arrival time at the FEL undulator changes because beam energy fluctuations are translated into temporal fluctuations in the magnetic elements of the beam compressor and transport system. The LCLS group has measured the jitter of the electron bunch near the undulator entrance to be ~ 50 fs rms, but only over 1–2 min; the jitter increases with longer integration time. An additional component of time jitter for SASE FELs may arise because the saturation point fluctuates pulse-to-pulse along the undulator length. This effect is expected to be small, particularly when the undulator length is larger than the saturation length.

The LCLS results show that the timing jitter in X-ray pulses with respect to the FEL clock and external lasers is important when the pulse duration is shorter than 50 fs. This problem can be overcome using one of the ideas developed to imprint an external

laser pulse on the electron bunch before it enters the undulator, thereby selecting the part of the bunch that will lase. One such example is ESASE [32]. In the case of a seeded FEL the synchronization can be achieved using the seed signal itself to synchronize with the same or with an additional laser [201].

This approach can be more difficult when the short pulse duration of the soft or hard X-ray pulse is obtained by reducing the electron bunch charge and length [130,140]. In this case no external laser pulse provides an imprint, and the electron bunch is so short, less than $1 \mu\text{m}$, that it would be hard to hit with an external laser. In this case, however, the coherent radiation of the electron bunch itself, having a length shorter than $1 \mu\text{m}$, when passing next to an aperture or a foil just before the undulator, can generate a signal to compare with that of the external laser to determine the relative time delay.

Testing these concepts on existing FELs is crucial to develop the pump–probe capability of the FEL to its full potential. While synchronizing two pulses at the femtosecond or sub-femtosecond level is a research challenge, there are promising several tools and techniques for its accomplishment.

11. Collective effects

Collective effects in FELs are caused by the electromagnetic field of the beam during its generation, acceleration, and transport through the system. They can degrade the beam quality and decrease the performance of the laser. Extensive studies of the collective effects in FELs have been carried out with the emphasis on coherent synchrotron radiation (CSR) wakefields, space charge wakes and resistive wall wakefields in the undulators.

Coherent synchrotron radiation (CSR) is one of the most challenging issues associated with the design of bunch compressor chicane required for X-ray free-electron laser amplifiers. Typically, CSR is emitted at wavelengths longer than the length of the electron bunch and leads to a detrimental tail–head interaction in bends [80–82]. CSR can also be emitted at wavelengths much shorter than the bunch length if the bunch charge density is modulated at these wavelengths. Computer simulations have shown that small density modulations can be significantly amplified by the CSR reaction force in bunch compressor chicane, giving rise to a microbunching instability [202]. This instability was studied [203–205] as it may impact the design of X-ray FELs calling for kiloampere, subpicosecond electron bunches.

In addition to the CSR wakefield, the longitudinal space charge (LSC) field can also contribute strongly to driving the microbunching instability in the linac [206–210]. The LSC effect dominates at the low-energy end of the accelerator and can be accumulated over large distances in drifts and accelerating structures. In contrast, the CSR effect is localized in bending magnets, and becomes important at higher energies.

Because the microbunching instability is very sensitive to the uncorrelated (local) energy spread of the electron beam, increasing that spread within the FEL tolerance (of order ρ) can provide strong Landau damping of the instability. Proposed countermeasures include a superconducting wiggler [2] and a laser heater that uses the resonant laser–electron interaction in an undulator [206]. A laser heater has been implemented in the LCLS [211], and has proved to be effective in improving performance of the FEL [212]. While the laser heater suppresses the microbunching instability in the LCLS, there remains a certain level of microbunching in the beam that generates coherent optical radiation that presents unique challenges to high-brightness beam diagnostics. The instability may become a more severe constraint in future higher-brightness accelerators and for seeded FELs that demand better control of the longitudinal phase space. Further theoretical and experimental studies of this instability on FEL testbeds will greatly reduce the risks for future light sources.

Collective effects can also play an important role in the undulator, where a typically small-gap vacuum chamber leads to strong interaction of the beam with resistive walls of the chamber. The standard theory of the resistive-wall impedance [213] must be modified to include effects of very short bunches [214]. In addition to wall resistivity in the undulator, the wall roughness can contribute to the overall impedance of the vacuum chamber [215–217]. Therefore surface finish must be controlled during manufacturing of the vacuum chamber.

The successful commissioning of the LCLS proved that the collective effects, to large extent, can be overcome in a hard X-ray FEL [218]. As an example, a comparison of the measured bunch length and the beam emittance after the second bunch compressor

in the LCLS from Ref. [218] is presented in Fig. 11. It shows a good agreement between the simulated and measured properties of the beam.

Recent analysis of geometric and resistive wall wakefields in the LCLS undulator [219] also shows a good agreement between the experimentally measured wakefields and the calculated ones.

Future development of the collective effects theory will be required to characterize and access collective effects in seeded beams which are characterized by a structured phase space of the beam. Generation of ultra-short electron bunches (in the femtosecond range of durations) will bring about collective effects in a parameter region in which many assumptions of the classical theory of wakefield and impedances breaks down. Therefore, new analytical tools must be developed. Finally, better CSR models must be implemented in fast simulation codes.

12. Simulation tools

12.1. Current status

Free-electron laser (FEL) simulation codes based on macro-particle models have become an integral part of the design of new devices and the analysis and diagnosis of existing machines. As one example of the interplay between simulation and experiment, Fig. 11 of Ref. [1] shows a comparison of the LCLS power measured at 1.5 Å with that predicted using the three-dimensional code GENESIS. The close agreement demonstrates that the present FEL codes can be used to predict the performance of new machines; alternatively, the codes can be used to cross-check experimental conditions, including beam emittance or energy spread.

The basic FEL algorithm integrates the equations of motion in the combined undulator and radiation fields for a distribution of macroparticles, each of which might represent thousands of real electrons. The radiation field generated by these simulation particles is obtained by solving the paraxial wave equation. Steady-state codes such as TDA [220] are seeded by a monochromatic external radiation field, and neglect variation along the electron bunch so that typical simulations involve a single slice of the beam chosen to be one radiation wavelength long. Within this approximation one can study such effects as the dependence of FEL gain on the beam parameters with minimal computational resources. In “time-dependent” simulation codes such as GINGER [221], GENESIS [222] and SIMPLEX [223] (which all can be run in the steady-state mode), the electron and radiation beam are divided into many slices, which are integrated in a manner similar to that of the single slice calculations. In this case, however, the radiation carries information between the slices through the slippage effect, which is typically implemented via a discrete slippage model. The model assumes the radiation and the electron bunch do not vary much over a radiation wavelength, which is valid because the FEL gain bandwidth is normally much narrower than the resonant fundamental frequency. Simulations that use the initial electron shot noise to seed the FEL interaction can be modeled by adding random deviations to the uniform distribution. These deviations are chosen so that the fluctuation level of the macroparticles matches that of the actual beam [224]. Fawley [225] devised a more sophisticated shot noise algorithm taking into account 6D phase space and harmonic bunching

The transverse profile of the radiation is typically solved using a discrete spatial mesh, which is taken to be axisymmetric in the code GINGER while GENESIS uses a fully three-dimensional cartesian grid. Both of these codes terminate the field at the transverse boundary, so the simulation grid must encompass the entire radiation field. To relax this constraint, the time-dependent

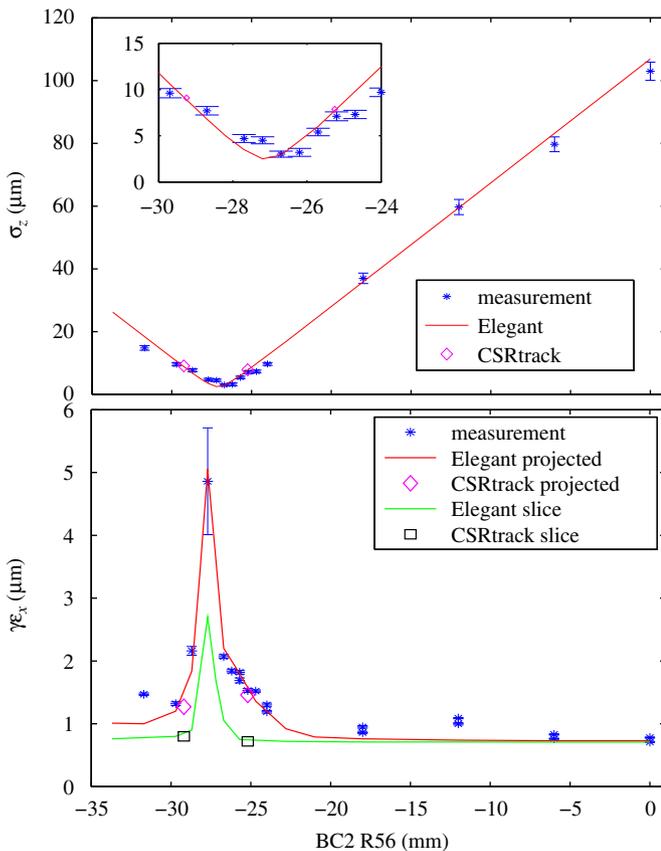


Fig. 11. Bunch length compression and emittance growth measurements and simulations. Top: compressed bunch length after the second bunch compressor at 250 pC. Bottom: horizontal emittance after the second bunch compressor at 250 pC.

code ALICE implements open boundary conditions for the paraxial wave equation and a mesh that is localized near the electron beam [226], while an integral representation of the paraxial wave equation is used in RON [227] and FAST [228] to allow for efficient field computation in the vicinity of the electron bunch. All of the above codes are based on the wiggler-period-averaged FEL equations, which assume the field develops slowly over many undulator periods. The harmonic emission in these codes is included *ad hoc*. A non-wiggler-averaged code, MEDUSA, has been developed to include higher harmonics from the beginning [229]. Finally, several codes are well-developed for realistic FEL simulations, especially for single-pass, high-gain systems. They can read as input 6D macroscopic electron distributions from accelerator tracking codes and take into account additional effects in the undulator such as wakefields, space charge and energy spread diffusion due to incoherent undulator radiation.

12.2. Principal challenges

With the rapid advances in FEL R&D, we illustrate a few principal challenges for advanced FEL simulations.

Many novel FEL schemes incorporate a combination of magnetic chicanes and undulator sections to manipulate the electron beam phase space in order to improve the properties of the subsequently generated radiation. For instance, optical klystron FELs use chicanes to enhance the FEL microbunching in a relatively small distance, while schemes such as high gain harmonic generation (HG) and echo-enhanced harmonic generation (EEHG) use chicanes to produce microbunching at a harmonic of a seeding laser. The current chicane model used in existing codes is rather crude and does not include transverse dynamics and radiation effects. These effects may be very important in determining the ultimate performance of these advanced FEL schemes.

Improvements in the generality of the X-ray transport and optics will also be important to properly model self-seeding schemes and to address the cavity requirements for X-ray FEL oscillators. Specifically, determining the effects of initial noise (in the case of self-seeding schemes) and to mirror imperfections and misalignments (for the X-ray oscillator) will be crucial for assessing the various effects of imperfections and for determining the design requirements on any potential machine.

Thanks to parallelization of several codes, run times are reasonable for the most demanding single-pass SASE FELs. For example, a full LCLS simulation with 4×10^8 gridpoints and 6×10^8 macroparticles requires about 200 CPU hours per run. Efforts to simulate the X-ray FEL oscillator using parallel codes have received less attention; while many simulations can be performed in reasonable time on a single processor, some applications may require improvements to the parallel scaling in order to take advantage of state-of-art supercomputers.

The discrete slippage model used in most time dependent codes is convenient for sequential progression through the electron bunch and requires very modest computational memory. However, it also imposes several limitations on the simulations; for example, within this model there is no simple way to exchange particles amongst different radiation slices, and the radiation bandwidth is restricted to a narrow range. In order to relax these limitations, a new algorithm or perhaps a new FEL code should be developed that takes full advantage of the passively parallel computer architectures to solve the time-dependent FEL equations. With increased computational resources one could have all the particles in memory simultaneously, thereby permitting the exchange of particles among the different radiation slices and the implementation of more general

slippage models. These developments would enable simulations of emerging concepts including echo-enhanced microbunching and the modeling of extreme beams such as those that might be produced with laser plasma accelerators.

Last but not least, these advanced codes should also improve the ability to take input from and give output to accelerator tracking codes such as ELEGANT/IMPACT. While certain particle data can be exchanged to perform start-to-end FEL simulations, improved interfaces should also more readily handle the transfer of the microscopic beam distribution. This option would permit a more accurate modeling of the creation and subsequent transport of electron pre-bunching that might be used in advanced seeding or harmonic generation schemes, and will also help assess the effects of particle noise on radiation generation.

13. Compression and transport

A substantial challenge in delivering a high brightness electron beam to an FEL undulator is preserving both the transverse and longitudinal emittances while compressing the bunch length and accelerating the electrons. Collective effects such as coherent synchrotron radiation (CSR) and space charge forces in magnetic bunch compressors, anomalous dispersion due to component misalignments, and strong wakefields of the RF accelerating structures, can quickly destroy the hard-earned brightness generated at the electron source. In addition, the final peak current and arrival time of the compressed beam can become very sensitive to RF phase and amplitude jitter, requiring care in the design and attention to tight stability tolerances. Nevertheless, systems are presently operating with sub-10- μm bunch length compression and high beam brightness [1]. Finally, rapid distribution of electron bunches to many separate beamlines introduces new stability challenges.

13.1. Coherent synchrotron radiation (CSR) and space charge (SC)

In recent years, the effects of CSR on electron beam quality have been studied extensively in theory and simulations (see Ref. [230] for a recent review). Examples of measurements can be found in Refs. [231–235,218]. Some of the most recent measurements at the LCLS [218] for two very different electron energies (0.25 and 4.3 GeV) suggest that the computer codes presently available do a reasonably accurate job of predicting the transverse dilution effects for a fairly wide variety of parameters and settings, even when applying a fast and simple 1D line-charge approach such as is available in the particle tracking computer code ELEGANT [236]. The much slower 3D codes are then used for occasional verification, or for applications with extreme parameters for which a self-consistent model may be required. In addition, strong space charge forces can dilute the beam brightness if extreme compression is applied at a low enough energy. Machine designs therefore usually incorporate more than one compression stage, attaining the final bunch length only after the electron energy is high enough. These issues frequently become a design limitation for highly compressed, high-brightness electron beams and require great care and judicious parameter choices. Velocity bunching methods have also been proposed and tested [237], compressing the electron bunch without using magnetic dipoles, thereby avoiding the effects of CSR, although space charge forces must be compensated as well. Typically, velocity bunching is used as a pre-compressor within the injector system, with at least one more magnetic compression system being employed at higher energy, where necessary.

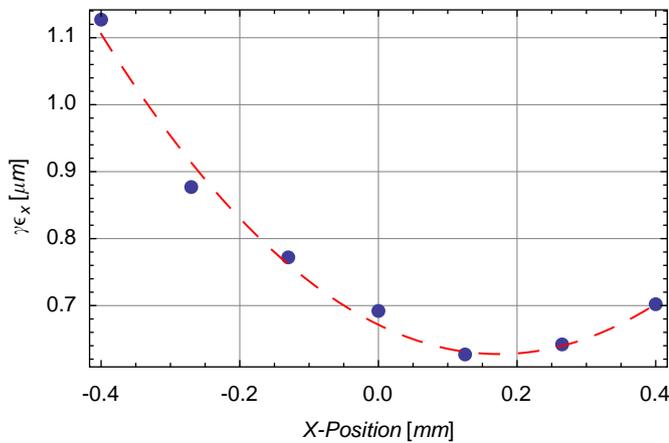


Fig. 12. Measured normalized emittance minimized by steering the electron beam through the small-iris X-band LCLS RF structure.

13.2. Wakefield emittance dilution

Another significant source of beam brightness degradation is produced with misaligned RF accelerating structures and the accompanying transverse wakefields. The transverse emittance growth increase with longer bunch length, higher bunch charge, and smaller iris (high gradient) RF structures. Fortunately, the wakefield emittance dilution can be minimized by steering the electron beam carefully through the linac using a beam-based alignment approach. An example is shown in Fig. 12 featuring a small iris X-band (11.4 GHz) RF structure.

13.3. Compression stability

The FEL gain is strongly enhanced by increasing the peak current of the electron beam, which is typically accomplished with bunch length compression. The strong compression results in a high sensitivity to RF phase (and amplitude) stability for the linac sections which generate the electron energy chirp just upstream of the compressor. The required phase stability of a single compression stage can be estimated using [238]

$$\Delta\varphi \leq \left| \frac{\varphi_0}{C_0} \frac{\Delta I_{pk}}{I_{pk}} \right| \quad (17)$$

where C_0 is the compression factor (e.g., $C_0=50$), $|\Delta I_{pk}/I_{pk}|$ is the tolerable relative peak current jitter (e.g., $< 10\%$), and φ_0 is the nominal RF chirp phase (e.g., $\varphi_0 = 20^\circ$). The example parameters lead to a required phase stability of 0.04° , which is typical. Multi-stage compressors and harmonic RF systems, which are also used to linearize the compression [239], can be used to relieve these shot-to-shot tolerances to some degree and beam-based feedback systems are required to maintain the peak current over longer time periods.

13.4. Beam switching

Electron beam switching to multiple undulator lines at a low repetition rate is typically not a problem. However, switching could be challenging at a repetition rates of 10 kHz or more. Recent advances in a technology for high voltage pulsers (see for example Ref. [240]) makes such rapid switching feasible, although the effects of any small residual fields on emittance should be measured.

14. Conclusions

With the well-established success of FEL oscillators in the longer wavelengths, the new science enabled by SASE FELs such as FLASH in the XUV, and the stunningly rapid commissioning of the hard X-ray FEL at LCLS, it has become clear that free electron lasers are ready to offer broad-based user communities the highest quality, tunable, intense, coherent light. These revolutionary photon sources will open up a vast scientific frontier to probe matter with substantially finer length, time, and energy resolution, where physical, chemical, and biological systems can be viewed on their critical temporal, spatial, and energetic scales—femtoseconds, nanometers, and millivolts.

The primary goals for future development are to move forward from this solid base to achieve the full promise of FEL performance and to provide a genuine user-focused infrastructure on a par with that of storage ring light sources. In soft and hard X-rays, high longitudinal coherence, in addition to full transverse coherence, will be the key performance upgrade, and ideas using laser or self-seeding or oscillators can be expected to qualitatively beat today's SASE sources. Short pulses, from femtoseconds to attoseconds, can be realistically envisioned. With high repetition rate electron sources coupled to superconducting radiofrequency linear accelerators, unprecedented average beam brightness will be possible and many users would be served simultaneously by a single accelerator complex.

Most of these goals are well within reach, and with a focused R&D program, facilities with these new capabilities should be achievable within this decade. In this paper, we have provided details of the scientific and technological basis that supports the feasibility of these goals and have laid out detailed plans for enhancing the experience base to proceed to facility construction. To reiterate, a vigorous, well-supported R&D program could bring transformational FEL-enabled science to a wide research community in this decade.

Specifically, there are a number of critical research areas that must be addressed now to move forward effectively toward major cutting edge user facilities or smaller scale revolutionary devices. The non-prioritized list includes:

- Low emittance, high repetition rate injectors, delivering a broad range of bunch charges. These will provide the critical technology for high repetition rate, high average power soft X-ray seeded FELs, as well as for X-ray oscillators and for advanced SASE schemes for more full coherent hard X-rays.
- Cost minimized SRF linac systems, including recirculation. SRF linacs offer the possibility of high average power and many simultaneous users. As the present cost/MeV is considerably higher than pulsed room temperature technology, progress in cost optimization will greatly enhance the value offered by SRF technology.
- Optical cavities for X-ray oscillators. At wavelengths from XUV to hard X-rays, losses must be minimized to allow effective lasing. Both Bragg crystal and high reflectivity mirror approaches are promising.
- High average power lasers for photocathode guns, and, more importantly, for laser seeding, high harmonic generation (HHG), and beam manipulation. Although the necessary single pulse energies have been achieved, at high repetition rates (above 10 kHz), the average laser powers required are beyond today's state of the art.
- Advanced techniques for synchronization and timing at the femtosecond level. To take full advantage of femtosecond and sub-femtosecond pulses, for example, for pump probe

- experiments, requires synchronization of the FEL with a variety of user subsystems.
- Multiplexing technology for FEL undulator farms. The ability to feed multiple users simultaneously will reduce cost per researcher dramatically.
 - Undulators optimized for FEL user facilities. Higher field and shorter wavelength undulators will translate directly into lower electron beam energies and lower costs.
 - Theory, modeling, and experiments in FEL and beam physics. Although LCLS has demonstrated that beam physics limits can be understood there is the potential for more cost effective system designs that can be realized if the beam physics, especially compression and collective effects, is better understood and the modeling more predictive. Promising innovations in FEL operation press the approximations made in existing FEL computer codes, especially with respect to limitations on longitudinal coherence, seeding, and oscillator operation.
 - Cost effective test bed facilities to allow direct experimental confirmation of innovative designs. Possibilities include a 1–2 GeV linac to examine options for ultra-short pulses (to sub-fs) and various seeding schemes and a lower energy CW linac facility to address high repetition rate issues.

Success in these areas will clear the path to the design and construction of FEL user facilities that will generate photon beams with an unprecedented level of performance. Progress can be rapid with resolute support.

Acknowledgements

The authors thank the US Department of Energy Office of Basic Energy Sciences for hosting a workshop to explore the status of the R&D needs for future light sources. J.B. Murphy and J. Bisognano served as the FEL working group co-chairs. We thank P. Channel for helpful comments and W. Fawley and S. Reiche for providing valuable input concerning computer simulation of FELs. We are very grateful to J. Chew for several critical readings of the manuscript.

References

- [1] P. Emma, for the LCLS Team, in: Proceedings of the 2009 Particle Accelerator Conference, IEEE, Vancouver, Canada, 2009.
- [2] Linac Coherent Light Source (LCLS) Conceptual Design Report, SLAC Report No. SLAC-R-593, 2002.
- [3] L. Elias, et al., Phys. Rev. Lett. 36 (1976) 717.
- [4] D.A.G. Deacon, et al., Phys. Rev. Lett. 38 (1977) 892.
- [5] Science and Technology of Future Light Sources, a White Paper by ANL, BNL, LBNL, SLAC collaboration, December 2008.
- [6] H. Motz, J. Appl. Phys. 22 (1951) 527.
- [7] R.M. Phillips, Trans. IRE Electron Devices 7 (1960) 231.
- [8] J.M.J. Madey, J. Appl. Phys. 42 (1971) 1906.
- [9] J.M.J. Madey, H.A. Schwettman, W.M. Fairbank, IEEE Nucl. Sci. 20 (1973) 980.
- [10] A.M. Kondratenko, E.L. Saldin, Part. Accel. 10 (1980) 207.
- [11] R. Bonifacio, C. Pellegrini, L. Narducci, Opt. Commun. 50 (1984) 373.
- [12] J.B. Murphy, C. Pellegrini, Laser Handbook, Elsevier, London, 1990, p. 9.
- [13] J.B. Murphy, C. Pellegrini, J. Opt. Soc. Am. B 2 (1985) 259.
- [14] E.T. Scharlemann, A.M. Sessler, J.S. Wurtele, Phys. Rev. Lett. 54 (1985) 1925.
- [15] G.T. Moore, Opt. Commun. 52 (1984) 46.
- [16] K.-J. Kim, Phys. Rev. Lett. 57 (1986) 1871.
- [17] S. Krinsky, L.-H. Yu, Phys. Rev. A 35 (1987) 3406.
- [18] K.-J. Kim, Characteristics of synchrotron radiation, in: M. Month, M. Dienes (Eds.), Proceedings of the US Particle Accelerator School, No. 184, AIP Conf. Proc. 565 (1989).
- [19] W. Barletta, C. Rizzuto, Riv. Nuovo Cimento 29 (6–7) (2007) 1.
- [20] R. Colella, A. Luccio, Opt. Commun. 50 (1984) 41.
- [21] W.B. Colson, Phys. Lett. 64 (1977) 190.
- [22] K.-J. Kim, Y. Shvyd'ko, S. Reiche, Phys. Rev. Lett. 100 (2008) 244802.
- [23] R.B. Palmer, IEEE Trans. Nucl. Sci. NS-28 (1981) 3370.
- [24] A.A. Zholents, K. Holldack, in: Proceedings of the Free Electron Laser Conference 2006, Berlin, 2006.
- [25] K.-J. Kim, in: X-ray Data Booklet, Section 2, LBNL/Pub-490, 2001.
- [26] R.W. Schoenlein, et al., Science March 24, 2000.
- [27] S. Khan, et al., Phys. Rev. Lett. 97 (2006) 074801.
- [28] A. Streun, et al., in: Proceedings of the EPAC06, THPLS062, Edinburg, 2006.
- [29] E.L. Saldin, E.A. Schneidmiller, M.V. Yurkov, DESY Report No. TESLA FEL-2003-02, 2003.
- [30] Z. Huang, M. Borland, P. Emma, J. Wu, C. Limborg, G. Stupakov, J. Welch, Phys. Rev. ST Accel. Beams 7 (2004) 074401.
- [31] P. Emma, et al., First results of the LCLS laser-heater system, PAC 2009, Proceedings, 2009.
- [32] A.A. Zholents, Phys. Rev. ST Accel. Beams 8 (2005) 040701.
- [33] A.M. Sessler, D.H. Whittum, L.-H. Yu, Phys. Rev. Lett. 68 (1992) 309.
- [34] A.A. Zholents, Phys. Rev. ST Accel. Beams 8 (2005) 050701.
- [35] P. Sonka, Part. Accel. 11 (1980) 45.
- [36] B.M. Kincaid, et al., AIP Conf. Proc. 118 (1984) 110.
- [37] R. Bonifacio, et al., Nucl. Instr. and Meth. A 296 (1990) 787.
- [38] L.-H. Yu, Phys. Rev. A 44 (1991) 5178.
- [39] L.-H. Yu, et al., Phys. Rev. Lett. 91 (2003) 074801.
- [40] E.L. Saldin, E.A. Schneidmiller, M.V. Yurkov, Opt. Commun. 202 (2002) 169.
- [41] J. Feldhaus, E.L. Saldin, E.A. Schneidmiller, M.V. Yurkov, Opt. Commun. 140 (1997) 31.
- [42] FERMI@Elettra Conceptual Design Report, available at <http://www.elettra.trieste.it/fermi/>.
- [43] J.N. Corlett, et al., FEL 2004, THPOS51, Trieste, Italy, 2004.
- [44] F. Krausz, M. Ivanov, Rev. Mod. Phys. 81 (2009) 163.
- [45] B. Dromey, et al., Nat. Phys. 2 (2006) 456.
- [46] T. Watanabe, et al., in: Proceedings of the FEL05, vol. 98, 2005.
- [47] M. Gullans, G. Penn, A.A. Zholents, Opt. Commun. 274 (2007) 167.
- [48] G. Stupakov, Phys. Rev. Lett. 102 (2009) 074801.
- [49] D. Xiang, G. Stupakov, Phys. Rev. ST Accel. Beams 12 (2009) 030702.
- [50] W.B. Colson, J. Blau, K. Cohn, J. Jimenez, R. Pifer, in: Proceedings of FEL2009, Liverpool, UK, 2009, p. 591.
- [51] S.W. van Amersfoort, et al., in: Proceedings of EPAC 1988, available at <www.jacow.org>, 1988, p. 430.
- [52] R.T. Jongma, et al., in: Proceedings of FEL2008, Gyongju, Korea, available at <www.jacow.org>, 2008, p. 200.
- [53] <http://www.magnet.fsu.edu/usershub/scientificdivisions/emr/facilities/fel.html>.
- [54] L.R. Elias, et al., Nucl. Instr. and Meth. A 237 (1985) 203.
- [55] M.J. van der Wiel, P.W. van Amersfoort, FELIX Team, Nucl. Instr. and Meth. V 331 (1–3) (1993) 30.
- [56] <http://www.rijnh.nl/n4/n3/n5/f1234.htm>.
- [57] N. Vinokurov, et al., in: Proceedings of the 2009 Accelerator Conference, Vancouver, Paper MO4PB102, 2009.
- [58] T.J. Orzechowski, et al., Phys. Rev. Lett. 57 (1986) 17.
- [59] D.C. Nguyen, et al., Phys. Rev. Lett. 81 (1998) 810.
- [60] M.J. Hogan, et al., Phys. Rev. Lett. 81 (1998) 4867.
- [61] S. Milton, et al., Science 292 (2001) 2037.
- [62] A. Tremaine, et al., Phys. Rev. Lett. 88 (2002) 204801.
- [63] L.-H. Yu, et al., Phys. Rev. Lett. 91 (2003) 074801-1.
- [64] X.J. Wang, et al., in: Proceedings of the FEL2006, 2006, p. 18.
- [65] E.L. Saldin, E.A. Schneidmiller, M.V. Yurkov, Opt. Commun. 148 (1998) 383.
- [66] J.B. Murphy, X.J. Wang, Synchrotron Radiat. News 21 (1) (2008) 41.
- [67] G. Lambert, et al., in: Proceedings of FEL08, Gyeongju, Korea, Paper TUPPH080, available at <www.jacow.org>.
- [68] L.-H. Yu, et al., Science 289 (2000) 932.
- [69] J. Wu, L.-H. Yu, Nucl. Instr. and Meth. A 475 (2001) 104.
- [70] FERMI@Elettra Conceptual Design Report, available at <http://www.elettra.trieste.it/fermi/>.
- [71] S. Benson, et al., Nucl. Instr. and Meth. A 582 (2007) 14.
- [72] D. Heche, et al., Appl. Surf. Sci. 253 (2007) 8041.
- [73] R. Rox Anderson, et al., Lasers Surg. Med. 38 (2006) 913.
- [74] G. Lupke, X. Zhang, B. Sun, A. Fraser, N.H. Tolk, L.C. Feldman, Phys. Rev. Lett. 88 (2002) 135501.
- [75] US Basic Energy Sciences Advisory Committee, Next-Generation Photon Sources for Grand Challenges in Science and Energy, DOE/BES, 2008.
- [76] J. Corlett, et al., Design studies for a VUV soft X-ray free-electron laser array, Synchrotron Radiat. News 22 (5) (2009).
- [77] J. Bisognano, et al., The Wisconsin free electron laser initiative, in: Proceedings of the 2009 Particle Accelerator Conference, 2009.
- [78] NLS design outline is available at <www.newlightsource.org>.
- [79] K. Honkavaara, et al., Status of the free-electron laser user facility FLASH, in: Proceedings of the 2009 Superconducting RF Conference, 2009.
- [80] J.B. Murphy, S. Krinsky, R.L. Gluckstern, in: Proceedings of the 1995 Particle Accelerator Conference, vol. 2980, 1996.
- [81] Y.S. Derbenev, et al., DESY Report No. TESLA-FEL 95-05, 1995.
- [82] J.B. Murphy, S. Krinsky, R.L. Gluckstern, Part. Accel. 57 (1997) 9.
- [83] L.-H. Yu, Phys. Rev. A 44 (1991) 5178.
- [84] E.J. Takahashi, et al., Generation of strong optical field in soft X-ray region by using high-order harmonics, IEEE J. Sel. Top. Quant. Elec. 10 (2004) 1315.
- [85] I. Ben-Zvi, K.M. Yang, L.-H. Yu, Nucl. Instr. and Meth. A 318 (1992) 726.
- [86] M. Liepe, Cornell ERL high power delivery and low level control systems, in: Proceedings of ERL09, June 2009.
- [87] <http://xfel.eu/>.

- [88] B. Newnam, Multi-facet metal mirror design for extreme-ultraviolet and soft X-ray free-electron laser resonator, in: Proceedings of the 17th Annual Symposium on Optical Materials for High Power Lasers, Boulder, CO, 1985.
- [89] J.C. Gallardo, R.C. Fernow, R. Palmer, C. Pellegrini, *IEEE J. Quant. Electron.* V 24 (8) (1988) 1557.
- [90] J.C. Goldstein, B. McVey, C.J. Elliot, *Nucl. Instr. and Meth. A* 272 (1988) 177.
- [91] J.C. Goldstein, B.D. McVey, B.E. Newnam, *Nucl. Instr. and Meth. A* 296 (1990) 288.
- [92] F. Curbis, et al., in: Proceedings of the FEL 2005, 2005, p. 473.
- [93] S.V. Benson, G.R. Neil, M.D. Shinn, in: Proceedings of SPIE 3931B, 2000, p. 243.
- [94] J.C. Goldstein, D.C. Nguyen, R.L. Sheffield, *Nucl. Instr. and Meth. A* 393 (1997) 137.
- [95] D.C. Nguyen, et al., *Nucl. Instr. and Meth. A* 429 (1999) 125.
- [96] B. Faatz, et al., *Nucl. Instr. and Meth. A* 429 (1999) 424.
- [97] B. Faatz, et al., *Nucl. Instr. and Meth. A* 483 (1999) 412.
- [98] N.R. Thompson, D.J. Dunning, in: Proceedings of FEL2007, Paper MOCAU01, available at <www.jacow.org>, 2007, pp. 182–187.
- [99] D.J. Dunning, B.W.J. McNeil, N.R. Thompson, *Nucl. Instr. and Meth. A* 593 (2008) 116–119.
- [100] A.E. Siegman, Lasers, University Science Books, Mill Valley, CA, 1986, ISBN 0-935702-11-5, p. 600.
- [101] D. Attwood, in: Soft X-rays and Extreme Ultraviolet Radiation, Cambridge University Press, Cambridge, UK, 1999.
- [102] H. Kumagai, K. Toyoda, K. Kobayashi, M. Obara, Y. Iimura, *Appl. Phys. Lett.* 70 (18) (1997) 2338–2340.
- [103] F.H. Fabreguette, R.A. Wind, Z.A. Sechrist, S.M. George, presented at AVS Topical Conference on Atomic Layer Deposition (ALD 2004), August 2004.
- [104] M.K. Tripp, F.H. Fabreguette, C.F. Herrmann, S.M. George, V.M. Bright, *Proc. SPIE* 5720 (2005) 241–251.
- [105] S. Bajt, et al., *Proc. SPIE* 4506 (2001) 2.
- [106] A. Hardouin, et al., in: Proceedings of the 8th International Conference of Physics of X-ray Multilayer Structures, Sapporo, March 2006.
- [107] <http://www.iof.fraunhofer.de/departments/optical-coatings>.
- [108] Z. Huang, R.D. Ruth, *Phys. Rev. Lett.* 96 (2006) 144801.
- [109] R. Thiele, et al., *Phys. Rev. E* 78 (2008) 026411.
- [110] S. Sahoo, et al., *Phys. Rev. E* 77 (2008) 046402.
- [111] A. Matsuda, et al., *STAM* 5 (2004) 511.
- [112] K.-J. Kim, Y. Shvyd'ko, S. Reiche, *Phys. Rev. Lett.* 100 (2008) 244802.
- [113] E. Burkel, *Rep. Prog. Phys.* 63 (2000) 171.
- [114] E.L. Saldin, E.A. Schneidmiller, Y.V. Shvyd'ko, M.V. Yurkov, *Nucl. Instr. and Meth. A* 475 (2001) 357.
- [115] B. Adams, G. Materlik, in: Proceedings of Free Electron Lasers 1996, II-24, North-Holland, Amsterdam, 1997.
- [116] R.M.J. Cotterill, *Appl. Phys. Lett.* 12 (1968) 403.
- [117] K.-J. Kim, Y. Shvyd'ko, *Phys. Rev. ST—Accel. Beams* 12 (2009) 030703.
- [118] M. Krisch, F. Sette, *Light Scattering in Solids IX*, Topics in Applied Physics, vol. 108, Springer, Berlin, 2007, pp. 317–370.
- [119] E. Gerdau, H. de Waard (Eds.), *Nuclear Resonant Scattering of Synchrotron Radiation, Hyperfine Interactions*, vols. 123–125, 1999 (special issue).
- [120] R. Röhlberg, *Nuclear Condensed Matter Physics with Synchrotron Radiation*, Springer, Berlin, 2004.
- [121] See papers in *Nucl. Instr. and Meth. A* 547 (2005) 1.
- [122] F. Parmigiani (private communication).
- [123] I. McNulty (private communication).
- [124] G. Dattoli, A. Marino, A. Renieri, F. Romanelli, *IEEE J. Quant. Electron.* QE 17 (1981) 1371.
- [125] P. Elleaume, *IEEE J. Quant. Electron.* QE 21 (1985) 1012.
- [126] R.R. Lindberg, K.-J. Kim, *Phys. Rev. ST Accel. Beams* 12 (2009) 070702.
- [127] R.R. Lindberg, W. Fawley, Y. Shvyd'ko, K.-J. Kim, in preparation.
- [128] K.-J. Kim, R.R. Lindberg, M. Borland, in preparation.
- [129] X.J. Wang, X.Y. Chang, *Nucl. Instr. and Meth. A* 507 (2003) 310.
- [130] J. Rosenzweig, et al., *Nucl. Instr. and Meth. A* 593 (2008) 39.
- [131] K. Togawa, et al., *Phys. Rev. ST Accel. Beams* 10 (2007) 020703.
- [132] T. Shintake, et al., *Phys. Rev. ST Accel. Beams* 12 (2009) 070701.
- [133] Y. Shvyd'ko, *X-ray Optics—High-energy Resolution Applications*, Springer Verlag, Berlin, 2004.
- [134] H. Mimura, et al., *Rev. Sci. Instr.* 79 (2008) 083104.
- [135] S. Stoupin, Y. Shvyd'ko, Anomalous thermal expansion of diamond crystals and low temperature, *Phys. Rev. Lett.* 104 (2010) 085901.
- [136] K.-J. Kim, ANL Technical Report No. AAI-PUB-2008-004, 2008.
- [137] S. Stoupin, F. Lenkszus, R. Laird, K. Goetze, K.-J. Kim, Y. Shvyd'ko, Nanoradian angular stabilization of X-ray optical components, *Rev. Sci. Instr.*, submitted for publication.
- [138] M. Borland, Accelerated Operations and Technical Report, AOP-TN-2008-057, 2008.
- [139] Y. Ding, et al., *Phys. Rev. Lett.* 102 (2009) 254801.
- [140] S. Reiche, P. Musumeci, C. Pellegrini, J.B. Rosenzweig, *Nucl. Instr. and Meth. A* 593 (2008) 45.
- [141] P. Emma, K. Bane, M. Cornacchia, Z. Huang, H. Schlarb, G. Stupakov, D. Walz, *Phys. Rev. Lett.* 92 (2004) 074801.
- [142] F. Krausz, M. Ivanov, *Rev. Mod. Phys.* 81 (2009) 163.
- [143] A.A. Zholents, W.M. Fawley, *Phys. Rev. Lett.* 92 (2004) 224801.
- [144] E.L. Saldin, E.A. Schneidmiller, M.V. Yurkov, *Opt. Commun.* 239 (2004) 161.
- [145] A.A. Zholents, G. Penn, *Phys. Rev. ST Accel. Beams* 8 (2005) 050704.
- [146] E.L. Saldin, E.A. Schneidmiller, M.V. Yurkov, *Phys. Rev. ST Accel. Beams* 9 (2006) 050702.
- [147] A.A. Zholents, M.S. Zolotarev, *New J. Phys.* 10 (2008) 025005.
- [148] Y. Ding, Z. Huang, D. Ratner, P. Bucksbaum, H. Merdji, *Phys. Rev. ST Accel. Beams* 12 (2009) 060703.
- [149] D. Xiang, Z. Huang, G. Stupakov, *Phys. Rev. ST Accel. Beams* 12 (2009) 060701.
- [150] G. Stupakov, *Phys. Rev. Lett.* 102 (2009) 074801.
- [151] A.A. Zholents, *Phys. Rev. ST Accel. Beams* 8 (2005) 040701.
- [152] A.A. Zholents, G. Penn, *Nucl. Instr. and Meth. A* 612 (2010) 254.
- [153] D. Dowell, et al., BES Workshop Cathode Paper, *Nucl. Instr. and Meth. Phys. A* (2010), to appear, doi:10.1016/j.nima.2010.03.104.
- [154] B.E. Carlsten, *Nucl. Instr. and Meth. A* 285 (1989) 313.
- [155] L. Serafini, J.B. Rosenzweig, *Phys. Rev. E* 55 (1997) 7565.
- [156] K.-J. Kim, Y. Shvyd'ko, S. Reiche, *Phys. Rev. Lett.* 100 (2008) 244802.
- [157] H. Tanaka, et al., Low-emittance injector at SCSS, in: Proceedings of the FEL 2006, 2006, pp. 769–776.
- [158] X.J. Wang, et al., *Nucl. Instr. and Meth. A* 507 (2003) 310.
- [159] P.N. Ostromov, K.-J. Kim, P. Piot, Development of Ultra-Low Emittance Injector for Future X-ray FEL Oscillator, in: Proceedings of LINAC 08 Conference, Victoria, BC, Canada, p. 676.
- [160] P.N. Ostromov, D. Capatina, K.-J. Kim, S.A. Knodarshev, B. Mustapha, A. Nassiri, Optimized design of an Ultra-Low Emittance Injector for Future X-Ray FEL Oscillator, in: Proceedings of PAC'09 Conference, Vancouver, BC, Canada.
- [161] I. Bazarov, C. Sinclair, *Phys. Rev. ST Accel. Beams* 8 (2005) 034202.
- [162] See for example: I.V. Bazarov, B.M. Dunham, C.K. Sinclair, *Phys. Rev. Lett.* 102 (2009) 104801.
- [163] C. Hernandez-Garcia, et al., in: Proceedings of 26th International Free Electron Laser Conference, Trieste, Italy, 2004, pp. 558–560.
- [164] B.M. Dunham, et al., in: Proceedings of 2007 IEEE Particle Accelerator Conference, Albuquerque, NM, USA, 2007, pp. 1224–1226.
- [165] G. Neil, P. Evtushenko, private communications.
- [166] M. Breidenbach, et al., *Nucl. Instr. and Meth. A* 350 (1994) 1.
- [167] D. Janssen, et al., FEL03, Tsukuba, Japan, September 8–12, 2003.
- [168] J. Rathke, et al., PAC'05, Knoxville, USA, May 16–20, 2005-09-15.
- [169] K. Zaho, et al., *Nucl. Instr. and Meth. A* 475 (2001) 564.
- [170] R. Legg, in: Proceedings of LINAC08, Victoria, BC, Canada, 2008, p. 658.
- [171] J. Hao, et al., in: Proceedings of the SRF2009 Conference, Berlin, Germany, 2009, p. 205.
- [172] J. Teichert, et al., in: Proceedings of the 2008 FEL Conference, Gyeongju, Korea, 2008.
- [173] D. Janssen, et al., EPAC04, Lucerne, Switzerland, July 5–9, 2004.
- [174] See for example: D.T. Palmer, et al., in: Proceedings of the 1997 Particle Accelerator Conference, Vancouver, BC, Canada, May 1997, p. 2846. Or: L. Staykov, et al., Proceedings of the 2007 FEL Conference, Novosibirsk, Russia, 2007, p. 138.
- [175] C. Limborg-Deprey, D. Dowell, J. Schmerge, Z. Li, L. Xiao, RF Design of the LCLS Gun, LCLSTN-05-3, February 2005.
- [176] P. Emma, et al., Lasing and saturation of the LCLS at 1.5 Angstrom, in: Proceedings of the 2009 FEL Conference, Liverpool, UK, 2009.
- [177] J.W. Staples, S.P. Virostek, S.M. Lidia, in: Proceedings of the 2004 European Particle Accelerator Conference, Lucerne, Switzerland, July 2004, pp. 473–475.
- [178] R.A. Rimmer, in: Proceedings of the 2005 Particle Accelerator Conference, Knoxville, TN, USA, May 2005, pp. 3049–3051.
- [179] J. Staples, F. Sannibale, S. Virostek, VHF-band photoinjector, CBP Tech Note 366, 2006.
- [180] K. Baptiste, et al., *Nucl. Instr. and Meth. A* 599 (2009) 9.
- [181] X. Chang, et al., PAC07, 2007, p. 2547.
- [182] D.C. Nguyen, et al., *Nucl. Instr. and Meth. A* 528 (2004) 71.
- [183] K. Baptiste, et al., Status and plans for the LBNL normal-conducting CW VHF photoinjector, in: Proceedings of FEL09, Liverpool, UK, 2009.
- [184] L. Serafini, M. Ferrario, *AIP Conf. Proc.* 581 (2001) 87.
- [185] K. Shirasawa, et al., High gradient tests of C-band accelerating system for Japanese XFEL project, in: Proceedings of the 2007 Particle Accelerator Conference, 2007, p. 2095.
- [186] Proceedings of the X-Band Accelerating Structure Design and Test Program Workshop, CERN, Geneva, Switzerland, 18–19 June 2007.
- [187] V.A. Dolgashev, et al., High power test of normal conducting single cell structures, in: Proceedings of the 2007 Particle Accelerator Conference, 2007, p. 2430.
- [188] H.A. Grunder, CEBAF commissioning and future plans, in: Proceedings of the 1995 Particle Accelerator Conference, 1995, p. 1.
- [189] P.A. McIntosh, et al., Superconducting options for the UK's new light source project, in: Proceedings of the LINAC08, Victoria, BC, Canada, 2008.
- [190] Abo-Bakr, et al., Start-to-end injector and linac tolerance studies for the Bessy FEL, in: Proceedings of the 2004 FEL Conference, 2004, pp. 104–107.
- [191] J. Bardeen, L.N. Cooper, J.R. Schrieffer, Theory of Superconductivity, Illinois U., Urbana, July 1957, *Phys. Rev.* 108 (1957) 1175–1204.
- [192] P. Kneisel, High gradient superconducting niobium cavities: a review of the present status, JLAB-ACT-98-07, in: Proceedings of the 1998 Applied Superconductivity Conference (ASC-98), Palm Desert, CA, USA, 13/09/1998.
- [193] R. Geng, Progress on improving SC cavity performance for ILC, JLAB-ACC-09-1044, in: Proceedings of the PAC09, Vancouver, BC, 2009.
- [194] R. Geng, Overview of high gradient SRF R&D for ILC cavities at Jefferson Lab, JLAB-ACC-09-1086, in: Proceedings of the 14th International Conference on RF Superconductivity (SRF2009), Berlin, Germany, September 20–25, 2009.

- [195] T.I. Smith, et al., Superconducting RF and RF control for energy-recovery linacs: summary of working group 3 at ERL 2007, in: Proceedings of the ERL07, Daresbury, UK, 2007.
- [196] The JLAMP VUV/soft X-ray user facility, DOE Proposal, December 2009.
- [197] B. Dunham, et al., First tests of the Cornell University ERL injector, in: Proceedings of the LINAC08, Victoria, BC, Canada, 2008.
- [198] M. Liepe, SRF system optimization process, in: Proceedings of the ERL09, Cornell University, Ithaca, NY, USA, June 8–12, 2009.
- [199] J. Marangos, et al., New light source (NLS) project: science case and outline facility design, 7/09 <<http://www.newlightsources.org>>.
- [200] G. Ciovati, High field Q-slope and the baking effect, in: Proceedings of the 14th International Conference on RF Superconductivity (SRF2009), Berlin, Germany, September 20–25, 2009.
- [201] A.A. Zholents, W. Fawley, P. Emma, Z. Huang, G. Stupakov, S. Reiche, in: Proceedings of FEL2004 Conference, 2004, p. 582.
- [202] M. Borland, et al., Nucl. Instr. and Meth. A 483 (2002) 268.
- [203] S. Heifets, G. Stupakov, Phys. Rev. ST Accel. Beams 5 (2002) 064401.
- [204] Z. Huang, K.-J. Kim, Phys. Rev. ST Accel. Beams 5 (2002) 074401.
- [205] E.L. Saldin, E.A. Schneidmiller, M.V. Yurkov, Nucl. Instr. and Meth. A 483 (2002) 516.
- [206] E.L. Saldin, E.A. Schneidmiller, M.V. Yurkov, Nucl. Instr. and Meth. A 528 (2004) 355.
- [207] T. Shaftan, Z. Huang, BNL Report No. BNL-71491, 2003.
- [208] J. Qiang, R.D. Ryne, M. Venturini, A.A. Zholents, I.V. Pogorelov, Phys. Rev. ST Accel. Beams 12 (2009) 100702.
- [209] M. Venturini, Phys. Rev. ST Accel. Beams 10 (2007) 104401.
- [210] M. Venturini, A.A. Zholents, Nucl. Instr. and Meth. A 593 (2008) 53.
- [211] Z. Huang, et al., Phys. Rev. ST Accel. Beams 7 (2004) 074401.
- [212] A. Huang, et al., Phys. Rev. ST Accel. Beams 13 (2010) 020703.
- [213] A.W. Chao, Physics of Collective Beam Instabilities in High Energy Accelerators, John Wiley & Sons, New York, NY, 1993.
- [214] K. Bane, G. Stupakov, SLAC-PUB-10707, 2004.
- [215] K. Bane, C. Ng, A. Chao, Report SLAC-PUB-7514, SLAC, 1997.
- [216] G. Stupakov, Phys. Rev. ST Accel. Beams 1 (1998) 064401.
- [217] G. Stupakov, et al., Phys. Rev. ST Accel. Beams 2 (1999) 060701.
- [218] K. Bane, et al., Phys. Rev. ST Accel. Beams 12 (2009) 030704.
- [219] A. Novokhatski, Paper WEPC22, FEL09, 2009.
- [220] T.-M. Tran, J.S. Wurtele, Comput. Phys. Commun. 54 (1989) 263.
- [221] W.M. Fawley, Report LBNL-49625, 2002.
- [222] S. Reiche, Nucl. Instr. and Meth. A 429 (1999) 243.
- [223] T. Tanaka, in: Proceedings of FEL2004, Trieste, Italy, August 2004.
- [224] C. Penman, B.W.J. McNeil, Opt. Commun. 90 (1992) 82.
- [225] W.M. Fawley, Phys. Rev. ST Accel. Beams 5 (2002) 070701.
- [226] I. Zagorornov, M. Dohlus, Proceedings of FEL2009, Liverpool, UK, August 2009, p. 71.
- [227] R.J. Dejus, O.A. Shevchenko, N.V. Vinokurov, Nucl. Instr. and Meth. A 429 (1999) 225.
- [228] E.L. Saldin, E.A. Schneidmiller, M.V. Yurkov, Nucl. Instr. and Meth. A 429 (1999) 225.
- [229] H.P. Freund, S.G. Biedron, S.V. Milton, IEEE J. Quant. Electron. QE-36 (2000) 275.
- [230] G. Bassi, et al., Nucl. Instr. and Meth. A 557 (2006) 189.
- [231] H.H. Braun, et al., Phys. Rev. Lett. 84 (2000) 658;
- [232] H.H. Braun, et al., Phys. Rev. ST Accel. Beams 3 (2000) 124402.
- [232] M. Borland, J. Lewellen, in: Proceedings of the 2001 Particle Accelerator Conference, Chicago, IL, IEEE, Piscataway, NJ, 2001, p. 2823.
- [233] P. Emma, et al., in: Proceedings of the 2003 Particle Accelerator Conference, Portland, OR, IEEE, New York, 2003, p. 3129.
- [234] F. Zhou, et al., Phys. Rev. ST Accel. Beams 9 (2006) 114201.
- [235] B. Beutner, et al., in: FEL2007 Proceedings, 2007, p. 18.
- [236] M. Borland, 'Elegant: A Flexible SDDS-Compliant Code for Accelerator Simulation', Advanced Photon Source LS-287, September 2000.
- [237] S.G. Anderson, et al., Phys. Rev. ST Accel. Beams 8 (2005) 014401.
- [238] M. Dohlus, T. Limberg, Bunch compression stability dependence on RF parameters, in: Proceedings of the 2005 FEL Conference, Stanford, CA, p. 250.
- [239] P. Emma, SLAC-TN-05-004, November 14, 2001.
- [240] <<http://www.fidtechnology.com/productsindex.htm>>.



Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima

X-ray sources by energy recovered linacs and their needed R&D

S. Benson^a, M. Borland^b, D.R. Douglas^a, D. Dowell^c, C. Hernandez-Garcia^a, D. Kayran^d, G.A. Krafft^{a,e,*}, R. Legg^f, E. Moog^b, T. Obina^g, R. Rimmer^a, V. Yakimenko^d

^a Thomas Jefferson National Accelerator Facility, 12000 Jefferson Avenue, Newport News, VA 23606, USA

^b Argonne National Laboratory, 9700 S. Cass Avenue, Argonne, IL 60439, USA

^c SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025, USA

^d Brookhaven National Laboratory, Upton, NY 11973, USA

^e Center for Accelerator Science, Old Dominion University, 4600 Elkhorn Avenue, Norfolk, VA 23529, USA

^f Synchrotron Radiation Center, University of Wisconsin-Madison, 3731 Schneider Drive, Stoughton, WI 53589, USA

^g High Energy Accelerator Research Organization: KEK, Tsukuba, Ibaraki 305-0801, Japan

ARTICLE INFO

Keywords:

Energy recovered linac
ERL
Advanced X-ray source
ERL X-ray source
ERL development
Recirculated linac X-ray source

ABSTRACT

In this paper we review the current state of research on energy recovered linacs as drivers for future X-ray sources. For many types of user experiments, such sources may have substantial advantages compared to the workhorse sources of the present: high energy storage rings. Energy recovered linacs need to be improved beyond present experience in both energy and average current to support this application. To build an energy recovered linac based X-ray user facility presents many interesting challenges. We present summaries on the Research and Development (R&D) topics needed for full development of such a source, including the discussion at the Future Light Sources Workshop held in Gaithersburg, Maryland on September 15–17, 2009. A first iteration of an R&D plan is presented that is founded on the notion of building a set of succeeding larger test accelerators exploring cathode physics, high average current injector physics, and beam recirculation and beam energy recovery at high average current. Our basic conclusion is that a reviewable design of such a source can be developed after an R&D period of five to ten years.

© 2011 Published by Elsevier B.V.

1. Introduction

Synchrotron radiation sources may be divided broadly into two classes. Sources of “spontaneous” radiation, including most storage ring light sources, involve independent emission of radiation by each electron. The other class, which includes free electron lasers, involves coherent emission of radiation by an appropriately bunched beam. Energy recovery linacs [1] (ERLs) can drive sources of both types. In the X-ray regime of photon frequencies the primary emphasis at present is on spontaneous radiation sources that promise to provide spectral brightness several orders of magnitude greater than that available from storage rings.

