Dielectric laser accelerators

R. Joel England,* Robert J. Noble, Karl Bane, David H. Dowell, Cho-Kuen Ng, James E. Spencer, Sami Tantawi, and Ziran Wu

SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, California 94025, USA

Robert L. Byer, Edgar Peralta, and Ken Soong

Department of Applied Physics, Stanford University, 450 Serra Mall, Stanford, California 94305, USA

Chia-Ming Chang, Behnam Montazeri, and Stephen J. Wolf

Department of Electrical Engineering, Stanford University, 450 Serra Mall, Stanford, California 94305, USA

Benjamin Cowan

Tech-X Corporation, 5621 Arapahoe Avenue, Boulder, Colorado 80303, USA

Jay Dawson

Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, California 94550, USA

Wei Gai

Argonne National Laboratory, 9700 South Cass Avenue, Argonne, Illinois 60439, USA

Peter Hommelhoff

Department of Physics, Friedrich Alexander University Erlangen-Nuremberg, Erlangen, Germany and Max Planck Institute for Quantum Optics, Garching, Germany

Yen-Chieh Huang

Department of Electrical Engineering, Department of Physics, Institute of Photonics Technologies, National Tsing Hua University, Hsinchu 30013, Taiwan

Chunguang Jing

Euclid TechLabs LLC, 5900 Harper Road, Number 102, Solon, Ohio 44139, USA

Christopher McGuinness

Department of Radiation Oncology, University of California San Fransisco, 55 Laguna Street, San Francisco, California 94102, USA

Robert B. Palmer

Brookhaven National Laboratory, 2 Center Street, Upton, New York 11973, USA

Brian Naranjo, James Rosenzweig, and Gil Travish

Department of Physics and Astronomy, University of California, 450 Hilgard, Los Angeles, California 90095, USA

Amit Mizrahi and Levi Schachter

Department of Electrical Engineering, Technion-Israel Institute of Technology, Technion City, Haifa 32000, Israel

0034-6861/2014/86(4)/1337(53)

Christopher Sears
KLA-Tencor Corporation, 1 Technology Drive, Milpitas, California 95035, USA

Gregory R. Werner
Center for Integrated Plasma Studies, University of Colorado, Boulder, Colorado 80309, USA

Rodney B. Yoder
Goucher College, 1021 Dulaney Valley Road, Baltimore, Maryland 21204, USA

(published 23 December 2014)

The use of infrared lasers to power optical-scale lithographically fabricated particle accelerators is a developing area of research that has garnered increasing interest in recent years. The physics and technology of this approach is reviewed, which is referred to as dielectric laser acceleration (DLA). In the DLA scheme operating at typical laser pulse lengths of 0.1 to 1 ps, the laser damage fluences for robust dielectric materials correspond to peak surface electric fields in the GV/m regime. The corresponding accelerating field enhancement represents a potential reduction in active length of the accelerator between 1 and 2 orders of magnitude. Power sources for DLA-based accelerators (lasers) are less costly than microwave sources (klystrons) for equivalent average power levels due to wider availability and private sector investment. Because of the high laser-to-particle coupling efficiency, required pulse energies are consistent with tabletop microJoule class lasers. Combined with the very high (MHz) repetition rates these lasers can provide, the DLA approach appears promising for a variety of applications, including future high-energy physics colliders, compact light sources, and portable medical scanners and radiative therapy machines.

DOI: 10.1103/RevModPhys.86.1337 PACS numbers: 41.75.Jv, 29.20.−c

CONTENTS

*england@slac.stanford.edu

I. INTRODUCTION

Because of electrical breakdown of metals in the presence of high electric fields, conventional particle accelerators, which consist of metal cavities driven by high-power microwaves, typically operate with accelerating fields of 10 to 30 MV/m. Charged particle devices based upon conventional radio-frequency (rf) technology are often large and expensive due to the accelerator length and total stored energy needed to accelerate particles to high energy. Size and cost reductions are required for many applications. Advanced acceleration concepts intended to address these concerns may be classified according to whether the acceleration occurs in vacuum or in plasma. If in vacuum, one may further divide the various concepts between near- or far-field schemes. The focus of this review is on near-field or structure-based devices operated at optical frequencies, which we refer to as *dielectric laser accelerators* (DLA). The basic physical considerations and limitations for structure-based accelerators at optical frequencies are covered in Secs. I and II of this review. Because of the high power loss in metals at optical frequencies, dielectrics are the only viable candidate for confinement of the electromagnetic energy in such schemes. By comparison, the damage thresholds of dielectric materials exposed to pulsed laser light are fluence limited to the order of 1 J/cm^2 below 2 ps pulse lengths which corresponds to axial accelerating fields in a dielectric structure in excess of 10^9 V/m (Soong *et al.*, 2011). These fields are 1 to 2 orders of magnitude above typical gradient limits in metallic structures.

Structure-based laser-driven linear accelerators have been the subject of intense investigations, primarily for use with relativistic particles, where obtaining high energies in compact geometries is desired. Early concepts proposed using lasers to accelerate particles by operating known radiative processes in reverse, including the inverse Cherenkov accelerator (Shimoda, 1962) and the inverse Smith-Purcell accelerator (Takeda and Matsui, 1968; Palmer, 1980). The latter effect has been demonstrated experimentally with nonrelativistic electrons using a metallic grating at sub-mm wavelengths (Mizuno *et al.*, 1987) and more recently at near-infrared wavelengths (Breuer and Hommelhoff, 2014). Energy modulation of relativistic electrons has also been observed in a laser field truncated by a thin downstream metallic film (Plettner *et al.*, 2005), and several successful experiments in optical acceleration have made use of the inverse free-electron laser (IFEL) interaction to attain a beam longitudinally bunched at the optical period and thereby produce a net acceleration of particles (Kimura *et al.*, 2001; Sears *et al.*, 2008).

The above results represent important steps in this field. However, in order to attain efficient acceleration that is scalable to high energies, it is necessary to take advantage of the high laser damage thresholds and low Ohmic losses of dielectric materials. To these ends, a variety of all-dielectric high-gradient laser-accelerator structures have been proposed in recent years (Lin, 2001; Mizrahi and Schachter, 2004a; Plettner, Lu, and Byer, 2006; Cowan, 2008; Naranjo *et al.*, 2012). Significant progress in the communication industry for high-power solid-state lasers, optical fibers, and photolithographic techniques developed by the semiconductor industry has enabled the first of these dielectric laser-accelerator prototypes to be fabricated at the micron scale and tested with particle beams, leading to the first high-gradient demonstrations of laser acceleration in these devices (Peralta *et al.*, 2013). These fabrication and experimental results are discussed in more detail in Sec. III. Most of these accelerator structure concepts assume speed-of-light particles, but can be adapted to operate in the subrelativistic regime, where they may be combined with low-emittance particle sources to create laser-driven injectors (Breuer and Hommelhoff, 2014). This important auxiliary topic and its associated challenges are the subject of Sec. IV.

Any novel acceleration concept intended to exceed the performance of rf-based accelerators must demonstrate gradients of the order of 1 GV/m or higher, as well as a reasonable efficiency. While the former is primarily limited by the field that the confining materials can sustain, the overall efficiency is the product of two different efficiencies: one is the so-called wall-plug power to light efficiency of the laser and the other is that of the acceleration process. Beyond the three parameters of gradient, efficiency, and luminosity, there are many others that must be taken into consideration in any practical design. Operation with one electromagnetic mode constrains the relevant apertures to less than the vacuum wavelength (typically 0.3 to 0.8 λ) of the laser field. While the lower limit is set by the beam emittance, and hence beam size, the upper is determined, on the one hand, by the need for a

reasonable gradient for a given laser power and, on the other hand, by the necessity to avoid nonlinear effects in the dielectric (Sec. III.G). With the immense progress of recent years in diode lasers, the efficiency of laser-driven systems, as discussed in Sec. V, is approaching 50%, similar to that of microwave sources (e.g., klystrons).

For high-energy physics colliders, an essential parameter is the luminosity $\mathcal{L}$ (units $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$), defined as the number of events per unit time per unit fundamental cross section. It is proportional to the product of the bunch populations and repetition rate and inversely proportional to the beam area. An undesired consequence of the beam-beam interaction is the emission of beamstrahlung (beam induced synchrotron radiation) which causes energy spread. Because of the unusual particle-beam pulse format (low-charge bunches at high repetition rate), the beam-beam interaction is much cleaner with low beamstrahlung energy losses, making this an attractive approach for a future multi-TeV collider. In addition, at optical frequencies, this bunch format corresponds to attosecond (10^{-18} s) scale bunch lengths, making attosecond free-electron lasers an attractive possibility. We conclude this review with a discussion of these and other applications in Sec. VI.

A. Acceleration in vacuum

1. Near and far fields

Acceleration of charged particles in vacuum generally requires an electric field parallel to and phase synchronous with the particle velocity. Only very special geometries and boundary conditions can lead to net acceleration. Lawson categorized the situations where acceleration to first order in the field amplitude could not occur. These conditions in the form of a theorem are powerful tools in that any valid acceleration concept must violate some aspect of the theorem. To understand the Lawson acceleration theorem (Lawson, 1975, 1979; Palmer, 1986) one must first understand the distinction between "near" (evanescent) and "far" (traveling) electromagnetic fields. Consider the field $\mathbf{E}(t)$ generated, in a vacuum, by a single point charge's motion. Jackson (1975a) gives the field as

$$\mathbf{E}(t) = e\left(\frac{\mathbf{n}_0-\boldsymbol{\beta}}{\gamma^2(1-\boldsymbol{\beta}\cdot\mathbf{n}_0)^3R^2}\right)_{\mathrm{ret}} + \frac{e}{c}\left(\frac{\mathbf{n}_0[(\mathbf{n}_0-\boldsymbol{\beta})\times\dot{\boldsymbol{\beta}}]}{(1-\boldsymbol{\beta}\cdot\mathbf{n}_0)^3R}\right)_{\mathrm{ret}}, \tag{1}$$

where $\mathbf{n}_0$ is the unit vector from the particle to the point of observation, R is the distance, γ is the Lorentz factor, and $\boldsymbol{\beta}$ and $\dot{\boldsymbol{\beta}}$ are the velocity and acceleration vectors normalized by c, and refer to the particle of charge e at the so-called retarded time $(t-R/c)$. This field consists of two parts. The first falls with the square of the inverse distance from the source ($\propto 1/R^2$) and is called the near field. The second part of Eq. (1) falls only as the inverse linear distance from the source ($\propto 1/R$) and is called the far field.

An equivalent way of seeing the existence of the two different components is to examine solutions to Maxwell's equations. In the absence of all sources and in vacuum, the solutions are sums of fields described by (Jackson, 1975b)

$$\mathbf{E}(\mathbf{r},t) = \mathbf{E}_0\exp(i\mathbf{k}\cdot\mathbf{r} - i\omega t) \tag{2}$$

with the additional constraint that in the far field region $\mathbf{E}$ must be perpendicular to $\hat{\mathbf{n}}$, where $\hat{\mathbf{n}}$ is the unit vector defining the direction of a field propagation vector $\mathbf{k} = k\hat{\mathbf{n}}$ [not the observation vector $\mathbf{n}_0$ denoted in Eq. (1)], ω is the frequency, and $i=\sqrt{-1}$. If the normalized complex propagation vector $\hat{\mathbf{k}} = \mathbf{k}/|\mathbf{k}|$ is real valued, then $\hat{k}_x^2+\hat{k}_y^2+\hat{k}_z^2 = 1$, and Eq. (2) describes the familiar plane parallel, transversely polarized electromagnetic waves. For instance, if $\hat{k}_x = \hat{k}_y = 0$, then the propagation is in the z direction and

$$E_y(z,t) = E_{y0}\cos\left(\omega\left[\frac{z}{c}-t\right]\right), \tag{3}$$

$$E_x(z,t) = E_{x0}\cos\left(\omega\left[\frac{z}{c}-t\right]\right), \tag{4}$$

which describes far-field waves with frequency ω, moving at a velocity c, with two possible polarizations E_x and E_y.

If $\hat{\mathbf{k}}$ is complex then one or two of $\hat{k}_x^2$, $\hat{k}_y^2$, or $\hat{k}_z^2$ can be negative, while another can be greater than 1. For instance, if $\hat{k}_y = 0$, $\hat{k}_x = i\chi$, where χ is real, then $\hat{k}_z = \sqrt{1+\chi^2}$, and Eq. (2), together with Maxwell's equations, gives

$$E_y = E_0\exp\left(-\frac{\chi y}{\lambda}\right)\cos\left(\omega\left[\sqrt{1+\chi^2}\,\frac{z}{c}-t\right]\right), \tag{5}$$

$$E_z = \chi E_0\exp\left(-\frac{\chi y}{\lambda}\right)\sin\left(\omega\left[\sqrt{1+\chi^2}\,\frac{z}{c}-t\right]\right). \tag{6}$$

These represent waves that propagate in the z direction at a velocity $v = c/\sqrt{1+\chi^2}$ which is less than the velocity of light c. They are uniform in the coordinate x, but their strength falls, or rises, exponentially in the direction y. We could imagine a planar source of such a wave in the plane with $y = y_0$, then the waves move close to, and across, that surface, but fall farther away from the surface: they are "evanescent," corresponding to the first part of Eq. (1).

If χ is large, then the reduction with y is rapid, the fields in the z direction are strong relative to those in y, and the propagation velocity is low relative to that of light. But as $\chi\to 0$, the rate of fall-off decreases, the fields in the z direction become small compared to those in the y direction, and the wave velocity approaches c. If $\chi = 0$ then the field reduces to the transverse wave described by Eq. (3), becoming a far field with no z component.

Most real fields will consist of sums of both types of fields, and a field that approximates a far field must satisfy the condition

$$\chi y/\lambda \gg 1, \quad \text{i.e.,} \quad y \gg \lambda/\chi. \tag{7}$$

The above equations apply only to fields in a vacuum where the distinction between near and far fields is unambiguous. The situation is slightly different if the medium is a dielectric with finite refractive index. The solutions are essentially identical to the above with the modification that the propagating velocity is reduced by the refractive index.

2. The acceleration theorem

The Lawson acceleration theorem can be used to rule out the possibility of particle acceleration in many electromagnetic situations where a few simple assumptions apply. If the particle

(a) is far from any source such as a dielectric, conductor, or plasma [satisfying Eq. (7)],
(b) is in a region free of any static external electric or magnetic field, and
(c) is in a vacuum, where acceleration is defined as being
(d) integrated from and to locations outside the field, and
(e) first order in the driving fields $\mathbf{E}$ (energy change $\Delta U \propto |\mathbf{E}|$), then the acceleration theorem (Lawson, 1975, 1979) says that the net acceleration is zero. We understand this physically from a simple energy conservation argument or from the actions of the fields.

Far from all sources, the electromagnetic fields are sums of transversely polarized plane waves traveling at the speed of light. Acceleration of any particle by these fields must involve the electromagnetic field losing an equal amount of energy. The only way energy can be taken out of such fields is if the accelerated particle radiates traveling waves that are coherent with the accelerating fields, with phases such that they interfere with the incoming fields and reduce their amplitude. For first-order acceleration (linear in $|\mathbf{E}|$) when integrated between field-free locations, a free particle far from all sources, in vacuum without external fields, does not radiate. It thus cannot cancel the incoming field, cannot take energy from it, and cannot be accelerated.

The other way to understand the theorem is from the direct actions of the fields on the particle motion. Far from all sources [set by Eq. (7)], the fields can only be sums of plane parallel, transversely polarized, traveling waves. The fields seen by the particle traveling at velocity $v = \beta c$ are oscillatory. If the particle moves at an angle θ to the field's direction of propagation, then the energy gain will be given by the integral

$$\Delta U = eE_0 \sin\theta \int_{-\infty}^{+\infty} \sum_{\omega} \{\cos(\omega t[\beta\cos\theta - 1])\} dz \qquad (8)$$

under the substitution $t \to z/v$. The waves can never stay in phase with the particle unless it is traveling at the same velocity, parallel to the wave. At all other angles, and all other velocities, the fields seen by the particle will oscillate and average to zero. In the one case where they are in phase, moving in the same direction at $v = c$, then the transversely polarized, electric field seen by the particle cannot accelerate it.

This simple acceleration theorem provides a convenient way to categorize all viable acceleration mechanisms, since each mechanism must violate one or more of its conditions. Only if one allows one or more of the conditions (a)–(e) to be violated is acceleration possible. We now consider these cases in turn.

3. Acceleration with violations of the theorem's conditions

a. The particle is not far from a conductor or other source of radiation

This is the most common and important violation. Far from any source, the only electromagnetic fields are those represented by transverse polarized traveling waves as given in Eqs. (3) and (4) that travel at $v = c$ and cannot remain in phase with a particle at a lower velocity or at an angle to it. But near a source, the fields are given by Eqs. (5) and (6) which travel at velocities $v < c$ with finite electric fields in the direction of propagation (longitudinal polarization). For example, such waves can exist over a periodically corrugated structure giving so-called "grating acceleration" (Shimoda, 1962; Takeda and Matsui, 1968; Palmer, 1980). Double-sided, dielectric grating structures powered by lasers have been fabricated and tested, and these are discussed in Sec. III.A.

Sums of these longitudinal waves over all azimuthal angles lead to cylindrically symmetric fields that can, for instance, be maintained in a dielectric loaded conducting tube, or dielectric Bragg structure (Sec. III.B). There is a special characteristic of these fields that is a consequence of the Panofsky-Wenzel theorem (Panofsky and Wenzel, 1956). In a circularly symmetric near-field structure, the longitudinal field is independent of the radial position when the phase velocity equals c, since in this case $\nabla_\perp^2 E_z = (k_z^2 - k_0^2)E_z = 0$, where k_z is the propagation wave number and $k_0 = \omega/c$. The radial fields rise linearly with the radius, and the ratio of acceleration to these radial fields becomes small if the radius of the cylindrical dielectric beam aperture is large relative to the wavelength. The most familiar example of near-field acceleration is the iris loaded linac. The evanescent field character is revealed by the fact that the acceleration, relative to local fields on the irises, falls as the iris radii become large compared with λ. In practical linacs, the iris radii are a fraction of the wavelength.

b. There is an external electric or magnetic field

When a charged particle moves through a transverse electric or magnetic field its motion becomes curved. In this case, the particle emits "synchrotron radiation," named because it was first observed in synchrotrons in which a transverse magnetic field bends the particles into circular orbits. Inversely, an initial electromagnetic field can be introduced similar to, but opposite in phase to the synchrotron radiation, which will interfere destructively with the incoming field and thus remove energy from it (Palmer, 1972). Since energy must be conserved, the particle will be accelerated. A simple example is an inverse free-electron laser in which an alternating external magnetic field bends the particles into undulating paths. The particles then emit radiation coherently in the forward direction within an angle $1/\gamma$. If a plane wave with correct frequency and phase is traveling in the same average direction as the particle, then there is a fluctuating, but always positive, acceleration.

c. The particle is not in a vacuum

Another case of an inverse radiation mechanism is inverse Cerenkov radiation. The argument is identical to case (b): a particle traveling through a medium with finite refractive index n_r, at a velocity $v/c = \beta > 1/n_r$, emits Cerenkov radiation at an angle $\theta = \cos^{-1}(1/\beta n_r)$. Introducing incoming radiation at that angle with the appropriate phase will interfere with the Cerenkov radiation and lead to acceleration. The forward velocity of a plane wave in a dielectric medium is c/n_r. This will remain in phase with a particle at $v = c$ if the

particle moves at an angle θ to the wave propagation $\beta\cos(\theta) = 1/n_r$. The acceleration will now depend on the relative phase ϕ between the undulating motion and the field $E_z = E_0 \sin(\theta)\sin(\phi)$, where E_0 is the strength of the radiation perpendicular to its direction of propagation.

A second case of acceleration that violates the vacuum condition is when the medium has free charges in the form of a beam, or a plasma. Local electrostatic fields can accelerate charged particles. Examples are plasma wake acceleration (Tajima and Dawson, 1979), in which an electron beam or laser excites a plasma oscillation that accelerates particles, and the plasma beat wave accelerator (Joshi *et al.*, 1985), in which two lasers with frequency difference equal to the plasma frequency, excites a plasma oscillation.

d. Acceleration is not integrated to locations outside the field

A relativistic charged particle traveling at a small angle θ to a transversely polarized wave in vacuum will see a slowly oscillating electric field. The energy gain over a distance $d\ell$ will be $|\mathbf{E}|\sin(\theta)d\ell \approx |\mathbf{E}|\theta d\ell$, and the wavelength of the field oscillation experienced by the particle will be $\lambda_{\mathrm{acc}} = \lambda/[1-\cos(\theta)] \approx 2\lambda/\theta^2$. The energy change over one-half of an oscillation will be proportional to $|\mathbf{E}|\lambda/\theta$ which, as θ goes to zero, can become large. But the acceleration will eventually reverse and the integral will remain zero unless obstructions, such as mirrors with holes in them, restrict the distance over which the fields and particles can interact. An example is the Stanford acceleration experiment of electrons in a laser field truncated with a thin downstream film (Plettner *et al.*, 2005). The film provided a laser termination plane, but equivalently can be viewed as an inverse transition radiation mechanism to cancel the laser field.

It has been claimed that first-order acceleration can occur at a laser focus far from all sources. It is certainly true that at a focus there can be significant acceleration within the depth of focus, but there is in fact deceleration both before and after the focus that exactly cancels the effect. Calculating these fields from the incoming waves, having the forms of Eqs. (3) and (4), these fields add linearly, so, to first order, there is no net acceleration. Acceleration can occur if the fields are strong enough to deflect the particle near the focus, but the acceleration is then nonlinear and falls into the following category.

e. The acceleration is of higher order in $|\mathbf{E}|$

A particle traveling in the same direction as a transversely polarized wave will see oscillating transverse electric and magnetic fields that will cause it to quiver in a so-called figure-eight–type of motion (Landau and Lifshitz, 1975). The transverse quiver motion causes it to emit radiation forward, at the same frequency as the traveling field. The radiated field will then interfere with the incoming field, reducing its power, and by energy conservation lead to particle acceleration. The acceleration comes from the interaction of the transverse electromagnetic fields and the transverse oscillatory motion of the particles, much like that in a free-electron laser. But the energy gain is nonlinear. It is proportional to the amplitude of the oscillation ($\propto |\mathbf{E}|$), and also to the strength of the field that accelerates ($\propto |\mathbf{E}|$). So the net energy gain is proportional to $|\mathbf{E}|^2$. It is a second-order acceleration, violating condition (e). Two examples are Compton scattering and radiation pressure.

4. Summary

We have examined the conditions under which acceleration can and cannot occur. Without external static fields, first-order acceleration far from sources cannot occur. Acceleration can occur near or inside a dielectric, conductor, or plasma, and if there is an external static field. It can occur locally at fields in free space, but it is canceled entering and exiting that locality, and it can occur to a higher order in the accelerating field $|\mathbf{E}|$. These exceptions to the conditions that disallow acceleration provide a useful tool for categorizing mechanisms that allow particle acceleration. Within these general constraints, one must also consider the fundamental field limitations of materials used for the accelerator, and the scaling laws for accelerating modes with respect to practical metrics such as efficiency, beam loading, synchronism with particle velocity, and focusing. These topics are addressed in the following sections with particular focus on dielectric laser accelerators.

B. Scaling laws for dielectric particle accelerators

An important feature common to all dielectric acceleration structures stems from the fact that they are not enclosed by conductive boundaries. As such, high order modes (higher frequency) which may disrupt the beam are not confined as in metallic microwave structures but rather they leak out. The frequency dependence of the dielectric material dramatically reduces the number of eigenmodes that may be supported by the structure (Schachter *et al.*, 2003). For example, if the system is designed to support a fundamental mode at 1 μm then the 20th mode may correspond to an x-ray wavelength where the material is virtually transparent and will not support a mode. For comparison, in a disk-loaded metallic structure, the bunch may excite hundreds of eigenmodes.

1. General laws for accelerating modes

There are useful scaling laws that reveal the relations between the group velocity, the surface field, and the accelerating fields for a speed-of-light (SOL) accelerating transverse magnetic (TM) mode. These relations are quite general and are independent of the detailed implementation of the guiding structure. It puts an upper bound on the ratio between the accelerating and the surface field for a given aperture size and geometry. It also sets an upper bound on the attainable group velocity given a specific ratio between the accelerating field and the surface field. The most fundamental laws that govern the relations between surface field, group velocity, and accelerator structure aperture are addressed here.

With the assumption of a speed-of-light guided TM mode, the Maxwell equations can be readily reduced to

$$\nabla_\perp \times \mathbf{E}_\perp = 0, \tag{9}$$

$$\nabla_\perp \times \mathbf{H}_\perp = i\omega\epsilon E_z \hat{\mathbf{z}}, \tag{10}$$

$$\nabla_\perp \times E_z \hat{\mathbf{z}} = 0, \tag{11}$$

$$\mathbf{E}_\perp = -Z_0 \hat{\mathbf{z}} \times \mathbf{H}_\perp. \tag{12}$$

Here $\nabla_\perp$ is the transverse gradient operator, perpendicular to wave propagation, $\hat{\mathbf{z}}$ is the unit vector along the direction of

propagation, $\mathbf{E}_\perp$ and $\mathbf{H}_\perp$ are the transverse electric and magnetic fields, E_z is the longitudinal electric field, ω is the frequency, $Z_0 = \sqrt{\mu_0/\epsilon_0} = 377\ \Omega$ is the free-space impedance, ϵ_0 is the free-space permittivity, μ_0 is the free-space permeability, $\epsilon = \epsilon_0\epsilon_r$ is the material permittivity, and $\epsilon_r = n_r^2$ is the relative permittivity or dielectric coefficient. From Eq. (11), E_z has a constant value across the transverse plane. Equation (9) implies that $\mathbf{E}_\perp$ can be written as $-\nabla_\perp\Phi$, where Φ is a scalar potential function. Using this expression and Eq. (12) to solve for $\mathbf{H}_\perp$ in terms of $\mathbf{E}_\perp$, Eq. (10) becomes

$$\nabla_\perp^2\Phi = -ik_0E_z, \tag{13}$$

where $k_0 = \omega\sqrt{\mu_0\epsilon_0} = \omega/c$ is the free-space propagation constant.

Because E_z is a constant in the x and y directions, Eq. (13) has the solution $\Phi = -A(x^2 + \alpha_e y^2)/2$, where A and α_e are constants. The field components of this TM mode are

$$\mathbf{E}_\perp = A(x\hat{\mathbf{x}} + \alpha_e y\hat{\mathbf{y}}), \tag{14}$$

$$\mathbf{H}_\perp = \frac{A}{Z_0}(x\hat{\mathbf{y}} - \alpha_e y\hat{\mathbf{x}}), \tag{15}$$

$$E_z = \frac{i}{k_0}A(1+\alpha_e), \tag{16}$$

where $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are the unit vectors for the transverse plane coordinates. From Eq. (14) the contours of a constant transverse electric field form ellipses that obey

$$x^2 + \alpha_e^2 y^2 = \frac{D^2}{4}, \tag{17}$$

where the characteristic ellipse size D has the dimensions of length.

The constant α_e can be interpreted as the eccentricity of the waveguide cross section, which is placed with a boundary at a constant transverse electric field. When $\alpha_e \to 1$ the cross section is a circular waveguide with a diameter D, and when $\alpha_e \to 0$ the cross section is a slab along the $\hat{\mathbf{y}}$ with a gap D along $\hat{\mathbf{x}}$. At the boundary, characterized by Eq. (17), the ratio between the accelerating field E_z and the transverse field magnitude can be written as

$$\frac{E_z}{|\mathbf{E}_\perp|} = \frac{1+\alpha_e}{\pi D/\lambda}, \tag{18}$$

where λ is the free-space wavelength, which is related to the propagation constant $k_0 = 2\pi/\lambda$. It is clear from Eq. (18) that smaller apertures (of the order of λ) are preferred, limited by beam transport constraints. Planar structures, on the other hand, may offer advantages of planar fabrication (lithography) and one-dimensional heat removal with concomitant structural stability.

Next we turn our attention to the group velocity of such a TM mode. The group velocity v_g is the speed by which the stored energy in the structure is moving; i.e., $v_g = P/U$, where P is the power flow through the structure and U is the stored energy per unit length in the guiding structure. P can be calculated from the transverse fields expressed explicitly in Eqs. (14) and (15), while integrating the resultant Poynting vector over the cross-sectional area with boundaries governed by Eq. (17). For the stored energy, one has to take into account not only transverse and longitudinal fields but also the stored energy beyond the boundaries of the guiding structure. The ratio of the longitudinal electric field and transverse magnetic field at the boundary is imaginary and positive, hence from the Poynting theorem, the stored magnetic energy is greater than the stored electric energy beyond that boundary.

One can readily verify that the difference between the two stored energies beyond the boundaries is exactly equal to the stored energy in the longitudinal electric field within the boundaries of the guiding structure. That difference represents the minimal possible stored energy beyond the boundaries. For strong confinement (i.e., the field decays rapidly beyond the boundaries), the assumption of this minimal stored energy is reasonable. In the following we make that assumption.

The maximum possible group velocity can be obtained if the stored energy beyond the boundary is due only to magnetic fields and is equal to that of the stored energy in the longitudinal electric field within the waveguide. A limiting expression for an upper bound for the group velocity is given by

$$v_g \le \frac{c}{1 + 2((1+\alpha_e)\lambda/\pi D)^2}, \tag{19}$$

where c is the free-space speed of light. With the help of Eq. (18), Eq. (19) can be put in the form

$$v_g \le \frac{c}{1 + 2(E_z/|\mathbf{E}_\perp|)^2}. \tag{20}$$

One has to be careful about the interpretation of Eq. (20) since the field ratio in this equation is for the fundamental space harmonic only.

We now look at the implications of Eqs. (18)–(20) on the parameters of an optical accelerator. With the assumption of an aperture $D = \lambda$, then $E_z/|\mathbf{E}_\perp| = (1+\alpha_e)/\pi$. For the planar geometry, $\alpha_e = 0$, implying $E_z/|\mathbf{E}_\perp| = 0.32$. With the assumption of a 2 GV/m allowable surface field the best gradient attainable with a planar structure is about 600 MV/m. However, to obtain this gradient the pulse length τ is typically limited to about 1 ps. Because the maximum possible group velocity associated with such a structure is $0.83c$, the structure length will be limited to $v_g\tau/(1 - v_g/c) \approx$ 1.5 mm. This argues for a distributed coupling and/or very short and efficient coupler designs. One should keep in mind that structures that use soft confinement, such as those fed by parallel channels (Naranjo *et al.*, 2012) and the phase reset-type structures discussed in Sec. III.A, do not obey the laws for group velocity presented above. In fact these structures were invented to avoid the limitations posed by these typical waveguide structures.

2. Efficiency and wakefields

Many accelerator applications involve high power and high average current, so energy efficiency is a major concern. Typical electron-positron collider designs at 500 GeV center

of mass require about 14 MW of average beam power assuming 1.7×10^{14} electrons/s. Calculations of achievable laser-to-electron efficiency in the tens of percent range combined with current state-of-the-art laser wall-plug efficiencies now approaching 50% indicate that total wall-plug power for a DLA-based collider design are comparable with conventional accelerators (see Sec. VI.A).

For a compact collider, it is necessary to generate maximum gradient at a given input power. Since both Maxwell equations and the materials involved are linear, the laser power P flowing in the acceleration mode is proportional to the square of the acceleration gradient E_{acc}. It is natural to define the characteristic interaction impedance $Z_c = (E_{\mathrm{acc}}\lambda)^2/P$, wherein c/λ is the operating frequency of the fundamental (accelerating) mode, with a higher Z_c value corresponding to a more efficient accelerator. The impedance Z_c is distinguished from the shunt impedance $R_{\mathrm{sh}} = (E_{\mathrm{acc}}\lambda)^2/P_{\mathrm{loss}}$ commonly used in metallic accelerators, which instead uses the power loss P_{loss} in place of the mode's power flow. The two impedances are related through the mode's attenuation coefficient α_{loss} defined in Table II of Sec. II.C.

As a bunch of relativistic electrons traverses a structure, it generates an electromagnetic wakefield exciting many modes in a broad range of frequencies. The electromagnetic energy stored in this wake comes at the expense of the kinetic energy of the bunch. Because of linearity, the decelerating field E_{dec} is proportional to the total charge q in the bunch, i.e., $E_{\mathrm{dec}} = k_l q$, where k_l is called the (total) wake loss coefficient. Among the modes excited by the bunch there is a "projection" on the fundamental accelerating mode and due to its decelerating character, this contribution is opposite in phase with the accelerating field. While the reduction of the acceleration field by the projected wake is $E_{\mathrm{load}} = k_l q$, by virtue of the fundamental beam-loading theorem, the projected value of the trailing wake is twice that value $E_{\mathrm{wake}} = 2k_{l1}q$ (Wilson, 1991). We refer to k_{l1} as the reduced wake coefficient, and it determines the efficiency of an acceleration section.

Figure 1 summarizes the configurations and parameters of four typical dielectric structures. Figure 1(a) is a transverse photonic band-gap (PBG) structure analyzed by Lin (2001) consisting of an array of cylinders bored in a uniform dielectric medium. The structure is uniform in the longitudinal direction. The central channel is enlarged to accommodate an electron beam. The interaction impedance is fairly low (19.5 Ω) because of the small aperture ($R = 0.68\lambda$) needed to tune a speed-of-light TM mode to the band gap, and the group velocity is $0.58c$ with a field ratio $E_{\mathrm{acc}}/E_{\mathrm{max}} \simeq 0.5$. The impedance increases with smaller aperture as roughly $(\lambda/R)^4$.

For enhancing the interaction impedance, the structure in Fig. 1(b) was designed to be periodic in the longitudinal direction, uniform in one transverse direction and periodic in the other (Cowan, 2003). This longitudinal PBG structure was designed for acceleration of elliptical bunches with a very large aspect ratio. When varying the width of the internal channel $2D_{\mathrm{int}} = (1.1-2.5)\,\lambda$, the interaction impedance (per unit length) may be enhanced by 1 order of magnitude while the group velocity varies between $0.2c$ and $0.6c$, whereas the field ratio may drop by a factor of 3 below the values calculated for the first structure, $E_{\mathrm{acc}}/E_{\mathrm{max}} = 0.15$ to 0.35.

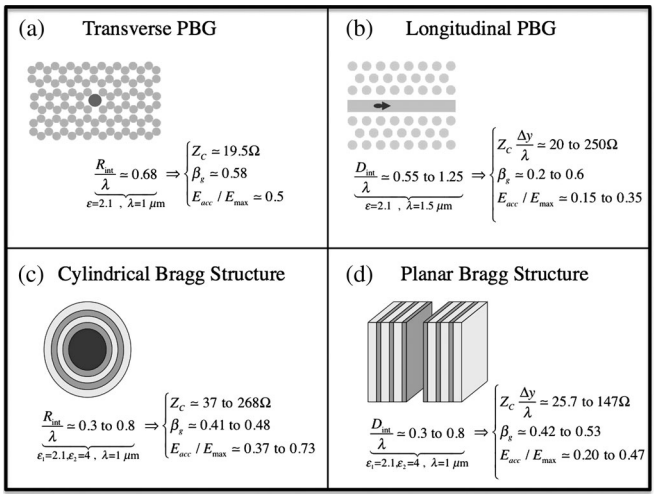

FIG. 1. Four different configurations of dielectric acceleration structures. (a), (c) The beam is cylindrical, whereas in (b), (d) the beam may be elliptical or flat (high aspect ratio). The characteristic parameters are also presented.

One-dimensional Bragg structures in Figs. 1(c) and 1(d) are also designed for higher impedances (Mizrahi and Schachter, 2004b). In the azimuthally symmetric case the interaction impedance may be enhanced by a factor of more than 10 comparing to Lin's structure (Lin, 2001) by varying the internal radius to smaller values and utilizing the appropriate indices of refraction of the Bragg layers. The field ratio varies between 0.37 and 0.73. In both cylindrical and planar Bragg structures the group velocity does not change dramatically while changing the width of the vacuum channel. However, in the planar case the field ratio approaches 0.47 but it can be as low as 0.20.

To obtain an estimate of the efficiency of an optical accelerator operating similarly to a microwave linear accelerator, consider a dielectric structure with a particle bunch in the form of a point charge passing through it (Schachter, 2004). Simultaneous with the acceleration process, the charge is decelerated by a wakefield proportional to the charge. For a cylindrical vacuum channel of internal radius R_{int}, the wake coefficient is $k_l = (2\pi\epsilon_0 R_{\mathrm{int}}^2)^{-1}$, whereas in the planar case, $k_l = (4\epsilon_0 D_{\mathrm{int}})^{-1}$, where D_{int} is the planar channel half-width (Schachter and Schieber, 1997). Although the structures are very different, from the perspective of an ultrarelativistic particle moving at c, what matters is only the vacuum-dielectric discontinuity since this is the only discontinuity that generates a reflected wave that may affect the point charge (short-range wake). Any reflection occurring farther away from the first discontinuity reaches the structure's axis only after the point charge has passed. Thus it may affect only trailing microbunches (long-range wake). The spatial behavior of the total electromagnetic power generated by such a point charge is $P \simeq cqE_{\mathrm{dec}}$ independent of the transverse variations beyond the vacuum-dielectric discontinuity.

The behavior of the trailing field on axis is strongly dependent on the transverse characteristics of the structure. In the case of a waveguide the spectrum is discrete according to the expression for the longitudinal field on axis $E_z(\tau = t - z/V) = 2qk_l \sum_{n=1}^{+\infty} W_n \cos(\omega_n \tau) h(\tau)$ (Schachter, 1997).

Here W_n are weighting functions that satisfy $\sum_{n=1}^{+\infty} W_n = 1$ with ω_n representing the discrete spectrum of frequencies and $h(u)$ is the step function. The first mode is for acceleration, having a phase velocity c, group velocity $\beta_g c$, and characteristic interaction impedance Z_c. These quantities are related to the wake coefficient k_l and, in particular, it is possible to establish the projection of the total deceleration on the fundamental mode (wavelength λ_o) represented by (Bane and Stupakov, 2003)

$$k_{l1} = k_l W_1 = \frac{\beta_g}{1-\beta_g} \frac{Z_c}{4\epsilon_0 \lambda_0^2 Z_0}. \tag{21}$$

The particle bunch experiences wake deceleration with a loss of kinetic energy $\Delta U_{\rm kin} = -q^2 k_l d$ over a length d. Including the acceleration, the net change in the kinetic energy of the bunch is given by $\Delta U_{\rm kin} = q(E_{\rm acc} - q k_l)d$. The total electromagnetic energy stored in the structure is $U_{\rm EM} = P_L \tau_{\rm EM}$, where $\tau_{\rm EM} = (d/c)(\beta_g^{-1} - 1)$, with the requirement that the charge and the laser pulse overlap. The efficiency of the acceleration process is $\eta = \Delta U_{\rm kin}/U_{\rm EM} = 4\eta_{\rm max} q(q-q_o)/q_o^2$, where $q_o = E_{\rm acc}/k_l$ is the charge for which the effective gradient vanishes, and the maximum efficiency is $\eta_{\rm max} = k_{l1}/k_l$, occurring for $q = q_{\rm opt} = q_o/2$. Typical corresponding optimal microbunch charges are in the range of 1–10 fC (Siemann, 2004).