To understand the potential advantage of ERLs, it is necessary to understand how the electron beam properties influence the X-ray brightness [2]. The brightness may be expressed as

$$B \sim \frac{I_e}{E_x E_y \sqrt{4\sigma_\delta^2 + \left(\frac{0.4}{hN_e}\right)^2}}, \quad (1)$$

where I_e is the electron beam average or peak current, N_e is the number of undulator periods, h is the undulator harmonic, σ_δ is the rms fractional momentum spread of the electron beam. In addition, E_x and E_y are the photon beam emittances in the horizontal and vertical plane, found by convolving the electron distribution with the single-electron photon distribution, yielding

$$E_q = \sqrt{\left(\varepsilon_q \beta_q + \frac{\lambda L}{8\pi^2}\right) \left(\frac{\varepsilon_q}{\beta_q} + \frac{\lambda}{2L}\right)} \quad (2)$$

where q is x or y , ε_q is the electron beam emittance, β_q is the electron beam beta function, λ is the radiation wavelength, and L is the undulator length. From this equation we observe that the single-electron photon distribution has an emittance of $\lambda/4\pi$ and a beta function of $L/2\pi$. If the electron beam is given the same beta function, then

$$E_q = \varepsilon_q + \frac{\lambda}{4\pi} \quad (3)$$

which is the smallest possible value for the photon beam emittance and yields the highest brightness for a given electron beam emittance. For a typical storage ring, L is 2–5 m, so that the optimum beta function is 0.3–0.8 m. Achieving such small values in a ring is difficult as it forces large beta functions elsewhere in the system, creating difficulties for dynamic aperture. This argument

* Corresponding author at: Thomas Jefferson National Accelerator Facility, 12050 Jefferson Avenue, Newport News, VA 23606, USA.
E-mail address: krafft@jlab.org (G.A. Krafft).

Table 1
Present advanced photon source beam parameters compared to ERL X-ray source parameters in high coherence and high flux modes [4].

Quantity	APS	ERL high coherence	ERL high flux
Beam energy (GeV)	7	7	7
Average current (mA)	100	25	100
Repetition rate (MHz)	6.5–352	1300	1300
Bunch charge (nC)	< 59	0.019	0.077
Horizontal emittance (geometric pm), [normalized (μm)]	3100 [42]	6 [0.08]	20 [0.27]
Vertical emittance (geometric pm), [normalized (μm)]	25–50 [0.35–0.70]	6 [0.08]	20 [0.27]
rms bunch length (ps)	> 20	2	1.7
rms energy spread (%)	0.1	0.015	0.014
Photon brightness ($10^{20}\text{p}/(\text{s mm}^2 \text{ mrad}^2 0.1\% \text{BW})$)	0.3	200	60

Photon brightness at 10 keV reported.

highlights the first possible advantage for ERLs as light sources: freedom to use the optimum beta function.

Assuming that $\beta_q = L/(2\pi)$, if we additionally have $\varepsilon_q < \lambda/(4\pi)$, the source is said to be diffraction limited. In this case, we obtain the highest possible brightness and full transverse coherence. One important goal of any spontaneous X-ray source is to provide diffraction limited radiation at, say, 1 Å. This requires $\varepsilon_q < 7$ pm, a value that is achieved in the vertical plane in many modern storage rings. However, present storage ring designs cannot approach this performance in both planes (see, however, Ref. [3]). These considerations highlight the second potential advantage of ERLs as light sources: the ability to supply ultra-low electron beam emittances and hence reach the diffraction limit in both transverse dimensions.

Returning to Eq. (1), we note that if $\sigma_\delta \rightarrow 0$, then $B \propto N_u$. If, however, σ_δ is large, the benefit of a longer device is diminished. In storage rings, typically $\sigma_\delta = 0.1\%$, whereas in a linac, σ_δ can be considerably lower, being determined by the bunch length and the rf frequency. This highlights the third potential advantage of ERLs as light sources, namely, the ability to better capitalize on long undulators to achieve higher brightness. Fourthly, the electron beam longitudinal distribution is easily manipulated in the ERL because of the absence of significant radiation damping in the time it takes the electrons to traverse the accelerator. This circumstance allows the possibilities of, by longitudinal compression of the electron bunches, high repetition rate short X-ray pulse fluxes beyond those possible in storage rings.

Finally, note that Eq. (1) contains the electron beam current. In storage rings, the average beam current can easily exceed 100 mA. In the past, high average current has been a decisive advantage in favor of storage ring light sources. The energy recovery concept directly addresses this issue by reducing the power requirements for running continuous beam from a linac. ERL beam conditions supporting two possible X-ray production modes are listed in Table 1, scaled to the same beam energy, and compared to those in the forefront storage ring source [4].

2. Accelerator physics and challenges of ERLs

While essential, the concept of energy recovery does not address all of the challenges facing linac-based spontaneous light sources, which are discussed in detail in this section of this paper. Much information on the challenges and their present status may be found in the proceedings of the biennial conferences on ERLs that have been held at Jefferson Laboratory [5], Daresbury Laboratory [6], and Cornell University [7], and in recent review papers at accelerator conferences [8–10]. Here, we briefly highlight some of the most difficult challenges, which begin at the cathode where the electrons originate, and continue throughout the system until the electron beam is decelerated and finally dumped.

In the injector, challenges include the production of ultra-low emittance beams with high average current that is sustained for long periods (e.g., days), as expected at a user facility. Significant improvements in cathode technology will be essential to success on this topic. The choice of gun type is not settled, particularly given some of the difficulties that have been encountered for very high-voltage DC guns.

Given the high average current and high beam energy, controlling beam loss is a critical issue, requiring dramatically better understanding of the mechanisms of halo generation and loss. Effective collimation systems that do not adversely impact beam brightness must be developed. As for any high-brightness beam, collective effects have significant potential for impacting performance. Among these are space charge, coherent synchrotron radiation (CSR), short- and long-range wakefields, ion trapping, and intrabeam scattering. An important issue related to collective effects is management of the energy spread after deceleration, which impacts beam loss and the ability to recover to a low beam energy.

Design of the overall system presents challenges of a different type, related to optimization of cost and performance. Issues such as the choice of maximum beam energy, choice of undulator length and the number of undulators, choice of the number of passes to use through the linac, and best design for preserving the emittance must be considered. Related issues are determining tolerances and development of adequate correction and feedback systems.

Experience with free-electron lasers has demonstrated convincingly the utility of accurate, start-to-end modeling. Most of these challenges will require a significant improvement in computational capabilities and sophistication. Not only are improved codes required, but benchmarking is essential to give the confidence required to build a full-scale facility.

2.1. Significant research and development topics needed for an advanced ERL-based X-ray source

2.1.1. Photocathode studies

The three basic types of cathode that have been used to generate electron beams for accelerator-based applications are semiconductor photocathodes, metallic photocathodes, and thermionic cathodes. At Jefferson Laboratory, Daresbury Laboratory, and Cornell University, a Cs:GaAs photocathode is used in a DC gun [11], while the Budker Institute FEL/ERL [12] utilizes a thermionic cathode. To date, no other type of cathode has delivered continuous beam for an ERL-based machine. The JLab FEL DC gun delivered over 900 h and 7000 Coulombs at 2–9 mA CW from a single GaAs wafer between 2004 and 2007. Cornell University has recently demonstrated 20 mA average DC current after the gun and 8 mA from their 6 MeV injector [13]. In 1991, the Boeing normal conducting RF gun demonstrated 32 mA with a K_2CsSb photocathode and still holds the record for the highest

average current from a photo cathode gun [14]. Cs₂Te cathodes have been in operation for 120 continuous days in a normal conducting RF gun at PITZ with minimal QE degradation. Thermionic cathodes have very good lifetime but gating the current in ps-long pulses is very difficult. The JAERI FEL used a thermionic cathode in a quasicontinuous, but low duty factor mode [15]. As a low repetition rate injector the CeB₆ gun works well, has successfully delivered stable 500-keV beams to the SCSS test accelerator for three years [16], and is now operating for various EUV-FEL user experiments. Such results indicate that 100 mA sustained sources are within reach. However, in order to more completely understand and quantify cathode lifetime limits, dedicated studies to understand the origin of the limits, occurring presently at Jefferson Lab, BNL, and Cornell University, will need to continue.

A primary issue regarding photocathode gun performance particularly important for ERL applications is whether it is possible to achieve required beam quality specifications. Requirements in transverse emittance are more than an order of magnitude better than has been achieved in CW electron sources to date. This issue is discussed in a companion paper to this one [17], and detailed simulations suggest that the emittances required for hard X-ray ERLs can be achieved [18,19]. Recent measurements of beam quality at Cornell University [20], in a DC photocathode gun designed to drive an ERL [21], are promising. Incomplete answers on both of these issues support the notion of measurements at photocathode test stands to perform dedicated R&D studies on long lifetime high beam quality photocathode arrangements. This idea is incorporated in the ERL R&D plan discussed below.

2.1.2. Drive laser

The photocathode drive laser for an ERL is integrally tied to the photocathode design. The challenges of the two complement each other. For example, a UV capable photocathode is more robust and has a longer lifetime but a UV photocathode drive laser is very challenging. If the cathode quantum efficiency is high, the drive laser gets easier but the cathode is now very hard to produce. The product of the laser power and quantum efficiency is a constant for a given laser wavelength and beam current. With very high quantum efficiency the laser power absorbed by the cathode can be reasonably low (though still requiring active cooling). For quantum efficiencies much less than about 1% for 100 mA beams, the power absorbed becomes a major engineering challenge [22]. The total power that can be absorbed by the photocathode is no more than 1 kW/cm² and with typical spot sizes this means that the drive laser cannot deliver more than 100 W to the cathode. The laser itself must put out more than this to provide for transport losses. This level of power is available today with custom laser systems but not in conjunction with the rest of the specifications detailed below. With the technology choices available today the mainstays of photocathode drive lasers are neodymium or ytterbium doped crystals. These lase at approximately 1060 nm in the near infrared and can be efficiently doubled to the green at 530 nm. With more difficulty they can be tripled or quadrupled to produce light near 350 or 265 nm. At high power levels the harmonic generation crystals suffer from rapid degradation. Laser systems operating in the visible are therefore greatly preferred. The drive laser must produce not just average power but a pulse train of highly stable shaped pulses with excellent contrast ratio [23]. The ideal pulse shape is that of a uniformly filled ellipsoid [24] but current technology can only produce uniformly filled cylinders with moderately sharp edges [25]. The power outside of these pulses must be as low as possible in both time and space. The pulses must also be nearly identical in shape, size, arrival time, and amplitude both on a fast time scale and over hours or days.

Diode pumped systems with feedback controls can provide reasonably stable laser systems but not yet at the levels required for an ERL machine (for example 0.1% amplitude at the harmonic wavelength and 100 fs timing stability). Finally the laser time structure must be variable so that a low-duty-factor, low average current, mode can be provided for accelerating tuning. There must then be some way to smoothly transition from the tune-up mode to full current operation in a reasonable time. Ideally this time structure control and the pulse shaping should be done at low power. The pulses would then be amplified and frequency converted into the final pulses. Due to the tendency of amplifiers to have much higher amplification at the beginning of a pulse, it is not possible to select out pulse trains at the input to the amplifier. This must be done at full power but can be accomplished using low loss Pockels cells and mechanical shutters. The longitudinal pulse shapes will probably change in the amplifiers as well so the ideal pulses are not necessarily at top hat at low power. Clearly there are many severe engineering design challenges that must be met to produce a reliable, stable, photocathode laser with the desired spatial and temporal specifications and the controls necessary to power a high current photocathode gun. Fortunately, the field of lasers is progressing rapidly so such a system is not out of the question for a high power ERL.

2.1.3. Emittance preservation

Preservation of beam quality in a large-scale ERL poses numerous interesting challenges [26,27]. An obvious issue is the management of the various transport lattice aberrations experienced by the beam during the acceleration, recirculation, and recovery cycle. By simulation calculation it has been shown [28,29] that one can preserve emittances at 1 μm (normalized) levels by using second order achromats [30,31] of sufficiently large bending radius. In these proofs-of-principle, the transport was also used to perform compression and decompression of the bunch length. Longitudinal effects during this process can have significant impact. There is a need to compensate for the effect of the RF curvature on the longitudinal phase space; such correction has been demonstrated using arc sextupoles and octupoles [32]. This correction must be completed with some care and caution, as it has been observed to drive intolerable phase space degradation in the aforementioned studies [33] with a devastating effect on the emittance. Though harmonic RF can be used for this compensation [34], it is a very costly option, inasmuch as the cost of the linac increases by more than 20% because the linearization process uses deceleration at the harmonic frequency and phase. In addition, it presents considerable opportunity for instability (via the impedances and wakes introduced by the harmonic RF) and operational difficulty (due to the aperture constraints imposed by the higher frequency cavities). The use of very high order achromats [35,36] or integrable lattices [37–40] has the potential to accommodate all functions in the arcs with preservation of the emittance: returning beam for deceleration, compensating the effect of RF curvature, compressing/decompressing the bunch length.

Incoherent synchrotron radiation (ISR) has been recognized as a performance limitation even in the earliest electron storage rings [41]. Recent existence proofs [42] demonstrate that the required performance can be achieved even in multiply recirculated systems, but considerable care must be taken and the resulting systems can be large and complex. In contrast, bunch self-interactions such as longitudinal space charge, microbunching instabilities, and coherent synchrotron radiation represent significant and ongoing challenges to contemporary linear-accelerator-driven light sources, particularly in high charge, short duration pulses. Success with the management of these effects in the Linac Coherent Light Source (LCLS) [43] provide promise that appropriate accelerator design and use of as-yet novel techniques

a month, with centroid shifts under this spread held for periods of days [70]. The beam path on recirculation in CEBAF changes at the 250 μm level (1°) due to daily temperature excursions [71]; these drifts are only periodically corrected. Measurements of micro-bunch phase stability have been performed at CEBAF and the Jefferson Lab FEL [72]. These measurements indicate tight control of the centroid fluctuation to the sub-100 fs level, on time scales up to 1 kHz. Fast position feedback systems, with update rates of several hundred Hz, are deployed at CEBAF. Position fluctuations at the 10 μm level, after correction, and angular fluctuations at the μ radian level are obtained with a system operating at 10–100 μA beam current [73]. As position signal levels with a typical ERL beam will be around three orders of magnitude larger, one anticipates position and angular stability, with feedback, at a level comparable to the storage ring source standard [74].

Perhaps the most difficult parameter to stabilize will be the average beam current because new beam is continuously injected into the ERL. Present experience with the Jefferson Lab FEL indicates current stability at the several per cent level. Advanced fast current locking systems have not been needed on existing recirculated linacs, and have not been developed. More information on this important topic should be forthcoming from ongoing work at Cornell.

2.1.7. High power beam diagnostics for ERLs

ERLs require exceptional beam quality in order to enable the science envisioned. To maintain the exceptional quality during user operational periods will require a suite of diagnostics capable of non-invasively monitoring the beam. Most of the bunch diagnostics available today were developed to operate in a single bunch or low power mode; see Refs. [75,76] for an overview. These diagnostics will need to be adapted to analyzing the electron beam in a non-destructive manner during production runs [77]. Non-invasive diagnostics which can monitor the beam quality during high power operation have been used on synchrotron storage rings for years [78], but depend on the fact that the lattice in a storage ring sets most of the beam properties, and that time averaging the result is desirable since that mirrors the user experience. In the ERL, by contrast, each pulse is produced separately and is only a function of the cathode and the intervening beam structures. Advanced synchrotron radiation monitors have been used in linear electron accelerators to monitor the transverse and longitudinal bunch dimensions [79,80], the energy spread [81], and get insight on the longitudinal phase space distribution [82]. Electro-optical diagnostics which utilize the electric field surrounding the beam to cause a Pockel's effect in adjacent crystals, have been implemented by various groups [83–85] with sub 100 fs resolution.

One of the resolution limits in these diagnostics comes from the $1/\gamma$ (γ being the Lorentz factor) opening angle of the bunch's electric field distribution. Such an energy-dependent resolution makes the electro-optical imaging technique not suitable to diagnose the bunch length at low energies (i.e. in the injector or beam dump region after deceleration). The resolution in these diagnostics is limited by the degree to which the bunch's electric field is perpendicular to the crystal, defined by the beam's energy, so this technique may not be applicable to the injector or beam dump region. The optical replica synthesizer method proposed and demonstrated at FLASH [86,87] shows great promise to allow single bunch analysis of the beam with minimal disruption. The technique can be extended to yield transverse slice emittance directly by placement of optical synchrotron radiation detectors with the proper phase advance along the beam path from the initial modulation. The resulting transverse profiles can be used to reconstruct the phase space of the beam. Many of these diagnostics depend on the ability to synchronize the diagnostic to the electron

beam with a resolution and stability much smaller than the bunch length. Synchronization systems with these specifications have been demonstrated [88] but will need further work to be integrated into the controls and diagnostic suites described. Further R&D will also be needed in order to take the research instruments cited above and re-engineer them for use on a user X-Ray ERL light source.

2.1.8. Undulators for ERLs

Undulators that are intended for use on an ERL can be similar to those presently in use at storage rings, or can take advantage of characteristics specific to an ERL to enhance the photon output. An ERL electron beam is smaller than a storage ring electron beam. The smaller beam allows the possibility of a smaller gap and higher field from a given undulator. That higher field translates into a tuning range that extends to lower photon energy, or can be used to make shorter period lengths possible without loss of tuning range. The shorter period undulator produces higher brilliance than a longer period (on the same harmonic), for all photon energies that it can reach. The ERL electron beam is also nearly round in cross-section, so magnetic components can be placed closer to the beam horizontally. This would allow the undulator to be turned so its field is horizontal and the resulting photons are vertically polarized. Users' mirrors would then deflect horizontally, keeping the entire beamline at the same height above the floor. The round beam will also increase the possibilities for polarizing undulators, because horizontally closer magnets and the resulting stronger horizontal field make stronger-field circular undulators possible. (See, for instance, the Delta [89] and APPLE-III [90] undulators.) Helical undulators with wire wound directly around the beam tube [91,92] also become possible. If one uses a design that puts permanent magnets closer to the beam and the source of stray radiation, consideration should be given to enhancing the radiation resistance of the permanent magnets, though today's high-coercivity magnet grades are much more radiation resistant than the magnet grades of the past.

The smaller electron beam energy spread of an ERL also offers possibilities. As the length of the undulator increases, the energy width of a harmonic peak decreases and the peak brilliance goes up, until the contribution from the energy spread of the electron beam becomes the dominant contributor [93]. Beyond that, there are no further gains in brilliance from increasing the undulator length. The smaller energy spread of an ERL electron beam extends the brilliance gains to longer undulators, so that the longest reasonable undulator might be $10 \times$ longer for an ERL than for a typical storage ring. There are challenges involved in building such a long undulator, however. R&D would be needed to devise and produce a means of keeping undulator segments in phase as the gap is changed. Also, the energy loss in the electron beam from such an undulator may be larger than the beam energy spread. This could lead to a long-undulator beamline affecting beam characteristics of a downstream beamline. RF cavities located downstream of the long undulator have been suggested as a correction [94], but this would require R&D. Also, there could be timing changes downstream as the gap of a long undulator is varied that would interfere with timing-sensitive applications. The smaller beam from an ERL also can present challenges to undulator tuning. With a very small spot size, users, especially microscopists, will become more sensitive to photon beam motion. Some variation in the kick at the end of an undulator is inevitable as the undulator gap is changed, and electron beam position monitoring is not sufficient to determine the photon beam position. Instead, the photon beam position itself must be monitored and used for feedback. Another possibility offered by an ERL is an increased coherent fraction in the photon beam as compared to a storage ring. This is most pronounced in the bending direction

because of the much smaller horizontal beam size in an ERL as compared to a storage ring. The various types of undulators presently being developed for storage rings would also find application on ERLs. In-vacuum, cryogenic permanent magnet, and superconducting undulators would offer enhanced capabilities, in similar ways to their possibilities on storage rings. The advantages of schemes to remove higher harmonic contamination from the spectrum, such as quasiperiodic [95,96] designs or the use of circular polarization would also apply to ERLs.

2.1.9. SRF guns for ERLs

Superconducting radio frequency (SRF) electron guns hold promise to produce beams of exceptional brightness as part of an ERL system. They do this by generating very large CW electric fields at the cathode, resulting in brighter beams at a given bunch charge [19]. Several implementations of SRF guns are in development world wide [97–101]. The devices currently under development [102] can be broadly divided by their operating frequency and shape. High frequency guns tend to be elliptical in design, while low frequency guns tend to be quarter wave resonator cavities, with the elliptical technology being more mature. Several mechanisms for mating a high temperature photocathode to the cavity have been employed. The difficulty is providing a thermal gap between the cathode and the cavity while making it appear to be a short circuit to the rf fields. This problem has been overcome using tuned structures between the cathode and the cavity. Many other areas of R&D remain, however. Some of those areas are cathode compatibility with the cavity, high power couplers and HOMs, particularly in cases where the device will be called on to produce relativistic beams at high average currents.

2.2. Significant computational requirements

Next, we briefly assess the status of existing computer codes relative to the physics challenges faced by ERLs. We attempt to identify where our modeling ability is weakest, and in particular to point out those phenomena which pose a significant risk to the success of an ERL light source but that are inadequately modeled presently.

Gun issues include modeling with space charge, cathode physics, and design-specific challenges such as insulator breakdown. Recent results [103] from the LCLS show that when sufficient care is taken in the modeling and engineering, results can be obtained that meet expectations. The development of improved cathode materials could have a significant payoff in terms of ERL brightness and feasibility as a user facility. One promising approach [104] is to use computation to model the electronic structure of candidate materials. This can be used to pre-select materials with the desired properties for experimental characterization. The state of code development for this effort appears to be adequate, but could benefit from streamlining and automation. Another significant challenge with DC guns is obtaining the required high voltage necessary to get ultra-low emittance [18]. Modeling tools could speed the development process and would have a high impact on the success of an ERL light source [105].

Loss of beam halo particles is a significant concern due to the high average current. There are many mechanisms for halo generation, most of which are poorly understood and modeled. In this latter group are phenomena such as field emission from the gun and linac, drive laser reflections and halo, cathode non-uniformity, and residual gas scattering. A few phenomena, such as Touschek scattering [106–109], external field nonlinearity, and space charge, are adequately covered, though application to ERLs has not necessarily been made. A related issue is design of effective collimation systems, which is adequately covered by combinations of existing tracking codes and Monte Carlo codes.

Computation of wakefields for picosecond bunches in long structures (e.g., a long insertion device chamber) is a challenge. The adequacy of existing higher-order electromagnetic codes, such as SLAC's T3P code [110], needs to be evaluated. Roughness wakes [111] and resistive wall effects [112] are also important and in need of more detailed computational study. Coherent synchrotron radiation effects are treated by several codes [113–115], using variants of a 1-dimensional model that has yielded good results for LCLS [103]. The parallel version of *elegant* [116] is capable of determining microbunching gain curves for a large ERL [117]. With the recent addition of shielding in BMAD [118], this subject is believed to be adequately modeled.

Start-to-end (S2E) modeling has proved very valuable in development and understanding of X-ray FELs [119,120]. The first application of S2E to an X-ray ERL [121] yielded some surprising results. A significant missing piece is fully integrated modeling of the laser system, including errors.

3. From R&D to facilities and evaluation of readiness

3.1. Key photon beam performance objectives

Up to present, discussions of ERLs as light sources have been specific to relatively hard photons, exceeding 1 keV, with photon fluxes exceeding only by small factors those present in existing storage ring sources, but with average and peak brilliances considerably above storage rings. Presently, serious proposals posit from 5 to 7 GeV in the ERL electron beam; to achieve high brilliance it is essential that exceptionally high average brilliance electrons be produced in the electron gun, and that the brilliance be increased during acceleration by the usual transverse betatron damping by acceleration. Thus ERLs are not so attractive at lower photon energies, both because competing storage rings have smaller damped normalized emittances there, and the advantage from betatron damping is not so great [122].

In order to fully utilize the higher electron beam brilliance, it will be necessary that the beam stability in transverse position at the insertion devices be a small fraction of the beam size there. Such small fluctuations are achieved at present day storage rings in the vertical direction, and lead to optimism that suitable feedback system designs can be developed starting with those deployed at rings [74].

3.2. Which topics could be addressed today?

Significant parts of the development process leading to the possibility of a high energy ERL-driven X-ray source can be, and are being addressed at present. For example, experiments can be performed in existing facilities at Cornell University, the Jefferson Lab energy recovering FEL, the Brookhaven National Lab ERL, and even at the CEBAF accelerator at Jefferson Lab, to elucidate ERL accelerator physics. Topics that could be investigated quantitatively with R&D support include: beam merging, emittance preservation in arc transport systems, CSR characterization and mitigation, quantitative ion trapping studies both through direct detection of accumulated ions and through detection of their effect on the electron beam dynamics, longitudinal space charge, beam stability in recirculated and energy recovered linacs, instability mitigation using transverse feedback systems, characterizing resistive wall effects, particularly in insertion devices, and benchmarking of codes with experiment.

A recurring theme in this workshop was the need to become more systematic in cathode studies. There are specific issues, particularly regarding cathode lifetimes in high average current

applications and space charge generated emittance limits [123], that can be addressed at the cathode surface physics laboratories being developed around the country [124,20]. Provisions for measurements of extracted beam quality at these laboratories is essential for future ERL development.

As shown in Fig. 1, Cornell University is completing an injector test facility whose overall goal is to demonstrate high average current electron beams of beam properties suitable for ERL light source applications [125,13]. The results from this test stand will be highly important in demonstrating suitable initial beam quality and control of the beam dynamics in the first parts of the accelerator. It will also allow, not as conveniently at the cathode physics laboratories, cathode studies with the photo-emitter integrated into a real operating environment.

3.3. Which topics could be addressed with a short term [few years], focused R&D program?

In the near term three high level goals could be achieved through a focused R&D program: (1) demonstrate production of high average brightness beams, (2) demonstrate suitable transport and phase space manipulation of high average brightness beams including quantifying injector halo, and (3) demonstrate requisite injector beam stability. Topic (1) includes demonstrating suitable beam current, and demonstrating a usable cathode lifetime for the ERL application in close-to-final injector configuration. Topic (2) includes direct checks of emittance growth during acceleration, demonstration of suitable emittance compensation schemes [25], and measurement and mitigation of any beam halos generated in the injector regions of the accelerator. In topic (3) one needs to demonstrate that the fluctuations in the beam bunch centroids in transverse position, angle, and longitudinal phase are small compared to the final injector bunch dimensions. Such measurements have been done in a cursory manner, and not necessarily with the precision required in a light source application, at existing facilities [70,73]. Presently, all of these issues are being addressed experimentally at the Cornell injector test facility, in a DC electron gun arrangement. The results of their studies will form a linchpin of all future ERL light source studies. In this paper it will be assumed that the Cornell injector will be supported to conclusion of these studies, and that the existing injector test stand can be made available for driving a small beam recirculation experiment afterwards.

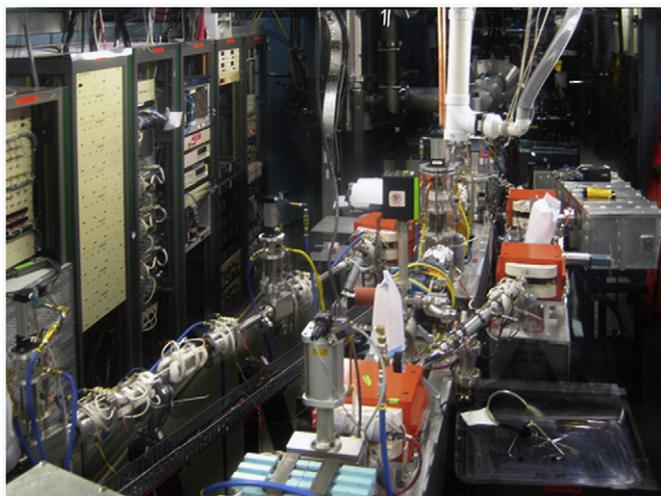


Fig. 1. Beam analysis lines of Cornell University's high average current ERL injector.

On the 3–6 year time horizon, two high level goals could similarly be accomplished: (1) develop an RF-gun-based alternative to the present DC guns used in ERL applications and (2) recirculate high average current beam through more than two accelerating and decelerating passes in an ERL configuration. It is well known that DC guns have technological limitations and difficulties that it would be nice to avoid [126]. There are reasons to think that RF guns, if they could be developed to operate in a CW mode, may produce beams with superior quality. As DC gun development is already proceeding at Cornell with National Science Foundation support, it may be wise that CW RF guns with high repetition rate be developed under Department of Energy stewardship.

In the future, when a final proposal for a light source is assembled, it will be necessary to know whether multiple accelerating and decelerating passes in the ERL configuration will be possible, because of possibilities of cost reduction. Because a highest energy recirculation must be done in any case, the question reduces to whether lower-energy recirculations contribute significantly to beam quality reduction, and whether average beam current limitations in the accelerator structures will be exceeded by multiple beam recirculations. These issues will be directly addressed by deploying an ERL test facility, which would take beam from the injector test facility and recirculate it multiple times around a few beam acceleration modules.

3.4. Which topics could be addressed in the longer term?

In the longer term several high level goals need to be achieved through the R&D program: (1) design of a full energy source using information obtained through prior R&D efforts, (2) establish the beam halo performance of the full design through simulations and experiments, (3) SRF cavity optimization regarding Q_0 , high order modes (HOMs), and frequency choices, (4) RF system optimization, and (5) cryogenics plant optimization. The final three topics are significantly related to the final operating costs for the full energy facility. Successful R&D on these topics could substantially reduce future facility operating costs.

The overall goal of any proposed development plan has to be to complete the studies needed to put forward a credible design for an ERL-based X-ray source. Up to now, existing source designs have utilized extant storage ring infrastructure, usually by having one of the turn-around arcs of the ERL consist largely of an existing storage ring. For comparison purposes, and to fully evaluate performance limitations imposed by such a selection, determining the performance of green field designs deploying the best available ideas could be highly useful in answering the question whether such choices are worthwhile. The answer to this question may evolve as designs are adjusted as more information becomes known.

The halo in the final machine, so important for machine protection, must be repeatedly addressed at each new level of device size. It is anticipated that studies on this particular topic will continue throughout the development process, and indeed, not receive rigorous resolution before the final source is built.

For next generation ERL-based light sources it is desirable to have the highest bunch rate possible to provide high average brightness. This requirement, combined with the need for high average accelerating gradient to keep the machines to manageable size, drives the requirement for CW superconducting RF (CW SRF) technology. Fortunately large-scale installations such as CEBAF [127] have shown that this technology is mature enough to make such facilities feasible. Existing ideas on light sources have been predicated on deploying superconducting RF cavities that were originally developed for High Energy Physics applications. Such cavities were not developed to be optimal for CW applications.

Although very much more efficient than normal-conducting RF, the operating costs of high gradient CW SRF are nevertheless

significant. Capital costs of installed SRF as well as large-scale cryogenic capability are cost drivers for major projects. A typical design requires more than 1 MW of RF power and 10 s of MW of power devoted to cryogenic cooling. The RF power may be supplied by a large number of relatively small sources, giving maximum flexibility of operation, or fewer higher power sources with more sophisticated control and distribution but a possible cost advantage. The best choice for any application will depend on many detailed factors.

ERLs require high circulating current, typically of the order of 100 mA per pass, placing a premium on HOM damping for beam stability. Multi-pass acceleration, as employed in CEBAF, offers the possibility of reduced SRF costs in exchange for the cost of recirculation arcs and spreaders and combiners provided the emittance can be preserved through these extra elements. The high average currents in ERLs require strict HOM damping [128,129], favoring designs that are shorter (5–7 cells) or have strong cell-to-cell coupling and cell-to-damper coupling. Such machines will have extremely tight stability requirements for energy and phase jitter [130], and require very low trip rates for user operations.

For CW SRF the most significant factor in operating cost is the efficiency of the accelerating system. For a given cavity geometry this translates to the highest possible quality factor Q_0 . The achievable Q_0 depends on many factors including material properties, processing history and surface morphology, as well as operating frequency and temperature. The well-known BCS theory [131] describes the ideal variation of superconducting cavity surface resistance with temperature and frequency and implies that the optimum operating point would be at the lowest practical temperature and frequency. However, the actual surface resistance of presently produced niobium cavities deviates from this ideal behavior and asymptotes to a residual surface resistance significantly higher than BCS theory predicts [131]. There is also a large variation in the practically achieved surface resistance in cavities that is as yet poorly understood. The typical average value is high enough above BCS that the theoretical gains from lower temperature and frequency are not realized in practice. For this and other reasons, most present or proposed CW machines remain at relatively high frequency (1.3–1.5 GHz). The origins of anomalous losses at typical operating gradients are the subject of ongoing investigation and any advancement in this area will pay large dividends in terms of usable gradient and overall facility costs. Recent excellent results with electro-polished cavities suggest that high operating Q_0 at the accelerating gradient of 20–25 MV/m may be reliably attained. If these results prove to be typical, they have the potential to shift the cost-optimal operating point to higher gradient. Field emission can seriously degrade the Q_0 and limit the usable gradient if cavities are imperfectly processed or mishandled after cleaning. Although great strides have been made in combating field emission, including recent tests to over 35 MV/m with no detectable X-rays, elaborate procedures will be necessary to ensure this can be achieved reliably for a large ensemble of cavities. Care must be taken in the design of magnetic shielding in the cryomodule [132] to achieve the full potential of the cavities. In practice the ideal shielding configuration is often compromised by the many penetrations necessary for tuners, couplers etc.

In practice the choice of frequency may be influenced by other factors besides SRF operating efficiency. Lower frequency cavities may support higher charge per bunch, but have longer RF buckets so depending on the detailed user beam requirements this may or may not be advantageous. Final choice of operating frequency, structure type and cryogenic temperature should be the result of a complex optimization and may be quite unique to a specific facility. A number of “light-source optimized” cavity designs are under development worldwide [133] and they vary considerably

in frequency, cell shape, number of cells, HOM damper type and power coupler configuration. An important milestone in cavity development will be beam test of these designs. The best verification of HOM damping, microphonics, power coupler etc. is with beam in a real machine or test facility. Several such facilities exist or are planned globally [134–136].

One of the fundamental aspects of successful energy recovery is the absence of a beam load on the RF system because the load is canceled by design [137]. Typically, the coupling in SRF cavities is chosen to match the beam current load, including provisions of extra RF drive and extra cavity bandwidth to allow for precise RF control of the fields in the cavity. When a beam load is not present, the possibility of significantly increasing the cavity fundamental mode Q_L arises, leading to less RF power required to drive the cavities [130]. At present, SRF cavities on the Jefferson Lab FEL/ERL have been successfully operated with a Q_L of 1.2×10^8 [138], including an energy recovered 5 mA beam passing through the cavities. The extra drive and bandwidth is required to ensure stability of the cavity gradient when the structure undergoes small tuning excursions due to external disturbances. This “microphonic” effect can be measured and to some degree mitigated by careful design, stiffening of the structure and good isolation or active feedback, but nevertheless places a practical limit on the maximum Q_L that can be operated stably. Further work on this subject, particularly in designing SRF cavities that are insensitive to coolant pressure fluctuations and noise pickup from the physical environment, and the design of quiet cryogenic systems which have smaller source terms for cavity resonance frequency fluctuations, could allow increases of the design Q_L still further. With success, the operating RF power requirements of the light source facility are proportionately lower. This research program is easily summarized by answering a simply stated question: what is highest practicable, fundamental mode Q_L obtainable in an SRF system with specific requirements for amplitude and phase control?

Research with the largest cost leverage would be that dedicated to improving the Q_0 at the operating gradient. Research is ongoing in this area, including studies of the high-field Q drop [139] and origins of residual resistance. Improvements in field emission free cavity processing would provide more confidence in higher operating gradients and could also allow cavity shapes to be contemplated that have lower operating losses but higher surface electric fields. Improvements in HOM damping, packing factor in the cryomodule, static losses and construction costs would all be worthwhile. Reduction of microphonics could allow for even higher operating efficiency in ERLs providing that nearly ideal energy recovery can be achieved. Success in this regard might ultimately allow solid-state amplifiers to be used with attendant simplifications and reliability.

Two recent examples of parametric studies have been performed for potential large light source facilities, the Cornell full-scale ERL source [140], and the UK NLS FEL project outline design report [141]. The optimal configurations for each are quite sensitive to detailed assumptions in the models, however, they share several common features. Both find a broad cost minimum (capital + 10 years operating costs), between about 15–25 MV/m and both chose gradients in the lower half of this range to be conservative. Both also end up choosing 1.3 GHz as the preferred operating frequency. Optimum operating temperature could be as low as 1.8 K depending on the assumed residual resistance. As might be expected the ERL study favors shorter (7-cell) but strongly HOM damped cavities, while the FEL study uses ILC-like 9-cell cavities. Some variables were not included in these studies, however, such as the variation of RF power costs, residual resistance and optimum operating temperature with frequency, and the relationship between end-use optical output and electron beam properties as influenced by all of the above. A comprehensive

evaluation including such variables should be undertaken for any new major facility.

As a final comment, further progress in increasing the energy efficiency of large cooling plants [142] is likely to occur over the next several decades. Dedicated basic research funding to optimize CW cryogenic plants may be appropriate, especially if it is determined that the final source is best run at elevated temperatures compared to today's standards.

4. Route to an ERL-driven X-ray source

Fig. 2 shows a schematic summary of a potential research path for an ERL X-ray source. We have left the duration of the individual steps undefined, the purpose of the diagram is to lay out parallelism and sequencing of the individual components one would expect to need for a full proposal to be developed. Arrows in the figure indicate completion dependencies between the various development projects most likely to lead to exceptional source performance. As indicated in Fig. 2, many parallel activities could start immediately: (1) measurement programs in existing facilities, (2) photocathode research devoted to high current photoinjector issues [17], (3) initiating the process of developing new SRF structures optimized for ERL light source applications, and (4) developing code focused on ERL beam dynamics.

As discussed in detail above, examples of potential measurements at existing facilities are quantitative 6-dimensional beam centroid stability measurements, ion trapping accumulation and mitigation measurements, beam halo generation, and even deploying and testing BBU mitigation hardware. Photocathode R&D work could be performed at the newly emerging cathode laboratories; for ERLs the issue of cathode lifetime needs serious attention in addition to the beam quality monitoring needed for other applications. Starting the process of deciding on the most optimal SRF accelerating cavity design for an ERL application could begin very soon.

After a short period, of order several years, we would expect, as is being presently accomplished at Cornell University as part of their National Science Foundation ERL X-ray source program, that a device called the injector test facility be completed. This device should be designed to fully support beam current and beam quality requirements needed for the eventual non-recovered portions of the final X-ray source. Demonstration of an injector capable to drive the

full current in the source with good beam quality would go a long way towards resolving quantitatively whether potential large brilliance gains will be possible by adopting ERL technology for an X-ray source application.

Before adoption of final source parameters, and utilizing the newly optimized SRF cavities, we propose that a natural next step would be an ERL test facility, where a small version of the final ERL accelerator would be built. It would consist of one or two cryomodules containing SRF cavities of the final design within an energy recovery loop, and would allow qualification of RF performance of the SRF cavities and controls in conditions very close to the final conditions that would be experienced in the final source. Completing this work would place one at the completion of Phase I of the original Cornell X-ray source plan [134]. But in contrast to this plan, we recommend that the test facility adopt at least two accelerating passes and two decelerating energy recovery passes in order to fully explore whether final facility costs can be reduced by multiple-pass recirculation. At the end of experiments demonstrating beam requirements after this stage, one should be able to make final design choices to construct the best X-ray source possible. Multiple-pass recirculation is being discussed in reference to advanced Free Electron Lasers [136], making the issue of multiple pass beam recirculation of broader relevance than to just the ERL X-ray source community.

In parallel with and in support of all the experimental activity, it is our expectation that codes describing all the relevant beam phenomena would continue to improve. It will be necessary to have benchmarked codes and their predicted results available to assemble a final proposal. Undulators are, of course, also important in determining the photon characteristics of an ERL, just as they are for storage rings. The development work presently underway for storage ring light sources, such as the work on superconducting undulators and long undulators, would therefore also enhance the capabilities of ERLs.

To conclude, it was the consensus of the meeting that R&D activity establishing the viability of an ERL X-ray source could be completed in a period of five to ten years, depending on the rate at which funding was available devoted to this purpose. The program at Cornell University will provide much useful guidance, and to a certain extent our discussions have repeated and reinforced the soundness of the existing plans there for developing ERL-based X-ray sources. Any R&D plan ultimately adopted by

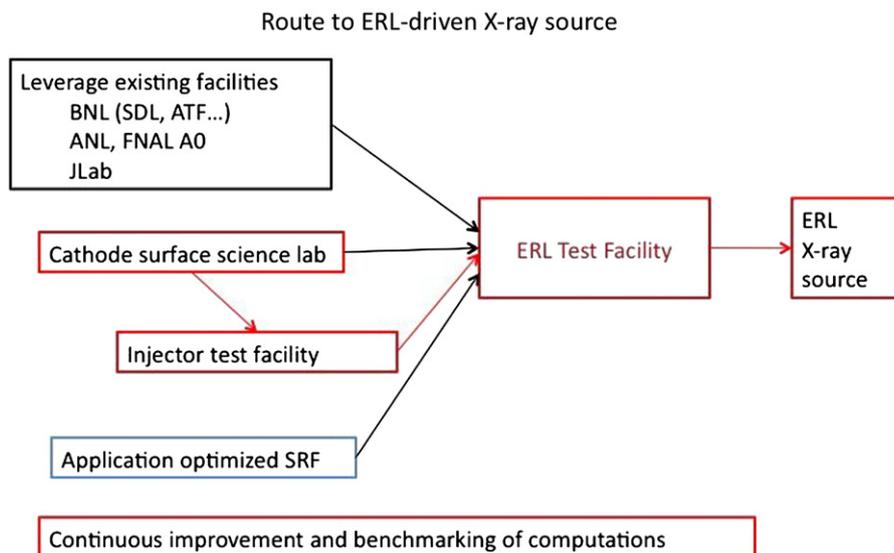


Fig. 2. Research path to an ERL X-ray source.

the Department of Energy for ERL X-ray source development should be highly integrated into the Cornell project, to avoid unnecessary duplication of effort. On the other hand, there are any number of issues, for example CW RF guns, long-term machine reliability, user operations, etc., where substantial Department of Energy support could greatly assist in ERL X-ray source development.

5. Summary

X-ray sources of a novel type, and with unique and interesting beam properties, can be built from multi-GeV scale energy recovered linacs. Energy recovered linacs are in their infancy, and high current multi-GeV devices will require a significant development effort to realize. There is a large body of interesting and substantial issues that could be addressed on existing facilities, even today, with R&D support. In addition to fully utilizing possibilities for measurements supporting ERL development at existing facilities, we foresee several phases in the development of large ERL machines including: demonstrating adequate photocathode performance for ERL applications at a cathode development facility, a high average current test injector that would demonstrate suitable beam performance characteristics for the large driver, and a high average current beam recirculation experiment with at least two accelerating and two decelerating beam passes. Concurrently with this effort, we believe that simulation software should be developed and improved to address physics in energy recovered linac accelerators and that R&D efforts be undertaken to develop SRF cavity technology better optimized to specific ERL applications. We believe that a proposal for a full energy X-ray source can be developed to level of detail suitable for such a major project, and construction of a full source begin in 5–10 years.

Acknowledgements

I. Bazarov, G. Hoffstaetter, and M. Tigner of Cornell University shared text, references, and Fig. 1. Their contributions improved this presentation and are gratefully acknowledged. This work supported by the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract nos. DE-AC02-06CH11357 and DE-AC03-76SF00515, and Office of High Energy and Nuclear Physics, under U. S. DOE Contract nos. DE-AC05-06OR23177 and DE AC02-98CH1-886, and the Commonwealth of Virginia. The U. S. Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce this manuscript for U. S. Government purposes.

References

- [1] M. Tigner, *Nuovo Cimento* 37 (1965).
- [2] K.-J. Kim, in: *AIP Conference Proceedings*, vol. 184, 1989, pp. 565–632.
- [3] M. Borland, Concepts and performance for a next-generation storage ring hard X-ray source, in: *Proceedings of the SRI09, Melbourne, 2009 AIP Conference Proceedings*, vol. 1234, 2010, p. 911.
- [4] G. Hoffstaetter, Paper WG203, in: *Proceedings of FLS2006, Hamburg, Germany, 2006*.
- [5] Conference Proceedings NIMA 557, <<http://conferences.jlab.org/ERL/2005/>>.
- [6] in: *Conference Proceedings of the ERL07, Joint Accelerator Conferences Website (JACoW), ABDW-ERL'07, 2007*. Available from: <<http://accelconf.web.cern.ch/accelconf/erl07/index.htm>>, <<http://www.astec.ac.uk/ERL07/S/>>.
- [7] in: *Proceedings of the ERL09, Joint Accelerator Conferences Website (JACoW), ABDW-ERL'09, Ithaca, NY, 2009*. Available from: <<http://www.lepp.cornell.edu/Events/ERL09/S/>>.
- [8] L. Merminga, in: *Proceedings of PAC2007, Albuquerque, NM, 2007*, pp. 22–26.
- [9] G.A. Krafft, in: *Proceedings of International LINAC Conference, Victoria, BC, 2008*, pp. 688–692.
- [10] R. Hajima, Status and future of ERLs, in: *Proceedings of PAC2009, Vancouver, B.C., 2009*.
- [11] C. Hernandez-Garcia, et al., Status of the Jefferson Lab ERL FEL DC photoemission guns, in: *Proceedings of ERL09, Ithaca, NY, 2009*.
- [12] N.G. Gavrilov, et al., *Nucl. Instr. and Meth. A* 575 (2007) 54.
- [13] B.M. Dunham, Challenges in guns and injectors, in: *Proceedings of ERL09, Ithaca, NY, 2009*.
- [14] D.H. Dowell, et al., *Appl. Phys. Lett.* 63 (1993) 2035.
- [15] E.J. Minehara, et al., *Nucl. Instr. and Meth. A* 445 (2000) 183.
- [16] H. Tanaka, et al., in: *Proceedings of FEL2009, Liverpool, UK, 2009*, pp. 758–765.
- [17] D. Dowell, et al., Cathode R&D for future light sources, BES Workshop on Accelerator Physics for Future Light Sources, Gaithersburg, MD, September 2009, companion paper.
- [18] I. Bazarov, C. Sinclair, *Phys. Rev. ST Accel. Beams* 8 (2005) 034202.
- [19] I. Bazarov, et al., *Phys. Rev. ST Accel. Beams* 11 (2008) 100703.
- [20] I. Bazarov, et al., *J. Appl. Phys.* 105 (2009) 083715.
- [21] B.M. Dunham, et al., in: *Proceedings of PAC 2007, 2007*, p. 1224.
- [22] S. Zhang, C. Hernandez-Garcia, Observation on heating of gas photo-cathode wafer by candidate drive laser, in: *Proceedings of HAPHB Workshop, UCLA, 2009*.
- [23] S. Zhou, D. Ouzounov, F.W. Wise, *Opt. Lett.* 31 (2006) 1041.
- [24] C. Limborg-DePrey, P.R. Bolton, *Nucl. Instr. and Meth. A* 557 (2006) 106.
- [25] I. Bazarov, et al., *Phys. Rev. ST Accel. Beams* 11 (2008) 040702.
- [26] G. Hoffstaetter, Cornell ERL Program, in: *Proceedings of ERL09, Ithaca, NY, 2009*.
- [27] M. Borland, ANL ERL Program, in: *Proceedings of ERL09, Ithaca, NY, 2009*.
- [28] V. Yakimenko, et al., in: *Proceedings of PAC2001, 2001*, p. 3344.
- [29] J.H. Wu, et al., in: *Proceedings of PAC2001, 2001*, p. 2866.
- [30] K. Brown, R. Servranckx, *Nucl. Instr. and Meth. A* 258 (1987) 480.
- [31] K.L. Brown, Second order Achromat, SLAC-PUB-3381, 1984.
- [32] Curvature correction during both bunching and energy recovery have been demonstrated in the JLab IR Demo and IR Upgrade (through octupole order in the latter case). See P. Piot, et al., *Phys. Rev. ST Accel. Beams* 6 (2003) 030702; D. Douglas, in: *Proceedings of LINAC2000, 2000*, pp. 716–720; and Benson and Douglas, U.S. Patent #7,166,973.
- [33] Yakimenko, et al., in: *Proceedings of PAC2001, 2001*, p. 3344; J.H. Wu, et al., in: *Proceedings of PAC2001, 2001*, p. 2866.
- [34] D.H. Dowell, T.D. Hayward, A.M. Vetter, in: *Proceedings of PAC1995, 1995*, pp. 992–994.
- [35] C. Wang, A. Chao, Analytic Second and Third order Achromat design, 1995.
- [36] W. Wan, M. Berz, Fifth-Order Achromatic, 1999.
- [37] L.M. Barkov, et al., *Part. Accel.* 31 (1990) 177.
- [38] V.V. Danilov, E.A. Perevedentsev, CERN Report No. SL/Note 94-21(AP); CERN Report No. SL/Note 94-74(AP).
- [39] E.M. McMillan, Report No. UCLRL-17795, University of California, Lawrence Radiation Laboratory, Berkeley, California, 1967.
- [40] L.J. Laslett, E.M. McMillan, J. Moser, Report No. NYO-1480-101, AEC Computing and Applied Mathematics Center, Courant Institute of Mathematical Sciences, New York University, 1968.
- [41] M. Sands, *The Physics of Electron Storage Rings*, SLAC-121.
- [42] Yakimenko, et al., in: *Proceedings of PAC2001, 2001*, p. 3344; Borland, PAC2009, Hoffstaetter, ERL2009.
- [43] K.L.F. Bane, et al., *Phys. Rev. ST A-B* 12 (2009) 030704.
- [44] D.C. Sagen, et al., *Phys. Rev. ST A-B* 12 (2009) 040703.
- [45] R. Kato, et al., *Phys. Rev. E* 57 (3) (1998) 57.
- [46] L. Groening, et al., Recent Experiments on the Effect of Coherent Synchrotron Radiation on the Electron Beam of CTF II, PAC, 2001, Chicago, p. 164.
- [47] M. Borland, X. Dong, Y. Li, in: *Proceedings of the SRI2009, Melbourne, AIP Conference Proceedings*, vol. 1234, 2009, 2010, p. 511.
- [48] M.P. Ehrlichman, G.H. Hoffstaetter, in: *Proceedings of the PAC2009, Vancouver, 2009*, pp. 3309–3311.
- [49] S. Benson, et al., PAC2007; G. Neil, FEL2008.
- [50] B. Buckley, G.H. Hoffstaetter, *Phys. Rev. ST A-B* 10 (2007) 111002.
- [51] A. Freyberger, in: *Proceedings of DIPAC Conference 2005, Lyon, France, 2005*, p. 12.
- [52] A. Bogacz, et al., in: *Proceedings of the 2003 Particle Accelerator Conference, 2003*, pp. 195–197.
- [53] R. Ursic, et al., in: *Proceedings of the 1995 Particle Accelerator Conference, 1995*, pp. 2652–2654.
- [54] E. Fischer, *CERN PS/Int. AR/60-14, 1960*.
- [55] K. Symon, Are clearing electrodes really necessary?, Private communication, UW-Madison, 1964.
- [56] E. Keil, B. Zotter, *CERN-ISR TH/71-58, 1971*.
- [57] Y. Baconnier, G. Brianti, *CERN/SPS/80-2, 1980*.
- [58] A. Poncet, *CERN/MT 90-1, CAS 1989, 1990*.
- [59] G. Hoffstaetter, M. Liepe, *Nucl. Instr. and Meth. A* 557 (2006) 205.
- [60] S. Van der Meer, *CERN PS/AA/Note 84-11, 1984*.
- [61] P. Emma, et al., in: *Proceedings of the Fourth EPAC, London, 1994*, pp. 1162–1164.
- [62] P.L. Colestock, et al., in: *Proceedings of the 2001 PAC, Chicago, 2001*, pp. 170–172.
- [63] T.O. Raubenheimer, in: *Proceedings of PAC95, Dallas, 1995*, pp. 2752–2754.
- [64] F. Zimmermann, et al., in: *Proceedings of PAC95, Dallas, 1995*, pp. 3102–3104.

- [65] G. Hoffstaetter, C. Spethmann, Phys. Rev. ST A-B 11 (2008) 014001.
- [66] T.P. Wangler, et al., in: Proceedings of LINAC 96, Geneva, 1996, pp. 372–374.
- [67] P.-Y. Beauvais, et al., in: Proceedings of PAC95, Dallas, 1995, pp. 3173–3175.
- [68] P. Krejčík, et al., in: Proceedings of PAC97, Vancouver, 1997, pp. 1626–1628.
- [69] F. Roncarolo, et al., in: Proceedings of EPAC08, Genoa, 2008, pp. 1680–1682.
- [70] G.A. Krafft, et al., in: Proceedings of International LINAC Conference, Monterey, CA, 2000, pp. 721–725.
- [71] M. Tiefenback, in: Proceedings of PAC01, 2001, pp. 553–555.
- [72] P. Evtushenko, et al., in: Proceedings of LINAC08, Victoria, B.C., Canada, 2008, pp. 606–608.
- [73] V.A. Lebedev, et al., Nucl. Instr. and Meth. A 408 (1998) 373.
- [74] C. Steier, Beam stabilization challenges in ERLs, in: Proceedings of ERL09, Ithaca, NY, 2009.
- [75] Beam Measurement, in: Proceedings of Joint US-CERN-Japan-Russia School on Particle Accelerators, Montreux, 1998. World Scientific, 1999.
- [76] H. Koziol, CERN/PS 2001-012 (DR), 2001.
- [77] S.G. Biedron, P. Piot, in: Proceedings of the 2009 High Average Current, High Brightness Beams Workshop, UCLA, 2009.
- [78] A. Lumpkin, Nucl. Instr. and Meth. A 557 (2006) 318.
- [79] S. Molloy, et al., in: Proceedings of PAC07, 2007, pp. 4201–4203.
- [80] D.X. Wang, et al., Appl. Phys. Lett. 70 (1997) 529.
- [81] P. Chevtsov, et al., Nucl. Instr. and Meth. Phys. Res.-A 557 (2006) 324.
- [82] S. Zhang, et al., in: Proceedings of FEL06, Berlin, 2006, pp. 740–743.
- [83] X. Yan, et al., Phys. Rev. Lett. 85 (2000) 3404.
- [84] T. Maxwell, P. Piot, Proposal for a non-interceptive spatio-temporal correlation monitor, in: Proceedings of PAC09, Vancouver, 2009.
- [85] A.L. Cavalieri, et al., Phys. Rev. Lett. 94 (2005) 114801.
- [86] E. Saldin, E. Schneidmiller, M. Yurkov, Nucl. Instr. and Meth. A 539 (2005) 499.
- [87] J. Bodewadt, et al., in: Proceedings of FEL08, Gyeongju, Korea, 2008, pp. 497–500.
- [88] T.R. Schibli, et al., Quantum Electronics and Laser Science (QELS) Postconference Digest (IEEE Cat No. CH37420-TBR), vol. 2, 2003.
- [89] A.B. Temnykh, Phys. Rev. Spec. Topics—Accel. & Beams 11 (2008) 120702.
- [90] J.B. Bahrtdt, et al., in: Proceedings of 2004 FEL Conference, Vancouver, 2004, p. 610.
- [91] Y. Ivanyushenkov, et al., in: Proceedings of PAC2005, 2005, p. 2295.
- [92] S.H. Kim, et al., in: Proceedings of PAC2007, 2007, p. 1136.
- [93] F. Finkelstein, et al., J. Phys. Chem. Solids 66 (2005) 2310.
- [94] M. Borland, et al., in: Proceedings of the PAC2009, Vancouver, 2009, pp. 44–48.
- [95] S. Sasaki, Exotic insertion devices, in: H. Onuki, P. Elleaume (Eds.), Undulators, Wigglers and Their Applications, vol. 237, Taylor & Francis, 2003.
- [96] R. Dejus, M. Jaski, S. Sasaki, in: Proceedings of the PAC2009, Vancouver, 2009, pp. 307–309.
- [97] T. Kamps, et al., in: Proceedings of SRF2009, Berlin, Germany, 2009, p. 164.
- [98] D. Janssen, et al., Nucl. Instr. and Meth. A 507 (2003) 314.
- [99] R. Legg, et al., SRF Photoinjector R&D at University of Wisconsin, in: Proceedings of the ERL09, Ithaca, NY, Proceedings of the ERL09, 2009, pp. 45–49.
- [100] A. Burrill, in: Proceedings of the ERL09, J.W. Lewellen, et al., BNL SRF Gun Development, Ithaca, NY, 2009, pp. 24–36.
- [101] J.W. Lewellen, et al., in: Proceedings of LINAC08, 2008, p. 495.
- [102] A. Arnold, J. Teichert, in: Proceedings of SRF2009, Berlin, Germany, 2009, p. 20.
- [103] Y. Ding, et al., Comparison of simulations and measurements for high brightness electron beams at LCLS, in: Proceedings of the ICAP2009, Proceedings of the PAC2009, Vancouver, 2009, pp. 2355–2357.
- [104] K. Harkay, et al., Photo cathode studies for ultra-low-emittance electron sources, in: Proceedings of the PAC2009, Vancouver, 2009, pp. 458–460.
- [105] M.N. Lakshmanan, I.V. Bazarov, T. Miyajima, Geometry optimization of DC and SRF guns to maximize beam brightness, in: Proceedings of ERL09, Ithaca, NY, 2009.
- [106] A. Xiao, et al., in: Proceedings of PAC2007, 2007, pp. 3453–3455.
- [107] A. Xiao, et al., in: Proceedings of Linac2008, 2008, p. 293.
- [108] A. Xiao, M. Borland, X. Dong, in: Proceedings of the ICAP2009, San Francisco, 2009, pp. 173–178.
- [109] G. Hoffstaetter, et al., in: Proceedings of EPAC2008, 2008, pp. 1631–1633.
- [110] C. Ng, et al., State of the art in EM field computation, in: Proceedings of EPAC2008, 2008, pp. 2763–2767.
- [111] M. Billing, Wake fields and impedance budgets for ERLs, in: Proceedings of ERL2009, Ithaca, NY, 2009.
- [112] N. Nakamura, Effects of longitudinal and transverse resistive-wall wake-fields on ERLs, in: Proceedings of ERL2009, Ithaca, NY, 2009.
- [113] M. Borland, Elegant: a flexible SDDS-compliant code for accelerator simulation, APS LS-287, 2000.
- [114] D. Sagan, Bmad: a relativistic charged particle simulation, Nucl. Instr. and Meth. A 558 (2006) 356.
- [115] J. Qiang, et al., J. Comp. Phys. 163 (2000) 434.
- [116] Y. Wang, et al., Recent progress on parallel elegant, in: Proceedings of the ICAP2009, THPsc054, San Francisco, 2009, pp. 355–358.
- [117] M. Borland, in: Proceedings of the Linac2008, TUP024, Victoria, 2008, pp. 444–446.
- [118] D. Sagan, et al., Phys. Rev. ST Accel. Beams 12 (2009) 040703.
- [119] M. Borland, et al., in: Proceedings of PAC2001, 2001, pp. 2707–2709.
- [120] S. Reiche, et al., in: Proceedings of PAC2001, 2001, pp. 2751–2753.
- [121] M. Borland, et al., Progress on an energy recovery upgrade to the advanced photon source, in: Proceedings of SRI09, to be published.
- [122] J. Bisognano, private correspondence.
- [123] I.V. Bazarov, B.M. Dunham, C.K. Sinclair, Phys. Rev. Lett. 102 (2008) 104801.
- [124] Appendix II of (D. Dowell, et al., Cathode R&D for future light sources, BES Workshop on Accelerator Physics for Future Light Sources, Gaithersburg, MD, September 2009, companion paper.).
- [125] B.M. Dunham, et al., in: Proceedings of XXIV Linear Accelerator Conference, 2009, p. 688.
- [126] See Reference (Proceedings of ERL09, Ithaca, NY, <<http://www.lepp.cornell.edu/Events/ERL09/>>, to be published.), particularly the Injector Working Group.
- [127] C.W. Leemann, D.R. Douglas, G.A. Krafft, Annu. Rev. Nucl. Part. Sci. 51 (2001) 413.
- [128] G.H. Hoffstaetter, I.V. Bazarov, C. Song, Phys. Rev. Spec. Topics—Accel. & Beams 10 (2006) 044401.
- [129] D.R. Douglas, et al., Phys. Rev. Spec. Topics—Accel. & Beams 9 (2006) 064403.
- [130] Y. Zhang, G.A. Krafft, Issue Editors, ICFA Beam Dynamics Newsletter #26, Recirculated Linac Light Sources, 2001, pp. 8–15.
- [131] J. Bardeen, L.N. Cooper, J.R. Schrieffer, Phys. Rev. 108 (1957) 1175.
- [132] E. Chojnacki, et al., in: Proceedings of the PAC2009, Vancouver, 2009, pp. 2781–2783.
- [133] F. Marhauser, JLAB high current cryomodule development, in: Proceedings of ERL09, Ithaca, NY, 2009.
- [134] S.M. Gruner, M. Tigner (Eds.), Phase I Energy Recovery Linac (ERL) Synchrotron Light Source at Cornell University, CHES Technical Memo 01-003 and JLAB-ACT-01-04, 2001.
- [135] P.A. McIntosh, et al., in: Proceedings of SRF09, 2009, p. 864.
- [136] F.E. Hannon, et al., in: Proceedings of FEL2009, Liverpool, UK, 2009, pp. 169–171.
- [137] L. Merminga, D.R. Douglas, G.A. Krafft, Annu. Rev. Nucl. Part. Sci. 53 (2003) 387.
- [138] M. Liepe, et al., in: Proceedings of PAC2005, 2005, pp. 2642–2644.
- [139] G. Ciovati, in: Proceedings of the 14th International Conference on RF Superconductivity, Berlin, Germany, 2009.
- [140] M. Liepe, SRF system optimization process, in: Proceedings of ERL09, Ithaca, NY, 2009.
- [141] J. Marangos, et al., New Light Source (NLS) Project Science Case and Facility Design, <<http://www.newlightsource.org>>, 2009.
- [142] V. Ganni, Design of Optimal Helium Refrigeration and Liquefaction Systems, Cryogenic Society of America, Oak Park, IL, 2009.



ELSEVIER

Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima

Review

The Potential of an Ultimate Storage Ring for Future Light Sources

M. Bei^a, M. Borland^b, Y. Cai^c, P. Elleaume^d, R. Gerig^b, K. Harkay^b, L. Emery^b, A. Hutton^e, R. Hettel^c, R. Nagaoka^f, D. Robin^{g,*}, C. Steier^g

^a Brookhaven National Laboratory, Upton, NY 11973, USA

^b Argonne National Laboratory, 9700 S. Cass Avenue, Argonne, IL 60439, USA

^c SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025, USA

^d European Synchrotron Radiation Facility, Polygone Scientifique Louis Néel, 6 rue Jules Horowitz, 38000 Grenoble, France

^e Thomas Jefferson National Accelerator Facility, 12000 Jefferson Avenue, Newport News, VA 23606, USA

^f Synchrotron SOLEIL, L'Orme des Merisiers Saint-Aubin, BP 48, 91192 Gif-sur-Yvette, France

^g Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

ARTICLE INFO

Article history:

Received 30 December 2009

Accepted 20 January 2010

Available online 29 January 2010

Keywords:

Storage ring

Synchrotron light source

Facility design

Research and development plan

ABSTRACT

This paper is the report of the working group on Ultimate Storage Rings at the Department of Energy's Basic Energy Sciences Workshop on Physics of Future Light Sources, which took place in Gaithersburg, Maryland on September 15–17, 2009. In this report we address the accelerator design issues related to the next generation of storage ring light sources, deemed “ultimate” storage rings. In our estimation, storage rings have the potential to provide an increase in photon brightness and coherent flux that is two orders of magnitude above that projected for rings currently under construction. In addition to photon brightness and coherent flux, we discuss other directions, such as shorter pulses, tailored bunches, and partial lasing, in which rings could evolve. For the most part we envision ultimate storage rings as an evolutionary advance from existing rings that faces no fundamental technological obstacles. Nevertheless we identify several important areas of R&D that should be pursued to enable the realization of the full potential of ultimate ring light sources.

© 2010 Elsevier B.V. All rights reserved.

Contents

1. Executive summary	519
2. Introduction	519
3. Brightness, flux, and coherence	520
3.1. Undulator brightness, flux, and coherence	521
3.1.1. Flux	521
3.1.2. Brightness	521
3.1.3. Coherence	521
3.2. Potential for increases in undulator flux, brightness, and coherence	521
3.2.1. Diffraction limited emittance	522
3.2.2. Reducing the horizontal emittance	522
3.3. Challenges and mitigation strategies	522
3.3.1. Dynamic and momentum aperture	523
3.3.2. Intrabeam scattering	523
3.3.3. Instabilities and impedance	524
3.3.4. Beam stability	526
4. Other directions	526
4.1. Shorter pulses and terahertz radiation	526
4.2. Tailored bunch operation	527
4.3. Partial lasing	527
5. Design challenges and R&D for ultimate rings	528
5.1. General recommendation—dedicated zero order design study	528

* Corresponding author.

E-mail address: DSRobin@lbl.gov (D. Robin).

5.2.	Optimizing average brightness	528
5.2.1.	Computational tools.....	528
5.2.2.	Instability mitigation.....	529
5.2.3.	Injection schemes	529
5.2.4.	Instrumentation.....	530
5.2.5.	Stability design and stabilizing systems	531
5.2.6.	Radio frequency system.....	531
5.2.7.	Insertion device developments	532
5.2.8.	High heat load photon optics design	532
5.3.	Additional capabilities.....	533
5.3.1.	RF deflecting cavities.....	533
5.3.2.	Tailored bunch operation.....	533
5.3.3.	Partial lasing.....	534
5.4.	Cost reduction	534
6.	Summary	534
	Acknowledgments.....	535
	References	535

1. Executive summary

Storage rings are the principle sources of high-brightness photon beams driving the majority of X-ray science experiments in the world today. There has been remarkable progress in developing these sources over the last two decades. Existing third generation light sources continue to upgrade their capabilities, while new light sources are coming on line with ever improved performance. As X-ray sources, storage rings combine many attractive features. They provide a wide, easily tunable energy spectrum from infrared to hard X-ray having high flux and average brightness. The beams are highly stable in energy, intensity, position, and size. Storage rings easily serve many experiments (> 40) simultaneously, are highly reliable, and offer low cost per user. With this combination of properties, storage rings are complementary to other light sources such as free electron lasers (FELs), which offer extremely high peak brightness in much shorter pulse durations but with typically far lower repetition rate. A broad class of X-ray science relies on the low peak brightness (to avoid over-exciting or even damaging samples) and high photon pulse repetition rates (to reach sufficient flux) provided by storage rings; such experiments simple cannot be conducted using the ultra-high peak brightness from FEL sources [1]. Storage rings will continue to be the workhorse sources for a large user community for the indefinite future.

While storage rings are a “mature” technology, they nevertheless have the potential for significantly enhanced performance. One can imagine an “ultimate” storage ring that produces high-brightness, transversely coherent X-rays while simultaneously serving dozens of beamlines and thousands of users annually. For such a source to maximize transverse photon coherence, the beam emittance must be extremely small in both transverse planes, approaching and even exceeding the wavelength-dependent diffraction limit. Storage ring sources have achieved diffraction limited emittances for hard X-rays in the vertical plane by minimizing horizontal-vertical electron beam coupling, but horizontal emittance must be reduced by a factor of 100 or more from the lowest emittance values achieved today to reach that limit. Groups studying possible designs for ultimate rings, which, given present technology, would necessarily have large circumferences to reach low emittance values, have identified no technological show-stoppers for such an implementation. However, focused R&D efforts for storage rings would reduce the cost and risk of a possible project while delivering a design with optimized performance. An ultimate storage ring would retain all the general strengths of today’s storage rings mentioned above

while delivering high transverse coherence up to the 10 keV energy regime. Ultimate rings would have brightnesses and coherent flux one or two orders of magnitude higher than the highest performance ring-based light sources in operation or presently being constructed.

In this paper we begin with an introduction to storage ring light sources and a discussion of their performance parameters and future performance potential. Next we examine the possibilities for enhancing the photon brightness, flux, and coherence by identifying the potential approaches and the challenges related to each. We then explore the possibilities for enhancing the performance in other directions such as short-pulse and partial lasing modes of operation. Finally we will present our assessment of the most beneficial R&D directions. We emphasize that the technical risks associated with realizing the performance of ultimate storage rings are comparatively modest.

To advance this promising choice for a 4th generation light source, we recommend funding a team of experts to define an optimized and realistic zeroth order design. Ideally this would be a joint international effort including at a minimum ANL, SLAC, ESRF, and SPring-8. In terms of specific R&D we recommend efforts in simulation and codes, further development of injection schemes, insertion devices, RF systems, bunch manipulation, and instrumentation.

2. Introduction

Storage ring-based light sources are among the most successful scientific instruments of the last 25 years. These facilities produce high-flux, high-brightness radiation that spans a large spectral region from far infra red to hard X-rays. Storage ring light sources have many attractive features, including high stability and reliability, with simultaneous cost-effective service to many users with diverse requirements. Because of these features, which are detailed below, storage rings have become essential tools for many fields of science, including medicine, protein crystallography, physical and chemical science, and environmental science. As an illustration, we note that in the last six years, four Nobel prizes have been awarded to scientists who have used synchrotron light sources to understand protein structures. Currently, the synchrotron light sources in the United States alone serve more than 9000 users per year.

Storage ring light sources have evolved considerably over the last 50 years. Starting from the first generation as parasitic operational modes on high energy physics accelerators to the present day, the brightness alone has increased by 15 orders of

magnitude [2]. The current, third generation of storage ring light sources were designed to enhance the brightness of radiation emitted from magnetic undulators, insertion devices that are dedicated to the production of quasi-monochromatic synchrotron radiation. These light sources first came on-line in the 1990s, beginning with the construction of several large, high electron-energy facilities optimized for hard X-rays, namely the ESRF in France (6 GeV), the APS in the US (7 GeV) and SPring-8 in Japan (8 GeV). These sources were complemented by the construction of the smaller, lower electron energy facilities optimized for soft X-rays, such as the ALS in the US (1.9 GeV), BESSY-II in Germany (1.7 GeV), the TLS in Taiwan (1.5 GeV), the PLS in Korea (2.5 GeV), ELETTRA in Italy (2.4 GeV) and the SLS in Switzerland (2.4 GeV). As insertion device technology improved, it was recognized that facilities with hard X-ray capabilities could be built in a more cost-effective manner at a medium electron energy around 3 GeV, having few-nanometer emittances. This has resulted in the construction of facilities such as SOLEIL in France (2.75 GeV), DIAMOND in the UK (3 GeV), and the SSRF in China (3.5 GeV) which have come on line in the 2000s. The TPS in Taiwan and the LNLS-2 in Brazil are facilities of similar scale that are now being designed. These and other existing facilities, including upgraded older facilities like the Photon Factory in Japan and SPEAR3 in the US, continue to evolve beyond what was envisioned at their onset.

New facilities, notably PETRA III in Germany, NSLS-II in the US and MAX IV in Sweden, will come on-line with further improved parameters, most notably with horizontal emittances of 1 nm-rad or less. These machines tout another performance metric, besides brightness, that has become steadily more important for X-ray science, especially for the imaging community: coherence. Anticipating the future beyond these advanced machines, where increased coherent flux as well as spectral brightness are performance goals, several groups have been conducting for storage rings that would have emittances on the 100 pm-rad level or less, towards the diffraction limit for multi-keV X-rays [3–8]. Representative brightness envelopes for such “ultimate” storage rings and present-day sources are shown for comparison in Fig. 1. The coherent fraction of photons in the multi-keV regime for these sources is several percent, approaching 100% for soft X-rays.

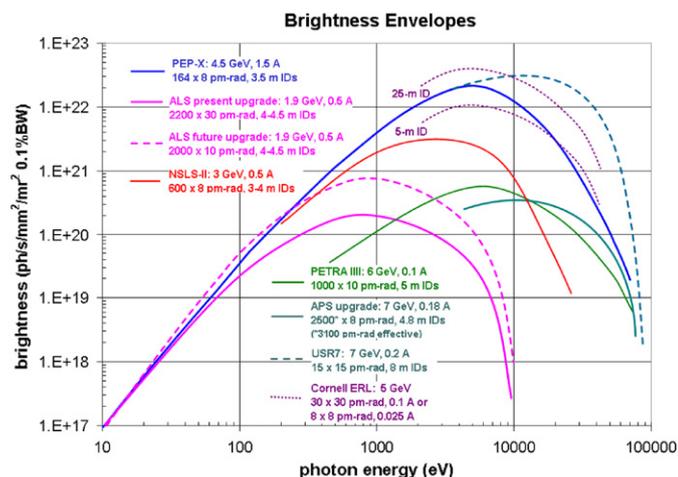


Fig. 1. Approximate brightness envelopes for representative existing and future ultimate storage ring designs assuming non-superconducting undulator lengths as specified in the figure. Higher brightnesses can be reached on these sources by using longer undulators and eventually with superconducting undulators. Partial lasing in long undulators may enhance the brightness for soft X-rays at energies < 1 keV. Approximate performance of the Cornell ERL [14] is shown for comparison.

It is reasonable to ask how far storage rings can evolve and where will they be competitive in the future with other types of sources.

Ultimate storage rings are commonly envisioned as sources that optimize average brightness and coherence by reducing horizontal emittance well below nanometer values. However, even within that definition there are several possible optimizations. For example one may consider using either high beam current (and, because of intrabeam scattering, somewhat enlarged emittance) to optimize photon flux (incoherent and coherent) together with the brightness, or lower beam current to realize higher coherent fractions of photons at the expense of flux. More broadly, ultimate storage rings could also be designed to provide an optimum suite of advanced capabilities, including very high repetition rate (up to 100 s of MHz) short pulses, tailored bunches, or partial lasing in the soft X-ray regime.

In this paper we explore the performance limits of ultimate storage rings and delineate what research and development would help in their actual implementation. Section 3 is a review of photon brightness, flux, and coherence, followed by a discussion of the potential improvements and challenges for rings. Section 4 describes other directions, such as shorter pulses, tailored bunches, or enhanced lasing, where rings also have potential for further development. Section 5 presents our recommendations of where R&D is needed to bring an ultimate ring to its full potential. Finally, Section 6 presents summary remarks.

3. Brightness, flux, and coherence

In a storage ring light source, an electron (or positron) beam circulates around the ring repetitively, generating electromagnetic radiation for the most part via spontaneous emission. The quality of the radiation is directly related to the quality of the electron beam and the radiating devices that the beam encounters—dipoles, wigglers, and undulators. Undulators provide the highest photon brightness, concentrated in wavelength bands around a design fundamental wavelength and its odd harmonics. Therefore, it is instructive to investigate the potential for maximizing undulator brightness.

The wavelength λ of radiation emitted by an undulator must satisfy the relationship

$$\lambda = \frac{\lambda_u}{2\gamma^2} \left(1 + \frac{K^2}{2} \right) \quad (1)$$

where λ_u is the undulator period, γ is the relativistic Lorentz factor, $K \approx \lambda_u B e / (2\pi m_0 c)$ (the “undulator parameter”), and B is the undulator peak magnetic field. For an undulator, $(1 + K^2/2)$ typically ranges between 1 and 5. The term $\lambda_u / (2\gamma^2)$ shows that one can arrive at the same wavelength using larger beam energies and longer periods, or lower beam energies and shorter periods. While improvements in undulator technology have resulted in higher performance, shorter period devices, choosing the optimum beam energy remains a complex problem. Many parameters enter, including cost, natural emittance, intrabeam scattering, beam instabilities, heat load on optics, magnetic material properties, and the desired photon wavelength range. In general, even with advanced undulator technology, it is still true that low energy rings (~ 2 GeV) allow optimized performance only at low photon energies (below 20 eV), intermediate energy rings (~ 3 GeV) provide excellent performance from a few 100 eV to above 10 keV, and, if higher photon energies are needed, higher electron energies are necessary (~ 4.5 –7 GeV).

In the following we review the three dominant performance metrics for light sources: flux, brightness, and coherence [9].

3.1. Undulator brightness, flux, and coherence

3.1.1. Flux

Spectral photon flux, $F(\lambda)$, is a measure of the rate of flow of photons within a bandwidth $\Delta\lambda/\lambda$ (typically 0.1%) at wavelength λ and is typically expressed in the following units:

$$F(\lambda) = \frac{\text{number of photons}}{(\text{s})(0.1\% \text{ bandwidth})}. \quad (2)$$

In an undulator

$$F(\lambda) \propto N_{und} I \quad (3)$$

where N_{und} is the number of undulator periods and I is the electron beam current. Clearly, the desire for high flux drives one towards larger current and favors ring designs with space for long insertion devices (i.e., accommodated in long straight sections).

3.1.2. Brightness

Spectral photon brightness, $B(\lambda)$, of light emitted at wavelength λ from an insertion device in a synchrotron light source obeys the following relationship:

$$B(\lambda) = \frac{F(\lambda)}{(2\pi)^2 \sigma_{Tx} \sigma_{Ty} \sigma_{Tx'} \sigma_{Ty'}} \quad (4)$$

where $F(\lambda)$ is the photon flux and σ_{Tx} , σ_{Ty} , $\sigma_{Tx'}$, and $\sigma_{Ty'}$ are the λ -dependent norms of the respective electron and photon beam sizes and divergences given in the horizontal plane, for example, by

$$\sigma_{Tx} = \sqrt{\sigma_x^2 + \sigma_r^2} \quad (5)$$

$$\sigma_{Tx'} = \sqrt{\sigma_x'^2 + \sigma_r'^2} \quad (6)$$

where σ_x (σ_x') is electron rms size (divergence) and σ_r (σ_r') is intrinsic photon size (divergence) at wavelength λ . In an insertion device of length L tuned to emit at a wavelength λ , σ_r and σ_r' are given by

$$\sigma_r = \frac{1}{2\pi} \sqrt{\frac{\lambda L}{2}} \quad (7)$$

$$\sigma_r' = \sqrt{\frac{\lambda}{2L}} \quad (8)$$

$$\varepsilon_r = \sigma_r \sigma_r' = \frac{\lambda}{4\pi} \quad (9)$$

$$\beta_r = \frac{\sigma_r}{\sigma_r'} = \frac{L}{2\pi}. \quad (10)$$

The photon brightness, $B(\lambda)$, is expressed in the following units:

$$B(\lambda) = \frac{F(\lambda)}{(\text{mm}^2)(\text{mrad}^2)} = \frac{\text{number of photons}}{(\text{s})(\text{mm}^2)(\text{mrad}^2)(0.1\% \text{ bandwidth})}. \quad (11)$$

For a given flux, maximum brightness is achieved when the electron beam emittance is reduced toward and beyond the intrinsic

diffraction-limited emittance $\lambda/(4\pi)$ of the photons and when the electron beam beta function is close to β_r . Maximal brightness is achieved when electron emittance is sufficiently small that the photon emittance dominates the value of the denominator in Eq. (4).

3.1.3. Coherence

Simply stated, coherence is a measure of the degree to which the radiation can exhibit interference patterns. The transverse coherence refers to the coherence of electromagnetic vibrations at two points perpendicular to the propagation direction. The transverse coherent flux at wavelength λ is given by

$$F_{\text{coh},T}(\lambda) = B(\lambda) \left(\frac{\lambda}{2}\right)^2. \quad (12)$$

The fraction f_{coh} of photon flux that is transversely coherent is related to the ratio of the intrinsic photon emittance to the total emittance of the photon beam:

$$\begin{aligned} f_{\text{coh}} &= \frac{F_{\text{coh},T}(\lambda)}{F(\lambda)} = \frac{\lambda/(4\pi) \lambda/(4\pi)}{\sigma_{Tx} \sigma_{Tx'} \sigma_{Ty} \sigma_{Ty'}} \\ &= \frac{1}{\sqrt{1 + \left(\frac{\sigma_x}{\sigma_r}\right)^2} \sqrt{1 + \left(\frac{\sigma_x'}{\sigma_r'}\right)^2} \sqrt{1 + \left(\frac{\sigma_y}{\sigma_r}\right)^2} \sqrt{1 + \left(\frac{\sigma_y'}{\sigma_r'}\right)^2}}. \end{aligned} \quad (13)$$

It should be noted that when the beta-matching criterion discussed in Section 3.1.2 is met and the electron emittance is equal to the intrinsic photon emittance, then $f_{\text{coh},T}(\lambda)$, the coherent fraction is 25%. Also, the coherent flux, $F_{\text{coh},T}(\lambda)$, from a high-current, low coherent fraction storage ring can exceed that from a low-current, high coherent fraction ring. Flux, brightness, coherent flux, and coherent fraction are not the only figures of merit for a synchrotron light source, but they are always important considerations. Which of the four quantities is most important depends upon each particular experiment.

3.2. Potential for increases in undulator flux, brightness, and coherence

A number of studies [3–8,10–13] have been carried out worldwide to advance storage ring technology for cost effectiveness and/or higher brightness. Table 1 summarizes the main design characteristics of advanced facilities under construction, in commissioning, or planned, as well as recent studies aimed at designing ultrasmall emittance lattices. The anticipated potential for the average brightness of spontaneous radiation in the soft and hard X-ray regime for advanced and ultimate storage rings is indicated in Fig. 1. Table 2 and Fig. 2 show the huge advances that ultimate storage rings promise to deliver in terms of the coherent fraction. The increases can be two to three orders of magnitude compared to existing sources, giving performance very similar to that promised by proposed spontaneous emission energy recovery linacs (ERLs) in their so-called high coherence mode.

Table 1
Advanced storage ring facilities that are either under commissioning, construction, or study.

Project	Energy (GeV)	Circumf. (km)	Horiz. emitt. (nm)	Current (A)	Lattice design	Status
PETRA III [10]	6	2.3	1	0.1	7/8 FODO + 1/8 DBA + DW	Under Commissioning
NSLS-II [11]	3	0.792	0.6	0.5	30 × DBA + DW	Under Construction
MAX IV [12,13]	3	0.528	0.24	0.5	20 × 7 BA + DW	MoU signed
USRLS [3]	7	2	0.3	0.5	50 × 4BA	Design Study
XPS7 [4]	7	2.2	0.08	1.0	80 × 6BA	Design Study
PEP-X [5]	4.5	2.2	0.1	1.5	DBA + TME + DW	Design Study
Tsumaki [6,7]	6	1.44	0.07	0.1	20 × 10BA, 4 × 5BA, 4 × LSS + DW	Design Study
USR7 [8]	7	3.1	0.015	0.2	40 × 10BA	Design Study

Table 2
Coherent fraction f_{coh} of existing and proposed or future storage ring light sources.

	f_{coh} 1 keV	f_{coh} 12 keV
Present (ALS, APS, ESRF)	0.01	0.0004
NSLSII	0.04	0.002
'Ultimate' future	0.75 (maybe more with coherent enhancement)	0.13

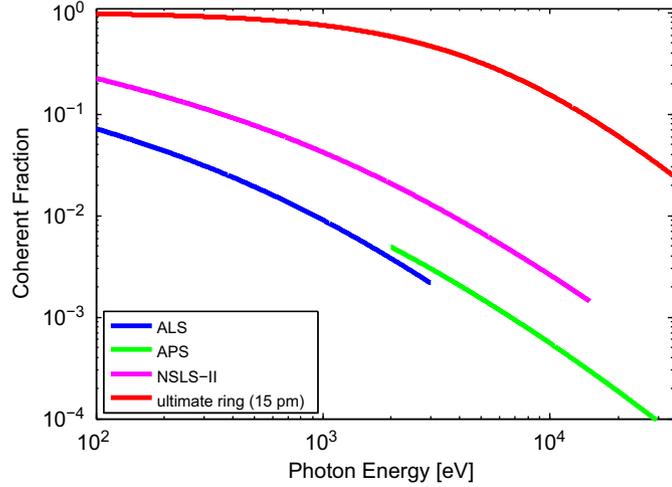


Fig. 2. Coherent fraction versus photon energy for some existing, planned, and ultimate storage rings.

Because of the generally higher flux of the ultimate ring based sources, the actual number of coherent photons may well be higher than is feasible in ERLs.

In the next section we discuss how such high brightness and coherence can be achieved in these new rings together with the challenges for doing so, where higher current a large number of periods in insertion devices are needed maximize flux, maximizing brightness and coherent flux requires both high flux and operating the rings closer to diffraction-limited emittances.

3.2.1. Diffraction limited emittance

The diffraction-limited light source performance for a given wavelength is met when $\varepsilon_T = \lambda/(4\pi) = \sigma_T \sigma_{T'}$ in each plane, where the subscript T indicates the total beam sizes and divergences derived from the convolution of the intrinsic photon and electron beam sizes and divergences (Eq. (5)). To approach diffraction-limited performance: (1) the electron beam emittance, $\sigma_{x,y}\sigma_{x',y'}$, must be comparable to or less than the intrinsic photon beam emittance, $\lambda/(4\pi)$ and (2) the phase space of the electron and photon beam sizes must be matched (i.e., $\sigma_r/\sigma_{r'} \approx \sigma_x/\sigma_{x'} \approx \sigma_y/\sigma_{y'}$). For hard X-ray radiation at 0.1 nm, $\lambda/(4\pi) = 8$ pm rad, while for soft X-rays at 1 nm the diffraction limit is 80 pm rad. For the best existing storage rings, the vertical electron beam emittances have closely approached the diffraction limit for hard X-rays. However, the horizontal emittances for present-day rings are nearly three orders of magnitude larger than the diffraction limit. Reducing the horizontal electron emittances as well as matching the electron and single photon phase spaces is essential in order to improve coherent fraction. Matching the phase spaces means that in addition to small emittances one would like to have relatively small β - functions in the undulator, i.e., $\beta_x \approx \beta_y \approx \beta_r = L/(2\pi)$. Since the β - function is s -dependent

and quadratically increases from the waist, having such a small β -function in the insertion device center will result in a large β -function at the ends of the insertion device. To minimize the β -function within the insertion device results in choosing a β -function of $\beta_y = L/2$ at the center. Minimizing the β -function everywhere also permits the vertical undulator gap to be minimized and photon flux to be maximized. In the following we will explore the limits on the horizontal emittance.

3.2.2. Reducing the horizontal emittance

The horizontal emittance ε_x results from the combined effect of an excitation of horizontal betatron oscillations, S_x , and its damping, τ_x :

$$\varepsilon_x = S_x \tau_x. \quad (14)$$

Both excitation and damping are induced by synchrotron radiation emitted within the accelerator magnets and are given by

$$S_x \approx E^5 \oint B^3 \frac{\eta_x^2 + \left(\beta_x \eta'_x - \frac{\beta'_x \eta_x}{2}\right)^2}{\beta_x} ds, \quad \frac{1}{\tau_x} \approx J_x E^3 \oint B^2 ds \quad (15)$$

where B is the magnetic field, E the electron energy, β_x the horizontal betatron function, $\beta'_x = d\beta_x/ds$ is the derivative of the betatron function versus the longitudinal coordinate s , η_x is the horizontal dispersion function and $\eta'_x = d\eta_x/ds$. J_x is the horizontal damping partition number, a dimensionless quantity typically close to one. The integrals along the longitudinal coordinate s in Eqs. (14) and (15) are non-zero only in the bending magnets and the insertion devices, while they are zero in the drift spaces (no field) and negligible in the quadrupole and sextupoles (negligible field close to the axis). To complete the analysis, one must remember that β_x and η_x are varying with the longitudinal coordinate s . In all machines built so far or in construction (with the exception of PETRA III and NSLS-II) the contributions of the insertion devices to S_x and τ_x are negligible and the essential contribution comes from the bending magnets. It is clear from Eqs. (14) and (15) that a small emittance requires the use of a lattice with small values of β_x and η_x in the bending magnet. How small these quantities can be is related to the length of the bending magnet. One is therefore driven to segment the bending magnets into as many pieces as possible and refocus both β_x and η between each pair of successive bending magnets. This is achieved by means of a triplet or quadruplet of quadrupoles. The resulting lattice is a multi-cell DBA or a multiple-bend achromat (MBA). The emittance from such a lattice scales asymptotically as

$$\varepsilon_x \propto \frac{E^2}{N^3} \quad (16)$$

where N is the number of bending magnets. The lattices [12,3,4,6,8] of Table 1 fall into this category. Therefore, to go to small emittances requires a large number of bending magnets which in turn leads to large ring circumferences, so that emittance scales roughly as $1/C^3$. The reason that the circumference tends to be large for lattices with many bending magnets is that more quadrupoles need to be added between the short bending magnets.

3.3. Challenges and mitigation strategies

The main challenges in terms of optimizing flux, brightness, and coherence are:

- Small dynamic aperture which has implications on injection and beam lifetime.
- Intrabeam scattering which has implications on the smallest achievable emittances.
- Low single bunch instability thresholds.

- Higher stability requirements due to smaller horizontal beam sizes.

3.3.1. Dynamic and momentum aperture

Because the bending magnets are relatively short and the focusing is frequent and strong in an ultimate ring lattice, the dispersion is small everywhere. All quadrupoles generate chromaticity which must be compensated by means of sextupoles, the efficiency of which scales with the dispersion. In other words, as the number of bending magnets is increased, stronger sextupoles are needed which result in more highly nonlinear betatron motion and reduction of the so-called dynamic and momentum apertures. A small dynamic aperture results in greater difficulty with injection, while a small momentum aperture results in reduced beam lifetime. Conventionally, difficult injection can be mitigated by improved engineering to reduce the separation between injected and already stored beams at the injection septum, and by reducing the beam emittance from the injector. Lifetime reduction can be compensated by top-up injection and in some cases with a bunch lengthening rf cavity. However, there are limits potentially encountered with ultimate ring lattices—particularly with conventional injection which requires sufficiently large dynamic aperture and betatron function to allow accumulation of multiple pulses in an rf bucket.

There are several routes worth investigating to increase the dynamic and momentum aperture or alleviate the operating consequences just noted. These approaches are not necessarily mutually exclusive, but may be combined:

- The horizontal betatron function in the injection straight can be enlarged to accommodate the space needed for the injected beam, the stored beam, and the thickness of the septum. One possible limit to this process is the breaking of periodicity of the lattice which can further reduce the dynamic aperture unless several of these modified straight sections are located symmetrically along the ring circumference.
- Another approach is to preserve the high periodicity of a lattice (keep the β -function low in the center of the injection straight) and to use a less conventional form of injection, namely, pulsed multipole injection [15–17]. With pulsed multipole injection it is possible to inject with a single injection kicker which has no field at the origin. This injection kicker could be placed at a location where the betatron function is large [18].
- Dynamic aperture can be improved by introducing extra families of chromatic sextupoles, harmonic sextupoles (in dispersion free regions), and octupoles. As the lattices become highly nonlinear, one may have to simultaneously optimize both linear optics and nonlinear dynamics including the position and setting of sextupoles. Typically, we strive to minimize the nonlinear chromaticity, width of resonances driven by sextupoles, and footprints in tune space. More recently, tracking-based methods using genetic algorithms have shown promise for direct optimization of the dynamic and momentum aperture [19,20]. In addition to a proper optimization of the sextupoles, one may add octupoles to control the tune shifts with betatron amplitude as carried out in the MAX IV lattice design [12,13].

For very small dynamic apertures (e.g., USR7), one cannot use the conventional accumulation-based injection using a closed kicker bump combined with a septum magnet. Even a pulsed multipole injection scheme (as discussed above) might not work. An additional option in this case is the on-axis injection. While this would seem to restrict operation to low current, this is not in

fact the case provided one is willing to replace all or part of the fill at short time intervals, using a concept known as “swap-out” [8]. The main benefit is the feasibility of an ultra-small horizontal and vertical emittance of 15 pm for a 3.1 km storage ring with 7 GeV beam energy, including emittance growth by intrabeam scattering. More details of this concept are discussed below.

Instead of a bending magnet-dominated lattice, one can design a damping wiggler dominated lattice where the field in the bending magnets is reduced, so that the dominant contribution to both excitation S_x and damping τ_x is from the wigglers rather than bending magnets. The benefit of damping wigglers is that for a similar field, S_x is reduced as the dispersion can stay much smaller than in a bending magnet. Compared to an MBA design, such a lattice will make different use of the circumference. It will save space as it does not need as many quadrupoles and sextupoles. On the other hand, it will require more space as the low-field bending magnets need to be longer to close the circumference, while a significant part of the circumference is occupied by the wigglers and the additional rf cavities to compensate for the energy loss from wigglers. Note that such a trend of shifting from bending magnet to damping wigglers to optimize the emittance is already observable on the PETRA III and NSLS-II lattices. A full comparison of an MBA conventional-type lattice and damping wiggler-dominated lattice both optimized for an ultra small emittance is still to be done. While wiggler-dominated rings may have a large horizontal aperture, they may also induce a reduction of the vertical aperture if the period is too short.

3.3.2. Intrabeam scattering

As already mentioned, intrabeam scattering (IBS) is one of the fundamental limitations in achieving ultrasmall emittances in storage rings. In preparation for linear collider damping ring designs, where IBS plays an important role as well, extensive studies were carried out [21,22]. The theory of intrabeam scattering is thus very well understood and with only small modifications to the theory developed decades ago for proton accelerators, the measurements agreed very well with calculations. The longitudinal emittance growth rate $1/T_p$ can be calculated as [21]:

$$\frac{1}{T_p} \approx \frac{r_e^2 c N_b (\log)}{16 \gamma^3 \epsilon_x^{3/4} \epsilon_y^{3/4} \sigma_s \sigma_p^3} \langle \sigma_H g(a/b) (\beta_x \beta_y)^{-1/4} \rangle \quad (17)$$

where r_e is the classical electron radius, c the speed of light, N_b the number of particles per bunch, (\log) the Coulomb log factor, γ the Lorentz factor corresponding to the beam energy, $\epsilon_{x,y}$ the transverse emittances, σ_s the bunch length, σ_p the energy spread, and $\beta_{x,y}$ the optical beta functions. The other factors in Eq. (17) are given by

$$\frac{1}{\sigma_H^2} = \frac{1}{\sigma_p^2} + \frac{\mathcal{H}_x}{\epsilon_x} + \frac{\mathcal{H}_y}{\epsilon_y} \quad (18)$$

$$a = \frac{\sigma_H}{\gamma} \sqrt{\frac{\beta_x}{\epsilon_x}}, \quad b = \frac{\sigma_H}{\gamma} \sqrt{\frac{\beta_y}{\epsilon_y}} \quad (19)$$

$$g(\alpha) = \alpha^{(0.021 - 0.044 \ln \alpha)} \quad (\text{for } 0.01 < \alpha < 1) \quad (20)$$

with $\mathcal{H} = (\eta^2 + (\beta\eta' - \frac{1}{2}\beta'\eta)^2)/\beta$ being the usual dispersion invariant. The transverse emittance growth rate $1/T_{x,y}$ can be calculated from the longitudinal one:

$$\frac{1}{T_{x,y}} \approx \frac{\sigma_p^2 \langle \mathcal{H}_{x,y} \rangle}{\epsilon_{x,y}} \frac{1}{T_p} \quad (21)$$

Intrabeam scattering decreases rapidly with higher beam energy, favoring higher energy storage ring designs. It also is inversely proportional to the damping time, favoring damping

wigglers. However, the emittance growth effect depends on the size of the \mathcal{H} -function at the location of a scattering event, favoring lattice designs that already have small natural emittance without damping wigglers.

The intrabeam scattering growth rates depend on the product of the horizontal and vertical emittances, which are related by $\varepsilon_x = \varepsilon_0/(1+\kappa)$ and $\varepsilon_y = \varepsilon_0\kappa/(1+\kappa)$, where κ is the emittance coupling. The product $\varepsilon_x\varepsilon_y$ is maximized when $\kappa = 1$, showing that round beams in general have a much smaller bunch density than flat beams, resulting in significantly less emittance growth due to intrabeam scattering. However, running a storage ring with $\kappa = 1$ is incompatible with conventional beam accumulation, since the large horizontal injected beam amplitude couples into the vertical plane, resulting in beam loss on the undulator chambers. However, if on-axis injection is used, it is possible to operate with fully coupled beams. Using on-axis injection, either by a full beam swap-out technique, by replacing individual bunches, or by replacing short bunch trains [23], can be an enabling technology to minimize the emittance growth due to intrabeam scattering.

Using round beams in ultimate storage rings to minimize emittance growth due to intrabeam scattering does not compromise the brightness performance as it would in current-generation storage rings. This is due to the fact that ultimate storage rings have a horizontal emittance close to the diffraction limit. So increasing the vertical emittance to be equal to the horizontal one is transparent to users, since the emittance is only increased up to the diffraction limit. Fig. 3 shows intrabeam scattering simulations of one example design for an ultimate storage ring at 7 GeV beam energy [24]. Using round beams, the emittance including intrabeam scattering effects stays virtually constant (close to the diffraction limit for hard X-rays at 10 keV), even for bunch charges larger than the nominal ones for high flux operation.

Let us explore this further by taking the USR7 design as an example. For full coupling, a Touschek lifetime of 4 h is predicted for 50 $\mu\text{A}/\text{bunch}$. With 4000 bunches, this gives a total current of 200 mA, which seems reasonable for a 7 GeV ring with undulators of up to 8 m length. We may imagine grouping the bunches into 200 trains of 20 bunches, each train separated by a short gap. With a 500 MHz rf system and a 3.16 km circumference, the bunch train gap would be 12 ns. It is not outside the realm of possibility to build kickers with rise and fall times this short, which would allow us to kick out a single 20-bunch train (towards a beam dump) and simultaneously replace it with a new bunch train from the injector. It is interesting to estimate the requirements on the injector, which we consider as a baseline for discussion. As described in Ref. [8], one basic parameter is the fractional depletion D of the bunch train that is acceptable before a bunch train is replaced. If we assume $D=0.1$ and a 200 mA store with a 4 h lifetime, we must replace a bunch train every

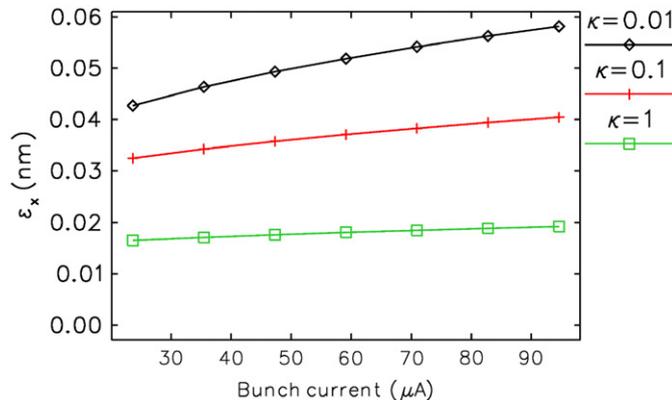


Fig. 3. Effect of intrabeam scattering emittance growth versus current for different values of emittance coupling [24].

7 s. The charge in the train is 10.5 nC (0.5 nC/bunch), giving an average injector current of 1.5 nA. The macropulse current in the baseline linac would be 260 mA. None of these parameters (plus the required booster synchrotron) are particularly challenging. However, this may become infeasible if sufficiently fast kicker rise and fall times cannot be achieved. In this case, another option is to use an accumulator ring to prepare a full or partial store that could be swapped with the store in the ring. Like some modern booster rings, the accumulator ring could be built in the same tunnel as the storage ring itself, but would be of much simpler design due to the lack of straight sections for insertion devices. Due to the amount of beam that would be swapped, the beam extracted from the storage ring would be returned to the accumulator and recycled. Details of this (more expensive) scheme, including issues such as the required dynamic aperture and emittance of the accumulator itself, need to be worked out.

3.3.3. Instabilities and impedance

Storage ring light sources are a mature technology for which collective instabilities and impedance are largely understood. Instabilities expected to be of importance for ultimate storage rings include collective interactions of the beam with self-induced wake fields due to the resistive wall impedance, chamber geometric effects, and coherent synchrotron radiation (CSR). However, the design parameters of an ultimate storage ring represent a new regime. The choice of the lattice and the machine parameters could potentially result in low current thresholds for instabilities as discussed below. The formulas in the following section are intended to give rough parameter dependence of the different types of instabilities. They are not meant to give a complete picture or allow exact quantitative evaluations since the full calculation of instability thresholds is significantly more complex and often involves the use of numerical codes. For example the interaction between the various collective effects usually leads to more relaxed instability thresholds than the simple estimates below would predict.

As described earlier in Section 3.2, the main objective of an ultimate storage ring is to increase brightness and coherence, which are inversely proportional to the emittance of the electron beam. A common practice in lowering the emittance is to make the machine circumference L_c larger, divide the dipole magnets into many pieces and make use of the $1/N^3$ dependence of the horizontal emittance, where N is the number of dipoles. As a consequence, however, the momentum compaction factor α , which acts on the longitudinal bunching of a beam, tends to become small, as the dispersion function η_x diminishes while the dipole radius of curvature ρ_0 tends to get larger, as seen in

$$\alpha = \frac{1}{L_c} \oint \frac{\eta}{\rho_0} ds. \quad (22)$$

The smaller momentum compaction in turn shortens the zero current bunch length σ_s , as it scales as the square root of the former:

$$\sigma_s = \sqrt{\frac{\alpha R}{J_i \rho_0} \frac{2\pi C_q}{(mc^2)^2} \frac{E^3}{e\omega_{rf} V_{rf} \cos\varphi_s}}, \quad (23)$$

where E is the machine energy, e the electron charge, c the speed of light, m the electron mass, $C_q = 3.84 \cdot 10^{-13}$ m, J_i the damping partition number, with ($i=x, y, \varepsilon$), V_{rf} the RF voltage, ω_{rf} the angular RF frequency, φ_s the synchronous phase, and R the machine radius.

A shorter bunch length lowers instability thresholds in general, but especially affects single bunch instabilities. This is due to the wider spectrum of the beam caused by the short bunches, creating a larger overlap with the ring impedance up to high frequencies, thus giving a larger effective impedance as seen by

the beam. Similarly, a smaller momentum compaction factor results in lowering the synchrotron tune ν_s ,

$$\nu_s = \sqrt{\frac{\alpha h V_{rf} \cos \varphi_s}{2\pi E/e}} \quad (24)$$

where h is the rf harmonic number. A smaller synchrotron tune can also enhance single bunch instabilities, as the coupling between different bunch modes, which are separated by ν_s , increases. In particular, the transverse mode coupling instability (TMCI) between the dipolar and the quadrupolar modes is a potential challenge for ultimate rings. In the simplest approximation the threshold can be estimated as the beam current for which the detuning of the dipole mode Δf_β , which can be calculated using [25]:

$$\frac{df_\beta}{dI} = -\frac{\beta_{avg}}{8\pi^{3/2}\sigma_s E/e} \cdot \Im(Z_T)_{eff} \quad (25)$$

roughly equals the synchrotron ν_s given in Eq. (24). Here, β_{avg} represents the average beta function and $\Im(Z_T)_{eff}$, the imaginary part of the effective impedance in the concerned transverse plane. It might be worth to point out that for ultra low emittance lattices, the average beta function β_{avg} is smaller than in higher emittance lattices, leading to smaller df_β/dI for a given impedance. Using a more complete description of TMCI, current thresholds have been calculated for several example designs of ultimate storage rings. For PEP-X the calculated TMCI threshold is 0.7–0.8 mA, well above the design bunch current of 0.5 mA, which corresponds to the high average design current of 1.5 A [26].

The threshold current I_{th} of the microwave instability, a longitudinal single bunch instability associated with a gradual increase in energy spread, is approximately given by [27]

$$I_{th} = \frac{3/2\omega_0^3\sigma_{th}^3 V_{rf} |\cos \varphi_s|}{\sqrt{2\pi\Re(Z_L)_{eff}/p_r}} \quad (26)$$

This threshold current again is expected to be lower [28] in ultimate storage rings, owing to the short bunches. Here, $\omega_0 = 2\pi f_0$, f_0 = revolution frequency and σ_{th} is the bunch length at the instability threshold, $\Re(Z_L)_{eff}$ and p_r denote, respectively, the effective real part of the longitudinal impedance and the revolution harmonic ($p_r = f_{res}/f_0$) for the resonant frequency of the impedance.

Another effect, affecting most existing light sources, that will also have an impact for an ultimate ring, is the overall small size of the vertical aperture of the vacuum chamber, due to small-gap insertion devices and to reduced aperture, strong-focusing quadrupoles used to reach the lower emittance. The resistive-wall impedance is enhanced as it scales cubically with the aperture, potentially bringing about resistive-wall instability in the multibunch filling, whose growth rate τ^{-1} for a given multibunch current I can be estimated by [28]

$$\tau^{-1} = \frac{\beta_{avg}\omega_0 I}{4\pi E/e} \frac{R}{b_{eff}^3} \sqrt{\frac{2\rho}{(1-\Delta\nu_\beta)\omega_0\epsilon_0}} \quad (27)$$

where ρ is the resistivity of the chamber material, $\Delta\nu_\beta$ is the fractional betatron tune, ϵ_0 is the dielectric constant, and b_{eff} denotes the effective half chamber aperture. In addition, the larger the machine circumference, the lower the resistive-wall instability threshold will be since the magnitude of the most dangerous spectral line of the resistive-wall impedance scales as $1/\sqrt{\omega_0}$. Reducing emittance again works unfavorably for this instability, although higher energy is beneficial. It must be noted that the known recipe of shifting the chromaticity to positive may not necessarily stabilize the beam as the higher-order head-tail modes are more likely excited by the broadband impedance, due again to the short bunch nature of the ultimate machine (Fig. 4). In view of the resistive-wall instability in multi-bunch and the

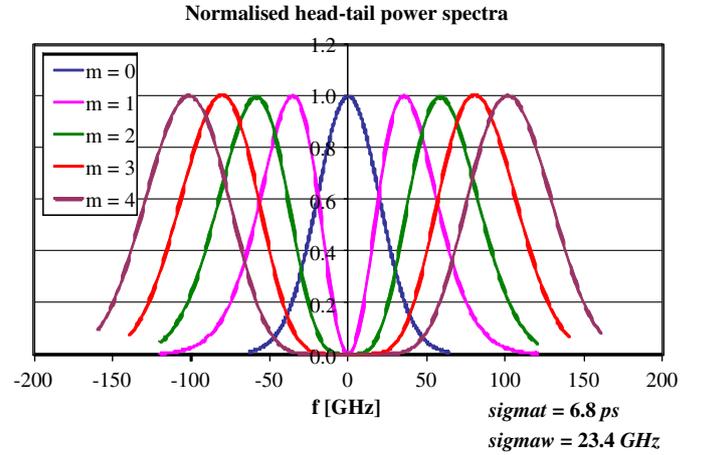


Fig. 4. Distribution of head-tail mode spectra [28].

TMCI in single bunch implementation of a bunch by bunch transverse feedback system will be indispensable. State-of-the-art multibunch feedback systems have demonstrated fast damping rates and very good noise performance and are likely to be sufficient to combat any multibunch instabilities in ultimate storage rings. The instability situation due to resistive wall impedance described above is already encountered in recently built storage rings such as SOLEIL [29]. Nevertheless, feedback systems, fill pattern manipulations and operation with large chromaticity allows to successfully achieve design performance.

In addition to the impedance-induced instabilities described above, beam-ion effects (particularly the fast ion instability) and possibly beam-electron cloud effects may prove to be sources of beam instability, diagnostics interference, or chamber heating, in ultimate rings. Again, the smallness of the emittance is likely to enhance two-beam interactions, as the growth rate of these instabilities basically scales inversely proportionally to the $\frac{3}{4}$ power of the product of horizontal and vertical beam emittance [30]. Another possible source of enhancing the beam-ion instability in an ultimate ring may be the out-gassing due to heating of local vacuum components caused by the interaction of short bunches with the resistive impedance, which is often appreciable at high frequencies. Intensive studies of those effects related to linear collider damping rings have been ongoing for years, with overall encouraging results. Further study in this area will be beneficial, but the effects are not likely to be a showstopper for ultimate storage rings [26].