The maximum efficiency k_{l1}/k_l is determined by the relative projection of the wake on the fundamental mode. This quantity may be enlarged by accelerating a train of microbunches whose spacing is the wavelength of the fundamental. Schachter (2004) presented some methods to enhance the efficiency and the total number of electrons accelerated. The three main concepts are (i) acceleration of a train of bunches with bunch spacing identical to the laser wavelength, (ii) a tapered laser pulse compensates for beam loading, and (iii) the wake trailing behind the train may be recycled and amplified for the acceleration of the next train of microbunches. Options (i) and (iii) have been calculated to have efficiencies approaching 40% to 50% for either scenario, comparable to conventional rf-based accelerator systems, as shown in Fig. 2 (Siemann, 2004; Na, Siemann, and Byer, 2005).

An alternative to a longitudinal bunch train is a bunch array or matrix of bunchlets near a single longitudinal location. Spencer (1998) calculated that this mediates the space-charge problem both within the bunchlet as well as between bunchlets, and indicated the conditions under which recombining the bunchlets into a single (higher charge) bunch at the interaction point (IP) would not negate the emittance gains from charge partitioning. Further simulations of lattice beams as well as an initial experimental trial with two bunchlets at the SLAC Next Linear Collider Test Accelerator (NLCTA) photoinjector were also performed (Noble and Spencer, 2007).

3. Luminosity

The rate of events per unit cross section is the main figure of merit of a high-energy collider, and this is measured in terms of the luminosity

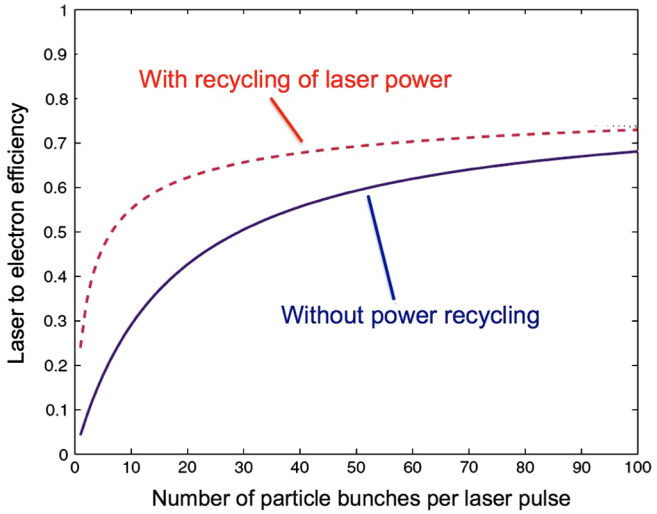

FIG. 2 (color online). Laser to electron coupling efficiency vs the number of electron microbunches per laser pulse. Adapted from Na, Siemann, and Byer, 2005.

$$\mathcal{L} = f_{\rm cross} H_D \frac{N_{\rm par}^2}{4\pi\sigma_x\sigma_y}. \tag{22}$$

In this expression $f_{\rm cross}$ is the crossing frequency of the particle bunches, $N_{\rm par}$ is the number of particles per crossing per beam, σ represents the transverse dimensions of the colliding microbunch, and H_D is the luminosity enhancement factor linked to the mutual focusing of the colliding beams. For useful physics, the required enhanced luminosity of a 3 TeV center-of-mass collider is about 2×10^{34} cm^{-2} s^{-1}.

The bunches arrive in bunch trains at frequency $f_{\rm train}$ with $N_{\rm train}$ microbunches per train and $N_{\rm micro}$ particles per microbunch. The microbunches are separated within the train by an integer multiple of the wavelength of the power source. In conventional radio-frequency accelerators, this separation is sufficiently large (one or more microwave periods) so that each microbunch interacts with only one microbunch of the opposing beam. In this case, in Eq. (22), $N_{\rm par} = N_{\rm micro}$ and $f_{\rm cross} = N_{\rm train} f_{\rm train}$. In the DLA scenario, on the other hand, the laser wavelength is 4 orders of magnitude smaller, and so for most cases of interest all microbunches within a given train will interact in a single crossing. Consequently, in Eq. (22), $N_{\rm par} = N_{\rm micro} N_{\rm train}$ and $f_{\rm cross} = f_{\rm train}$. This is a relevant distinction for the calculation of luminosity for the DLA case, which is discussed in more detail in Sec. VI.A, where parameters are worked out for a 3 TeV collider example. Such an optical collider is at least 10 times shorter as the average gradient is expected to be at least 1 GV/m, compared with conventional rf accelerators (30 MV/m) and the proposed Compact Linear Collider (CLIC) high-gradient microwave accelerator (100 MV/m). Optical microbunching is discussed further in Sec. IV.B.

4. Electromagnetically induced stress

The high-energy electromagnetic field guided within an optical acceleration structure gives rise to forces on the dielectric materials referred to as "radiation pressure" and polarization effects. But up to the fields where laser fluence will damage the dielectric, the radiation pressure is not

sufficient to break the structure and form cracks. As an example without loss of generality, consider a planar acceleration structure guiding laser light and examine the time-averaged forces on this structure. The total transverse force exerted on the planar structure is given by integrating the time-averaged Maxwell stress tensor over a closed surface $\langle F_x \rangle = -\langle T_{xx} \rangle = \frac{1}{4}\epsilon_0 |G_0|^2$ (Stratton, 1941). Hence, for a given accelerating gradient G_0, the total pressure is repelling (positive sign) and is independent of the details of the structure. For the gradient of $G_0 = 1$ GV/m, the total pressure is $\langle F_x \rangle = 2.2 \times 10^6$ N/m^2. For a Bragg structure of dimensions 1 mm square the total transverse force is 2.2 N. It is reasonable to assume that a pressure below Young's modulus E_Y may be sustained without damage to the structure (Lawn, 1995). Young's modulus for SiO_2 is 72.6 GN/m^2, whereas for Si it is 162 GN/m^2 (Bass *et al.*, 1995). The total pressure in our case is of the order of 1×10^6 N/m^2, and the electromagnetic forces are about 4 orders of magnitude below the theoretical threshold E_Y.

5. Thermal considerations

Optical fields may have a significant effect on the acceleration structure via thermal processes if the optical energy is absorbed by the dielectric and causes expansion (Karagodsky, Mizrahi, and Schachter, 2006). For a planar optical Bragg acceleration structure, the temperature distribution is determined by the 1D steady state diffusion. The maximal temperature in a silica-zirconia structure is $3.56 \times 10^9 \tan(\delta_1)$[K] and the maximal heat flux is $9.22 \times 10^9 \tan(\delta_1)$[W/cm^2], while in the silica-silicon structure these quantities are $6.64 \times 10^8 \tan(\delta_1)$[K] and $1.79 \times 10^9 \tan(\delta)$ [W/cm^2], where $\tan(\delta_1)$ is the loss tangent of the first layer (silica). If we assume a loss factor similar to that for communication optical fibers, 1 dB/km[$\tan(\delta) \sim 10^{-11}$], the temperature increase will be less than 1 K and the heat flow will be a fraction of a W/cm^2, which is more than 3 orders of magnitude lower than the known technological limit for heat dissipation. Similar one-dimensional cooling requirements are familiar in the laser industry where 1 kW/cm^2 heat removal rates are typical (Eggleston *et al.*, 1984; Rutherford *et al.*, 2000). In addition, silicon has a thermal expansion coefficient below 1×10^{-8}/K at cryogenic temperatures, which could allow for stable accelerator structures even when pumped with 10 kW of laser power per meter of accelerator length.

II. ACCELERATION IN DIELECTRIC STRUCTURES

A. Photonic crystals

To accelerate particles efficiently, electromagnetic waves must be guided or confined to the region in which the particles travel. To accelerate particles over significant distances, the wave phase velocity (in the direction of particle travel) must be synchronized with the particle velocity, and the waves must have an electric field component in the direction of acceleration. As explained in Sec. I, free plane waves cannot satisfy these conditions.

Traditionally, electromagnetic waves have been confined to a vacuum beam channel by metal structures, with periodic corrugations or a dielectric liner to reduce the phase velocity to approximately the speed of light (i.e., with features to adjust the mode frequency and wave number so $v_{\rm ph} \approx c$ to accelerate high-energy particles). Field confinement and synchronism can also be achieved through many partial reflections from surrounding dielectric elements. Each dielectric interface causes a partial reflection, so that together the dielectric elements reflect most of the outgoing wave energy back toward the beam channel. If the dielectric is arranged so that reflections from different elements interfere constructively (Bragg reflection), while transmitted waves interfere destructively, confinement can be nearly perfect (Yeh and Yariv, 1976).

Many different dielectric arrangements have been explored to optimize interference of reflected and transmitted waves from the dielectric interfaces. The most widely explored confinement technique is based on PBGs in photonic crystals (PhCs), which are regular arrays or lattices of dielectric elements (Joannopoulos *et al.*, 2008). Confinement due to Bragg reflection, coherent scattering that maximizes reflection, can be considered a consequence of PBGs. For example, a Bragg mirror is a 1D finite-layer version of a PBG crystal (Sec. III.B) with the reflection becoming perfect only in the limit of an infinite number of layers. However, some photonic quasicrystal (PQC) and aperiodic dielectric configurations exhibit strong confinement as discussed later in this section.

The confinement properties of an infinite, periodic structure can be found through numerical computation of the band structure of allowed modes, which reveals any PBGs. However, the inverse problem of designing a structure with the desired PBG is very difficult. For simple (especially quasi-1D) systems, analytic conditions for maximum constructive or destructive interference can lead directly to an optimal structure design, even for finite and aperiodic structures. Optimizing scattering phases from individual dielectric elements and the relationship with PBGs are discussed (mostly for telecommunication applications) by Duguay *et al.* (1986), Archambault *et al.* (1993), Rowland, Afshar V, and Monro (2008), and Sakai and Suzuki (2011). These concepts are applied to two-dimensional (2D) structures in White *et al.* (2002), Litchinitser *et al.* (2003), Couny *et al.* (2007), and Murao, Saitoh, and Koshiba (2011). Because Bragg reflectors can be understood through both interference conditions and the PBG condition, their application to acceleration, first explored by Mizrahi and Schachter (2004a), is particularly illuminating and is discussed in detail in Sec. III.B.

A dielectric photonic crystal is essentially a dielectric medium that varies periodically in space: $\epsilon_r(\mathbf{x} + \mathbf{a}_j) = \epsilon_r(\mathbf{x})$ for lattice basis vectors $\mathbf{a}_j$ ($j = 1, 2, 3$ for a 3D lattice). Just as atomic crystals lead to a band structure for electron eigenstates, photonic crystals lead to a band structure for electromagnetic eigenmodes. The modes of this periodic structure satisfy Maxwell's equations (assuming oscillation as $e^{-i\omega t}$)

$$\begin{aligned} &\nabla \cdot \epsilon_0 \epsilon_r(\mathbf{x}) \mathbf{E} = \nabla \cdot \mathbf{D} = 0, \\ &\nabla \cdot \mathbf{H} = 0, \\ &\nabla \times \mathbf{E} = i\omega\mu_0 \mathbf{H}, \\ &\nabla \times \mathbf{H} = -i\omega\epsilon_0 \epsilon_r(\mathbf{x}) \mathbf{E}, \end{aligned} \tag{23}$$

where the periodic dielectric is determined by the permittivity $\epsilon = \epsilon_0 \epsilon_r(\mathbf{x})$. The two Maxwell curl equations combine to yield

the eigenvalue equation for the field $\mathbf{H}$, thus relating frequency to wave number,

$$\Theta\mathbf{H} = \nabla \times [\epsilon_r^{-1}\nabla \times \mathbf{H}] = \frac{\omega^2}{c^2}\mathbf{H}. \tag{24}$$

The operator Θ is linear, positive semidefinite, and Hermitian, so its eigenvectors are orthogonal with real, non-negative eigenvalues (and therefore, frequencies ω are real). For anisotropic (birefringent) materials, ϵ_r can be a symmetric 3×3 tensor.

The eigenmodes of the periodic system are characterized by Bloch's theorem (cf. Floquet's theorem), which groups eigenmodes into frequency bands $\omega_m(\mathbf{k})$, where the Bloch wave vector (or crystal momentum) $\mathbf{k}$ describes the periodicity of the eigenmode m with respect to the lattice. A Bloch eigenmode can be written as

$$\mathbf{H}(\mathbf{x}) = \mathbf{u}_\mathbf{k}(\mathbf{x})e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{25}$$

where $\mathbf{u}_\mathbf{k}$ has the same periodicity as ϵ_r. Since $\nabla \cdot \mathbf{H} = 0$, $\mathbf{u}_\mathbf{k}$ must satisfy the transversality constraint $\nabla \cdot \mathbf{u}_\mathbf{k} + i\mathbf{k} \cdot \mathbf{u}_\mathbf{k} = 0$. For a given $\mathbf{k}$, substituting Eq. (25) into Eq. (24) shows that $\mathbf{u}_\mathbf{k}$ is an eigenmode of

$$\Theta_\mathbf{k}\mathbf{u}_\mathbf{k} = (\nabla + i\mathbf{k}) \times [\epsilon_r^{-1}(\nabla + i\mathbf{k}) \times \mathbf{u}_\mathbf{k}] = \frac{\omega^2}{c^2}\mathbf{u}_\mathbf{k} \tag{26}$$

over the bounded region of a lattice unit cell with periodic boundary conditions. Instead of solving Eq. (24) over all space, we can solve Eq. (26) over a *unit cell* (defined by the vectors $\mathbf{a}_j$) for a range of $\mathbf{k}$. The spectrum of the bounded system is discrete, with eigenmodes typically labeled by a positive integer m: $\mathbf{u}_{m,\mathbf{k}}$.

Bloch's theorem connects the periodicity in coordinate space with the periodicity it imposes on the *reciprocal lattice space* (also called $\mathbf{k}$ space, wave-number space, or crystal momentum space). This imposed periodicity is described by reciprocal lattice vectors $\mathbf{K}$, which satisfy $\exp(i\mathbf{K} \cdot \mathbf{a}_j) = 1$ for all j. The functions $\exp(i\mathbf{K} \cdot \mathbf{x})$ have the same periodicity as ϵ_r and, in fact, form a Fourier basis for all periodic functions on the lattice. Therefore the operator $\Theta_\mathbf{k}$ has the same set of eigenmodes as $\Theta_{\mathbf{k}+\mathbf{K}}$. As $\mathbf{k}$ moves from $\mathbf{k}_0$ to $\mathbf{k}_0 + \mathbf{K}$, the set of eigenmodes $\mathbf{u}_\mathbf{k}$ and the set of eigenfrequencies vary periodically with $\mathbf{k}$, yielding a *band structure*.

Because of the periodicity with respect to $\mathbf{k}$, Bloch wave vectors $\mathbf{k}$ and $\mathbf{k} + \mathbf{K}$ are equivalent (which distinguishes Bloch wave vectors and crystal momentum from ordinary wave vectors and momentum). A mode with Bloch wave vector $\mathbf{k}$ has the periodicity $\mathbf{H}(\mathbf{x} + \mathbf{a}_j) = \mathbf{H}(\mathbf{x})e^{i\mathbf{k}\cdot\mathbf{a}_j}$, the same as a mode with $\mathbf{k} + \mathbf{K}$ for any reciprocal lattice vector $\mathbf{K}$. Without loss of generality, we can label such modes with $\mathbf{k}$ belonging to a unit cell of the reciprocal lattice (e.g., the first Brillouin zone). For example, for a 1D lattice with lattice constant a, the reciprocal lattice vectors are $K = 2\pi n/a$ for all integers n (i.e., the Fourier wave numbers of a function with period a). The first Brillouin zone is the interval $B_1 = (-\pi/a, \pi/a]$. Any wave number can be written uniquely as $k + K$, for $k \in B_1$ and some reciprocal lattice vector K.

As noted, the periodicity in $\mathbf{k}$ space leads to a frequency band structure that illuminates general properties of PhC eigenmodes. For each eigenmode m of Eq. (26), the eigenfrequency $\omega_m(\mathbf{k})$ is a periodic (and generally smooth) function of $\mathbf{k}$, which means it is a bounded function, forming a frequency band. The eigenmodes $\mathbf{H}_{m,\mathbf{k}}(\mathbf{x})$ satisfy

$$\Theta\mathbf{H}_{m,\mathbf{k}}(\mathbf{x}) = [\omega_m(\mathbf{k})/c]^2\mathbf{H}_{m,\mathbf{k}}(\mathbf{x}), \tag{27}$$

$$\mathbf{H}_{m,\mathbf{k}}(\mathbf{x} + \mathbf{a}_j) = \mathbf{H}_{m,\mathbf{k}}(\mathbf{x})e^{i\mathbf{k}\cdot\mathbf{a}_j}. \tag{28}$$

The periodic band structure has significant consequences. For example, when ω_m is bounded and smooth, its derivatives must vanish at its extrema. Therefore the group velocity is zero at the band edges. Regardless of the details of $\omega_m(\mathbf{k})$, we can conclude that a periodic dielectric has some zero-group-velocity modes.

The boundedness of frequency bands means that a PBG may open between bands, such that the frequencies in bands $m \leq M$ are all below $\omega_{M,\max}$ and the frequencies in $m \geq M + 1$ are all above $\omega_{M+1,\min} > \omega_{M,\max}$. The mode frequencies most sensitive to perturbation by the periodic dielectric are those near the Brillouin zone edges and frequencies in adjacent bands will move apart. For example, in one dimension this occurs for $k \approx \pi/a$ since $k = \pi/a$ and $k = \pi/a - K$ have the same wavelength for $K = 2\pi/a$. Physically, the mode just below the gap has its power concentrated in the high-index dielectric while the mode above the gap is concentrated in the low-index material (Joannopoulos *et al.*, 2008).

There are no propagating modes with frequencies ω in the gap. A plane wave at frequency ω, impinging on a (half-infinite) PhC with a band gap containing ω, cannot couple (or transfer energy) to any propagating modes in the PhC. The plane wave must therefore be completely reflected. Only evanescently decaying fields penetrate the PhC, and they transport no energy. A material with a photonic band gap thus acts like a frequency-selective mirror, completely reflecting frequencies in the band gap. It is remarkable that there exist 2D PhCs with complete band gaps for 2D wave propagation, and 3D PhCs with complete band gaps for 3D propagation, including all polarizations (John, 1987; Yablonovitch, 1987).

If a so-called *defect* is introduced into the lattice, say by removing or altering one element of the lattice, then any fields excited at frequencies within the band gap will be trapped near the defect, because they cannot propagate away. The term defect does not refer to a manufacturing error, but rather to a deliberate, localized change that breaks the lattice symmetry. The trapped fields decay evanescently into the crystal. As long as the surrounding crystal is infinite, the evanescent fields carry away no energy. However, if the surrounding crystal is finite, then fields can "tunnel" through the crystal (analogous to quantum mechanical tunneling through a barrier), emitting a propagating wave outside the crystal that does carry away energy ("diffractive" or "radiative" losses). The evanescent fields decay roughly as $\exp(-\Gamma r)$, where r is the distance through the crystal, and in general the decay constant Γ depends on the crystal structure, being larger for frequencies closer to the center of the band gap and for wider band gaps. For a dielectric contrast (the ratio of largest to smallest ϵ_r) around 4 or 5, the surrounding PhC may need to be 10 or 20 wavelengths thick to substantially reduce diffractive losses. PhC waveguides and cavities tend therefore to be much wider

transversely than metallic structures, which are usually on the order of a single wavelength across. Section III reviews different examples of contemporary PBG structures being investigated for particle acceleration.

For an optical accelerator, the PBG defect is typically a linear hollow channel, and the accelerating mode must have a longitudinal electric field along the axis of the particle trajectory, and a phase velocity equal to the particle velocity. In accelerator terminology, these are called transverse magnetic modes because the magnetic field is predominantly transverse to the beam direction, with a strong longitudinal electric field. Except for modes with cylindrical symmetry and modes with $k_z = 0$ along the axis, PhC waveguide modes do not decouple strictly into TM and transverse electric (TE) modes. However, in the central vacuum defect region, away from the first dielectric layer, the accelerating modes behave like TM modes, with electromagnetic properties very similar to those of accelerating modes in conventional metal cavities.

The desire to improve confinement as well as to reduce the number of lattice layers and the transverse size of PBG acceleration structures has motivated recent work on so-called truncated photonic crystals and PQCs. Much less is understood about PQC confinement for accelerators, with no self-consistent theory to calculate the optimal dielectric arrangement. The first numerical simulations to achieve accelerating mode confinement in an irregular structure were carried out by Bauer, Werner, and Cary (2008) with subsequent investigations of PQCs by DiGennaro *et al.* (2009). The concept of a band gap does not apply to these aperiodic structures, yet their special arrangement of dielectric elements can in some cases provide stronger confinement than finite PBG structures with the same number of elements.

For a defect mode in a 2D triangular PhC lattice of dielectirc rods, the fields would decay exponentially transversely away from the defect, carrying no energy. Ignoring material losses, the Q factor (stored energy or energy loss per radian cycle) of the mode would be infinite. If we truncate the crystal after a finite number of hexagonal layers of rods, some energy will tunnel through the crystal and escape, yielding a finite Q factor. Because the fields decay exponentially within the crystal, Q increases by roughly the same factor for each additional layer added. For example, Table I shows how the Q factor increases as the lattice is truncated at increasing sizes. In this case, adding one more lattice layer surrounding the defect increases Q by about a factor of 8 (for $\epsilon_r = 10$).

TABLE I. The Q factor for a truncated triangular lattice. As the number of rods surrounding a defect cavity increases, less radiation escapes, and the Q factor (not including dielectric losses) increases. These results are for $\epsilon_r = 10$ rods with radius-to-lattice-constant ratio $r/a = 0.179$, with the central rod removed to create a cavity with resonant frequency $\omega_0 = 2.6c/a$.

No. of rings of rods	No. of rods	Q
2	18	1.3×10^2
3	36	1.1×10^3
4	60	9.7×10^3
5	90	8.0×10^4
6	126	6.6×10^5

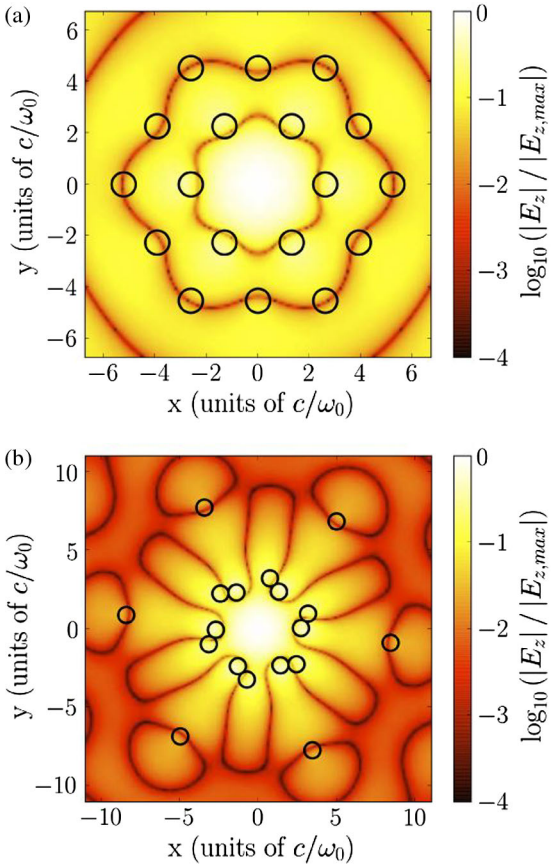

FIG. 3 (color online). (a) Axial electric field strengths in a 2D hexagonal lattice and (b) the optimized photonic 2D quasicrystal cavity. From Bauer, Werner, and Cary, 2008.

In Bauer, Werner, and Cary (2008), the rods of a two-dimensional photonic crystal were allowed to move away from their lattice positions in directions that increased the Q factor (due to radiative losses only). The resulting Q factors were orders of magnitude higher than in lattice structures with the same number of rods. For example, a rod-radius-to-lattice-spacing ratio $r/a = 0.179$ yielded the highest Q factor, 130, for a triangular lattice structure with 18 rods of $\epsilon_r = 10$. In contrast, only 18 rods, with optimized positions, yielded a Q factor of 1.1×10^4. Figure 3 shows that the field values outside the optimized structure are significantly lower than the fields outside the lattice structure. A PhC lattice structure would need at least 60 rods to attain a Q factor higher than the 18-rod optimized PQC structure.

Optimization of dielectric elements in this example was extremely successful, yielding orders-of-magnitude improvements. However, the structure with minimal radiative losses proved to be more sensitive to rod positions, so such a structure would have to be fabricated with tighter tolerances. On the other hand, in some cases a sensitive Q factor might be an advantage, since a drastic change in Q could find application as a switch.

B. Accelerating and focusing forces

Dielectric laser accelerators are limited by dielectric breakdown effects to accelerating fields in the GV/m range for

picosecond pulse lengths. With the need to confine and precisely shape the electromagnetic fields for particle acceleration, one employs photonic design principles (Sec. II.A). The structure in general may display a rich array of longitudinal spatial harmonics, as are found in standard rf linear accelerators (linacs), and which play a critical role in the dynamics of high-field optical accelerators, as discussed below. There will also be transverse spatial harmonics. The symmetry of the transverse PBG cross section employed for confinement permits a choice for the transverse beam profile in the channel. Axisymmetric, 2D rectangular, hexagonal, and flat 1D PBG structures are all possible. For a wide aspect ratio Cartesian structure, the significant expanding of one transverse dimension (unlike rf linacs having approximate axisymmetry) permits the acceleration of relatively large fluxes of charge despite the short wavelength employed. Without this innovation, optical accelerators required for large current throughput would not store significant energy per unit length, and the longitudinal wakefields (beam loading) will greatly limit the charge accelerated (Sec. I.B.2). Without loss of generality, we concentrate here on the accelerating and focusing fields for the wide aspect ratio, quasi-2D Cartesian symmetry case. As shown in Sec. I.B.1, one can straightforwardly extend the discussion of TM-like fields to the higher dimension case.

In a DLA with one wide transverse dimension (for example, the x dimension) the beam may be highly asymmetric, taking on a ribbon shape, and thus reflecting the structure symmetry. In practice, one may wish to have 1 to 2 orders-of-magnitude difference between the two rms beam sizes $\sigma_x \gg \sigma_y$. Given constraints on the relative focusing in x and y, this also implies that the emittances are likely highly asymmetric in this case $\epsilon_x \gg \epsilon_y$. This scenario yields two additional effects beyond mitigation of beam-loading limits. The first is the mitigation of quasistatic transverse beam self-forces (space charge), which are highly disruptive at low energy, by the geometric spreading of the beam in x. Also destabilizing transverse wakefields, which present a significant challenge in scaling accelerators from microwave to optical-IR frequencies, are greatly diminished due to geometric effects inherent in the Maxwell equations (Tremaine, Rosenzweig, and Schoessow, 1997). Wakefields in slab-symmetric dielectric structures with highly asymmetric beams have recently been explored experimentally (Andonian *et al.*, 2012).

The dynamics of electrons in DLAs present a number of distinct challenges. The longitudinal dynamics (bunching and acceleration) can be classified by the normalized vector potential $\alpha = eE_{\mathrm{acc}}/mc\omega$ of the resonant accelerating wave, where ω is the frequency, E_{acc} is the peak accelerating field, and m is the particle rest mass (Rosenzweig, 2003). In standard rf electron linacs, α is of the order of unity, while ion linacs have $\alpha \ll 1$ (usually $< 10^{-3}$) due to the higher mass involved. The longitudinal dynamics associated with electron linacs are simpler because the phase motion essentially stops, as relativistic velocities are achieved early in the acceleration process, and in the absence of bending chicanes, there is negligible change in phase. In contrast, ion linac dynamics are more intricate and are described in terms of stable, localized regions in longitudinal phase space around a resonant (nonrelativistic) velocity that are termed *buckets*. In an infrared DLA at the GV/m level, one has $\alpha \simeq 10^{-4}$, due to the high frequency. This motion in a DLA is not, however, strictly comparable to that of an ion rf linac, because in the DLA case the electron Lorentz factor $\gamma = U/m_e c^2$ greatly exceeds unity.

1. Transverse dynamics

With longitudinal periodicity, we can employ the traditional method used in rf linacs for field description, based on a Floquet form. In a system that is periodic in z with period d, the electric field can be written as

$$\mathbf{E}(\mathbf{x}, t) = E_0 \exp[i(k_z z - \omega t)]\mathbf{u}(\mathbf{x}). \tag{29}$$

Here the spatial function $\mathbf{u}(\mathbf{x})$ has the property that $\mathbf{u}(\mathbf{x}_\perp, z + d) = \mathbf{u}(\mathbf{x}_\perp, z)\exp(i\psi)$, where ψ is the Floquet phase, which can be written in the form $\psi = p\pi/q$, with p and q integers ($p \le q$, and $q = N_c - 1$ with N_c being the number of periodic cells in the structure).

Assuming the accelerating fields have little transverse variation (negligible in the wide direction x with subwavelength offsets assumed in the vertical direction $k_z y \ll 1$) we can write the longitudinal electric field in Eq. (29) as

$$E_z = E_0 \sum_{n=-\infty}^{+\infty} a_n \exp(ik_{z,n}\zeta_n), \tag{30}$$

where $k_{z,n} = (2\pi n + \psi)/d$, and we made the substitution $\zeta_n = z - v_{\phi,n}t$, with $v_{\phi,n} = \omega/k_{z,n}$. The field in Eq. (30) is decomposed in terms of spatial harmonics, each having the same temporal frequency, but a different phase velocity. Acceleration is achieved by resonant or secular interaction of a beam electron having velocity v_b nearly equal to one of the $v_{\phi,n}$. The resonant spatial harmonic is traditionally chosen to be the fundamental ($n = 0$) component, with $v_{\phi,n} = \omega d/\psi$, as this harmonic has a much larger amplitude than all the others, with $a_0 \approx 1$ and $a_0 \gg |a_n|(n \neq 0)$. We discuss later the option of resonantly accelerating with a nonfundamental spatial harmonic.

With a wide structure, we ignore all x dependence of E_z, and therefore there are no forces in x. Assuming E_z is symmetric about the $y = 0$ plane, and examining the lowest order field dependence away from $y = 0$, we obtain $E_y \approx -y\partial E_z/\partial z$ and $B_x \approx -y\partial E_z/c^2\partial t$. Substituing Eq. (30) gives

$$\begin{aligned} E_y &\approx -iyE_0 \sum_{n=-\infty}^{+\infty} a_n k_{z,n} \exp(ik_{z,n}\zeta_n), \\ B_x &\approx iyE_0/c^2 \sum_{n=-\infty}^{+\infty} a_n k_{z,n} v_{\phi,n} \exp(ik_{z,n}\zeta_n). \end{aligned} \tag{31}$$

For a particle resonant with the mth spatial harmonic, $v_b \approx v_{\phi,m}$, the transverse (y) Lorentz force on the electron is thus

$$\begin{aligned} F_y &= -e(E_y + v_b B_x) \\ &\approx -iyeE_0 \sum_{n=-\infty}^{+\infty} a_n k_{z,n}(1 - v_b v_{\phi,n}/c^2) \\ &\quad \times \exp(ik_{z,n}z - ik_{z,m}z + \phi), \end{aligned} \tag{32}$$

where ϕ is the phase set by the initial timing of the electron entry into the periodic structure.

The contribution of the resonant harmonic is vertically defocusing for phases that are accelerating with phase focusing [having a negative gradient $\partial(-eE_z)/\partial\zeta$ at the phase of interest] $F_y \approx -ya_m k_{z,m} eE_0 \sin(\phi)/\gamma^2$. This result is expected from the Earnshaw theorem (Earnshaw, 1842) from which one deduces that a phase-stable synchronous acceleration should be accompanied by a net defocusing effect. The amplitude of this force indicates a maximum defocusing strength equivalent to a defocusing betatron wave number $\kappa_{\beta,\max} \simeq \alpha_m^{1/2} k_{z,m}/\gamma^{3/2}$, where $\alpha_m \simeq a_m eE_0/k_{z,m} m_e c^2$. As the resonant normalized vector potential is of the order of 10^{-4} for moderate energy beams ($\gamma < 100$), the defocusing length is of the order of 1000 rad (< mm) of the resonant optical wave. For rf linacs, such an effect may be easily counteracted by use of external focusing. In high-field optical accelerators, however, the defocusing length is too short, and other remedies must be found.

One may consider using a dielectric structure excited in a quadrupole mode that is resonant with the electron beam. The Panofsky-Wenzel theorem (Panofsky and Wenzel, 1956), however, states that $\vec{\nabla}_\perp F_z = \partial \vec{F}_\perp/\partial\zeta$, thus relating the change of acceleration in the transverse direction to the phase-dependent quadrupole focusing. Writing the field in a form reminiscent of the radio-frequency quadrupole $F_z \simeq -eE_0[(y^2 - x^2)/L_q^2]\exp(ik_z\zeta)$, where we define the distance L_q over which the field varies from maximum to zero, the vertically focusing quadrupole field at optimum phase is found to be $F_\perp = -2eE_0 y/k_z L_q^2$. Equating this to the defocusing one must overcome, and taking the maximum accelerating fields to be equal in the accelerating and quadrupole structures, we find that $k_z L_q = \gamma/\sqrt{2}$. This relation implies that at low energy ($\gamma \approx 1$) the useful acceleration region occupies a fraction of an optical wavelength. This makes the approach more suitable at higher energies. In addition, to obtain focusing in both x and y, the quadrupole fields must be varied periodically to give second-order focusing (alternating phase focusing). Xie (2000) proposed an alternating phase scheme producing both second-order (in field strength) focusing and acceleration.

It was recently proposed to utilize the nonresonant spatial harmonics of a periodic DLA structure, as shown in Fig. 4 (Naranjo *et al.*, 2012), which give alternating gradient fields that provide second-order focusing (Rosenzweig and Serafini, 1994; Reiche, 1997). The second-order effect generally produces focusing associated with the nth harmonic proportional to $[qa_n E_0/\gamma(k_{z,m} - k_{z,n})]^2$, and the effect of the spatial harmonics can be enhanced by using harmonics with $k_{z,n}$ near to $k_{z,m}$. Further, one may utilize the fundamental harmonic for focusing and accelerate with a relatively weaker harmonic, overcoming the resonant wave defocusing contribution.

2. Longitudinal dynamics

As indicated previously, the condition $\alpha \ll 1$ implies that the stable particle motion due to the resonant wave is located near in momentum to the resonant value, that is the fractional momentum error $\delta p/p_0 \ll 1$, where $p_0 = \gamma_0 m_e v_0$ is the momentum of the reference particle with velocity $v_0 = v_\phi$. In this case it is possible to use a perturbative analysis for

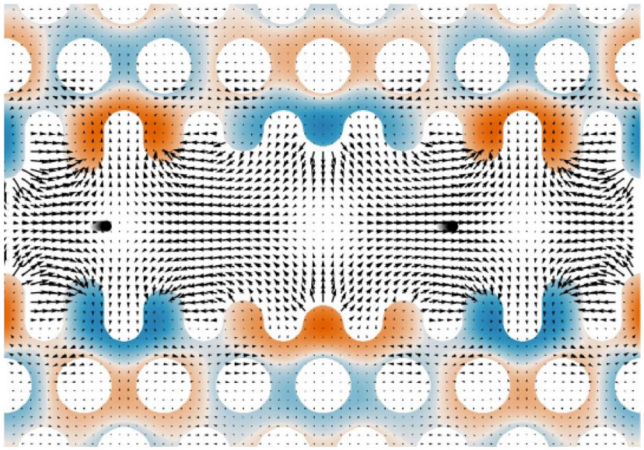

FIG. 4 (color online). Cross section of a biperiodic dielectric laser accelerator based on photonic confinement and optimized for strong second-order focusing. From Naranjo *et al.*, 2012.

describing the motion in DLAs. This approach assumes that the nearby spatial harmonics do not perturb the motion inside of the stable region, or bucket, of longitudinal phase space $(\zeta, \delta p)$. This assumption permits the writing of an approximate resonant Hamiltonian and visualization of the motion through Poincare plots. The validity of the resonant wave analysis is discussed later.

The longitudinal equations of motion in the resonant wave approximation can be written in terms of normalized variables as

$$\frac{d}{d\tilde{z}}\left(\frac{1}{\gamma_0^2}\frac{\delta p}{p_0}\right) = \frac{\alpha_m}{\beta_0\gamma_0^3}(\sin\phi - \sin\phi_0),$$
$$\frac{d\phi}{d\tilde{z}} = \frac{1}{\gamma_0^2}\frac{\delta p}{p_0}, \tag{33}$$

where $\tilde{z} = k_{z,m} z = \omega t$ and $\phi = k_{z,m}\zeta$. The resonant phase $\phi_0 = \pi$ for a nonaccelerating bucket and $\pi/2 < \phi_m < \pi$ for a stable accelerating bucket (Fig. 5). The synchrotron frequency, the frequency of small amplitude oscillations near the fixed point $(\phi_0, 0)$, is given by $\omega_s = \omega\sqrt{\alpha_m|\cos\phi_0|/\beta_0^2\gamma_0^3}$ (Rosenzweig, 2003). The trajectories of these equations are derivable from the resonant wave Hamiltonian,

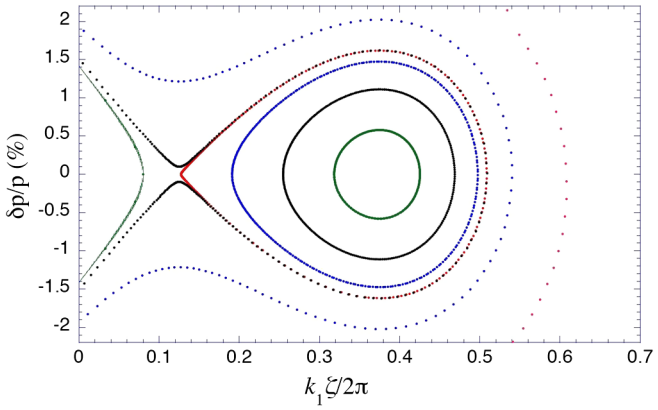

FIG. 5 (color online). Longitudinal phase space showing stable accelerating bucket and unbound trajectories.

$$H(\zeta,\delta p) \simeq m_e c^2 \left\{ \frac{\beta_\phi^2}{2\gamma_\phi p_0^2}(\delta p)^2 + \alpha_m[\cos(k_{z,m}\zeta) - k_{z,m}\zeta\sin(\phi_0) + 1] \right\}. \quad (34)$$

In the case of the nonaccelerating, or stationary bucket, one obtains from the Hamiltonian the bucket height or maximum excursion in momentum at the separatrix $\delta p/p_0 = 2\sqrt{\alpha_m \gamma_0}/\beta_0$. The stationary bucket case is of interest here in the context of adiabatic capture and bunching.

In order not to induce excessive filamentation of the phase space during capture, one should adiabatically raise the resonant wave field. The adiabaticity condition is expressed as the relative rate of change being slower than the synchrotron frequency, quantified as $\dot{\alpha}_m(t)/\alpha_m(t) \ll \omega_s$. When this is obeyed, the area in phase space (the action) that particle orbits occupy is a constant of the motion. In the case of an ion linac, when capture can occur over tens of periods at most, the adiabaticity is poor. In DLA, however, one may perform capture over tens of thousands of periods and still have a process that extends only over about a centimeter in length. In this sense, the capture dynamics in a DLA are similar to those following beam injection into a ring.