In summary, single-bunch current or peak current instability thresholds present a potentially significant challenge for implementing ultimate rings, especially with lattices with very low momentum compaction and small bunch length. An integrated design approach is needed whereby the lattice parameters and current thresholds are considered together in order to optimize the overall performance. This should include looking at schemes for lengthening the electron bunch by reducing rf frequency and/or using harmonic cavities to increase microwave and TMCI current instability thresholds.

Regarding impedance analysis, existing computational tools and resources have generally been adequate in calculating the geometric impedance in rings, the small-gap undulator chambers in particular, for present-day rings. It must be pointed out that the above is true thanks to the great progress of the code development made in the last decades in calculating electro-magnetic fields excited by the beam in general 3D structures using parallel-processing techniques. In several existing light sources, the

impedance of vacuum components were minimized at the design stage through such numerical studies, and computed geometric and resistive wall impedances have been shown to be consistent with observed single- and multi-bunch instability thresholds.

For future ultimate storage rings, on the other hand, these computational tools may not be adequate in predicting the impedance, since the ratio of the bunch length to the length of the vacuum structures (e.g., low-gap insertion device chambers and cavities) has decreased significantly. The main difficulty lies in the computational power required to calculate the impedance up to several tens of GHz of long vacuum ducts possessing simultaneously sub-millimeter structures due to RF fingers, flanges, pumping slots, and BPM electrodes. The small mesh size required renders the computation impracticable with present-day tools. Furthermore, the effect of surface roughness may become significant and demands further study. Also, the contribution of the coherent synchrotron radiation (CSR) impedance may be non-negligible. In addition to the effort of evaluating and minimizing the magnitude of different impedance contributions, it would also be important to study their combined effect, such as between the resistive-wall and geometric impedances, which has not been understood well so far. The overall risks associated with impedance and instability limitations are estimated to be medium and it is expected that they can be mitigated with appropriate R&D.

3.3.4. Beam stability

Beam stability is one of the most important properties for the users of a synchrotron light source: stability of orbit, beam size, current (lifetime), energy, and energy spread. Beam stability is a specific strength of storage ring based light sources largely because of the inherent intensity constancy, particularly if top-up injection is being used. As light sources increasingly generate higher brightness, add fast-switching variable polarization devices, and produce smaller sources sizes, there is a necessity for continuous improvements in beam stability.

The requirements for beam stability for synchrotron radiation experiments vary widely, depending on the sensitivity of the user experiments to beam parameter fluctuations [31]. Specifically, the sensitivity is a function of the beam line and endstation configuration, the detector and measurement methods and time scales, the experimental sample characteristics and the photon beam properties. To establish general goals for the design and operation of facilities, the diverse requirements can be condensed into a generic list of stability goals. As one can show, for most experiments the relevant stability goals are expressed relative to the beam size or divergence, typically a few percent of those dimensions. Therefore, future light sources with smaller, micron-level beam sizes, will require better stability than today's facilities. In addition, for storage rings, lattice amplification factors usually increase if the lattices are optimized for smaller emittances. This also increases the demands for designing highly stable accelerator and beam line components, and implementing advanced active stabilizing systems. The developments necessary to provide the required better beam stability in ultimate storage rings are evolutionary and very similar to what is required for FEL and ERL based future light sources. In fact, the required improvements compared to the current state-of-the-art are probably smaller for ultimate storage rings, because certain sources of instability, namely the laser system of the photon guns as well as the complete injector chain, are not relevant for storage rings. Nevertheless, the development programs for beam stabilization provide many opportunities for cooperation among the various source types.

4. Other directions

Besides increasing the spontaneous flux, brightness, and coherence in a storage ring, there are other directions that are worth discussing to describe areas of potential improvement. They include:

- Shorter pulses and terahertz generation.
- Tailored bunches to simultaneously satisfying high flux and dynamics uses.
- Partial lasing.

4.1. Shorter pulses and terahertz radiation

The duration of the X-ray pulse generated at a synchrotron light source is typically tens of picoseconds. For many applications shorter pulses are highly desired by the users. In addition, shortening the electron bunch length gives rise to an enhancement in the production of terahertz radiation—which is also desirable for some users. In this section we briefly discuss the potential and challenges for generating shorter pulses and enhanced terahertz radiation.

Similarly to the electron emittance, the equilibrium electron bunch length is a balance between quantum excitation and radiation damping. At low current the bunch length is proportional to the momentum compaction α , the energy spread σ_δ , and inversely proportional to the synchrotron tune ν_s :

$$\sigma_s \propto \frac{\alpha}{\nu_s} \sigma_\delta. \quad (28)$$

Several methods have been proposed to shorten the X-ray pulse radiated from a synchrotron. They can be divided into three categories. The first category includes those approaches that either vary or modulate the longitudinal phase space parameters, such as increasing the rf voltage, installing a higher harmonic rf system, lowering the momentum compaction factor [32], or modulating the rf phase or voltage [33–35]. Methods of the second category make use of the short duration of a laser pulse. Thomson scattering [36] and femtosecond laser slicing [37,38] fall into this group. The third group of methods take advantage of the smaller vertical beam size in a storage ring. The rf deflecting cavity method [39] and vertical kick method [40] belong to this category.

In general it is difficult in a ring to get to very short pulse lengths with high flux. The experience has been that it is possible to reduce the bunch length down to the picosecond level with lower current by reducing the momentum compaction factor. This short electron bunch length is accompanied by an enhancement in terahertz radiation that in some cases has extended as low as 200 μm . The limitation appears to be unavoidable impedance issues. Due to the low current and large emittances incipient with low momentum compaction operation, this mode of operation is not desirable for high brightness X-ray users. A mode of operation that produces short pulses for some users and is compatible with high brightness operation for other users is the femtosecond slicing method. This technique allows for very short (~ 100 fs) pulse lengths with the penalty of much lower flux and has been successfully demonstrated at a number of facilities. One drawback of the femtosecond slicing technique is that it is less effective at higher electron energy rings.

One technique—rf deflecting cavities—offers the possibility of producing few-picosecond pulses with higher flux. In the rf deflection concept [39], illustrated in Fig. 5, transverse deflecting rf cavities ('crab cavities') are used to impose a correlation ('chirp') between the longitudinal position of an electron within the bunch and its vertical momentum. A chirped X-ray beam is then produced in a downstream undulator, evolving into a spatial

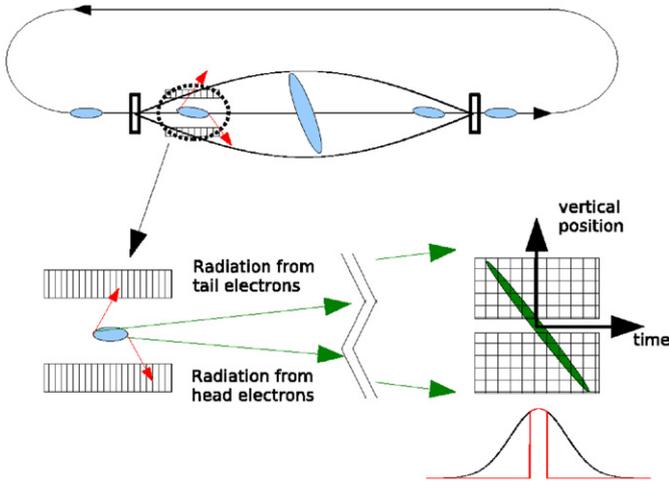


Fig. 5. Illustration of Zholents' scheme [39] for short X-ray pulse production using deflecting rf cavities in a storage ring.

chirp in the beam line so that a strong correlation appears in the X-ray pulse between arrival time and vertical coordinate. Use of vertical slits then permits filtering the pulse in the time dimension, allowing production of an X-ray pulse that is shorter than the electron pulse. This technique, which in principle is compatible with normal high brightness operation, has yet to be practically demonstrated, largely due to the cost of the deflecting cavities, which must be superconducting for CW operation, or pulsed at sub-kHz repetition rates if they are normal-conducting. In spite of the potential for short bunch production, this scheme might not be compatible with ultimate storage ring operation for several reasons. First, as discussed earlier, ultimate rings are likely not to have the dynamic aperture to support the beam manipulations inherent in this scheme. Second, the strong sextupoles needed in such a ring will very likely lead to large emittance growth [41] using this scheme. Finally, short bunch operation in future rings may not be so important in a world where other 'non-ring-based' sources such as FELs will provide the short pulses needed by the community.

While rf deflecting cavities may not be compatible with future ultimate rings, their use on present-day X-ray storage rings is highly desirable, and thus R&D on rf deflecting cavities is recommended. Should the technique prove to be successful it would be available to future general-purpose facilities.

4.2. Tailored bunch operation

Typically storage ring light sources operate with the maximum number of bunches as possible with one or more small gaps for ion clearing. However, this mode may not be desirable for users doing dynamics or time-of-flight experiments using the photons from a single bunches, each bunch followed by a gap, to excite a time-dependent response in an experimental sample. Often these users would like the gaps between bunches to be large, implying that only a few bunches should be stored in the ring (the number being more or less linearly dependent on ring circumference). The standard approach to satisfying these users is to have separate running times with only a few bunches in the ring—which is not a desirable mode for high flux users. Large rings may use a so-called "cam-shaft" mode with a single, isolated intense bunch together with many weak bunches on the opposite side of the ring, but this is less workable in smaller rings. The difficulty of satisfying both high flux applications and time of flight applications simultaneously is one of

the main limitations of synchrotron light sources as they are now operated.

In principle it is possible to tailor the properties of individual bunches in such a way as to satisfy both classes of users simultaneously [42,43]. The basic theme for these schemes is that one 'tailors' the characteristics of individual bunches. By changing, for example, the orbit or energy of one bunch with respect to the remaining bunches, it is possible to separate the light from that bunch either geometrically or chromatically.

In addition to tailored intensity, it may be possible to do more. For instance it may be possible to generate a shorter pulse length in individual bunches and separate the radiation from radiation from the long bunch train. The ability to generate short pulses is of great interest to many dynamics users. There already exist a few schemes for generating short pulses in storage rings as discussed above, including using femto-slicing on selected bunches. It may be possible to extend this or some other technique to generate short pulses for some users while storing high current for other users. This versatility would greatly extend the utility and capabilities of usefulness and capabilities of storage ring light sources.

4.3. Partial lasing

For a light source relying on the spontaneous radiation in undulators, the ultimate attainable brightness is on the order of 10^{23} , in the usual units, at keV photon energies. One method to exceed this barrier is to utilize the amplification process occurring in an FEL. It is well known that high electron brightness is critical for FEL operation. In particular, in the transverse planes, the emittance must be close to or less than $\lambda/4\pi$. As we have discussed earlier, this condition could be satisfied up to a wavelength of $\lambda \approx 0.1$ nm in an ultimate storage ring. This leads to a possibility of FEL lasing at soft and hard X-ray wavelengths in such a storage ring. For simplicity, let us consider the one-dimensional FEL model. The power gain length L_G in an undulator having a period of λ_u is given by

$$L_G = \frac{\lambda_u}{4\pi\sqrt{3}\rho} \quad (29)$$

where ρ is the Pierce parameter:

$$\rho = \left[\frac{1}{8\pi I_A} \left(\frac{K[J]}{1+K^2/2} \right)^2 \frac{\gamma\lambda^2}{(\pi\sigma_x)^2} \right]^{1/3} \quad (30)$$

Here \hat{I} is the peak current, I_A is the Alfvén current (about 17 kA), $[J]$ is the Bessel function factor associated with a planar undulator

$$[J] = J_0 \left(\frac{K^2}{4+2K^2} \right) - J_1 \left(\frac{K^2}{4+2K^2} \right) \quad (31)$$

σ_x is the electron beam size, and λ is the FEL resonance wavelength given by Eq. (1). Aside from having small enough emittance, the FEL is operated most efficiently when the electron beam size σ_x is well matched to the size σ_r

$$\sigma_r = \sqrt{\frac{\lambda}{4\pi} \frac{\lambda_u}{4\pi\rho}} \quad (32)$$

of the FEL photon beam [44]. This implies that the beta function β_x at the undulators should approximately be $\lambda_u/4\pi\rho$ (on the order of a few meters for λ_u of a few centimeters). Moreover, the energy spread σ_δ has to be smaller than ρ or otherwise there will be a significant increase of the gain length.

With a typical λ_u of a few centimeters, one needs $\rho \approx 10^{-3}$ in order to have a reasonable gain length. This requires a sufficient peak current \hat{I} as indicated in Eq. (30). In the PEP-X study [5] for

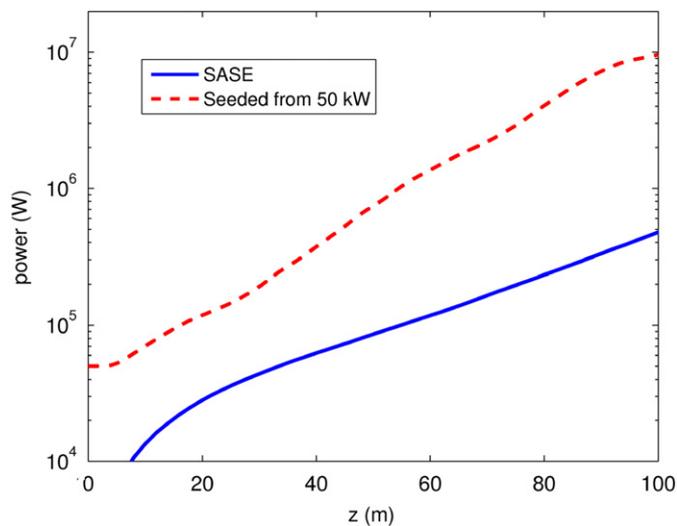


Fig. 6. Power evolution for a soft X-ray SASE FEL at $\lambda_r = 3.3$ nm from a single bunch having peak current of 270 A in a 4.5 GeV ultimate ring with 0.6- μ m normalized emittance. The radiating undulator has a 5 cm period with $K = 4.3$. The average power from a single bunch is 0.7 W. A seeded FEL simulation is shown for comparison.

the soft X-ray wavelengths, the required peak current is several hundreds of amperes. Electron beam requirements can be relaxed or radiated power increased if seeding is used. The relatively large energy spread in storage rings will most likely limit lasing to soft X-ray wavelengths. Once the conditions for lasing are satisfied, one can either operate the FEL with stored beam or use a switched bypass for the FEL. In the former case, full single-pass saturation in the FEL is not achieved. Instead, the lasing process reaches an equilibrium where emission is enhanced by one or two orders of magnitude from the spontaneous case (Fig. 6), regulated by the increase in energy spread of the stored beam that acts to terminate lasing well before saturation [45].

5. Design challenges and R&D for ultimate rings

Clearly there is an opportunity for improving the performance of storage ring light sources well beyond even the most advanced sources, such as NSLS-II and PETRA III, that are being built today. Studies have shown that there is the potential for orders of magnitude increase in brightness, coherent fraction and coherent flux. There is also room for development in other areas such as short pulse operation at high repetition rates (MHz) or more flexible operational modes using bunch tailoring. There is even the possibility of enhanced emission at lower photon energies using partial lasing. Since the performance of a synchrotron light source cannot be optimized in all these different areas simultaneously, the definitions of “ultimate” storage rings must remain varied. However, no matter what performance parameters acquire particular emphasis—such as optimizing the brightness and coherent flux based upon spontaneous emission—more study is required to understand precisely what is possible, and research and development efforts will be needed to realize maximal performance of ultimate rings.

Because there is a complex tradeoff in parameters, our first recommendation is that a dedicated study be undertaken to give a detailed global understanding of where the important issues lie. At the same time, it is clear that we can already begin with R&D in several directions, which we will list below. There are areas where targeted R&D will be beneficial for maximizing photon flux,

brightness, and coherence. There are areas of development where targeted R&D could further enable short pulse operation, tailored bunch operation, and low-level lasing. In all areas benefits can be gained by improved simulation tools, in some cases beyond the state-of-the-art, as well as performing experiments on existing facilities. Many of the developments, namely simulation codes, insertion devices, stabilization systems, and instrumentation to just name a few, are cross-cutting and will equally benefit other future light sources based on FELs or ERLs. It is important to note that many of these R&D areas will be of benefit not only for future ultimate rings, but also for improving the performance of existing and planned storage rings.

In the following sections, we group the areas of study and topics for R&D for ultimate rings into four categories:

- General studies.
- Optimization of average brightness and coherence.
- Producing advanced photon properties other than average brightness coherence, such as short or tailored pulses, partial lasing, etc.
- Cost optimization and value engineering, which in turn can lead to higher performance accelerators.

Given sufficient resources, it is expected that the R&D needed to arrive at an optimized design for an ultimate storage ring can be completed in a 5-year timeframe. Because storage rings are such a mature technology, this R&D is not of a fundamental nature. Rather, it is an extension and further exploration of well-known topics. As such, there is relatively high confidence of ultimate success.

5.1. General recommendation—dedicated zero order design study

Our primary recommendation is to fund a team of experts to define an optimized and realistic zeroth order design. Ideally this would be a joint international effort including at a minimum scientists and engineers from ANL, SLAC, ESRF, and SPring-8 and perhaps other laboratories. In particular the design optimization should include considerations of beam energy, lattice design, dynamic aperture, emittance, beta function, total current, instability thresholds, injection schemes, bunch lengthening and manipulation schemes. Such a study would help to prioritize the specific areas where targeted R&D would be most beneficial. This design study mostly requires effort and computing resources.

5.2. Optimizing average brightness

As previously explained, the main method for maximizing brightness and coherent fraction, reducing emittance, has many side effects that need to be mitigated. The R&D necessary to achieve ultimate brightness and coherence by necessity also includes all measures to mitigate the effects of ultrasmall emittances. Those measures include improving the small dynamic and momentum aperture, enabling injection, and raising bunch current instability thresholds. To be able to make full use of the ultimate brightness one also needs to achieve improved stability. Finally all of those optimizations need advanced computational tools. In some cases, code developments beyond the current state-of-the-art are necessary.

5.2.1. Computational tools

As discussed above, further development of existing codes will be necessary to conduct comprehensive beam dynamics simulations for ultimate rings. For example, to correctly predict single bunch instability thresholds and optimize the machine design relative to such thresholds, as well as optimize the geometry of

individual vacuum system components, impedance calculations up to the highest frequencies contained in the bunch spectrum are necessary. To carry those out, development is required of advanced simulation codes for short bunches in long structures including, for example, parallel 3-D computation codes having higher order accuracy that extends the maximum frequency range up to many tens of GHz [46–49]. Another effect that currently cannot be modeled with the necessary accuracy is the shielding of the coherent synchrotron radiation (CSR), limiting the ability to calculate the potentially detrimental effects of CSR on the electron beam. Specifically, advanced codes are needed to compute CSR shielding on an element-by-element basis. Ideally, these new codes would extend beyond existing codes the types of vacuum chamber geometries for which the shielding effect can be calculated.

In order to be able to find the optimum lattice design that maximizes brightness, raises instability thresholds, improves stability, and maximizes the dynamic and momentum apertures, it would be very beneficial to further develop direct high-performance methods for simultaneously optimizing the linear and nonlinear lattice properties. This would include the development of methods for efficient, automated evaluation of large numbers of candidate designs. Finally, one wants to carry out extensive beam dynamics studies including the ring lattice, realistic impedance and ion models, feedback, Landau cavities, and both single-bunch and coupled-bunch effects due to both impedance and ions. These studies will help to simultaneously optimize the lattice both for nonlinear single particle dynamics and for increased instability thresholds.

5.2.2. Instability mitigation

Mitigating electron bunch instability is a central challenge for ultimate storage rings due to the fact that ultrasmall emittance lattices inherently have small momentum compaction factors and therefore potentially small bunch current instability thresholds.

Maximizing single bunch and multibunch instability thresholds: Maximization of bunch current instability thresholds must be considered an integral part of the lattice and machine optimization process. There are many potential conflicting factors in an optimization and the trade-offs between them need to be studied in detail. Design options begin with the lattice, namely the optimum choice of momentum compaction factor, chromaticity, and zero-current bunch length while considering single-bunch current thresholds, anomalous bunch lengthening, and energy spread blowup. It might be helpful to evaluate options for increasing the momentum compaction factor without significantly increasing the emittance. This could be achieved by soft bending magnets, damping wigglers, or even chicanes. To raise instability thresholds further, one could use Landau cavities to lengthen the bunch and thereby stabilize single-bunch and multibunch instabilities. However, one needs to evaluate the effect of long bunches on ring performance and user requirements. Bunch-by-bunch feedback is an important mitigation measure and its performance need to be considered in realistic simulations. This is especially important in the case that, despite the optimization listed above, one is still left with relatively fast growth rates due to classical impedance-induced and/or fast-ion instabilities, or the case of higher-order head-tail mode excitations. An important issue in this context is that the noise performance of such feedback must be extremely good so as not to spoil the extremely small emittance by external noise. Once this study is completed, the design should be well documented and the mitigation schemes should be tested on existing rings as much as possible.

Ion- and electron-cloud effects: As has been realized in the design of many of the early 3rd generation light sources and even

more so the damping rings for the international linear collider, ion- and electron-cloud effects can be very important for high-current beams having very small emittances. While these instabilities are generally not a significant problem for existing rings, the quantitative understanding of the observed phenomena is not sufficient to make definitive predictions for rings having much smaller emittance. Therefore, one should further analyze trapped-ion and fast-ion effects in existing storage rings, and carry out experiments to allow extrapolation to ultimate storage ring parameters. The experimental data should be benchmarked with codes, and improve codes as needed. Studies of electron cloud effects in future rings should also be carried out, although the effects are expected to be rather weak for electron rings and may only be a source of noise in the beam position monitor diagnostics or heating of superconducting chamber surfaces by direct scattering of the electron cloud under acceleration by the beam. This work can be carried out in very close collaboration with work already ongoing for damping ring development, for example at CESR-TA [50].

Reduced impedance: Complimentary to the improvement of instability thresholds by lattice design, bunch lengthening, or other design choices is the reduction of the impedances driving instabilities. Modeling shows that for many light sources the dominant contribution to the transverse broadband impedance comes from the small-gap undulator chambers. It is therefore of great interest to see if one can minimize the impedance of these small-gap undulator chambers through advanced modeling and automated design. Taking a step further, one could optimize the chamber aperture for the entire ring, considering chamber transition impedance versus resistive wall impedance. To complement and improve the computation and design, it is important to continue developing beam-based impedance measurement techniques and use data to benchmark theory and computations.

With high beam current and/or short bunches, the heating of components can become significant. Therefore, one should analyze the beam-induced heating of vacuum components. One also needs to analyze the impedance budget for high average current. An issue that has had significance for X-ray FEL designs is also becoming more important for storage rings: the surface roughness. This is especially relevant in the context of small-gap insertion devices, as well as NEG coated vacuum chambers, which are desirable because of cost and vacuum performance. The effect of chamber surface roughness needs to be modeled to determine specifications for the ultimate storage ring. There could be cross-cutting collaboration with FEL and ERL projects on this topic.

Finally, once all of the impedance design studies are progressed far enough, it is essential to develop hardware designs with reduced impedances, build prototypes and test them in existing rings as needed.

5.2.3. Injection schemes

When the horizontal emittance of an ultimate storage ring is reduced to the 30 pm level or less it can be effectively operated in a fully coupled mode to produce a round beam (which minimizes intrabeam scattering) having a very high transverse coherent photon fraction, comparable with proposed ERLs. As discussed above, the dynamic aperture for this type of lattice will not be sufficient for off-axis beam injection and accumulation, and on-axis injection with bunch replacement (“swap-out”) will be required. There are several technical issues associated with swap-out injection that must be studied in detail, including fast kicker performance and real estate requirements. If kicker system requirements cannot be met, the fall-back scheme of accumulator rings must be explored at the same time. Candidate designs should be developed, and make detailed modeling of the

swap-out process between accumulator and storage ring. A hybrid injector consisting of a low-energy accumulator followed by a linac or booster may also be advantageous, particularly in achieving the high charge per pulse needed for long bunch trains.

For ultimate ring implementations having sufficient dynamic aperture for off-axis injection and accumulation, methods to reduce the disturbance to the stored-beam orbit during injection should be studied. One such method is to use a single pulsed multipole magnet in place of traditional closed-bump kicker magnets [18], as described below.

Fast dipole kickers for bunch replacement: Swap-out injection requires fast kicker magnets that can support the replacement of either the entire ring current or bunch trains separated by gaps. Kicker requirements are generally similar to those needed for the linear collider damping rings: kicker rise- and fall-times must be short, flat-tops must be highly constant, and pulse-to-pulse stability must be excellent. The requirements differ depending on the exact injection scheme used. For example, swap-out of the full ring would require kickers having excellent stability along a long flat-top but moderate rise- and fall-times, whereas swap out of short bunch trains would require very short rise and fall times so that the gaps between bunch trains can be minimized in order to maximize stored current. Similar kickers might also be developed for spreader designs for FELs having multiple end stations.

Pulsed multipole injection: Pulsed multipole injection kickers may have a large payoff in that they may enable transparent off-axis injection in top-up mode, with negligible induced transient to the stored beam, as well as potentially allowing injection into straight sections with small beta functions. However, many technical obstacles need to be overcome. They include the design of strong enough pulsers with fast rise- and fall-times and good shot-to-shot stability. The design of the actual pulsed magnet is equally challenging, particularly the field shaping to minimize focusing effects both for the stored and the injected beam, as well as reaching the field strength necessary for the scheme to work. Finally, depending on the actual magnet and vacuum chamber design, impedance and beam stay clear issues, as well as beam heating issues, need to be addressed. Many ideas exist, including classical multipole magnets, in-vacuum wire magnets, slotted kickers, and a magnets with flat geometries [15–17]. An overall design optimization is necessary and prototypes should be tested on test benches as well as on existing rings.

5.2.4. Instrumentation

Future ultra-low emittance storage rings will require electron and photon beam position and profile monitors capable of resolving beam displacement and size variations to a few percent of nominal beam dimensions that, at the photon source point, are on the order of $10\ \mu\text{m}$ rms in transverse size, and $1\ \mu\text{rad}$ rms in transverse divergence, $10\ \text{ps}$ rms in length, and 0.1% rms in energy spread. While this performance is already being achieved at many present-day light sources that have reduced the vertical beam emittance to diffraction-limited values, it is only reached within limited bandwidths and for relatively constant bunch fill patterns, bunch intensities and insertion device gap settings. With respect to bandwidth, the value of high resolution beam monitors having much higher bandwidth, up to the bunch frequency ($\approx 500\ \text{MHz}$), is already recognized for carrying out studies in transverse and longitudinal beam dynamics and nonlinear lattice phenomena. This instrumentation capability will be even more valuable for future rings having near-diffraction-limited emittances in both planes, highly nonlinear lattices, flexible fill patterns, short electron beam lifetimes, and rapidly changing insertion device parameters.

Electron BPMs: Present-day electron beam position monitor (BPM) systems for storage rings have sub-micron resolution in a bandwidth of order $1\ \text{kHz}$ for beam currents of a few mA or more and are used to stabilize beam orbit to 10% or less in that bandwidth. The most modern BPM systems are capable of detecting the position of first-turn electron bunches having a few tens of pico-coulomb charges with millimeter resolution, and turn-by-turn orbit position (averaged over many or all of the electron bunches in the ring) with a resolution of order $10\ \mu\text{m}$. Future ultra-low emittance storage rings would benefit with an order of magnitude increase in resolution for first-turn and turn-turn orbit measurements, as well as with the capability to measure bunch-by-bunch position on each turn with micron resolution for the purposes of measuring and stabilizing transverse bunch dynamics. BPM resolution and stability on the order of $100\ \text{nm}$ or better will continue to be required to stabilize beam orbit in a kilohertz bandwidth. Long-term stability will likely necessitate careful mitigation of mechanical motion of BPM electrodes and vacuum chamber assemblies, including using mechanical motion monitors to help stabilize long-term drift as is done presently in some light sources. Quadrupole and sextupole modulation systems will be needed to establish BPM electrical centers with respect to quadrupole and sextupole magnetic centers, the latter being critically important due to the restricted dynamic aperture of ultimate ring lattices. Sensitivity to bunch fill pattern and current, which plagues present-day BPM processing system, must be reduced to the order of a percent of transverse beam dimensions to accommodate the flexible fill patterns and rapidly decaying bunch currents (few-hour lifetimes) expected for future ultimate rings. Today there is only one type of advanced electron BPM electronics that is commercially available, but this system which has several limitations that would make it insufficient for use on ultimate rings. The development of a next generation digital electron beam position monitor is therefore a high priority.

X-ray BPMs: Future ultra-low emittance storage rings are likely to have long beam lines, of order $100\ \text{m}$ or more. X-ray BPMs (XBPMs), possibly two or more, will be needed to improve photon beam stability before and after optical components in each of these long beam lines. While existing XBPMs work very well for broad spectrum, fixed gap wigglers and bending magnet beam lines, they are less successful when used for variable gap undulators, and particularly elliptically polarized undulators (EPUs). XBPMs for wigglers and dipoles work well because their radiation fans are wide horizontally, allowing an unused portion of the fan to be dedicated to a fully intercepting monitor that only needs to operate for the vertical plane. In contrast, XBPMs for the much more horizontally collimated radiation from undulators, with photon spectra concentrated in fundamental and harmonic wavelengths, typically intercept only the fringes of the radiation pattern, allowing the central core of the beam to pass unimpeded. The beam line radiation fringe pattern contains not only the radiation from the undulator but also the radiation from upstream and downstream dipole magnets. As the undulator gap is changed, not only will its pattern and wavelength content change, especially for EPUs, but also the relative mix of undulator and dipole radiation will change. Both effects can cause changes in the intercepted fringe pattern and effective centroid, potentially resulting in an apparent and erroneous change in detected core-beam position. The response of the XBPM to changing undulator wavelength can also contribute to this erroneously detected position. While some of these problems have been mitigated using look-up tables and other ‘smart BPM’ methods to partially correct for gap-dependent errors [51] for planar IDs, the EPU XBPM problem has not yet been solved. Continued XBPM development is needed to improve performance and some breakthrough technology will be needed for EPUs. Finally, the very small horizontal emittance and high horizontal collimation of the

undulator beam implies that XBPMs will need to operate in both transverse planes. As for electron BPMs, mechanical motion sensors and mitigation methods will most likely be needed for XBPMs to maintain precise photon beam pointing stability over long beam line distances. As is also the case for electron BPMs, high bandwidth XBPMs will be valuable for detecting high frequency beam dynamics.

High resolution profile monitors: Methods developed to measure transverse diffraction-limited X-ray beam emittances (of order 10 pm rad or even less) include using X-ray pinholes, visible light interferometers, Fresnel zone plate lenses, and a method that measures the vertically polarized profile of dipole radiation [52]. All of these methods can reach a resolution of a few microns, sufficient for measuring diffraction-limited beam sizes as long as the radiation source points are at lattice locations having sufficiently high beta functions (typically ≥ 10 m). Micron and less transverse electron beam size can also be measured using a scanning laser wire [53], but the method is slow compared with optical techniques. As is the case with electron and photon BPMs, high bandwidth monitors, capable of detecting bunch-by-bunch beam profile changes, will be valuable for detecting beam dynamical phenomena.

5.2.5. Stability design and stabilizing systems

Stability design requires a multi-disciplinary approach, which starts from the site selection and characterization, proceeds through the design of the foundation and buildings, vibration minimization of conventional systems, mechanical design, power supplies, and beam diagnostics, and finally ends with feedback systems to stabilize the beam. All of those areas have shown continuous improvements over the years and R&D is beneficial to make further progress in all of them. Of particular interest are low amplification mechanical (e.g. girder) designs, progress in beam position monitor systems, faster orbit feedback electronics and improved photon beam diagnostics. As discussed above, there are no good solutions as yet for XBPMs serving the increasing proliferate EPU and electron BPM processing must be improved for ultimate ring implementations. One key technology for improved beam position monitors and faster orbit feedbacks are field programmable gate arrays (FPGAs) and real-time communication systems. Excellent FPGA expertise exists in several national laboratories, however, a collaborative, focused effort would hold significant potential for stability improvements.

An important point is the fact that correction of insertion device effects is not limited to orbit feed-forward of feedback systems, but may need to include beam size and optics correction as well. Systems to accomplish that have evolved significantly over time and nowadays often involved dozens of quadrupoles as well as several skew quadrupoles for a single undulator. However, future lattice may be more sensitive and the smaller emittances may require even better correction than is possible with existing systems. This might require active consideration of magnet hysteresis. It is also possible that sextupole or higher multipole magnets might need to be included in the compensation schemes to minimize effects on the already small dynamic and momentum aperture.

Finally, an important step in stability design is to optimize the whole system of accelerator, photon beamline and endstation simultaneously, which all too often is not being accomplished. If done correctly, it will allow ultimate stability performance, but will require much closer collaboration between beamline designers, accelerator physicists and people designing experiments and endstations. Such a front-to-end design optimized for stability could also profit from better simulation tools linking accelerator, photon optics and end station simulations together.

Control system and data rates: Device control and readback data rates need to be sufficient to achieve the desired component performance, ranging from rates on the order of 1 Hz for slow devices (e.g. temperature and pressure monitors) to many kHz (e.g. for hundreds of digitally controlled orbit feedback power supplies and BPM processors), and MHz for wideband devices. The technology that should be used to achieve these very high data rates in a distributed (and deterministic) network is an active area of R&D and further development is very desirable.

Orbit and beam line feedback design: Advanced orbit and beam line feedback systems will be needed to achieve the requisite level of electron and photon beam stability ($< 10\%$ of photon beam dimensions) in any future light source. An integrated effort from the accelerator and beam line designers will be needed to maintain stability integrity in all aspects of hardware and control system design. It is likely that high-resolution (≈ 100 nm or better) mechanical motion/position survey sensors will be needed for critical components in the accelerator (e.g. user BPMs) and beam line (e.g. optical components, small apertures and collimators, etc). Some of these devices may require cutting-edge technology (e.g. ‘telescope technology’ such as the laser-Doppler stabilization system used for atomic force microscopes and the X-ray Nanoprobe at the APS). Maintaining the beam pointing and position stability at user experimental stations located > 100 m from the photon source is an engineering challenge.

Fast (multibunch) feedback systems: High frequency electron bunch motion, driven by accelerator transverse and longitudinal impedances will be controlled with bunch-bunch feedback systems. Feedback kicker bandwidth must be at least half of the chosen rf frequency to affect all bunches. Longitudinal instability caused by rf voltage phase and amplitude noise, including that caused by ripple in the high voltage power supply at harmonics of 60 Hz (or higher frequencies for switched mode supplies), must be controlled with a combination of the low-level rf and longitudinal multibunch feedback systems. In general, those systems will be more difficult in ultimate storage rings compared to current rings, because of generally lower instability thresholds, potentially faster instability growth rates, as well as the requirement to not dilute the smaller emittances, requiring smaller noise from the feedback systems. Another area, where progress would be desirable would be transverse feedback systems that could simultaneously suppress multibunch instabilities, as well as the transverse mode coupling instability. Finally, harmonic bunch lengthening cavities can complicate fast rf and longitudinal feedback system performance, requiring further development of those systems.

5.2.6. Radio frequency system

The choices that must be made for the rf systems for an ultimate ring include the rf frequency, the type of cavity, and the technology of the high power transmitters. In all cases the development time for a totally new high-power rf system is so long that most it is likely the design choices will be based on technology that exists at the time. In this respect, it is important to continue evolutionary R&D of rf systems so that the resulting innovations will be the basis upon which future light sources will be built.

RF frequency: The large majority of existing light sources have selected a radio frequency between 350 and 500 MHz. As a result, any new project selecting the same frequency benefits from existing knowledge, considerably reducing system development time and the risk. Also, the availability of rf cavity designs and transmitters are often dominant in the final choice of the rf frequency. A notable exception to the trend of using 350–500-MHz, MAX IV has chosen 100 MHz [54] in order to have a longer

bunch with higher instability thresholds, as discussed earlier. The penalty for this choice is the inability of producing high current, short bunches. An Instead of using such a low-frequency system, it might also be possible to use a two-frequency system, e.g. 450 and 500 MHz. Such a solution might allow the operation with longer bunches, which is desirable from an instability standpoint, but still retain some of the advantages of higher frequency systems (e.g., compactness and high power).

RF cavities: Once the rf frequency has been selected, a choice must be made concerning the type of cavity (warm or superconducting) and whether prototyping and testing of the selected rf cavity must be started. Besides the optimization of the shunt impedance for the main accelerating mode, one of the most crucial issues is the shunt impedance of the cavity higher order modes (HOMs), which may drive coupled bunch mode instabilities. The 100-MHz rf cavities mentioned above are warm or capacitively loaded designs, which pushes up the frequencies of the HOMs where their influence is diminished due to a reduced beam longitudinal form-factor. For the 500 MHz regime, two directions have been followed. One is the development of superconducting cavities having no HOMs. Such superconducting rf cavities are technologically complex and have a long development time. As a result, after many years there exist only a few designs world-wide which are considered mature and ready to be produced by industry. The other direction is to use normal-conducting cavities having several HOM dampers. The development time of such cavities is usually much shorter and less expensive, but their effectiveness in reducing the HOM shunt impedance tends to be not as good as superconducting cavities. Also the operating cost of normal conducting cavities tends to be higher due to the power deposited in the cavity walls. An important issue for accelerators with high beam currents and gaps in the fill pattern (for ion clearing, injection, ...) is the transient effect due to beam-loading. This effect can be minimized with optimized designs or in the case of normal conducting cavities by the addition of energy storage cavities.

RF transmitters: The next important issue is the choice of technology for the rf transmitter. High-power klystron transmitters have been used for many years and are considered mature. However, nowadays industry only produces such high-power klystrons in limited quantities, and only for particle accelerators. As the market is small, the number of industrial suppliers has been reduced world-wide and only limited R&D effort has been devoted to improving the technology. Recent projects have used an alternative which combines several moderate-power inductive output tubes (IOTs). IOTs are mass-produced by a variety of vendors for the UHF television broadcasting market. Another promising alternative that warrants further development is the solid state amplifier which has become a viable option only recently with the availability of high-power transistors in the few-hundred megahertz frequency range [55].

5.2.7. Insertion device developments

As discussed above, the technology of insertion devices tailored to the specific needs of X-ray science applications has been the subject of a tremendous development over the past 20 years. Several directions for further development have been identified which can significantly benefit the performance as well as reduce the cost of insertion devices for use on an ultimate storage ring source as well on existing sources.

The use of a short undulator period allows the production of hard X-rays from a reduced electron energy, saving on the infrastructure cost. To maintain a reasonable magnetic field, short period permanent magnet undulators are only possible with a small magnetic gap. How small the gap can be is set by the

internal aperture required for the electron beam in order to maintain an acceptable lifetime and to minimize the local electron losses. Another determining factor can be the impedance of the narrow vacuum chamber and the transitions at the end, which increases dramatically with a smaller gap. In-vacuum, permanent-magnet (typically NdFeB) undulators are the most efficient technology presently in use, initiated on a large scale at SPring-8 [56] and now in use all over the world. Further increase in the magnetic field (and therefore the potential for shrinking the period) is expected by cooling the permanent magnets to cryogenic temperatures [57]. Such a technology must be further developed and tested on existing light sources. An alternative to the in-vacuum permanent magnet undulator is the superconducting undulator [58] for which significant R&D and testing is needed. Such devices promise larger gaps for the same performance as the permanent-magnet types and, with the use of Nb₃Sn, they promise magnetic fields well beyond the reach of cryogenic permanent-magnet undulators.

Other exotic types of insertion devices have been identified and require R&D. These include fast-switching polarizing undulators which are likely to require a combination of permanent magnet and electromagnets. Another interesting development that reduces the heat load at the lower energy tuning of the fundamental is an undulator that is tuned by varying the period rather magnetic field strength, a feature that is possible using superconducting technology [59] and other means. A device that may be of interest for pulsed applications, such as pump-probe experiments, is the rf undulator, which uses a travelling wave mode in a long cavity structure to undulate the electron beam [60]. Such a device can be switched in polarity very quickly (potentially at multi-kilohertz rates).

In order to fully utilize advanced undulator parameters in the design of new facilities, it is important that any new insertion device development be carried out early, before the actual project starts in earnest, since it might directly impact the optimization of beam energy, lattice and other accelerator parameters, and therefore the scale of the facility.

5.2.8. High heat load photon optics design

The preservation of photon beam emittance, coherence and stability in the presence of the potentially unprecedented photon beam power generated by an ultimate ring light source poses a significant challenge for beam line and photon optics design. For example, insertion devices on high-current ultimate rings such as PEP-X [5] generate on the order of 75 kW of radiated power with a peak angular power density of $\sim 1 \text{ MW/mrad}^2$. While it is possible to use present-day beam optics technology by reducing photon power densities and levels to manageable levels ($\sim 10 \text{ MW/mm}^2$ or less at the first mask and $\leq 0.5 \text{ kW}$ on LN-cooled monochromators)—using long drift lengths, aggressive aperturing, and filtering—the current technology cannot fully preserve the extraordinary low emittance properties of future sources even at low power. Future development of X-ray optical components that can preserve emittance as well as handle higher power densities to reduce beam line length and increase flexibility in beam line configuration will be critical for exploiting future high current ultimate ring sources.

Front-end mirrors: The first optical component in a beam line beyond upstream masks and apertures may be a flat coated mirror at a few milliradian angle of incidence that provides power-filtering for downstream components by suppressing reflection of photon energies higher than the mirror cut-off energy (e.g. 23 keV for a Rh-coated mirror operating at a 2.7 mrad angle of incidence). In addition to power filtering, such a

mirror can perform an essential radiation shielding function by physically separating the synchrotron radiation from the forward-directed gas Bremsstrahlung thus facilitating gas Bremsstrahlung termination inside the storage ring shielding. More generally, a variable cutoff energy mirror system can be devised using two anti-parallel mirrors with variable angle of incidence. Such mirrors may absorb a few hundred watts in a high-current ring, with a power density on the order of 0.5 W/mm², causing thermal deformation of a water-cooled mirror approaching 1 μ radrms, a significant fraction of 10- μ rad-level horizontal photon beam divergence, and of the same order as the vertical beam divergence, thus degrading emittance and coherence. The development of cryogenically cooled mirrors and, for correcting long wavelength figure error, adaptive mirror technology would address this issue. The long optical lever arm associated with mirrors located many tens of meters from the beam source imposes mirror stability requirements that will necessitate active mirror pointing feedback.

Monochromators: Even with aggressive aperturing and power filtering, the power transmitted to a monochromator located a reasonable distance from the first optical element on a high-current ring will be on the order of 100–150 W, and the power density will be of order 10–15 W/mm². Finite element analysis of the response of a typical LN-cooled monochromator to this power load indicates a Si crystal surface temperatures on the order of 100 K and thermal deformation of a few μ radrms, again causing emittance degradation. Possible fruitful areas for R&D to address this issue include improved crystal materials (e.g., isotopically pure diamond) and enhanced cryogenic cooling of silicon whereby the crystal is maintained in a more isothermal state at a temperature closer to that of zero thermal expansion (i.e., \sim 130 K). The later approach could involve alternative cryogens to LN and/or thermal load feedback loops to maintain the diffracting volume in a low net strain state. Ultra-high energy resolution post-monochromators involve exotic crystal geometries which push crystal fabrication and monochromator design state of the art.

Downstream optics: While beam power is no longer an engineering issue downstream of the monochromator, the performance of present-day focusing technology, including Kirkpatrick–Baez (KB) mirrors, compound refractive lenses and zone plates, is not sufficient to preserve the ultra-low emittance of future ultimate ring light sources. Even today, focusing optical elements of various types are the subject of intense development work as manifest by the rapidly evolving state of the art. For example, reasonably state-of-the-art mirrors can attain slope errors in the 0.25- μ radrms range (e.g., the hard X-ray offset mirror system used for the LCLS FEL). The point-spread function for a mirror system of this quality in a typical application degrades the delivered beam emittance approximately two-fold. Moreover mirror surface height variations spoil beam coherence. Applying the Maréchal Criterion for preserving optical wavefronts limits the acceptable wavefront distortion to $\lambda/14$, which, for a two-mirror system, constrains the mirror surface long-wavelength height variation to $\lambda/(28\alpha\sqrt{2})$ rms, where α is the beam's incident angle on the mirror surface. Assuming 10 keV radiation and 2.7 mrad incident angle yields a mirror surface height control of 12 Årms, which it at or just beyond current mirror fabrication state of the art for reasonable mirror dimensions. Similarly fabrication limitations on zone plate line widths dictate the ultimate focus achieved, fabrication technology constrains the useful acceptance of refraction optics, etc.

Optics and related systems R&D: Given the afore-mentioned emittance and phase distortion effects with current beam line technology, it is appropriate to list areas where technological

improvements could deliver important beam line cost and/or performance advantages:

- Improved thermal designs could reduce masking costs and provide more beam line layout flexibility.
- Improved mirror cooling technologies could reduce emittance degradation and improve beam stability.
- Improved mirror polish/figures would reduce emittance and coherence degradation.
- Advanced beam position and shape monitors would enhance beam stability when incorporated into feedback systems.
- Reduced thermal deformation of monochromator crystals through cooling improvements and/or alternative crystals would reduce emittance and coherence degradation as well as improve beam stability. Substantial improvements in power management could eliminate the need for power filtering mirrors. Though not explicitly discussed above, grating monochromators for VUV and soft X-ray beam lines would derive similar benefits from improved thermal performance.
- Advances in micro-focusing optics, such as smaller zone plate line widths, would enhance microscope resolution.
- Improvements in optics support and experimental hall floor stability would reduce beam instability.
- Improved sample manipulation systems providing finer resolution and increased stability would enhance microscope performance.

5.3. Additional capabilities

5.3.1. RF deflecting cavities

Superconducting deflecting cavities are being considered for existing light sources as a means of producing short X-ray pulses [61]. Several cavity design options are being considered including a single-cell [62] and a number of multi-cell structures [63] (see Fig. 7).

The R&D needed to demonstrate technical feasibility for this approach includes: cavity and cryomodule design including couplers and dampers, low-level rf characterization, testing in a storage ring, and X-ray optics development.

5.3.2. Tailored bunch operation

Tailored bunch operation can take several forms—changing the orbit, energy, or pulse length of individual bunches. In terms of changing the orbit of a single bunch, development of fast dipole kickers—allowing the possibility of kicking a single bunch with such amplitude as to put it on a different closed orbit that is sufficiently separated from the remaining bunches. In addition the development of programmable fast coupled bunch digital transverse feedback would allow the possibility of feeding back on different bunches in different ways.

One could also imagine changing the energy of a single bunch relative the energy of the other bunches. This could be done by adjusting the functional dependence of the path length versus energy. For instance if the dependence has two zero crossings there will be two fixed points. Each of the fixed points will be at different energies. This is illustrated in Fig. 8.

In the figure, the slope of the curve at the two fixed points is different indicating that the momentum compaction factors are different and thus the bunch lengths are different. One can express this curve as a power series

$$\frac{\Delta T}{T} = \alpha_1 \delta + \alpha_2 \delta^2 + \alpha_3 \delta^3 + \alpha_4 \delta^4 + \dots, \quad (33)$$

where α_1 , α_2 , α_3 , and α_4 , are the first, second, third, and fourth order momentum compaction factors. To achieve this in practice requires the development of lattices that have the feasibility of

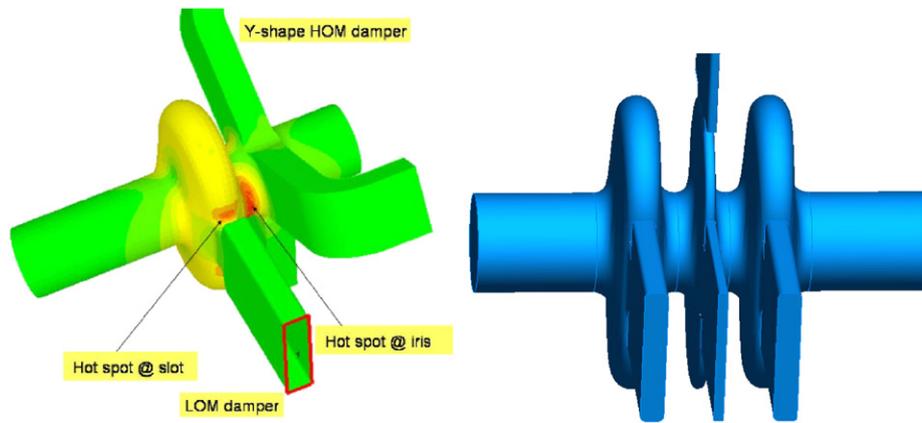


Fig. 7. Left: A single-cell cavity with waveguide damping. Right: A 3-cell cavity with waveguide damping.

controlling these terms as well as ways of feeding back on the terms to keep them stable. It may be possible to attempt this in existing storage rings with some modifications.

5.3.3. Partial lasing

In order to evaluate the potential of partial lasing, more realistic simulations need to be carried out combining all relevant beam dynamics effects in the storage ring with the simulation of the FEL gain process. The results of such studies would then guide further R&D efforts.

5.4. Cost reduction

A large storage ring facility will be expensive to build and operate. However, we believe these costs can be reduced compared to what might be expected from extrapolation of recently built facilities. Some examples of possible cost-reducing measures that should be explored include:

- Use of combined function magnets, e.g., combined function quadrupoles and sextupoles.
- Where appropriate, use of permanent magnets, perhaps combined with low-power, low-cost electromagnetic trims.
- Beam energy optimization based on the needs of the user community.
- Choice of the number of beamlines based on a realistic assessment of the user community.
- Use of standard components (e.g., employing an RF frequency for which there are existing designs and equipment providers).
- Optimization of injection schemes.
- Use of highly optimized, high performance insertion devices to reduce the beam energy and potentially the cost.
- Optimization of beam-stay-clear and vacuum chamber impedance to allow smaller and more cost effective magnets.
- Use of distributed NEG coating or other techniques may enable cheaper vacuum systems. However, issues of surface roughness and resistive wall impedance need further study.

It should be noted that, although expensive to build, the storage ring is inherently cost effective due to the large number of beam lines available. In considering a new facility, the number of beam lines needed by the community should be considered. If the need is great, a ring providing straight sections for insertion devices around its whole circumference can be implemented. However, given the very large circumference of some ultimate rings, 30–50 insertion devices, enough to serve a user community of a few

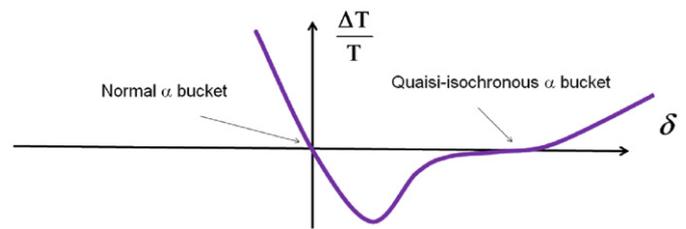


Fig. 8. Momentum compaction factors adjusted to have a large and small momentum compaction bucket.

thousand, could be accommodated and consolidated in certain regions of the ring circumference, reducing experimental hall costs and allowing other lattice cell types (e.g. low-emittance TME cells) to be used in other regions of the ring that do not require straight sections (e.g. as proposed for PEP-X [5]). DESY exploited this approach for PETRA III in order to save construction costs, creating a new DBA lattice for insertion devices in one octant of the ring and using the original lattice components in the remaining octants [10].

6. Summary

Storage ring light sources are among the most successful scientific instruments of the last 25 years. These facilities produce high-flux, high-brightness synchrotron radiation that spans a remarkably large spectral region, from far infra red to hard X-rays. These light sources have many attractive features, including high stability, excellent reliability, and the ability to simultaneously serve many users with diverse requirements.

Despite the long history of performance evolution of storage rings, further large increases in performance are possible. One can imagine an ultimate storage ring producing transversely coherent X-rays and simultaneously serving dozens of beamlines and thousands of users annually. Groups studying possible ultimate storage ring designs have identified no show-stoppers. However, focused R&D efforts for storage rings would reduce the cost and risk of a possible project and deliver a design with optimized performance. Such an ultimate storage ring would retain all the general strengths of today's storage rings while delivering high transverse coherence up to 10 keV. Some ultimate ring implementations meet and even exceed proposed ERL performance. In this paper we have outlined the underlying physics limitations and design tradeoffs, as well as a focused R&D program which will help to minimize the risks, optimize the performance, and

ultimately lead to the realization of what we believe will be an exciting 4th generation storage ring-based light source.

Acknowledgments

We wish to acknowledge the contribution from Tom Rabedeau from SLAC on high heat load photon optics. This work was supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract nos. DE-AC02-05CH11231 (LBNL), DE-AC02-06CH11357 (ANL) and DE-AC02-76SF00515 (SLAC).