The adiabaticity condition is obviously violated at the very start of the capture process, and thus the initial turn-on period sets the action values for the subsequent motion. This effect, along with filamentation (see Fig. 6), dictates the amount of longitudinal emittance growth suffered during capture. We note at this point that nonadiabatic capture has been explored experimentally in laser accelerators, in particular, inverse free-electron lasers (Kimura *et al.*, 2001) but the adiabatic case has not.

Adiabatic capture presents great promise for future applications. The example of capture given in Fig. 6 illustrates the case of electron injection at $\gamma = 6$ in the GALAXIE DLA-driven x-ray FEL project (Rosenzweig, 2012). In this case, a 1 A unbunched (at the optical wavelength level) beam is captured, and accelerated to 1 GeV ($\gamma = 2000$). After adiabatic capture, the peak current is enhanced by about 13 times. In Fig. 7, however, one sees in the final phase space a shortening by another factor of about 80, giving a peak current exceeding 1 kA. This effect is termed "adiabatic compression"

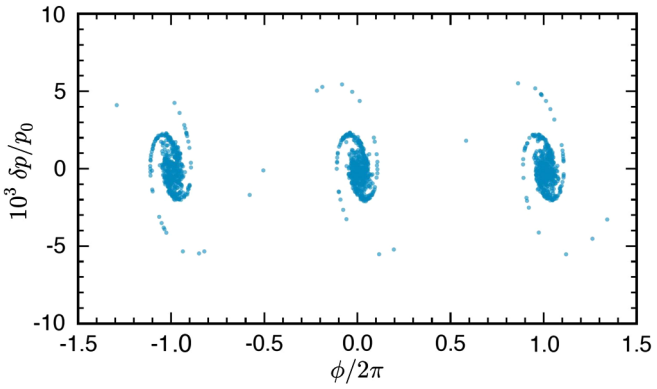

FIG. 6 (color online). Longitudinal phase space during adiabatic capture and bunching.

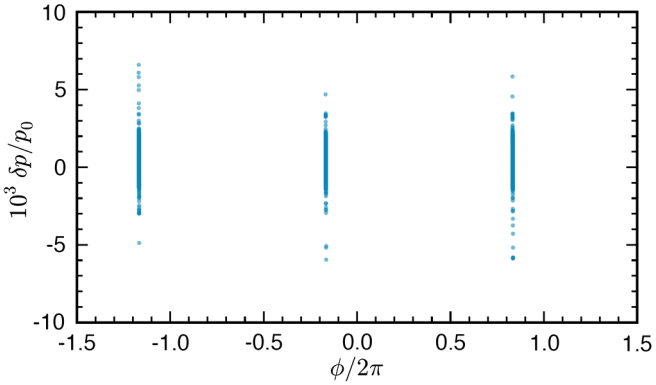

FIG. 7 (color online). Longitudinal phase space after strong acceleration.

and results from the tendency of particle motion to stretch in $\delta p/p_0$ at higher γ. Conservation of longitudinal action then implies that the longitudinal oscillation amplitude scales as $\delta\zeta_{\max} \propto \gamma^{-3/4}$, which in this example is close to 80. Beyond adiabatic capture, adiabatic compression may be crucial in obtaining the peak currents needed in applications such as x-ray FELs based on DLA. In linear accelerators, this new effect is most pronounced in DLA-based approaches. For ion linacs, there is no significant increase in γ for the usual low energies encountered in typical applications. For an IFEL, the longitudinal acceleration interaction scales as $1/\gamma$, and this effect does not occur with notable strength (Musumeci, 2005).

C. Wakefields, impedance, and efficiency

When a charged particle bunch passes near a material boundary, electromagnetic fields generated by the head of the bunch affect following bunch particles through the short-range wakefields or those of following bunches through the long-range wakefields. The wakefields give the voltage change experienced by a test particle within or behind the bunch, per unit charge of the driving bunch. The energy supplied to the wakefields comes at the cost of kinetic energy lost by the bunch particles.

In the case of a metallic boundary, both the electrical resistance of the metal and the geometric variation of the boundary contribute to the wakefields (Chao, 1993; Wangler, 2008). In the case of a dielectric boundary, the Cherenkov radiation condition will be satisfied when the particle velocity $\beta = v/c > 1/n_r = 1/\sqrt{\epsilon_r}$, in which case the beam will generate wakefields. In the case of metallic accelerating structures, the resistive wall contribution to the wakes is considerably weaker than that of the geometrical variation of the structure wall. Similarly, wakefields due to the presence of dielectric material in metallic waveguide are usually much stronger than wakes from the resistive metal wall. Wakefields have both longitudinal and transverse components, which describe the direction of the forces acting on the beam particles. Longitudinal wakefields on average decelerate the bunch, and they distort the voltage gain along the bunch from the ideal harmonic form of the externally applied rf (beam loading). Transverse wakefields are a major source of beam instabilities, such as head-tail instabilities and emittance growth.

Since inception of the DLA concept, its wakefield effects have been intensively studied (Lin, 2001; Schachter, Byer, and Siemann, 2003; Mizrahi and Schachter, 2004b; Siemann, 2004). Normally the complete wakefield solution calls for solving the inhomogeneous wave propagation equations with a point current source (Dirac delta function). This solution is the wake function (or wake potential) with units of volt/coulomb. This is the so-called Green's function or impulse response in a generalized system. The standard approach to obtain the Green's function is to represent the Maxwell equations with the scalar and vector potentials and find the solution for each of them. The wakefield solution of a dielectric-lined waveguide using this approach is given by Ng (1990) and Rosing and Gai (1990). Such a waveguide consists of a cylindrical tube of dielectric with inner and outer radii a and b, respectively, with a conducting boundary located at $r = b$ and vacuum occupying the region $r < a$. Although this configuration would be undesirable for use at optical wavelengths, due to the presence of a conducting boundary, it can provide a simple example for order-of-magnitude analytical calculation of the wakefields. The electromagnetic field radiated by a particle with charge q moving at velocity $v \approx c$ along the axis of the tube has the following longitudinal electric field (E_z) and wake function ($W_\parallel$), summed over the modes j excited by the particle:

$$E_z(0,z_0) = \frac{4q}{\epsilon_r a}\sum_j \left(\frac{F_{00}(k_2 a)Y_0(k_2 b)}{(d/dk_2)[F'_{00}(k_2 a)Y_0(k_2 b) - (k_2 a/2\epsilon_r)F_{00}(k_2 a)Y_0(k_2 b)]}\right)_{k_2=k_{2j}} \cos(\omega_j z_0/c), \tag{35}$$

$$W_\parallel(t) = (1/q)\int_{-\infty}^{+\infty} E_z(0,z_0)dz, \tag{36}$$

where k_{2j} are the roots which satisfy the condition

$$F'_{00}(k_2 a) - \frac{k_2 a}{2\epsilon_r}F_{00}(k_2 a) = 0,$$

and $z_0 = z - vt$. The function F_{00} and its derivative are defined by

$$F_{nn}(k_2 r) = J_n(k_2 r) - \frac{J_n(k_2 b)}{Y_n(k_2 b)}Y_n(k_2 r),$$
$$F'_{nn}(k_2 r) = J'_n(k_2 r) - \frac{J_n(k_2 b)}{Y_n(k_2 b)}Y'_n(k_2 r), \tag{37}$$

where J_n and Y_n are the cylindrical Bessel functions. To obtain the wake of a bunch, one performs the convolution $W_\parallel(z_0) = \int_0^\infty W_\parallel(z)\lambda(z_0 - z)dz$, with $\lambda(z_0)$ being the longitudinal charge distribution.

For a DLA structure with circular cross section, there will be no transverse focusing force from the on-axis particles since there is no radial dependence of E_z (in the vacuum channel) in Eq. (35). That means that the transverse beam profile will not be influenced by the wakefields of the monopole modes. For particles traveling off axis at r_0, the longitudinal wakefield expression is more complicated but can be solved. The results are given in Ng (1990) and Rosing and Gai (1990). The transverse forces can be directly calculated from the longitudinal field E_z by using the Panofsky-Wenzel theorem (Panofsky and Wenzel, 1956), with $F_r = q(E_r - \beta B_\theta) = q\int(\partial E_z/\partial r)dz$ and $F_\theta = q(E_\theta + \beta B_r) = q\int r^{-1}(\partial E_z/\partial\theta)dz$.

The DLA structure is essentially an externally powered, constant impedance traveling-wave accelerating structure. Its wakefields can be calculated using another important parameter, the loss factor k_l, which is linked to energy loss of the driving particle $U_l = k_l q^2$, as well as the wakefield at the origin $W_\parallel(0+) = |E_z|_{\max} = 2k_l q$. The loss factor is calculated using

$$k_l = \sum_j k_{lj} = \sum_j \frac{\omega_j}{4}\left[\frac{R_{\rm sh}}{Q}\right]_j \frac{1}{(1-\beta_{gj})}, \tag{38}$$

where the subscript j enumerates the waveguide modes excited by the traversing particle (Chojnacki *et al.*, 1993; Bane and Stupakov, 2003; Jing, 2004). The shunt impedance per unit quality factor $R_{\rm sh}/Q$ is a figure of merit to compare different accelerating structures, with units of ohms per meter, and β_g is the group velocity of the propagated wakefield divided by the speed of light. The wakefields can be obtained by solving the homogenous wave equation in the structure and calculating the loss factor for each guided mode that has a field pattern matching with the corresponding wakefield mode. For example, the TM_{0n} modes dominate the longitudinal wakefield excited by the on-axis particles in a dielectric-lined waveguide. Assuming a Gaussian bunch, its longitudinal wakefield can be calculated as

$$W_\parallel(r=0,z_0) = \frac{q}{\sqrt{2\pi}\sigma_z}\sum_j k_{lj}\int_{-\infty}^{z_0}\cos(\omega_j z_0/c)\exp\{-[(z_0-z)/2\sigma_z]^2\}dz. \tag{39}$$

For convenience, a few commonly used parameters to characterize a rf accelerating structure and its relation to the most important figure of merit, the characteristic interaction impedance Z_c, is summarized in Table II. Consider the example of the Lin fiber (Lin, 2001), which has $Z_c = 19.5\ \Omega$ at $\lambda = 1\ \mu$m and $\beta_g = 0.58$. Assuming the single mode wakefield excitation on axis by a Gaussian bunch with bunch length of $\lambda/20$ and charge of 10^{-15} C, the fundamental mode loss factor $k_{l0} = 2$ GV/pC. The wakefield is plotted in Fig. 8.

However, unlike the conventional traveling-wave rf accelerator in which the wakefields consist of only discrete waveguide modes, wakefields in DLA structures, due to lack of metal boundary, include not only the discrete waveguide modes but also the broadband radiation which is not confined

TABLE II. Commonly used parameters for rf and dielectric laser accelerators.

rf accelerator	DLA
Characteristic impedance Z_c (Ω)	$G^2\lambda^2/P$
R_{sh}/Q (Ω/m)	$Z_c\beta_g/(2\pi\lambda)$
Attenuation coefficient α_{loss} (1/m)	$\ln(10)L_{att}(dB/m)/10 = 4\pi\mathrm{Im}(n_{eff})/\lambda$
Shunt impedance R_{sh} (Ω/m)	$Z_c/(2\alpha_{loss}\lambda^2)$
Quality factor Q	$\pi/(\alpha_{loss}\lambda\beta_g)$
Loss factor k_{l0} (fundamental)	$Z_c\beta_g/[4\epsilon_0 Z_0\lambda^2(1-\beta_g)]$

in the optical fiber. The total loss factor in this case is different from Eq. (38) and calculated by Schachter, Byer, and Siemann (2002), $k_l = 1/2\pi\epsilon_0 R_{eff}^2$, where R_{eff} is the effective radius of the channel that the beam traverses. Again, taking the example of the Lin fiber, which has a radius of the beam hole $R = 0.678\lambda$, the total loss factor k_l is calculated to be 39 GV/pC. It should be pointed out that this total loss factor significantly increases the short-range wake but has much less effect on the long-range wake, since the broadband radiation wake decays exponentially in time and space.

Alternatively, the wakefields can be solved in the frequency domain. We introduce the wake impedance, the quantity that characterizes the frequency content of the wakefields. The wake function and wake impedance are a Fourier transformation pair. The wake impedance in DLA structures is defined as

$$Z_w = |E_{dec}\lambda|^2/P_w = Z_0\lambda^2/2\pi R_{eff}^2, \tag{40}$$

where $P_w = qcE_{dec}$ is the power generated by particles with charge of q, and E_{dec} is the self-decelerating field of the particles (Schachter, Byer, and Siemann, 2002). Using the expression for k_l, we have $Z_w = 2\pi\lambda k_l/\omega$.

All the well-developed theories in rf accelerators regarding beam loading and efficiency can be applied to the DLA. We introduce the beam-loading enhancement factor $B = k_l/k_{l0}$, which is defined as the ratio of the total energy loss of the particle to the energy being transferred to the fundamental mode. Then from Wilson (1991), the maximum rf-to-beam efficiency in the case of single pointlike bunch acceleration is

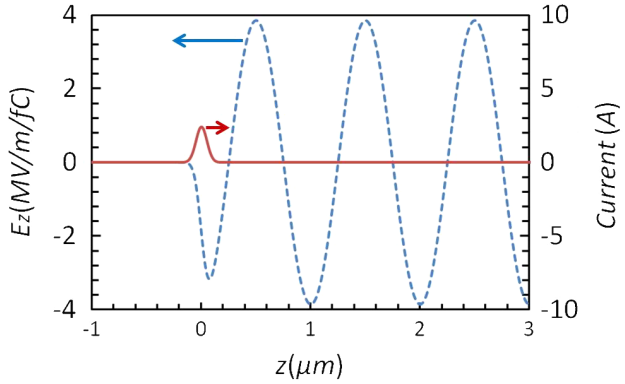

FIG. 8 (color online). Fundamental mode of the longitudinal wakefield (left axis, solid curve) generated by a Gaussian bunch (right axis, dashed curve) in the Lin fiber accelerator.

$$\eta_{s-max} = \cos^2\theta/B, \tag{41}$$

and the beam induced electric field is

$$E_{b0} = 2k_{l0}q = G_0\cos\theta/B, \tag{42}$$

where θ is the phase of the external electrical field just before the charge arrives at the structure, and G_0 is the initial gradient set by the input laser. Again, using the loss factors of the Lin fiber obtained earlier, we calculate the maximum single bunch efficiency to be 5% at the bunch charge $q_{opt} = G_0/2k_l$. With the assumption of a 7.4 kW laser input (based on the damage threshold of the Lin fiber), the charge to achieve the best efficiency is about 5 fC.

For a Gaussian distributed bunch, the above efficiency is modified by a form factor F to become $\eta_{s-max,G} = F^2\cos^2\theta/B$, where $F = \exp[-(2\pi\sigma_z/\lambda)^2/2]$ for a Gaussian spatial bunch distribution $f(z) = (q/\sqrt{2\pi}\sigma_z)\exp(-z^2/2\sigma_z^2)$.

When a single electron bunch passes through a DLA structure, it excites not only the fundamental accelerating mode but also other higher-order modes and broadband unconfined radiation. However, if this single bunch breaks into multiple microbunches, only the mode which has the same frequency as a multiple of beam frequency (microbunch repetition rate) can be coherently enhanced. Other modes will be out of resonance and will detune. Therefore, the beam loading and efficiency calculation is more complicated in the case of bunch train acceleration. The maximum rf-to-beam efficiency in acceleration by a bunch train is given by (Schachter, 2004)

$$\eta_{t-max} \simeq 12(1-\beta_g)\beta_g^2\xi(1-\xi)/(\xi^2+3\beta_g^2) \tag{43}$$

at the optimal total charge of the bunch train $q_{opt} \simeq G_0\xi/k_l$, where $\xi = 3\beta_g^2(\sqrt{1+\beta_g^{-2}}-1)$. Equation (43) reveals that the best efficiency has a strong dependence on the group velocity of the structure and so can be improved significantly. For example, the efficiency of the Lin fiber can reach 35% by using a bunch train with a total charge of 6 fC. This efficiency is at a level comparable to conventional rf accelerators. In order to increase the total accelerated charge while maintaining or further improving the efficiency, a laser recovery system is needed (Schachter, 2004; Siemann, 2004; Na, Siemann, and Byer, 2005).

D. Laser damage thresholds of materials

The laser damage threshold is a fundamental limit for any laser-driven accelerator and is set by the material properties of the accelerator structure. Quantitatively stated, the maximum accelerating gradient of an accelerator structure scales as $G \propto \sqrt{F_{th}}$, where G is the gradient and F_{th} is the laser damage threshold fluence (energy per unit area). Compared to metals, dielectric materials can tolerate orders-of-magnitude higher electric fields before material breakdown occurs, making dielectric structures ideal for a high-gradient laser-driven accelerator (Soong *et al.*, 2012b). However, even among dielectric materials, there is a significant variation in the material breakdown limit. Therefore, a thorough understanding

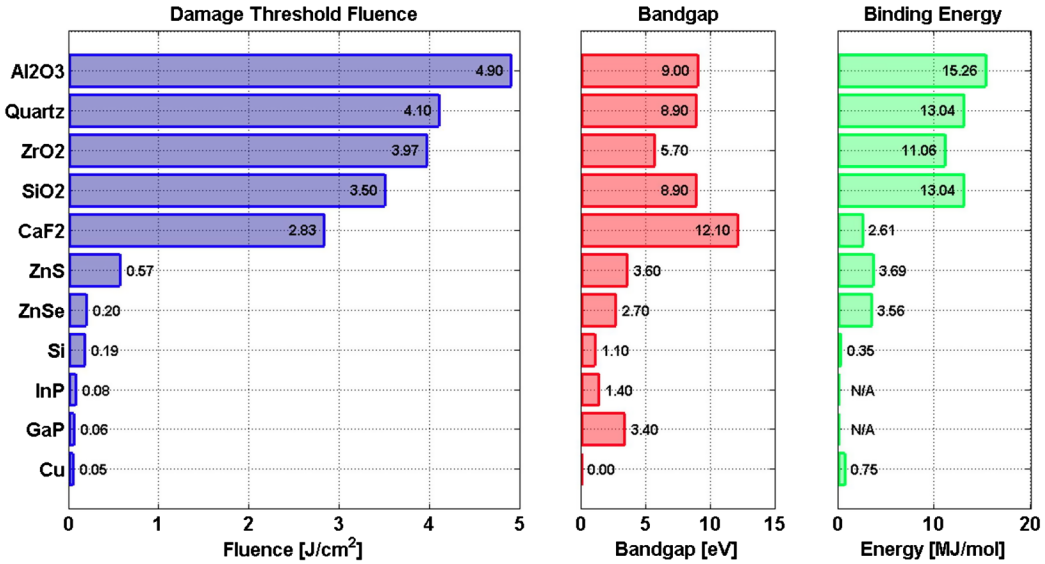

FIG. 9 (color online). The laser damage threshold of a variety of optical materials, as well as copper, for comparison. Each of these measurements were conducted with a 1 ps, 800 nm, 600 Hz Ti:sapphire laser. The data indicate a strong positive correlation between the damage threshold, the material band gap, and the material's lattice binding energy. From Soong *et al.*, 2012b.

of the relationship between the dielectric intrinsic properties and laser damage threshold is critical to the development of an optimal dielectric laser-driven accelerator.

For the application of a dielectric laser-driven accelerator, the specific interest lies in material damage from repeated ultrafast laser pulses, typically in the infrared regime. Such material data are uncommon in the published literature, which necessitates a dedicated laser damage experiment specifically designed to characterize materials for laser-driven accelerators, such as the experiment described by Soong *et al.* (2011). In this experiment, repeated 800 nm, 1 ps laser pulses from a 600 Hz Ti:sapphire laser are incident normal to the material sample under investigation. The power of the Ti:sapphire laser is varied at different locations across the sample, while a low power HeNe laser is used as a probe for surface damage. Through large sample sets, the laser damage threshold fluence is statistically determined, with the laser damage threshold fluence defined as the laser fluence where damage to the sample has a 50% probability of occurrence. The aggregated results of these laser damage threshold tests on polished bulk materials are qualitatively shown in Fig. 9. The data show a strong positive correlation between the laser damage threshold fluence and the material band gap and lattice binding energy, although no quantitative relation is presented.

In addition to the variation of laser damage threshold with intrinsic material properties, the laser damage threshold is also a function of the drive-laser parameters, namely, laser pulse length and laser wavelength. For laser pulses between 1 ns and 10 ps, the laser damage threshold has been well characterized and a relation between the laser pulse length and the laser damage threshold fluence is well known (Stuart *et al.*, 1995), as shown in Fig. 10. In this long pulse regime, the laser damage threshold scales with the root of the laser pulse length $\tau^{1/2}$, which is characteristic of the transfer and diffusion of electron kinetic energy to the material lattice. The laser damage can be described as a thermal melting and boiling of the material surface and occurs over the entire area of laser irradiation.

The deviation from the $\tau^{1/2}$ dependence at shorter pulses is the result of a change in the primary breakdown mechanism. At pulse lengths below 10 ps, the damage mechanism shifts from thermal heating to surface ablation associated with multiphoton ionization. Correspondingly the damage morphology in the short pulse regime appears as pitted craters and is limited to a small region where the laser intensity is highest. Soong *et al.* (2012b) described the damage process as a resonance between the electric fields of the incident laser pulse and the multiphoton-ionized electrons in the material. In this model, a strong coupling between the laser fields and the electrons occurs when the frequency of the drive laser (ω_{laser}) matches the electron plasma frequency (ω_p),

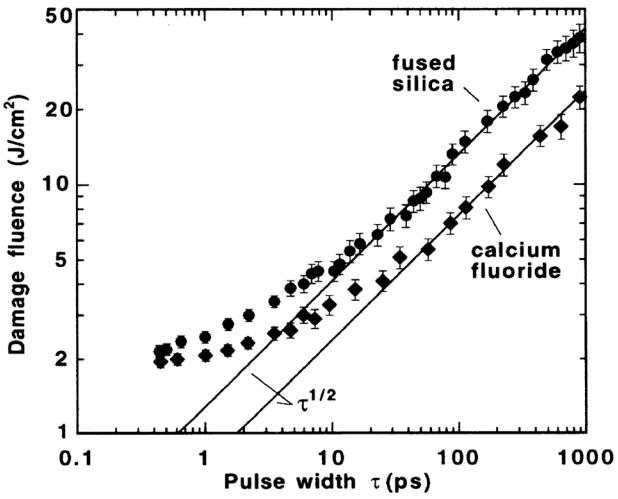

FIG. 10. Observed values of laser damage threshold at 1053 nm for fused silica (circle) and CaF_2 (diamond). Solid lines are $\tau^{1/2}$ fits to long pulse results. Estimated uncertainty in the absolute fluence is about 15%. From Stuart *et al.*, 1995.

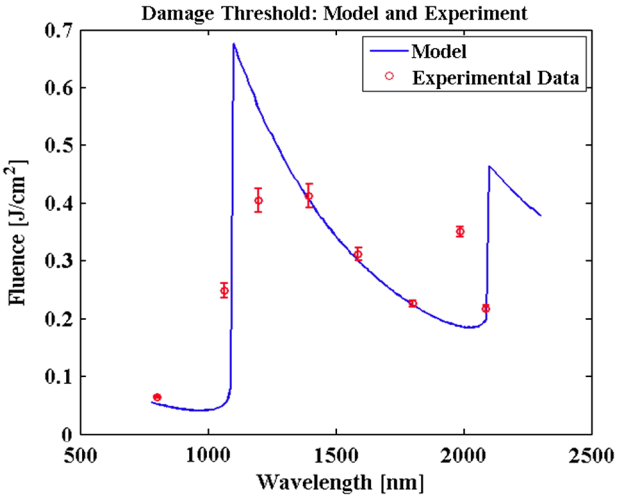

FIG. 11 (color online). Laser damage threshold of silicon for picosecond laser pulses in the near to short wavelength infrared. The damage model (solid curve) is superimposed on top of the experimental data. From Soong *et al.*, 2011.

$$\omega_{\text{laser}} = \omega_p = \sqrt{\frac{4\pi n_e e^2}{m^*}}, \tag{44}$$

where n_e is the plasma density, e is the electron charge, and m^* is the effective electron mass. This model of the laser damage mechanism explicitly relates the laser damage fluence to the laser wavelength and has been experimentally supported, as shown in Fig. 11.

III. DIELECTRIC ACCELERATOR STRUCTURES

A. Grating accelerator

The transmission-based, dual-layer grating structure, initially proposed by Plettner, Lu, and Byer (2006), consists of two gratings on either side of a confined vacuum channel. Laser light is incident transversely on the grating with the laser field parallel to the channel. The structure is not a PBG arrangement but operates as a phase-reset device in which the gratings determine the relative phase slippage of the laser front for each period longitudinally, insuring the correct accelerating phase on axis. This design eliminates the pulse-slippage problem (electron beam at speed $v \simeq c$ outrunning the laser pulse traveling at group velocity $v_g < c$) encountered in conventional waveguide-based accelerators by allowing the structure to be optically pumped perpendicular to the electron beam channel. Since the fields traverse very little material, dispersive effects are minimized, allowing the structure to be driven by ultrashort high-power laser beams. Also, because the incident field is allowed to leave the vacuum gap and reform as a plane wave at the opposite grating, it is not a resonant structure and there is no energy storage (no optical quality factor Q) so the exposure time to such high fields is reduced, enabling sustained operation at high gradients.

The structure acts as a longitudinally periodic phase mask that produces a speed-of-light space harmonic on axis. Each grating pillar adds a π phase shift with respect to the adjacent vacuum space so the resulting fields accelerate a passing electron bunch when the phase of an input TM field is properly chosen. Figure 12 shows snapshots of the longitudinal field in a single grating period, calculated with a finite difference time domain (FDTD) code, taken every quarter cycle of a 800 nm wavelength beam incident from the bottom. An electron bunch (shown as a black and white circle), launched at optimal phase, traverses the 400 nm gap structure from left to right. One can see that the grating minimizes the beam's exposure to the decelerating phase of the field so it experiences a net acceleration every period.

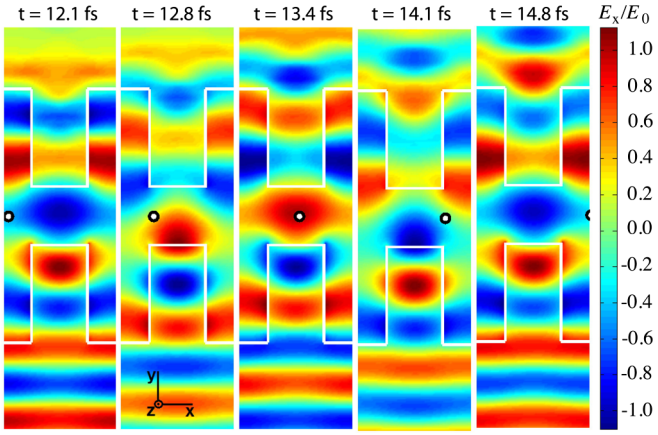

FIG. 12 (color online). Snapshots of the longitudinal field (E_x) in the grating structure at five points in time across one period, showing its principle of operation: each grating pillar provides a π phase delay such that a correctly phased particle experiences net acceleration. From Peralta *et al.*, 2012.

1. Phase synchronicity and forces

The grating periodicity allows the fields to be written as a superposition of space harmonics. The TM wave's magnetic field in the vacuum channel is

$$B_z(x,y,t) = e^{(i\omega t - i\phi)} \sum_{m=-\infty}^{\infty} W_m(y) e^{imk_p x}, \tag{45}$$

where $\omega = ck$ is the laser angular frequency, ϕ is the phase between the laser and the particle, $W_m(y)$ describes the amplitude of the mth diffraction order in the grating with wave vector k_p, and the coordinates are defined with $\hat{z}$ being parallel to the grating grooves, $\hat{x}$ being perpendicular to the grooves but along the vacuum channel where the electron beam travels, and the laser beam incident along y. With $\nabla \times B = -\mu\epsilon i\omega E$, we find similar expressions for $E_x(x,y,t)$ and $E_y(x,y,t)$ but with $U_m(y) = -(ic/k)dW_m(y)/dy$ and $V_m(y) = -(cmk_p/k)W_m(y)$ as their amplitude in the sum.

The resulting Lorentz force is then calculated as $F = q(E + v \times B)$, where the particle velocity is $v = \beta c(\hat{x}\cos\alpha_g + \hat{z}\sin\alpha_g)$ in the general case of a grating tilt angle α_g. The ideal acceleration case corresponds to $\alpha_g = 0$. The average acceleration force over some distance $s(T) = \beta cT$ traveled by the particle is

$$\overline{F_x}(y,\phi) = q\Re\left\{\frac{e^{-i\phi}}{s(T)}\int_0^{s(T)} ds \sum_{m=-\infty}^{\infty} U_m(y) e^{i\psi_m(s)}\right\}, \tag{46}$$

where $\psi_m(s) = s(mk_p\cos\alpha_g + k/\beta)$ is the path-dependent phase of a given mode which must be zero to yield a nonzero

contribution to the integral, giving the phase-synchronicity condition

$$mk_p + k/(\beta \cos \alpha_g) = 0. \tag{47}$$

Solutions to Eq. (46) for various grating configurations with dual laser illumination are given by Plettner, Byer, and Montazeri (2011). For the case of single laser illumination and a relativistic beam ($\beta \simeq 1$) traveling perpendicular to the grating grooves ($\alpha_g = 0$) it is evident that the first-order mode ($m = -1$) provides the acceleration force once the grating period is matched to the laser wavelength. This condition means that the usual form of $W_m(y) = w_m^+ e^{\Gamma_m y} + w_m^- e^{-\Gamma_m y}$, with $\Gamma_m = \sqrt{(mk_p)^2 - k^2}$, from separation of variables using Eq. (45) in the wave equation, must be replaced with $W_{-1}(y) = c_1 + c_2 y$ since $\Gamma_{-1} = 0$. This can be written as $W_{-1}(y) = w_m^+(1 - \kappa y) + w_m^-(1 + \kappa y)$ and substituted back into Eq. (46) to yield an average acceleration force $\overline{F_x}(y, \phi) = cq(\kappa/k)(w_m^- - w_m^+) \sin \phi$ which is invariant in the transverse coordinates and depends only the relative phase between the electron and laser beams.

2. Design and simulations

Computation of Eq. (46) to determine the optimal structure geometry that maximizes the resulting acceleration gradient has been carried out using a 2D FDTD code as well as 3D finite element frequency domain simulations. The gradient calculation is done for a single relativistic particle traveling through the vacuum channel center ($y = 0$) of a single period fused silica grating structure with periodic boundary conditions and matched to the drive wavelength of a Ti:sapphire laser ($\lambda_p = \lambda = 800$ nm). Snapshots of the calculated longitudinal fields are shown in Fig. 12 with the resulting optimization curves $\overline{F_x}(0, \phi_{\max G})$ and $\overline{F_y}(0, \phi_{\max G})$ shown in Fig. 13.

Figure 13(a) shows that there are optimal pillar heights for a given gap size (e.g., $0.825\lambda_p$ when the gap is $\lambda_p/2$). It is also evident from Fig. 13(b) that the acceleration gradient decreases as the gap increases, while the corresponding unwanted deflection forces (which push the electron beam toward the grating surface) remain fairly constant at around 5% of the magnitude of the acceleration force. This small, nonzero deflection is expected from our treatment above since the instantaneous force behaves as $F_y(x, y, t, \phi) = cq(k_p/k - \beta)[(w_{-1}^- + w_{-1}^+) + (w_{-1}^- - w_{-1}^+)y]e^{i(wt - k_p x - \phi)}$ and

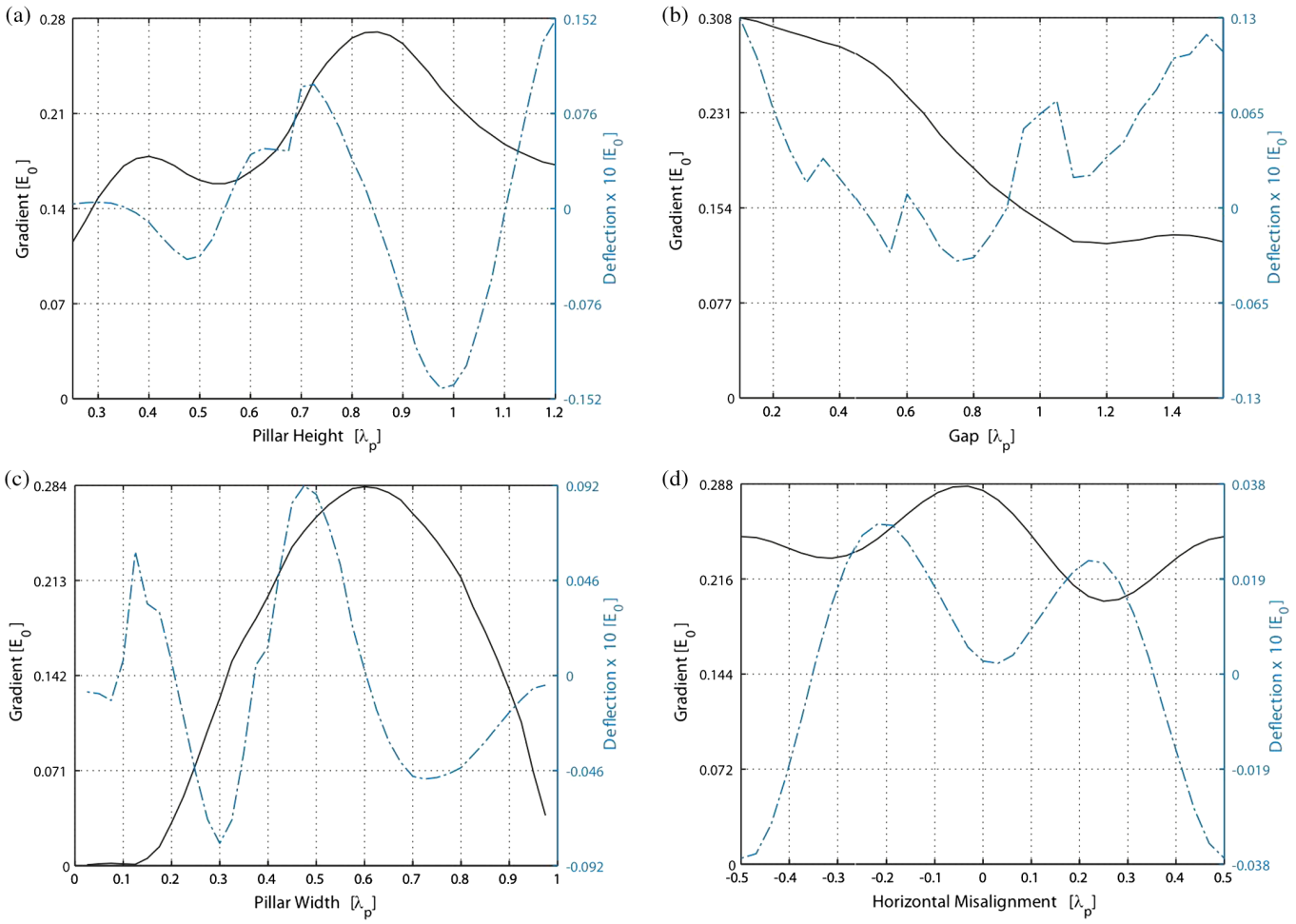

FIG. 13 (color online). FDTD optimization of acceleration gradient (solid curves, left axis) with corresponding net deflection force (dashed curves, right axis) as a function of (a) grating pillar height with a fixed gap of 400 nm, (b) grating-to-grating gap with an optimal pillar height, (c) grating pillar width with a fixed gap of 400 nm and pillar height of 620 nm, and (d) grating-to-grating longitudinal misalignment with a fixed gap of 400 nm, pillar height of 620 nm, and pillar width of 480 nm.

even though the period is matched to the drive wavelength ($k_p = k$), β is not exactly 1. Nevertheless, the phase of maximum gradient ($\phi_{\max G}$) coincides almost perfectly with one that shows zero net deflection ($\phi_{\min D}$) which could also be used at the expense of a reduced gradient.

Figure 13(c) confirms that the optimal pillar widths at the optimized pillar height for a given gap are near the expected $\lambda_p/2$ but the gradient improves with slightly wider pillars for the smaller gaps (i.e., $0.6\lambda_p$ when the gap is $\lambda_p/2$). The effect of the longitudinal misalignment between the top and bottom gratings on the gradient was also investigated and is shown in Fig. 13(d), where it can be seen that the two gratings need not be perfectly aligned in order to obtain a significant acceleration gradient (i.e., for a $\lambda_p/2$ gap structure the worst misalignment reduces the gradient only by a factor of 1/4). From these results we can estimate that in an ideal scenario with an aligned, 100 nm gap structure using 10 fs pulse-front tilted laser pulses one can achieve loaded gradients of $G_L \sim 2.4$ GeV/m, while a more realistic experiment with a misaligned, 1200 nm gap structure driven by 3.3 ps flat pulse-front laser pulses can still give a gradient of $G_L \sim 28$ MeV/m. This opens the possibility of manufacturing such grating structures without strict subwavelength alignment tolerances if lower gradients are acceptable.

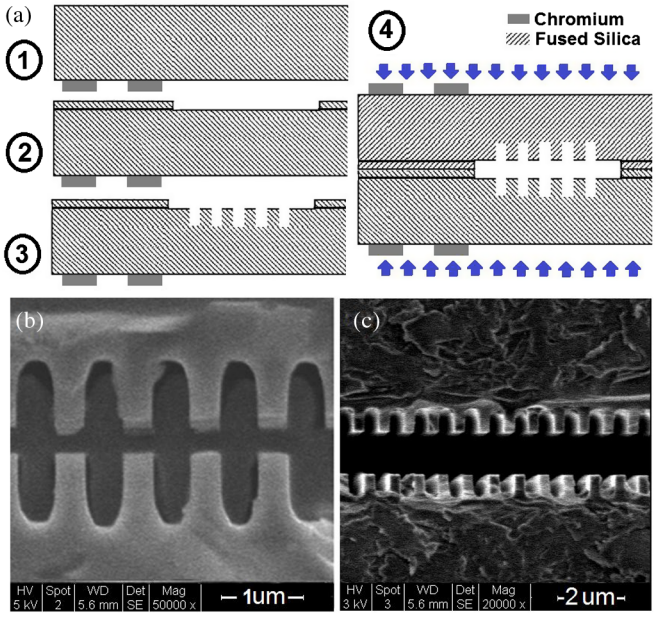

FIG. 14 (color online). Fabrication process for the (a) dual-layer grating structure, and cross-sectional SEM images of resulting (b) 400 nm and (c) 1200 nm gap structures. Adapted from Peralta *et al.*, 2012.