References

- [1] ANL, BNL, LBNL, SLACNAL, Science and Technology of Future Light Sources—A White Paper, SLAC R-917 <<http://www-ssrl.slac.stanford.edu/aboutssrl/documents/future-x-rays-09.pdf>>, 2008.
- [2] H. Winick, The Synchrotron Radiation Sources, A Primer, World Scientific Publishing Co., Singapore, 1995.
- [3] A. Ropert, et al., in: Proceedings of the European Particle Accelerator Conference 2000, Vienna, 2000.
- [4] M. Borland, Nucl. Instr. and Meth. A 557 (2005) 230.
- [5] R. Hettel, et al., Concepts for the PEP-X light source, in: Proceedings of the Particle Accelerator Conference 2009, Vancouver, 2009.
- [6] K. Tsumaki, N. Kumagai, Nucl. Instr. and Meth. A 565 (2006) 394.
- [7] K. Tsumaki, et al., in: Proceedings of the European Particle Accelerator Conference 2006, vol. 3362, Edinburgh, 2006.
- [8] M. Borland, in: Proceedings of SRI 2009, Melbourne, 2009.
- [9] K. Kim, Opt. Eng. 34 (1995) 342–352 (Also Erratum 34(4) (1995) 1243).
- [10] K. Balewski, et al., PETRA III: a low emittance synchrotron radiation source, Technical Design Report, DESY 2004-035, 2004.
- [11] J. Ablett, et al., NSLS-II, Conceptual Design Report <<http://www.bnl.gov/nsls2/project/CDR/>>, 2006.
- [12] M. Eriksson, et al., Using multi-bend achromats in synchrotron radiation, in: Proceedings of the European Particle Accelerator Conference 2008, vol. 2007, Genova, 2008.
- [13] S. Leemann, et al., Phys. Rev. Special Topics Accel. Beams 12 (2009) 120701.
- [14] G. Hoffstaeter, Toward an energy recovery Linac X-ray source at Cornell University, Presentation for the 37th ICFA Advanced Beam Dynamics Workshop on Future Light Sources, Hamburg, 2006.
- [15] K. Harada, et al., Phys. Rev. Special Topics Accel. Beams 10 (2008) 123501.
- [16] H. Takaki, et al., Beam injection system by use of a pulsed sextupole magnet at the photon factory storage ring, in: Proceedings of the European Particle Accelerator Conference 2008, vol. 2204, Genova, 2008.
- [17] P. Kuske, et al., in: Proceedings of the European Particle Accelerator Conference 2008, vol. 2067, Genova, 2008.
- [18] D. Robin, C. Steier, L. Yang, Using novel injection schemes for enhanced storage ring performance, in: Proceedings of the Particle Accelerator Conference 2009, Vancouver, 2009.
- [19] M. Borland, in: Proceedings of the Particle Accelerator Conference 2009, Vancouver, 2009.
- [20] M. Borland, in: Proceedings of ICAP 2009, San Francisco, 2009.
- [21] K. Bane, et al., Phys. Rev. Special Topics Accel. Beams 5 (2002) 084403.
- [22] C. Steier, et al., in: Proceedings of the Particle Accelerator Conference 2001, vol. 2938, Chicago, 2001.
- [23] L. Emery, M. Borland, Possible long-term improvements to the advanced photon source, in: Proceedings of the Particle Accelerator Conference 2003, vol. 256, 2003.
- [24] A. Xiao, Intrabeam scattering calculations for the USR7 lattice, Presentation of L. Emery at AP for future Light Sources Workshop, 2009.
- [25] K. Satoh, Y. Chin, Transverse mode coupling in a bunched beam, KEK 82-2, 1982.
- [26] K. Bane, et al., Conceptual Design Report of PEP-X, 2010, SLAC Publication SLAC-PUB-13999.
- [27] C. Limborg, Ultimate brilliance of storage ring based synchrotron radiation facilities of the 3rd generation. Potential of storage ring based sources in the production of short and intense X-ray pulse, Dissertation, Université Joseph Fourier, Grenoble, France, 1996.
- [28] R. Nagaoka, Update of instability thresholds for the future machine, ESRF Note 05-02/Theory, 2002.
- [29] R. Nagaoka, et al., Beam instability observations and analysis at SOLEIL, in: Proceedings of the Particle Accelerator Conference 2007, Albuquerque, 2007.
- [30] T. Raubenheimer, F. Zimmermann, Phys. Rev. E 52 (5) (1995) 5487.
- [31] R. Hettel, Beam stability issues at light sources, in: Proceedings of the 25th ICFA Advanced Beam Dynamics Workshop, Shanghai, 2002.
- [32] J. Feikes, K. Holdack, P. Kuske, G. Wüstefeld, in: Proceedings of the European Particle Accelerator Conference 2004, vol. 1954, Lucerne, 2004.
- [33] D. Li, et al., Phys. Rev. E 48 (1993) 1638.
- [34] M. Bai, et al., Phys. Rev. Special Topics Accel. Beams 3 (2000) 064001.
- [35] G. Decker, et al., Phys. Rev. Special Topics Accel. Beams 9 (2006) 120702.
- [36] R. Schoenlein, et al., Science 274 (5285) (1996) 236.
- [37] A. Zholents, M. Zolotarev, Phys. Rev. Lett. 76 (1996) 000912.
- [38] R. Schoenlein, et al., Science 287 (5461) (2000) 2237.
- [39] A. Zholents, P. Heimann, M. Zolotarev, J. Byrd, Nucl. Instr. and Meth. A 425 (1999) 385.
- [40] W. Guo, B. Yang, C.-X. Wang, K. Harkay, M. Borland, Phys. Rev. Special Topics Accel. Beams 10 (2007) 020701.
- [41] M. Borland, Phys. Rev. Special Topics Accel. Beams 8 (7) (2005) 074001, doi:10.1103/PhysRevSTAB.8.074001.
- [42] G. Portmann, et al., Creating a pseudo single bunch at the ALS—first results, in: Proceedings of BIW 2008, Berkeley, 2008.
- [43] D. Robin, G. Portmann, F. Sannibale, W. Wan, Novel schemes for simultaneously satisfying high flux and dynamics experiments in a synchrotron light source—ALS, in: Proceedings of the European Particle Accelerator Conference 2008, Genova, 2008.
- [44] Z. Huang, K.-J. Kim, Phys. Rev. Special Topics Accel. Beams 10 (2007) 034801.
- [45] Z. Huang, et al., Nucl. Instr. and Meth. A 593 (2008) 120.
- [46] Y.-C. Chae Min, P.F. Fischer, Spectral-element discontinuous Galerkin simulations for bunched beam in accelerating structures, in: Proceedings of the Particle Accelerator Conference 2007, vol. 3435, 2007.
- [47] Y.-C. Chae Min, P.F. Fischer, Wake fields for TESLA cavity structures: spectral element discontinuous Galerkin simulations, in: Proceedings of RF Superconductivity 2007, vol. 207, 2007.
- [48] E. Gjonaj, et al., Large scale parallel wake field computations for 3D-accelerator structures with the PBCI Code, in: Proceedings of the International Computational Accelerator Conference 2006, 2006.
- [49] S. Schnepp, F. Wolfheimer, E. Gjonaj, T. Lau, T. Weiland, New J. Phys. 8 (2006) 285.
- [50] M. Palmer, et al., The conversion and operation of the cornell electron storage ring as a test accelerator (CesrTA) for damping rings research and development, in: Proceedings of the 2009 Particle Accelerator Conference, 2009.
- [51] O. Singh, G. Decker, Operational experience with X-ray beam position monitors at the advanced photon source, in: Proceedings of the Particle Accelerator Conference 2001, 2001, pp. 539–544.
- [52] Å. Andersson, V. Schlott, M. Rohrer, A. Streun, O. Chubar, Electron beam profile measurements with visible and X-ray synchrotron radiation at the Swiss light source, in: Proceedings of the European Particle Accelerator Conference 2006, vol. 1223, Edinburgh, 2006.
- [53] Y. Honda, et al., Nucl. Instr. and Meth. A 538 (2005) 100.
- [54] Å. Andersson and others, The 100 MHz RF system for MAX-II and MAX-III, in: Proceedings of the European Particle Accelerator Conference 2002, vol. 2118, Paris, 2002.
- [55] P. Marchand, et al., Operation of the SOLEIL RF systems, in: Proceedings of the Particle Accelerator Conference 2007, vol. 2050, Albuquerque, 2007.
- [56] H. Kitamura, et al., J. Synchr. Radiat. 5 (1998) 184.
- [57] T. Hara, et al., Phys. Rev. Special Topics Accel. Beams 7.
- [58] R. Rossmannith, et al., in: Proceedings of the European Particle Accelerator Conference 2002, vol. 2628, Paris, 2002.
- [59] A. Bernhard, P. Peiffer, D. Wollmann, T. Baumbach, R. Rossmannith, Superconductive insertion devices with variable period length, in: Proceedings of EPAC08, 2008, pp. 2231–2233.
- [60] S. Tantawi, RF undulators for storage rings, SLAC Internal Report, 2008.
- [61] A. Nassiri, Deflecting cavities for light sources, in: Proceedings of the ICFA Beam Dynamics Mini-Workshop on Deflecting/Crabbing Cavity Applications in Accelerators, Shanghai, China, 2008.
- [62] J. Shi, et al., Superconducting RF deflecting cavity design and prototyping for short X-ray pulse generation, in: Proceedings of the European Particle Accelerator Conference 2008, Genova, 2008.
- [63] G. Waldschmidt, private communication, 2008.



Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima

New source technologies and their impact on future light sources

B.E. Carlsten^{a,*}, E.R. Colby^b, E.H. Esarey^c, M. Hogan^b, F.X. Kärtner^d, W.S. Graves^d, W.P. Leemans^c, T. Rao^e, J.B. Rosenzweig^f, C.B. Schroeder^c, D. Sutter^g, W.E. White^b

^a Los Alamos National Laboratory, P.O. Box 1663, Los Alamos, NM 87545, USA

^b SLAC National Accelerator Center, 2575 Sand Hill Road, Menlo Park, CA 94025, USA

^c Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

^d Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02142, USA

^e Brookhaven National Laboratory, P.O. Box 5000, Upton, NY 11973, USA

^f University of California, Los Angeles, 405 Hilgard Ave., Los Angeles, CA 90095, USA

^g University of Maryland, College Park, MD 20742, USA

ARTICLE INFO

Available online 17 June 2010

Keywords:

Free-electron laser
Laser-plasma acceleration
Plasma-wakefield acceleration
Inverse-Compton scattering
High-harmonic generation
Direct-laser acceleration

ABSTRACT

Emerging technologies are critically evaluated for their feasibility in future light sources. We consider both new technologies for electron beam generation and acceleration suitable for X-ray free-electron lasers (FELs), as well as alternative photon generation technologies including the relatively mature inverse Compton scattering and laser high-harmonic generation. Laser-driven plasma wakefield acceleration is the most advanced of the novel acceleration technologies, and may be suitable to generate electron beams for X-ray FELs in a decade. We provide research recommendations to achieve the needed parameters for driving future light sources, including necessary advances in laser technology.

© 2010 Published by Elsevier B.V.

1. Introduction

Light sources are entering a new era, with the success of the linac coherent light source (LCLS) as the first free-electron laser (FEL) operating in the hard X-ray regime [1]. LCLS represents a very significant achievement, demonstrating new levels of electron beam and wiggler control. LCLS is a major facility, with standard S-band travelling-wave cavities extending over 1 km, to accelerate an electron beam up to 14.5 GeV. Facility users will compete for time at the limited number of experimental stations. Because of the cost and size of this type of machine using conventional microwave accelerator technology, where the maximum accelerating gradient is limited to less than 100 MeV/m, it is hard to imagine that more than a few will be built, and the small number will create a severe bottleneck for enabling discovery science using coherent X-rays. However, remarkable new source and acceleration technologies are now emerging that will have significant impacts on future light sources. Specifically, laser-driven plasmas [2–4] and structures [5], as well as electron-beam driven plasmas [6,7], have the promise to provide compact and cheap generation and acceleration of electron bunches that can provide a new paradigm of X-ray light sources in universities

and small laboratories worldwide as well as significantly cheaper national-level high average flux hard X-ray sources.

These emerging acceleration technologies, along with alternative photon-generation techniques, are reviewed in this paper. A short summary of the physics behind laser-driven plasma wakefield acceleration, electron-beam driven plasma wakefield acceleration and laser-driven structure acceleration is provided. The technology challenges and overall technology readiness is also discussed for these beam generation and acceleration technologies. A comparison of them is made in the conclusion, considering 5-year, 10-year, and greater than 10-year horizons.

As alternatives to X-ray generation through the FEL interaction, inverse Compton scattering (ICS) [8,9] and laser high-harmonic generation (HHG) [10,11] are also considered. Although betatron motion by electron beams injected into plasma [12] or formed in laser-driven plasma [13] can be used as hard X-ray sources these concepts are not separately reviewed, but rather included in the discussions on laser-driven plasma and electron-beam driven plasma technologies. It is important to note that some of these sources are coherent (FEL and HHG) and others are not (ICS and betatron emission). As with the electron beam generation and acceleration technologies, a short summary of the physics behind these technologies and their technology readiness is included, and recommendations for R&D for these technologies are presented. Conventional laser technology is also reviewed, particularly considering advances required for applications for FEL technologies and the emerging electron source technologies. Areas are identified where

* Corresponding author. Tel.: +1 505 667 5657; fax: +1 505 667 8207.
E-mail address: bcarlsten@lanl.gov (B.E. Carlsten).

laser research can lead to immediate increases in performances. Finally, research recommendations are included that would advance these technologies the most in areas critical for future light sources.

As these emerging technologies mature, they each can fill a unique, and very useful, X-ray source niche. These niches vary from relatively inexpensive high-flux incoherent X-ray sources capable of servicing multiple users to ultra-compact, low-average flux coherent university-scale sources, as well as significantly reducing the cost of X-ray FELs by replacing most of the rf accelerator by advanced accelerator technology for both high- and low-average fluxes.

2. Emerging electron beam source and acceleration technologies

2.1. Laser-driven plasma wakefield acceleration (LPA)

The technology of accelerators that power today's light sources such as the LCLS was developed many decades ago, with presently only incremental improvements. As the LCLS becomes operational, it will provide scientists with unprecedented capabilities to answer key scientific questions. The LCLS and other existing (or under construction) machines will determine the performance requirements for the next generation of light sources, which will be enabled by advanced accelerator R&D. Technologies must be developed to make accelerators more compact and economical to address the future needs of users.

Laser-plasma accelerators (LPAs) [3,4] have demonstrated accelerating gradients on the order of 10–100 GV/m, three orders of magnitude beyond conventional accelerators, and have produced high quality electron bunches at the 1 GeV level [3], having normalized emittances of a few mm-mrad with acceptably high charge, > 10 pC. This makes LPAs promising candidates for the next generation of compact accelerators. In addition, the wavelength of the accelerating field (the plasma wavelength) in an LPA is unprecedentedly short, on the order of 10–100 μm . Hence, LPAs are intrinsic sources of ultrashort (fs) particle bunches, with high current. If the high-brightness electron bunch from an LPA is used to drive a radiation source, such as an X-ray FEL, then this facility could deliver synchronized pulses of fs radiation, particles, and laser light, all from one compact machine.

2.1.1. LPA physics

In an LPA as shown in Fig. 1, an intense ($> 10^{18} \text{ W/cm}^2$), short (tens of fs) laser pulse is focused into plasma to drive a large amplitude plasma wave [2]. Such laser-driven plasma waves can sustain electric fields in excess of $E_0[\text{V/m}] \cong 96(n_0[\text{cm}^{-3}])^{1/2}$ with a wavelength $\lambda_0[\mu\text{m}] \cong 3.3 \times 10^{10}(n_0[\text{cm}^{-3}])^{-1/2}$, where

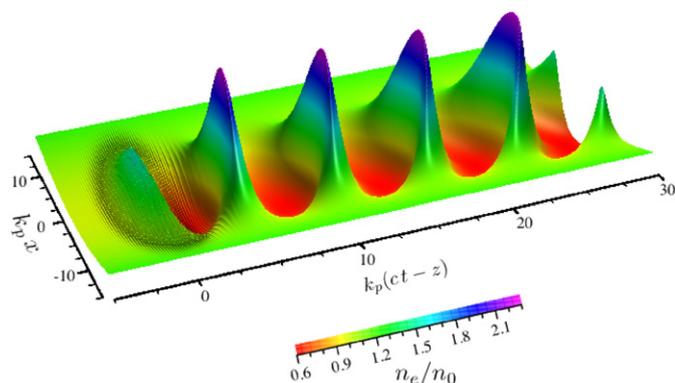


Fig. 1. Numerical simulation of a large amplitude plasma wave (plasma electron density perturbation) generated by an intense laser pulse, with electrons trapped in the wave being accelerated.

n_0 is the plasma density. For typical experimental densities $n_0 = 10^{17} - 10^{19} \text{ cm}^{-3}$, fields ranging from 30 to 300 GV/m are produced with a wavelength 10–100 μm .

In 2006, experiments at LBNL achieved a milestone with the production of high quality (few % energy spread, few mrad divergence, tens of pC) electron bunches at the 1 GeV level [14], using a 40 TW laser and a 3-cm long plasma with density near 10^{18} cm^{-3} . As the source size is on the order of a micron, the normalized emittance obtained from the plasma is, as stated above, at the mm-mrad level. As the beam pulse length is $\sim \text{fs}$, the beams obtained have tens of kA, and the beam brightness may be estimated as $B_e \sim 10^{16} \text{ A/(m-rad)}^2$, which exceeds that of the LCLS photoinjector source by an order of magnitude.

In terms of experimental methods and technology, it is important to note that in these experiments, the electron bunch was self-trapped and accelerated from the rest of the background plasma. A plasma channel was used to guide the laser pulse (prevent diffraction) and extend the acceleration length by an order of magnitude (cm scale) compared to previous experiments [15] (mm scale), as well as lower the plasma density by an order of magnitude. The single-stage energy gain in an LPA scales inversely with plasma density. Hence, plasma channel technology is crucial to producing the long and relatively low density plasmas needed to obtain GeV energy gains and beyond.

LPA beams, with their inherently high brightness and ultrafast pulse structure, are well suited for radiation generation across the electromagnetic spectrum. Single-cycle multi-THz radiation can be generated via coherent transition radiation using a foil or a plasma-vacuum transition. Incoherent, broadband hard X-rays can be generated by betatron (synchrotron) radiation in plasma (as many as $10^8 - 10^9$ photons over 1 keV per shot with commercial 30 TW, 10 Hz lasers [16]). Directed beams of gamma-rays can be generated by Thomson scattering using a counter-propagating laser pulse. In addition, using the same laser technologies as needed for LPAs, coherent VUV and soft X-ray radiation can be generated directly by the laser interacting with a gas and/or plasma (e.g. high harmonic generation (HHG)). This technique is now employed in seeding short wavelength FELs [17]; in fact, the high brightness bunches from an LPA are well suited for producing high peak brightness coherent X-ray radiation via the FEL mechanism, provided the bunch quality is sufficiently high [18]. Since LPAs may have as yet unprecedented brightness, the FEL gain length can be significantly reduced, and, hence, LPA electron beams are not only generated from a compact accelerator, but can use a shorter undulator length to achieve saturation. Further compactness (reduced saturation length) can be achieved by seeding the FEL using the above-mentioned HHG radiation from the drive laser. For example, a 0.5 GeV LPA generated electron beam can be used to coherently amplify a HHG seed in a FEL, generating $> 10^{13}$ photons/pulse at 30 nm in less than 3 m using a conventional undulator [19].

For driving light sources, key benefits of LPAs include compactness (i.e., a few to tens of cm plasma channel driven by a laser occupying a small laboratory-room footprint), relative low cost, intrinsic synchronization between the drive laser pulses, electron bunches and radiation pulses, and the ultrashort (fs) duration of the electron and radiation pulses. An LPA could be the basis for a hyperspectral source, spanning THz to gamma-ray regimes that would be ideal for pump-probe experiments and enable new applications in many fields of science.

2.1.2. LPA performance challenges and overall readiness

The laser intensity required to drive an LPA is on the order of $10^{18} - 10^{19} \text{ W/cm}^2$, which is now routine with solid state lasers using the technique of chirped-pulse amplification, initial devel-

oped during the 1990s. The repetition rate of high peak power 0.1–1 PW laser systems is currently limited to about 1–10 Hz, implying an average power on the order of 10–100 W. Development of higher average power laser systems is crucial to many applications of LPAs.

Several groups around the world, including the Berkeley Lab Laser Accelerator (BELLA) Project at LBNL, have plans to explore LPA-physics issues using petawatt (PW) laser systems with up to 10 Hz repetition rates. Spurring on the worldwide LPA progress is the fact that commercial companies (most notably in France) have developed the expertise to build sophisticated multi-100 TW and even PW-class systems, a feat that previously has been accomplished in only a few select institutions in the world. These laser systems are compact, occupying only a few m² in area. With PW-class short-pulse systems in operation (e.g., BELLA), it is anticipated that 10 GeV-class beams will be generated in meter-scale LPAs. These high-brightness beams have the potential to drive an FEL for coherent hard X-ray generation, with greatly reduced undulator length owing to the high peak currents (tens of kAs) produced by LPAs, as further progress is made reducing the emittance and energy spread of these beams.

Within the next 5 years, LPA electron beams at 10 GeV should be demonstrated from a 10-cm plasma channel driven by a PW laser system (now commercially available at 1 Hz rep-rate) with the required beam quality to drive an FEL in the 1–10 nm range. Techniques for improving the bunch quality will be explored, such as tailoring the profiles of the laser pulses and the plasma channels, as well as laser-based injection techniques for initiating the electron bunches from the background plasma, such as the use of density down ramps and colliding laser pulses. Techniques for producing long (10 cm or greater) plasma channels with adjustable density profiles will be developed. Time-resolved diagnostics need to be developed and implemented with the ability to measure such properties as the slice emittance with fs accuracy.

Initial experiments on producing undulator radiation with an LPA electron beam are underway [19]. To proceed to coherent radiation production via the FEL mechanism, demonstrating an electron beam energy spread on the order of 0.1% (presently at the 1% level) is a necessary step in the development of an LPA-driven FEL and is likely to be achieved in the next few years. Initial experiments on the staging of two or more LPA modules are underway. This includes development of compact methods (e.g., plasma mirrors) of laser in-coupling to the plasma structure. The in-coupling methods must be compact, i.e., less than the length of plasma structure, so as not to significantly degrade the overall acceleration gradient. Radiation generation driven by LPAs would also benefit from research in short-period, compact undulator technology. Understanding the details of the LPA physics requires the continued development of analytical and numerical modeling, in particular, efficient methods for simulating LPAs on massively parallel computers. Laser technology needs to continue its rapid pace of development to allow higher average power systems with improved efficiency.

2.2. Electron-beam driven plasma wakefield acceleration (PWFA)

Demonstration of the PWFA concept was first achieved at ANL in the late 1980s with experiments that mapped the plasma wakefield produced by a drive electron beam by measuring the energy gain (< 1 MeV) of a time-delayed witness electron beam [20]. Since that time progress for beam-driven plasma accelerators has been remarkable with the maximum energy gained in the plasma exceeding 40 GeV in 2007 [7], enabled by the capability of the SLAC linac to deliver high intensity bunches.

2.2.1. PWFA physics

As shown in Fig. 2, a high-amplitude oscillation can be set up in plasma by the wakefield of a drive electron bunch. The oscillations are set up by the expulsion of the plasma electrons as the drive bunch traverses the plasma. The drive bunch needs to be shorter than the plasma wavelength and for symmetric bunches, the maximum accelerating field is twice the field decelerating the drive bunch. Plasma wakefield experiments have also demonstrated that high gradients, ~ 50 GeV/m, can be sustained over meter-scale distances [7]. Just as with LPAs, this gradient is roughly three-orders of magnitude greater than in standard S-band linacs. In addition, a variety of other effects has been demonstrated in the SLAC experiments, such as the generation of betatron X-rays from a few keV to tens of MeV energy, and the acceleration of electrons from the plasma itself with extremely high acceleration gradients.

In addition to reducing the size and cost of future high-energy physics machines, beam-driven plasma acceleration may also enable more compact accelerators to drive X-ray FELs. In perhaps the simplest application, when added to an existing linac, a short plasma afterburner could boost (e.g., double for a Gaussian drive bunch) the energy of the beam on the scale of a few meters and extend the wavelength reach of an accompanying FEL (e.g., a factor of four for an energy doubler).

The PWFA is an attractive technology because existing microwave accelerators can efficiently produce high current bunches well suited for driving plasma wakes with fields over 10 GeV/m. The PWFA then acts as a transformer, converting one or more high current, low energy bunches into one or more relatively low current, high energy bunches. This process can be characterized by the ratio of the peak accelerating field to the peak de-accelerating field in the plasma wake, called the transformer ratio [21]. The transformer ratio can be manipulated by tailoring the longitudinal profile of the beams driving and sampling the plasma wake. Experiments to date with Gaussian shaped bunches have operated with transformer ratios between one and two. Recent analytic and numerical models have predicted that by optimizing the longitudinal profile, transformer ratios of five may be attained [22], e.g. a 1 GeV drive bunch with several nC of charge could boost the energy of a 1 nC, 1 GeV bunch to an energy of 5 GeV on the scale of a meter. In the example studied in Ref. [23], a 5-nC, 0.56-ps, 1-GeV drive bunch is able to accelerate a 0.35-nC, 23-fsec, 1-GeV trailing bunch to 5 GeV with an energy spread of less than 1%, with an energy conversion efficiency of 35% from drive to accelerated electrons and with preserved emittance. If these calculations prove out, this technology has the potential to reduce the length of LCLS by a factor of five with the same basic microwave accelerator technology, or even more significantly, lead to a vastly more compact system if the GeV drive beam is generated in a modern X-band accelerator, like the next linear collider test accelerator (NLCTA) at SLAC [24], with an overall footprint on the same order of magnitude as in a 10-GeV LPA. These types of preliminary experimental and numerical results, plus the high average-power capability of

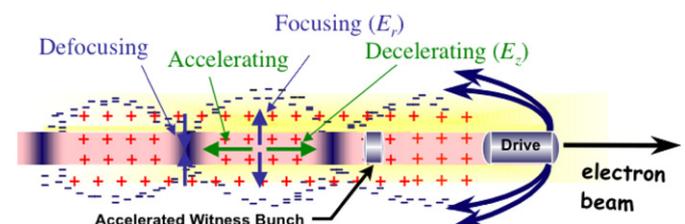


Fig. 2. PWFA schematic, indicating plasma oscillations set up by drive electron bunch expelling plasma electrons from the path of the drive bunch.

conventional rf accelerators, lead to the observation that, with sufficient R&D support, this technology may potentially drive light sources that match the anticipated high average flux of future X-ray FELs and ICS systems, more so than other novel accelerator technologies.

Plasma induced betatron radiation is also an attractive potential incoherent and broadband alternative to a traditional X-ray FEL. In an ion channel the incoming electron beam expels the plasma electrons and an ion-focusing channel is formed. The betatron motion of the individual electrons in the channel leads to synchrotron radiation, and potentially, to self-amplified spontaneous emission in an ion channel laser [25] with wavelengths that are up to 100 times shorter than that from conventional undulators (depending on plasma density and beam energy). Plasmas provide stronger radial fields and shorter periods than conventional undulators, in principle enabling an ion channel to operate at a given wavelength with a comparatively short plasma and lower energy electron beam. Additionally, the high effective wiggler strength in plasma leads to very high harmonic generation. In a first step, beam driven plasma experiments have measured spontaneous emission in the X-ray region (6.4 keV) due to the betatron motion of 30 GeV electrons in plasma of density $\sim 2 \times 10^{14} \text{ cm}^{-3}$. These results were later extended to Gamma-rays with multi-MeV energy produced in plasma with density $\sim 10^{17} \text{ cm}^{-3}$. Stimulated emission relies upon the coherent interaction between electrons in the beam and the radiation field. Numerical simulations using a combination of PIC and modified FEL algorithms are needed to study these concepts and define an experimental program.

Recent experiments operating with field ionized plasmas and multi-GeV/m gradients have measured trapped plasma electrons with multi-GeV energy, mm-mrad emittance and multi-kilo-ampere peak currents. The emittance of these trapped electrons scales inversely with the plasma density. Future experiments will continue to investigate the physics of the trapping process while optimizing the measured brightness.

2.2.2. PWFA performance challenges and overall readiness

After the initial concept demonstrations of beam driven plasma acceleration, the challenge now is to develop these techniques into useful acceleration methods for light source and high-energy physics applications. Among the issues that need to be resolved are:

- Efficient high-gradient acceleration of mono-energetic beam bunches with a narrow energy spread.
- Shaped single bunches and/or multiple bunches to increase the transformer ratio, and the energy gained per stage, including stability of the drive bunch train.
- High demagnification focusing of electrons.
- Study of emittance degradation due to matching, hosing, and ion motion for electrons.
- Study of plasma stability and heating or damage effects with a multi-bunch or high repetition rate beam.

The Facilities for Accelerator Science and Experimental Test Beams (FACET) at SLAC is being constructed to address these issues. Appropriate investment in these second generation facilities will allow the continued development of the plasma acceleration concepts that hold so much promise for future compact accelerators.

2.3. Direct laser acceleration (DLA)

Direct laser acceleration in dielectric structures is the optical extension of conventional rf structure acceleration. The

extraordinarily high electric fields in TW to PW lasers have led to concepts using laser fields in vacuum structures to accelerate electron beams, with the promise of equivalently much higher gradients. So far, there has not been much work in this field, far less than what would be required to realize these objectives. There has been a series of experiments beginning with an inverse FEL (IFEL) experiment at BNL at 10 μm and a series of experiments at Stanford and SLAC at 1 μm . The SLAC experiments are directed at achieving high accelerating gradients and high coupling efficiencies and are described below.

2.3.1. DLA physics

Acceleration was generated by crossing laser beams in vacuum at the Laser Electron Acceleration Project (LEAP), leading to a net longitudinal electric field, of the order of 1 GV/m. Although there were substantial experimental issues with electron-beam jitter and resolution of diagnostics, LEAP successfully demonstrated the interaction, with 15 keV energy modulation on the electron beam in roughly 1.5 mm of interaction [5]. The interaction gradient (10 MV/m) was limited by breakdown thresholds on the optics used to cross the laser beams. Also high-harmonic inverse free-electron laser acceleration was demonstrated [26]. The accelerating efficiency (photon-to-electron) in each case was limited to about 5×10^{-4} , with much lower wall-plug efficiency.

To address the low conversion efficiency and gradient limitations of the LEAP concept, new dielectric accelerator structures were developed that would provide high coupling efficiencies between laser energy and electron beam acceleration. SLAC Experiment E-163 accelerated ~ 1 ps FWHM electron bunches from the NLCTA using IFEL and 800 nm light from a Ti:sapphire laser, producing on the order of 100 keV rms beam energy modulation, followed by a compact chicane to produce microbunching, as shown in Fig. 3. The stable, controllable preparation of attosecond-class bunches was demonstrated [27]. These microbunches were accelerated in a staging experiment using an Inverse transition radiation (ITR) accelerator [28], demonstrating that the required 100 attosecond-class timing stability could be obtained using straightforward engineering techniques (mass, rigidity, enclosures to reduce air turbulence) and simple monitoring on the timescale of seconds to control for slow thermal drifts.

Three categories of dielectric structure are currently being examined at SLAC as candidate high-efficiency accelerators (see Fig. 4): (1) photonic band gap fibers, following the work of Lin [29], (2) photonic band gap crystals, following the work of Cowan [30,31], and (3) transmission-mode grating structures, following the work of Plettner et al. [32]. Fiber-based structures have already been produced by industry that yield synchronous TM-like modes (e.g. Crystal Fiber's HC-1060 fiber supports a family of TM-like modes in the 1–2 μm

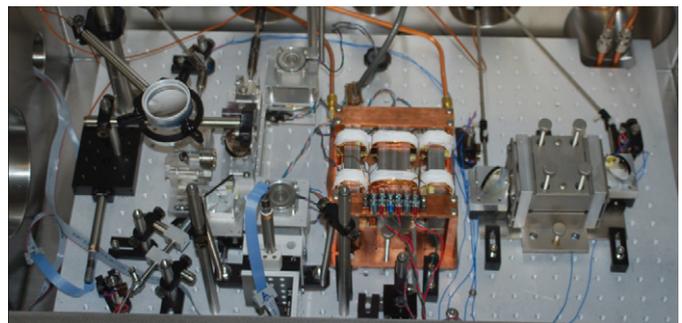


Fig. 3. Inside the primary test vacuum chamber for the E-163 experiment. The electron beam enters from the right, passing through the 3-period permanent IFEL, and the compressor chicane to form optically bunched pulse trains.

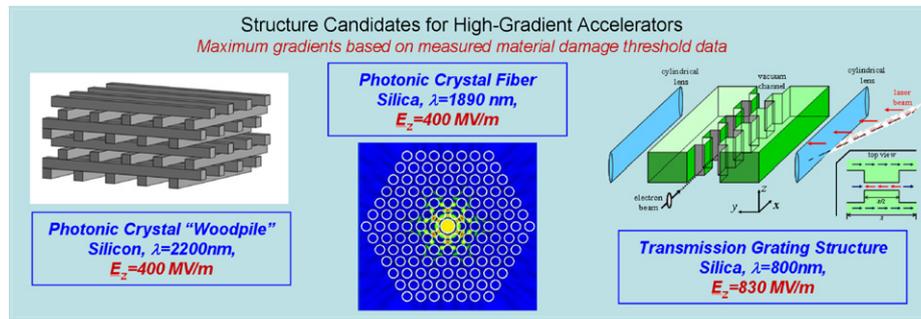


Fig. 4. Sketches showing the three types of dielectric laser-driven accelerators being investigated at Stanford and SLAC. Expected accelerating gradients are derived from measured material damage threshold data at the specified wavelength on flat substrates.

range), and tests are underway at the present time to demonstrate acceleration using these commercially produced fibers. Given the measured damage threshold of fused silica in the 1–2 μm range, accelerating gradients of 200–400 MeV/m are expected to be achievable. Photonic crystal structures called “Woodpile” structures [33] are being fabricated at the Stanford Nanofabrication Facility in silicon [34], with others under development at UCLA with experimental testing planned at the SLAC E-163 facility. These structures are attractive for their comparative geometric simplicity, potential to sustain ~ 200 –400 MeV/m accelerating gradients, ability to be mass produced by conventional semiconductor lithographic techniques, and their ability to be integrated with silicon-based lasers on a single wafer to form a fully integrated “accelerator-on-a-chip”. Transmission-grating-type structures made from fused silica are attractive for their geometric simplicity, larger aperture, and expected ability to sustain gradients of 830 MeV/m for very short pulses. Of the three types of structures, the first two (photonic fibers and crystals) are used as waveguides and hence are subject to group velocity dephasing; the third is used as a transverse phase mask and can, in principle, escape this limitation.

High gradient acceleration at high efficiencies necessarily means low bunch charge for these very high impedance structures. Conversion efficiency of laser power into electron beam energy has been shown to reach in excess of 60% with optimized selection of bunch charge (typically 10–20 fC total in a pulse train), bunch train format (typically 50–100 bunches, spaced at the optical wavelength), and laser power recirculation efficiency [35]. Structures are under study that increase the accelerating aperture to permit high charge, while still achieving reasonable power conversion efficiencies. It is important to point out that these charges are at least an order of magnitude smaller than bunch charges associated with the other technologies discussed here. Because of these low charges, we anticipate this technology is best suited for niche applications where fewer coherent X-rays are needed, in ultra-compact architectures.

These structures have also been examined from the perspective of producing ultra-compact undulators for the production of X-rays [36]. With modest geometric modifications, a synchronous deflecting field in GV/m range can be produced in the transmission-grating type structures that can be used as the basis for an undulator. Simulations for a laser-driven grating undulator for a hard X-ray FEL show generation of 120 keV photons.

2.3.2. DLA performance challenges and overall readiness

The accelerating structure technology is as yet immature, with only a few of very basic proof-of-principle experiments so far completed. Key next steps are the preparation of “long” structures

($\sim 1000\lambda$) providing significant energy gains, and providing a testbed for establishing the achievable accelerating gradients from this technology. Integration of stages into high energy-gain sections, and the development of high brightness supertip electron sources such as described in Ref. [36] are key steps to realizing a laser-driven electron source in the few GeV energy range that would be of interest to users. Since the electron pulse trains are naturally bunched at the optical scale, radiation pulses produced from these accelerated bunches will have attosecond duration.

Significant industry R&D proceeds on the core technologies used in fabricating these structures—photonic fiber production, semiconductor lithography, and on advancing the state of the art in laser performance. Yb-doped fiber laser achieving average powers well into the tens of kilowatts are available in the commercial cutting/welding market that achieves wall-plug efficiencies reaching 30%. Thulium-doped fibers operating at 2 μm are coming into the market now that offer the dual prospect for both enlarging the beam aperture (by a factor of 4 over the 1 μm structures under test) and very high wall-plug efficiencies.

Significant accomplishments include demonstration of stable attosecond-class beam bunching and diagnosis, the preliminary fabrication of silicon-based structures, and development of a facility for conducting electron–laser interaction physics in the 1–3 μm range. Major issues that remain include demonstrating dielectric damage thresholds and effects from the driving laser, structure damage from beam emittance growth and halo interception, charging of the structure by stray electrons, transport of attosecond electron beam, and X-ray pulse lengths. Present beam technologies already provide beam brightnesses that are an order of magnitude brighter than are needed for this application [37,38]. Proof-of-principle demonstrations of key concepts that lead to a totally optically driven system need to be completed. Beam quality preservation also needs to be demonstrated in this context. It is worth pointing out that this work has already led to a short-term spinoff experiment to demonstrate a new mechanism of seeding called echo-enabled harmonic generation (“EEHG”) [39]. Seeding in X-ray FELs is an important tool for reducing the wiggler length and generate longitudinal coherency, and is discussed in the accompanying review paper on X-ray FEL technology.

The horizon before this technology that can be directly applied to light sources is greater than 10 years. However, with the continued proliferation of high-power short-pulse laser systems to universities and small laboratories, there most likely will be specific facility niches (low cost, low yield, small size) that this technology will be able to address better than existing technologies.

3. Alternative photon source technologies

3.1. Inverse Compton scattering (ICS)

ICS X-ray sources show promise for delivering synchrotron-like X-ray performance at a cost and physical size appropriate to a university, hospital, or industrial lab. This is made possible by advances in electron beam performance and, especially, high average power lasers. The physics of ICS is similar to synchrotron light generation, except that the magnetostatic fields of an undulator is replaced with the electromagnetic fields of a laser pulse. The resulting factor of 10,000 reduction in magnetic period from a few cm for an undulator to $\sim 1 \mu\text{m}$ for a laser gives a factor of 100 reduction in electron beam energy to produce a photon of the same energy. For example a 30 MeV beam used for ICS can produce 15 keV photons, similar to a ~ 10 GeV electron beam using a conventional undulator. The cost of the accelerator drops by the same scale, from \$1B to \$10M, with a similar reduction in size. Furthermore, the ICS source is well suited to production of ultrashort X-ray pulses at the 100 fs scale or shorter. The X-ray output, while not coherent, is tunable in wavelength and polarization.

The basic physics of ICS X-ray generation is well understood and measured X-ray properties accurately track numerical predictions [9], at least in the linear regime. A representative ICS source [40] is shown in Fig. 5. A laser is focused to a waist onto an electron beam, producing back-scattered photons that are upshifted in energy from the initial laser wavelength by $\lambda_{\text{ICS}} = \lambda_{\text{laser}}(1 + \gamma^2 \theta^2)/4\gamma^2$, where γ is the electrons' relativistic mass factor and θ is the observation angle relative to the interaction axis.

X-rays from ICS are similar to incoherent undulator radiation since they are produced by spontaneous emission without gain, but due to the low electron energy they have a relatively wide bandwidth and opening angle since both vary with angle of order $1/\gamma$.

The challenges for development of light sources based on ICS are to tailor the electron and laser beam properties to produce the best possible X-ray beam. In general, very low emittance electron beams are required to produce both large X-ray fluxes, due to the small interaction spot sizes required to produce a sufficient number of scattered photons, as well as high spectral brightness, due to the strong dependence of the X-ray beam spectral width on the electron beam divergence at the interaction point. Additionally, relatively short pulse lengths for both the electron and laser beams are desired to maximize the interaction within the Rayleigh diffraction length of the laser beam. The optimization of high average flux, high spectral brightness inverse Compton scattering X-ray sources requires electron beams with low emittance (< 1 mm-mrad) and short pulse duration (< 1 ps), and tightly focused ($< 5 \mu\text{m}$), short pulse (< 1 ps) lasers.

3.1.1. ICS physics

Assuming Gaussian laser and electron beam profiles, an analytic expression for the total number of X-rays produced by a head-on inverse Compton scattering interaction is given by [41]

$$N_x = \frac{N_e N_\gamma \sigma_T}{2\pi(x_L^2 + x_e^2)} [\sqrt{\pi\alpha} e^{\alpha} \text{erfc}(\sqrt{\alpha})] \quad (1)$$

where

$$\alpha \equiv \frac{2(x_L^2 + x_e^2)}{c^2(\Delta t_L^2 + \Delta t_e^2) \left((x_L^2/z_0^2) + (x_e^2/\beta_e^2) \right)} \quad (2)$$

In Eqs. (1) and (2), σ_T is the total Thompson cross section, N_e is the total number of electrons, N_γ is the total number of photons in the laser beam, x_L and x_e are the laser and electron beam rms spot sizes at the interaction point, Δt_L and Δt_e are the rms pulse durations, z_0 and β_e are the laser Rayleigh range and electron beta function at interaction. The term in square brackets in Eq. (1) is a form factor that is always less than unity, and represents the degradation of the interaction efficiency for cases where the pulse durations exceed the interaction diffraction lengths of the laser and electron beams.

For most experimental parameters, the on-axis spectral width of the scattered X-ray beam will be dominated by the electron beam emittance. In this limit, the time averaged on-axis spectral brilliance, in units of photons per second per unit area per unit solid angle per 0.1% bandwidth, can be approximated by the expression.

$$B_{\text{avg}} \approx 1.5 \times 10^{-3} \frac{N_e N_\gamma \sigma_T \gamma^2}{(2\pi)^3 \varepsilon_{nx}^2 x_L^2} F \quad (3)$$

where ε_{nx} is the rms normalized electron beam emittance, γ is the electron beam Lorentz factor, and F is the interaction repetition rate. Note that due to the tradeoff between the source spot size, X-ray divergence, and flux in this limit, the brilliance is independent of the electron beam spot size.

In order for Eq. (3) to be valid, both the electron and laser pulse durations must be short enough such that $\alpha \gg 1$. Additionally, x_L must not be so small that nonlinear effects begin to degrade the scattered X-ray spectrum. It can be shown that for $\lambda = 1 \mu\text{m}$, and assuming $a_{0\text{max}} \sim 0.1$ to control the nonlinear interaction contributions to the bandwidth [42], pulse durations on the order of a picosecond and interaction spot sizes on the order of a few microns will be desired to achieve optimum X-ray beam brilliance. Eq. (3) yields the following expression for the approximate optimum X-ray brilliance in units of photons/s/ $\mu\text{m}^2/0.1\%$ b.w.:

$$B|_{\text{opt}} \approx 1.2 \times 10^6 \frac{a_{0\text{max}}^2 \sigma_e(nC) \sqrt{W_\gamma(\text{Joules})}}{\varepsilon_{nx}^2(\text{mm-mrad}) \sqrt{\lambda(\mu\text{m})}} F \quad (4)$$

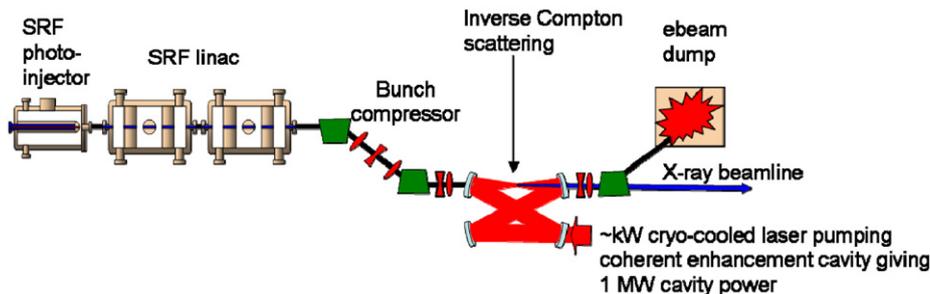


Fig. 5. Inverse Compton scattering layout showing photoinjector, short linac, high power laser, bunch compressor, and interaction point.

In order for this optimum brilliance to be approached, the electron beam pulse duration will need to be on the order of or shorter than the optimum laser pulse length. This dependence can be seen in the numerical simulation results shown in Fig. 6 for the interaction of a 25 MeV, 0.1 nC electron beam with a 0.5 ps, 10 mJ laser pulses. In particular, it is seen that there is an optimum laser spot size of a few microns, with significant degradation in the X-ray beam brilliance for electron beam pulse durations exceeding one picosecond. Note that no nonlinear effects [43] were included in this computation. Consideration of these effects limits the practical laser energy to less than a joule at 1 μm wavelength [42].

In short, it can be seen that the optimization of high average flux, high spectral brilliance inverse Compton scattering X-ray sources will require low emittance (< 1 mm-mrad), short pulse duration (~ 1 ps) electron beams and small interaction spot sizes (~ 5 μm).

3.1.2. ICS performance challenges and overall readiness

The basic ICS technology is mature, with both laboratory and industrial ICS demonstrations. Increasing the performance of an ICS source depends on generating a very low emittance electron beam (0.1–1 mm-mrad) with relatively high average current (1–100 mA), parameters that are similar to requirements for electron beams driving X-ray FELs or ERLs, although the ICS operates at much lower final energy. The development of a CW photoinjector and its associated photocathode is important. The photoinjector can be either normal conducting (NC) RF (as the 100-mA, 700-MHz one currently being commissioned at LANL [44] or the mA-class photoinjector being designed at LBNL [45]), SRF (as the 0.5-A photoinjector being developed at BNL [46]), or DC (typically 10 s of PC from the JLAB injector [47]). The cathode may be normal conducting but may need to be compatible with a SRF cavity, have high quantum efficiency, low intrinsic emittance, rapid photo-response time (fs), and long lifetime. While superconducting technology is well matched to the CW nature of this type of machine, it is not clear at this point if superconducting injectors will be able to match the beam brightness from NC injectors, which is needed. It is still an open research question if NC or SC will be the preferred injector technology of the future. While SRF technology has undergone rapid development for the large accelerators, the small-scale ICS would benefit from some changes in technical direction. The primary issues are the capital and operating costs, and size, for an SRF facility operating at 2 K. Successful development of CW, high-field SRF cavities operating at 4 K would have a significant impact on the size and cost of the ICS source. Such cavities must operate at lower frequencies

(< 500 MHz) than the conventional elliptical cavities developed for the large electron accelerators. R&D into novel SRF structures, surface studies and surface preparation of superconducting materials for CW use at high field strength, and new approaches to cryogenic design for minimum energy use should be pursued. There is a need for compact and efficient liquid Helium refrigerators with a capacity of a few hundred watts. Other key performance parameters such as stability in energy and position, and fs timing synchronization are shared with the large facilities.

The development of high average power, stable mode-locked lasers is important to ICS performance. Diode-pumped, cryogenically cooled Yb:YAG lasers show promise for scaling to kilowatt average power with the required stability and beam quality. This laser is similar to those proposed for high repetition rate HHG generation for seeding X-ray FELs. Due to the superior quantum efficiency (91%) of Yb:YAG, and the possibility of extracting the full stored energy of the crystal at liquid nitrogen temperature while maintaining the high beam quality, this source can develop 1 kW of average optical power using about 3 kW electrical wall-plug power. The high power pulse stream is then coherently coupled into an enhancement cavity with a finesse of $F=3000$, i.e. a power or energy enhancement of 1000, such that a 10 mJ, 1 ps pulse is continuously maintained in the cavity. The average power in the cavity will be 1 MW. The coupling of the electron beam into the interaction region and of the X-rays out of the interaction region is currently planned via a 1 mm diameter hole in the cavity focusing mirrors.

There are a couple of approaches being considered to generate longitudinally coherent X-rays using ICS. Improved future electron generation from integrated systems of nanostructured photocathodes and stabilized lasers may generate spatially and temporally coherent electron beams locked to optical drive lasers with attosecond precision, allowing generation of fully coherent X-ray laser light from an ICS source. Concurrently novel methods of generating quasi-crystalline electron beams [48] directly from the cathode using nanostructured field emitters and coherent optical control with attosecond timing stability are being explored. This latter research has a longer time horizon, but if successful could result in extremely compact (cm length scale) sources of fully coherent X-rays. A second coherent generation concept is shown in Fig. 7, where an ultraintense, short pulse laser drives a very short electron bunch off a nanometer foil [49,50].

3.2. Laser high-harmonic generation (HHG)

In the past decade, there has been a considerable progress in the field of HHG due to significant developments in high energy, ultrashort pulses from near IR laser systems. Commercial stable oscillators with average powers of ~ 1 W and pulse durations in the range of 10 fs as well as amplifiers with repetition rates in the range of kHz, 50 W of average power and single pulse energy up to 20 mJ and pulse duration < 25 fs can be readily purchased. At low repetition rates such as 10 Hz, 100 mJ, and J-level systems are available as well. This progress has led to commercial EUV sources, albeit with low (microwatt) output power. Single pulse energies in the order of micro joules at EUV wavelengths up to 30 eV have been obtained in several laboratories [51,52]. Photon energies up to ~ 0.5 keV and beyond have been generated [53,54] using ultrafast lasers, although the efficiency for higher photon energies are significantly lower than those achieved for the photon energy range < 100 eV. Most recently the use of MID IR driver pulses led to the generation of harmonic radiation at 300 eV with 5×10^{-8} efficiency [52]. Some of the highest efficiencies of the HHG process as a function of the harmonic energy are shown in Fig. 8.

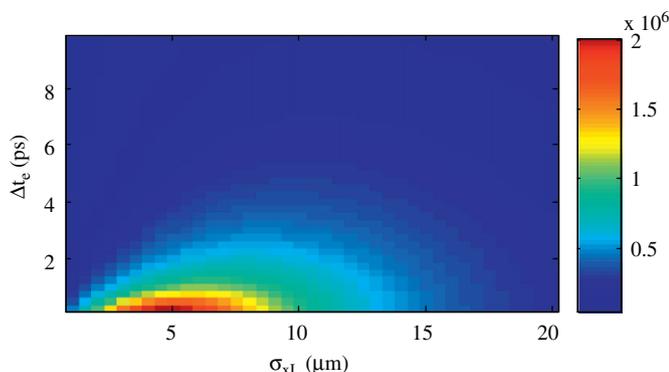


Fig. 6. Numerically calculated on-axis brilliance divided by repetition rate vs. laser spot size and electron beam pulse duration assuming $\varepsilon_{nx}=1$ μm , $x_e=10$ μm , $\Delta t_e=0.5$ ps, $\lambda=1$ μm , $Q_e=0.1$ nC, and $W_l=10$ mJ.

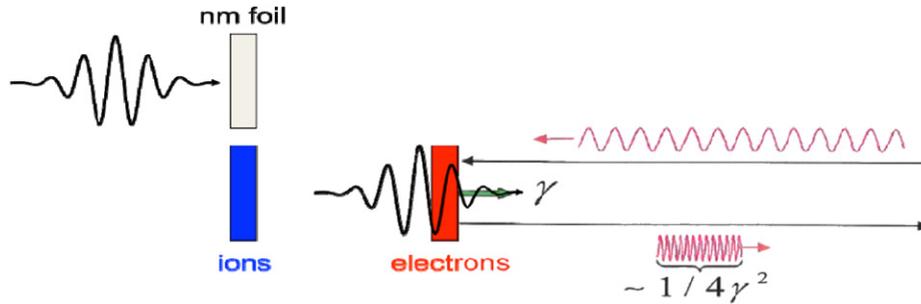


Fig. 7. Coherent scheme using an ultrarelativistic electron mirror: An ultraintense short pulse laser interacts with a nanometer foil, driving out all electrons. If the laser field is much stronger than the total restoring force exerted by the ions, the electrons can break out in a dense, ultra-short bunch, moving with a relativistic factor γ . This bunch can reflect an incoming second laser pulse and frequency upshift it by $4\gamma^2$.

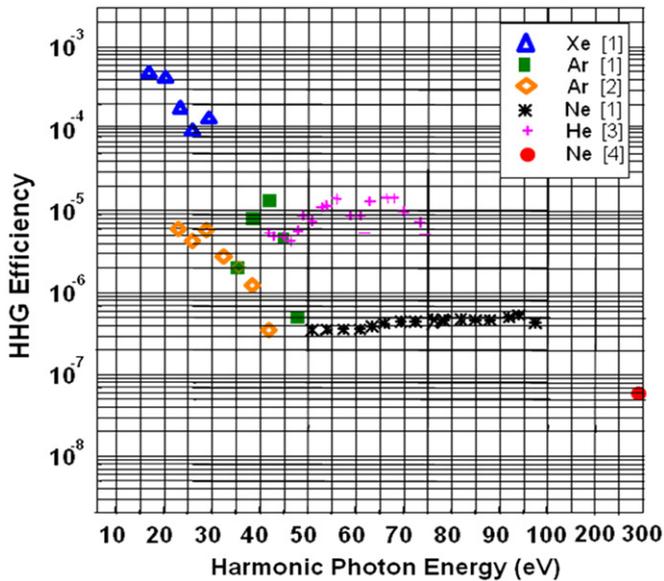


Fig. 8. Efficiencies for HHG published in Refs. [51,52,55,56].

Kim et al. [57] and Gibson et al. [58] have shown that the soft X-ray beam generated by this technique is fully coherent. The output at shorter wavelengths is limited both by the phase mismatch between the fundamental drive wavelength and the generated harmonics and the absorption of the short wavelength radiation by the nonlinear medium. Dephasing can be minimized by adjusting the pressure of the gas, using different geometries such as long gas jets, hollow capillaries, and optimizing the position of the laser focus with respect to the medium [57–60]. The necessary ultrashort pulses for HHG have been generated either directly from the oscillator and preserved in the amplifier or a longer pulse from the system has been spectrally broadened and then compressed [55] to deliver short pulses. With the latter approach the sensitivity of the beam transport to the large bandwidth of the laser beam is reduced.

3.2.1. HHG physics

The principle behind high harmonic generation (HHG) is best explained in terms of a semi-classical approach [10,11]. When an atom is exposed to the intense electromagnetic field associated with an ultrashort, intense laser pulse, its coulomb barrier can be suppressed and the atom can be tunnel-ionized. The freed electron is then accelerated by the EM field during the first optical cycle after its release, gaining energy. This energy may be released as harmonics of the fundamental laser wavelength when the electron returns to the nucleus and eventually

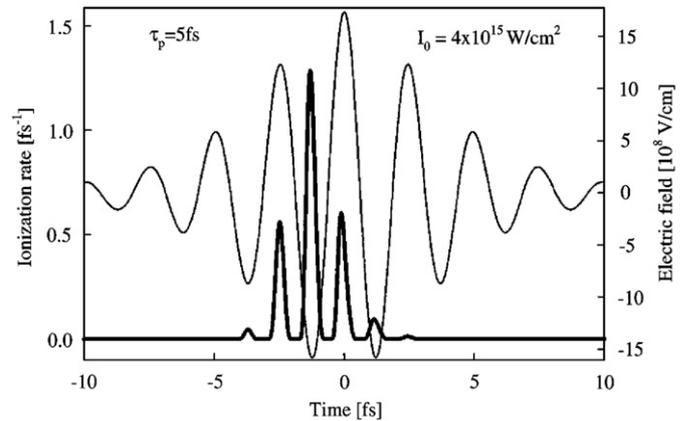


Fig. 9. Evolution of the electric field in a linearly polarized 5 fs pulse of peak intensity $4 \cdot 10^{15} \text{ W/cm}^2$ at 750 nm (thin line and ionization rate induced by this pulse in He (thick line). Barrier suppression field is $1.03 \times 10^9 \text{ V/cm}$.

recombines. Fig. 9 shows the ionization rate induced by ultra-short laser beam with a pulse duration of 5 fs and intensity of $4 \times 10^{15} \text{ W/cm}^2$. As can be seen from Fig. 9 [53], the barrier suppression field is exceeded during only a fraction of the optical cycle of the drive laser field, opening up the possibility of generating harmonics with pulse durations significantly shorter than a single optical cycle of the fundamental radiation. In this non-perturbative process, the highest photon energy achievable I_{max} can be expressed [54] as $I_{\text{max}} = I_p + 3.17 U_p$, where I_p is the ionization potential of the medium, and U_p is the ponderomotive energy of the electron, $U_p (\text{eV}) = 0.93 \times 10^{-13} I_s (\text{W/cm}^2) \lambda^2 (\mu\text{m})$, and I_s is the saturation intensity when the ensemble is fully ionized, $U_p \sim I_p$, where λ is the wavelength of the drive laser.