3. Fabrication

Prototyping of the dual-grating phase-reset structure was recently carried out at Stanford University and SLAC National Accelerator Laboratory. The process to fabricate the dual-layer grating structures was purposely designed to employ standard complementary metal-oxide-semiconductor (CMOS) techniques. Avoiding *e*-beam lithography and focused ion beam etching and using instead optical lithography and reactive ion etching techniques reduces the complexity and expense of the process, allowing for high throughput and reproducibility. Initial efforts were based on a monolithic approach (Peralta, 2011), however, difficulties depositing the required film thicknesses and the realization that longitudinal alignment is not as critical as previously believed gave way to the wafer-bonding approach (Peralta *et al.*, 2012) shown in Fig. 14(a).

The fabrication process is relatively simple: (1) A layer of chromium is deposited by *e*-beam evaporation and alignment marks are defined via lift-off, then (2) a trench defining the vacuum channel is etched onto the fused silica to a depth of half the desired structure gap. Step (3) is to etch the grating inside this trench via magnetically enhanced reactive ion etching and the final step (4) is to bond two such wafers together. Figures 14(b) and 14(c) show two cross-sectional scanning electron microscopy (SEM) images of grating structures made with this process. The measured period is 796 ± 4 nm, measured grating pillar width is 353 ± 9 nm, and measured gaps are 365 ± 14 and 1142 ± 21 nm, respectively.

4. Gradient demonstration experiments

The first DLA demonstration experiments were conducted recently using grating-based accelerators, including the demonstration of high-gradient (300 MV/m) speed-of-light electron acceleration in a dielectric structure (Peralta *et al.*, 2013) and the first demonstration of nonrelativistic acceleration at optical wavelengths (at 25 MV/m) in a similar structure (Breuer and Hommelhoff, 2013). In the former experiment, the dual-grating structures described in Sec. III.A.3 were diced into individual samples and acceleration beam tests were carried out at SLAC's NLCTA facility using a 60 MeV preaccelerated electron beam. The experimental layout is illustrated in Fig. 15(a). In these experiments, the signal of interest was a change in the energy spectrum of the transmitted distribution that was correlated with the presence of the IR laser incident on the structure. Since the electron beam was ~1 ps long, as opposed to being a small fraction of the laser wavelength (~attosec duration), the electrons sampled all phases of the laser field, and the interaction presented itself as an energy modulation or broadening of the transmitted energy spectrum. By comparison of the modulated energy spectrum with both simulated and analytical predictions, the corresponding accelerating gradient was determined. A plot of measured gradient at various laser pulse energies is shown in Fig. 15(b), yielding the expected linear dependence on incident laser electric field E_0 (or square root dependence on pulse energy), and showing maximum achieved gradients up to 300 MV/m, which is an order of magnitude beyond typical operating gradients in conventional accelerators. At Max Planck Institute for Quantum Optics, these results have been extended to the nonrelativistic regime using an open grating operating at the third spatial harmonic of the same (800 nm) drive-laser wavelength, and yielding observed gradients as high as 25 MV/m (Breuer and Hommelhoff, 2013) with a 28 keV electron microscope as the particle source. These results are important for the development of fully integrated systems with a compact injector that can accelerate particles from near rest to ultrarelativistic energies. When combined with DLA-matched electron sources these techniques can permit phase-stable net acceleration of

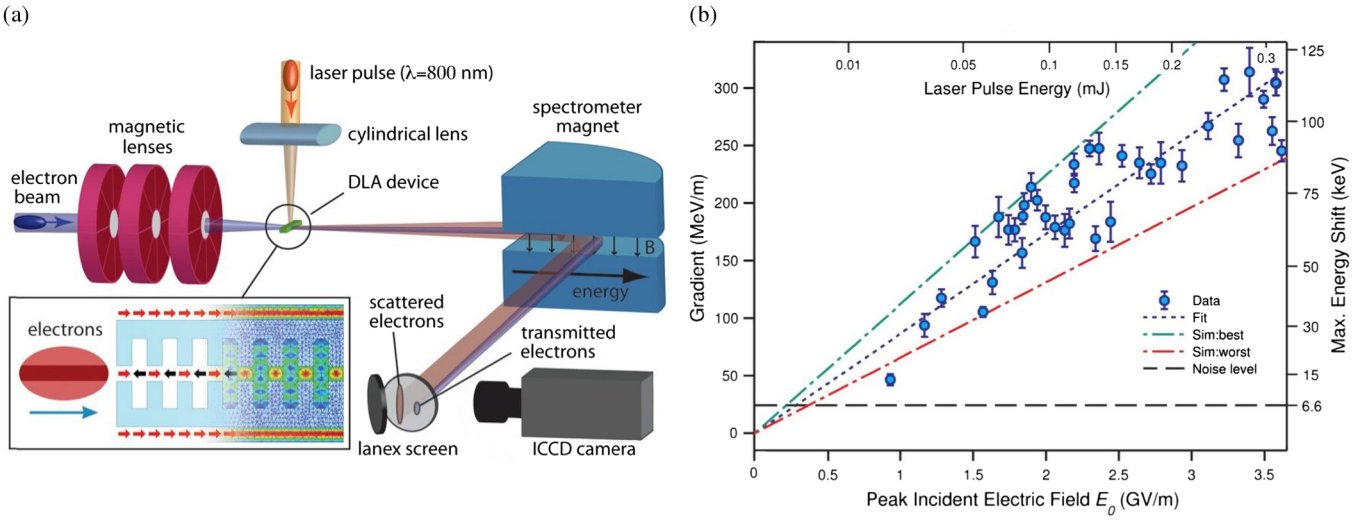

FIG. 15 (color online). (a) The experimental setup and (b) measured accelerating gradient as a function of laser electric field for the dual-grating acceleration experiment at SLAC. Each data point in (b) corresponds to a fitted curve of energy modulation vs relative timing of the laser and electron beam. From Peralta *et al.*, 2013.

synchronous attosecond microbunches over many accelerating stages.

As indicated in Sec. III.A.2, GV/m gradients should be achievable by this approach with further refinement of the material and structure design. In one example, a monolithically fabricated silicon variant of the grating concept referred to as a *buried grating accelerator* was proposed by Chang and Solgaard (2014). Making the structures out of silicon provides access to a wide range of thin-film materials and deposition methods for modifying the surface chemistry in order to create atomically smooth, defect-free, high-band-gap surfaces with increased breakdown fields.

B. Bragg and omni-guide accelerators

Bragg reflection waveguides are one-dimensional periodic structures, designed to guide light in a low refractive index surrounded by alternating layers of high refractive index. Two possible configurations to be considered here: one is the cylindrical Bragg fiber and the other is the planar Bragg reflection waveguide, depicted in Figs. 1(c) and 1(d), respectively. The theory of Bragg reflection waveguides was developed by Yeh and Yariv (1976), and recently there has been a growing interest in using such hollow cylindrical structures as low-loss optical fibers in long distance telecommunications. As noted earlier, Bragg waveguides are now recognized as finite-layer PBG structures in which reflection and mode confinement are perfect only when the number of layers becomes infinite.

Most of the studies carried out on Bragg reflection waveguides dealt with configurations where all dielectric layers are transverse quarter-wave thick. However, controlling the dispersion properties for confined modes in Bragg fibers was demonstrated by creating a defect in the form of changing one or more of the layer widths. In this section, a systematic adjustment of the layer adjacent to the core defect region will be shown to change the waveguide properties to fulfill specified requirements of an acceleration structure.

Motivated by the requirement $v_{\mathrm{ph}} = c$, Mizrahi and Schachter (2004a) have developed a general method for designing the Bragg waveguide for a given phase velocity, given the core dimension and a set of dielectric materials. The core dimension itself may be dictated by other considerations, such as the maximum field allowed to prevent material breakdown, beam dynamics, and the interaction efficiency. The modes of interest are the symmetric TM modes, summarized in Table III. The examples presented in this section are for waveguides made from dielectric materials with refractive indices 1.6 and 4.6, and the core is a vacuum with an index of 1. In addition, we neglect the diffractive losses due to the finite cladding, assuming the multilayer structure is a good approximation to an infinite PBG structure.

1. Matching layer

We consider the planar Bragg reflection waveguide (uniform in the y direction), with core half-width D_{int}, as depicted in Fig. 1(d). We assume that at some specific wavelength λ_0 with a corresponding angular frequency ω_0, this waveguide is required to support a symmetric TM mode with a specific phase velocity $v_{\mathrm{ph}} = \omega_0/k_z$, with k_z being the longitudinal wave number. Equivalently, a specific field distribution in the

TABLE III. Hollow-core symmetric modes for the general and speed-of-light (SOL) synchronous cases. The transverse wave number is $k_\perp = \sqrt{\omega^2/c^2 - k_z^2}$. In the planar case $k_x = k_\perp$ and in the cylindrical case $k_r = k_\perp$.

Mode type	General mode $\times \exp(-ik_z z)$	SOL mode $\times \exp(-i\omega z/c)$
Planar TM	$E_z = E_0 \cos(k_\perp x)$	$E_z = E_0$
	$E_x = i(k_z/k_\perp)E_0 \sin(k_\perp x)$	$E_x = i(\omega x/c)E_0$
	$H_y = i(\omega/ck_\perp)E_0/Z_0 \sin(k_\perp x)$	$H_y = i(\omega x/c)E_0/Z_0$
Cylindrical TM	$E_z = E_0 J_0(k_\perp r)$	$E_z = E_0$
	$E_r = i(k_z/k_\perp)E_0 J_1(k_\perp r)$	$E_r = i(\omega r/2c)E_0$
	$H_\phi = i(\omega/ck_\perp)E_0/Z_0 J_1(k_\perp r)$	$H_\phi = i(\omega r/2c)E_0/Z_0$

hollow core may be required, determined by the transverse wave number $k_x = \sqrt{\omega^2/c^2 - k_z^2}$, as shown in Table III. Expressing the phase velocity in terms of the transverse wave number in the core, we obtain $v_{\mathrm{ph}} = c/\sqrt{1-(k_x\lambda_0/2\pi)^2}$. Note that for a speed-of-light mode $k_x = 0$, which implies a uniform field in the core.

The electromagnetic field components (modulo the usual space-time phase factor $\exp[i\omega t - ik_z z]$) in the layer adjacent to the core (the matching layer), which has a dielectric coefficient ε_1, are given by

$$E_z = A_1 \exp(-ik_1 x) + B_1 \exp(ik_1 x), \tag{48}$$

$$E_x = -(k_z/k_1)[A_1 \exp(-ik_1 x) - B_1 \exp(ik_1 x)], \tag{49}$$

$$H_y = -Z_1^{-1}[A_1 \exp(-ik_1 x) - B_1 \exp(ik_1 x)], \tag{50}$$

where the transverse wave number is $k_1 = \sqrt{\varepsilon_1\omega_0^2/c^2 - k_z^2} = k_z\sqrt{\varepsilon_1\beta_{\mathrm{ph}}^2 - 1}$ and the transverse impedance is $Z_1 = Z_0 k_1 c/\omega_0\varepsilon_1$. The required electromagnetic field in the vacuum core E_0 dictates the amplitudes in the first layer, by virtue of the continuity boundary conditions on E_z and H_y at the interface between the core and the adjacent dielectric layer. Imposing the boundary conditions on these two fields at $x = D_{\mathrm{int}}$, the amplitudes are found to be

$$A_1/E_0 = (B_1/E_0)^* = \frac{1}{2}\exp(ik_1 D_{\mathrm{int}})\cos(k_x D_{\mathrm{int}}) - i\frac{k_1}{2\varepsilon_1 k_x}\exp(ik_1 D_{\mathrm{int}})\sin(k_x D_{\mathrm{int}}). \tag{51}$$

The electromagnetic field required in the core entails the amplitudes of the outgoing and incoming transverse waves in the first layer, as calculated earlier. It is now our goal to ensure that the complete structure, including the Bragg reflector, indeed supports the required field at the given wavelength as an eigenmode. This requires a matching of the first layer fields to those in the Bragg reflector layers.

The Bragg reflector can be analyzed from the perspective of a pure periodic structure according to the Bloch (or Floquet) theory of Sec. II.A. This analysis gives the eigenvectors and the eigenvalues of the periodic structure, and determines the band gaps of the system, where the waves are evanescent. The strongest exponential decay is found to be when each material of the two is chosen to be one-quarter of wavelength thick ($\pi/2$ phase shift). In the case of a Bragg reflection waveguide it is one-quarter of the transverse wavelength, meaning that this thickness of layer ν with dielectric coefficient ε_ν is given by

$$\Delta_\nu = \pi/2\sqrt{\varepsilon_\nu\omega_0^2/c^2 - k_z^2} = \pi/2k_{\perp\nu}. \tag{52}$$

The eigenvalues in this optimal case are determined by the relative values of the two transverse impedances in the two materials. In the optimal confinement case, each one of the electromagnetic field components either peaks or vanishes at the interface between any two dielectrics. Since the amplitudes in the first dielectric layer are already known, the interface between the first and the second dielectric layers (at $x = D_{\mathrm{int}} + \Delta_1$) may be considered as an entrance to a periodic structure, with the boundary conditions

$$E_z(x = D_{\mathrm{int}} + \Delta_1) = 0, \quad \text{if } Z_1 > Z_2 \text{ (i.e., } n_1 < n_2) \tag{53}$$

or

$$\frac{\partial E_z}{\partial x}(x = D_{\mathrm{int}} + \Delta_1) = 0 \quad \text{if } Z_1 < Z_2 \text{ (i.e., } n_1 > n_2), \tag{54}$$

where Δ_1 is the first layer width, and Z_1, Z_2 are the transverse impedances of the first and second layers, respectively (Mizrahi and Schachter, 2004a). Setting the first layer width according to the above condition will ensure that the required mode at the given wavelength will indeed be supported by the waveguide. The first layer may therefore be perceived as a matching layer between the core vacuum region and the subsequent periodic structure.

Given the amplitudes in Eq. (51) for a TM mode of arbitrary phase velocity, the corresponding first layer widths for which E_z or its derivative vanishes are

$$\Delta_1^{\mathrm{TM}} = k_1^{-1}\arctan\left(\frac{\varepsilon_1 k_x}{k_1}\cot(k_x D_{\mathrm{int}})\right) \quad \text{if } Z_1 > Z_2, \tag{55}$$

$$\Delta_1^{\mathrm{TM}} = k_1^{-1}\arctan\left(\frac{-k_1}{\varepsilon_1 k_x}\tan(k_x D_{\mathrm{int}})\right) \quad \text{if } Z_1 < Z_2. \tag{56}$$

In the above expression, the smallest positive value of the arctan function is chosen. A special case of this expression is $k_x D_{\mathrm{int}} = \pi$ for $Z_1 > Z_2$ or $\pi/2$ for $Z_1 < Z_2$. Then the matching layer is also transverse quarter-wavelength thick like the outer layers. For the case where the phase velocity equals the speed of light ($k_z = \omega_0/c$ and hence $k_x = 0$ in the core), the expression for the first layer width reads

$$\Delta_1^{\mathrm{TM}} = k_1^{-1}\arctan\left(\frac{Z_0 c}{Z_1\omega_0 D_{\mathrm{int}}}\right) \quad \text{if } Z_1 > Z_2, \tag{57}$$

$$\Delta_1^{\mathrm{TM}} = k_1^{-1}\arctan\left(\frac{-Z_1\omega_0 D_{\mathrm{int}}}{Z_0 c}\right) \quad \text{if } Z_1 < Z_2. \tag{58}$$

Figure 16 illustrates in the lower curve the first layer width as a function of the core half-width D_{int} for a planar TM mode with $v_{\mathrm{ph}} = c$. The first layer was set to have a refractive index of $n_1 = \sqrt{\varepsilon_1} = 1.6$, and the other material was taken to be of refractive index $n_2 = \sqrt{\varepsilon_2} = 4.6$. The first layer width is normalized by $\Delta_q = \lambda_0/4\sqrt{\varepsilon_1 - 1}$, which is the transverse quarter-wavelength width in the case $v_{\mathrm{ph}} = c$. The choice of placing the lower refractive index first means that Eq. (57) is used in the calculation of the planar TM first layer width. For reference the TE mode is also illustrated in both cases.

To summarize, the design procedure is as follows: According to the required k_z, a planar Bragg mirror is designed so that all layers are a transverse quarter-wavelength thick. In order to match the mirror to the desired core field, the

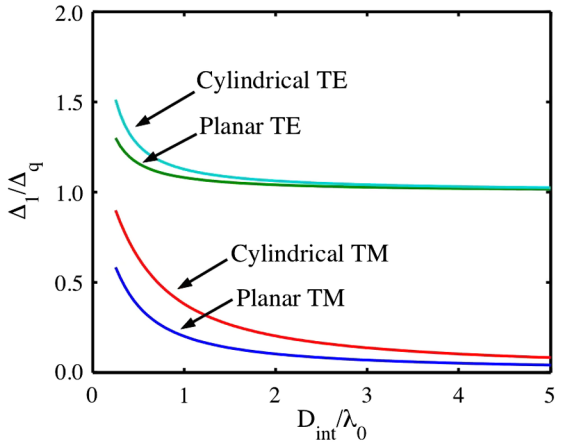

FIG. 16 (color online). First layer width for Bragg TM and TE modes with speed-of-light phase velocity, normalized by $\Delta_q = \lambda_0/4\sqrt{\varepsilon_1 - 1}$. The layer adjacent to the core has a refractive index of 1.6 and the other material has an index of 4.6. From Mizrahi and Schachter, 2004a.

layer adjacent to the core is adjusted to the width given by the analytic expressions above. A similar approach can be employed for cylindrical structures; see Mizrahi and Schachter (2004a) for details.

2. Field distribution

With the same given set of two dielectric materials and a given core dimension, the above design procedure makes it possible to achieve different phase velocities and, correspondingly, different field distributions across the core. As a demonstration of the ability to control the field behavior in the core, the symmetric planar TM mode is considered here. Figure 17 presents different configurations, where in all cases the core half-width is $D_{\text{int}} = \lambda_0$, and the two materials used have refractive indices of 1.6 and 4.6. In all cases, E_z is marked by a solid line whereas H_y is marked by a dashed line. The low index dielectric layers are depicted in light gray, and the higher refractive index layers are indicated by the dark gray. The central core (white) is vacuum.

For the special case of high-energy particle accelerators, we consider a structure which supports a mode with a phase velocity equal to c. The transverse impedance for the first layer takes the form $Z_1 = Z_0\sqrt{\varepsilon_1 - 1}/\varepsilon_1$ when $k_z = \omega_0/c$. For the materials chosen here, this entails that the higher refractive index material has the lower transverse impedance and vice versa. We locate the material with the lower refractive index adjacent to the hollow core as the matching layer. Setting the matching layer width according to Eq. (57), the field profile depicted in Fig. 17(a), in which E_z is uniform

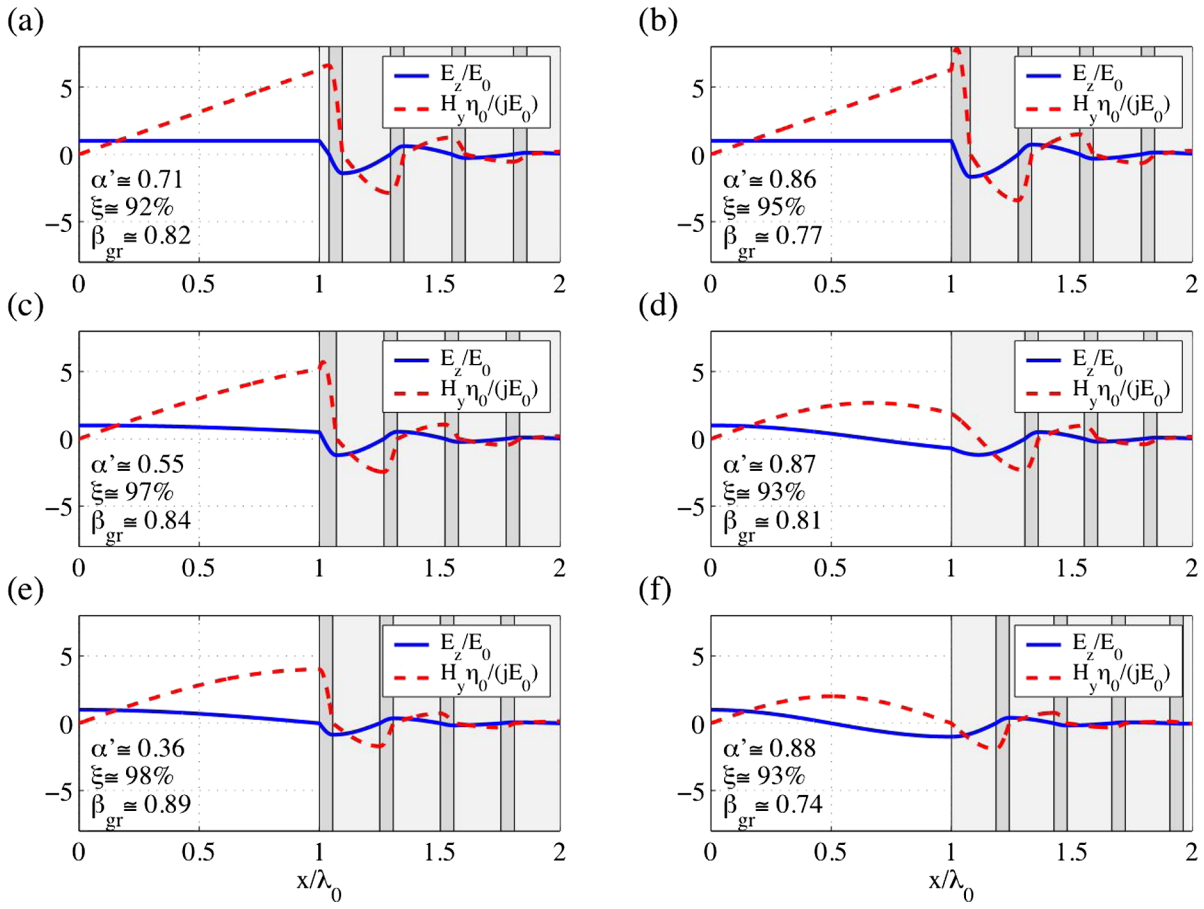

FIG. 17 (color online). Planar TM profiles for vacuum core Bragg structures. (a) $k_x D_{\text{int}} = 0$ and low refractive index first. (b) $k_x D_{\text{int}} = 0$ and high refractive index first. (c) $k_x D_{\text{int}} = \pi/3$. (d) $k_x D_{\text{int}} = 3\pi/4$. (e) $k_x D_{\text{int}} = \pi/2$ (metallic-like walls). (f) $k_x D_{\text{int}} = \pi$ (magnetic-like walls). The normalized decay parameter $\alpha' = \alpha\lambda_0/\tan\delta$, and ξ is the ratio of the power flowing in the core to the total power. From Mizrahi and Schachter, 2004a.

across the core, is obtained. As indicated by Eq. (53) the longitudinal electric field vanishes and the transverse magnetic field peaks at the interface between the first and the second dielectric layers, identically to the case of a metallic wall located at that interface. Maintaining the same field distribution in the core itself, Fig. 17(b) illustrates the case where the material with the higher refractive index borders the core. The second case of Eq. (58) is used, and the picture obtained is as if a perfect magnetic wall is placed at the interface between the first and the second layers.

As examples of arbitrary field profiles that can be achieved setting the matching layer width according to Eqs. (55) and (56) with the phase velocity not equal to c, we present Figs. 17(c) and 17(d), where the transverse wave numbers were chosen to be $k_x D_{\mathrm{int}} = \pi/3$ and $3\pi/4$, respectively. Finally, as a special case of Eqs. (55) and (56), the field distributions when all the layers are transverse quarter-wave thick are shown. Figure 17(e) illustrates the case $k_x D_{\mathrm{int}} = \pi/2$, in which the field in the core behaves as if the core boundary is a metallic wall. Figure 17(f) illustrates the case $k_x D_{\mathrm{int}} = \pi$, which has an identical field to a perfect magnetic wall at the core boundary. Common to all cases presented is that at every interface between any two dielectrics, each one of the fields either peaks or vanishes.

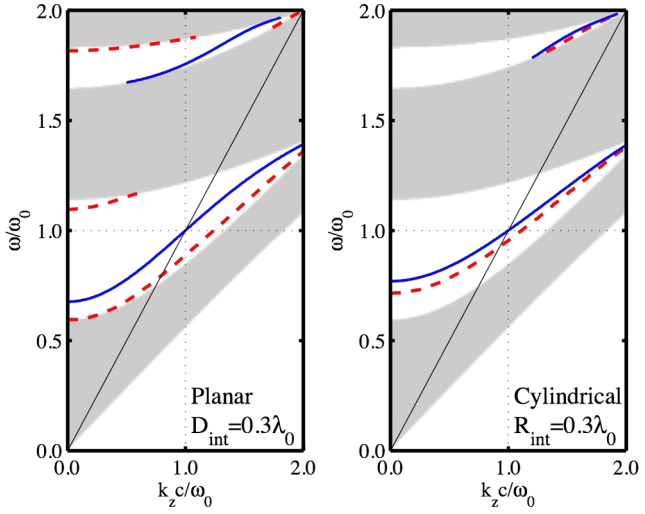

FIG. 18 (color online). Symmetric TM mode dispersion diagram. Left frame: Planar waveguide with $D_{\mathrm{int}} = 0.3\lambda_0$. Right frame: Cylindrical waveguide with $R_{\mathrm{int}} = 0.3\lambda_0$. In both cases the dashed curves are obtained with no design procedure (first layer identical to the remainder of the structure), and the solid curves correspond to a $v_{\mathrm{ph}} = c$ design procedure. From Mizrahi and Schachter, 2004a.

3. Dispersion curves

So far we considered only the electromagnetic field behavior at the specific wavelength λ_0, for which the waveguide was designed. As demonstrated next, adjusting the first layer width may have a significant effect on the dispersion curve. The dispersion points are determined by searching numerically for the zeros of the dispersion function, which has an analytical expression in the planar case. For the cylindrical case, the transfer matrix method is harnessed to determine the reflection coefficient from the outer layers. Taking a relatively large number of layers, the reflection coefficient within the band gap represents that of an infinite number of layers, i.e., its absolute value is unity for all practical purposes.

Our next step is to investigate the symmetric TM mode of both planar and cylindrical Bragg waveguides with $D_{\mathrm{int}} = 0.3\lambda_0$ and $R_{\mathrm{int}} = 0.3\lambda_0$, which are smaller core widths than the example of Sec. III.B.2. For the layer adjacent to the core, the material with the lower refractive index was chosen. In the left frame of Fig. 18, a band diagram is shown, where the allowed transverse propagation areas are indicated in gray, and the dispersion curves of the symmetric TM mode in the planar case are depicted for two configurations. In the first configuration, all the layers have thickness $\lambda_0/4\sqrt{\varepsilon_1 - 1}$, meaning that the Bragg mirror is designed for $v_{\mathrm{ph}} = c$, but without using a matching layer to match between the mirror and the core field. The result is that the dashed dispersion curve does not intersect the point $(\omega/\omega_0 = 1, ck_z/\omega_0 = 1)$ as required. Nevertheless, this dispersion curve intersects the light line at a lower frequency. Operating the waveguide at that frequency is not desirable since the mirror is not optimal, i.e., the transverse exponential decay is weaker than could be achieved. When the first layer is adjusted according to the design procedure described earlier, the solid dispersion curve is obtained. It is seen that changing the first layer thickness shifted the dispersion curve so that there is now an intersection with the point $(\omega/\omega_0 = 1, ck_z/\omega_0 = 1)$. A similar picture is obtained for the cylindrical case shown in the right frame of Fig. 18. More details regarding possible applications of Bragg structures for laser-driven acceleration structures can be found in Mizrahi and Schachter (2004b, 2006).

C. 2D PBG fiber accelerator

A 2D photonic band-gap accelerator has a periodic structure in the two transverse dimensions for confinement and is uniform in the third dimension, being hollow and open ended to accommodate a particle-beam channel (a linear defect embedded in the lattice) and the laser accelerating field. This naturally leads to a round or elliptical particle beam similar to that in a conventional rf waveguide accelerator, although the PBG periods in the two transverse directions can be different. The cylindrical Bragg structure in Sec. III.B is a special case of an axisymmetric 2D PBG structure, making it effectively one dimensional transversely. The 2D PBG geometry typically means that numerical simulations are needed to design the structures, and analytic expressions similar to those of a Bragg accelerator are not generally available. El-Dahshory, Attiya, and Hashish (2007) developed empirical design equations for certain 2D PBG structures by interpolating results from plane-wave expansion calculations, although this has not been applied to accelerators.

For 2D PBG accelerators, the traverse periodicity can be obtained with either dielectric rods suspended in open space or vacuum holes in a dielectric medium such as glass. The 2D PBG confinement of TM modes for particle acceleration was first described by Kroll, Smith, and Schultz (1992) in 11 GHz rf structures composed of sapphire rods supported at the two ends by metallic boundaries. Subsequently, the combined

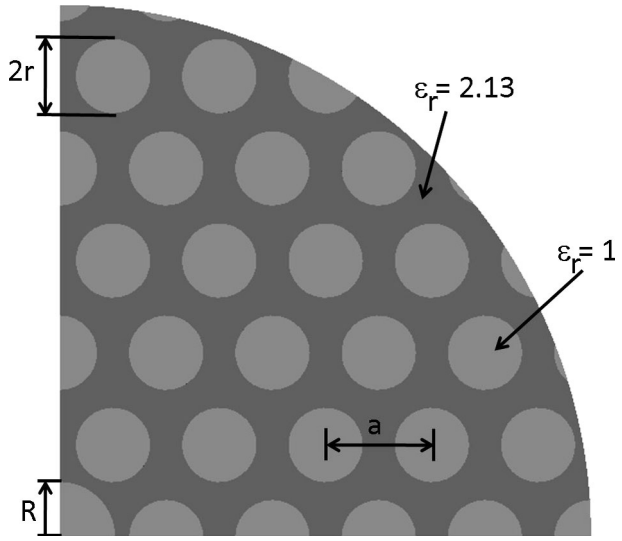

FIG. 19. Quarter cross section of the Lin PBG fiber accelerator, consisting of a lattice of vacuum holes (light gray) in a glass substrate (dark gray) with lattice period a and hole radii $r = 0.35a$. The periodic lattice holes serve to confine the accelerating mode, and the particle beam is accelerated within the central defect, radius $R = 0.52a$, propagating perpendicular to the page. From Ng *et al.*, 2010.

advantages of slow-wave ($v \leq c$) confinement and gigavolt per meter gradients in all-dielectric structures led to the proposal by Lin (2001) to form an accelerator with a single glass fiber permeated by a hexagonal lattice of vacuum holes surrounding a central defect hole for the particle beam, as shown in Fig. 19.

This latter PBG structure is self-supporting, is common for hollow-core telecommunications fibers, and lends itself to the industrial stack-and-draw method based on close-packed, glass tubes heated in a furnace and drawn down to micron-size dimensions. Prototype holey fibers specifically designed for speed-of-light TM modes have been recently fabricated in the industry, and Fig. 20 shows a photograph of the transverse glass structure for an accelerating fiber drawn with a 12 μm central defect (Noble, Spencer, and Kuhlmey, 2011).

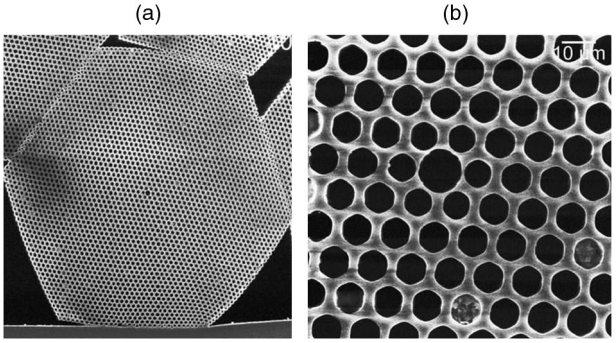

FIG. 20. Photographs showing (a) 700-μm-wide cross section of a hollow-core PBG fiber drawn from borosilicate glass, and (b) magnified view of the 12-μm central defect hole, where the air holes show up as black against the light gray of the glass matrix.

Although this review focuses on dielectric structures at optical frequencies, there is a large body of theoretical and experimental work on metallic PBG structures for acceleration at GHz frequencies, for example, as described by Smith *et al.* (1994), McCalmont *et al.* (1996), Shapiro *et al.* (2001), Smirnova (2005), Smirnova *et al.* (2005), Marsh, Shapiro, and Temkin (2007), and Munroe *et al.* (2013). Much of this research involved test structures at 17 GHz, and progress was made on accelerators designed to both allow wakefields to diffract away, thus avoiding deleterious beam effects, and with reduced surface magnetic fields to limit pulse heating. The metallic PBG structures have achieved gradients comparable to standard disk-loaded copper cavities at similar frequencies (>100 MV/m).

1. Structure design

The 2D PBG structure has frequency band gaps (ω vs k) that are two dimensional in wave-number space. Being homogeneous in the longitudinal direction there are no interference effects to create a band gap determined by k_z alone. The structure naturally supports so-called out-of-plane propagating modes (Joannopoulos *et al.*, 2008), unimpeded in the longitudinal direction. This is unlike a 3D PBG structure (Sec. III.D) in which the modes are forbidden within the frequency gap for all choices of wave vectors ($k_\perp$, k_z).

On the projected band-gap diagram, ω vs k_z, of a 2D PBG lattice, there can be choices of the longitudinal wave number k_z for which modes of a given $k_\perp$ may or may not be reflected by the crystal lattice. Gaps can open and close on this diagram as k_z is varied, not due to k_z directly, but from the relative effects of $k_\perp$ and k_z on the dispersion curves $\omega(k_\perp, k_z)$. Usually the bands cover a wide range of frequencies as $k_\perp$ is varied, and no gaps form when $k_z \ll k_\perp$. But for larger k_z the frequency ω is mostly determined by k_z, and the dispersion curves then flatten as functions of $k_\perp$. Band gaps between these flattened bands may open. This is why 2D glass fibers of low dielectric contrast (ratio of highest to lowest permittivity) usually have band gaps for modes with larger k_z but no gaps for small propagation constants. Figure 21 shows the projected band-gap diagram corresponding to the Lin fiber of Fig. 19 with lattice period a, hole radii $r = 0.35a$, and refractive index $n_r = 1.45$ ($\epsilon_r = 2.13$) for silica glass. Only for $k_z a > 7$ do band gaps open for this structure. If a glass of higher refractive index were used, band gaps will open for smaller values of $k_z a$.

For a simple round-hole, hexagonal lattice the four constants that determine the PBG structure properties are the transverse hole spacing or lattice period a, which sets the wavelength scale of the system, the ratio of hole radius to lattice spacing r/a, the ratio of central defect radius to lattice spacing R/a, and the relative permittivity (dielectric coefficient) ϵ_r of the glass matrix. Because of the high degree of symmetry, hexagonal lattices exhibit the widest band gaps (good mode separation) compared to other regular lattices, and their natural close packing makes them the simplest array to manufacture with glass tubes. The band diagram in dimensionless units of ka is determined by the ratio r/a and the permittivity (analogous to adjusting the 1D Bragg layers to be quarter-wave reflectors). Generally as the relative amount of

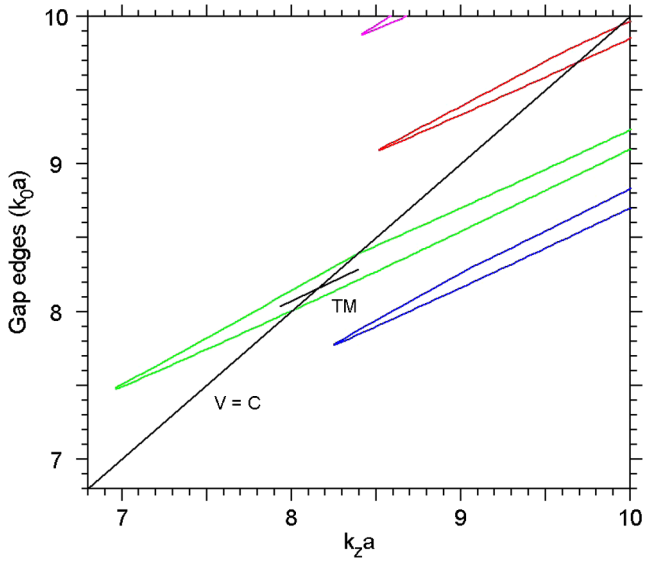

FIG. 21 (color online). The lowest band gaps near the light line ($v = c$) for the Lin fiber plotted on the frequency vs wave-number dispersion plane as calculated with R-Soft BANDSOLVE photonics software (R-Soft, 2011). When a defect hole of the correct size is introduced, a TM mode with dispersion line crossing the light line appears in the band gap. From Noble, Spencer, and Kuhlmey, 2011.

dielectric increases (smaller air holes) or the permittivity increases (higher ϵ_r), the bands shift lower in frequency with the band diagram consisting of several well-separated gaps at relatively smaller values of $k_0a = \omega a/c$. The group velocity of modes also decreases with smaller holes and higher permittivity. The Lin fiber in Fig. 19 has about one-half the matrix volume being glass. The particular hole size $r/a = 0.35$ is chosen so that a band gap crosses the light line at about the point where the gap is the widest, which generally improves trapped mode confinement.

2. Accelerating mode properties

The defect hole radius R is specifically chosen so that a TM-like defect mode crosses the light line near the center of the band gap, insuring good confinement and mode separation. The physical effect of adjusting the defect radius to obtain a speed-of-light, TM mode in the hexagonal PBG accelerator is analogous to the procedure of adjusting the matching layer in a 1D Bragg accelerator (Sec. III.B). Because of the complicated web of glass and holes immediately surrounding the defect, no analytic expressions are known, and numerical simulations are needed. When the defect size is adjusted to $R/a = 0.52$, the accelerating mode shown in Fig. 22 resides in the band gap crossing the light line near $k_0a \simeq 8.2$, corresponding to $\lambda = 0.77a$. For example, the choice of lattice spacing $a = 1.3$ μm yields an accelerating mode with a wavelength $\lambda = 1$ μm. The longitudinal field is radially uniform within the defect, a direct result of the SOL condition ($v_p/c = k_0/k_z = 1$), and both E_r and H_θ are proportional to r. Within the central defect region, the fields are to first order the same as listed in Table III for the cylindrical Bragg structure (Sec. III.B).

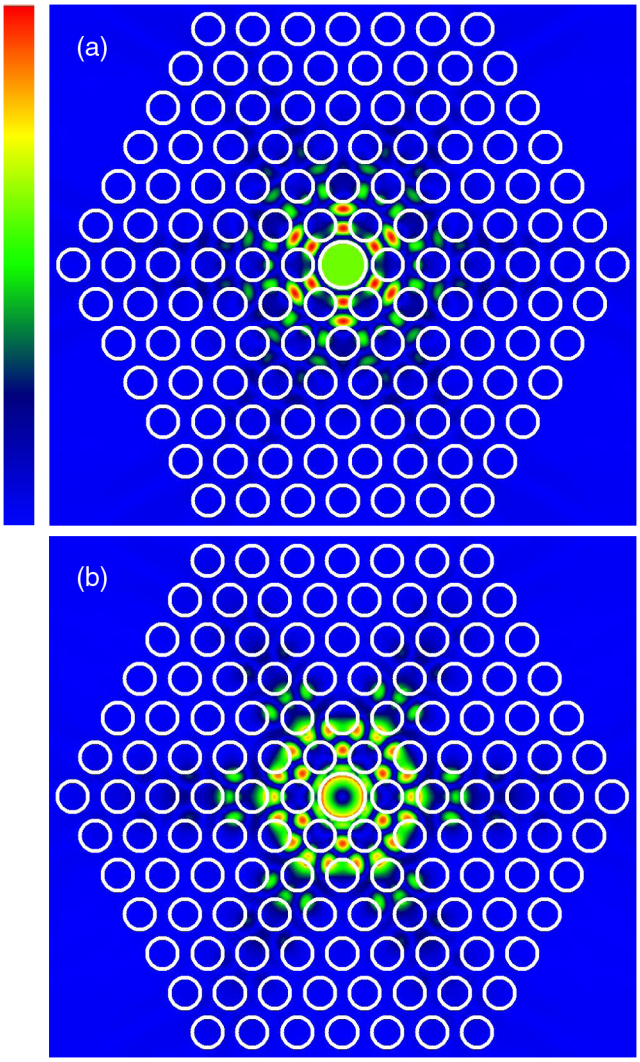

FIG. 22 (color online). (a) Longitudinal and (b) radial electric field intensities of the defect mode crossing the light line in the PBG fiber, as calculated with a multipole solver (Kuhlmey *et al.*, 2002). The white circles indicate the hole boundaries. A rainbow color scheme (the legend) is used to represent linear magnitude of the relative field intensity with the different colors being the maximum (1) and minimum (0). From Noble, Spencer, and Kuhlmey, 2011.

This mode is reasonably efficient for acceleration with a characteristic impedance $Z_c = E_{\rm acc}^2\lambda^2/P = 19.5\ \Omega$, a ratio of accelerating gradient to maximum field of $E_{\rm acc}/E_{\rm max} \simeq 0.5$, and a group velocity of $0.58c$. The defect size $R/a = 0.52$ is the smallest value for which a TM mode appears in this gap for the lattice choice $r/a = 0.35$. A larger defect would be better for beam transport, but as the defect size increases the TM mode shifts to higher frequency and out of the gap. New TM modes may enter the band gap from below, but they will typically have lower characteristic impedance and gradient. The impedance Z_c scales approximately like $(kR)^{-4}$ since the power P in the denominator of Z_c is proportional to the integral of $\int E \times Hrdr$. For a metallic pillbox cavity where fields terminate at the boundary, this scaling relation is very good. For a PBG dielectric fiber, the scaling is only approximate since much field energy is outside the defect, and this

leads to an exponent of typically between 3 and 4. Because of this strong dependence, the value of Z_c for a TM mode is mainly determined by kR when the mode is centered in the gap.

D. 3D PBG woodpile accelerator

1. Design and properties

A three-dimensional, dielectric accelerating structure was developed with several key features in mind (Cowan, 2008). First, to achieve high efficiency and take advantage of similarities to conventional structures, an accelerating waveguide is desired. Second, we want to have a structure with materials and geometry amenable to lithographic fabrication, leveraging the immense knowledge of microfabrication tools and techniques developed by the integrated circuit and microelectromechanical systems (MEMS) industries over the last half century (Plummer, Deal, and Griffin, 2000). Third, we like a structure that allows integration of ancillary devices, such as power couplers, beam diagnostics, and possibly even laser power amplifiers.

To meet these objectives, a 3D photonic crystal waveguide structure was chosen. The design process followed the procedure outlined in Sec. II.A for the design of general photonic crystal devices, namely: (1) Find a photonic crystal lattice with a band gap, (2) introduce a defect, (3) find the point along the dispersion curve of the accelerating mode with speed-of-light phase velocity, and (4) adjust the parameters of the waveguide if necessary. The lattice chosen was the so-called "woodpile" geometry, a well-established three-dimensional photonic crystal lattice designed to provide a complete photonic band gap in a structure with a straightforward lithographic fabrication process (Ho *et al.*, 1994). The woodpile lattice exhibits an *omnidirectional* band gap, a range of frequencies in which no mode, of any wave vector or polarization, exists. This feature enables confinement of electromagnetic fields in waveguide networks, even around sharp bends, since scattering through the lattice is prohibited even with sudden changes in wave vector. Thus, the lattice itself provides for integration of ancillary devices, including low-loss couplers as described later.

The lattice consists of layers of dielectric rods in vacuum, with the rods in each layer rotated 90° relative to the layer below and offset half a lattice period from the layer two layers below. We form a waveguide by removing all dielectric material in a region which is rectangular in the transverse dimensions and extends infinitely in the beam propagation direction. In addition, there are two modifications: In order to avoid deflecting fields in the accelerating mode, we make the structure vertically symmetric by inverting the upper half of the lattice so it is a vertical reflection of the lower half. Second, we extend the central bars into the waveguide by $0.124a$, where a is the lattice constant, on each side in order to suppress quadrupole fields. The geometry, with a defect waveguide introduced, is shown in Fig. 23. The inversion of half the lattice introduces a planar defect where the two halves meet, but this waveguide still supports a confined accelerating mode. Indeed, the mode is lossless to within the tolerance of the calculation, placing an upper bound on the

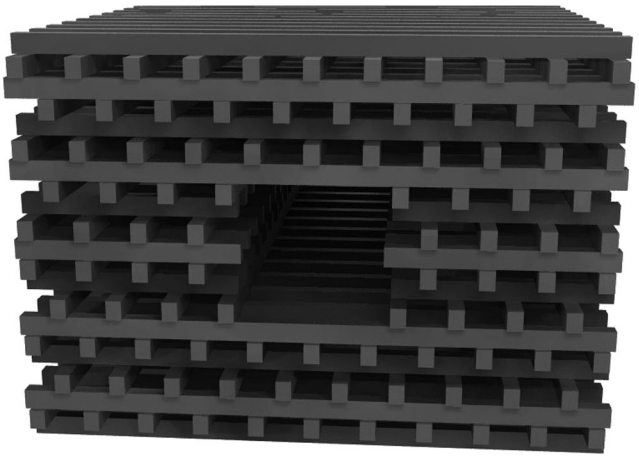

FIG. 23. The geometry of a woodpile accelerating waveguide structure. From Cowan *et al.*, 2010.

loss of 0.48 dB/cm. Its fields are shown in Fig. 24. For this mode $a/\lambda = 0.364$. If we use $\lambda = 1550$ nm, where many promising sources exist (see Sec. V), this determines $a = 565$ nm. The individual rods are then 158 nm wide by 200 nm tall.

This structure was investigated in detail by Cowan (2008). Using some preliminary studies of optical breakdown in silicon, it was estimated that the structure could achieve an unloaded gradient of about 300 MV/m. It has the highest characteristic impedance among the structures discussed here of 484 Ω; it was estimated that by recycling the drive laser pulse in an optical cavity, as discussed in Sec. I.B, the structure could achieve optical-to-beam efficiency in excess of 75%, accelerating 12 fC charge per optical bunch train. The investigation of this structure also introduced a novel optical focusing technique, in which a modified structure is used as an electromagnetic quadrupole by running $\pm\pi/2$ out of phase. It was found that this technique could enable stable transverse

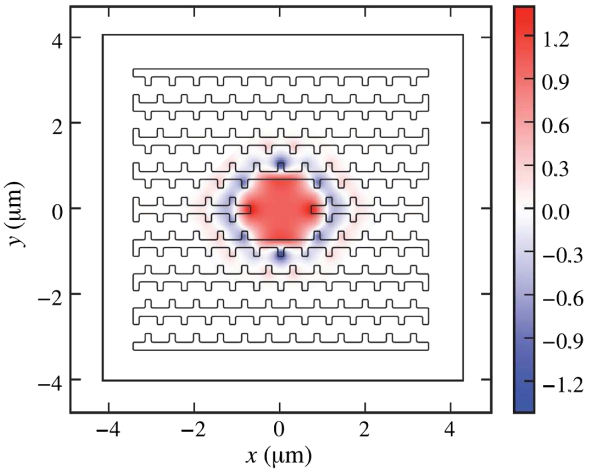

FIG. 24 (color online). The accelerating field seen by a speed-of-light particle, averaged over a lattice period, normalized to the accelerating field on axis, shown with structure contours for a transverse slice at $z = 0$. The inner rectangle denotes the interface between free space and the absorbing boundary in the simulation. From Cowan, 2008.

TABLE IV. Target structure parameters for the final woodpile accelerator structure fabricated at Stanford. The rod width is chosen to maximize the band gap. The layer thickness is chosen to maintain the fcc lattice, such that it is $h = a/2\sqrt{2}$. The wavelength corresponds to the free-space wavelength where the accelerating mode crosses the light line.

	Period (μm)	Rod width (nm)	Layer thickness (nm)	Wavelength (μm)
Scaling relation	a	$0.28a$	$0.354a$	$2.72a$
Final dimensions	1.27	356	449	3.45

motion, with focusing forces orders of magnitude stronger than conventional magnets and exhibiting modest emittance growth. This technique is discussed further in Sec. II.B.

2. Fabrication

a. Layer-by-layer fabrication method

The fabrication process developed for the woodpile structure described here is based upon the method developed by Lin *et al.* (1998). Standard semiconductor fabrication processes are used to pattern each of the layers. The structure is built up layer by layer. This process is appealing due to its simplicity. Standard semiconductor processing tools are used for each of the steps, making it possible to scale up to large production at foundries if desired. The layer-by-layer process impose challenges when trying to fabricate structures with many layers. Stress induced by thin-film deposition was a major limit to the number of layers that could be fabricated on a single wafer. Ultimately, the structures presented here were fabricated in two halves, consisting of eight and nine layer structures, which were then aligned and bonded. This is described in detail by McGuinness (2012).

Several rounds of fabrication were performed with slight variations to the dimensions for each iteration. The parameters for the final structure are shown in Table IV. The process involves depositing layers of silicon dioxide and poly silicon and is illustrated in Fig. 25. The silicon dioxide serves as a matrix into which the desired features for the given layer are patterned and filled with silicon. When all the layers are complete, a selective etch removes the silicon dioxide, yielding a free-standing structure of silicon and air.

b. Wafer stacking and alignment

The design for the woodpile accelerator structure calls for a minimum of 17 layers. This includes a seven layer defect and five cladding layers on top and bottom. Rather than fabricating all 17 layers on a single wafer, the structure was fabricated in two halves. There were many reasons for doing this including time, yield, and layer stress.

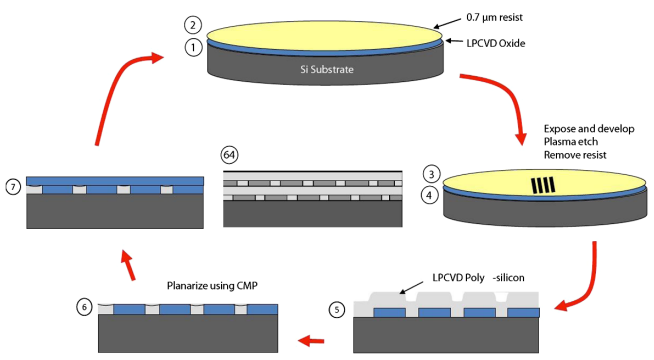

FIG. 25 (color online). Visual outline of the woodpile fabrication process. From McGuinness, 2012.

The alignment strategy used for the final structures was based on visual alignment marks for coarse alignment and a diffraction pattern for fine alignment. A large bonding pad was included in the design and used as a mechanical support to reduce the number of degrees of freedom required in the alignment. As the two half structures were aligned and brought into contact with each other, the bonding pads provided a rigid surface for each half to contact, mechanically determining the separation between the two halves, and tip and tilt of the wafers. A blackbody source was used to back illuminate the samples. A near-infrared (NIR) objective and an NIR camera was used to image the transmitted light. The two half samples were mounted onto translation stages and aligned using the vernier patterns from the front and back samples. With perfect image capabilities it would be possible to achieve alignment accuracy down to 10 nm using this method. In reality the imaging system was limited, and alignment accuracy on the order of 50 nm was expected.

Once the coarse alignment was achieved using the vernier patterns, fine alignment was performed using diffraction from a grid, where the period of the diffraction grid is the same as the lattice period for the woodpile structure. For TE and TM polarized light the zeroth order spot is a minimum for perfect overlap and a maximum for a half period translational offset. The first-order spots are a maximum for perfect overlap and a minimum for one half period offset. Once the fine alignment was complete the wafers were permanently joined using a quick setting epoxy. The alignment was monitored as the epoxy was curing. Once it was cured, the sample was removed from the alignment stage. A cross-section SEM image of the final aligned and bonded structure is shown in Fig. 26. These images are for the front and back of one of the samples. It demonstrates very good angular alignment. Transverse alignment is within approximately one-third of the lattice period or 430 nm.

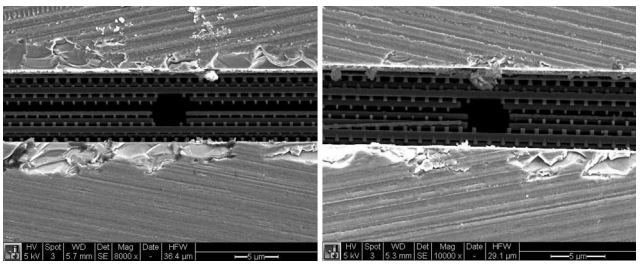

FIG. 26. Front and back images of an aligned and bonded structure. A slight transverse offset can be seen, corresponding to approximately one-third of a lattice period. Angular alignment of this structure was very accurate. From McGuinness, 2012.

Fabrication of 17-layer woodpile accelerator structures was successfully achieved using the layer-by-layer approach combined with an alignment strategy presented here. Fabrication tolerances were within 5–8% of the layer thickness, rod width, layer-to-layer alignment, and taper angle. Final alignment of the two half structures reveals a one-third period offset. Improved accuracy and automation in the alignment system is important for future fabrication runs.

E. Focusing and diagnostic structures

In a conventional rf accelerator, the essential accelerator components can be coarsely grouped into four categories based on their functions: acceleration, deflection, focusing, and diagnostics. In order to build a laser-driven dielectric accelerator beam line, we require each of these components. However, there is a large dimensional mismatch between conventional meter-scale devices and the proposed micron-scale dielectric accelerator elements. Combining these two technologies at the disparate scales is extremely difficult and would likely negate the size reduction and nanometer-class alignment accuracy inherent in the dielectric structures from the lithographic fabrication process. This motivates the design of new dielectric, micron-scale compatible structures, which are capable of beam deflection, focusing, and diagnostics.

One of the great advantages of using lithography techniques to fabricate a particle accelerator is the ability to consecutively stage together several hundred accelerating structures with nanometer-class alignment accuracy. In order to extend this alignment accuracy to the deflecting structures and focusing structures, it is necessary that they have compatible fabrication processes, which is most readily achieved when the accelerating, deflecting, and focusing structure all share a similar geometry.

For the case of the grating accelerator design detailed in Sec. III.A, Plettner and Byer (2008) described one implementation of a deflecting structure that consists of a pair of tilted gratings, oriented at an angle α relative to the electron beam trajectory, as discussed in Sec. VI.B.2 and Fig. 39(a). Through an eigenmode decomposition method, Plettner details the resulting force pattern from the tilted grating geometry, when driven with a laser in the same manner as the accelerator structure. Because of the high levels of geometrical similarity with the accelerating structure, this deflecting structure can be fabricated using the methods described in Sec. III.A.3 with little modification.

Similarly, Soong *et al.* (2012a) described one implementation of a focusing structure, which relies on a modification to the incident drive laser, rather than modifying the accelerating structure geometry. In this implementation, depicted in Fig. 27, the drive laser is focused on a near diffraction limited waist at both faces of the grating structure and subsequently diverges as it fills the vacuum channel. The force pattern from the diverging laser fields establishes a net radial focusing force with field strength on the order of 10^6 T/m.

Although ultrahigh levels of alignment accuracy between various accelerator structures can be expected from fabrication on a single wafer, such alignment accuracies are unlikely from wafer to wafer. Therefore, the ability to measure beam position to nanometer-scale accuracies will be essential in order to correct for wafer-to-wafer misalignments. Traditional beam

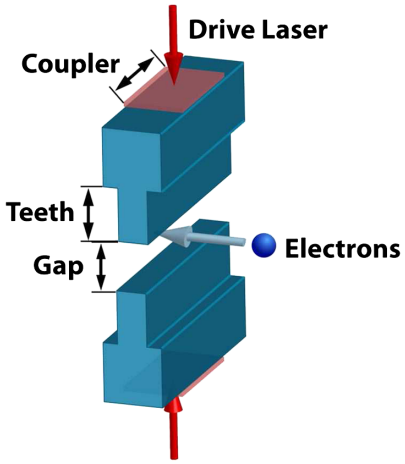

FIG. 27 (color online). A single period of the focusing structure design, with optimizable parameters labeled. From Soong *et al.*, 2012a.

position monitors (BPMs) are unsuitable for this purpose. As with the deflecting and focusing components, traditional BPMs are far too large to be incorporated with the dielectric laser-driven accelerator structures, and attempting individual alignment of traditional BPMs with dielectric structures is impractical. Furthermore, traditional BPMs have measurement precision on the order of microns, which is insufficient for the required nanometer-scale alignment accuracy.

Soong and Byer (2012) and Soong *et al.* (2014) described the theory, fabrication, and experimental demonstration of a novel BPM which is particularly well suited for the application of dielectric laser-driven accelerators. The monitor operates on the principle of an inverse grating accelerator, and has the unique ability to map a particle-beam position to a measurable wavelength. In essence, the BPM consists of a pair of opposing fan gratings, whose grating period varies linearly in one direction. The direction of the variation is chosen to correspond to either horizontal or vertical offsets of the particle beam. When a particle beam traverses this BPM structure, it generates radiation at a wavelength corresponding to the grating period at the beam location, as shown in Fig. 28. Prototypes of these structures have been fabricated at Stanford

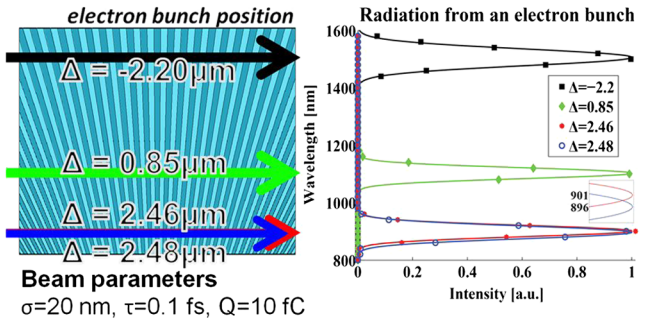

FIG. 28 (color online). Four independent electron bunches (left) traversing the grating BPM structure at four unique positions along the BPM, as viewed from the top, and the resulting radiation spectrum (right) from each of the four electron bunches. The inset demonstrates the resolvability of the two closely separated spectral distributions. From Soong and Byer, 2012.

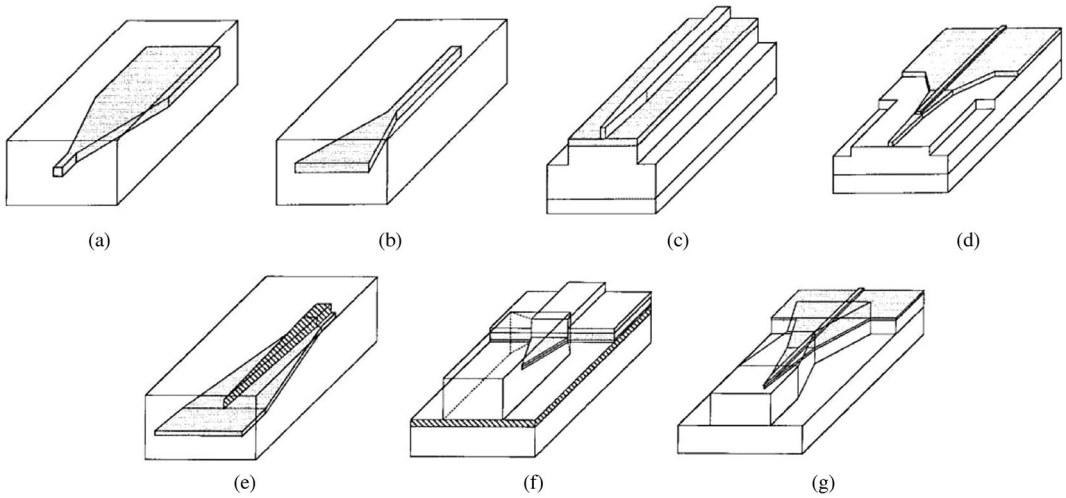

FIG. 29. Schematics of lateral taper designs. From Moerman, VanDaele, and Demeester, 1997.

University with a continuous wavelength range from 800 to 1050 nm, and proof-of-principle demonstrations carried out at SLAC using a conventional electron linac (Soong *et al.*, 2014). Combined with a DLA-generated submicron-sized electron beam and a typical 0.1 nm resolution spectrograph, it is estimated that this technique could be employed to resolve beam position in a DLA with 0.75 nm precision.

F. Input power coupling to structures

1. External laser-to-structure matching

To accelerate electrons in a DLA device, laser pulses must be illuminated into the structure with appropriate phase matching conditions. Such illumination can be accomplished by free-space propagation or guided wave propagation. In free-space coupling, laser pulses are propagated to the accelerators using free-space optics, including collimating lenses and focusing lenses. Phase matching can then be fulfilled by a pulse-front tilt technique, in which a 45° pulse-front tilt is suitable for a speed-of-light electron (Bor and Racz, 1985; Hebling, 1996; Plettner, Lu, and Byer, 2006). If materials with large refractive indices are needed for accelerators, the reflection at the vacuum and material interface will decrease the input laser power. An antireflection coating is required to couple most laser power into the accelerators.

Laser pulses can also be delivered to the accelerators through on-chip optical waveguides. One advantage of integrated waveguides is that the laser power can be distributed through waveguide network and waveguide splitters. The phase of the laser pulses arriving at the accelerators can therefore be precisely controlled by the design of the waveguide network (e.g., by having different waveguide length or by adding phase controlled devices). To deliver maximum power to the accelerators through the waveguide network, it is essential to first couple laser pulses efficiently to the waveguides. For the 2D and 3D waveguide-based accelerator cases, another coupler from the waveguide network to the accelerator structure is required to deliver power to the accelerating mode effectively. This accelerating mode coupler needs to be compact. Because of the group velocity slippage effect discussed in Sec. I.B, coupling of a new laser pulse into the waveguide must occur on a length scale shorter than the distance over which the laser pulse slips off the particle bunch. For a 1 ps long laser pulse and accelerating wavelength at 1.55 μm, this length is about 100 μm (Cowan *et al.*, 2010).

Many techniques have been developed for external laser to waveguide coupling (Hunsperger, 2009), mainly falling into four categories: transverse coupling, grating coupling, prism coupling, and vertical tapering (Liu and Prather, 2004). A focusing reflector based method has been proposed and experimentally verified as well (Dillon *et al.*, 2008). The most straightforward transverse method is an "end-fire" coupler that directly launches the optical field at the input facet of the waveguides. Based on the Poynting theorem and overlap integral, the coupling efficiency of the end-fire method depends on mode matching, alignment, and facet quality. Theoretically, highly efficient coupling (> 90%) can be achieved by matching the modes and eliminating the alignment errors. Practically, the waveguide mode does not completely match the incoming laser through a lensed fiber or an objective lens, and alignment errors occur due to tilt and transverse and longitudinal offsets. These effects may increase the coupling loss.

Another commonly used coupling method is the tapered coupler, in which the waveguide gradually changes in transverse size so that the waveguide mode can match the input laser mode, as shown in Fig. 29 (Moerman, VanDaele, and Demeester, 1997). An up-taper coupler (i.e., the taper size is larger than the waveguide size) is widely used because in most cases the waveguide geometry and mode size is smaller than the input laser mode. Recently, the inverted taper has received much attention because it can provide efficient coupling with small feature size. In the inverted taper, the taper size is smaller than the standard waveguide size so that the optical mode is weakly confined, resulting in a large mode size that can match the laser mode. Coupling efficiency of about 90% has been demonstrated for fiber to silicon waveguide coupling (McNab, Moll, and Vlasov, 2003).

Grating couplers offer some advantages, including the fact that the coupling is achieved from the top surface of the

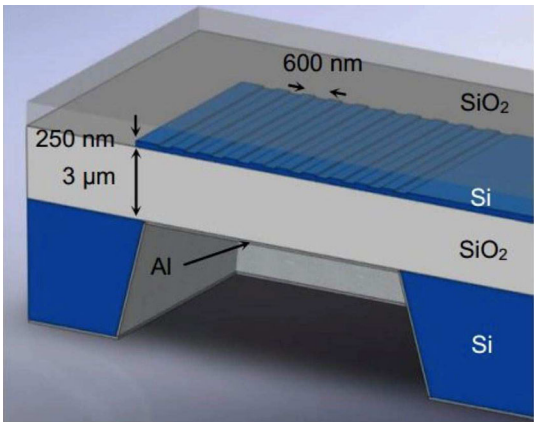

FIG. 30 (color online). Schematics of the grating coupler with bottom metal mirror. From Sfar-Zaoui *et al.*, 2012.

waveguide. Gratings can provide phase matching between the incoming optical wave and the guided waveguide mode. The limitations on alignment error and facet quality are mitigated. The coupling efficiency is wavelength dependent and is limited by the losses due to transmission, reflection, and scattering at the grating. With a bottom metal mirror (Fig. 30), experimental coupling efficiency of about 70% was demonstrated on the silicon-on-insulator platform (Sfar-Zaoui *et al.*, 2012). With subwavelength microstructures the mode field overlap could reach 94% with minimal backreflection, and 43% coupling has been experimentally achieved (Halir *et al.*, 2010).

Prism-based coupling takes advantage of the leaked optical wave through the tunnel layer, into the guided mode of the thin-film waveguide, as discussed by McNab, Moll, and Vlasov (1970) and Liu and Prather (2004). Because of its parallel coupling nature, polishing of the waveguide input face, transverse profile matching, and precision alignment become unnecessary. The silicon-based prism has the potential to miniaturize the structure to millimeter scale for integrated optical circuits. Coupling efficiency of 77% is theoretically predicted, and 46% has been experimentally achieved (Liu and Prather, 2004).

For the 2D and 3D waveguide-based accelerators, the fundamental accelerating modes are usually TM like with local Poynting flux maxima outside the waveguide aperture (Lin, 2001; Cowan, 2008), which is essentially different from the incoming Gaussian-like dielectric waveguide modes. Based on the mode profile matching, the end-coupling efficiency between these modes would be poor. An end-fire coupler would also potentially obstruct the particle beam unless it is open on the axis. Therefore a side-coupling scheme, with laser power directed to the local Poynting maxima, is preferred for the accelerating mode coupler (Schachter, 2011). Recent simulations following such a scheme, e.g., a 2D PBG fiber accelerator with silicon guide or defect lattice guide as a side coupler (England *et al.*, 2011), and a 3D woodpile accelerator with T-junction coupler made out of a transverse woodpile waveguide (Wu *et al.*, 2011; Cowan *et al.*, 2010), have shown power coupled into the fundamental accelerating modes, but how to obtain the highest efficiency is still under study.

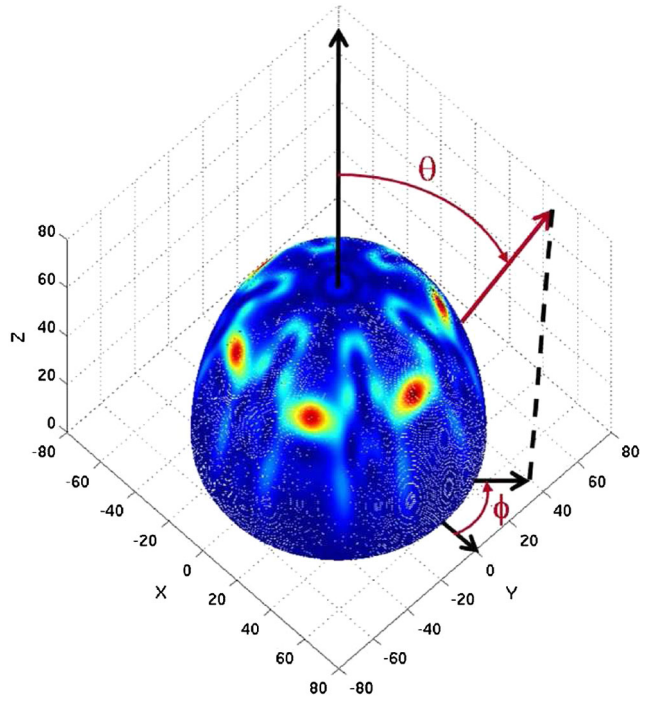

FIG. 31 (color online). Poynting flux distribution from the accelerating mode radiating from the end of a PBG fiber accelerator (located at origin $x = y = z = 0$) on a hemispherical surface of radius 80λ. From Ng *et al.*, 2010.

To circumvent the difficulties of coupling power into a 2D PBG fiber accelerator using optical waveguides, the coupling of laser power can be directed from free space onto the fiber end face. However, the low efficiency of this end-coupling technique makes it likely useful only for proof-of-principle experiments. The mechanism of achieving free-space coupling has been investigated through an inverse process by studying the far-field radiation pattern from the propagation of the accelerating mode and its radiation from the end of the PBG fiber (Ng *et al.*, 2010). The far-field radiation pattern focuses in six well-defined spots (Fig. 31), which respect the hexagonal symmetry of the lattice of the Lin fiber (Lin, 2001). Matching of external laser light into the fiber in an experimental setup might therefore be accomplished using either an array of six linearly polarized laser beams, whose polarizations are rotated relative to each other in increments of 60°, or by nonparaxial focusing of a single laser mode with the requisite radial polarization and hexagonal symmetry. Modification of this scheme by cleaving the end of the PBG fiber at an angle has shown that the far-field radiation changes to a simpler pattern. By varying the angle gradually, the far-field radiation pattern evolves from its original sixfold symmetry to two well-defined spots in the upper hemisphere when the angle is 45°. This has the potential of simplifying the configuration of laser beams needed to excite the accelerating mode via end coupling.

Practical fabrication difficulties limit the layout of the side-coupled guide to the accelerating guide, e.g., inserting a silicon guide into the PBG fiber remains a challenge (England *et al.*, 2011). On the other hand, for the 3D PBG accelerator case, silicon guides may be naturally integrated to the silicon

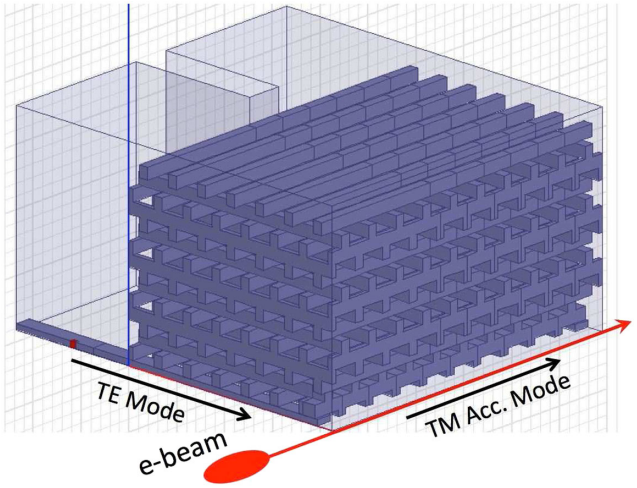

FIG. 32 (color online). Simulation geometry for high-efficiency transverse power coupling to the woodpile structure from a TE mode silicon guide to a TM mode accelerating channel where the electron beam (ellipse) travels.

woodpile structure by photolithography. However, similar to the conventional rf cavity-disk loaded (RFQ) accelerator structure, a woodpile defect waveguide exhibits longitudinal periodicity. Because of this translational variant nature, a cross-sectional matched incident mode profile will be insufficient to launch the propagating mode in the woodpile waveguide properly, resulting in a significant standing wave. Standing waves would tighten tolerance to fabrication errors and make the structure susceptible to high-power breakdown.

A traveling-wave launch method based on scattering matrices was previously developed for RFQ accelerator structures (Nantista, Tantawi, and Dolgashev, 2004), and the method can be applied to the woodpile waveguide. A mode launcher made of a two-dimensional waveguide, such as a silicon guide in the woodpile accelerator case, with a perturbation to form a coupling iris, can excite all the space harmonics needed to compose the propagating mode. A one-quarter cutaway cross section of such a scenario is illustrated in Fig. 32. This figure shows transverse coupling from a silicon light guide of TE operation to a TM mode accelerating channel in the woodpile structure. An embedded piece of perfect conductor, or other material, inside the silicon guide (rectangle on the left of Fig. 32) serves as a coupling iris to optimize the traveling wave match to the TE guide, which has been optimized in simulation to better than 95% coupling efficiency (Wu *et al.*, 2014).

2. Optical waveguides

The integrated waveguide approach has received increasing attention because it offers several advantages over other approaches. Compared to the free-space approach that requires bulky optics, the waveguide approach can uniformly distribute the laser pulses to different accelerator sections by using the waveguide splitters and waveguide network. A schematic of the power distribution network is shown in Fig. 33, in which the 1:8 split waveguide network distributes the pulses to different sections of the accelerator. The pulse arrival time at different accelerator sections can be precisely controlled by the optical delay lines. Another advantage of this approach is that the fabrication of the waveguides can leverage well-developed processes. For example, the size and the geometry of the waveguide devices are defined by lithography and etching. Therefore, the optimal device design can be implemented by minimizing the process variations.

One challenge of the waveguide network design is the power handling capability. Since the accelerator is operated close to the damage threshold of the materials, for a 1:8 split waveguide network, the input bus waveguide should be able to handle power that is 8 times higher, a requirement that is challenging to meet due to material damage. In addition to the power handling capability, the mode profile and mode size of the waveguides are important factors for delivering pulses to the accelerator. The choice of mode sizes depends on several considerations, including accelerating gradient, single mode conditions, power density, and nonlinear properties of the waveguide. In the damage threshold limited case, large mode size is preferred.

Many waveguide techniques have been developed for integrated optics. Some of these techniques have the potential to be implemented for the power distribution network to the accelerator. Of particular interest are silica waveguides, silicon waveguides, and antiresonant reflecting optical waveguides (ARROWs). Silica waveguides have been extensively studied for many years (Li and Henry, 1996; Miya, 2000). Because of low-index contrast (refractive index ratio), the mode size of a silica waveguide can be designed to be $8 \times 8\ \mu m^2$, which is similar to the mode size of a single mode fiber. Therefore, the propagation loss and fiber coupling loss can be lower than 0.01 dB/cm and 0.1 dB, respectively (Miya, 2000). Another advantage of the silica waveguide is the power handling capability since silicon dioxide has a much higher damage threshold than other materials. One possible issue of the silica waveguide is the large bending radius that may increase the size of the power splitter. The silica waveguides can be fabricated on the silica-on-silicon platform, which makes

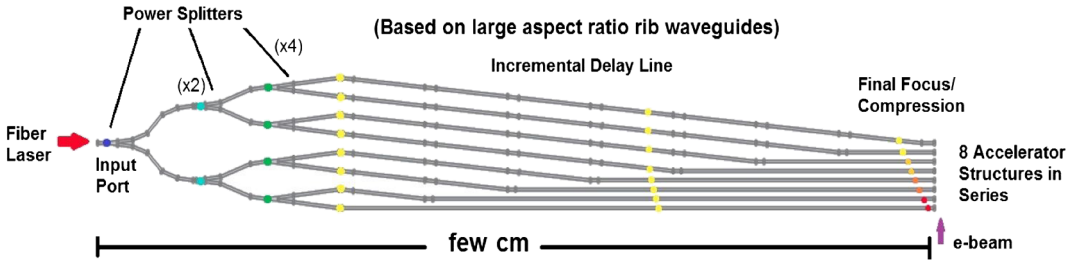

FIG. 33 (color online). Schematics of the waveguide network with waveguide splitters and optical delay lines. From Peralta *et al.*, 2012.

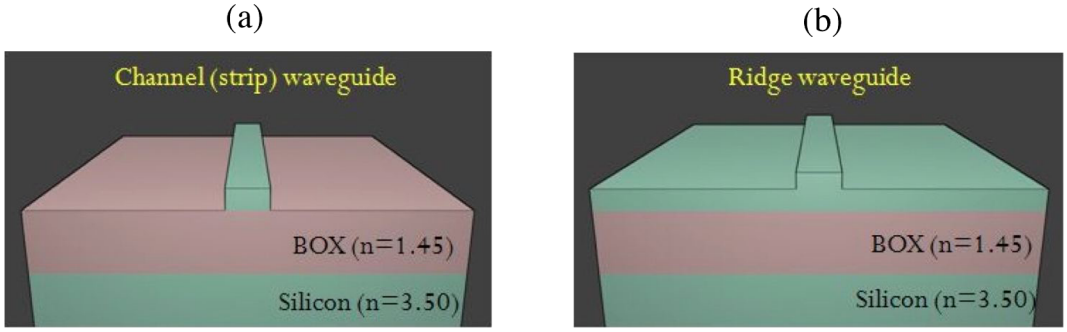

FIG. 34 (color online). Schematics of (a) conventional SOI waveguides: channel and ridge waveguides; (b) monolithic silicon waveguide in standard silicon.

the integration of the waveguide network and accelerator possible.

The possibility of integrating CMOS electronics and photonics has generated wide interest in silicon photonics. Unlike silica waveguides that have low-index contrast, silicon waveguides provide high-index contrast between the core and the cladding, resulting in strong light confinement, compact device size, and small bending radius. As shown in Fig. 34(a), these waveguides are usually implemented on a silicon-on-insulator (SOI) platform by using CMOS and MEMS compatible processes. They can also be fabricated in standard silicon [see Fig. 34(b)] (Chang and Solgaard, 2013a, 2013b, 2013c), facilitating integration with the accelerator.

Although the mode size of the silicon waveguide is generally small, a large mode size is achievable in silicon ridge waveguides, in which the higher-order modes are leaky, resulting in a quasisingle mode with the mode size of several microns (Schmidtchen *et al.*, 1991; Soref, Schmidtchen, and Petermann, 1991; Rickman, Reed, and Namavar, 1994). However, the power handling capacity of the silicon waveguide is much worse compared to the silica waveguide. Therefore, the silicon waveguides may be suitable only for the power limited case.

In addition to conventional waveguides by total internal reflection, ARROWs can also be used as the basis for a power distribution network (see Fig. 35) (Zmudzinski *et al.*, 1995). The ARROWs contain alternating regions with different materials so that the input mode can be laterally distributed to other regions based on the resonance condition (Duguay *et al.*, 1986). Since the pulses are guided in low-index material, the mode size can be several microns. The lateral coupling to adjacent regions makes the power distribution uniform (Zmudzinski *et al.*, 1995). The ARROWs can be fabricated on a silicon wafer with buried SiGe layer and ridge structures.

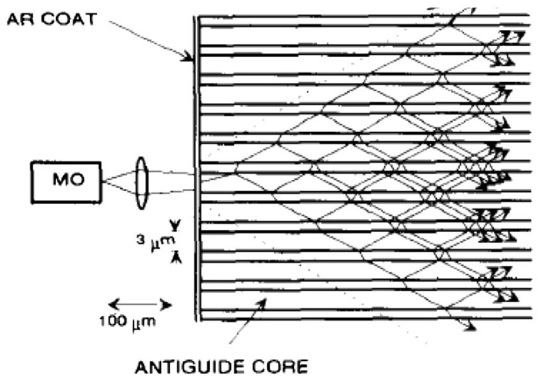

FIG. 35. Schematics of the power distribution network using ARROW structure. From Zmudzinski *et al.*, 1995.