Since the ensemble of freed electrons interacts coherently with the field, the generated harmonics are fully coherent, i.e., coherent both longitudinally and transversely. The polarization is preserved as well in an isotropic nonlinear medium, typically a gas jet. Since this is a non-perturbative process, the phase matched photon yield at the higher harmonic does not decrease with increasing order as rapidly as in a perturbative process, but rather stays constant. There is a plateau in the energy of generated harmonics, modulated primarily by the absorption in the nonlinear medium. The ultimate yield can be optimized by maximizing the interaction region, minimizing absorption and matching the phase between fundamental and the harmonic radiation over the entire interaction volume. Fine tuning of the harmonic photon energy is achieved by tuning or chirping the fundamental radiation while adjacent harmonics provide the coarse tuning. The step size of the coarse tuning can be further reduced by

incorporating both the fundamental and second harmonic in the nonlinear process.

The HHG process can provide a standalone soft X-ray/EUV source for a number of experiments. With the soft X-ray photon flux currently available from the HHG processes it is now possible to perform pump probe experiments with attosecond time resolution, EUV spectroscopy and phase contrast and lensless imaging. The highly polarized photons with flux levels of 10^{10} photons/pulse (10 Hz PRF) and 10^{10} photons/s (1 MHz PRF) can be used to perform a number of polarization sensitive measurements. With further improvements in drive laser technology, nonlinear medium, and geometry of the interaction region, the EUV/XUV radiation from HHG can be an efficient seed for X-ray FELs.

3.2.2. HHG performance challenges and overall readiness

EUV/XUV radiation with significant flux and photon energy up to ~ 300 eV has been generated in a number of laboratory systems. Commercial units capable of delivering $< \mu\text{W}$ are currently available. Significant research is needed to further scale the EUV/XUV output from such sources to meet the requirements of either stable and reliable standalone sources or seed sources for X-ray FELs. With focused research, it is possible to develop reliable laser systems capable of delivering ~ 20 mJ energy, in 5–20 fs pulse duration, and at repetition rate up to 1 kHz within 5 years. In this timescale, it is also possible to optimize the interaction region, fully characterize the HHG beam and integrate the source, sample, diagnostic, and detector as a unit for standalone measurements and seeding of FELs. Over a 10 year period, the repetition rate of these laser systems can be increased to multiple 10 kHz and efficient HHG can be pushed to wavelengths shorter than 10 nm.

4. Conventional lasers

4.1. Advances required for electron source technologies

Conventional lasers are critical for advanced light sources in the production and manipulation of high brightness electron beams, seeding of FELs, and as either a pump or probe in experiments. While next generation light sources will rely on evolutionary advances beyond current capabilities in areas such as temporal and spatial control, or frequency conversion, revolutionary advances will be required in the area of rep-rate and the associated average power of ultrafast lasers. As light sources make the move from warm to superconducting accelerators, rep-rates will push to the 100 s of kHz or even MHz range. This will push development of high-reliability and high quantum-efficiency cathodes, as laser sources for metal cathodes (e.g. as used at the LCLS) would require 3–4 orders of magnitude increase in average power beyond what is currently commercially available.

On the electron side of the machine, this is most important for the production of high brightness electron bunches in RF photoguns. If we consider scaling a gun like the one used in the LCLS to MHz rep-rates we find the first challenge. This gun is based on a Cu cathode and temporally and spatially shaped picosecond pulses at 253 nm are used to generate the electron bunches. Typically about 200 μJ of UV light on the cathode is required to generate 1 nC of charge. The UV pulses are generated by first amplifying a Ti:sapphire laser to mJ levels, and then converting to the third harmonic with roughly 10% efficiency. Spatial shaping, transport, diagnostics, and controls further reduce the energy with a throughput of roughly 10%. This means the required energy in the fundamental is roughly 100 times the UV energy needed on the cathode. For the LCLS example, this would mean at least 20 mJ

would be required at the fundamental wavelength. At the modest rep-rate of the LCLS (120 Hz) this requires an average power of only 2.4 W, which is quite achievable with current technology. However, scaling this type of gun to next generation rep-rates of 100 kHz or even 1 MHz requires 3–4 orders of magnitude increase in the average power level to many kW, which is well beyond the capabilities of today's technology. In this case, the laser issues of scaling to kW powers can be mitigated by development of higher efficiency cathode materials that can use existing laser technologies as a drive source. This is an active area of R&D that is critical for next generation sources and discussed elsewhere [61,62].

In the case of LPAs, we are once again faced with a need for orders of magnitude increase in the average power while maintaining state of the art energy per pulse. The current state of the art is 40 J at 5 Hz or 200 W. Scaling of this accelerator technology to 1–10 kHz will require 3–4 orders of magnitude increase in average power. Most of these systems are currently based on Chirped Pulse Amplification in Ti:sapphire pumped by flashlamp pumped frequency doubled Nd:YAG. In this case, the pump lasers provide the largest hurdle to scaling the rep rates to kHz rates. Commercial flashlamp pumped Nd:YAG lasers have been available for 25 years and average powers of individual lasers have increased by 1–2 orders of magnitude during that period. Current state of the art systems use multiplexing of multiple lasers to increase the average power. It is clear that Nd:YAG pumped Ti:sapphire will not be scalable by 3–4 orders of magnitude and new technology will be required. This will require entirely new laser materials and system configurations. Development of new materials can be a slow and expensive process. Since the discovery of Ti:sapphire more than 20 years ago, many new materials have been discovered that appeared to be the next generation broadband solid-state laser material yet none succeeded at becoming commercially viable. Because development of new laser materials is so slow and risky, commercial companies tend not to invest R&D money in this area. Significant government investment in broadband laser materials will be necessary if future light sources are to rely on laser plasma based accelerators.

The other technology limiting these high energy/average power lasers is the gratings used to re-compress the pulses. Existing systems already use the largest gold coated diffraction gratings commercially available. At 40 cm widths, it is unlikely that gold grating technology will be scalable by another 2 orders of magnitude. Dielectric gratings are a promising technology for replacing gold gratings but have been in development for 15–20 years and are still not capable of handling the necessary bandwidths for these systems. Significant improvements in this or some other disruptive technology will be necessary for kW level, 40 J short pulse lasers to become reality. This is also a very expensive area for R&D with most recent advances driven by national laboratories such as LLNL.

4.2. Advances required for FELs

At the FEL end of next generation light sources, conventional lasers will be used to enhance the FEL process. Techniques such as ESASE, EEHG seeding, and HGHG seeding will all require advances in laser technology, once again because as accelerator rep-rates increase, the associated lasers must increase in rep-rate and therefore average power. In the case of ESASE, the beam manipulation is performed with the mid-IR fundamental. In this case the average power required is very close to currently available systems. However for EEHG and HGHG, there is a huge decrease in efficiency due to the harmonic conversion process. Fig. 8 shows that for energies above 50 eV, there is > 6 order of magnitude loss between fundamental and harmonic.

It is clear that for higher photon energies and higher rep rates, current commercially available lasers are 3–4 orders of magnitude below the required performance. It is predicted that 10 s of nJ of energy will be required for seeding. From Fig. 8 we can see that if we need seed photons above 50 eV, mJ of energy would be required in the fundamental, implying 10 s of Watts at 10 kHz and 100 s of Watts at 100 kHz. If EEHG works as predicted, there is significant multiplication of the seed frequency, meaning seeding could be accomplished with harmonic efficiencies in the 10^{-3} – 10^{-4} range. This would allow seeding with laser powers that are becoming available now. As such, EEHG R&D could potentially have an enormous return on investment to the DOE. In lieu of EEHG, it is clear that the laser technology needed for high rep-rate direct seeding is pushing the state of the art. In the near term seeded FELs will be forced to trade off rep-rate and photon energy to stay within the available laser capabilities. In this case seeding at 10 s of eV and 10 s of kHz is plausible with fundamental average power of 10–100 W, which will be available within the next 5 years. It should be pointed out that today's state of the art lasers may provide the necessary laser beam parameters but will require significant engineering effort to make them compatible with operations at a light source user facility. Seeding at higher photon energies and rep-rates of 100 kHz–1 MHz will require kW level lasers systems will require disruptive technology. Several developments enabling such light sources are currently pursued and discussed in the next section.

4.3. Performance advances achievable soon with immediate impact

Recent advances in ultrashort pulse fiber lasers [63] show great promise in providing high rep-rate drivers for higher efficiency next generation photocathodes as well as sources for harmonic seeding in the regime of lower photon energy described above. Ideally, this technology will begin moving into commercial products and be available within 5 years. Cryogenically cooled Yb-doped lasers are already today capable of producing multi-kW power levels with 30% wall-plug efficiency in continuous wave operation and hundreds of Watts with picoseconds pulse format. Other broadband Yb-doped laser materials allowing for femtosecond pulse generation, and with similar scaling of its thermo-optical properties as Yb:YAG have been identified, such as Yb:YLF. These systems are currently under development and may enable kW-class femtosecond laser systems in the near future, probably less than 5 years if an aggressive development effort is supported. Such highly efficient laser systems are also ideal for repetition rate scaling of LPWA acceleration. Therefore, significant and long term investment in laser and materials R&D would have a tremendous impact both in the short and long term in multiple areas of BES.

5. Conclusions

5.1. Evaluation of state of readiness of the alternative sources

The technologies presented in this review represent a broad spectrum of maturity levels, capabilities, performance limitations, and needed technology R&D. The two plasma-based acceleration technologies, LPA and PWFA, have both demonstrated 10–100 GV/m gradients and most likely over the next 5 years will demonstrate on an individual pulse basis the range of electron beam parameters for driving an XFEL. For LPA, lasers presently exist for driving an XFEL at 10 Hz, and the achievement of higher average flux requires further research on high average power, high peak intensity laser systems. Techniques for precision

control and adjustment of the laser pulse intensity profile are also needed. Improving the average power from the present level of 100 W to the 100 kW and beyond require substantial investments in new technology. Increased wall-plug efficiency, such as by using diode pump lasers, also requires a high level of investment. PWFAs, in contrast, can rely on conventional accelerator technology capabilities for high average power and wall-plug efficiencies. However, PWFA transformer ratios on the order of 5 need to be demonstrated to consider PWFA as an acceleration mechanism for a 10–30 GeV XFEL driver. A first demonstration is foreseeable within 5 years. While both of these technologies might accelerate electron beams for hard X-ray FELs, mostly likely LPA could be scaled down to university-sized XFELs more easily than PWFAs.

The DLA technology is very immature. Because of the low inherent bunch charges accelerated in a DLA, it is probably unlikely that a DLA would be capable of driving an LCLS-scale XFEL. However, particularly when coupled to a laser-driven undulator, a DLA could be a useful small laboratory source of coherent hard X-rays. This capability is beyond a 10 year horizon.

ICS is a mature technology that can produce ultrashort incoherent X-ray pulses with $\sim 1\%$ of the flux of LCLS at $\sim 1\%$ of the cost in a high peak power, low repetition rate configuration. Alternatively it can produce average flux and brightness similar to a second generation synchrotron beamline in a high-repetition rate configuration. This is a very compelling technology if X-ray coherency is not needed and moderate X-ray energy spread can be tolerated. Using current technology, an ERL-based ICS source serving multiple beamlines can be built (Fig. 10) that can serve as a compact user facility at reduced cost relative to a large ring.

Additional research on ICS over the next 10 years will likely demonstrate concepts leading to longitudinal coherent X-rays, leading to additional applications. While the lower electron energy of ICS sources restricts the ultimate power available to less than that of a high energy FEL, a coherent ICS source would offer significant capabilities at a fraction of the cost, and whereas there may be a limited number of national XFEL facilities, there could be many regional ICS X-ray sources with varying degrees of capabilities. This is an important complementary technology to XFELs that has significant offerings to BES.

HHG sources offer a viable alternative to FELs in the UV–soft X-ray range for low-average power applications. However HHG output is limited to a small fraction of the drive laser power and cannot compete with FELs for applications that are flux-driven. HHG efficiency drops sharply with photon energy so that FELs are likely to remain the dominant sources of hard X-rays for the

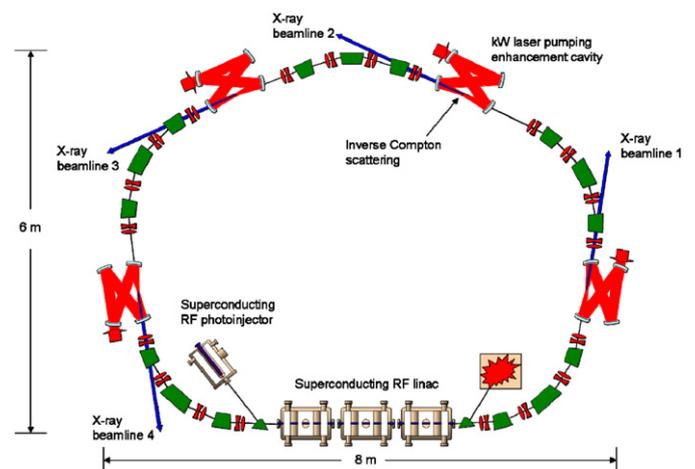


Fig. 10. Synchrotron-type ICS source with multiple X-ray beamlines as a user facility.

foreseeable future. HHG sources are important for providing seed radiation to X-ray FELs, which can upconvert and amplify the HHG seed to fully coherent hard X-rays.

5.2. Research recommendations

Because nearly all the technologies described here, as well as conventional XFELs, intrinsically depend on lasers, advancing laser technology would provide the most broad and significant capability improvements. The specific work identified for plasma-based accelerators could also bring these technologies to states capable of reducing the costs of LCLS-scale XFELs. Ancillary optical technology also requires commensurate development. Recommendations based on the technology reviews are then:

- R&D focused on increased high-power laser technology as an underpinning foundation for light-source technologies.
- R&D for LPA technology for the key demonstrations needed for driving LCLS-scale XFELs. This includes production of beams at the 10 GeV level, enhancing the beam quality (emittance and energy spread) and staging of LPAs, production of beams with energies greater than 10 GeV.
- R&D for PWFA technology for the key demonstrations needed for driving LCLS-scale XFELs. This includes acceleration of a short witness electron bunch by a PWFA with high beam quality, transformer ratios of 5 or greater (which in effect enhances the acceleration gradient in the drive linac by a factor of 5 or greater), and compact methods for in-coupling of beams for multiple PWFA stages.
- Demonstration of partially coherent X-rays from ICS sources.

The specified work identified could bring plasma-based accelerator technologies to states capable of supporting the conceptual design of an XFEL national user facility in about 10 years. Adding capabilities to ICS technology, and in particular, demonstrating techniques for generating partially coherent X-rays from ICS sources would add an inexpensive complementary capability to X-ray FELs. Coupled with advances in SRF technology, this approach could support regional incoherent high-flux ICS user facilities within 5 years and regional enhanced-capability ICS user facilities within 10 years.

In addition, DLA and laser-driven wiggler technologies could proliferate next generation light source performance to smaller university laboratories, and their technology development should be supported. These recommendations would lead to the capability of building university-scale coherent X-ray sources beyond 10 years.

Acknowledgements

The authors benefited from discussions and contributions from W.J. Brown. The authors wish to acknowledge the participants at the Accelerator Physics of Future Light Sources workshop sponsored by DOE BES September 15–17, 2009: B.E.C., E.H.E., F.X.K., W.S.G., W.P.L., T.R., J.B.R., C.B.S., D.S., and W.E.W.

References

- [1] <<http://lcls.slac.stanford.edu/>>.
- [2] T. Tajima, J.M. Dawson, *Phys. Rev. Lett.* 43 (1979) 267.
- [3] E. Esarey, C.B. Schroeder, W.P. Leemans, *Rev. Mod. Phys.* 81 (2009) 1229.
- [4] W. Leemans, E. Esarey, *Physics Today* 63 (3) (2009) 44.
- [5] T. Plettner, R. Byer, E. Colby, B. Cowan, C. Sears, J. Spencer, R. Siemann, *PR-STAB* 8 (2005) 121301.
- [6] P. Chen, J.M. Dawson, R.W. Huff, T. Katsouleas, *Phys. Rev. Lett.* 64 (1985) 693.
- [7] I. Blumenfeld, C.E. Clayton, F.J. Decker, M.J. Hogan, C. Huang, R. Ischebeck, R. Iverson, C. Joshi, T. Katsouleas, N. Kirby, W. Lu, K.A. Marsh, W.B. Mori, P. Muggli, E. Oz, R.H. Siemann, D. Walz, M. Zhou, *Nature* 445 (2006) 741.
- [8] L.M. Brown, R.P. Feymann, *Phys. Rev.* 85 (2) (1952) 231.
- [9] W.J. Brown, F.V. Hartemann, *PR-STAB* 7 (2004) 060703.
- [10] P.B. Corkum, *Phys. Rev. Lett.* 71 (1993) 1994.
- [11] M. Lewenstein, P. Balcou, et al., *Phys. Rev. A* 49 (1994) 2117.
- [12] D.H. Whittum, A.M. Sessler, J.M. Dawson, *Phys. Rev. Lett.* 64 (1990) 2511.
- [13] A. Rousse, K.T. Phuoc, R. Shah, A. Pukhov, E. Lefebvre, V. Malka, S. Kiselev, F. Burgy, J.-P. Rousseau, D. Umstadter, D. Hulin, *Phys. Rev. Lett.* 93 (2004) 135005.
- [14] W.P. Leemans, B. Nagler, A.J. Gonsalves, C. Toth, K. Nakamura, C.G.R. Geddes, E. Esarey, C.B. Schroeder, S.M. Hooker, *Nat. Phys.* 2 (2006) 696.
- [15] C.G.R. Geddes, C. Toth, J. Van Tilborg, E. Esarey, C.B. Schroeder, D. Bruhwiler, C. Nieter, J. Cary, W.P. Leemans, *Nature* 431 (7008) (2004) 538.
- [16] F. Albert, R. Shah, K.T. Phuoc, R. Fitour, F. Burgy, J.-P. Rousseau, A. Tafzi, D. Douillet, T. Lefrou, A. Rousse, *Phys. Rev. E* 77 (2008) 056402.
- [17] HHG seeding has been demonstrated at SPARC in the VUV (67 nm), L. Giannessi, Personal communication.
- [18] C.B. Schroeder, W.M. Fawley, F. Gruner, M. Bakeman, K. Nakamura, K.E. Robinson, C. Toth, E. Esarey, W.P. Leemans, Free-electron laser driven by the LBNL laser-plasma accelerator, in: C.B. Schroeder, E. Esarey, W. Leemans (Eds.), *Advanced Accelerator Concepts*, vol. 1086, AIP, New York 2009, pp. 637–642.
- [19] Fuchs, et al., *Nat. Phys.* 5 (2009) 826.
- [20] J.B. Rosenzweig, D.B. Cline, B. Cole, H. Figueroa, W. Gai, R. Konecny, J. Norem, P. Schoessow, J. Simpson, *Phys. Rev. Lett.* 61 (1988) 98.
- [21] T. Katsouleas, *Phys. Rev. A* 33 (1986) 2056.
- [22] C. Huang, et al., *Phys. Rev. Lett.* 99 (2007) 255001.
- [23] C. Huang, et al., *J. Comput. Phys.* 217 (2006) 658.
- [24] R. Ruth et al., Test accelerator for the next linear collider, SLAC PUB 6293, 1993.
- [25] A.M. Sessler, D.H. Whittum, L.-H. Yu, *Phys. Rev. Lett.* 68 (1992) 309.
- [26] C. Sears, E. Colby, B. Cowan, R. Siemann, J. Spencer, R. Byer, T. Plettner, *PRL* 95 (2005) 194801.
- [27] C. Sears, E. Colby, R. Ischebeck, C. McGuinness, J. Nelson, R. Noble, R. Siemann, J. Spencer, D. Walz, T. Plettner, R. Byer, *PRST-AB* 11 (2008) 061301.
- [28] C. Sears, E. Colby, R. England, R. Ischebeck, C. McGuinness, J. Nelson, R. Noble, R. Siemann, J. Spencer, D. Walz, T. Plettner, R. Byer, *PRST-AB* 11 (2008) 101301.
- [29] X. Lin, *PRST-AB* 4 (2001) 051301.
- [30] B. Cowan, *PRST-AB* 6 (2003) 101301.
- [31] B. Cowan, *PRST-AB* 11 (2008) 011301.
- [32] T. Plettner, P. Lu, R. Byer, *PRST-AB* 9 (2006) 111301.
- [33] J. Rosenzweig, A. Murokh, C. Pellegrini, *Phys. Rev. Lett.* 74 (1995) 2467.
- [34] C. McGuinness, E. Colby, R. Byer, *J. Mod. Opt.* 19 (2009).
- [35] N. Na, R. Siemann, R. Byer, *PRST-AB* 8 (2005) 031301.
- [36] T. Plettner, R. Byer, *PRST-AB* 11 (2008) 030704.
- [37] P. Hommelhoff, C. Kealhofer, M. Kasevich, *Mod. Trends Laser Phys.* 19 (4) (2009) 736ff.
- [38] E. Colby, et al., An electron source for a laser accelerator, in *Proceedings of the PAC2005*, 2005, p. 2101ff.
- [39] D. Xiang, G. Stupakov, *Phys. Rev. ST Accel. Beams* 12 (2009) 030702.
- [40] W.S. Graves, W.J. Brown, F.X. Kaertner, D.E. Moncton, *Nucl. Instr. and Meth.* 608 (i.1) (2009) s103.
- [41] W.J. Brown, F.V. Hartemann, *AIP Conf. Proc.* 737 (2004) 839.
- [42] J.B. Rosenzweig, O. Williams, *Int. J. Mod. Phys. A* 23 (2007) 4333.
- [43] G.A. Krafft, *Phys. Rev. Lett.* 92 (2004) 204802.
- [44] D.C. Nguyen, N.A. Moody, G. Bolme, L.J. Castellano, C.E. Heath, F.L. Krawczyk, S.I. Kwon, R. McCrady, F.A. Martinez, P. Marroquin, M. Prokop, P. Roybal, W.T. Roybal, T.L. Tomei, P.A. Torrez, W.M. Tuzel, L.M. Young, T. Zaugg, Endpoint energy measurements of field emission current in a CW normal-conducting RF injector, submitted for publication to *PRST-AB* (2010).
- [45] K. Baptiste, J. Corlett, S. Kwiatkowski, S. Lidia, J. Qiang, F. Sannibale, K. Sonnad, J. Staples, S. Virostek, R. Wells, *Nucl. Instr. and Meth.* 599 (2009) 9.
- [46] A. Burrill, I. Ben-Zvi, R. Calaga, X. Chang, H. Hahn, D. Kayran, J. Kewisch, V. Litvinenko, G. McIntyre, A. Nicoletti, D. Pate, J. Rank, J. Scaduto, R. Tao, K. Wu, A. Zaltsman, Y. Zhao, H. Bluem, M. Cole, M. Falletta, D. Holmes, E. Peterson, J. Rathke, T. Schultheiss, A. Todd, R. Wong, J. Lewellen, W. Funk, P. Kneisel, L. Phillips, J. Preble, D. Janssen, V. Nguyen-Tuong, *Nucl. Instr. and Meth.* 557 (2006) 75.
- [47] C. Hernandez-Garcia et al., Performance and modeling of the JLAB IR FEL upgrade injector, in: *Proceedings of the 26th International Free-electron Laser Conference*, Trieste, Italy, 2004, p.558.
- [48] J.B. Rosenzweig, M.P. Dunning, E. Hemsing, G. Marcus, P. Musumeci, A. Marinelli, M. Ferrario, Quasicrystalline beam formation in RF photoinjectors, in: *Proceedings of the 30th International Free-Electron Laser Conference*, Gyeongju, Korea, 2008.
- [49] Kulagin, et al., *Phys. Rev. Lett.* 99 (2007) 124801.
- [50] O. Klimo, J. Psikal, J. Limpouch, V.T. Tikhonchuk, *PRST-AB* 11 (2008) 031301.
- [51] J.-F. Hergott, M. Kovacev, et al., *Phys. Rev. A* 66 (2002) 021801(R).
- [52] E. Takahashi, Y. Nabekawa, et al., *IEEE J. STQE* 10 (2004) 1315.
- [53] M. Schnürer, Spielmann Ch., et al., *Phys. Rev. Lett.* 80 (1998) 3236.
- [54] Z. Chang, A. Rundquist, et al., *Phys. Rev. Lett.* 79 (1997) 967.
- [55] I.J. Kim, C.M. Kim, et al., *Phys. Rev. Lett.* 94 (2005) 243901-1-4.
- [56] E.J. Takahashi, et al., *PRL* 101 (2008) 253901.
- [57] H.T. Kim, I.J. Kim, et al., *IEEE J. STQE* 10 (2004) 1329.
- [58] E. Gibson, X. Zhang, et al., *IEEE J. STQE* 10 (2004) 1339.

- [59] M. Nisoli, E. Priori, et al., *Phys. Rev. Lett.* **88** (2002) 33902-1.
- [60] S. Kazamias, D. Douillet, et al., *Phys. Rev. Lett.* **90** (2003) 193901-1.
- [61] D. Dowell et al., Cathode R&D for future light sources, *Nucl. Instr. and Meth.*, submitted.
- [62] N. Moody, et al., *J. Appl. Phys.* **102** (2007) 104901.
- [63] See for example A. Tunnermann, J. Limpert, S. Nolte, *Ultrashort pulse fiber lasers and amplifiers*, in: *Femtosecond Technology for Technical and Medical Applications*, Springer, Berlin, ISSN 1437-0859, 2004.



ELSEVIER

Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima

Review

Enabling instrumentation and technology for 21st century light sources

J.M. Byrd^{a,*}, T.J. Shea^b, P. Denes^a, P. Siddons^c, D. Attwood^d, F. Kaertner^e, L. Moog^f, Y. Li^f,
A. Sakdinawat^d, R. Schlueter^a

^a Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA

^b Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

^c Brookhaven National Laboratory, Upton, Long Island, New York 11973, USA

^d University of California Berkeley, Berkeley, CA 94720, USA

^e Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

^f Argonne National Laboratory, Argonne, Illinois 60439, USA

ARTICLE INFO

Article history:

Received 3 June 2010

Accepted 19 June 2010

Available online 17 July 2010

Keywords:

Light sources

Instrumentation

Synchrotron radiation

ABSTRACT

We present the summary from the Accelerator Instrumentation and Technology working group, one of the five working groups that participated in the BES-sponsored Workshop on Accelerator Physics of Future Light Sources held in Gaithersburg, MD September 15–17, 2009. We describe progress and potential in three areas: attosecond instrumentation, photon detectors for user experiments, and insertion devices.

© 2010 Elsevier B.V. All rights reserved.

Contents

1. Introduction	911
2. Attosecond instrumentation	911
2.1. Ultrashort electron and photon bunch length measurements	911
2.2. High precision timing distribution	912
2.2.1. Optical master oscillator	913
2.2.2. Timing distribution via length stabilized fiber links	913
2.2.3. Femtosecond synchronization techniques: optical to RF and optical to optical	913
2.3. Optics	913
2.3.1. Optics and diffractive structures for lens based and lenseless nanoscale imaging with femtosecond/attosecond FEL pulses	914
2.3.2. Mirrors for femtosecond/attosecond pulses	914
3. Detectors	915
4. Insertion devices	916
4.1. Present status	916
4.2. Undulator technology options for future light sources	917
4.3. Principal superconducting undulator (SCU) development challenges and readiness	917
4.4. Undulator R&D tasks	918
5.1. Attosecond instrumentation	919
5.2. Detectors	919
5.3. Insertion devices	919
References	919

1. Introduction

As part of the Workshop on Accelerator Physics of Future Light Sources sponsored by the Department of Energy Office of Basic Energy Sciences, a working group was organized to examine the state of the art of accelerator instrumentation and technology for

* Corresponding author. Tel.: +1 510 486 6329.
E-mail address: JMByrd@lbl.gov (J.M. Byrd).

future light sources and to recommend a few topics for which directed R&D funding could help enable the tremendous potential of the next generation. These future light sources will achieve significant improvements in brightness, peak brightness, time resolution and stability. To reach these goals, advances are required in accelerator instrumentation and technology in many diverse areas, such as: RF acceleration, component alignment and stability, attosecond instrumentation and optics, photocathodes, pulsed power components, photon detectors, halo monitors, collimators, lasers, insertion devices, noninvasive profile monitors, high resolution position monitors, trapped ion diagnostics, and feedback systems.

To cull a few R&D topics from this long list, we applied the following criteria: impact, viability, uniqueness, and applicability. Technological developments with high impact will significantly enhance the performance and scientific output of future machines. Viable R&D programs should show results in five years and lead to deployable systems within about 10 years. Many of the technologies listed in the previous paragraph would be developed specifically for certain light source types and several are covered in publications by other working groups. We identified unique topics that were not fully addressed by the other working groups, and that were applicable to multiple types of future light sources. By consensus of the working group, our final selections are: attosecond instrumentation and optics, detectors, insertion devices, and photocathodes. The first three will be discussed in the following sections while the photocathode technology was singled out for a more detailed treatment in a separate paper.

2. Attosecond instrumentation

Free electron lasers are emerging as the 21st century source for high brightness ultrafast X-rays. To date, two facilities are operational with several more planned to come online in the next few years. With the recent operation of the LCLS with electron bunch lengths of less than 10 fs, the possibility of sub-fsec, or attosecond, pulses is approaching. We have identified three areas where development is needed to be able to take advantage of these pulses. These include electron bunch length, photon pulse length and spectral diagnostics, timing and synchronization, and X-ray optics. Each of these is discussed in further detail in the following subsections.

2.1. Ultrashort electron and photon bunch length measurements

One of the challenges facing the next generation of ultrafast X-ray FELs is the characterization of electron and photon pulses with femtosecond time scales. This includes both measurement of the longitudinal current distribution and energy spread for the electron bunches and the time and spectral distribution of the photon pulses. Since both of the pulses are expected to reach below 10 fs in the near future, sub-fsec, or attosecond, resolution will be required. In addition, a measurement of the arrival time of each of the pulses is necessary with respect to the pump laser in a pump/probe experiment. Furthermore, the ideal measurement is nondestructive and is made on every electron and photon pulse.

For measurement of electron bunches, several approaches are continuing development to address the above needs. These approaches include electro-optic sampling [1,2], coherent synchrotron terahertz radiation [4], streak cameras [8,9], transverse deflecting structures [3], and fluctuational interferometry [5–7]. A comparison of these techniques is beyond the scope of this paper. As an example, we examine below the resolution of the transverse deflecting structures (TDS).

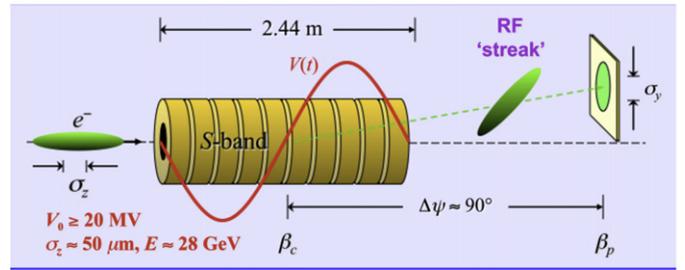


Fig. 1. Schematic view of a transverse deflecting structure for bunch length measurement.

A schematic view of the transverse deflecting structure is shown in Fig. 1. A high-frequency electromagnetic field exerts a time-dependent transverse force on the electrons, analogous to the sawtooth voltage in the oscilloscope, and converts the temporal profile of the bunch into a transverse (here vertical) streak on an observation screen. The bunch charge density profile can thus be measured single shot. Furthermore, appropriate variations of quadrupole strengths in the beam line upstream of the TDS allow for time-resolved horizontal phase space tomography. A crucial quantity that can be deduced from such a measurement is the horizontal slice emittance. A second screen mounted behind a dipole magnet is utilized to measure the energy distribution along the bunch axis and to carry out a longitudinal phase space analysis.

The longitudinal resolution of the TDS can be approximated as the vertical beam size at the screen divided by the vertical deflection along the bunch given by the cavity. This can be written as [3]

$$\sigma_{\text{long}} = \frac{\sqrt{\varepsilon_y}}{\sqrt{\beta_{yTDS} \sin \Delta \phi_y}} \frac{\lambda_{TDS} E / e}{V_{TDS}} \quad (1)$$

where ε_y is the vertical beam emittance, β_{yTDS} is the vertical beta function at the center of the TDS, $\Delta \phi_y$ is the betatron phase advance between the TDS and observation screen, λ_{TDS} is the wavelength of the TDS, and V_{TDS} is the vertical deflecting voltage. A given configuration is optimized by maximizing the beta function in the center of the TDS and by choosing a betatron phase advance of an odd multiple of 90° . From this point, the longitudinal resolution can be increased by reducing the RF wavelength of the TDS and increasing the deflecting voltage. An example of a measurement at FLASH [3] is shown in Fig. 2.

For measurement of X-ray photon pulses, several approaches are continuing development to address the above needs. These approaches include optical streak cameras [10,11], conventional RF streak cameras [8,9], and fluctuational interferometry [12]. As in the case for electron bunches, a comparison of these techniques is beyond the scope of this paper. As an example, we examine in more detail recent advances in streak cameras operating at THz frequencies.

Streak cameras are proven tools in ultrashort pulse metrology and have single-pulse capability. In conventional streak cameras, photocathodes are used to generate electron bunches with temporal structures identical to that of the light pulses. The electrons are accelerated, transversely deflected by a rising electric field and then detected on a phosphor screen. Such schemes are limited in their time resolution to a few hundred femtoseconds. This limitation is mainly due to the spread of the initial momenta of the electrons released from the photocathode, which leads to a significant temporal broadening of the wave packet upon propagation to the deflector. This limitation can be overcome by using techniques recently developed for attosecond metrology [10]. A photoemitter is immersed in an electromagnetic field as shown in Fig. 3, transforming the time of

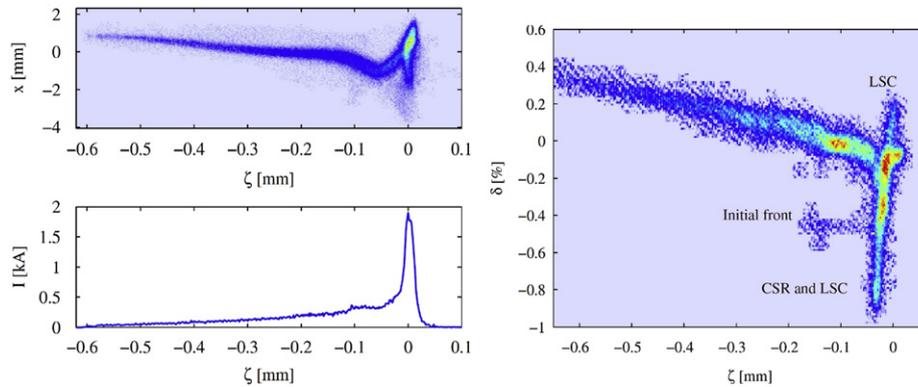


Fig. 2. Example longitudinal distribution from FLASH measured with the TDS (from Ref. [3]). The longitudinal phase space reconstructed from the measurement is shown in the right.

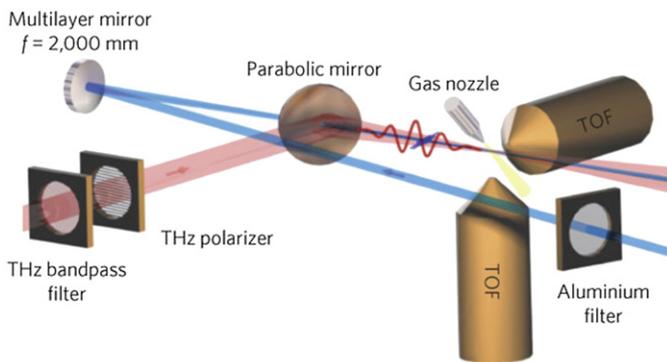


Fig. 3. Horizontally polarized soft X-ray (blue beam) and vertically polarized terahertz (red beam) pulses are focused and collinearly superimposed in a krypton gas target (from Ref. [10]). Photoelectrons are detected with two time-of-flight (TOF) spectrometers, one parallel and one perpendicular to the terahertz polarization. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

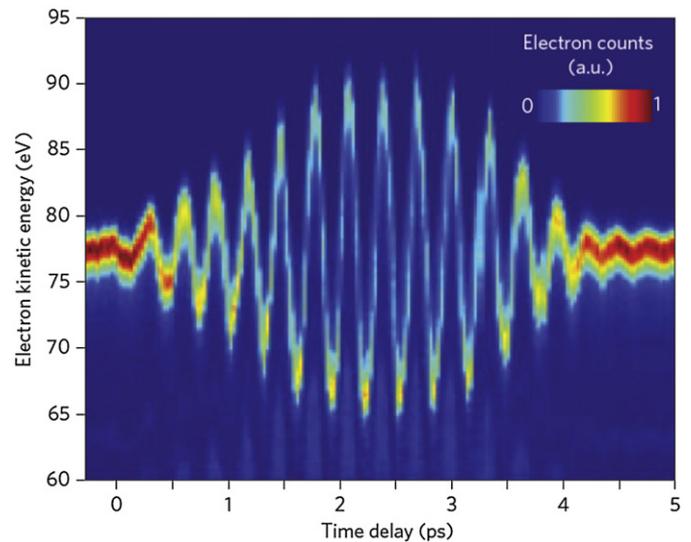


Fig. 4. Series of kinetic energy spectra of 4p photoelectrons detached from krypton atoms by a 13.5-nm soft X-ray pulse in the presence of an intense pulsed terahertz field (false-colour representation). The energy shift of the electrons versus the X-ray/terahertz delay directly represents the strength of the X-ray field (from Ref. [10]).

prompt ionization into an energy shift of the resulting photoelectrons. In this example, a deflecting field in the THz regime allows measurement of pulses up to tens of fs long. A measurement of this deflecting THz field is shown in Fig. 4.

2.2. High precision timing distribution

Fourth-generation light sources such as seeded FEL require a whole array of femtosecond lasers and synchronization techniques between low-level RF-systems, photo-injector laser, seed lasers as well as potential probe and diagnostic lasers. A layout of a generic seeded FEL facility and its synchronization needs is shown in Fig. 5. One of the main challenges in reaching the level of synchronization required for next generation light sources is transmission of a timing signal over a relatively large facility. For example, in a facility of a kilometer in length, diurnal temperature variation results in cable length variation from several hundred ps to a nanosecond. The master clock for the overall facility is an ultrastable oscillator. This could be either an ultra-low noise master microwave oscillator or a mode-locked fiber laser, locked to a microwave oscillator. The second option combines the superior high frequency noise characteristics of the fiber laser with the superior low-frequency noise characteristics of the microwave oscillator. The timing signals are distributed over stabilized fiber links throughout the facility and used to derive secondary synchronized sources, and lock critical optical and RF sub-systems [13]. Several approaches have been used for stabilizing the fiber links. The two approaches that have been implemented at light sources are typically referred to as “pulsed” and “CW” (continuous-

wave). In the pulsed approach, optical pulses from the master fiber laser are transmitted directly on the fiber and stabilization is achieved by locking the reflected pulse repetition rate to the master clock. RF timing signals are derived from harmonics of the pulse repetition rate. In the CW approach, each link comprises one arm of an optical Michelson interferometer which senses the variation in the link. RF timing signals are transmitted as modulations of the optical carrier with a phase adjusted by the correction sensed with the interferometer. Alternatively to the use of a mode-locked laser as the optical master oscillator also a highly stable continuous wave (cw)-laser could be used to length stabilize the optical fiber links and for transmission of optical as well as microwave signals [14,15].

Rapid advances over the last few years in frequency metrology based on ultrafast lasers and, therefore, also in laser stabilization and synchronization, show that the requisite low timing jitters between different laser and rf-systems can be achieved and maintained over long times and distances of several hundred meters [16].

2.2.1. Optical master oscillator

Over the last years high repetition rate (200–250 MHz), 100 fs fiber lasers have been developed [17] and are also commercially

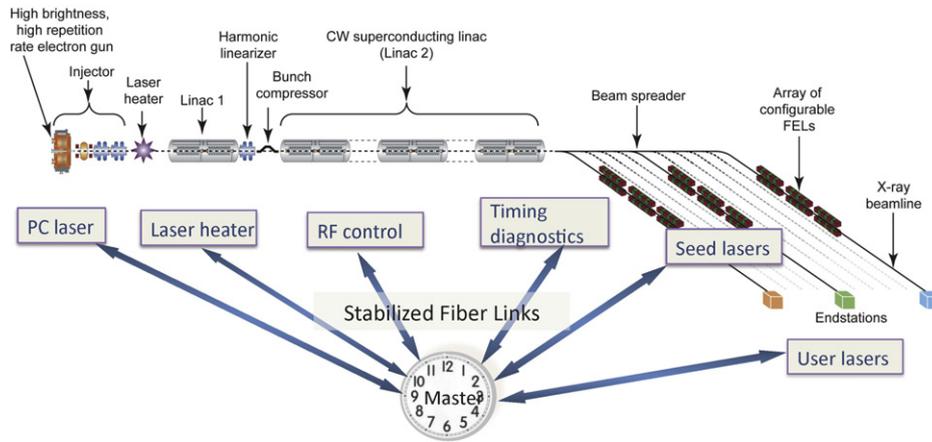


Fig. 5. Schematic outline of the timing distribution and synchronization for a seeded FEL facility.

available with several hundred mW of output power after amplification. These laser sources are, in terms of pulse parameters such as pulsewidth, repetition rate and output power, ideally suited to serve as the master oscillator for the intended facility. There high frequency timing jitter is below 1 fs and therefore these sources are well suited for timing distribution with 1 fs jitter level or even below [18].

2.2.2. Timing distribution via length stabilized fiber links

The use of optical signals as a means for timing delivery in an accelerator environment has many advantages compared to conventional temperature-stabilized coaxial cables, such as better robustness against electromagnetic interference (EMI), ease of installation, and space efficiency. Furthermore, the use of pulse trains enables direct stabilization of the group-delay of the fiber link while suppressing of Brillouin scattering and residual reflections. It also adds more flexibility in the operation and diagnostics of FELs by using the delivered ultralow-jitter pulse trains for direct seeding of optical amplifiers or down-conversion of microwave signals. Most importantly, optical cross correlation can be used to detect drifts in the length of the fiber link with high precision and robustness. Therefore, such drift detectors can be used to feed back on the length of the fiber link and keep it constant with better than 10fs precision over many days of operation [13] demonstrated sofar for links up to 300 m in length. Compact single-crystal balanced cross-correlators for timing error detection and long-term stable timing link stabilization have been developed [19].

2.2.3. Femtosecond synchronization techniques: optical to RF and optical to optical

Tight synchronization is necessary not only for all the ultrafast lasers in the FEL facility, but also for the RF sources driving the accelerator sections. The electron beam dynamics is controlled by the microwave fields in the accelerator cavities. Therefore, highly stable microwave signals, tightly synchronized with each other in different accelerator sections, are an indispensable prerequisite for the control of electron beams with higher timing accuracy. High-quality RF signals can be extracted from the optical pulse trains delivered by timing-stabilized fiber links. However, the extraction of drift-free RF signals, which is tightly locked with the pulse trains, is a highly nontrivial task. Excess noise in the photodetection processes and thermal drifts of photodetectors [20] seriously compromise the achievable timing stability of RF-signals. On the other hand, tight synchronization of a mode-locked laser to a microwave frequency standard (Fig. 5) is also necessary for the optical master oscillator.

These issues have been addressed by the development of the balanced optical-microwave phase detector (BOM-PD) [21]. This device is based on a differentially biased Sagnac-loop interferometer for sensitive timing detection with electro-optic sampling. This PLL can be operated either by using an optical pulse train as a reference and a voltage-controlled oscillator (VCO) as a slave oscillator (optical-to-RF synchronization) or by using a RF signal as a reference and a mode-locked laser as a slave oscillator (RF-to-optical synchronization), see Fig. 6.

Use of the BOM-PDs has allowed synchronization of a 10.225-GHz VCO to a 200.5-MHz optical pulse train from an Er-fiber mode-locked laser. The measured short-term jitter was about 5 fs (1 Hz–1 MHz), and the long-term stability is below 7 fs rms, integrated over 10 h [13]. When all necessary components and sub-systems are well synchronized, the final issue is to precisely measure and monitor the achieved stability at critical points in the facility. For example, the electron beam stability at the bunch compressor and the phase stability of the microwave fields driving the accelerator structures must be continuously monitored. Availability of ultralow-jitter pulse trains at many positions in the facility allow the demonstrated techniques to support these diagnostic tools. For example, an electron bunch arrival time monitor [22] can be implemented based on electro-optic sampling. The down-conversion of microwave signals in the GHz range using BOM-PDs can be used to verify synchronism with the pulse trains at various points in the RF-system of the facility.

2.3. Optics

Experiments at new soft and hard X-ray free electron lasers (FELs) will require the use of specialized optics that are both tailored to the unique qualities of the FEL beam and to the specific experiment being performed. Crystal optics for hard X-rays will require special attention to assure that absorbed energy in ultrashort pulses does not cause short term heating sufficient to affect spectral selectivity. Optical components for focusing soft X-rays, such as zone plates, diffractive structures for holography, polarization control optics, femtosecond/attosecond mirrors, and pulse shaping optics will need to be properly designed and optimized for use with the intense soft X-ray FEL beam. Currently, there has been limited development and utilization of optics with ultrafast EUV sources such as laser high harmonic generation (HHG) and the FLASH FEL facility. In order to be prepared for the wide range of new scientific opportunities, much research and development in the optimal design, material selection, fabrication, efficiency, and radiation damage thresholds of the optics is needed. Indeed there is some experience in the EUV region, but

essentially nothing in the soft and hard X-ray regions, and little quantitative studies of distortion and damage effects in any of these regions.

2.3.1. Optics and diffractive structures for lens based and lensless nanoscale imaging with femtosecond/attosecond FEL pulses

Wavefront preserving focusing optics can be used in many ways: lens-based full field microscopy, keyhole coherent diffractive imaging, and formation of a nanometer-scale, intense probe. The main advantage of using a lens-based imaging system is that the image of a complex sample can be obtained directly and with high resolution. However, a typical configuration for a transmission X-ray microscope at 3rd generation synchrotrons, where Fresnel zone plate objectives are used with partially coherent illumination, cannot be directly translated for use in the spatially coherent beam from the FEL since coherent artifacts would affect the image. Alternatives include direct imaging using diffractive optics which can utilize the spatially coherent X-rays or use of an optic to reduce the spatial coherence of the illuminating beam within a single shot. Imaging using diffractive optics that can utilize the spatially coherent X-rays has been demonstrated with the DIC, spiral, and Zernike zone plates, shown in Fig. 7, using spatially coherent X-rays at the Advanced Light Source in Berkeley, CA. These zone plates are sensitive to both the amplitude and phase properties of the sample. For example, the DIC zone plate has been used to image phase contrast in magnetic samples. These zone plates are single-element imaging objectives in the microscope and are trivial to align. SEM images of the zone plates are shown in Fig. 7. The zone plates shown in Fig. 7 can utilize spatially coherent light from the FEL to form high resolution full-field real-space images. They are phase sensitive and can be used to detect phase contrast in a sample, potentially reducing radiation dose to the sample. Experiments designed to quantitatively describe conditions for survivability of these types of zone plates using various imaging geometries are required to predict their best use. Initial calculations suggest that

survivability is possible in situations such as full-field zone plate imaging where only the direct zeroth order beam is avoided and a lens of many zones is used so as to minimize absorbed energy per unit mass. On the other hand, such lenses can play a valuable role even when destroyed in use, such as the zone plate lens array used in sequel keyhole coherent diffractive imaging experiments at FLASH. Examples of a lens before use, and portions of the array showing absent lens positions after use, are shown in Fig. 8. Mass production methods for producing large arrays of disposable zone plates cheaply should be investigated for such studies.

An additional form of diffractive structure likely to play a very useful role in nanoscale, ultrashort pulse imaging is the uniformly redundant array (URA), as seen in Fig. 9. The URA is designed for holographic soft X-ray imaging, offering the advantage of increased reference beam intensity, equal or greater than the object beam, while maintaining the high spatial resolving capability. These attributes are important for image quality, linearity, and accurate image fidelity, while making better use of available coherent photon flux. Again, flux related distortion and damage thresholds must be understood well in advance of experimental planning to achieve the best scientific results.

2.3.2. Mirrors for femtosecond/attosecond pulses

Conventional multilayer mirrors are generally not optimized for FEL sources and ultrafast experiments in the femtosecond/attosecond time domain. According to Heisenberg Uncertainty limits, very short pulses require appropriately wide spectral bandwidths, as expressed below.

$$\Delta E \Delta t_{FWHM} \leq 1.82 \text{ eV-fs} \quad (2)$$

To support such very short pulses mirrors with appropriately broad spectral bandpass are required, generally broader than typical required for longer pulse experiments. Tradeoffs between overall reflectivity and bandwidth need to be carefully considered. Fig. 10 shows an example of a multilayer mirror designed to

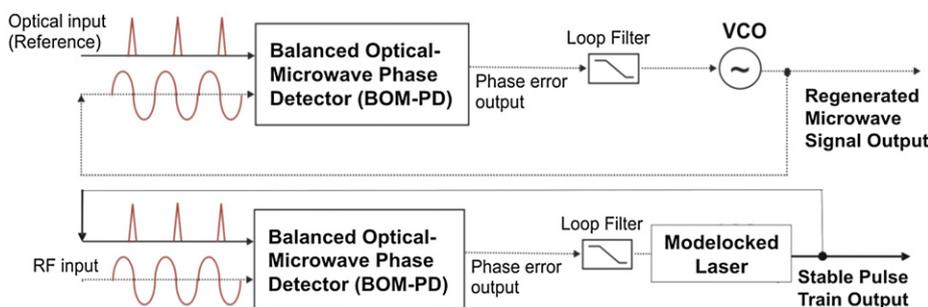


Fig. 6. Schematic of optical-to-RF and RF-to-optical synchronization using a balanced optical-microwave phase detector (BOM-PD).

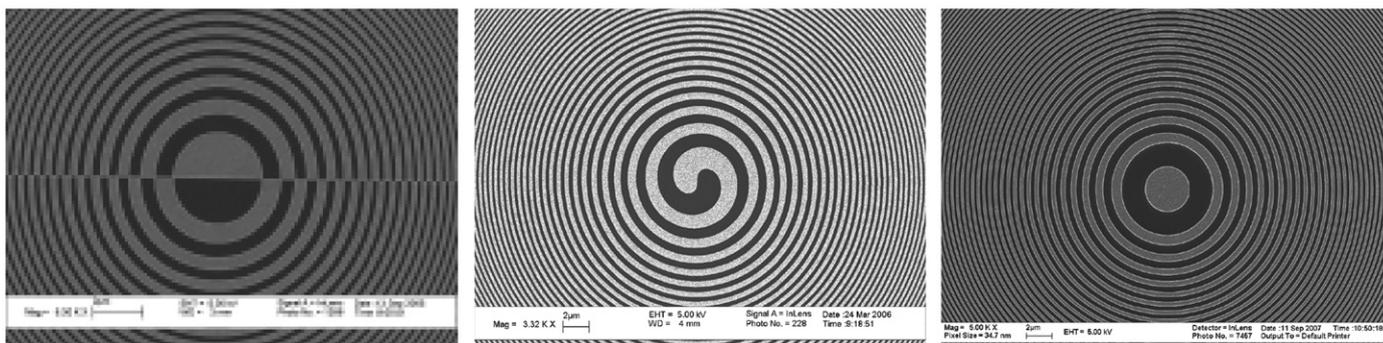


Fig. 7. SEM images of DIC (left), Spiral (center), and Zernike (right) zone plates. (Sakdinawat).

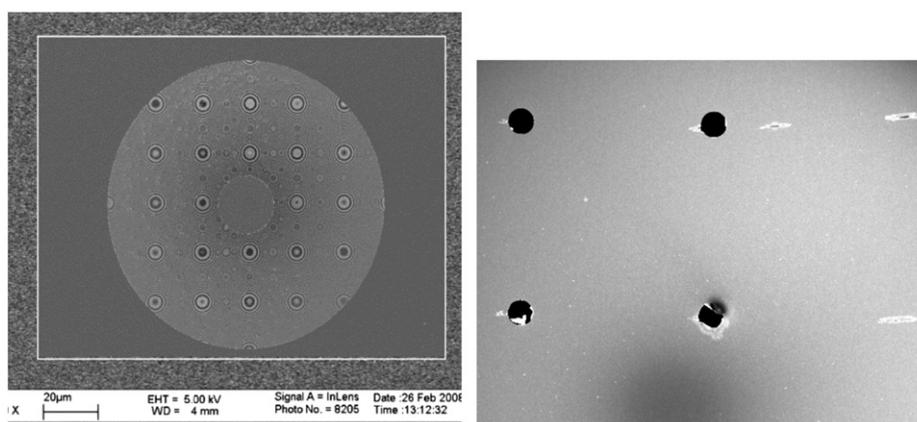


Fig. 8. SEM image (left) of one of over 200 zone plates fabricated in an array using electron beam lithography for keyhole coherent diffractive imaging. These zone plates were used in single-shot CDI experiments at the FLASH FEL. The right image shows the location of destroyed zone plates after use in the direct FEL beam. (Sakdinawat).