G. Nonlinear dielectric effects at high fields

When laser intensity is high enough in a dielectric, a number of nonlinear optical effects can take place that affect the particle acceleration in a DLA. The second-order nonlinearity exists only in a material structure without inversion symmetry. Most DLA dielectrics under consideration, such as silica, silicon, sapphire, and diamond, have only the third-order nonlinearity, which is responsible for phenomena such as Raman scattering, Brillouin scattering, optical Kerr effects, and four-wave mixing (FWM).

Raman scattering can transfer the pump laser energy to frequency downshifted Stokes waves or upshifted anti-Stokes waves. The frequency shift is equal to one of the vibration-mode frequencies in a dielectric. Since Raman scattering is an inelastic scattering process, stimulated Raman scattering (SRS) not only depletes the pump laser power but also generates heat that could damage the dielectric. As SRS does not require phase matching, SRS can occur easily with a high enough pump power. The SRS threshold power can be defined as the pump laser power required to generate a comparable Stokes power (Smith, 1972), given by $P_{\mathrm{SRS}} = 16A_{\mathrm{eff}}/fg_R l$, where f is the material fill factor of a DLA, A_{eff} is the effective cross-sectional area of the laser beam, g_R is the Raman gain, and l is the DLA length. Because pump depletion occurs at the so-defined SRS threshold, l can be considered as a Raman-induced pump-depletion distance for a DLA, given by $l = 32Z_0/fn_r g_R E^2$, where $Z_0 = 377\ \Omega$ is the free-space impedance, n_r is the refractive index, and E is the peak laser-field amplitude. For a silica DLA driven by a 1 μm laser, this pump-depletion distance is about 17 cm for $f = 50\%$, $E = 1$ GV/m, $n_r = 1.45$, and $g_R = 10^{-13}$ m/W.

Stimulated Brillouin scattering (SBS) is a process similar to SRS, except that the pump laser is reflected in the backward direction with its frequency downshifted by an amount equal to the frequency of a generated acoustic phonon in the forward direction. The SBS gain is characterized by a narrow linewidth equal to the inverse of the phonon lifetime in a material. In most materials, the phonon lifetime is longer than 1 ns and

SBS rarely occurs for a picosend to femtosecond pump laser (Agrawal, 1989).

The optical Kerr effect is related to the intensity-dependent refractive index in a dielectric, given by $n_r = n_0 + n_2 E^2$, where n_0 is the unperturbed linear refractive index and n_2 is linearly proportional to the third-order nonlinear optical coefficient. This intensity-dependent refractive index can introduce an additional phase shift to the drive laser beam through self-phase modulation (SPM). The amount of phase error accumulated over a distance l is equal to $\Delta\phi = 2\pi f n_2 E^2 l/\lambda$, where λ is the pump laser wavelength. The condition $|\Delta\phi| = \pi$ defines a dephasing length of a DLA, beyond which particle deceleration takes place. For a silica DLA with $f = 0.5$ and $n_2 = 5.2 \times 10^{-23}$ $(\mathrm{V/m})^{-2}$ (Milam and Weber, 1976), this dephasing length is 2 cm for $E = 1$ GV/m, $n_0 = 1.45$, and $\lambda = 1$ μm. In practice, an optical phase shifter can be installed every dephasing length to correct the phase error caused by SPM. It is also possible to correct the phase error by detuning the pump laser frequency slightly. For example, approximately 10 ppm (parts per million) frequency detuning can compensate the SPM-induced phase error of a PBG-DLA working at a peak acceleration field of 0.38 GeV/m (Lin, 2001).

The optical Kerr effect does not always act adversely for particle acceleration. For example, a Gaussian laser beam with peak intensity on the axis can undergo self-focusing in a dielectric with a positive n_2. At a laser waist, self-guiding occurs when diffraction is balanced by self-focusing at a critical power $P_{\mathrm{cr}} = \lambda^2/16\pi n_2 Z_0$ (Saleh and Teich, 2007). For fused silica, this critical power is about 1 MW at $\lambda = 1$ μm. Similar to the self-guiding in the transverse dimension, a constant laser pulse duration can be maintained through the interplay of the optical Kerr effect and material dispersion at a critical laser field $E = \sqrt{-D_v \lambda/n_2}/2\pi\tau_0$, where D_v is the group velocity dispersion of a laser pulse in a dielectric and τ_0 is the characteristic laser pulse width (Hasegawa and Tappert, 1973a, 1973b). Such a constant-duration laser pulse is called a soliton pulse. To form a soliton pulse, $-D_v/n_2$ has to be positive. For instance, in fused silica, $D_v = -0.25$ ps/m THz at $\lambda = 1.6$ μm. A 1 GV/m laser field is sufficient to form a 14-fs soliton pulse at 1.6 μm in fused silica.

FWM is a third-order parametric scattering process, which manifests itself as energy exchanges among four mixing waves in a dielectric. For FWM to be efficient, it requires phase matching among the mixing waves. In most cases, phase mismatch from material dispersion can discourage FWM of different frequency components.

Among all the nonlinear effects, it is most important to suppress SRS. Multiple Stokes waves from SRS can add several E^2 terms at different frequencies into the optical Kerr effect $n_r = n_0 + n_2 E^2$ and induce so-called cross phase modulation (XPM) to further dephase particle acceleration. Multiple Stokes components can also build up FWM. The joint effects from SRS, SPM, XPM, and partially phase-matched FWM can convert the pump laser into an ultra-broadband radiation called supercontinuum (Alfano and Shapiro, 1970a, 1970b, 1970c).

To increase the SRS threshold for a DLA, a straightforward approach is to select a dielectric with a small Raman gain. For example, silica has a Raman gain 10 to 20 times smaller than that of sapphire (Grasyuk *et al.*, 1998) and 2 to 3 orders of magnitude smaller than that of silicon (Rong *et al.*, 2005) and diamond (Kaminskii *et al.*, 2007). It has also been demonstrated with success that one can design a stop band in a PBG fiber to fully suppress SRS in silica (Taru, Hou, and Knight, 2007; Blin *et al.*, 2009). In principle, the same SRS-suppression scheme can be applied to the design of a PBG-DLA. A small or negligible material filling factor f can be very helpful to reduce the nonlinear effects. The filling factor is virtually zero for a reflection-type DLA (Huang, 1998; Huang and Byer, 1998), as the pump laser does not propagate into the material. Unlike a transmission-type DLA (Lin, 2001; Cowan, 2008), a side-pumped DLA (Huang and Byer, 1996; Yoder and Rosenzweig, 2005; Plettner, Lu, and Byer, 2006) does not have a laser continuously propagating along the accelerator axis and the effective f is relatively small.

In conclusion, the adverse nonlinear effects in a DLA are mostly manageable. For example, SBS cannot occur due to the long phonon lifetime in comparison with the short laser pulse. FWM requires phase matching and is less likely to build up in a material with dispersion. Suppression of SRS, by designing a suitable stop band or adopting a small filling factor, or doping some impurity to broaden the Raman linewidth for a DLA, helps avoid XPM, FWM, and supercontinuum. The SPM-induced phase error can be corrected by installing an optical phase shifter every dephasing length or by detuning the pump laser frequency. The optical Kerr effect sometimes benefits transverse and longitudinal laser guiding in a DLA structure.

IV. SOURCES, BUNCHING, AND INJECTION

A. High-brightness electron microsources

1. Requirements on electron beam quality

Dielectric laser acceleration with driving wavelengths in the optical and near-infrared region requires electron beams with normalized emittance of a few nanometers ($\epsilon_n = 0.001$ mm mrad $= 1$ nm) so that the beams can be fed into and kept inside the structure channel. In addition, for most applications, particularly for free-electron laser applications, high peak currents are mandatory, requiring high-brightness electron beams. The source parameters for the Stanford DARPA Advanced X-ray Integrated Systems (AXiS) project, for example, are shown in Table V. Other

TABLE V. Example requirements for DLA: the injector electron beam parameters of the Stanford AXiS project, funded by DARPA. Numbers in parentheses refer to DARPA-BAA-11-11 as of 17 December 2010.

Parameter	To be met
Injection energy	1 MeV
Peak current	0.06 A
Bunch charge	20 fC (< 1 pC)
Bunch length	330 fs (< 1 ps)
Emittance (norm)	1 nm (< 50 nm)
Normalized brightness	4×10^{13} A/m^2
Energy spread	0.1%
Beam radius at DLA entrance	0.4 μm
Pulse repetition rate	25 MHz (> 10 MHz)

DLA approaches will likely put similarly stringent requirements on their beam parameters when driven with near-infrared optical wavelengths. This is because the driving wavelength sets the typical dimensions of the structure through which the electron beam must travel.

The beam brightness given in Table V is the energy invariant peak brightness B_n defined as the geometric brightness B divided by the relativistic factor $\beta^2\gamma^2$ (Reiser, 1994),

$$B_n = \frac{B}{\beta^2\gamma^2} = \frac{2I}{\pi^2\epsilon_n^2}, \tag{59}$$

where I is the electron bunch peak current.

The DLA beam brightness requirement is approximately the same as the brightness already demonstrated by both rf and dc photocathode guns. For example, the beam from a photocathode rf gun operating with a bunch charge of 20 pC and a peak current of 5 A has a measured normalized emittance of 140 nm (Ding *et al.*, 2009). This beam has a normalized brightness of 5×10^{13} A/m^2. This is slightly better than the brightness to be met by the AXiS project.

Transverse beam dimensions must be matched to the acceptance of an optical-scale accelerator structure. The bunch length out of the gun must also be microbunched by a factor of 1000 down to the attosecond scale. Attosecond bunching is discussed in Sec. IV.B. Here we discuss the low emittance and related submicron transverse size that are essential to transporting the beam through the small diameter aperture of the DLA.

2. Current status of conventional photocathodes

Conventional photocathodes used in today's accelerator facilities usually consist of a flat metal or semiconductor surface inside an accelerator microwave cavity. Typically, UV laser pulses are fired at the surface to set free electrons via the photoelectric effect. These electrons are then accelerated by the microwave field. These so-called rf guns achieve larger acceleration fields (~100 MeV/m) than dc guns, which help to boost electrons to relativistic energies, thus mitigating harmful collective effects such as Coulomb repulsion. A similar concept applies to tip-based sources, as discussed next. The electron source diameter is given by the spot size of the laser beam, and the initial acceleration field is given by the (microwave) field applied to the source surface (Chao and Tigner, 2006). Electron injectors in modern accelerator facilities have improved to the point that high-brilliance electron beams are now limited by the intrinsic emittance of the photocathode (Ding *et al.*, 2009).

3. Tip-based ultralow emittance sources

Although development in the conventional, flat photocathodes continues, new approaches arose based on laser-driven photoemission from sharp metal tips (Hommelhoff *et al.*, 2006). Sharp metal tips emitting electrons via so-called Schottky and cold field emission have long been and continue to be standard dc electron sources in high-resolution electron microscopes, owing to their extremely high brightness (Spence, 2009). A reduced brightness of $B_r = 5.1 \times 10^7$ A/(m^2 sr V) has been measured from sharp tungsten tips, and $B_r = 3 \times 10^9$ A/(m^2 sr V) from carbon nanotube emitters (DeJonge *et al.*, 2002). Even larger values may be obtained for single-atom tip emitters, but due to their minimal geometrical source size they are likely of no use for accelerator applications. Jensen *et al.* (2010) found analytically that the normalized (rms) emittance of a single tip equals $\epsilon_n \approx 360\hbar/mc = 0.13$ nm.

Using sharp tips as electron sources is advantageous for a number of reasons. The source size can be extremely small, and the emittance is even smaller than what can be expected from the geometric source size, because in photoassisted field emission electrons leave the tip nearly perpendicular to the surface. If the electron trajectories are traced back, they seem to originate from a virtual source of smaller area. It is this virtual source that determines the electron optical properties (Spence, Qian, and Silverman, 1994).

Because the tip surfaces can be made atomically clean, surface fields limited only by the onset of dc field emission can be applied. For tungsten, this means that a dc field as large as 2 GV/m can be applied to the tips. Much like in conventional photocathodes, the large field is advantageous to accelerate the electrons quickly to high energies and so mitigate initial-energy–to–timing jitter effects and collective repulsion effects. However, unlike conventional photocathodes, the field drops in front of the tip on a length scale given by the tip radius. Thus, in order to efficiently accelerate electrons, the field configuration should be such that the field at the tip is at ~2 GV/m, after which it is reduced to a constant dc field whose magnitude is set by surface breakdown, similar to conventional photocathodes (Hoffrogge *et al.*, 2014). Note that collective effects only set in once the electrons have left the tip, so the tip radius can be used as a free parameter whose optimum value can be found by considering the required bunch charge.

4. Focusing requirements

Much of the behavior of particle beams can be understood using the beam envelope theory developed by Reiser, Lawson, and others (Reiser, 1994). This theory gives valuable insight into basic beam dynamics since it includes external focusing forces as well as space-charge forces. Here we use it to estimate the space-charge emittance and focusing needed to confine the beam from a tip cathode.

For the analysis of tip cathode beam dynamics, we recall the radial beam envelope equation given by

$$r_m'' + \frac{\gamma' r_m'}{\beta^2\gamma^2} - \frac{\gamma'^2}{\beta^2\gamma^2} r_m + k_r r_m - \frac{\epsilon_n^2}{\beta^2\gamma^2 r_m^3} - \frac{K}{r_m} = 0. \tag{60}$$

This is the equation of motion for the beam's radial envelope at radius r_m. The envelope radius r_m is consistent with the emittance ϵ_n. The first term is the radial acceleration of the envelope radius r_m. The second and third terms are the first- and second-order forces due to the beam's acceleration. The fourth term is for the focusing of a lens using magnetic or electrostatic fields. For a magnetic solenoid lens with an axial field B_{sol},

$$k_{r,\text{solenoid}} = \left(\frac{eB_{\text{sol}}}{2mc\beta\gamma}\right)^2, \tag{61}$$

while for a lens with effective length l_{eff} and focal length f, the focal strength is

$$k_{r,\text{lens}} = \frac{1}{l_{\text{eff}} f}. \tag{62}$$

The fifth term gives the outward pressure force due to the beam's emittance, i.e., divergence. The sixth term gives the radial acceleration due to electrostatic forces between the electrons, often referred to as space-charge forces. The space-charge defocusing force is given in terms of the generalized perveance K defined as

$$K \equiv \frac{2}{\beta^3\gamma^3}\frac{I}{I_0}, \tag{63}$$

where I_0 is the characteristic current, $I_0 = 17$ kA.

Multiplying the envelope equation by $\beta^2\gamma^2 r_m^3$ gives an equation relating the normalized emittance to the space-charge and external focusing terms. If we assume the beam is coasting at constant velocity $\gamma' = 0$, and there is an external focusing lens that balances the space-charge and emittance terms, then $r_m'' = 0$. For such a beam there is an equilibrium (constant) radial envelope obtained when the focusing of the external lens balances or counteracts the outward expansion produced by the emittance and space-charge forces. This equilibrium condition is called *balanced* or *Brillouin flow* and occurs when

$$k_{r,\text{balanced}}\beta^2\gamma^2 r_m^4 - \epsilon_n^2 - \frac{2I}{\beta\gamma I_0} r_m^2 = 0. \tag{64}$$

By inspection of Eq. (64), we can see that characteristic space-charge emittance can be defined as

$$\epsilon_{n,\text{sc}} = r_m\sqrt{\frac{2I}{\beta\gamma I_0}}. \tag{65}$$

Since the field enhancement extends only about one tip radius away from the tip surface, we estimate the electron energy of this distance from the tip as half the surface field times the tip radius. For a surface field of 2 GV/m and a radius of 0.4 μm, the electron energy is approximately 400 eV with velocity relative to c of $\beta = 0.04$. Applying the AXiS parameters in Table V gives $\epsilon_{n,\text{sc}} = 11$ nm. The lens strength necessary to maintain the beam to 0.4 μm size, and the 22 nm emittance can be estimated by equating the focusing and space-charge terms in Eq. (64) and solving for the lens strength,

$$k_{r,\text{balanced}} = \frac{2}{\beta^3\gamma^3}\frac{I}{I_0}\frac{1}{r_m^2}. \tag{66}$$

The lens strength needed to confine the beam grows rapidly as the envelope radius is decreased. Using the parameters in Table V, we find that within 1 μm from the cathode when the beam is at 400 eV, $k_{r,\text{balanced}} = 5.7\ \mu\text{m}^{-2}$. The electrostatic lenses proposed for focusing beams from tip cathodes have effective lengths of approximately 1 μm and voltages of 10 to 100 eV produce the required micron-length focal lengths in a micron-long gap (Helfenstein *et al.*, 2011; Mustonen *et al.*, 2011). Electrostatic focusing is required since magnetic lenses are much less effective than electrostatic lenses at these low beam velocities. For example, the magnetic field needed to give the same $k_{r,\text{balanced}} = 1\ \mu\text{m}^{-2}$ is 170 T. This is an impractically high magnetic field.

5. Scaling of transverse emittance

The total normalized emittance is the quadratic sum of a variety of emittance generating effects. For tip cathodes three of the largest contributions are the intrinsic emittance, the space-charge emittance (as just discussed), and the field enhancement emittance,

$$\epsilon_n = \sqrt{\epsilon_{n,\text{intrinsic}}^2 + \epsilon_{n,\text{sc}}^2 + \epsilon_{n,\text{field}}^2}. \tag{67}$$

The intrinsic emittance results from the material properties such as the work function and the character of the electron density of states of the cathode (Dowell and Schmerge, 2009). It can be shown that intrinsic emittance is given by

$$\epsilon_{n,\text{intrinsic}} = \sigma_x\sqrt{\frac{\hbar\omega - \phi_{\text{eff}}}{3mc^2}}, \tag{68}$$

with σ_x the rms laser spot diameter on the photocathode, $\hbar\omega$ the driving laser's photon energy, mc^2 the electron's rest energy, and ϕ_{eff} the effective barrier height (material work function reduced by the Schottky effect due to applied high-voltage field) given by

$$\phi_{\text{eff}} = \phi_W - \phi_{\text{Schottky}} = \phi_W - \sqrt{\frac{e\beta_{\text{field}}F_0}{4\pi\epsilon_0}}. \tag{69}$$

Here ϕ_W is the cathode material work function, F_0 is the externally applied electric field far from the tip where there is no field enhancement, and β_{field} is the field enhancement factor of the tip with respect to a flat cathode. For tips with micron radii and smaller, the field enhancement factor is 10 and larger. Therefore field enhancement at the tip easily increases the surface field into the GV/m range, resulting in large Schottky potentials which reduce the effective work function. For example, a 1 GV/m surface field has $\phi_{\text{Schottky}} = 1.2$ eV. This lowering of the effective work function increases the intrinsic emittance approximately a factor of 2 with respect to a flat cathode for a typical metal work function of 4.5 eV and a photon energy of 4.9 eV.

From Eq. (68) it is clear that the intrinsic normalized emittance can, in principle, approach zero if either the laser spot size approaches zero or the laser photon energy equals the effective barrier height (Hauri *et al.*, 2010). In the latter case, however, the quantum efficiency of the emission process can be shown to approach zero faster than ϵ_n, hence this is not a useful approach. The former approach (decrease σ_x) is limited

theoretically to about one-half the driving wavelength, or, for practical purposes, to about two wavelengths.

The field emittance $\epsilon_{n,\mathrm{field}}$ is produced by the divergent electric field produced by a nonflat cathode. The field diverges in order to meet the boundary condition of no transverse electric field at the surface of a conductor. For large area flat cathodes commonly used in rf guns, this effect manifests itself via the surface roughness which is minimized to keep the field emittance negligible (Xiang *et al.*, 2007). However, for tip cathodes this can be an important effect. We estimate the field emittance from a tip of radius 10 nm by numerically integrating an analytic expression for the tip's electric field to obtain the electron trajectories out to a few microns from the tip. Using this tip geometry at 100 MV/m applied field, we have $\beta_{\mathrm{field}} = 22$ and $\epsilon_{n,\mathrm{field}} = 0.02$ nm due to the very small rms emission size of 2.9 nm.

6. Current status of tip-based and tip array-based approaches

A number of approaches have been investigated using laser-triggered electron emission from sharp tips (Boussoukaya *et al.*, 1989; Garcia and Brau, 2002; Hommelhoff, Kealhofer, and Kasevich, 2006; Hommelhoff *et al.*, 2006; Barwick *et al.*, 2007; Ropers *et al.*, 2007; Yanagisawa *et al.*, 2009). Ganter *et al.* (2008) performed photofield emission tests from a single ZrC needle. With 16 ps long laser pulses at 266 nm, they obtained electron pulses with up to 150 pC and a peak current of 2.9 A at 60 keV in a pulsed high-voltage operation, at a repetition rate of 30 Hz and a tip radius $r \approx 50$ μm. The estimated emittance was below 0.05 mm mrad, without taking into account the virtual source size being much smaller than the geometric source size, which would further reduce it.

Other approaches are based on nanofabricated field emission arrays, also known as Spindt emitters (Spindt *et al.*, 1976). These cathode array sources underwent massive development until about a decade ago as a means to develop flat screen cathode ray tubes but were outperformed by other display technologies. Nanofabricated field emission arrays usually consist of an array of sharp emitters gated with an anode layer about a micron away from the needles. The group at Paul Scherrer Institute has started to pursue this route (Tsujino *et al.*, 2008; Mustonen *et al.*, 2011). When 1.2×10^5 Mo tips with $r \approx 10$ nm and a pitch of 5 μm were illuminated with 50 fs laser pulses at a 800 nm center wavelength, a total charge of 5.2 pC per pulse was drawn from the array.

In a more complex setup, Graves *et al.* (2012) proposed to take advantage of the many parallel beamlets emitted from gated cathode arrays to obtain a longitudinally microbunched beam, as discussed in Sec. IV.B. Taking into account electrostatic lens abberations in their double-gated cathode array, they calculated a single-tip emittance of 0.008 to 0.020 nm at the cathode exit, which is 1 to 2 orders of magnitude larger than the initial emittance at the tip, for a charge of 10 fC per tip.

In summary, several groups have started pursuing the development of tip-based electron sources for accelerator applications. Recent technological progress in 3D nanoscale engineering has a direct impact. Because of the new effects such as plasmon assisted laser-tip (array) interaction and new degrees of freedom (including but not limited to tip radius, gate design, emitter material), the future of tip-based sources seems promising.

B. Attosecond scale bunching techniques

1. Optical density modulation

In the context of optical particle acceleration schemes, microbunching of a longer beam is needed to fit the particles into individual acceleration buckets (typically a fraction of a wavelength). It is a means for both improved system efficiency and better beam quality of the accelerated beam. Microbunching of an electron beam is a process through which a much longer pulse of electrons is divided into the order of 100 to 1000 consecutive subpulses. The timing interval between the individual pulses can range from picoseconds to the subfemtosecond regime. Optical microbunching can often arise in particle accelerators and storage rings as an unwanted, deleterious effect brought on by coherent synchrotron radiation in bending magnets (Bergher, 1997; Stupakov and Heifets, 2002) limiting the maximum possible bunch compression and/or current. Microbunching is also a necessary consequence of FELs in which longitudinal density modulation of the electron beam occurs simultaneously with the onset of exponential amplification of the laser radiation (Brau, 1990).

Several successful experiments in optical acceleration made use of the IFEL interaction to attain a bunched beam (Kimura *et al.*, 2001; Musumeci, 2005; Sears *et al.*, 2008). In the case where a bunched beam is the end goal, an intense laser field is provided at the start of a much shorter undulator. The beam is energy modulated with an amplitude greater than the initial-energy spread.

After the IFEL, the electrons pass through a dispersive chicane (or magnetic compressor) to transform the energy modulation into a longitudinal density modulation. The equations for bunching of electron beams were first formulated for nonrelativistic beams in klystrons (Webster, 1939). In klystrons, bunching is from velocity modulation and time of flight. Here the chicane replaces the velocity modulation and instead varies the path length of electrons based on their modulated energy. Since both processes are linear, however, the same equations apply. The longitudinal density modulation ρ can be written as a cosine expansion with Bessel function coefficients,

$$\rho(z) = \rho_0 \left[1 + 2 \sum_{n=1}^{\infty} J_n(n\beta_b) \cos(nkz) \right]. \tag{70}$$

In Eq. (70), $\beta_b = kD_z\eta/\gamma_0$ describes the degree of bunching where η is the normalized amplitude of IFEL modulation, γ_0 is the normalized average beam energy (energy per rest mass), $k = 2\pi/\lambda$ is the laser wave number, and ρ_0 is the initial density of the beam. D_z is the longitudinal dispersion (or momentum compaction) of the chicane, which can be defined through magnetic measurements. Equation (70) can be further refined (Baccaro, DeMartini, and Ghigo, 1982; Luccio *et al.*, 1990) to include an initial-energy spread σ_γ,

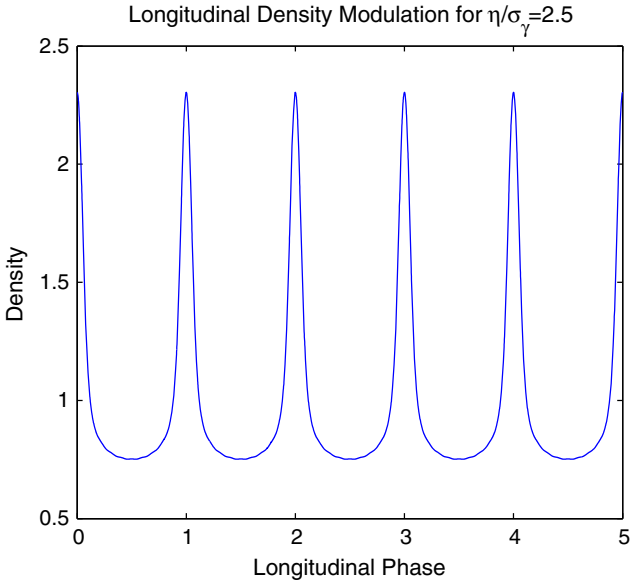

FIG. 36 (color online). Plot of Eq. (71) with $\eta/\sigma_\gamma = 2.5$ and $\beta_b = 1$. Density is relative to the initial density ($\rho_0 = 1$). From Sears, 2008.

$$\rho(z) = \rho_0 \left[1 + 2\sum_{n=1}^{\infty} b_n \cos(nkz)\right], \tag{71}$$

$$b_n = J_n\left(nkD_z \frac{\eta}{\gamma_0}\right) \exp\left[-\frac{1}{2}\left(\frac{nkD_z\sigma_\gamma}{\gamma_0}\right)^2\right]. \tag{72}$$

The Fourier series coefficients b_n are termed the *bunching coefficients* and are often quoted as a figure of merit for bunching in FELs and IFELs. Optimal bunching occurs when the bunching coefficients are maximized, which in turn occurs for $\beta_b = 1$ to 1.5 depending on the harmonic number. For $\beta_b = 1$, we can write the expression in the exponent as $nkD_z/\gamma_0 = n/\eta$. We then see that when the IFEL modulation η is less than the initial-energy spread σ_γ the microbunching is suppressed. Further attenuation terms can be included to account for electron beam divergence and laser wave-front curvature. These, however, tend to be much smaller corrections compared to the initial-energy spread term. Figure 36 shows a plot of the longitudinal distribution for maximum bunching and $\eta/\sigma_\gamma = 2.5$.

2. Capture efficiency and multicolor bunchers

In theory, one need not have a separate microbunching stage at all. An initially unbunched beam with a pulse length much longer than the accelerating wavelength, sent through an optical accelerator, will emerge with a very large energy spread and an energy collimator could be used to select a small portion for further use. However, this is a very inefficient use of the initial charge produced and may produce unacceptable radiation to the surroundings and even damage to the accelerator depending on the total current and other system parameters. By bunching the beam, the captured fraction can jump from a few tens of percent in a 5% energy window to $> 80\%$ with an energy spread of $< 1\%$ (Duris, Musumeci, and Li, 2012). The microbunches also emerge with higher peak current, which is an essential factor in some applications such as light sources (Brau, 1990).

Optical microbunching can be further improved by using multiple wavelengths (Pottorf and Wang, 2002; Duris, Musumeci, and Li, 2012). Whereas a single frequency modulation will impart a sinusoidal energy modulation on a beam, by using additional frequencies the energy modulation can more closely emulate a triangle wave. When the modulated beam is then sent into a dispersive system (chicane), the longitudinal phase space is rotated with nearly all of the charge packed into microbunches. Figure 37 shows a diagram of an IFEL based, multifrequency (or *multicolor*) buncher and the resulting longitudinal phase space. For a multicolor buncher to be successful, the relative phase and intensity of each optical frequency must be well controlled. The multicolor buncher results in a modest capture efficiency increase [98.5% in Duris, Musumeci, and Li (2012)], but perhaps, more importantly, can result in substantially improved energy spread approaching 0.1%. This is a result of the improved bunch density at the start of the optical accelerator. Such low-energy spread is also a key parameter in coherent light sources (Brau, 1990).

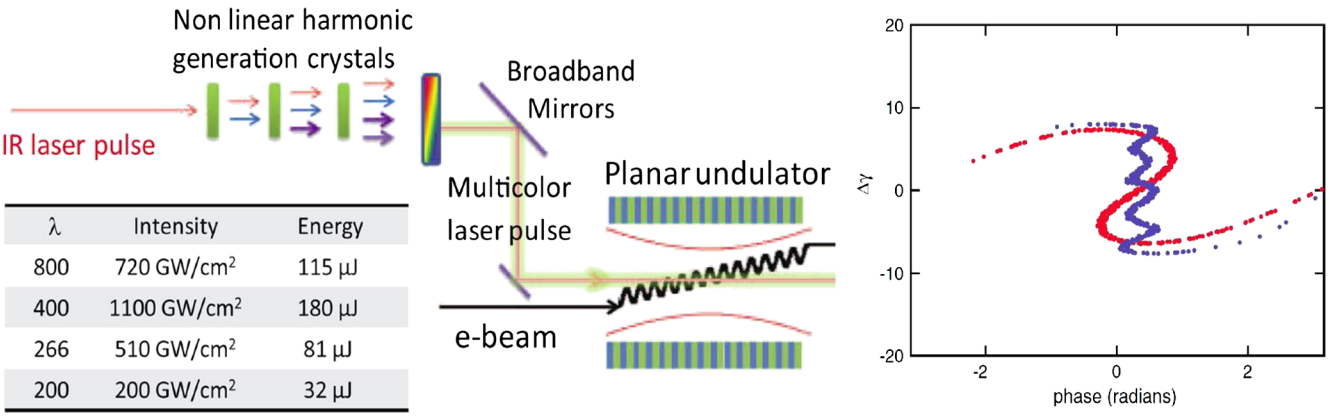

FIG. 37 (color online). Layout of a multicolor IFEL microbuncher and resulting longitudinal phase space. From Duris, Musumeci, and Li, 2012.

3. Bunching via emittance exchange

An altogether different approach to bunching involves transverse-longitudinal phase space exchange. In this case, the transverse phase space of a beam is modulated, for example, by sending a beam through a slotted aperture array. If the slotted aperture is in turn located in a region of transverse-longitudinal correlation (spatial chirp), for example, inside a magnetic chicane, the beam is bunched. Muggli *et al.* (2008) used this technique to generate a beam modulated at subpicosecond time scales for terahertz generation, although the technique could be extended to optical frequencies either by tuning the slot and chicane geometries to directly produce an optically bunched beam or by using bunch compression techniques to compress the subpicosecond modulation into the optical range.

Graves *et al.* (2012) generated a spatially modulated beam at the moment of emission using a field emitter array. This is then accelerated and the phase space rotated in an emittance-exchange transfer line (Cornacchia and Emma, 2002). This work targets bunching at 13 nm for extreme ultraviolet lithography applications, but with small adjustments could equally be applied to optical acceleration.

C. Low-velocity injection and preacceleration

All accelerators require the creation and injection of a particle beam, with properties (size, time structure, and energy) that must match the dimensions and fields of the high-energy accelerating structure. Since DLAs require beam parameters that are quite distinct from those of conventional machines, their further development requires a new class of sources and injectors. The distinct beam properties needed for DLAs generally include attosecond pulses with very low charge (∼1 fC) per bunch, very high repetition rates (MHz to GHz), and extremely small spot sizes (∼10–100 nm). Promising candidate mechanisms for microscale electron sources, and techniques for bunching them, were discussed in Sec. IV.B. However, unless these beams can be successfully injected into a DLA structure, the problem is only partly solved.

The most significant challenge to successful injection is the mismatch in energy between candidate electron sources and the near-relativistic energy regime in which DLAs operate most effectively. At the micron scale, there is no equivalent to the conventional rf gun, in which electrons are brought to relativistic energies in just one or two wavelengths. With laser-driven devices, the energy gain per optical period is very small (typically on the order of a few keV, at most) and the bunch will have $\beta = v/c$ well below unity for tens or hundreds of periods. Any DLA concept must therefore include a method for subrelativistic preacceleration.

Although a conventional rf preacceleration stage may be conceivable, it will undercut many advantages of the DLA: the increased size of that stage would make true miniaturization impossible (and would be mismatched in scale to the DLA itself), the added complexity would lead to synchronization challenges, and rf technology would likely limit the achievable repetition rates. Therefore, it is desirable to find subrelativistic or "low-beta" DLA-like structures capable of optically accelerating electrons from tens of keV to roughly 1 MeV. Development of low-beta acceleration schemes would also have applications in structure-based proton and ion acceleration. Possible schemes can be divided into two approaches: those in which the electron velocity and optical phase are synchronous, and those in which they are not. Note that in either approach, the particle velocity is continually increasing (for example, from $0.3c$ to $0.8c$) and the structure parameters must also change, either continuously or stepwise, to maintain continuous acceleration over that velocity range.

For synchronous acceleration schemes, the phase velocity must match the particle velocity, which is well below c. In general, a synchronous interaction for $v < c$ particles can only be created using periodic material boundaries near the beam path. This scheme is especially suited to nonresonant, near-field structures in which the optical phase can be controlled to a high level by detailed adjustment of the structure periodicity and coupling. Alternately, in a near-field structure without strong cell-to-cell coupling, portions of the beam path can be made nearly field free, thus shielding the beam from off-phase deceleration as is done in a drift-tube linac.

By contrast, the symmetry of Maxwell's equations means that resonant, standing-wave devices can have phase velocities only less than c if they include field-shaping boundaries, which would destroy many of the advantages of these open structures. In resonant, open-boundary DLAs, an asynchronous scheme must be used, in which a deliberate mixing of optical wavelengths leads to net acceleration, even though the particle and phase velocities are never synchronous. This scheme requires a subharmonic periodicity to be introduced, either with multiple-wavelength illumination or with monochromatic illumination of a subwavelength structure.

1. Examples

Two examples of low-beta DLA designs illustrate these approaches further. As an example of a synchronous approach, Breuer and Hommelhoff (2012) extended the grating acceleration concept described by Plettner and Byer (2008) to the low-beta regime. This design is a near-field structure in which a transversely illuminated double grating structure supports evanescent fields near its inner surfaces, which can provide synchronous acceleration when phased appropriately. In principle, extending the design to the low-β regime can be accomplished merely by changing the grating period from λ to $\beta\lambda$, or a multiple of it. However, the efficiency of coupling to the evanescent wave goes down rapidly as β decreases; the ratio of the accelerating field to the incident laser amplitude is only a few percent for $\beta < 0.8$, as compared to more than 0.5 for $\beta = 1$. Nevertheless, acceleration gradients of up to 50 MeV/m have been demonstrated for 27 keV electrons in simulation, and experimental testing is underway.

An asynchronous low-beta accelerator based on the micro-accelerator platform (MAP) (Travish and Yoder, 2011) was designed at UCLA (McNeur *et al.*, 2012) and is being fabricated. As the MAP is a resonant, open-boundary structure, the phase velocity cannot easily be lowered, and a subharmonic periodicity must be generated. This can be done effectively by introducing a greatly increased structure periodicity that is an integer multiple of both the laser wavelength λ

and $\beta\lambda$. In that case, the accelerating field becomes strongest directly under the coupling slots, and correct phasing of the electron bunch allows net acceleration even though the bunch experiences some decelerating fields. In simulation, a MAP-like structure with periodicity $10\lambda = 30\beta\lambda$ shows acceleration gradients of 400 MeV/m for 25 keV electrons.

2. Focusing

In general, adapting accelerating structures for $v < c$ operation introduces defocusing forces on the beam. In the case of the relativistic MAP structure, the translational symmetry and consequent speed-of-light accelerating mode are fundamental to the operation of the accelerator, and in the relativistic limit deflecting forces vanish due to that symmetry. For nonrelativistic operation, the symmetry is necessarily broken, with the consequence that simultaneous transverse focusing and longitudinal stability are no longer possible. External magnetic focusing is one possible solution, although of course it introduces additional complexity, as is systematic structure perturbation inspired by alternating focusing and defocusing (FODO)-type lattices (McNeur *et al.*, 2010). The evanescent accelerating mode in the grating structure also includes a defocusing component, which becomes stronger as the beam approaches the grating. Increasing the beam standoff distance keeps deflection manageable but also reduces the accelerating field experienced by the beam.

3. Open issues

A number of potential challenges remain to be explored in the area of low-beta laser acceleration. As the structure properties must be matched to the beam velocity, which changes significantly only on a scale of tens or hundreds of periods, simulations to date have been carried out only at single values of β. For the asynchronous resonant structure, it is still possible that an alternate approach using a traveling rather than standing wave may improve the low-beta interaction, although a traveling-wave approach introduces challenges with coupling and power density. A reduction in gradient, as with the grating structure, may also be necessary to reduce defocusing to a manageable level, although this would of course limit the overall efficiency of the device.

Beam dynamics issues in both types of low-beta accelerators remain to be fully explored. Considerations include trapping efficiency and the possibility of structure damage by untrapped particles, beam expansion due to space charge at low β, and beam-loading effects, which would be expected to vary along the interaction length.

V. LASERS FOR DIELECTRIC PARTICLE ACCELERATORS

A. Requirements

The laser requirements for a DLA-based accelerator reflect the stringent power and efficiency requirements for future linear colliders, outlined in Sec. VI.A, as well as the unusual pulse format of the electron beam: namely, very high repetition rates with low per-pulse energy but high average power. This compares favorably to much higher and more difficult to attain pulse energies needed for other laser-based particle accelerator methods. In addition, because each laser pulse can drive an entire bunch train in the DLA scenario, subpicosecond pulse lengths are not required. We discuss the laser requirements, the state of the art in fiber lasers (the recommended laser technology for this application) next. We also present a nominal baseline design for a modular system designed to drive multiple accelerator stages.

The laser requirements for four examples of DLA accelerator structures are presented in Table VI, corresponding to estimated requirements for a future linear collider consistent with beam parameters and pulse format outlined in Secs. III and VI.A. For the DLA applications, the pulse envelope is of the order of 1 ps. Thus carrier envelope phase locking (CEP) may not be required, as the pulse envelope is very long compared to an optical cycle. However, the optical phase of the carrier wave needs to be locked to the phase of the accelerating electron beam in order to ensure each consecutive laser pulse is in phase with the microbunches and thus accelerating the beam. The nominal laser type will be a fiber laser because of its efficiency, power scaling capabilities, and reliable operation (Dawson *et al.*, 2008). To achieve a uniform acceleration of the electron packet, a flat top super-Gaussian pulse in time will be used to maintain a constant electric field across the pulse. Polarized light is necessary for most structures; however, the polarization extinction ratio is not a critical parameter and use of polarization-maintaining fiber should be more than sufficient to realize this goal. While the power requirements of the laser seem quite high, especially the repetition rate, the current state-of-the-art systems are actually quite rapidly approaching such performance.

B. State of the art for high peak power lasers

To date fiber laser systems offer the best potential to attain the combination of reliability and efficiency that would be ultimately required to make a laser-based particle accelerator. The beam quality of fiber lasers is typically superior compared to other lasers of similar power and pulse energy. High-power fiber laser systems (Gapontsev, Fomin, and Yusim, 2009; Goodno, Book, and Rothenberg, 2009; Richardson, Nilsson, and Clarkson, 2010) leverage the waveguide properties of optical fiber in order to achieve exceptional wall-plug efficiencies ($> 30\%$) and diffraction limited beam quality with high average output powers (> 10 kW). Pulsed fiber laser systems with pulse widths of a few nanoseconds are limited to around 4 MW peak power in a single fiber waveguide due to self-focusing (Cheng *et al.*, 2005). This limit has been attained with 1 ns pulses, and high quality beams with > 4 mJ output have been demonstrated from roughly 100 μm core diameter fiber rods (DiTeodoro *et al.*, 2010).

Ultrafast fiber lasers have demonstrated significant powers and pulse energies approaching the limits discussed with regard to nanosecond pulses above. Commercial fiber laser systems with up to 50 μJ pulse energies and subpicosecond pulses are currently available from a number of vendors (e.g., Raydiance, IMRA, Calmar Laser, Polar Onyx, Fianium). Given the current rate of development, one could reasonably expect to see mJ-class commercial systems with subpicosecond pulses available in the next 5 years.

To date, chirped pulse amplification (CPA) fiber laser systems with no pulse picking (direct amplification of a mode-locked oscillator) have attained average powers as high as 830 W at 78 MHz (10.6 μJ) (Eidam *et al.*, 2010). Similar systems with 80 W of power have been demonstrated with carrier envelope phase-locked oscillators (Ruehl *et al.*, 2010) suggesting the required amplification process is not inconsistent with low-phase-noise pulse trains. Another promising technology in this area is thin disk lasers, which have shown the ability to generate very high repetition rate pulse trains ($>$ 1 GHz) (Pekarek *et al.*, 2012) and generate reasonably high average powers ($>$ 100 W) (Baer *et al.*, 2010).

Recent results for CPA in fiber rods with pulse picking (lower repetition rate output pulse formats attained by only selecting one of every N pulses) have demonstrated 11 W of average power with 2.2 mJ pulse energies and $<$ 500 fs pulses (Eidam *et al.*, 2011). Systems with a very small amount of chirp ($<$ 100 ps) have been shown to achieve a few hundred nanojoules of pulse energy in $<$ 250 fs pulses with excellent pulse fidelity (Dawson *et al.*, 2009). Chirped volume Bragg gratings, a promising technology for simplifying CPA systems, have been demonstrated in a 0.5 μJ, 72 MHz system with compressed pulse widths of $<$ 200 fs (Rever *et al.*, 2009).

At present, there are two common fiber lasers that operate in the wavelength region longer than the Yb:fiber laser at 1 μm. One is the ytterbium-erbium (Yb, Er) system at around 1550 nm, and the other is the Tm-doped system at about 2000 nm. The Er-doped system is limited to a maximum optical efficiency of 65%, although in practice due to losses in the energy-transfer process the efficiency tends to fall in the 30–40% range. Tm-doped fibers are more promising, because it is possible to pump the Tm ions at around 800 nm, where efficient diodes are readily available. One can in theory obtain a pump to laser efficiency of 82%. In practice, efficiencies in the 60–70% range have been obtained in multi-hundred-watt lasers (Moulton *et al.*, 2009). We expect the wall-plug Tm: fiber laser efficiency to approach 40% with careful attention paid to pump coupling and power supply efficiency.

1. Baseline design and options for DLA

An outline of one possible baseline design for the DLA laser system for a TeV scale accelerator is shown in Fig. 38. The design is modular to enable easy scaling to the TeV level and facilitating timing synchronization across a long accelerator. These challenges would be somewhat reduced for a 100–1000 MeV application and the design should be directly applicable to those applications as well.

The baseline design begins by producing a carrier (envelope) phase-locked oscillator with its repetition rate matched to a stable rf reference frequency source in the range of 100 MHz to 1 GHz, with 1 GHz being the target. This oscillator will serve as the clock for the accelerator. The global oscillator or clock will be distributed via optical fiber to local oscillators, which are phase locked to the global oscillator. Each structure will require a phase control loop to allow for acceleration through successive structures. Both small-fast (acoustic, $<$ 10 kHz, $<$ 1 μm) and large-slow control (thermal time scale of seconds to hours, $>$ 10 μm) of the phase will be

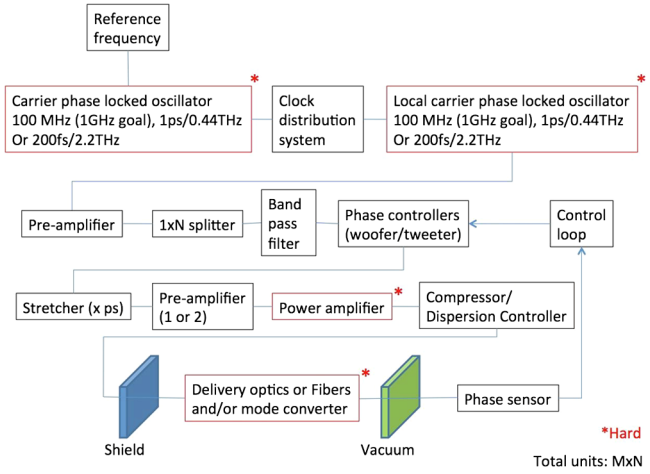

FIG. 38 (color online). The laser system baseline design. The outlined boxes highlight the challenging portions of the system. A total of M local oscillators will be built and split N times, giving a total of $M \times N$ laser coupled accelerator structures. From Bermel *et al.*, 2014.

necessary (Augst *et al.*, 2007). By monitoring the energy linewidth as well as the timing of the electron bunches, successful acceleration through the structures may be confirmed.

In order to get to the pulse energy necessary per structure for TeV scale acceleration (200 nJ to 10 μJ), the pulses will undergo CPA (Maine *et al.*, 1988). The design discussed here uses dispersion control immediately following the power amplifier. In this configuration, the dispersion controller, whether it is fiber (Galvanauskas *et al.*, 1995), grating pairs (Maine *et al.*, 1988), chirped mirrors (Szipocs *et al.*, 1994), or chirped volume Bragg gratings (Chang *et al.*, 2009), will set the dispersion so that after passing through the remaining elements, the pulse will be compressed to the optimal pulse length in the accelerator structure.

The baseline design looks to be a manageable system, with the toughest challenges coming from the requirements for the oscillators, the power amplifier, and the delivery optics. In addition, it will be necessary to repeat the local system multiple times, with each local system phase locked to the global oscillator.

2. Research and development challenges

Power delivery and coupling to accelerator structures as well as timing issues are unique requirements for DLA, and thus they are key areas for targeted research and design investments. The other areas such as kW power scaling of short pulses, compact stretchers and compressors, and cost control have synergies with other laser applications. In these areas, coordination of efforts with other communities interested in development of short pulse fiber lasers would be beneficial to everyone from an overall cost perspective.

a. Timing issues

As the acceleration process of DLA is linear with the electric field, the optical phase must be well controlled. Poor

synchronization would result in either a decrease of efficiency or an electron energy spreading or even defocusing. Frequency comb technologies can detect and control both the repetition rate of the delivered pulses and the CEP. The technology used to generate frequency combs in ultrahigh finesse Fabry-Pérot cavities is able to control phase noise in the range 0.01 Hz to 100 kHz. Further stabilization will necessitate control systems operating above 100 kHz and requires effort in feedback-loop electronics as well as ultrafast, ultralow noise detectors. This is a special need for the DLA application but no fundamental obstacles are foreseen.

b. Short pulses at kW average power

Depending on the accelerating technology adopted, driving lasers should deliver pulses from 100 fs to 10 ps with average powers ranging between a few 10 W to 10 kW. For grating or resonant structures where wavelength is not restricted, Yb-doped fiber laser and amplifier technology at 1 μm is rather close to fulfilling the requirements in terms of average power, pulse energy, and duration. Thulium-doped materials are probably the best candidates around 2 μm, and average power in excess of 1 kW (cw operation) was recently reported for a Tm-doped fiber laser. Further research is needed to produce and amplify femtosecond to picosecond pulses at such high average power in Tm-doped fibers, but this is a topic that other communities beyond DLA have an interest in to solve as well.

c. Stretcher and compressor

Limiting the intensity during the amplification process requires first enlarging the beam diameter in a wider fiber and/or temporally stretching the pulse (to reduce power density), and then recompressing the pulse via CPA. Dispersing optical devices are widely available at 1 μm but do not exist at longer wavelength and therefore will require specific developments. Conventional gratings, chirped volume Bragg gratings, prisms, or bulk materials are expected to offer workable solutions.

Coupling power into the structure efficiently is by far the biggest challenge to be addressed. Care must therefore be taken in the coupler design to avoid impedance mismatch, which would lead to localized regions of intense electric field. Initial results in simulating such couplers for the woodpile structure using SOI waveguides indicate coupling efficiencies from the input waveguide to the accelerating mode close to 100% (Wu *et al.*, 2011).

3. Potential breakthroughs in laser technologies

Fiber lasers have had the fortunate advantage of constantly witnessing revolutionary technology developments. However, other types of laser systems have also experienced some promising developments in recent years. The development of ceramic gain media has resulted in a significant decrease in cost, compared to the traditional crystal gain media (Kumar *et al.*, 2004). Material engineering to increase the thermal conductivity of ceramic gain media is a foreseeable breakthrough that would result in a lower price per watt of laser power. Similarly, material engineering to increase the doping levels of ceramics would also be an avenue to high laser power.

TABLE VI. Laser requirements for four DLA structures, with goal parameters for a future linear collider.

Requirement	Woodpile	Fiber	Grating	MAP
Pulse energy	200 nJ	2 μJ	10 μJ	1–10 μJ
Average power	200 W	1 kW	10 kW	1 kW
Wavelength	> 2 μm	> 1 μm	> 1 μm	> 1 μm
Pulse length	1 ps	1 ps	0.1–0.2 ps	1.8–10 ps
CEP locking	$< 1°$	$< 1°$	$< 1°$	$< 1°$
Repetition rate (MHz)	100–1000	100–1000	100–1000	100–1000
Wall-plug efficiency	30%–40%	30%–40%	30%–40%	30%–40%

For DLA applications, the development of a longer-wavelength fiber laser source would be a major breakthrough. While efficient high-power fiber lasers at 1 μm have already been well developed, the practical limitation of nanofabrication (as well as laser damage considerations for silicon) would dictate a preference to operate at a wavelength longer than 1.5 μm. Alternatively, improvements in lithography techniques (and material choices) would make current 1 μm fiber lasers a viable source and drastically change the focus of the laser development. Additionally, the current scaling limit for single aperture cw fiber lasers is expected to be < 36 kW (Dawson *et al.*, 2008). However, this limit may be exceeded, for example, by changing the fiber geometry from round to rectangular (Drachenberg *et al.*, 2013). A rectangular fiber geometry with a propagating higher-order mode may also be inherently compatible with some DLA structures such as the grating scheme, where the mode emitted by the fiber could have the desired periodic π phase changes.

4. Outlook

The laser requirements for DLA are challenging, but are believed to be attainable without the need for revolutionary advances beyond current state of the art. The main areas of development needed to achieve the laser requirements specified in Table VI are timing accuracy and distribution (combined with phase sensing and feedback at the point light is coupled to the electron beam), power scaling of longer-wavelength fiber lasers, and beam transport and coupling to the accelerator structure. Pulse energies, pulse widths, and repetition rates for the DLA applications are well within what has already been demonstrated to date by fiber laser technology. Further, fiber laser technology offers a compact, robust form factor that is naturally compatible with the demanding reliability requirements for an accelerator facility. The next efforts in fiber laser technology for DLA-based systems are needed in 2 μm systems, timing control, laser beam transport, average power scaling, and power coupling to accelerator structures.

VI. SOME APPLICATIONS OF COMPACT ACCELERATORS

A. High-energy collider

To reach 10 TeV center-of-mass energies, a next generation lepton collider based on traditional rf microwave technology

would need to be over 100 km in length and would likely cost tens of billions of dollars to build. Because of the inverse scaling of the interaction cross section with energy, the required luminosity for such a machine would be as much as 100 times greater than proposed 1–3 TeV machines [International Linear Collider (ILC) and CLIC], producing a luminosity goal of the order of 10^{36} cm^{-2} s^{-1}. By attempting to meet these requirements in a smaller cost and size footprint using advanced acceleration schemes, the increased beam energy spread from radiative loss during beam-beam interaction (beamstrahlung) at the interaction point becomes a pressing concern. Since the beamstrahlung parameter is proportional to bunch charge, a straightforward approach to reducing it is to use small bunch charges, with the resulting quadratic decrease in luminosity compensated by higher repetition rates. This is the natural operating regime of the DLA scheme, with the requisite average laser power (> 100 MW) and high (> 10 MHz) repetition rates to be provided by modern fiber lasers.

In Table VII, we compare strawman parameters for a DLA-based collider and Higgs factory with expected parameters for CLIC at 3 TeV. In these examples, DLA meets the desired luminosity and with a significantly smaller beamstrahlung energy loss. Other advanced collider schemes such as beam-driven plasma and terahertz also rely upon a traditional pulse format for the electron or positron beam and would therefore compare similarly to CLIC in this regard. Although the numbers in Table VII are merely projections used for illustrative purposes, they highlight the fact that due to its unique operating regime DLA is poised as a promising technology for future collider applications. Calculations in Table VII for the DLA cases assume the full bunch train charge per crossing. This assumption is valid for the geometrical luminosity if the beta function (wavelength of beam envelope oscillation$/2\pi$) at the interaction point is larger than the bunch train length (true for most configurations of interest). It is valid for the beamstrahlung and luminosity enhancement (here a factor of 6) only if the microbunching is undone after acceleration but prior to the IP, such as by introducing a dispersive section to smear out the longitudinal modulation. The final focus technology matched to the DLA scale is a separate topic requiring development, but the DLA quadrupole focusing technology under development (Sec. III.E) is a possible path.

TABLE VII. Strawman parameters for the DLA Linear Collider.

Parameter	Units	CLIC	DLA 3 TeV	DLA 250 GeV
Center-of-mass energy	GeV	3000	3000	250
Bunch charge	e	3.7×10^9	30 000	38 000
Bunches per train		312	159	159
Train repetition rate	MHz	5.0×10^{-5}	20	60
Bunch train length	ps	26 005	1.0	1.0
Single bunch length	μm	34.7	0.0028	0.0028
Design wavelength	μm	230 609	2.0	2.0
Invariant X emittance	μm	0.66	0.0001	0.002
Invariant Y emittance	μm	0.02	0.0001	0.002
IP X spot size	nm	45	1	2
IP Y spot size	nm	1	1	2
Beamstrahlung energy loss	%	28.1	1.0	0.6
Enhanced luminosity/top 1%	cm^{-2}/s	2.0×10^{34}	3.2×10^{34}	1.3×10^{34}
Beam power	MW	14.1	22.9	7.3
Wall-plug efficiency	%	4.8	12.2	9.5
Wall-plug power	MW	582	374	152
Gradient	MV/m	100	1000	1000
Total linac length	km	42.0	3.0	0.3

Obtaining wall-plug efficiencies suitable for linear collider applications will additionally require the development of integrated couplers with high efficiency, fed by a network of waveguides that split the laser power from a common feed among various accelerator components. Methods for achieving efficient coupling were discussed in Sec. III.F. The power distribution scheme is then envisioned as a coupler that brings a pulse from an external fiber laser onto the integrated chip, distributes it between multiple structures via guided wave power splitters, and then recombines the spent laser pulse and extracts it from the chip via a mirror-image output coupler (Colby, England, and Noble, 2011), after which the power is either dumped or, for optimal efficiency, recycled. For the DLA calculations in Table VII, accelerator to electron coupling efficiencies of 40% are assumed, consistent with the estimates of Siemann (2004). Assumed laser wall-plug efficiency is 40%, feasible using current solid-state thulium-doped fiber laser technology, which has already achieved wall-plug efficiencies above 30% (Moulton *et al.*, 2009). Maintaining phase synchronicity of the laser pulse and the accelerated electrons between many separately fed structures would be accomplished by fabricating the requisite phase delays into the lengths of the waveguide feeds. Prototype grating to SOI couplers has recently been produced for the 1.5 μm wavelength operation.

B. X-ray light sources

1. Operating regime of DLA x-ray FEL

During the initial development of the DLA, it was already recognized (Huang and Byer, 1997) that an x-ray free-electron laser (XFEL) driven by a GeV DLA could generate comparable or higher output powers than does the Linac Coherent Light Source (LCLS), while having a size 1–2 orders of magnitude smaller than the LCLS. As the development of a DLA has progressed rapidly in the past 15 years, tabletop and even chip-scale XFEL driven by MeV DLA are emerging from the horizon. For what follows, we consider only single-pass XFEL generating coherent radiation from self-amplified spontaneous emission (SASE).

In a SASE FEL, the initial spontaneous radiation from the electrons is fed back to bunch the electrons at the resonant wavelength. The bunching process builds up coherent radiation power with an exponential gain. In an XFEL, when the emitted photon energy is a significant fraction of the electron energy, quantum recoil of electrons can introduce electron energy spread and push the electrons outside the FEL gain bandwidth. A DLA-driven XFEL could fall into the undesirable quantum regime, if a low-energy beam and a short undulator period are chosen to achieve compactness for a radiation device. The crucial parameter that distinguishes the

classical and quantum regimes for an FEL is the so-called quantum fundamental FEL parameter (Bonifacio *et al.*, 2008)

$$\bar{\rho} = \gamma mc\rho/\hbar k, \tag{73}$$

where $\hbar k$ is the x-ray photon momentum, γ is the electron energy normalized to the electron rest mass energy, m is the electron rest mass, and c is the vacuum speed of light. In Eq. (73), ρ (without overbar) is the fundamental FEL parameter or the energy spread of the electrons at saturation, given by

$$\rho = \frac{1}{2\gamma}\left[\frac{I}{I_A}\left(\frac{\lambda_u A_u}{2\pi\sigma_r}\right)^2\right]^{1/3}, \tag{74}$$

where I is the peak beam current, $I_A = 17$ kA is the Alfven current, λ_u is the undulator period, and σ_r is the beam radius. Here $A_u = a_u$ for a helical undulator and $a_u[J_0(\xi) - J_1(\xi)]$ for a planar undulator, with $\xi = a_u^2/2(1 + a_u^2)$ and $a_u = \gamma\psi_0$ being the undulator parameter, where ψ_0 is the maximum angular deflection in the undulator.

In the quantum FEL regime, $\bar{\rho} < 1$ and the quantum recoil induced energy spread is larger than the energy acceptance bandwidth ρ. Less than one x-ray photon is generated from an electron at saturation. The major disadvantages for a quantum XFEL include a reduced power growth rate and a small output photon flux at power saturation.

2. Laser undulator

A DLA is built upon the semiconductor microfabrication technology, which allows mass production through lithographic patterning. It is highly desirable to integrate a chip-scale XFEL undulator to a DLA in a fabrication process. A possible and convenient approach is to design an undulator structure utilizing the drive-laser field as the undulator field.

A laser itself in vacuum is an undulator for Thomson scattering, having an undulator period comparable to the laser wavelength. For example, to generate hard x rays, Thomson backscattering needs only an MeV electron beam to collide with a 1 μm laser pulse, although the quantum effects cannot be ignored for a stimulated Thomson scattering process using such a low-energy beam. A recent study shows a design for a centimeter long, 300-kW Thomson XFEL at 20 Å driven by a 6 MeV beam with 10% quantum-reduced power growth rate (Sprangle, Hafizi, and Penano, 2009). However, the 10^{-4} energy spread required for the MeV-beam Thomson XFEL imposes a challenge to a DLA.

To avoid the FEL gain reduction in the quantum regime, one approach is to increase the drive beam energy. The radiation wavelength of an FEL satisfies the well-known FEL synchronism condition, given by

$$\lambda_r = \lambda_u(1 + a_u^2)/2\gamma^2, \tag{75}$$

where λ_r is the radiation wavelength, and $a_u = 0.093\lambda_u(\mathrm{cm})B_{u,\mathrm{rms}}(\mathrm{kG})$ with $B_{u,\mathrm{rms}}$ being the rms undulator field. A higher energy beam means a longer undulator period. Significant progress has been made toward the development of a long-period laser undulator that can be microfabricated on the same DLA platform. Figure 39(a) shows a laser-pumped slanted grating structure that may serve as a long-period laser undulator and at the same time a particle accelerator (Plettner and Byer, 2008). The deflecting force to the electrons is derived from summing the transverse fields of the space harmonics in the periodic structure with an optimized tilt angle α for the grating or for the coupling laser wave front. The undulator period is the phase-flip distance between the electron and the optical field, given by

$$\lambda_u = \lambda/|\beta_e^{-1} - \beta_p^{-1}|, \tag{76}$$

where β_e is the electron velocity, and β_p is the effective axial phase velocity of the mode field in the structure, both normalized to the vacuum speed of light. Since the undulator

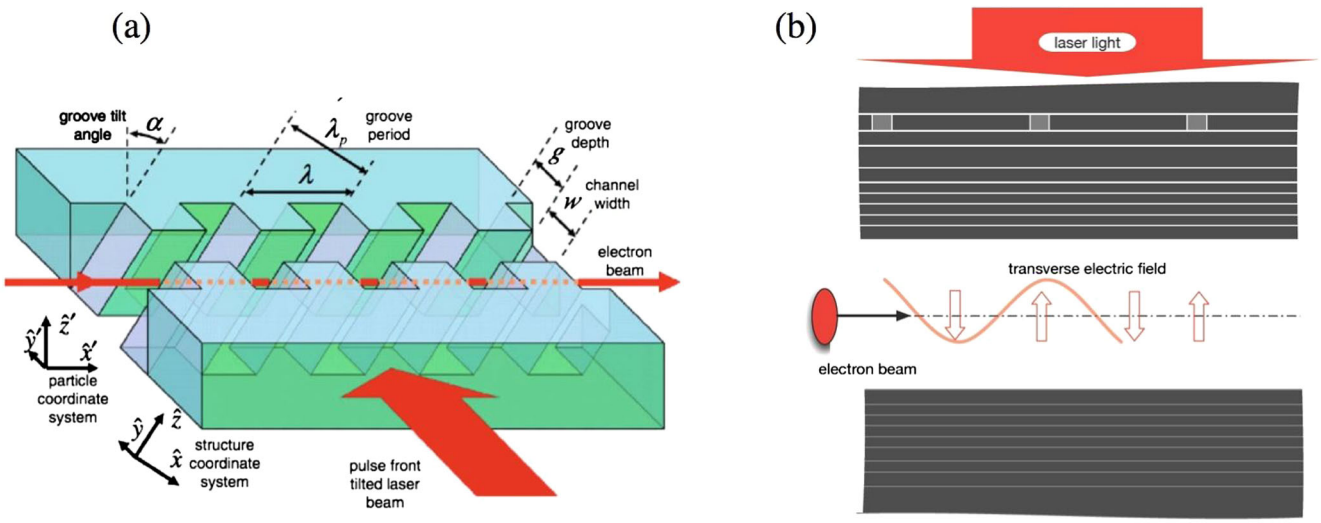

FIG. 39 (color online). A laser undulator employing a laser side-coupled (a) slanted grating (Plettner and Byer, 2008) or (b) Bragg waveguide to derive the undulator deflecting force with an undulator period much longer than the laser wavelength. From Travish and Yoder, 2011.

is assembled from cascaded grating sections with individual optical phase control, the effective phase velocity β_p of the mode field can be manually tuned to achieve an undulator period much longer than the laser wavelength λ.

Figure 39(b) shows another laser undulator based on the same phase-flip concept to derive a long undulator period from a short drive laser wavelength (Travish and Yoder, 2011). This laser undulator has a smooth electron channel in a Bragg waveguide. The laser is side coupled into the structure with external phase and polarization control. The smooth wall has less concern of wakefield generation from and radiation loss to an electron beam. This laser undulator can be integrated with the similarly designed resonant-type DLA in the so-called MAP.

It should be pointed out that all the laser undulators require a uniform laser field over some temporal duration to provide a useful length of the undulator. From the FEL synchronism condition (75), it is straightforward to show that the required laser-field uniformity has to be better than the saturation energy spread of a FEL or $\Delta E/E < \rho$, where E is the laser envelope field of a laser undulator. Usually ρ is about 10^{-3} for an XFEL. The line-by-line laser pulse shaping technique developed recently (Jiang *et al.*, 2007) might be helpful in the future to synthesize a uniform enough laser-field envelope containing an optical carrier for the laser undulator.

3. Microchip-scale XFEL

With a dielectric material in a laser undulator, the undulator field is limited to the damage field of about 1 GeV/m for 100-fs laser pulses (Stuart *et al.*, 1995). At the material damage threshold, the equivalent peak undulator field is 3.3 T. With a known undulator field, the undulator period can be calculated from the FEL synchronism condition (75) for a given beam energy and radiation wavelength. Figure 40(a) shows the undulator period versus beam relativistic energy factor γ for $\lambda_r = 0.1$ and 1 nm, indicating an undulator period shorter or comparable to a millimeter for an XFEL driven by an electron beam below 0.5 GeV energy. By using Eqs. (73) and (74), and assuming a beam radius of $\sigma_r = 100$ nm, we plot in Figs. 40(b) and 40(c) the operation regimes of an XFEL on the I-γ plane for $\lambda_r = 0.1$ and 1 nm, respectively. The colored areas correspond to different radiation regimes extending from $\bar{\rho} < 1$ (quantum regime) to $\bar{\rho} > 100$ (classical regime). To have negligible quantum effects, the XFEL is preferred to operate in the classical regions of the plot. Overlapped on the plots are constant-gain-length curves of the XFEL calculated from the 1D SASE FEL model (Bonifacio, Pellegrini, and Narducci, 1984; Kim, 1986), $L_g = \lambda_u/4\pi\sqrt{3}\rho$, with their values marked on the curves in units of mm.

Although the actual performance of an SASE FEL is also influenced by the thermal spread (emittance and energy spread) and focusing of the beam, in view of the superior beam quality of a DLA in Sec. IV.A, Fig. 40 can serve as a first-order estimate for the performance of a DLA-driven XFEL adopting a dielectric laser undulator. For example, to saturate a hard x-ray FEL ($\lambda_r \sim 0.1$ nm) in a 5–10-cm-long dielectric laser undulator, it is desirable to have a beam energy above 300 MeV, an undulator period of about 100 μm [see bottom curve in Fig. 40(a) at $\gamma \sim 600$], and a beam current of

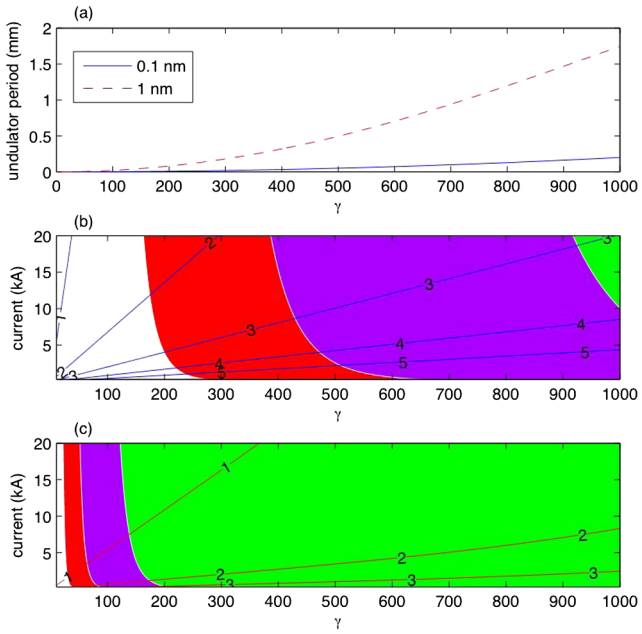

FIG. 40 (color online). (a) Undulator period for 1 nm (dashed) and 0.1 nm (solid) radiation wavelengths vs electron energy subject to a 1-GV/m laser damage field on a dielectric laser undulator. Quantum (white) and classical (colored) regimes on the current vs gamma plane for an XFEL radiating at (b) 0.1 nm and (c) 1 nm. Constant-gain-length curves are also shown with their values marked in units of millimeters (leftmost section: $\bar{\rho} < 1$, second from left section: $1 < \bar{\rho} < 10$, third from left section: $10 < \bar{\rho} < 100$, rightmost section: $\bar{\rho} > 100$).

10–20 kA [shaded area in Fig. 40(b) with gain length about 3 mm]. For soft x-ray FEL, one can avoid the disadvantageous quantum regime as long as the beam energy is larger than a few tens of MeV. By using a 50 MeV, 5 kA DLA beam in a centimeter long, 20 μm period laser undulator (gain length $L_g = 1$ mm), a soft x-ray FEL can operate in the classical regime ($\bar{\rho} \simeq 37$, the shaded area near $\gamma = 100$ and $L_g = 1$ mm in Fig. 40(c)) with an estimated peak output power of 22.5 MW ($\rho = 9 \times 10^{-4}$) at 1 nm wavelength. The average x-ray power from the soft XFEL can be in the milliwatt range with a MHz electron pulse rate from the DLA. Therefore, a dielectric laser undulator driven by a DLA is promising to realize a portable or even hand-held XFEL in the future.

C. Medical devices

1. Current health care uses of accelerators and beams

The use of particle beams and x rays has a rich history in modern medicine (Thwaites and Tuohy, 2006). The discovery of x rays a century ago was immediately recognized as having biological implications for both imaging and treatment, often with dubious claims of efficacy. As evidence-based treatments prevailed over speculative ones, the uses of x rays narrowed. The advent of particle accelerators opened new areas for the responsible use of radiation (Ginzton and Nunan, 1985). The primary health care uses of particle beams and relativistic beam-generated x rays are currently (1) oncology therapy, (2) advanced imaging (nonclinical), (3) isotope production, and (4) sterilization (nonclinical) (Jongen, 1997).

In radiology, conventional imaging is widely employed today in the clinic using tube-based x-ray sources. In oncology, while relativistic electron beams are perhaps the most prevalent, beams of protons, antiprotons, ions (e.g., helium, carbon, oxygen), neutrons, and other particles have all been seriously considered or are in current use. Linacs are most commonly found in oncology and experimental imaging, whereas cyclotrons are most often employed for isotope production (Schlyer, 2005). For sterilization, which requires only cost-effective dose delivery, all manner of accelerators have been considered and several variants have been deployed. Because of the broad history of electron beams for health care and the ongoing consideration of the efficacy of protons and ions, we restrict the remainder of this section to relativistic electron beam accelerators in clinical settings.

2. Challenges and needs

Research uses of accelerators span several orders of magnitude in size, cost, beam energy, and number of users. However, in clinical settings, advanced medical treatments tend to have a narrow range of acceptable size (footprint), capital cost, cost per treatment, and number of patients that can be treated. For instance, a device might have to fit in a clinic basement room, at a total capital expense of 3 to 5 million dollars, require less than 1000 dollars operating expense to provide a fractionated treatment, and handle one patient per hour. There are exceptional treatment centers that provide unique capabilities and can be well outside the nominal range of size and cost, but the majority of hospitals and clinics operate under space and funding restrictions which do not allow for exotic facilities. There are several factors that determine the acceptable size and cost (e.g., population density and total population served, wealth of patients or donors, etc.), but in general hospital administrators are sensitive to these practical considerations.

One of the primary reasons accelerator technology finds slow and limited penetration into health care is that the actual accelerator is a small component of the complete system and total costs. For instance, a shielded treatment room with all the requisite safety considerations retrofitted into a hospital basement can often cost more to build than the machine it is intended to house. The staffing costs of a treatment center typically greatly surpass the operating costs of the treatment machine. Nevertheless, the demand for hardware remains strong: in the oncology sector alone, the worldwide market for therapy machines was already 2 billion dollars in 2005, with about 700 machines sold annually. Imaging represents a multibillion dollar growing market worldwide. However, given the investment in existing infrastructure, it is not surprising that new accelerator technologies often have to demonstrate vastly superior characteristics before finding their way into medical devices. Before we consider the potential for DLAs to meet this threshold for inclusion in health care, we briefly examine the current challenges in imaging and radiation therapy.

a. Imaging

X-ray production for medical imaging has a rich history beginning over a century ago when Wilhelm Röntgen used various vacuum tubes to produce short-wavelength radiation. While significant engineering improvements have been made (e.g., the rotating anode), the fundamentals of these x-ray sources remain the same. An electron current from a cathode strikes an anode, which produces x rays through bremsstrahlung. These sources, despite a variety of improvements, produce broadband, wide-angle output, which is generally considered low brightness. Nevertheless, these tubes are the workhorses of medical imaging, including the expanding area of computed x-ray tomography (CT).

At the other end of the technological complexity spectrum are light sources based on synchrotrons and undulator devices. These sources provide x rays several orders of magnitude higher in brightness than tubes and can serve dozens of users simultaneously. Their relatively high cost (about a billion dollars), large size (about 1 km in diameter), and operational complexity make them impractical for local clinical use. The advent of so-called fourth-generation light sources, based on free-electron lasers, has pushed the peak brightness up by several more orders of magnitude, but does little to address clinical needs such as machine size, capital cost, and ease of use.

Despite the widespread and successful use of x-ray tubes, it is recognized that these devices produce output far from ideal: Bremstrauhlung spectra contain more absorbed low-energy photons (even after filtering) than the useful high-energy photons, the cone-beam distribution produces geometric artifacts (from cone-beam projections) and constrains the imaging geometry, and the use of polychromatic absorption imaging for tissue means that only low contrast can be obtained (e.g., 2:1).

b. Oncology

Roughly half of all cancer patients in the U.S. are treated with radiation, typically in electron or x-ray accelerator facilities, of which there are several thousand currently in service (Mackay, Jemal, and Lee, 2006). Existing radiation oncology treatments rely largely on external radiation sources, although brachytherapy (introduction of controlled radioactive sources directly into the body) has obvious advantages for preserving adjacent tissues and limiting damage to surrounding organs. Several commercial brachytherapy products exist, but the majority incorporate naturally radioactive materials that cannot be turned off and on, are complicated and invasive to use, and must be well shielded and controlled for safe operation.

Typical accelerator-based radiotherapy devices operate around 6–12 MeV electron energy, with energies as high as 25 MeV used in some cases. Direct electron irradiation is used only for skin or superficial treatment and in the past was used for interoperative treatment. The irradiation volume and dose is tightly controlled and exposure is typically from multiple directions to avoid damage to critical organs and structures. Doses below 2 Gy per treatment are typically used. A series of technological improvements from intensity modulated radiation therapy to image guided radiation therapy have been employed to improve efficacy and reduce morbidity (Verellen, Ridder, and Storme, 2008). Damage to healthy tissue, patient and staff safety, and high capital and operating costs are all shortcomings of existing external beam radiation therapy (EBRT).

3. Opportunities for DLA devices

DLAs come in many configurations. For the health care applications, we focus on a self-contained source of relativistic electron beams. For most applications, these electrons are used to produce intense x rays. The two areas of need seeming to be particularly suitable to future DLA-based systems are oncology and imaging.

a. Minimally invasive and robotic therapy

A self-contained DLA-based device, as a medical radiotherapy device, offers consumers and facilities new options for minimally invasive cancer treatments with increased flexibility of use. An encapsulated DLA can be placed within a tumor site using standard endoscopic methods, allowing it to deliver the same radiation dose provided by current EBRT, without damage to surrounding tissue. The manufacturing and operating costs of a radiation treatment machine based on disposable DLA accelerators are anticipated to be much lower than those for conventional EBRT machines.

As suggested in Fig. 41, the DLA module can be built onto the end of a fiber-optic catheter and attached to an endoscope, allowing a doctor to deliver electronically controlled, high-energy radiation directly to organs, tumors, or blood vessels within the body. Electrons with 1–3 MeV have a range of about a centimeter, allowing for irradiation volumes to be tightly controlled. This approach has the potential to provide a more effective and more affordable radiation treatment with fewer side effects and better patient quality of life than current practice. Radiation energy and dose is directly controlled by the laser source, which enables the surgeon to treat a wide range of tumor morphologies. Such treatment is ideally suited to the precise and remote control afforded by robotic surgery systems.

b. Advanced imaging

A high-brilliance, compact x-ray source has been the holy grail of radiation production for decades. In medical imaging, a monochromatic and directional source would allow a move from low-contrast absorption imaging toward phase-contrast (Arfelli *et al.*, 1998) and possibly diffraction-based imaging (Chapman *et al.*, 1997). Efforts to date have included synchrotron-based light sources (Arfelli, 2000; Suortti and Thomlinson, 2003), inverse Compton scattering (ICS) sources (Bech *et al.*, 2009; Yamada *et al.*, 2009), and FEL concepts (Karagodsky, Schieber, and Schachter, 2010).

Compactness alone is not a compelling enough benefit of DLA-based sources for medical imaging. Instead, the combination of small size with high beam brightness may enable new methods to deliver x rays. For the energies of use in medical imaging (e.g., 20 to 120 keV for planar radiology and up to 240 keV for CT), only ICS-based schemes produce sufficiently hard x rays with reasonably low (e.g., 30 MeV) electron beams. Implementing an ICS-based system using DLA techniques does not necessarily deliver a performance advantage, compared to a conventional accelerator approach. However, the repetition rate and bunch format inherent to DLAs may be better suited to the matching laser of the ICS. A system based on low charge and a MHz or even GHz repetition rate would provide a near-cw flux of high-brightness, narrow-band x rays suitable for phase-contrast imaging in a clinic-sized system (Hoheisel, 2006).

Finally, work has been done to use MeV photons for medical imaging, such as might be produced by a few MeV electron beam with bremsstrahlung. However, while for use in oncology, this approach is a strong departure from conventional radiology and requires extensive study to understand the benefits and safety concerns.

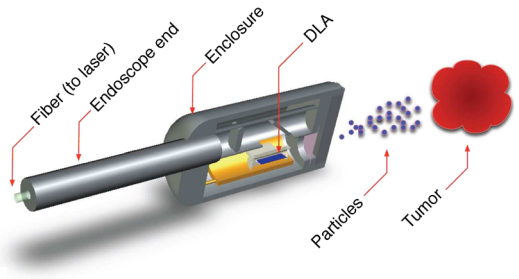

FIG. 41 (color online). Schematic of a self-contained DLA encapsulated as a fiber endoscope.