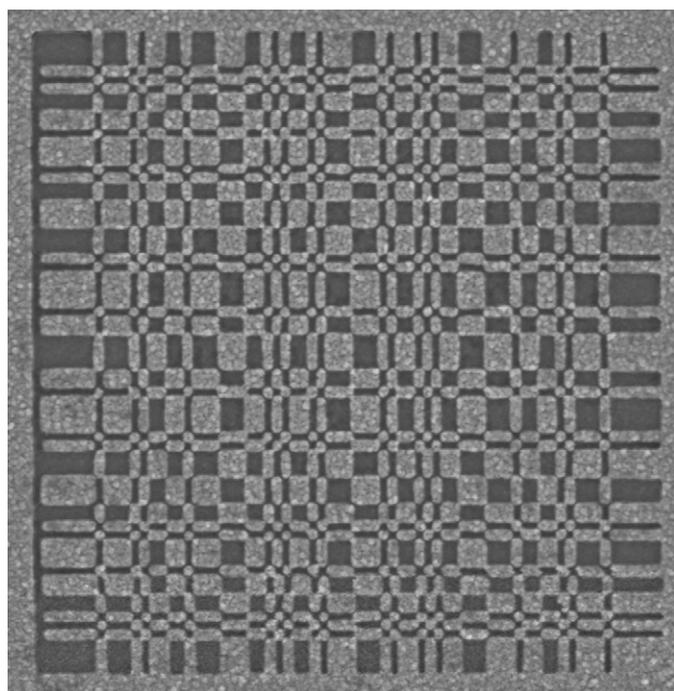


Fig. 9. An SEM image of a uniformly redundant array diffractive element used as a reference for high resolution holographic imaging. The highly parallel nature of this type of reference enables an improvement in resolution over pinhole-based Fourier transform holography. (Sakdinawat and Marchesini).

have sufficient bandwidth to support 100as pulses at a photon energy of just under 100 eV. Further optimizations are required to explore the available material combinations for ultrashort EUV to X-ray pulses for a wide range of applications, including mirrors for two-color pump–probe experiments, etalons for sequential pulse generation at specified sub-femtosecond separation.

3. Detectors

New (and existing) synchrotron light sources can benefit enormously from R&D on X-ray detectors [23]. Whereas older experiments simply accepted the limitations of existing, commercial detectors, many experiments at newer synchrotrons simply cannot work without detectors beyond the commercial state of the art. Experiments at the latest FEL facilities illustrate

escalating detector requirements even more starkly. At the Linac Coherent Light Source, diffract and destroy experiments must record two-dimensional diffraction patterns on a shot-by-shot basis, with a wide dynamic range over the detector, and no memory of the previous pulse [24]. At the European X-ray Free Electron Laser, the complex bunch structure requires storing and tagging X-rays, and then reading them out between bunches [25]. In both of these examples, micro-electronic enabled detector technologies, originally developed for High Energy Physics, were at the root of the solution. Such detectors are likely to continue to provide beyond-the-state-of-the-art solutions to new (and existing) sources. The development of such sophisticated systems is not a trivial undertaking, and typically demands many years to fully complete. For the machines currently under construction it is already too late to start a development program to have things ready for day-one operations. The LCLS and XFEL projects fortunately have programs in place, but as we learn about these exciting new sources and how to best use them, we will need to build different systems from those conceived currently.

For non-FEL sources, it is becoming clear that the next generation of X-ray detectors will add spatial resolution in two dimensions to whatever other properties they may have. Detectors having excellent energy resolution or good time resolution already exist, but the next generation must be multi-dimensional. For example, energy resolving detectors currently have relatively few elements. The next generation will offer megapixel designs with per-pixel energy resolution similar to that of current single-element devices. Similar comments will relate to other property combinations, such as position and time. In order to achieve these goals, direct detection will become mandatory, and current devices relying on indirect detection via a scintillator will become obsolete. Achieving these goals will inevitably demand more complex circuitry be compressed into smaller pixel areas. The physics of charge collection will require intelligent reconstruction of the charge cloud in order to achieve good energy resolution and/or spatial resolution. This in turn will demand even more pixel complexity. All of this will require innovation in sensor physics, integrated circuit technology and device packaging. In order to reach the required level of sophistication, we must build infrastructure which can support it. This means bringing the US sensor foundries up to modern standards, providing the best software tools to US chip designers, and bringing in and educating new talent to take us forward. With the exponential progress in semiconductor processing, modern detectors are increasingly based on direct detection in semiconductors. Hard and soft X-rays present challenges for which R&D on materials and processes are needed.

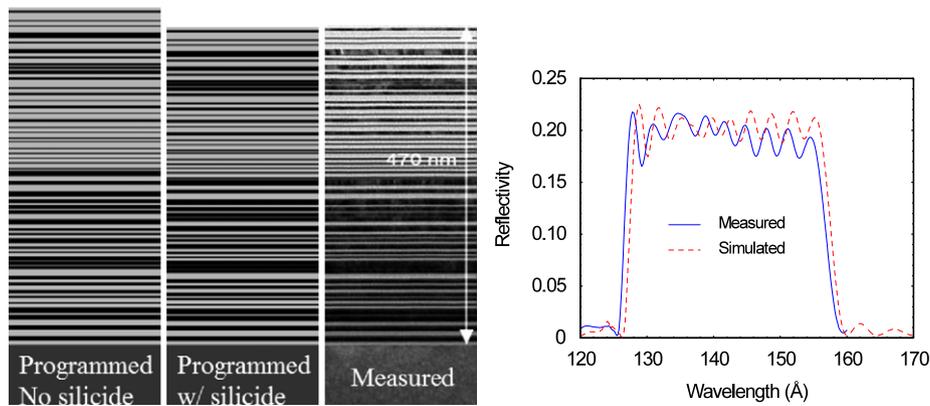


Fig. 10. An aperiodic multilayer mirror with 18 eV bandwidth, wide enough to support a 100 as pulse at 88 eV photon energy (Aquilari, Liu and Gullikson).

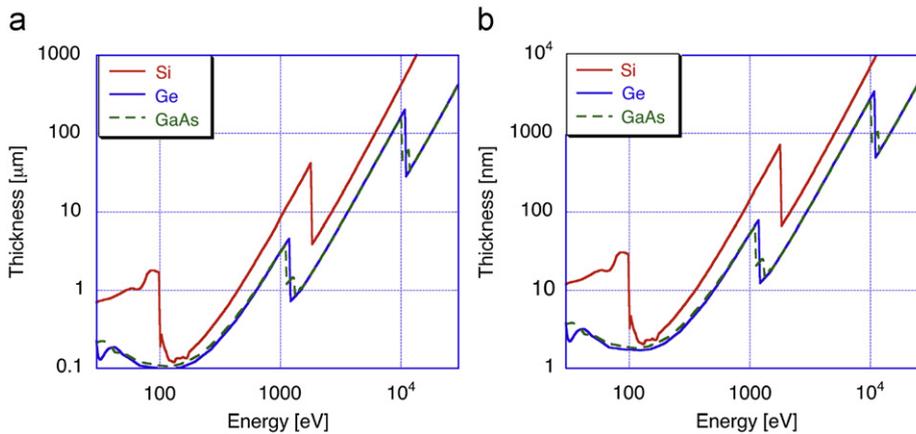


Fig. 11. (a) Thickness needed to absorb 95% flux and (b) window thickness to transmit 95% flux.

Fig. 11a shows the required detector thickness to absorb 95% of incident X-rays as a function of energy for three typical semiconductor detector materials: Si, Ge and GaAs. Typical silicon detector thicknesses are 200, 300 μm , showing that other materials are needed for harder X-rays. Conversely, Fig. 11b shows the maximum thickness of inert detector material allowable in order that 95% of incident X-rays are not absorbed, demonstrating that R&D on thin window implants is essential for soft X-ray detectors. Both of these areas are in need of significant R&D if the sophisticated detector systems being imagined by current instrument designers are ever to become a reality.

The issue of circuit complexity on the readout chips of pixelated detectors will require the use of deep submicron integration and even beyond, into the area of 3D stacking of circuitry to achieve optimal combinations of technologies and higher densities, and new approaches to thermal management, since all of this intelligence will consume significant power [26].

In addition to the primary data-collection devices, we should not neglect beam and beamline diagnostics. Devices to measure and control the photon beam position are significantly more challenging than the equivalent instruments for determining the electron beam position in an FEL or storage ring. The charged particle beam can interact directly with pickup electronics, whereas the photon beam must first be converted into an electrical signal; a process which inevitably introduces non-linearity to the problem. Measurements of photon bunch length below 1 ps (the current state of the streak-camera art [27]) is extremely challenging, and conventional streak-camera technology may not provide the final answer. Novel approaches, however speculative, should be tried.

4. Insertion devices

4.1. Present status

Well-established, high-performance undulator technologies include (a) planar permanent magnet (PM) undulators [28], (b) elliptically polarizing undulators (EPUs), e.g. the widespread Apple-IIs [29], and (c) in-vacuum undulators (IVIDs) [30], now standard in many synchrotrons. Under development are (a) new varieties of polarizing undulators, (b) quasi-periodic devices, (c) cryogenic PMs [30,31], and (d) superconducting undulators (SCUs), including planar designs and beyond [33]. Undulator designs specific for FELs and ERLs, but unsuitable for storage rings include (a) those with poles close to beam horizontally, e.g. vertically polarizing planar devices, some designs for variable polarization, e.g. Delta [34] and Apple-III [35] designs, and some superconducting designs, e.g., helical windings on a round beam tube [36,37], (b) ultra-small-gap, (< 4 mm) devices [38], and (c) specialty designs for small emittance, e.g. crossed undulators capable of fast polarization switching [39].

4.2. Undulator technology options for future light sources

An undulator technology's inherent strength and polarization capabilities impact both (a) nominally attainable FEL output capability, including spectral range, polarization, tuning range, and brightness and (b) overall [undulator + accelerator] system design footprint and cost. Furthermore, undulator technology choice also heavily impacts the practical design aspects of (a) field

error control, e.g. where smaller undulator periods are problematic, (b) tunability ease and precision, (c) polarization flexibility, where added complexity provides added capability, and (d) reliable operability, e.g. avoidance of instabilities, device heating, radiation damage, and quenching.

Finally, undulator technology options present tradeoffs in performance risk vs. both enhanced performance and reduced cost. Planar permanent magnet undulators, PM EPU and PM IVIDs are well-tested in storage rings, whereas Cryo-PMs and SCUs are still at or beyond state of the art, require further R&D, and are seen to entail greater risk.

Various SCU designs do in fact have the potential to vastly outperform all other undulator technologies (see Figs. 12, 13 and 14) [40]. Moreover, requirements for future light sources may also alter relative risks insofar as SCUs may prove less problematic in e.g. FEL facilities because of relative inherent ease of (a) field error control with device miniaturization and (b) spectral range and polarization control the macroscopic moving parts required by PM devices. Specifically, PM machining, assembly, shimming, and gap/polarization positional control become more problematic for

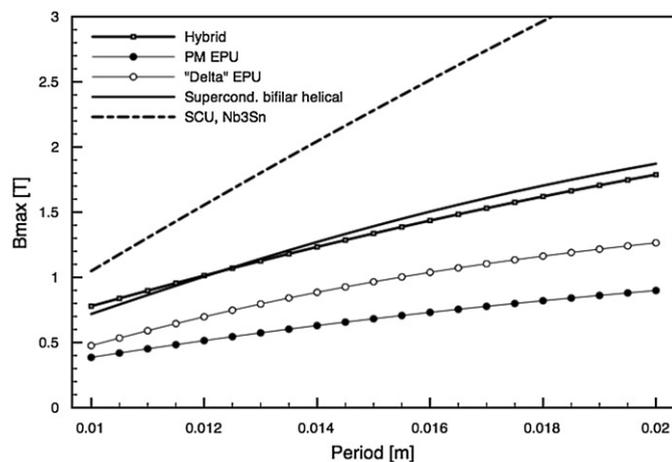


Fig. 12. Performance comparison between undulator technologies: PM-hybrid, APPLE EPU, Delta-EPU, and superconducting bifilar, for a vacuum aperture of 4 mm. Calculations of the bifilar helical SCU data assume an iron-free system with $J_E = 1500 \text{ A/mm}^2$, neglect $J_c(B)$ dependence, and are only reasonably valid for $B < 2 \text{ T}$. PM-based devices assume $Br = 1.35 \text{ T}$. SCU data assume Nb3Sn material. Fields are maximum on-axis values.

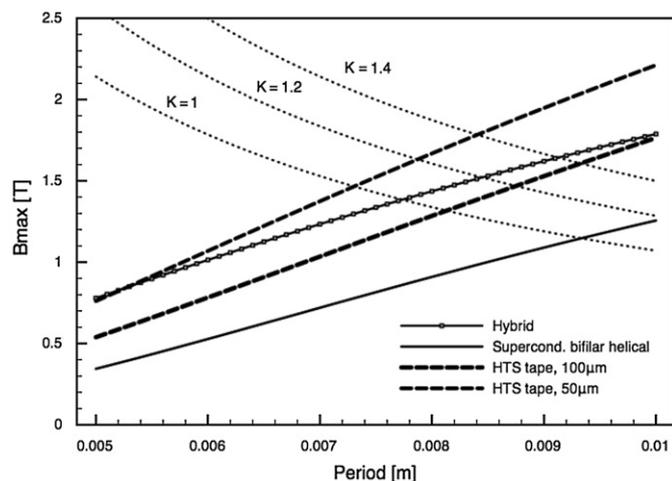


Fig. 13. Performance comparison between in-vacuum version of the PM-hybrid from Fig. 12, the HTS PM-hybrid, and the HTS tape concept. Calculations for two tape thicknesses (50 and 100 μm) are provided, both operating at 4.2 K.

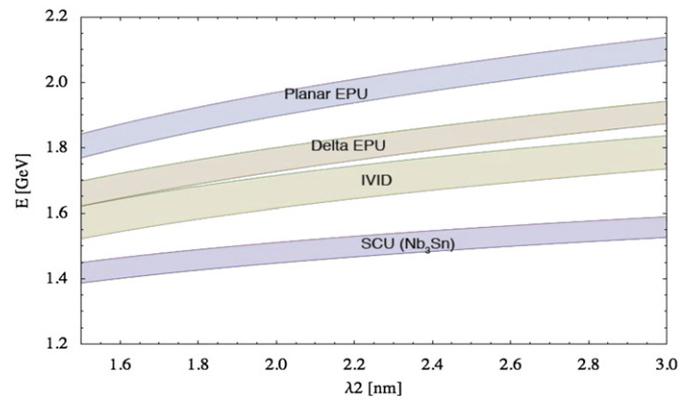


Fig. 14. Performance comparison: Machine electron energy (E) needed to produce radiation in the range $1.5 \text{ nm} < \lambda < 3 \text{ nm}$ for various undulator technologies. Planar modes assumed. Bands represent beam stay-clear apertures of 4–5 mm, illustrating gap-dependence. Band vertical position indicates required E for each undulator technology. For each technology, band width shows E sensitivity to vacuum gap. Band slope shows required E sensitivity to tunability capability.

smaller gaps and periods. Also, beam stay-clear requirements limit how small the gap can become, thus limiting the field strength achievable, e.g. from PM technology. Nb3Sn technology utilizes existing precision winding capabilities and offers variable strength operation with no moving parts. It also offers higher field strength, as compared with PM technologies, for the same beam stay-clear aperture. The high temperature superconductor (HTS) concept [40] utilizes accurate existing micromachining capability, and offers ease of assembly, and low device cost, all important for large-scale FEL applications.

R&D resources thus far invested in SCUs are an order of magnitude less than that devoted to IVIDs before their acceptance in third generation synchrotron facilities, and are significantly less than has been devoted recently to cryogenic permanent magnet systems in Europe, Japan and even the US [31,32,41–43]. Though the family of non-cryogenic PM devices is already poised to serve as baseline design options for future light sources, it would nonetheless be prudent to expedite SCU R&D to enable ultimate performance potential of future light sources.

4.3. Principal superconducting undulator (SCU) development challenges and readiness

Numerous groups are engaged in R&D aimed at overcoming practical limitations of other advanced undulator options, e.g. cryogenic-PMs [31,32]; issues for such devices include phase-shift/shake as a function of (a) gap variation, i.e. change in magnetic force state coupled with the mechanical structure (this is particularly true for EPU), and (b) temperature, in particular for cryogenic devices. Another worthy goal is to ready the very highest performance capability devices, namely the SCU family of undulators, for implementation into future light source plans and designs.

Many key SCU developmental issues have already been addressed. Preliminary readiness has been demonstrated in various SCU prototypes [33,44–48], including specifically demonstration of (a) highest-performing strength capability of all candidate undulator technologies, (b) tuning strength capability technique for phase error correction, (c) in-situ cryogenic tuning control for maintaining phase, (d) attaining near ($> 90\%$) short-sample fields in Nb3Sn undulators, (e) winding, fabrication, and assembly of Nb3Sn devices, and (f) development in industry of thin ceramic insulators with adequate coverage and insulator thickness quality control for long-length conductors [49].

Furthermore, conceptual and prototype designs have already been developed for (a) superconducting elliptically polarizing undulators (SC-EPU) [50], (b) stacked HTS undulators [40], (c) micro-undulators [40], and (d) helical SCUs [36,37], all aimed to meet specific needs for ultimate performance capability at future light source applications.

Principal remaining R&D issues that should be addressed early so as to have the most favorable impact on future light source system design, cost, and ultimate performance capability include (a) fabrication method details, including coil winding and treatment, of various SCU design types, (b) vacuum and wake-field design accommodation and heating accommodation for various operating environments, (c) specifics of in-situ cryogenic field tuning and manipulation, and (d) cryogenic magnetic measurements.

4.4. Undulator R&D tasks

A listing of R&D tasks needed to be ready to incorporate the highest performance devices in a future light source facility is given in the following paragraph. Priority should be given to those that closely match the particular needs and result in optimal performance of a proposed future light source.

Reliable winding and potting processes have been demonstrated for NbTi [42] but they, along with reliable reaction processes, remain to be fully demonstrated for Nb3Sn-based planar and bifilar helical SCUs (Figs. 15 and 16). An in-situ trajectory correction method remains to be honed, and a cryogenic magnetic measurement system needs to be developed.

For stacked high temperature superconductor (HTS) undulators it remains to (a) demonstrate attainment of effective current density (J), (b) evaluate image-current issues, (c) determine field quality and trajectory drivers, (d) verify current path accuracy, i.e.

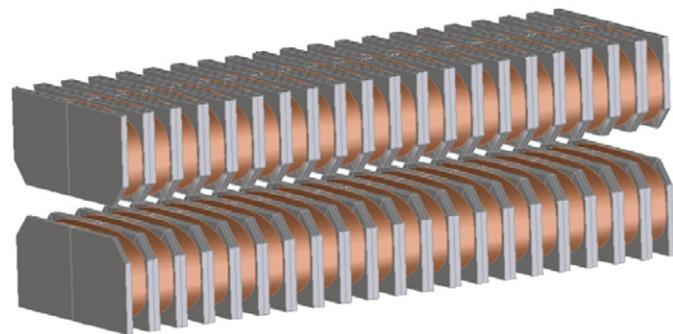


Fig. 15. Nb3Sn SCUs are thermally robust, and outperform all other technologies in the 10–20 mm period range, gap > 3 mm.

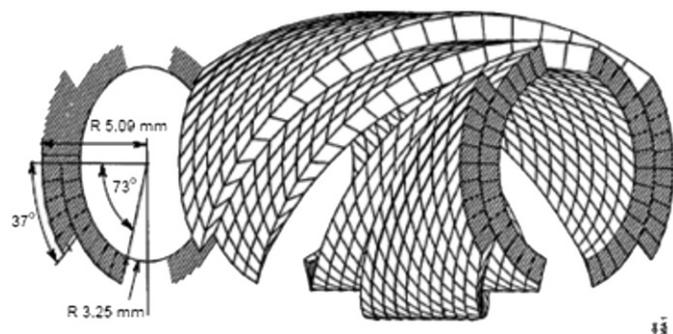


Fig. 16. A bifilar helical superconducting undulator would enable a shorter gain length and thus shorter FEL undulator length.

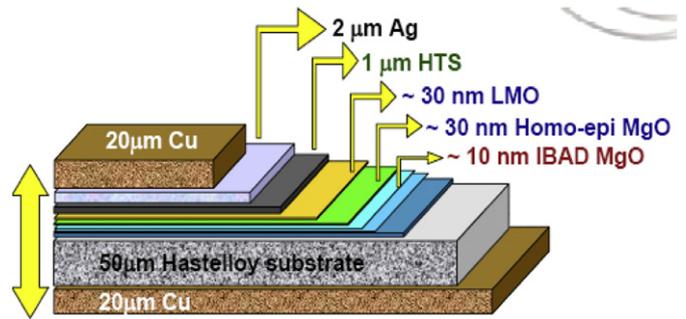


Fig. 17. Diagram (not to scale) of SuperPower Inc.'s YBCO tape. The material can be purchased with or without the Cu cladding. Similar conductors are available from other vendors.

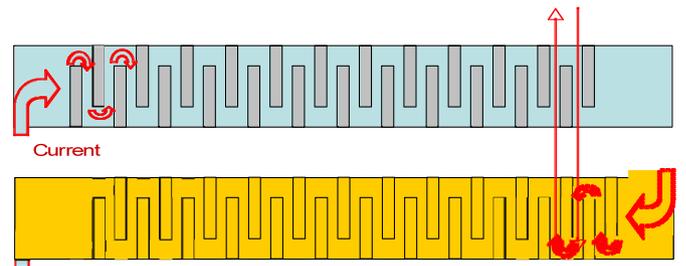


Fig. 18. HTS tape undulator concept. Laser cuts (in gray, not to scale) define the current path by destroying the YBCO superconducting layer in a thin strip of material, without impacting the underlying Hastelloy substrate. The current flows from left to right in the first layer (top); the current transitions to the next layer (bottom) on the right. The cuts are aligned to produce additive magnetic fields as the current flows back to the left.

the $J(x,y)$ distribution, (d) qualify an accurate stacking technique, and (e) develop field correction methods, e.g. use of an outer layer devoted to field correction [46] (Figs. 17,18).

For the Stacked HTS Micro-undulator it remains to (a) demonstrate ability to micro fabricate $5\mu\text{m}$ stacked layers, (b) demonstrate attainment of effective current density (J), and (c) evaluate image-current issues.

For SC-EPU designs it remains to (a) develop an integrated switch network and (b) demonstrate performance in a prototype (Fig. 19).

For FEL/ERL long-undulators it remains to develop fast shifters/chicanes between FEL sections. Other candidate design types can also be considered including cryo-PM undulators and microwave undulators. The R&D issues associated with these devices are not treated here.

5. Summary and conclusions

5.1. Attosecond instrumentation

Developments in attosecond instrumentation will primarily benefit future FELs. For measurement of electron and photon bunch length, techniques that can vastly exceed the performance of current streak cameras is required. For some of the techniques under consideration, technical advances could be applied to both electron and photon diagnostics and therefore, it may be advantageous to embark on a coordinated research program. Measurements at FLASH indicate shot to shot variations in longitudinal profile that could impact certain experiments. Since future X-ray sources will probably exhibit similar behavior, an

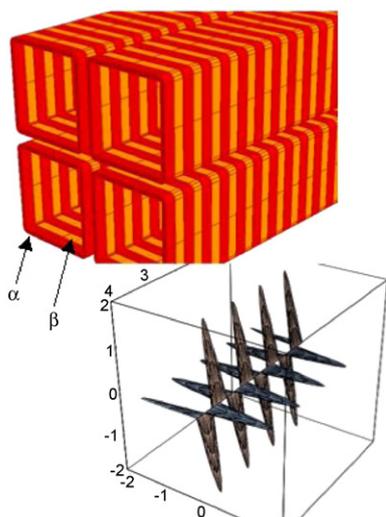


Fig. 19. Two interlaced sets of four-quadrant coil arrays (see top sketch), operated with four power supplies, can provide full variable linear and elliptic polarization control. As an example, the bottom figure maps the α and β fields for a case where the two planar fields are defined to be orthogonal, yielding variable elliptic polarization by varying the relative field strengths.

emphasis on single shot, nondestructive measurements would prove most beneficial.

Today, 10-fs jitter level synchronization over several hundreds of meters is possible using the techniques described above and has been demonstrated over several days of continued operation in the laboratory in operation. Prototype systems that can be deployed in facilities are in development and will be used over the next few years. Continued development over a time span of 5, 10 and 15 years is necessary to push the precision to the femtosecond and finally sub-femtosecond level over distances of up to 10km and more and to go from laboratory systems to fully engineered systems deployable in accelerator and light source facilities. Such predictions would have been laughed at 10 and even 5 years ago. The progress in frequency metrology over the last 10 years shows clear pathways of how such precision might be achievable, and dedicated R&D programs could bring it to fruition.

In preparation for future experiments with intense, ultrashort pulses of soft and hard X-rays, optical components will require development on several fronts. From experiments at HHG sources and the FLASH facility, there is some experience in the EUV region, and this must be extended into the X-ray region. Theoretical and experimental studies of distortion and damage effects should be pursued for several types of optical components, include various types of zone plates and other diffractive structures such as uniformly redundant arrays. As photon pulses approach the attosecond regime, and photon energy increases, mirrors must support a very broad spectral pass band. Improvements in mirror technology would benefit virtually all experiments at future short pulse FELs.

5.2. Detectors

A coordinated R&D program in detector development could increase scientific output from all types of current and future light sources. To identify priorities for this broad program, a workshop similar to that summarized in Ref. [23] should be held. Some technical challenges cut across several light source types and may be more efficiently addressed by programs that are independent of particular facilities. Therefore, the synchrotron focus of the original workshop should be expanded to also include applications at FELs, ERLs, and additional novel sources. As discuss

above, the detector community requires progress in several areas, including: improvements to US foundries, training of new talent, and developments in detector and window materials. Recent challenges that demanded sophisticated readout electronics have been addressed by piggy-backing on HEP developments. Ultimate detector performance may eventually require high performance integrated circuitry that is designed specifically for light source applications.

5.3. Insertion devices

Key performance metrics of candidate technologies show that various superconducting undulator (SCU) designs have the potential to vastly outperform all other undulator technologies. The unique requirements of future light sources may alter relative risks of candidate technologies, particularly to the advantage of SCU designs featuring no permanent magnets or macroscopic moving parts. Several SCU conductor technologies, including Nb₃Sn-based and HTS-based designs have the potential to significantly enhance undulator performance. In addition, various superconducting designs incorporating special polarization features, including helical SCUs and SC-EPUes could play key roles in customization and optimization of future light sources.

Many key SCU R&D issues have already been addressed in completed SCU prototypes in-house at DOE laboratories and elsewhere. It would be prudent to address remaining development issues of these technologies, as discussed herein, so as to enable technology readiness for maximizing ultimate performance and low cost, with manageable risk for future light sources.

References

- [1] G. Berden, et al., Phys. Rev. Lett. 99 (2007) 164801.
- [2] J. Van Tilborg, et al., Phys. Rev. Lett. 96 (2006) 014801.
- [3] M. Röhrs, C. Gerth, H. Schlarb, B. Schmidt, P. Schmüser, Phys. Rev. Spec. Top.—Accel. and Beams 12 (2009) 050704.
- [4] O. Grimm, in: The Proceedings of Particle Accelerator Conference, PAC 07, Albuquerque, New Mexico, 25–29 June 2007, p. 2653.
- [5] M. Zolotarev, G. Stupakov, in: Proceedings of the Particle Accelerator Conference, Vancouver, BC, Canada, (1997 IEEE, New York), 1997, p. 2180.
- [6] P. Catravas, et al., Phys. Rev. Lett. 82 (1999) 5261.
- [7] F. Sannibale, G.V. Stupakov, M.S. Zolotarev, D. Filippetto, L. Jägerhofer, Phys. Rev. Spec. Top.—Accel. and Beams 12 (2009) 032801.
- [8] J. Feng, H.J. Shin, J. Nasiatka, W. Wan, A. Young, G. Huang, A. Comin, J.M. Byrd, H. Padmore, Appl. Phys. Lett. 91 (2007) 134102.
- [9] J. Qiang, J.M. Byrd, J. Feng, G. Huang, Nucl. Instr. and Meth. A 598 (2009) 465.
- [10] U. Fruhling, et al., Nature Photonics 9 (2009) 523.
- [11] E. Goulielmakis, M. Uiberacker, R. Kienberger, A. Baltuska, V. Yakovlev, A. Scrinzi, Th. Westerwalbesloh, U. Kleineberg, U. Heinzmann, M. Drescher, F. Krausz, (27 August 2004) Science 305 (5688) (2004) 1267.
- [12] M. Yabashi, et al., Phys. Rev. Lett. 97 (2006) 084802.
- [13] J. Kim, et al., Large-scale timing distribution and rf-synchronization for FEL facilities, FEL 04, Trieste, Italy, September 2004.
- [14] R. Wilcox, J.M. Byrd, L. Doolittle, G. Huang, J. Staples, Opt. Lett. 34 (2009) 3050.
- [15] R.B. Wilcox, J.W. Staples, Systems design concepts for optical synchronization in accelerators, PAC 07, Albuquerque, New Mexico, USA, June 2007.
- [16] J. Kim, J.A. Cox, J.J. Chen, F.X. Krtner, Nature Photonics 2 (2008) 733.
- [17] J. Chen, J.W. Sickler, E.P. Ippen, F.X. Krtner, Opt. Lett. 32 (2007) 1566.
- [18] J. Kim, J. Chen, J. Cox, F.X. Krtner, Opt. Lett. 32 (2007) 3519.
- [19] J. Kim, et al., Opt. Lett. 32 (2007) 1044.
- [20] E.N. Ivanov, S.A. Diddams, L. Hollberg, IEEE Trans. UFFC 54 (2007) 736.
- [21] J. Kim, F.X. Krtner, F. Ludwig, Opt. Lett. 31 (2006) 3659.
- [22] F. Löhler, V. Arsov, M. Felber, K. Hacker, W. Jalmuzna, B. Lorbeer, F. Ludwig, K.-H. Matthiesen, H. Schlarb, B. Schmidt, P. Schmüser, S. Schulz, J. Szewinski, A. Winter, J. Zemella, Electron bunch timing with femtosecond precision in a superconducting free-electron laser, Phys. Rev. Lett. 104 (14) (2010) 144801, doi:10.1103/PhysRevLett.104.144801.
- [23] A. Thompson, et al., A program in detector development for the U.S. synchrotron radiation community, White paper based on Workshop in Washington, DC, 30–31 October 2000. (Full text available at: <http://www.osti.gov/bridge/servlets/purl/787153-XUP8Mj/native/787153.PDF>).
- [24] <http://www-ssrl.slac.stanford.edu/lcls/papers/lcls_experiments_2.pdf>.
- [25] <http://www.xfel.eu/documents/european-xfel-tdr_eng.pdf>.
- [26] Proceedings of Pixel 2008 Workshop, Journal of Instrumentation, IOP.

- [27] J. Feng, H.J. Shin, J.R. Nasiatka, W. Wan, A.T. Young, G. Huang, A. Comin, J. Byrd, H.A. Padmore, *Appl. Phys. Lett.* 91 (2007) 134102.
- [28] R.D. Schlueter, *Wiggler and Undulator Insertion Devices, Synchrotron Radiation Sources: a Primer*, World Scientific, February 2009.
- [29] S. Sasaki, et al., *Nucl. Instr. and Meth. A* 331 (1993) 763.
- [30] S. Yamamoto, et al., *Rev. Sci. Instr.* 63 (1992) 400.
- [31] T. Hara, et al., *Phys. Rev. Spec. Top.—Accel. and Beams* 7 (2004) 050702.
- [32] J. Chavanne, et al., Construction of a cryogenic permanent magnet undulator at the ESRF, in: *Proceedings of EPAC08, Genoa, Italy, 2008*, p. 2243.
- [33] S.O. Prestemon, et al., Design and evaluation of a short period Nb3Sn superconducting undulator prototype, in: *Particle Accelerator Conference 03, Portland, 12–16 May 2003*, pp. 1032–1034.
- [34] A.B. Temnykh, *Phys. Rev. Spec. Top.—Accel. and Beams* 11 (2008) 120702.
- [35] J.B. Bahrdt, et al., in: *Proceedings of 2004 FEL Conference, 2004*, p. 610.
- [36] L.R. Elias, J.M. Madey, *Rev. Sci. Instr.* 50 (11) (1979) 6.
- [37] S. Caspi, et al., High-field strong-focusing undulator designs for X-ray linac coherent light source applications, in: *1995 Particle Accelerator Conference, Dallas, May 1995*.
- [38] S.O. Prestemon, R.D. Schlueter, Undulator options for soft X-ray free electron lasers, in: *31st International FEL Conference, Liverpool, August 2009*.
- [39] Y. Ding, Z. Huang, *Phys. Rev. Special Topics Accel. and Beams* (2008) 030702.
- [40] S.O. Prestemon, et al., High performance short-period undulators using high temperature superconductor tape, in: *2009 PAC Proceedings, 2009*.
- [41] T. Tanaka, T. Hara, R. Tsuru, D. Iwaki, X. Marechal, T. Bizen, T. Seike, H. Kitamura, in: *Proceedings of the 27th International Free Electron Laser Conference, 2005*, p. 370.
- [42] T. Tanabe, J. Ablett, L. Berman, D.A. Harder, S. Hulbert, M. Lehecka, G. Rakowsky, J. Skaritka, A. Deyhim, E. Johnson, J. Kulesza, D. Waterman, in: *Proceedings of the SRI 2006 Conference, 2006*.
- [43] C.A. Kitegi, J. Chavanne, D. Cognie, P. Elleaume, C. Penel, B. Plan, F. Revol, M. Rossat, in: *Proceedings of EPAC 2006, 2006*.
- [44] A. Madur, et al., Superconducting switch concept applied to superconducting undulator phase-error correction, in: *Synchrotron Radiation Instrumentation Conference SRI09, Melbourne, October 2009*.
- [45] R. Schlueter, S. Marks, S. Prestemon, D. Dietderich, Superconducting undulator research at LBNL. *Synchrotron Radiation News*, 17 (1) (2004). <<http://www.informaworld.com/10.1080/08940880408603073>>.
- [46] Y. Ivanyushenkov, et al., Status of R&D on a superconducting undulator for the APS, in: *The Proceedings of PAC09, Vancouver, B.C. 2009*. <<http://www.JACoW.org/P09/PAPERS/MO6PFP078.PDF>>.
- [47] S.H. Kim, C. Doose, R.L. Kustom, E.R. Moog, and I. Vasserman, R&D of short-period NbTi and Nb3Sn superconducting undulators for the APS, in: *The Proceedings of PAC 2005, Knoxville, TN, 2005*, p. 2419.
- [48] D. Wollman, et al., *Phys. Rev. Spec. Top.—Accel. and Beams* 11 (2008).
- [49] SBIR <http://www.sc.doe.gov/sbir/awards_abstracts/sbirsttr/cycle25/phase2/053.htm> Solicitation—under development in industry.
- [50] S. Prestemon, et al., *IEEE Trans. Appl. Supercond.* 16 (2) (2006) 1873.



Cathode R&D for future light sources[☆]

D.H. Dowell^{a,*}, I. Bazarov^b, B. Dunham^b, K. Harkay^c, C. Hernandez-Garcia^d, R. Legg^e,
H. Padmore^f, T. Rao^g, J. Smedley^g, W. Wan^f

^a SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025, USA

^b Cornell University, Cornell Laboratory for Accelerator-Based Sciences and Education (CLASSE) Wilson Laboratory, Cornell University, Ithaca, NY 14853, USA

^c Argonne National Laboratory, 9700 S. Cass Avenue, Argonne, IL 60439, USA

^d Thomas Jefferson Laboratory, 12000 Jefferson Ave, Free Electron Laser Suite 19 Newport News, VA 23606, USA

^e University of Wisconsin, SRC, 3731 Schneider Dr., Stoughton, WI 53589, USA

^f Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley, CA 94720, USA

^g Brookhaven National Laboratory, 20 Technology Street, Bldg. 535B, Brookhaven National Laboratory Upton, NY 11973, USA

ARTICLE INFO

Available online 18 March 2010

Keywords:

Cathode research
Electron source
Photocathodes
Quantum efficiency
Thermal emittance
High average current
Energy recovery linacs

ABSTRACT

This paper reviews the requirements and current status of cathodes for accelerator applications, and proposes a research and development plan for advancing cathode technology. Accelerator cathodes need to have long operational lifetimes and produce electron beams with a very low emittance. The two principal emission processes to be considered are thermionic and photoemission with the photocathodes being further subdivided into metal and semi-conductors. Field emission cathodes are not included in this analysis. The thermal emittance is derived and the formulas used to compare the various cathode materials. To date, there is no cathode which provides all the requirements needed for the proposed future light sources. Therefore a three part research plan is described to develop cathodes for these future light source applications.

Published by Elsevier B.V.

1. Introduction

The development of the photocathode gun has become a significant enabling technology for X-ray free electron lasers and other 4th generation light sources. As the first X-ray FEL user facility, the performance of LCLS is impressive, lasing 10-orders of magnitude higher in peak energy than previous X-ray light sources [1]. And there are opportunities for improving even this performance. The emission processes of the cathodes used in the LCLS gun are not completely understood. The quantum efficiency needs to be made reliable and the low-charge, thermal emittance is nearly a factor of two higher than given by theory. In addition, it operates at a low repetition rate (120 Hz) and it is anticipated that future applications will require repetition rates of 100 kHz and higher with CW operation. Therefore a principal technical challenge for ERL's as well as for other high repetition rate light sources will be the production of LCLS-like beams in a lower peak field but high average power gun producing up to 100 mA of average current. The combination of high average current and ultra-low emittances required by the ERL and X-ray FEL oscillator

has never been achieved in a CW gun. An area requiring significant support is photocathode R&D since there are presently no cathodes meeting the known requirements. Thus there is a strong motivation for two overlapping lines of cathode R&D: one of cathodes for low-repetition rate and ultra low emittance guns like LCLS, and a second of cathodes for high-average current guns to be used in ERL's and other CW applications.

2. Drive laser and cathode requirements

The operating range of the injector and the corresponding drive laser system can be divided into three distinct regimes: $< 1 \mu\text{A}$, $1 \mu\text{A} - 1 \text{ mA}$ and $> 1 \text{ mA}$. The cathode and drive laser requirements are presented in Fig. 1 where the average drive laser power is given vs. the cathode quantum efficiency (QE). Lines of constant average current are plotted in the log-log graph. Three shaded regions schematically show the QE range for metal, antimonide and Cs:GaAs cathode types along with the vacuum required for them to survive several hours. For the low average current injectors ($< 1 \mu\text{A}$), metal photocathode irradiated by UV laser provide ultra high brightness beams as evidenced by the LCLS. Properly conditioned metal photocathodes such as Mg or Pb, along with a few watts of UV would be able to service the injector in the current range of $1 \mu\text{A}$ to 1 mA, delivering peak brightness comparable to LCLS. For high average current injectors, in order to

[☆] Written for Accelerator Physics of Future Light Sources, A Workshop Sponsored by DOE Basic Energy Sciences, September 15–17, 2009.

* Corresponding author.

E-mail address: dowell@slac.stanford.edu (D.H. Dowell).

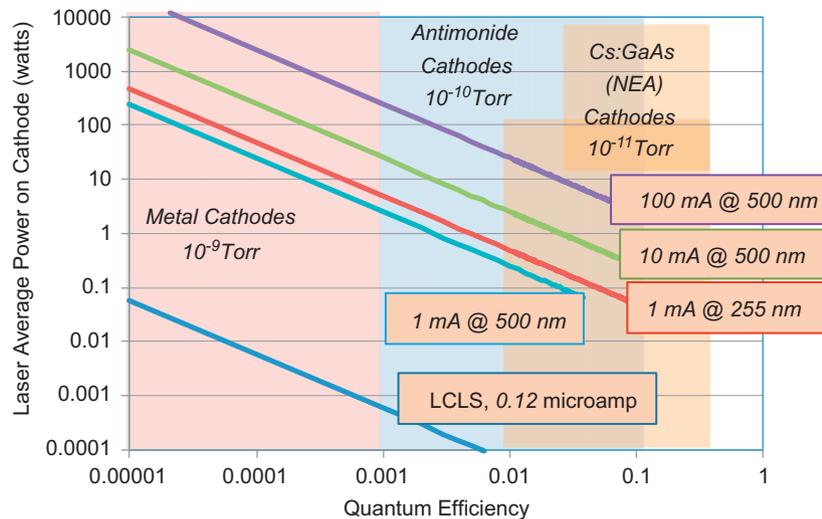


Fig. 1. Plot of the average laser power vs. quantum efficiency to produce various average beam currents. The QE ranges for the general cathode types are shown along with their vacuum requirements.

make the drive laser practical, it is necessary to limit its average power to less than approximately 25 W, operate it at visible or near IR wavelengths and require cathodes with QE's of 1–10%. Although there has been significant progress in the high average power lasers such as diode pumped fiber lasers that deliver up to 100 W in IR, at present, these lasers are generally less reliable, and the stringent beam shaping and stability necessary to produce a bright electron beam would necessitate additional laser R&D. One to ten percent QE's can be reached using Cs:GaAs or K_2CsSb ; however their reliable fabrication and operation at 100 mA in a gun has not been demonstrated and represents a major technical challenge.

Having the drive laser operating at either visible or near-IR wavelengths allows a more equitable sharing of the technical risks between the laser and cathode for high average current injectors. While cathodes at the longer wavelengths are more difficult, the laser challenges at UV wavelengths are greater. Current UV laser systems begin with an IR beam which is then frequency-tripled or quadrupled using non-linear conversion crystals. Good conversion efficiency requires focusing the laser to a small spot in the crystal which can damage the crystal. In addition, shaping the laser pulse is difficult and inefficient at UV wavelengths as are the diagnostics. The option of frequency up conversion followed by power amplification is viable at specific wavelengths, but has not yet been fully investigated. In this approach, the beam shaping can be accomplished at lower average power UV beam that is subsequently amplified. In this scheme, the losses due to the shaping can be compensated by the amplification. The major drawbacks of this path are the UV gain media that are limited to specific wavelengths such as 248 nm and the repetition rate is limited to < 10 kHz. With the current laser technology, the maximum commercially obtainable UV average power is ~ 2 W resulting in a maximum obtainable average current with a metal photocathode to hundreds of microamperes. A UV laser and enhanced metal or CsTe cathodes with ~ 1 –10% QE are viable alternatives for a few mA-tens of mA injectors, especially since excimer lasers operating at multi kHz can be very efficient amplifiers, having very large bandwidth and gain. Other laser options are on the horizon such as diode-pumped, cryogenically cooled Yb:YAG, which show promise for scaling to kilowatt average power with the required stability and beam quality. Due to the superior quantum efficiency (91%) of Yb:YAG, and the possibility of extracting the full stored energy of

the crystal at liquid nitrogen temperature while maintaining the high beam quality, this source can develop average optical power of hundreds of watts using about 3 kW electrical wall-plug power while maintaining the ultra short pulse duration.

Additional complications of using the UV radiation stem from transport optics. Typically, the AR coating of UV lenses and dielectric coating of mirrors are also much more susceptible to damage. For example, the vacuum windows on the laser transport tube for LCLS were prone to damage even at the low average fluencies at LCLS, forcing a re-design of the optics to increase the beam size on the windows to eliminate the damage. In general, UV optics are more sensitive to laser damage and are less efficient requiring an even higher power laser to make up for the losses. The additional laser power can be significant. Optical damage thresholds in these applications are limited typically by the peak power and not the average power. However, since most of the laser transport line would be in vacuum to minimize the beam fluctuation, the damage threshold and air absorption can be minimal.

The desired characteristics for the drive laser are sub-ps stability, micron level position stability, uniform transverse and longitudinal beam profiles are required for cathodes and gun types. There is one distinct difference between low and high average current systems and that is the amount of allowed photocurrent outside a nominal temporal and spatial window of the laser. For example, in a typical RF gun, laser light reaching the cathode ± 20 degRF or more away from the nominal laser-RF phase will produce off energy and different trajectories, potentially producing beam halo with significant average power for a high average current system. Laser-related beam halo is also produced by scattered laser light striking the cathode either at the wrong time or wrong location. The allowed laser-related beam halo is typically 1 part in 10^{-6} of the total beam current.

The desired laser parameters at low average current can be achieved through the use of diode-pumping of the Titanium-Sapphire laser medium [2]. This type of laser system is also very reliable, operating for 18 months with better than 97% uptime. However one technical issue deserving further research is temporal pulse shaping. The desired three dimensional pulse shape has been difficult to attain even with state-of-the-art techniques.

Due to this interdependence of the cathode and laser, there should be parallel laser R&D concentrating on developing reliable

and stable high-repetition rate systems with the capability of pulse shaping in three-dimensions. As noted above, the cathode drive laser is expected to operate at visible to near IR wavelengths and should have limited bandwidth to minimize the production of electron beam instabilities, unless the drive laser is also being used to seed the FEL. Laser pulse shaping allows pre-forming of the electron bunch to maintain linear space charge forces and to manipulate space charge waves.

Beyond the high-QE and its survival in a gun environment, the cathode has to emit a beam of exceptional quality with very little dark current. Recent results for the LCLS gun and the Cornell injector show significant progress in the practical applications of emittance compensation to control space charge effects, and the generation of near perfect RF and magnetic fields to eliminate optical aberrations. Combined with the overall advantages of operating at lower bunch charge, we are now reaching the thermal emittance as the limit to increasing the beam brightness. However, the effective thermal, or “intrinsic”, emittance depends on several effects, including the crystallinity, surface roughness, surface impurities and QE non-uniformity. Thus it is very challenging to measure and combine all these phenomena into a complete and useful physical model. For example, it is relevant to note that the measured thermal emittance from the LCLS gun is nearly twice the theoretical value [3,4] and the source of this difference is not understood. If the thermal emittance had the theoretical value, the already excellent LCLS emittance would be still 20% lower. Therefore the second major challenge for cathodes is to understand the origins of the low charge emittance and its interaction with the space charge forces.

3. The three part cathode R&D plan

In order to address these and other cathode related issues we suggest the following R&D program consisting of three inter-related parts:

1. Studies of optimal cathode formation methods and cathode emission characteristics, using available surface and material diagnostics.
 2. Modeling of cathode emission physics and electron dynamics near the cathode.
 3. Operational testing in the gun and injector system and validating models.
1. *Studies of optimal cathode formation methods and cathode emission characteristics, using available surface and material diagnostics* Optimal performance of a cathode can only be achieved with understanding of the material properties, such as surface and bulk crystallinity, band structure, surface morphology, material optical properties and surface chemistry. Such understanding will provide feedback to allow optimization of growth and processing of cathodes, and will provide performance data that will be used to validate modeling codes and ultimately predict cathode behavior. These techniques should supplement a program including direct measures of cathode performance, such as spectral response, lifetime (both dark and operational) and sensitivity to chemical poisoning by gases typically found in an injector. With the advent of modern user facilities (principally light sources and nanocenters), techniques such as diffraction, photoemission spectroscopy and high resolution imaging are available to explore these material properties. New growth methods should be investigated for the creation of accelerator cathodes, such as atomic layer deposition and molecular beam epitaxy (MBE is already used for GaAs cathodes). For the longer-term, advances in the synthesis of novel materials, nano-engineering in particular,

raises the possibility of designing photocathode materials optimized for specific properties [5] after validation of the design tools based on the data from existing cathodes. Such cathodes should be studied using analytic tools such as Density Functional Theory (DFT) analysis [6] with promising candidates being fabricated and characterized [7].

2. *Modeling of cathode emission physics and electron dynamics near the cathode* The fine details of the emission process need to be included to the electron simulation codes. This should include the physics of the emission process, using models such as Spicer’s three step model of photoemission [8] or the exact one-step model [9]. The results of these models should be used to predict thermal emittance values based on full energy and angular distribution curves of the emitted electrons. Given the improvements made in RF and magnetic optics, and emittance compensation, the next step is for the computational dynamacists to put the physics learnt in the cathode labs into the particle codes. These enhanced codes can then be used to simulate and design the complete injector and be verified in the gun and injector studies part of the R&D plan. Genetic algorithms have already been employed in some areas of electron beamline design [10]; a program to integrate this capability into modeling codes along with a complete emission physics package could lead to much better optimization of cathode, gun and laser properties.
3. *Operational testing in the gun and injector system.* It is of course essential that these lab results and computer studies be tested in an operating gun. Some experiments can be performed in a low duty factor system but will also require testing in a CW gun. Among the current photocathode guns being proposed: DC, NCRF and SCRf, all are viable platforms for cathode testing since each has its own application niche. Some tests of cathode performance, such as thermal emittance, operational lifetime and response time, are best made in an injector.

4. Cathode technology

The semiconductor photocathodes in wide use today as high-brightness electron sources for accelerators derive in large part from cathode R&D performed decades ago. Much of the development work on photocathodes was focused on photoemissive detectors, where the most important criteria are quantum efficiency, reliability, low dark current, spectral response, and response time. The emission distribution, or emittance, was not a high priority, and thus these cathodes were not optimized for ultra-high brightness. In order to meet the requirements for future light sources, a new wave of R&D is needed with collaboration among accelerator physicists and materials and surface scientists.

Thermionic cathodes can deliver thousands of coulombs reliably and have been used in several FELs around the world. To reduce their large cathode emittance though, grid structures must be eliminated, making pulse generation difficult. To offset that problem and to increase the peak charge density that can be extracted from the cathode, a pulsed DC structure is used at SCSS with a CeB₆ cathode. The interesting approach of Spring8 SCSS FEL is to use pulsed HV which can more easily reach 500 kV, use a single crystal thermionic cathode and modulate the beam energy at high frequency with a RF cavity and slice out short bunches using energy slits in a magnetic chicane. The CeB₆ gun has been very reliable and successfully delivered stable 500-keV beams to the SCSS test accelerator for three years, and is now operating for various EUV-FEL user experiments [11].

Metal photocathodes are predominately used in high gradient RF guns, in particular the s-band BNL/SLAC/UCLA gun. Improvements in this gun design were incorporated into the LCLS gun to produce the high brightness beam for LCLS [12]. The relatively low thermal emittance of metals contributes to the low gun emittance, especially at low charge. Metal cathodes are more tolerant to vacuum contamination and unlike other photocathodes can be transferred and installed at atmospheric pressures and thus do not require a load-lock. The disadvantage of the metal cathode is its low QE and need for a UV drive laser, which limits these cathodes for applications requiring ≈ 1 mA average current. Fortunately recent advances in laser technology have greatly improved the reliability of these lasers through the use of diode pumping of the gain medium. Therefore, while costly, fully integrated laser systems are commercially available.

Photoemission and thermionic cathodes are currently being used in ERL-based FELs. The Energy Recovery Linac-based FELs at Jefferson Laboratory in the US and at Daresbury Laboratory in the UK use a Cs:GaAs photocathode in a DC gun illuminated with laser pulses at ~ 532 nm [13,14], while the BINP FEL/ERL and the HEPL Recyclotron used a thermionic cathode [15,16]. To date, no other type of cathode has delivered beam for an ERL-based machine. The JLab FEL DC gun delivered over 900 h and 7000 C at 2–9 mA CW from a single GaAs wafer between 2004 and 2007 with a lifetime of 550 C or 30 h at an average current of 5 mA CW [17]. In 1992 the Boeing normal conducting RF (NCRF) gun demonstrated 32 mA with a K_2CsSb photocathode and still holds the record for the highest average current from a photo cathode gun [18]. Cs_2Te cathodes have been in operation for 120 continuous days in a normal conducting RF gun at PITZ with minimal QE degradation [19].

5. Materials science analysis and modeling of cathodes

Numerous material analysis tools are available to assist in cathode development. These can be broken into three broad classes—those that analyze the structure of the cathode, those that analyze the chemical makeup of the cathode and any contaminants and those that evaluate its function as an electron emitter. For structural analysis, surface imaging techniques such as atomic force microscopy and scanning electron microscopy (SEM) provide surface roughness values; the SEM can also provide spatially imaged chemical data via energy dispersive X-ray spectroscopy (EDS) and crystalline makeup of the cathode via electron backscattered diffraction (EBSD). Other electron diffraction techniques provide surface crystalline information, including local reconstruction due to surface termination (such as hydrogen on diamond). X-ray diffraction (XRD) is capable of providing crystalline information on both the surface and the bulk (by varying angle of incidence). XRD can be used to determine the grain size of grown cathodes (alkali antimonides and tellurides), and this can in turn be used to optimize the growth parameters to improve grain size and orientation. Diffraction imaging techniques such as X-ray topography can “see” strain in crystalline cathodes, possibly providing insight into damage caused by ion bombardment in GaAs.

Chemical analysis techniques include X-ray photoemission spectroscopy (XPS), secondary ion mass spectroscopy (SIMS) and X-ray absorption spectroscopy (XAS). These techniques can provide feedback to the growth process of grown cathodes and can provide data on adsorbed contaminants on all cathodes [20]. SIMS is capable of providing a depth profile of the chemical makeup of a cathode, allowing variations in the cathode make-up to be observed. XAS and XPS are sensitive to chemical bonding in

addition to elemental make-up, and can be used to distinguish similar chemical forms (K_2CsSb and KCs_2Sb , for example).

The function of a cathode as an electron emitter can be evaluated using a variety of photoemission spectroscopy techniques, including photoemission electron microscopy (PEEM) and angle-resolved photoemission spectroscopy (ARPES). The PEEM allows the emitted electrons to be spatially imaged, uncovering variation in material work function and spatial non-uniformity. ARPES provides the energy and angular distribution of the emitted electrons. Together, these tools provide the spatial, angular and energy distribution of the beam. This represents all of the data necessary to determine the initial phase space volume of the beam from the cathode—the “thermal” emittance. Density Functional Theory (DFT) is a fully quantum mechanical approach for solving the electronic structure of solid surfaces. Many contributed to the idea of using density as the basic variable for the description of the energies of electronic systems. Kohn and Sham [21] demonstrated that the electron density of a fully interacting system could be rigorously obtained from a simple one-electron theory. Much current understanding of metal surfaces comes from using the simplest DFT approximation of plane-wave pseudopotentials within the local density approximation (LDA). Other approximations are suitable for studying strongly correlated systems [22].