4. Outlook

There are physics, engineering, and practicality considerations that must be overcome in order for dielectric laser accelerators to find a role in health care. The physics considerations include the bunch time scale, repetition rate, and spot size. These DLA parameters tend to be in regimes rather different from those found in conventional accelerators and it is as yet unclear what benefits can be obtained for medical applications. For instance, the medical applications discussed here do not take advantage of the time structure of the DLA-produced bunches (e.g., subfemtosecond). Engineering efforts are necessary to produce a fully integrated system, at a reasonable cost, one that complies with the performance demand of invasive medical devices (Cherry and Seshadri, 2000), and one whose damage mechanisms are well understood. The operational reliability and performance degradation from radiation damage may require the use of disposable devices (Du *et al.*, 1994; Goulding *et al.*, 1996). Finally, practical considerations include the cost, the need for long efficacy trials, and slow adoption rates. In order for a new class of devices to be accepted by the medical community, not only do efficacy and cost savings need to be shown, but the benefits must be overwhelming compared to existing solutions in order to overcome institutional inertia and existing infrastructure investment. Significant work remains to identify these benefits and to demonstrate the feasibility of a DLA-based medical device.

VII. CONCLUSIONS

The DLA concept leverages well-established industrial fabrication capabilities and the commercial availability of tabletop lasers to reduce cost, while offering significantly higher accelerating gradients, and therefore a smaller footprint. Power estimates for the DLA scenario are comparable with conventional rf technology, assuming that similar power efficiency (near 100%) for guided wave systems can be achieved, 40% wall-plug laser efficiencies (feasible with solid-state thulium fiber laser systems), and 40% laser to electron beam

coupling (consistent with published calculations). This research has significant near and long-term applications beyond energy frontier science, including radiation production for compact medical x-ray sources and university-scale free-electron lasers. These additional applications are beginning to be explored. Since DLA structures operate optimally with optical-scale electron bunch formats, high repetition rate (10 s of MHz) attosecond scale pulses are a natural combination. The inherently ultrafast aspect of DLAs is relevant, given that many of the chemical, biological, and material processes that can be probed using x-ray light sources involve subnanometer systems moving on femtosecond time scales.

Dielectric laser acceleration is a multidisciplinary field, drawing upon expertise in infrared laser technology, materials science, beam dynamics, semiconductor fabrication techniques, and accelerator physics. Progress toward an energy scalable DLA architecture requires a research and design focus on fabrication and structure evaluation to optimize existing and proposed concepts, and development of low-charge high-repetition rate electron sources that can be used to evaluate performance over many stages of acceleration. Following initial proof-of-principle demonstrations of dielectric laser acceleration, which are taking place now, the challenge will be to develop this technique into a useful acceleration method. Among the issues that need to be resolved are understanding IR laser damage limits of semiconductor materials at picosecond pulse lengths, development of high (near 100%) efficiency schemes for coupling fiber or free-space lasers into DLA structures, development of low-velocity accelerators for injectors into the high-energy structures, developing integrated designs with multiple stages of acceleration, and understanding phase stability issues related to temperature and nonlinear high-field effects in dielectrics. Answering these questions requires strong collaborations between university and national laboratory groups and industry partners.

ACKNOWLEDGMENTS

We thank Robert H. Siemann, our late colleague and friend, for his many years of dedication and pioneering contributions to this work. For many useful discussions, we thank James Harris, Rasmus Ischebeck, Boris Kuhlmey, Chunghun Lee, Kenneth Leedle, Minghao Qi, Brian Schwartz, Olav Solgaard, and Dieter Walz. Key work in the field of DLA accelerators originated with contributions from our colleagues Eric Colby, Norman Kroll, Xintian Eddie Lin, and Tomas Plettner, and many valuable discussions with them occurred during the years. This work was supported by the U.S. Department of Energy Grants No. DE-AC02-76SF00515, No. DE-FG06-97ER41276, No. DE-FG02-13ER41964, No. DE-FG-98ER45693, No. DE-FG-02-04ER41317, No. DE-FC02-07ER41499, No. DE-SC0008920, No. DE-SC0000839, No. DE-SC0007579, and No. DE-FG02-13ER41970, the Office of Naval Research Contract No. ONR N00014-06-1-0925, DARPA Contracts No. N66001-11-1-4197 and No. N66001-11-1-4199, the U.S. Defense Threat Reduction Agency (DTRA) Grant No. HDTRA1-09-1-0043, the DFG Cluster of Excellence Munich Center for Advanced Photonics, and the National Science Council (Taiwan) Grant No. NSC-99-2112-M-007-013-MY3.

REFERENCES

Agrawal, G. P., 1989, *Nonlinear Fiber Optics* (Academic, New York).
Alfano, R. R., and S. L. Shapiro, 1970a, Phys. Rev. Lett. **24**, 584.
Alfano, R. R., and S. L. Shapiro, 1970b, Phys. Rev. Lett. **24**, 592.
Alfano, R. R., and S. L. Shapiro, 1970c, Phys. Rev. Lett. **24**, 1217.
Andonian, G., *et al.*, 2012, Phys. Rev. Lett. **108**, 244801.
Archambault, J.-L., R. J. Black, S. Lacroix, and J. Bures, 1993, J. Lightwave Technol. **11**, 416.
Arfelli, F., 2000, Nucl. Instrum. Methods Phys. Res., Sect. A **454**, 11.
Arfelli, F., *et al.*, 1998, Phys. Med. Biol. **43**, 2845.
Augst, S. J., J. K. Ranka, T. Y. Fan, and A. Sanchez, 2007, J. Opt. Soc. Am. B **24**, 1707.
Baccaro, S., F. DeMartini, and A. Ghigo, 1982, Opt. Lett. **7**, 174.
Baer, C. R. E., C. Krankel, C. J. Saraceno, O. H. Heckl, M. Golling, R. Peters, K. Petermann, T. Sudmeyer, G. Huber, and U. Keller, 2010, Opt. Lett. **35**, 2302.
Bane, K. L. F., and G. Stupakov, 2003, Phys. Rev. ST Accel. Beams **6**, 024401.
Barwick, B., C. Corder, J. Strohaber, N. Chandler-Smith, C. Uiterwaal, and H. Batelaan, 2007, New J. Phys. **9**, 142.
Bass, M., E. W. V. Stryland, D. R. Williams, and W. L. Wolfe, 1995, *Handbook of Optics*, Vol. II (McGraw-Hill, New York).
Bauer, C. A., G. R. Werner, and J. R. Cary, 2008, J. Appl. Phys. **104**, 053107.
Bech, M., O. Bunk, C. David, R. Ruth, J. Rifkin, R. Loewen, R. Feidenhans'l, and F. Pfeiffer, 2009, J. Synchrotron Rad. **16**, 43.
Bergher, M., 1997, Nucl. Instrum. Methods Phys. Res., Sect. A **395**, 259.
Bermel, P., R. L. Byer, E. R. Colby, B. M. Cowan, J. Dawson, R. J. England, R. J. Noble, M. Qi, and R. B. Yoder, 2014, Nucl. Instrum. Methods Phys. Res., Sect. A **734**, 51.
Blin, S., l. Provino, N. Traynor, A. Mugnier, D. Pureur, and T. Chartier, 2009, in *Conference on Lasers and Electro-optics/ Europe, Munich* (IEEE, Piscataway, NJ), p. CJ1.5.
Bonifacio, R., C. Pellegrini, and L. Narducci, 1984, Opt. Commun. **50**, 373.
Bonifacio, R., N. Piovella, M. M. Cola, L. Volpe, A. Schiavi, and G. R. M. Robb, 2008, Nucl. Instrum. Methods Phys. Res., Sect. A **593**, 69.
Bor, Z., and B. Racz, 1985, Opt. Commun. **54**, 165.
Boussoukaya, M., H. Bergeret, R. Chehab, B. Leblonde, and J. LeDuff, 1989, Nucl. Instrum. Methods Phys. Res., Sect. A **279**, 405.
Brau, C., 1990, *Free Electron Lasers* (Academic Press, Boston, MA).
Breuer, J., and P. Hommelhoff, 2012, in *Advanced Accelerator Concepts, 15th Workshop, 2012*, AIP Conf. Proc. Vol. 1507, edited by R. Zgadzaj (American Institute of Physics, New York), pp. 431–434.
Breuer, J., and P. Hommelhoff, 2013, Phys. Rev. Lett. **111**, 134803.
Breuer, J., and P. Hommelhoff, 2014, Phys. Rev. ST Accel. Beams **17**, 021301.
Chang, C. M., and O. Solgaard, 2013a, in *Proceedings of the Conference on Lasers and Electro-Optics (CLEO)* (IEEE, Piscataway, NJ), p. JTu4A.52.
Chang, C. M., and O. Solgaard, 2013b, Opt. Express **21**, 27 209.
Chang, C. M., and O. Solgaard, 2013c, IEEE Micro **33**, 32.
Chang, C. M., and O. Solgaard, 2014, Appl. Phys. Lett. **104**, 184102.
Chang, G., M. Rever, V. Smirnov, L. Glebov, and A. Galvanauskas, 2009, Opt. Lett. **34**, 19.
Chao, A. W., 1993, *Physics of Collective Beam Instabilities in High Energy Accelerators* (Wiley, New York).

Chao, A. W., and M. Tigner, 2006, *Handbook of Accelerator Physics and Engineering* (World Scientific, Singapore).
Chapman, D., W. Thomlinson, R. E. Johnston, D. Washburn, E. Pisano, N. Gmür, Z. Zhong, R. Menk, F. Arfelli, and D. Sayers, 1997, Phys. Med. Biol. **42**, 2015.
Cheng, M. Y., Y. C. Chang, A. Galvanauskas, P. Mamidpudi, R. Changkakoti, and P. Gatchell, 2005, Opt. Lett. **30**, 358.
Cherry, J., and S. Seshadri, 2000, Radiology Management **November/December**, 42.
Chojnacki, E., R. Konecny, M. Rosing, and J. Simpson, 1993, in *Proceedings of the 1993 Particle Accelerator Conference, Washington, DC* (IEEE, New York), p. 815.
Colby, E. R., R. J. England, and R. J. Noble, 2011, in "2011 Particle Accelerator Conference Proceedings, New York, NY" (IEEE), p. 262.
Cornacchia, M., and P. Emma, 2002, Phys. Rev. ST Accel. Beams **5**, 084001.
Couny, F., F. Benabid, P. Roberts, M. Burnett, and S. Maier, 2007, Opt. Express **15**, 325.
Cowan, B., M. C. Lin, B. T. Schwartz, E. R. Colby, R. J. England, R. J. Noble, J. E. Spencer, R. L. Byer, and C. M. McGuinness, 2010, in "Proceedings of the 2010 International Particle Accelerator Conference (IPAC2010), Kyoto, Japan" (ACFA), p. 4077.
Cowan, B. M., 2003, Phys. Rev. ST Accel. Beams **6**, 101301.
Cowan, B. M., 2008, Phys. Rev. ST Accel. Beams **11**, 011301.
Dawson, J. W., M. J. Messerly, R. J. Beach, M. Y. Shverdin, E. A. Stappaerts, A. K. Sridharan, P. H. Pax, J. E. Heebner, C. W. Siders, and C. Barty, 2008, Opt. Express **16**, 13 240.
Dawson, J. W., M. J. Messerly, H. H. Phan, J. K. Crane, R. J. Beach, C. W. Siders, and C. P. J. Barty, 2009, IEEE J. Sel. Top. Quantum Electron. **15**, 207.
DeJonge, N., Y. Lamy, K. Schoots, and T. H. Oosterkamp, 2002, Nature (London) **420**, 393.
DiGennaro, E., C. Zannini, S. Savo, A. Andreone, M. R. Masullo, G. Castaldi, I. Gallina, and V. Galdi, 2009, New J. Phys. **11**, 113022.
Dillon, T., J. Murakowski, S. Y. Shi, and D. Prather, 2008, Opt. Lett. **33**, 896.
Ding, Y., *et al.*, 2009, Phys. Rev. Lett. **102**, 254801.
DiTeodoro, F., M. J. Hemmat, J. Morais, and E. C. Cheung, 2010, in *Fiber Lasers VII: Technology, Systems, and Applications*, SPIE Proceedings Vol. 7580 (SPIE-International Society for Optical Engineering, Bellingham, WA), p. 758006.
Dowell, D. H., and J. F. Schmerge, 2009, Phys. Rev. ST Accel. Beams **12**, 074201.
Drachenberg, D. R., M. J. Messerly, P. H. Pax, A. Sridharan, J. Tassano, and J. W. Dawson, 2013, Opt. Express **21**, 11 257.
Du, D., X. Liu, G. Korn, J. Squier, and G. Mourou, 1994, Appl. Phys. Lett. **64**, 3071.
Duguay, M., Y. Kokubun, T. Koch, and L. Pfeiffer, 1986, Appl. Phys. Lett. **49**, 13.
Duris, J. P., P. Musumeci, and R. K. Li, 2012, Phys. Rev. ST Accel. Beams **15**, 061301.
Earnshaw, S., 1842, Trans. Cambridge Philos. Soc. **7**, 97 [https://archive.org/details/transactionsofca07camb].
Eggleston, J. M., T. J. Kan, K. Kuhn, J. Unternahrer, and R. L. Byer, 1984, IEEE J. Quantum Electron. **20**, 289.
Eidam, T., S. Hanf, E. Seise, T. V. Andersen, T. Gabler, C. Wirth, T. Schreiber, J. Limpert, and A. Tunnermann, 2010, Opt. Lett. **35**, 94.
Eidam, T., J. Rothhardt, F. Stutzki, F. Jansen, S. Hadrich, H. Carstens, C. Jauregui, J. Limpert, and A. Tunnermann, 2011, Opt. Express **19**, 255.
El-Dahshory, M. A., A. M. Attiya, and E. A. Hashish, 2007, Prog. Electromagn. Res. **74**, 319.
England, R. J., C. K. Ng, R. Noble, J. Spencer, Z. Wu, and D. Xu, 2011, Report No. SLAC-PUB-14440.
Galvanauskas, A., M. E. Fermann, D. Harter, K. Sugden, and I. Bennion, 1995, Appl. Phys. Lett. **66**, 1053.
Ganter, R., R. Bakker, C. Gough, S. C. Leemann, M. Paraliev, M. Pedrozzi, F. LePimpec, V. Schlott, and L. Rivkin, 2008, Phys. Rev. Lett. **100**, 064801.
Gapontsev, V., V. Fomin, and A. Yusim, 2009, in *Proceedings of the 22nd Solid State and Diode Laser Technology Review Technical Digest* (Directed Energy Professional Society, Newton, MA), p. 142.
Garcia, C. H., and C. A. Brau, 2002, Nucl. Instrum. Methods Phys. Res., Sect. A **483**, 273.
Ginzton, E. L., and C. S. Nunan, 1985, Int. J. Radiat. Oncol. Biol. Phys. **11**, 205.
Goodno, G. D., L. D. Book, and J. E. Rothenberg, 2009, Opt. Lett. **34**, 1204.
Goulding, R. H., S. J. Zinkle, D. A. Rasmussen, and R. E. Stoller, 1996, J. Appl. Phys. **79**, 2920.
Grasyuk, A. Z., S. V. Kurbasov, L. L. Losev, A. P. Lutsenko, A. A. Kaminskii, and V. B. Semenov, 1998, Quantum Electron. **28**, 162.
Graves, W. S., F. X. Kartner, D. E. Moncton, and P. Piot, 2012, Phys. Rev. Lett. **108**, 263904.
Halir, R., P. Cheben, J. H. Schmid, R. Ma, D. Bedard, S. Janz, D. X. Xu, A. Densmore, J. Lapointe, and I. Molina-Fernndez, 2010, Opt. Lett. **35**, 3243.
Hasegawa, A., and F. Tappert, 1973a, Appl. Phys. Lett. **23**, 142.
Hasegawa, A., and F. Tappert, 1973b, Appl. Phys. Lett. **23**, 171.
Hauri, C. P., R. Ganter, F. LePimpec, A. Trisorio, C. Ruchert, and H. H. Braun, 2010, Phys. Rev. Lett. **104**, 234802.
Hebling, J., 1996, Opt. Quantum Electron. **28**, 1759.
Helfenstein, P., E. Kirk, K. Jefimovs, T. Vogel, C. Escher, H. W. Fink, and S. Tsujino, 2011, Appl. Phys. Lett. **98**, 061502.
Ho, K. M., C. T. Chan, C. M. Soukoulis, R. Biswas, and M. Sigalas, 1994, Solid State Commun. **89**, 413.
Hoffrogge, J., J. P. Stein, M. Kruger, M. Forster, J. Hammer, D. Ehberger, P. Baum, and P. Hommelhoff, 2014, J. Appl. Phys. **115**, 094506.
Hoheisel, M., 2006, Nucl. Instrum. Methods Phys. Res., Sect. A **563**, 215.
Hommelhoff, P., C. Kealhofer, and M. A. Kasevich, 2006, Phys. Rev. Lett. **97**, 247402.
Hommelhoff, P., Y. Sortais, A. Aghajani-Talesh, and M. A. Kasevich, 2006, Phys. Rev. Lett. **96**, 077401.
Huang, Y. C., 1998, Opt. Commun. **157**, 145.
Huang, Y. C., and R. L. Byer, 1996, Appl. Phys. Lett. **69**, 2175.
Huang, Y. C., and R. L. Byer, 1997, in *Proceedings of the Eighteenth International Free Electron Laser Conference*, edited by G. Dattoli and A. Renieri (Elsevier Science B.V., North-Holland, Amsterdam), Pt. II., p. 37.
Huang, Y. C., and R. L. Byer, 1998, Rev. Sci. Instrum. **69**, 2629.
Hunsperger, R. G., 2009, *Integrated Optics: theory and technology* (Springer, New York), 6th ed.
Jackson, J. D., 1975a, *Classical Electrodynamics* (Wiley, New York), 2nd ed., p. 657.
Jackson, J. D., 1975b, *Classical Electrodynamics* (Wiley, New York), 2nd ed., p. 269.
Jensen, K. L., P. G. O'Shea, D. W. Feldman, and J. L. Shaw, 2010, J. Appl. Phys. **107**, 014903.
Jiang, Z., C. B. Huang, D. E. Leaird, and A. M. Weiner, 2007, Nat. Photonics **1**, 463.
Jing, C., 2004, Argonne National Laboratory, WF-Note-222.

Joannopoulos, J. D., S. G. Johnson, J. N. Winn, and R. D. Meade, 2008, *Photonic Crystals: Molding the Flow of Light* (Princeton University Press, Princeton, NJ), 2nd ed.
John, S., 1987, Phys. Rev. Lett. **58**, 2486.
Jongen, Y., 1997, *Proceedings of the 1997 Particle Accelerator Conference, Vancouver* (IEEE, Piscataway, NJ), Vol. 3, p. 3770.
Joshi, C., C. E. Clayton, C. Darrow, and D. Umstadter, 1985, in *Laser Acceleration of Particles*, AIP Conf. Proc, No. 130 (AIP, New York), p. 99.
Kaminskii, A. A., R. J. Hemley, J. Lai, C. S. Yan, H. K. Mao, V. G. Ralchenko, H. J. Eichler, and H. Rhee, 2007, Laser Phys. Lett. **4**, 350.
Karagodsky, V., A. Mizrahi, and L. Schachter, 2006, Phys. Rev. ST Accel. Beams **9**, 051301.
Karagodsky, V., D. Schieber, and L. Schachter, 2010, Phys. Rev. Lett. **104**, 024801.
Kim, K. J., 1986, Nucl. Instrum. Methods Phys. Res., Sect. A **250**, 396.
Kimura, W., *et al.*, 2001, Phys. Rev. Lett. **86**, 4041.
Kroll, N., D. R. Smith, and S. Schultz, 1992, in *Photonic Band Gap Structures: A New Approach to Accelerator Cavities*, AIP Conf. Proc. No. 279 (AIP, New York), p. 197.
Kuhlmey, B. T., T. P. White, G. Renversez, D. Maystre, L. C. Botten, C. Martijnde Sterke, and R. C. McPhedran, 2002, J. Opt. Soc. Am. B **19**, 2331.
Kumar, G. A., J. Lu, A. A. Kaminskii, K. I. Ueda, H. Yagi, T. Yanagitani, and N. V. Unnikrishnan, 2004, IEEE J. Quantum Electron. **40**, 747.
Landau, L. M., and E. M. Lifshitz, 1975, *The Classical Theory of Fields* (Pergamon Press, Oxford), 4th revised English edition, p. 118.
Lawn, B., 1995, *Fracture of Brittle Solids* (Cambridge University Press, Cambridge, England), 2nd ed.
Lawson, J., 1975, Rutherford Appleton Laboratory Report No. RL-75-043.
Lawson, J., 1979, IEEE Trans. Nucl. Sci. **26**, 4217.
Li, Y. P., and C. H. Henry, 1996, IEE Proceedings - Optoelectronics **143**, No. 5, 263.
Lin, S. Y., J. G. Flemming, D. L. Hetherington, B. K. Smith, R. Biswas, K. M. Ho, M. M. Sigalas, W. Zubrzycki, S. R. Kurtz, and J. Bur, 1998, Nature (London) **394**, 251.
Lin, X. E., 2001, Phys. Rev. ST Accel. Beams **4**, 051301.
Litchinitser, N. M., S. C. Dunn, B. Usner, B. J. Eggleton, T. P. White, R. C. McPhedran, and C. Martijnde Sterke, 2003, Opt. Express **11**, 1243.
Liu, Z. L., and D. W. Prather, 2004, Opt. Lett. **29**, 1748.
Luccio, A., G. Matone, L. Miceli, and G. Giordano, 1990, Laser Part. Beams **8**, 383.
Mackay, J., A. Jemal, and N. C. Lee, 2006, *The Cancer Atlas* (American Cancer Society, Atlanta, GA).
Maine, P., D. Strickland, P. Bado, M. Pessot, and G. Mourou, 1988, IEEE J. Quantum Electron. **24**, 398.
Marsh, R. A., M. A. Shapiro, and R. J. Temkin, 2007, in *Proceedings of the 2007 Particle Accelerator Conference* (IEEE, New York), p. 3005.
McCalmont, J. S., M. Sigalas, G. Tuttle, K. M. Ho, and C. M. Soukoulis, 1996, Appl. Phys. Lett. **68**, 2759.
McGuinness, C., 2012, Particle Accelerator on a Chip: Fabrication and Characterization of Woodpile Accelerator Structure, Ph.D. thesis (Stanford University).
McNab, S. J., N. Moll, and Y. A. Vlasov, 1970, J. Opt. Soc. Am. **60**, 1325.
McNab, S. J., N. Moll, and Y. A. Vlasov, 2003, Opt. Express **11**, 2927.
McNeur, J., E. R. Arab, R. Dusad, Z. Hoyer, J. B. Rosenzweig, G. Travish, N. Vartanian, J. Xu, J. Zhou, and R. B. Yoder, 2010, in *Proceedings of the 2010 International Particle Accelerator Conference*, edited by A. Shirakawa and C. Petit-Jean-Genaz (IEEE, New York), pp. 4387–4389.
McNeur, J., N. Carranza, G. Travish, H. Yin, and R. B. Yoder, 2012, in *Advanced Accelerator Concepts, 15th Workshop, 2012*, edited by R. Zgadzaj, AIP Conf. Proc. Vol. 1507 (AIP, New York), pp. 464–469.
Milam, D., and M. J. Weber, 1976, J. Appl. Phys. **47**, 2497.
Miya, T., 2000, IEEE J. Sel. Top. Quantum Electron. **6**, 38.
Mizrahi, A., and L. Schachter, 2004a, Opt. Express **12**, 3156.
Mizrahi, A., and L. Schachter, 2004b, Phys. Rev. E **70**, 016505.
Mizrahi, A., and L. Schachter, 2006, Phys. Rev. E **74**, 036504.
Mizuno, K., J. Pae, T. Nozokido, and K. Furuya, 1987, Nature (London) **328**, 45.
Moerman, I., P. P. VanDaele, and P. M. Demeester, 1997, IEEE J. Sel. Top. Quantum Electron. **3**, 1308.
Moulton, P. F., G. A. Rines, E. Slobodtchikov, K. F. Wall, G. Frith, B. Sampson, and A. Carter, 2009, IEEE J. Sel. Top. Quantum Electron. **15**, 85.
Muggli, P., V. Yakimenko, M. Babzien, E. Kallos, and K. P. Kusche, 2008, Phys. Rev. Lett. **101**, 054801.
Munroe, B. J., A. M. Cook, M. A. Shapiro, R. J. Temkin, V. A. Dolgashev, L. L. Laurent, J. R. Lewandowski, A. D. Yeremian, and S. G. Tantawi, 2013, Phys. Rev. ST Accel. Beams **16**, 012005.
Murao, T., K. Saitoh, and M. Koshiba, 2011, Opt. Express **19**, 1713.
Mustonen, A., P. Beaud, E. Kirk, T. Feurer, and J. L. Shaw, 2011, Appl. Phys. Lett. **99**, 103504.
Musumeci, P., 2005, Phys. Rev. Lett. **94**, 154801.
Na, Y. C. N., R. Siemann, and R. L. Byer, 2005, Phys. Rev. ST Accel. Beams **8**, 031301.
Nantista, C., S. Tantawi, and V. Dolgashev, 2004, Phys. Rev. ST Accel. Beams **7**, 072001.
Naranjo, B., A. Valloni, S. Putterman, and J. B. Rosenzweig, 2012, Phys. Rev. Lett. **109**, 164803.
Ng, C. K., R. J. England, L. Q. Lee, R. J. Noble, V. Rawait, and J. Spencer, 2010, Phys. Rev. ST Accel. Beams **13**, 121301.
Ng, K. Y., 1990, Phys. Rev. D **42**, 1819.
Noble, R. J., and J. E. Spencer, 2007, in *Proceedings of the 2007 Particle Accelerator Conference* (IEEE, New York), p. 3112.
Noble, R. J., J. E. Spencer, and B. T. Kuhlmey, 2011, Phys. Rev. ST Accel. Beams **14**, 121303.
Palmer, R. B., 1972, J. Appl. Phys. **43**, 3014.
Palmer, R. B., 1980, Part. Accel. **11**, 81 [http://cds.cern.ch/record/1107986/files/p81.pdf].
Palmer, R. B., 1986, in *Proceedings of the US-CERN School on Particle Accelerators, South Padre Island, October 1986*, Lecture Notes in Physics Vol. 296 (Springer-Verlag, Berlin), p. 607.
Panofsky, W. K. H., and W. A. Wenzel, 1956, Rev. Sci. Instrum. **27**, 967.
Pekarek, S., A. Klenner, T. Sudmeyer, C. Fiebig, K. Paschke, G. Erbert, and U. Keller, 2012, Opt. Express **20**, 4248.
Peralta, E. A., 2011, in *Proceedings of the 2011 Particle Accelerator Conference* (IEEE, New York), p. 280.
Peralta, E. A., E. Colby, R. J. England, C. McGuinness, B. Montazeri, K. Soong, Z. Wu, and R. L. Byer, 2012, in *Proceedings of the 2012 Advanced Accelerator Concepts Workshop*, AIP Conf. Proc. No. 1507 (AIP, New York), p. 169.
Peralta, E. A., *et al.*, 2013, Nature (London) **503**, 91.
Plettner, T., and R. L. Byer, 2008, Phys. Rev. ST Accel. Beams **11**, 030704.

Plettner, T., R. L. Byer, E. R. Colby, C. M. S. Sears, J. Spencer, and R. H. Siemann, 2005, Phys. Rev. Lett. **95**, 134801.
Plettner, T., R. L. Byer, and B. Montazeri, 2011, J. Mod. Opt. **58**, 1518.
Plettner, T., P. P. Lu, and R. L. Byer, 2006, Phys. Rev. ST Accel. Beams **9**, 111301.
Plummer, J. D., M. D. Deal, and P. B. Griffin, 2000, *Silicon VLSI Technology: Fundamentals, Practice and Modeling* (Prentice Hall, Upper Saddle River, NJ).
Pottorf, S., and X. J. Wang, 2002, in *Quantum Aspects of Beam Physics, 18th Advanced ICFA Workshop* (World Scientific Publishing Co., Singapore), p. 232.
Reiche, S., 1997, Phys. Rev. E **56**, 3572.
Reiser, M., 1994, *Theory and Design of Charged Particle Beams* (Wiley, New York).
Rever, M., S. Huang, C. Yahus, V. Smirnov, E. Rotari, I. Cohanoshi, S. Mokhov, L. Glebov, and A. Galvanauskas, 2009, in *Proceedings of the Conference on Lasers and Electro-Optics and Quantum Electronics and Laser Science (CLEO/QELS) 2009, Baltimore, MD* (IEEE, New York), pp. 313, 314.
Richardson, D. J., J. Nilsson, and W. A. Clarkson, 2010, J. Opt. Soc. Am. B **27**, B63.
Rickman, A. G., G. T. Reed, and F. Namavar, 1994, J. Lightwave Technol. **12**, 1771.
Rong, H., A. Liu, R. Jones, O. Cohen, D. Hak, R. Nicolaescu, A. Fang, and M. Paniccia, 2005, Nature (London) **433**, 292.
Ropers, C., D. R. Solli, C. P. Schulz, C. Lienau, and T. Elsaesser, 2007, Phys. Rev. Lett. **98**, 043907.
Rosenzweig, J., 2003, *Fundamentals of Beam Physics* (Oxford University Press, New York).
Rosenzweig, J. B., 2012, in *Proceedings of the 2012 Advanced Accelerator Concepts Workshop*, AIP Conf. Proc. No. 1507 (AIP, New York), p. 493.
Rosenzweig, J. B., and L. Serafini, 1994, Phys. Rev. E **49**, 1599.
Rosing, M., and W. Gai, 1990, Phys. Rev. D **42**, 1829.
Rowland, K. J., S. Afshar V, and T. M. Monro, 2008, Opt. Express **16**, 17 935.
R-Soft, 2011, BandSolve photonics software (R-Soft Design Group, Ossining, NY).
Ruehl, A., A. Marcinkevicius, M. E. Fermann, and I. Hartl, 2010, Opt. Lett. **35**, 3015.
Rutherford, T. S., W. M. Tulloch, E. K. Gustafson, and R. L. Byer, 2000, IEEE J. Quantum Electron. **36**, 205.
Sakai, J.-i., and Y. Suzuki, 2011, J. Opt. Soc. Am. B **28**, 183.
Saleh, B. E. A., and M. C. Teich, 2007, *Fundamentals of Photonics* (John Wiley and Sons, Hoboken), 2nd ed.
Schachter, L., 1997, *Beam-Wave Interaction in Periodic and Quasi-Periodic Structures* (Springer-Verlag, Berlin), p. 51.
Schachter, L., 2004, Phys. Rev. E **70**, 016504.
Schachter, L., 2011, *Beam-Wave Interaction in Periodic and Quasi-Periodic Structures* (Springer, New York), 2nd ed.
Schachter, L., S. Banna, R. L. Byer, E. Colby, B. Cowan, M. Javanmard, A. Mizrahi, R. Noble, T. Plettner, and R. H. Siemann, 2003, *Proceedings of the 2003 Particle Accelerator Conference* (IEEE, New York), p. 722.
Schachter, L., R. Byer, and R. Siemann, 2002, in *Proceedings of the 2002 Advanced Accelerator Concepts Workshop*, AIP Conf. Proc. No. 647 (AIP, New York), p. 310.
Schachter, L., R. Byer, and R. Siemann, 2003, Phys. Rev. E **68**, 036502.
Schachter, L., and D. Schieber, 1997, Nucl. Instrum. Methods Phys. Res., Sect. A **388**, 8.
Schlyer, D. J., 2005, in *Handbook of Radiopharmaceuticals* (John Wiley & Sons, Ltd, Chichester, UK), pp. 1–70.
Schmidtchen, J., A. Splett, B. Schuppert, and K. Petermann, 1991, Electron. Lett. **27**, 1486.
Sears, C. M. S., 2008, Production, Characterization, and Acceleration of Optical Microbunches, Ph.D. thesis (Stanford University).
Sears, C. M. S., *et al.*, 2008, Phys. Rev. ST Accel. Beams **11**, 101301.
Sfar-Zaoui, W., M. Felix-Rosa, W. Vogel, M. Berroth, J. Butschke, and F. Letzkus, 2012, in *Proceedings of the European Conference on Optical Communication (ECOC), Amsterdam, Netherlands* (IEEE, Piscataway, NJ), Paper Tu.1.E.2.
Shapiro, M. A., W. J. Brown, I. Mastovsky, J. R. Sirigiri, and R. J. Temkin, 2001, Phys. Rev. ST Accel. Beams **4**, 042001.
Shimoda, K., 1962, Appl. Opt. **1**, 33.
Siemann, R., 2004, Phys. Rev. ST Accel. Beams **7**, 061303.
Smirnova, E. I., 2005, Novel Photonic Band Gap Structures for Accelerator Applications, Ph.D. thesis (Massachusetts Institute of Technology).
Smirnova, E. I., A. S. Kesar, I. Mastovsky, M. A. Shapiro, and R. J. Temkin, 2005, Phys. Rev. Lett. **95**, 074801.
Smith, D. R., S. Schultz, N. Kroll, M. Sigalas, K. M. Ho, and C. M. Soukoulis, 1994, Appl. Phys. Lett. **65**, 645.
Smith, R. G., 1972, Appl. Opt. **11**, 2489.
Soong, K., and R. L. Byer, 2012, Opt. Lett. **37**, 975.
Soong, K., R. L. Byer, E. R. Colby, R. J. England, and E. A. Peralta, 2012a, in *Proceedings of the 2012 Advanced Accelerator Concepts Workshop*, AIP Conf. Proc. No. 1507 (AIP, New York), p. 310.
Soong, K., R. L. Byer, E. R. Colby, R. J. England, and E. A. Peralta, 2012b, in *Proceedings of the 2012 Advanced Accelerator Concepts Workshop*, AIP Conf. Proc. No. 1507 (AIP, New York), p. 511.
Soong, K., E. R. Colby, C. McGuinness, R. L. Byer, and E. A. Peralta, 2011, in *Proceedings of the 2011 Particle Accelerator Conference* (IEEE, New York), p. 277, paper MOP095.
Soong, K., E. Peralta, R. J. England, Z. Wu, E. R. Colby, I. Makasyuk, J. P. MacArthur, A. Ceballos, and R. L. Byer, 2014, Opt. Lett. **39**, 4747.
Soref, R. A., J. Schmidtchen, and K. Petermann, 1991, IEEE J. Quantum Electron. **27**, 1971.
Spence, J., 2009, *High Resolution Electron Microscopy*, Monographs on the Physics and Chemistry of Materials (Oxford University Press, New York).
Spence, J., W. Qian, and M. P. Silverman, 1994, J. Vac. Sci. Technol. A **12**, 542.
Spencer, J. E., 1998, Int. J. Mod. Phys. A **13**, 2479.
Spindt, C. A., I. Brodie, L. Humphrey, and E. R. Westerberg, 1976, J. Appl. Phys. **47**, 5248.
Sprangle, P., B. Hafizi, and J. R. Penano, 2009, Phys. Rev. ST Accel. Beams **12**, 050702.
Stratton, J. A., 1941, *Electrodynamic Theory* (McGraw-Hill, New York).
Stuart, B. C., M. D. Feit, A. M. Rubenchick, B. W. Shore, and M. D. Perry, 1995, Phys. Rev. Lett. **74**, 2248.
Stupakov, G., and S. Heifets, 2002, Phys. Rev. ST Accel. Beams **5**, 054402.
Suortti, P., and W. Thomlinson, 2003, Phys. Med. Biol. **48**, R1.
Szipocs, R., K. Ferencz, C. Spielmann, and F. Krausz, 1994, Opt. Lett. **19**, 201.
Tajima, T., and J. M. Dawson, 1979, Phys. Rev. Lett. **43**, 267.
Takeda, Y., and I. Matsui, 1968, Nucl. Instrum. Methods **62**, 306.
Taru, T., J. Hou, and K. C. Knight, 2007, in *European Conference and Exhibition on Optical Communication, Berlin* (VDE Verlag/IET, Berlin), p. 771.
Thwaites, D. I., and J. B. Tuohy, 2006, Phys. Med. Biol. **51**, R343.
Travish, G., and R. B. Yoder, 2011, in *Laser Acceleration of Electrons, Protons, and Ions*, SPIE Proceedings Vol. 8079

(SPIE-International Society for Optical Engineering, Bellingham, WA), p. 80790K.
Tremaine, A., J. Rosenzweig, and P. Schoessow, 1997, Phys. Rev. E **56**, 7204.
Tsujino, S., P. Beaud, E. Kirk, T. Vogel, H. Sehr, J. Gobrecht, and A. Wrulich, 2008, Appl. Phys. Lett. **92**, 193501.
Verellen, D., M. D. Ridder, and G. Storme, 2008, Radiotherapy and Oncology **86**, 4.
Wangler, T. P., 2008, *RF Linear Accelerators* (Wiley, New York), 2nd ed.
Webster, P. L., 1939, J. Appl. Phys. **10**, 501.
White, T. P., R. C. McPhedran, C. Martijnde Sterke, N. M. Litchinitser, and B. J. Eggleton, 2002, Opt. Lett. **27**, 1977.
Wilson, P. B., 1991, Stanford Linear Accelerator Laboratory Report No. SLAC-PUB-2884.
Wu, Z., R. J. England, C.-K. Ng, B. Cowan, C. McGuinness, C. Lee, M. Qi, and S. Tantawi, 2014, Phys. Rev. ST Accel. Beams **17**, 081301.
Wu, Z., C. Ng, C. McGuinness, and E. Colby, 2011, in *Proceedings of the 2011 Particle Accelerator Conference, New York* (IEEE, New York), p. 241.
Xiang, D., Y. Du, W. H. Huang, R. K. Li, Y. Lin, C. X. Tang, L. X. Yan, J. H. Park, and S. J. Park, 2007, in *Proceedings of the 2007 Particle Accelerator Conference* (IEEE, New York), p. 1049.
Xie, M., 2000, in *Proceedings of the 2000 European Particle Accelerator Conference, Vienna, Austria* (AIP, Melville, NY), p. 895.
Yablonovitch, E., 1987, Phys. Rev. Lett. **58**, 2059.
Yamada, K., R. Kuroda, H. Toyakawa, H. Ikeura-Sekiguchi, M. Yasumoto, M. Koike, F. Sakai, K. Mori, H. Mori, and N. Fukuyama, 2009, Nucl. Instrum. Methods Phys. Res., Sect. A **608**, S7.
Yanagisawa, H., C. Hafner, P. Dona, K. Klockner, D. Leuenberger, T. Greber, M. Hengberger, and J. Osterwalder, 2009, Phys. Rev. Lett. **103**, 257603.
Yeh, P., and A. Yariv, 1976, Opt. Commun. **19**, 427.
Yoder, R. B., and J. B. Rosenzweig, 2005, Phys. Rev. ST Accel. Beams **8**, 111301.
Zmudzinski, C., D. Botez, L. J. Mawst, A. Bhattacharya, M. Nesnidal, and R. F. Nabiev, 1995, IEEE J. Quantum Electron. **1**, 129.