DFT analysis has been used to compute the work function of various metal crystals and the agreement with experimental values is reasonable (within 10%) [23]. The surface bands computed in this model give the highest-energy partially occupied bands that fall below the Fermi level, and the electrons in this “Fermi pool” have a bounded surface-parallel momentum, k_{max} (i.e., the transverse momentum in the accelerator physics convention). The laser energy determines what fraction of the Fermi pool can be photoemitted. Most notably, the measured angular distribution of photoelectrons has been shown to correspond with the calculated k_{max} , e.g., see Ref. [24]. These results suggest that DFT analysis or other analytical methods are promising tools in studying candidate ultralow emittance photocathodes, both single crystals and more complex structures. This is an area that requires R&D.

Initial investigations were made of MgO monolayers on Ag, a well-studied material in catalysis [25]. DFT computations suggest that the surface-parallel momenta in the surface band for this system are well limited, and the corresponding emittance can potentially be reduced below 0.1 mm-mrad [7,26] Furthermore, thin oxide films induce a significant change in the work function [27]. For the MgO monolayers on Ag, a reduction in the work function of > 1 eV relative to Ag is both computed and observed [28]. A possible practical device based on this material or a similar principle should be developed and studied.

The Spicer 3 step model is widely used to compute the QE of a cathode and seems to do a reasonable job for normal incidence light. However, even in this case there are ambiguities that can significantly affect the results. The most important perhaps is the fact that many real cathodes are polycrystalline, but have in reality a preferred crystallographic texture. Evaporation of thin films onto non-comensurate substrates often leads to this effect; for example, Al on glass has a $\langle 111 \rangle$ texture with only a few degree variation from the surface normal. In the case of the LCLS Cu cathode for example, micro-XRD has shown that the surface consists of an equal mixture of 111 and 110 grains. At the LCLS injector photon energy and field, the 111 grains would emit roughly 30 times less than the 110 grains, and so modeling really ought to take into account the statistical distribution of grains and corresponding work functions. A more intrinsic deviation from the Spicer model is seen when using p-polarized light off normal incidence. Recent work on Cu(111) has shown that 0.5 eV above

the work function, the p-polarized QE peaks at around 70° off-normal incidence, at 14 times the normal incidence yield [29]. Earlier work on Cu showed the same effect in polycrystalline Cu [30]. The same effect has been observed in Al, and in Mo, the enhancement is around 40. In recent work on annealed and ion-damaged Cu, the effect can clearly be associated with sharpness of the metal–vacuum interface. Such effects are qualitatively predicted from theory [31] when accurate models of the surface potential are taken into account, but so far there is no universal predictive model. The general point is that the very rapid change in electric potential from outside to inside the surface causes sufficient uncertainty in electron momentum that many more initial and final states can be coupled, thus increasing yield. We need to advance to a point where details of the electron structure, electron transport and emission are all taken into account within a self consistent framework so that these complicated phenomena can be completely understood.

6. The theoretical thermal emittance

In order to compare the various cathode types it is necessary to first define the thermal emittance for each emission process. If the electrons from a cathode are assumed to have no correlation between location of emission and the emission angle then the normalized thermal emittance per unit beam size, ε_n/σ_x , with units of microns/mm(rms) can be written as

$$\frac{\varepsilon_n}{\sigma_x} = \frac{\sqrt{\langle p_x^2 \rangle}}{mc} \quad (1)$$

Here σ_x and p_x are the rms transverse beam size and momentum, respectively. The rms momentum is obtained from the electron distributions (the electron density of states) for each of the emission processes and reflects the electronic properties for that emitter.

The electron distribution for a thermionic emitter is given by the Maxwell–Boltzmann distribution which leads to the well-known thermal emittance in terms of the electron temperature

$$\frac{\varepsilon_{th,n}}{\sigma_x} = \sqrt{\frac{k_B T}{mc^2}} \quad (2)$$

For photoemission from a metal, the electron distribution is assumed to be Fermi–Dirac distribution at zero temperature convoluted with a uniform density of states. In this case, the emittance is given in terms of the effective work function and the photon energy [4]

$$\frac{\varepsilon_{pe,n}}{\sigma_x} = \sqrt{\frac{\hbar\omega - \phi_{eff}}{3mc^2}} \quad (3)$$

The effective work function includes the effect of Schottky reduction of the barrier in the presence of an applied electric field, E_a

$$\phi_{eff} = \phi_W - \phi_{Schottky} = \phi_W - e\sqrt{\frac{eE_a}{4\pi\varepsilon_0}} \quad (4)$$

For the comparison purposes of this paper the Schottky work function, $\phi_{Schottky}$, is assumed to be zero.

At this point it is necessary to discuss an important approximation leading to the simple form of Eq. (3). The emittance derivation involves angular and energy integrations constrained by the surface boundary condition conserving the transverse momentum across the cathode–vacuum boundary. Eq. (3) is simple because the initial state electrons are assumed to be in s-wave states with an energy distribution given by the Fermi–Dirac function. The s-wave assumption gives a simple isotropic

angular distribution and the Fermi–Dirac function for zero temperature electrons (a very good approximation at 300 K) becomes the Heaviside step function. The combination leads to Eq. (3). However in general, the electron density of states will be more complicated, involving states with higher angular momentum such as d- and p-wave states oriented by the crystalline planes and having a structured energy distribution. This complication is significant in non-ideal electron gas metals such as lead [32]. It is also relevant in the interpretation and conversion of the electron energy distribution curves (EDC's) measured in a laboratory surface science chamber to the normalized photoelectric (thermal) emittance. Simply measuring the energy spectrum does not provide enough information, therefore determining the emittance requires knowing the correlated angular-energy distributions. Hence angular resolved photoelectron spectra (ARPES) at the operating photon energies will be necessary. These same comments also apply to the emittance analysis of semiconductors.

For semiconductors it is necessary to consider prompt and delayed emission separately. In prompt emission, the emittance is assumed to be determined by the electrons' excess energy in the vacuum. For example, the excess energy (ignoring the Schottky work function) of a metal is

$$E_{excess,metal} = \hbar\omega - \phi_W \quad (5)$$

A simple semiconductor with a band gap energy of E_G and an electron affinity of E_A is shown in Fig. 2. If we assume most of the excited electrons come from the valence band, $E_G + E_A$ correspond to the material work function described above and the excess energy for a semi-conductor becomes

$$E_{excess,semi} = \hbar\omega - E_G - E_A \quad (6)$$

Thus it follows that the emittance for emission from a semiconductor can be approximated by

$$\frac{\varepsilon_{semi,n}}{\sigma_x} = \sqrt{\frac{\hbar\omega - E_G - E_A}{3mc^2}} \quad (7)$$

In delayed emission the excited electrons have time to thermally equilibrate with the lattice. Hence a special situation exists for delayed photoemission from semiconductor cathodes, especially negative electron affinity (NEA) cathodes such as Cs:GaAs. In these cathodes the excited electrons easily scatter with the lattice phonons, reaching thermal equilibrium with the ambient temperature phonons before escaping. Since the electrons are all thermal, the expression for thermionic emission should be used. Thus for Cs:GaAs cathodes, one should use the emittance formula for thermionic emission corresponding to the ambient temperature of the lattice

$$\frac{\varepsilon_{GaAs,n}}{\sigma_x} = \sqrt{\frac{k_B T}{mc^2}} \quad (8)$$

It is important to note that the expression for the cathode emittance given by Eq. (8) only applies for emission with low energy photons, near 880 nm for Cs:GaAs. Emission with higher energy photons will lead to a mixture of prompt and delayed emission in which both Eqs. (7) and (8) apply. Because it is uncertain to know this mixture, for consistency Eq. (7) will be used to compute the thermal emittance of all semi-conductor

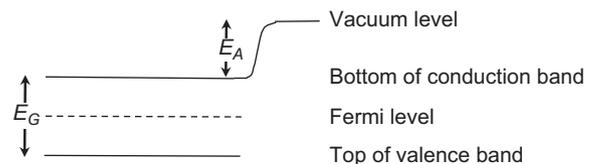


Fig. 2. The energy levels of a simple semiconductor [33].

cathodes, including GaAs, in Table 3. In addition, the comments made earlier concerning how to interpret the EDC of metals in terms of the thermal emittance also apply to semi-conductor cathodes. The calculation for semi-conductors is complicated by the addition of electron–phonon scattering and the presence of the electron affinity energy level, both of which can be ignored in metallic photoemission. Electrons which leave from the bottom of the conduction band have energy with respect to vacuum equal to the magnitude of the electron affinity. Since they are likely to scatter elastically during emission, their momentum is not always normal to the surface. This effect is expected to be the primary source of thermal emittance from the diamond amplifier (see below), and it will be non-negligible for GaAs as well. The thermal emittance for photoemission using EDC's has been obtained for Cs₂Te [35] and GaAs [36].

7. Description of cathode properties

The above definitions will be used to estimate the emittances in the tables below. That is, Eq. (2) for the thermionic emittance, Eq. (3) for photoemission from metals, Eq. (7) for prompt photoemission from a semiconductor.

7.1. Thermionic cathodes

Table 1 gives the emission properties for CeB₆ which is used in the SCSS pulse high voltage gun. The thermionic emittance, Eq. (2), has been used to compute the emittance from the temperature. Since a thermionic cathode naturally produces a DC beam, it is necessary to determine how long a bunch can be fit into the longitudinal acceptance of a RF cavity. The bunch length needed for a desired bunch charge, Q , is estimated by

$$t_{\text{bunch}} = \frac{Q}{4\pi\sigma_x^2 J_{\text{thermal}}} = \frac{Q}{\pi R_c^2 J_{\text{thermal}}} \quad (9)$$

Using the surface current density, J_{thermal} , for CeB₆ and the emission size listed in Table 1 gives a bunch length needed for 250 pC as 84 ps. The 84 ps bunch can be sliced from the DC beam by producing an energy chirp and then sending it through a chicane to bunch and slice out the bunch with energy slits. The long bunch is then velocity compressed before injection into higher frequency accelerator sections. This is the scheme used for the SCSS FEL.

7.2. Metal photocathodes

Metal photocathodes are commonly used in high gradient, high frequency RF guns and are the mainstay of the BNL/SLAC/UCLA s-band gun. The technological descendent of this device, the LCLS gun, has produced the bright beam needed for the first hard X-ray FEL. Due to the high work function UV photons are needed for reasonable QE, which makes them impractical for high average current applications such as ERLs. However, they are the most robust of all the photoemitters and can survive for years at the high cathode fields required to produce a high brightness beam. The current copper cathode installed in the LCLS gun has operated

nearly continuously as the electron source for the X-ray FEL for over a year.

An interesting and significant increase in the QE occurs when the metal is coated with just a few angstroms of CsBr. Although CsBr alone has a large band gap energy, the metal acts as a reservoir of electrons which can be excited into new states formed by color centers at the interface between the two materials. Experiments show the QE can be enhanced by a factor of 50-times in copper and 350-times in niobium [37,38]. Since the CsBr coating is a few to tens of angstroms thick, it is much thinner than the coherence length of the Cooper pairs in a superconductor. Thus niobium with a CsBr coating would retain its superconductivity and one would effectively have a superconducting photocathode. This could significantly simplify present SCRF gun designs by eliminating the need for thermal isolation between a warm cathode and the cryogenic RF cavity. A QE amplification of 350-times for niobium is enough for initial testing with CW RF at low charge per bunch. Further research may lead to high QE superconducting cathodes which would greatly simplify future SCRF gun designs (Table 2).

7.3. Semiconductor cathodes

Table 3 lists properties of many of the known semi-conductor materials which are possible candidates for study in the cathode R&D plan. In all cases the thermal emittances have been computed using Eq. (7) combined with the photon, gap and electron affinity energies given in the table.

Besides having good QE and low thermal emittance, the ideal photocathode should also have low thermionic emission. While the thermionic and photoelectric work functions are the same in metals, they differ in semiconductors. The thermionic work function is the energy difference between the Fermi and vacuum levels. Referring to Fig. 2, it can be seen that the thermionic work function for semiconductors is then $1/2(E_G + E_A)$. One could apply this relation to the cathodes listed in Table 3 to form an estimate of the thermionic emission, however it is better to use experimental values instead, since the emission is dependent upon many other factors such as defects and surface condition. Sommer also lists the thermionic emission at room temperature for some of the materials shown in Table 3. In particular he notes that the K₂CsSb has one of the lowest emissions of the bi-alkali cathodes, 10^{-11} μA/cm² [42], adding to its suitability as a photocathode for use in RF and DC guns.

7.4. NEA cathodes for ERL's

A large variety of photocathodes are employed for production of bright electrons: from metallic cathodes (Mg, Cu, Nb, and Pb) typical of RF and superconducting RF guns [41,32,39] to high quantum efficiency alkali-antimonide and multi-alkali cathodes (Cs₂Te and K₂CsSb) [46] as well as III–V semiconductor photocathodes activated to negative electron affinity [47,48]. For the production of high average current beams as required for ERLs only high QE cathodes are practical, with those having a good response in the visible being preferred to keep the requirements on the laser system realistic (typically frequency doubled high

Table 1
Properties of the SCSS thermionic cathode.

Thermionic cathodes	Typical temperature, T (°K), $k_B T$ (eV)	Emission radius (mm)	Surface current density (A/cm ²)	Work function, ϕ_w (eV)	Thermal emittance (microns/mm(rms))
CeB ₆ Single crystal	1723 K, 0.15 eV	1.5	42	2.3	0.54

Table 2
Properties of metal photocathodes.

Metal cathodes	Wavelength & energy: λ_{opt} (nm), $h\omega$ (eV)	Quantum efficiency (electrons per photon)	Vacuum for 1000 h operation (Torr)	Work function, ϕ_w (eV)	Thermal emittance (microns/mm(rms))	
					Eq. (3)	Expt.
Bare metal						
Cu	250, 4.96	1.4×10^{-4}	10^{-9}	4.6 [34]	0.5	1.0 ± 0.1 [39] 1.2 ± 0.2 [40] 0.9 ± 0.05 [3]
Mg	266, 4.66	6.4×10^{-4}	10^{-10}	3.6 [41]	0.8	0.4 ± 0.1 [41]
Pb	250, 4.96	6.9×10^{-4}	10^{-9}	4.0 [34]	0.8	?
Nb	250, 4.96	$\sim 2 \times 10^{-5}$	10^{-10}	4.38 [34]	0.6	?
Coated metal						
CsBr:Cu	250, 4.96	7×10^{-3}	10^{-9}	~ 2.5	?	?
CsBr:Nb	250, 4.96	7×10^{-3}	10^{-9}	~ 2.5	?	?

The thermal emittances are computed using the listed photon and work function energies in Eq. (3) and expresses the thermal emittance as the normalized rms emittance in microns per rms laser size in mm. The known experimental emittances are given with references.

Table 3
Properties of semiconductor cathodes.

Cathode type	Cathode	Typical wavelength & energy, λ_{opt} (nm), (eV)	Quantum efficiency (electrons per photon)	Vacuum for 1000 h (Torr)	Gap energy+ electron affinity, $E_G + E_A$ (eV)	Thermal emittance (microns/mm(rms))	
						Eq. (7)	Expt.
PEA: mono-alkali	Cs ₂ Te	211, 5.88	0.1	10^{-9}	3.5 [42]	1.2	0.5 ± 0.1 [35]
		264, 4.70	–	–	“	0.9	0.7 ± 0.1 [35]
		262, 4.73	–	–	“	0.9	1.2 ± 0.1 [43]
	Cs ₃ Sb	432, 2.87	0.15	?	$1.6 + 0.45$ [42]	0.7	?
	K ₃ Sb	400, 3.10	0.07	?	$1.1 + 1.6$ [42]	0.5	?
	Na ₃ Sb	330, 3.76	0.02	?	$1.1 + 2.44$ [42]	0.4	?
PEA: multi-alkali	Li ₃ Sb	295, 4.20	0.0001	?	?	?	?
	Na ₂ KsB	330, 3.76	0.1	10^{-10}	$1 + 1$ [42]	1.1	?
	(Cs)Na ₃ KsB	390, 3.18	0.2	10^{-10}	$1 + 0.55$ [42]	1.5	?
	K ₂ CsSb	543, 2.28	0.1	10^{-10}	$1 + 1.1$ [42]	0.4	?
	K ₂ CsSb(O)	543, 2.28	0.1	10^{-10}	$1 + < 1.1$ [42]	~ 0.4	?
	NEA	GaAs(Cs,F)	532, 2.33	0.1	?	1.4 ± 0.1 [42]	0.8
		860, 1.44	0.1	?	?	0.2	0.22 ± 0.01 [44]
GaN(Cs)		260, 4.77	0.1	?	$1.96 + ?$ [44]	1.35	1.35 ± 0.1 [45]
GaAs(1-x)Px		532, 2.33	0.1	?	$1.96 + ?$ [44]	0.49	0.44 ± 0.1 [44]
$x \sim 0.45$ (Cs,F)							
S-1	Ag–O–Cs	900, 1.38	0.01	?	0.7 [42]	0.7	?

The thermal emittances are computed using the listed photon, gap and electron affinity energies in Eq. (7) and expresses the thermal emittance as the normalized rms emittance in microns per rms laser size in mm.

power IR lasers). While good quantum efficiency is an important consideration for the new generation of high current high brightness sources, it is by far not the only figure of merit. Other key factors are longevity of the photocathodes during the operation as well as a short (picosecond or less) temporal response of photoemitted electrons and low transverse intrinsic (thermal) emittance [49]. The longevity aspect of photocathodes itself has several components to it, including the vacuum condition, the state of the surface, especially for cesiated photoemitters, operational conditions such as beam losses downstream of the gun and ion back bombardment. Even though other physical mechanisms can dominate cathode lifetime in a particular setup, it is the ion back bombardment which sets the ultimate limit to the photocathode longevity. To meet the longevity requirement, therefore, a two-prong approach is necessary: (1) improving the photoemitting materials by e.g. using stoichiometric compounds or large gap materials with stronger binding of the cesiated layer, or even eliminating Cs and achieving the NEA condition through delta-doping techniques

[50]; and (2) improvement of operational conditions through achieving better vacuum, halo and beam loss minimization in the gun vicinity. Additionally, we note that a care should be exercised when reporting photocathode lifetime values so that the main cause of the degradation is properly identified and correctly attributed to (which may have very little to do with a particular photocathode material choice).

The accelerator community so far has primarily been users of known photocathode materials when employing them for high brightness beam production. Since the understanding of the requirements and demands on the photocathodes for new high brightness high current electron sources has grown tremendously over the last decade, there exists a well-defined incentive for accelerator scientists to be engaged in the effort of obtaining a comprehensive understanding, which will ultimately lead to the creation of better photocathodes. A notable example worth emulating in this regard is the use of GaAs for the production of polarized electrons. Once the need for photocathodes delivering a higher degree of polarization was identified, the accelerator

community stayed engaged in the process of improving their performance, contributing to the creation of strained superlattice photocathodes now operating close to the theoretical limit of polarization (over 90% degree of polarization improved from the initial 30% for the bulk GaAs) [51]. A similar need remains to be addressed for high brightness high current unpolarized beams by providing careful photocathode characterization (transverse and longitudinal energy distributions, photoemission response time performed in a systematic and well-controlled environment) and then using these experimental data as an input for development of comprehensive and verifiable theoretical models, which will eventually allow engineering of new photocathode materials with the desired properties.

As a motivation reiterating the need for a systematic approach, one could point out the need for better understanding of photocathode thermal emittance, and in particular the thermal emittance of high quantum efficiency materials. For example, three different III–V materials were investigated at Cornell University over a wide range of photon energies (GaAs, GaAsP, and GaN), and a large variation of thermal transverse energies has been observed. GaN has been found to have a surprisingly large thermal emittance as compared to GaAs when excited with photons of energies above the bandgap by a similar amount. GaAsP has demonstrated very long response times as well as strong surface condition dependence on both thermal emittance and the response time. No such strong dependence was observed for GaAs and GaN. While several possible causes explaining these phenomena have been proposed [47,45], the results still remain to be quantitatively explained.

Another puzzling fact concerns the transverse momentum conservation and the role of the reduced mass on the thermal emittance. Some groups [36] have reported seeing the effect of momentum conservation at the surface from electrons thermalized to the Γ valley in GaAs and the resulting narrow cone emission (and sub-thermal intrinsic emittance as a result) by the virtue of the effective mass ratio for electrons inside the semiconductor and in vacuum (analog of Snell's law in optics). However, other measurements [52,47] show that the emitted electrons essentially do not experience the energy spread reduction due to the effective mass change, and an energy spread equal to the lattice temperature is obtained for a near band gap photon illumination. This is in contrast with simple theoretical models which predict much smaller values in the case of GaAs with effective mass ratio of 0.067 (effective mass in Γ valley to vacuum mass). The surface morphology and preparation techniques may be the deciding factors behind these seemingly conflicting observations, and these need to be accounted for in a systematic fashion. All this underscores the importance of bringing thermal emittance and response time of photoemitted electrons into quantitative agreement with a comprehensive theory.

7.5. Development of theoretical models for NEA photocathodes

As pointed out earlier, critical parameters for high brightness photocathodes, such as thermal emittance and response time, can be largely understood in the framework of Spicer's three-step model [53]. It should be noted that electron thermalization to the bottom of the conduction band in NEA photocathodes occurs quickly (e.g. 10^{-13} to 10^{-12} s in GaAs), thus, it is possible in principle to have both cold and sufficiently prompt electrons for ERL applications where a picosecond cathode response typically suffices.

The Spicer model has been extremely fruitful in explaining a wide range of photoemission-related phenomena. Its use has been typically limited to explaining quantum efficiency dependence on the wavelength using parameterized expressions with one or more adjustable parameters [54,55]. The diffusion model has also

been proven useful in explaining temporal response from GaAs [56,57,47]. Recently, Jensen et al. [58] demonstrated excellent quantitative agreement of QE versus photon energy for Cs₃Sb as computed in the framework of the three-step model without relying on adjustable parameters. The model is also being used to explain thermal emittance from metals and cesiated metallic surfaces [4,59]. The next logical step is to explain the transverse, longitudinal energy distributions and response time from high quantum efficiency photocathodes. Development of a sophisticated model should be a research goal of ERL photocathode R&D, the model will be refined with time as new data and theoretical insights become available. The ultimate objective is to obtain sufficient predictive power from the modeling to allow band structure and basic geometry engineering (active layer thickness, etc.) of new photocathodes to achieve favorable properties for high brightness average current photoinjectors. The essential model includes photon absorption, electron transport (diffusion and inelastic phonon scattering), as well as basic surface interface interaction [58,60,36]. It also becomes necessary to include multiple conduction band minima and the intervalley scattering between them as the effect has been shown to matter for a number of photocathodes. This proves to be important for indirect gap photocathodes such as GaP and is known to increase the thermal emittance [61]. Similarly, the effect of intervalley scattering may play a significant role in direct gap photocathodes when indirect conduction band minima are insufficiently separated in energy. Surface and geometry effects become critical when the photon absorption depth is comparable to the band-bending region (e.g. band-bending region is about 10 nm for GaAs). The spatial variation of the potential near the surface then needs to be incorporated into the model. Similarly, effects of geometry matter when modeling epitaxially grown thin layer photocathodes [57]. Additionally, empirical parameters describing the effects of the surface roughness and surface states will have to be added to better account for these phenomena.

7.6. Cathodes by design

Beyond the basic cathode types just discussed there are cathode systems designed to meet the needs of a particular application. In the simplest form, a laser producing photoelectrons from a thermionic cathode is a synthesis of two emission phenomena which can improve the QE and cathode robustness.

By far the most novel and technically challenging cathode is the diamond amplified cathode proposed by BNL [62]. In this scheme, a K₂CsSb cathode in transmission-mode is encapsulated in a single stage electron multiplier with a few KV across a ~ 1 mm gap formed between the cathode and a thin diamond film exit window. A laser generates electrons which accelerate in the gap which in turn produce secondary electrons in the diamond film. The secondary electrons escape and are accelerated in the gun. Multiplication of the photo-electrons by a factor of > 100 is expected and ~ 40 has been measured [63]. These cathodes are being developed for a SRF gun to drive a demonstration ERL. Very high average current densities (> 10 A/cm²) have been transported through diamond.

The properties of metallic cathodes can be significantly enhanced if they are designed so that plasmons can couple to light. Plasmons are a quasi-particle of coherent oscillations of electron density but do not normally couple to light due to momentum mismatch. This momentum mismatch can be made up either by coupling with a grating on the emitting surface, or by back surface illumination through a prism. In both cases, the momentum matching manifests itself in a giant increase in yield

through a small angular range of the incoming light. In the case of Al, it has been shown that for backface (Kretschmann) geometry, for a photon energy of 0.5 eV above the work function, the increase in electron yield in this matched condition is remarkably around 100 [64]. This increase comes from firstly a localization of field at the vacuum–metal interface, and a very large increase in absorption. Even though Al is a free electron metal, under these conditions the reflectivity becomes zero. Electron yield enhancements of 3500, 1000, 2500 and 50 over the bulk emission have been measured from Ag, Au, Cu and Al, respectively using multiphoton process with 100 fs laser operating at 625 nm wavelength and 50 mW average power [65]. In the grating coupled case large QE increases have also been seen in the Ag–O–Cs system around 900 nm [66] and around 350 nm wavelength [67]. Plasmons can also be generated at sharp discontinuities and then trapped in nano-cavities. The plasmons generated can exist up to very high wave-vector and thus be trapped in very small cavities of the order of 10 nm in size [68] This may be useful for localizing emission on a grid to reduce the effects of stochastic electron–electron interaction, or simply to design a certain emission profile in a dot-matrix arrangement. When illuminated with short (10's of fs) pulses, plasmon induced field localization can cause significant electron acceleration. In one example [69], 400 eV electrons with 50% energy bandwidth were created using only 1.5 nJ 27 fs 800 nm pulses on gold in prism coupled geometry using multiphoton excitation. This is both a good and bad result for photocathodes! On the one hand we must be careful while using plasmon enhancement to increase QE not to 'damage' the intrinsic momentum spread of the material. But on the other hand plasmon acceleration might offer a way to impulsively accelerate electrons to high energy, offering a way to avoid the most damaging effects of electron–electron scattering at low energy. In recent work it has been shown that by control of the carrier envelope phase of the laser and use of small emission points or strips, quasi monochromatic emission at high energy can be obtained [70].

8. Summary and conclusions

8.1. Physics challenges

The challenges for accelerator-based cathode R&D have two principal aspects. One is to continue improving the peak brightness of beams at low repetition rate. It is advances in this area which lead to the success of LCLS. The second is to develop cathodes for use in CW, high-average current accelerators for ERLs and high-average-power FELs.

Previous accelerator R&D provided the development of high peak brightness guns as a key enabling technology for the 4th generation, X-ray FEL light sources. Further R&D leading to even lower emittance beams will allow future FELs to be built at lower beam energy and shorter high energy accelerators which will greatly reduce the overall facility cost. High-peak brightness also widens the possibilities for using advanced FEL concepts and ideas, such as the production of fully coherent attosecond X-ray pulses, in new user facilities. Already one can see the advantages where the brighter than expected beam from the LCLS gun allows the X-ray FEL to saturate in less than half of the constructed undulator length, and at low charge produce a few micron long X-ray pulses which are nearly 100-times shorter than originally expected. If this was known before LCLS was built, it may have reduced the cost of the undulator system and allowed for a more aggressive FEL design.

On the other hand, the high-average current applications such as ERLs want both the very low emittance and high-average

current. The ERL will require a cathode current density of approximately 1.3 mA/mm^2 , QEs greater than 4% at visible or longer wavelengths, and an operational charge of several kilo-Coulombs from a single cathode. The cathode should also have low thermal emittance as well as low thermionic and field emission to minimize beam halo. These needs present the greatest challenge for cathode technology and are perhaps where the most intense R&D should be performed.

The issues of beam halo and dark current deserve additional attention given their impact on high average current applications. Beam halo is associated with the photocurrent and results from a poorly shaped laser beam, scattered laser light, space charge interactions with the beamline impedance and poorly matched electron optics. Beam halo is minimized with good laser shaping and laser transport, and mitigation of wake fields in the low energy section of the accelerator. Dark current is produced without any laser light and is mainly due to thermionic and field emission. Typically for cathodes operating at or below ambient temperature the thermionic emission is small. (See value for K_2CsSb above.) However the field emitted current can be as high as a few mA from a high gradient gun, and comes not only from the cathode, but also from any surface at high electric field and low work function cathodes may be more problematic in such an environment. Recent work at the Photo injector Test Facility at DESY, Zeuthen (PITZ) shows that cleaning the gun surfaces with dry ice can reduce the dark current by an order of magnitude [15].

In addition, achieving the high peak brightness and high current simultaneously will require meticulous three-dimensional shaping of the cathode drive laser pulse. The laser requirements are more stringent than for the low-duty factor beams because CW operation implies these guns will have lower accelerating fields making the space charge forces much more dominant, especially near the cathode. Therefore there should also be laser R&D in conjunction with the cathode research.

8.2. Summary of cathode R&D plan

Many of the details of the cathode R&D plan have been described both earlier in this paper and in outline form in Appendix 1, therefore only a concise summary is presented here.

The proposed cathode R&D plan consists of the following three interrelated parts:

1. Studies of optimal cathode formation methods and cathode emission characteristics, using available surface and material diagnostics.
2. Modeling of cathode emission physics and electron dynamics near the cathode.
3. Operational testing in the gun and injector system.

In general terms, the first part can be performed in national lab and university surface science laboratories, likely in collaboration with user facilities. This research provides an excellent opportunity for the education of Ph.D. level students. As its product, it will fill in the knowledge gaps as indicated in cathode tables above and provide detailed information on the thermal emittance, QE, lifetime, etc. And will provide important engineering design requirements such as vacuum, temperature stability and other specifications needed to engineer an accelerator cathode system.

The second part will also use the data produced by (1) to incorporate improved physics into the electron beam simulation codes. These enhanced codes in turn will be utilized to provide more realistic beam dynamics from the cathode through the injector and entire accelerator to preserve the low cathode

emittances. One result will be a physics design which can be used to engineer future guns and injectors.

The third part is then to use the knowledge gained from the previous two for testing cathodes in an operating gun and injector. If possible these tests should be performed in all three styles of guns: DC, NCRF and SCRF. For example, the test facility at JLab is already available to compare the performance of K_2CsSb and $Cs:GaAs$ cathodes in a DC gun. And cathode studies in a DC gun can also be performed at Cornell. The VHF gun now being constructed at LBNL provides an excellent opportunity for cathode studies in a NCRF gun. The 700 MHz SRF injector in construction and 1.3 GHz operational SRF gun at BNL provide excellent test beds for cathode studies in superconducting RF environment.

8.3. Activities that should be supported within the next 5 years

In the near term adequate support (M&S and effort) is needed to continue cathode surface science activities such as reliable production of K_2CsSb cathodes. This should include development of load-lock capabilities so that cathodes can be characterized at existing user facility beamlines. This characterization should include measuring the quantum efficiency of various photocathode materials as a function of wavelength and testing performance for incident laser pulses with a variety of photon energies and temporal and transverse distributions. Analysis of the experimental results with density functional theory and other models will yield further understanding of photoemission processes which can be applied to future engineered photocathodes.

Within 3–5 years, a variety of cathodes should be tested in normal and superconducting RF and high voltage DC guns, coupled to a beam dynamics characterization line to verify cathode survivability and the required beam quality for ERLs and FELs. These studies can be performed at injector test stands which already exist or are under construction at ANL, BNL, Cornell, JLab and LBNL. In order to facilitate these tests a standard load lock system for transferring cathodes between the various cathode and accelerator labs should be developed. This will assure compatibility among the various research facilities.

Acknowledgements

Authors gratefully acknowledge valuable discussion and constructive comments of Karoly Nemeth (ANL/Northern Illinois U.) and Klaus Attenkofer (ANL). We also recognize the contributions of Emanuele Pedersoli, Mike Greaves and Weishi Wan (LBNL). D. Dowell is supported by Department of Energy Contract DE-AC03-76SF00515 and a private contribution. K. Harkay's work supported by US Department of Energy, Office of Basic Energy Sciences under Contract no. DE-AC02-06CH11357. H. Padmore is supported by Department of Energy Contract # DE-AC02-05CH11231. T. Rao and J. Smedley's work are supported by Department of Energy Contract # DE AC02-98CH1-886.

Appendix 1. Outline of near and long term cathode research and development

Near term R&D

- Compile existing cathode database
 - Use the database to select promising cathode materials for various applications.

- Perform experiments to fill in the “data holes”.
- Choose a prototype cathode material
 - Good QE in visible and can survive in expected operating gun environment.
 - Develop fabrication techniques to reliably produce good QE.
 - Characterize material properties of “good” cathodes, both to improve reproducibility and optimize growth parameters.
 - Study surface characteristics: Lifetime, thermal emittance, robustness,...
 - Set specifications for survival in operational gun environment.
 - Analysis using surface science theory, like Density Functional Theory,...
- Implement emission and surface properties into beam simulation codes
 - Study beam dynamics near cathode and beyond to include injector designs.
- Pursue advanced cathode materials R&D in surface science lab
- Use results of cathode studies to define specifications for injector cathode system
 - Programming new physics into existing particle simulation codes.
 - Photoemission models are being put into codes like IMPACT and Parmela
 - Load lock specifications and gun vacuum requirements.
 - Cathode fabrication/transfer system for operational gun and injector.
 - Develop a versatile load lock design that can be shared and copied by all the labs to encourage transfer of cathode materials and ideas between labs
 - Load lock should be compatible with both the lab surface science chambers and the gun to both fabricate new cathodes and test used ones in the surface science lab
- Initial testing in gun within 1 to 2 years

Long term R&D

- Operation testing at full duty factor in support of beam physics experiments at injector test facility.
- Further develop and implement advanced cathodes.
- R&D in support of injector operations.

Appendix 2. Ongoing cathode research

Cathode research at national laboratories

Currently three laboratories, ANL, BNL and LBNL have active R&D programs on K_2CsSb cathode preparation. This type of cathode can be driven with laser light at visible wavelengths and has demonstrated robust operation in vacuum environments in the 10^{-9} to 10^{-10} Torr range, typical of normal conducting radio frequency guns. Jefferson Lab has developed $GaAs$ photocathodes for the CW 10 KW IR-FEL and has operated them at 10 mA of average current.

Argonne National Laboratory: Several groups are collaborating on cathode research for applications such as light sources and high-energy physics. For light sources, R&D on both cathode physics and high-brightness injector designs are being pursued. An ARPES lab is being commissioned that will be used to carry out fundamental photocathode studies for ultra-low emittance, including the benchmarking of DFT analyses. For emittance compensation, laser pulse shaping schemes have been developed and tested on the bench, and tests in the Injector Test Stand (ITS) are planned. Various injector design studies are underway. A thermionic RF injector is being studied that combines the

SCSS-type CeB₆ thermionic cathode with a very-low frequency (~100 MHz) LBNL-type RF gun [P.N. Ostroumov, K.-J. Kim, P. Piot, Proc. 2008 Linac Conference, 676 (2009)]. To potentially shorten the pulse length in such an injector, studies of laser-gated emission using a standard APS thermionic RF injector are underway in the ITS. For high-energy physics and other applications, high-QE cathode preparation and study are underway for very-high-charge injectors and for photodetectors; these cathodes include Cs₂Te, bialkali, and III–V semiconductors (see also photon detection collaboration under University research).

Brookhaven Nation Laboratory: BNL researchers have been aggressively working for a number of years on a variety of photocathode related topics including development, optimization and characterization of metal, semiconductor and superconducting cathode materials, investigation of various photoemission processes including multi photon, surface plasmon and photofield assisted emission and modeling of photoemission. They have been recently concentrating their effort on K₂CsSb and its use in a diamond amplified cathode. In this scheme, the cathode in transmission-mode is encapsulated in a kind of single stage electron multiplier with a few KV across a ~1 mm gap formed between the cathode and a thin diamond film. A laser generates primary electrons which accelerate in the gap and produce secondary electrons in the diamond film. The secondary electrons escape and are accelerated in the gun. Multiplication of the photoelectrons by a factor of >100 is expected and 40 has been measured. These cathodes are being developed for an SRF gun as part of an ERL demonstration.

BNL has an extensive material characterization effort dedicated to cathode development, utilizing resources at both the National Synchrotron Light Source and the Center for Functional Nanomaterials. This effort includes surface morphology and chemical analysis with scanning electron microscopy (including X-ray analysis), atomic force microscopy, near-edge X-ray absorption fine structure and X-ray photoemission spectroscopy. Crystallinity studies are performed with X-ray diffraction, X-ray topography and electron diffraction. Angle resolved photoemission and total-yield spectroscopy are used to measure band structure and electron affinity of cathodes. Bulk impurities in cathode materials such as diamond have been investigated using IR spectroscopy, Raman imaging and photoluminescence. X-ray micro-beam mapping has been used to identify the causes of spatial non-uniformity in diamond cathode response, and high-flux X-ray beams have been used to test the high current density performance of diamond.

BNL has three major photoinjector based accelerators (Accelerator Test Facility, Source Development Lab, Laser-Electron Accelerator Facility) in operation, and a fourth under construction (Energy Recovery Linac—ERL). The BNL ERL, with a design goal of 0.5 A average current, will be a test bed for the high-current ERLs needed for light-source applications. The ATF injectors operating at normal conducting mode at 2.856 GHz have been used extensively in testing metal photocathodes to deliver electron beams of very high brightness and low average current. Copper and magnesium cathodes tested in this gun has led the way to a number of very high brightness injectors including the LCLS injector. The superconducting RF injectors operating at 1.3 GHz, and 700 MHz and 112 MHz injectors currently under construction will be used for testing cathodes capable of producing very high brightness and very high average current.

Jefferson National Laboratory: Jefferson Lab's FEL is developing the next generation DC photoemission gun based on inverted insulators for reliable operation at 500 kV and with the capability of photocathode change out via load-locked system. The aim is to study various types of photocathodes in the same gun environment, in particular K₂CsSb and Cs:GaAs. The gun will be located in

the FEL's Gun Test Stand facility which has a 600 kV DC power supply, a drive laser system with the flexibility for 3D pulse shaping, and a diagnostics beam line that needs to be upgraded for measuring transverse and longitudinal emittance.

Lawrence Berkeley National Laboratory: LBNL is working on three aspects of photocathode research: (1) understanding the fundamental aspects of the interaction of light with the electronic system of a metal surface, leading to production of electrons; (2) understanding the chemistry and reliable production of alkali antimonide photocathodes for FEL applications and (3) the design of plasmonic metal surfaces for enhanced production of photoelectrons. A lab dedicated to alkali antimonide photocathode production and R&D and a second lab for characterization of cathodes using surface science techniques such as Angle Resolved Photo-Electron Spectroscopy (ARPES). There is also a UV PEEM to examine the microstructure of emitting surfaces, and synchrotron radiation surface analytical techniques at the ALS. The lab is in the final stages of construction of a 20 MV/m RF photogun that will be used as a facility for photocathode testing. The photocathode production and transfer system for this system is under design and construction.

Cathode research at universities

Maryland University: The photocathode research group at Maryland University focuses on studying dispenser cathodes for robust performance at high average current. This type of cathode is based on a metal substrate where cesium is diffused through the bulk of the cathode replenish constantly the quantum efficiency. In collaboration with the *Naval Research Laboratory* mathematical and computational models of density functional theory for emission from cesium-coated metals and some semiconductor materials have been developed for many years.

Cornell University: As a part of ongoing ERL R&D effort, Cornell University is in the process of establishing a dedicated gun and photocathode research laboratory. The photocathode research program involves the study of high quantum efficiency photocathodes, their properties as they pertain to high brightness electron beam creation, namely thermal emittance and response time. So far the photocathodes under investigation have belonged to III–V semiconductor group (GaAs and GaAsP [47,71], GaN [45]). Alkali-antimonide photocathodes are being evaluated, and a setup that allows the growth of Cs₃Sb, K₂CsSb, and Na₂KSb has been designed and now being constructed. In addition to the time resolving diagnostics which now allows characterization of photocathode response times to a 0.1 ps level, a dedicated setup is being built for simultaneous characterization of transverse and longitudinal energy spectra outside of high voltage DC gun environment based on the method originally implemented at Max Plank Institute [72]. Another research direction is developing Monte-Carlo models incorporating the photocathode physics to explain the salient features of the semiconductor photocathodes and their dependence on the laser wavelength, band-gap structure, and electron transport parameters in materials. One of the early successes of such modeling has been qualitative explanation of wavelength dependence of the response time from GaAs photocathode [56,57]. In collaboration with EE Department at Cornell University and SVT Associates, the work is underway to identify new promising structures with good quantum efficiency when excited with visible light suitable for low emittance beam operation for trial measurements. Independently from the gun and photocathode laboratory, the operating 10 MeV ERL injector prototype accelerator allows investigations in realistic running conditions of high average current performance of existing and new photocathodes. The highest average current obtained so far

was 8 mA at 6 MeV (20 mA DC beam demonstrated from GaAs after the high voltage gun), and operation at significantly higher currents is planned for this year. Detailed lifetime studies at high average current will commence then also.

Vanderbilt University: The group within the Physics Department and in collaboration with the Electrical Engineering Department studies field-emitter arrays (FEAs) cathodes for production of bright electron beams for compact free-electron lasers (FELs). This type of cathodes are rugged, require no laser driver, and generate little heat, which makes them attractive to test in normal conducting RF guns, but have only been tested in ~50 kV DC guns. The group has developed two methods to fabricate diamond FEAs, in the first method, pyramids are formed on a Si substrate and sharpened by microlithography and then coated with CVD nanodiamond. Typically, tip radii on the order of hundreds of nanometers are formed on 20- μm pyramids. In the second method, all-diamond pyramids are formed by a mold-transfer process in which they become sharpened from an oxide layer in the mold process, with tip radii smaller than 10 nm formed on 10- μm pyramids.

Old Dominion University: Through the applied research center in the Jefferson Lab Campus, the Electrical Engineering Department has performed over the last decade many surface analysis studies on GaAs photocathodes for both the Continuous Electron Accelerator Facility polarized electron gun and for the Free Electron Laser un-polarized, high current DC photoemission gun. The director of the applied research center has several publications on cesiated GaAs photocathode studies and has a strong graduate research program.

The College of William & Mary: The Applied Physics Department has various surface science laboratories located in the Applied Research Center, also in the Jefferson Lab Campus. Establishing a collaboration with the College of William & Mary to conduct R&D on photocathode preparation and surface analysis would be very easy. Some of the equipments potentially available include Time of Flight Ion Mass Spectrometer, Scanning Probe Microscope, Atomic Force Microscope, DekTak Surface Profilometer, Fourier Transform-Infrared Spectrometer, Scanning Electron Microscope, Hirox High Resolution Digital Microscope, Variable-angle Spectroscopic Ellipsometer, Scanning Tunneling Microscope, etc.

Cathode Research for Photon Detection Applications

Additional efforts in photocathode research are being performed by an interdisciplinary collaboration that is devoted to the development of Large Area Photodetectors. Under the leadership of Argonne National Laboratory, the University of Illinois Urbana Champaign, the University of Illinois Chicago, Washington University, and the Space Science Laboratory of the University Berkeley, these institutions are working on novel design concepts of fast and robust photocathodes with high quantum efficiency, low dark current, and long life time. The efforts address engineering, design, simulation, and industrial production aspects of standard alkali and III–V based photocathodes as well as nano-engineered materials. There is sizable overlap in this R&D effort with future light source accelerator applications.

References

- [1] J. Murphy, Plenary Presentation at BES Workshop, this issue.
- [2] D.H. Dowell et al., "LCLS Injector Drive Laser", in: Proceedings of the 2007 Particle Accelerator Conference.
- [3] Y. Ding, et al., PRL 102 (2009) 254801.
- [4] D.H. Dowell, J.F. Schmerge, Phys. Rev. Spec. Top. Accel. Beams 12 (2009) 074201.
- [5] J.G. Endriz, et al., Appl. Phys. Lett. 25 (1974) 261.
- [6] A. Michaelides, M. Scheffler, in: K. Wandelt (Ed.), Surface and Interface Science, vol. 1, Wiley-VCH, 2010, submitted for publication.
- [7] K. Harkay et al., in: Proceedings of the 2009 Particle Accelerator Conference, Vancouver, BC, submitted for publication.
- [8] W.F. Krolikowski, W.E. Spicer, Phys. Rev. 185 (1969) 882.
- [9] S. Hüfner, in: Photoelectron Spectroscopy, Springer, 1995, pp. 244.
- [10] I.V. Bazarov, C.K. Sinclair, Phys. Rev. Spec. Top. Accel. Beams 8 (2005) 034202.
- [11] K. Togawa et al., in: Proceedings of the 2003 PAC, Portland, OR, 2003, pp. 3332–3334.
- [12] Akre, et al., PRST-AB 11 (2008) 030703.
- [13] C. Hernandez-Garcia et al., A high average current DC GaAs photocathode gun for ERLs and FELs, in: Proceedings of the 2005 Particle Accelerator Conference, Knoxville, TN, May 16–20 2005, pp. 3117–3119.
- [14] L.B. Jones et al., Photocathode preparation system for the ALICE Photoinjector, in: AIP Conference Proceedings August 4, vol. 1149, 2009, pp. 1089–1093.
- [15] V. Volkov et al., in: Proceedings of EPAC08, Genoa, Italy, 2008, pp. 2213–2215.
- [16] C.M. Lyneis, et al., IEEE Transactions on Nuclear Science NS-26 (3) (1979).
- [17] C. Hernandez-Garcia et al., Status of the Jefferson Lab ERL FEL DC photoemission guns, in: Proceedings of the 2009 ERL Workshop, Ithaca, NY June 2009, in press.
- [18] D.H. Dowell, et al., Appl. Phys. Lett. 63 (15) (1993) 2035.
- [19] C. Boulware et al., Latest results at the upgraded PITZ facility, in: FEL2008 Conference Proceedings, Gyeongju, Korea, August 2008.
- [20] S. Lederer et al., in: FEL2007 Conference Proceedings, Novosibirsk, Russia, vol. 457, 2007.
- [21] W. Kohn, L. Sham, Phys. Rev. Lett. 140 (1965) 1133.
- [22] I.A. Nekrasov, et al., Eur. Phys. J. B 18 (2000) 55.
- [23] Y. Ikuno, K. Kusakabe, e-J. Surf. Sci. Nanotechnol. 6 (2008) 103.
- [24] S.D. Kevan, Phys. Rev. B 28 (1983) 2268.
- [25] M. Sterrer, et al., Phys. Rev. Lett. 98 (2007) 206103 and references therein.
- [26] K. Nemeth, K.C. Harkay, M. van Veenendaal, L. Spentzoutis, M. White, K. Attenkofer, G. Srajer, Phys. Rev. Lett. 104 (2010) 046801.
- [27] U. Martinez, L. Giordano, G. Pacchioni, J. Chem. Phys. 128 (2008) 164707.
- [28] T. König, et al., J. Phys. Chem. C 113 (26) (2009) 11301.
- [29] E. Pedersoli, et al., Appl. Phys. Lett. 93 (18) (2008) 183505.
- [30] E. Pedersoli, et al., Appl. Phys. Lett. 87 (8) (2005) 081112.
- [31] P.J. Feibelman, Phys. Rev. Lett. 34 (17) (1975) 1092.
- [32] J. Smedley, T. Rao, J. Sekutowicz, Phys. Rev. Spec. Top. Accel. Beams 11 (2008) 013502; G. Jezequel, A. Barski, P. Steiner, F., Solal, P. Roubin, R. Phys. Rev. B 30 (1984) 4833.
- [33] A.H. Sommer, in: Photoemissive Materials, John Wiley & Sons, New York, 1968, p. 7.
- [34] A.H. Sommer, in: Photoemissive Materials, John Wiley & Sons, New York, 1968, pp. 21–26.
- [35] D. Sertore, D. Favia, L. Monaco, P. Pierini, Cesium telluride and metals photoelectron thermal emittance measurements using a time-of-flight spectrometer, in: Proceedings of EPAC 2004, Lucerne, Switzerland, 2004, pp. 408–410.
- [36] Z. Liu, Y. Sun, P. Pianetta, R.F.W. Pease, J. Vac. Sci. Technol. B 23 (2005) 2758.
- [37] Juan R. Maldonado, et al., Phys. Rev. Spec. Top. Accel. Beams 11 (2008) 060702.
- [38] J. R. Maldonado, P. Pianetta, D.H. Dowell, J. Smedley, P. Kneisel, J. Appl. Phys. 107 (2010) 013106.
- [39] W.S. Graves, L.F. DiMauro, R. Heese, E.D. Johnson, J. Rose, J. Rudati, T. Shaftan, B. Sheehy, Measurement of thermal emittance for a copper photocathode, in: Proceedings of the 2001 Particle Accelerator Conference, 2227pp.
- [40] J.F. Schmerge et al., Emittance and quantum efficiency measurements from a 1.6 cell s-band photocathode RF gun with Mg cathode, in: Proceedings of the 2004 FEL Conference, pp. 205–208.
- [41] X.J. Wang, M. Babzien, R. Malone, Z. Wu, Mg cathode and its thermal emittance, in: Proceedings of LINAC2002, Gyeongju, Korea.
- [42] A.H. Sommer, in: Photoemissive Materials, John Wiley & Sons, New York, 1968, p. 130.
- [43] M. Miltchev et al., Measurements of thermal emittance for cesium telluride photocathodes at PITZ, in: Proceedings of 2005 FEL Conference, 2005, pp. 56–563.
- [44] I.V. Barazov et al., Thermal emittance measurements from negative electron affinity photocathodes, in: Proceedings of PAC 2007, 2007, 1221pp.
- [45] I.V. Bazarov, B.M. Dunham, X. Liu, M. Virgo, A.M. Dabiran, F. Hannon, H. Sayed, J. Appl. Phys. 105 (2009) 083715.
- [46] Michelato et al., Cs₂Te photocathodes for the TTF injector II, in: Proceedings of European Particle Accelerator Conference, Barcelona, Spain, 1996, 1510pp.
- [47] I.V. Bazarov, B.M. Dunham, Y. Li, X. Liu, D.G. Ouzounov, C.K. Sinclair, F. Hannon, T. Miyajima, J. Appl. Phys. 103 (2008) 054901.
- [48] N. Yamamoto, et al., J. Appl. Phys. 102 (2007) 024904.
- [49] I.V. Bazarov, B.M. Dunham, C.K. Sinclair, Phys. Rev. Lett. 102 (2009) 104801.
- [50] D. Bell, Recent progress in non-cesiated III-Nitride photocathodes, in: Presented at 1st workshop on photocathodes: 300–500 nm, University of Chicago, 20 July 2009.
- [51] T. Maruyama, et al., Appl. Phys. Lett. 85 (2004) 2640.
- [52] V.V. Bakin, et al., JETP Lett. 77 (2003) 167.
- [53] W.E. Spicer, A. Herrera-Gomez, Modern theory and applications of photocathodes, in: K.J. Kaufmann (Ed.), in: Proceedings of the SPIE, vol. 2022, Photodetectors and power meters, pp. 18–35.
- [54] W.E. Spicer, Phys. Rev. 112 (1958) 114.
- [55] D.H. Dowell, et al., Phys. Rev. Spec. Top. Accel. Beams 9 (2006) 063502.
- [56] P. Hartmann, et al., J. Appl. Phys. 102 (2007) 024904.
- [57] K. Aulenbacher, et al., J. Appl. Phys. 92 (2002) 7536.

- [58] K.L. Jensen, B.L. Jensen, E.J. Montgomery, D.W. Feldman, P.G. O'Shea, *J. Appl. Phys.* 104 (2008) 044907.
- [59] K.L. Jensen, N.A. Moody, D.W. Feldman, E.J. Montgomery, P.G. O'Shea, *J. Appl. Phys.* 102 (2007) 074902.
- [60] G. Vergara, A. Herrera-Gomez, W.E. Spicer, Escape probability for negative electron affinity photocathodes: calculations compared to experiments, in: *Proceedings of the SPIE*, vol. 2550, 1995, 142pp.
- [61] A.W. Baum, W.E. Spicer, R.F. Pease, K.A. Kostello, V.W. Aebi, *Proc. SPIE* 2550 (1995) 189.
- [62] J. Smedley, I. Ben-Zvi, J. Bohon, X. Chang, R. Grover, A. Isakovic, T. Rao, Q. Wu, Diamond amplified photocathodes, in: *Diamond Electronics—Fundamentals to Applications II*, Mater. Res. Soc. Symp. Proc. 1039, Warrendale, PA, 2007, 1039-P09-02.
- [63] X. Chang, I. Ben-Zvi, A. Burrill, J. Kewisch, T. Rao, J. Smedley, Y.-C. Wang, Q. Wu, First Observation of an Electron Beam Emitted from a Diamond Amplified cathode, PAC 09, Vancouver, Canada.
- [64] T.A. Callcott, E.T. Arakawa, *Phys. Rev. B* 11 (8) (1975) 2750.
- [65] T. Tsang, T. Srinivasan-Rao, J. Fischer, *Phys. Rev. B* 43 (1991) 8870.
- [66] J.G. Endriz, *Appl. Phys. Lett.* 25 (1974) 261.
- [67] F. Sabary, et al., *Appl. Phys. Lett.* 58 (12) (1991) 1230.
- [68] J. Le Perche, et al., *Phys. Rev. Lett.* 100 (6) (2008) 066408.
- [69] S.E. Irvine, A. Dechant, A.Y. Elezzabi, *Phys. Rev. Lett.* 93 (18) (2004) 184801.
- [70] P. Dombi, P. Rác, *Opt. Express* 16 (5) (2008) 2887.
- [71] I.V. Bazarov et al., *Phys. Rev. ST Accel. Beams* 11 (2008) 040702.
- [72] D.A. Orlov et al., *Appl. Phys. Lett.*, 78 (2001) 2721.

Accelerator Physics for Future Light Sources

A collection of reports from the September 15–17, 2009, workshop held in Gaithersburg, Maryland, sponsored by the Department of Energy, Office of Basic Energy Sciences, and co-chaired by W. A. Barletta (Massachusetts Institute of Technology) & J. N. Corlett (Lawrence Berkeley National Laboratory)



U.S. DEPARTMENT OF
ENERGY

Office of
Science

DISCLAIMER: This report was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government.