Advances in Optics and Photonics

Miniature light-driven nanophotonic electron acceleration and control

Roy Shiloh,[1,*] Norbert Schönenberger,[1,2] Yuval Adiv,[3,4] Ron Ruimy,[3,4] Aviv Karnieli,[3,4,5] Tyler Hughes,[6] R. Joel England,[7] Kenneth James Leedle,[8] Dylan S. Black,[8] Zhexin Zhao,[8] Pietro Musumeci,[9] Robert L. Byer,[6] Ady Arie,[5] Ido Kaminer,[3,4] and Peter Hommelhoff[1,2]

[1]Department of Physics, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Staudtstraße 1, Erlangen 91058, Germany
[2]Max Planck Institute for the Science of Light, Staudtstrasse 2, 91058 Erlangen, Germany
[3]Department of Electrical Engineering, Russell Berrie Nanotechnology Institute, Technion–Israel Institute of Technology, Haifa 32000, Israel
[4]Solid State Institute, Technion–Israel Institute of Technology, Haifa 32000, Israel
[5]School of Electrical Engineering, Fleischman Faculty of Engineering, Tel Aviv University, Tel Aviv 69978, Israel
[6]Department of Applied Physics, Stanford University, Stanford, California 94305, USA
[7]SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, California 94025, USA
[8]Department of Electrical Engineering, Stanford University, Stanford, California 94305, USA
[9]UCLA Department of Physics and astronomy, 475 Portola Plaza, Los Angeles, California 90095, USA
*roy.shiloh@fau.de

Received April 19, 2022; revised September 22, 2022; accepted September 26, 2022; published 23 December 2022

Dielectric laser accelerators (DLAs) are fundamentally based on the interaction of photons with free electrons, where energy and momentum conservation are satisfied by mediation of a nanostructure. In this scheme, the photonic nanostructure induces near-fields which transfer energy from the photon to the electron, similar to the inverse-Smith–Purcell effect described in metallic gratings. This, in turn, may provide ground-breaking applications, as it is a technology promising to miniaturize particle accelerators down to the chip scale. This fundamental interaction can also be used to study and demonstrate quantum photon-electron phenomena. The spontaneous and stimulated Smith–Purcell effect and the photon-induced near-field electron-microscopy (PINEM) effect have evolved to be a fruitful ground for observing quantum effects. In particular, the energy spectrum of the free electron has been shown to have discrete energy peaks, spaced with the interacting photon energy. This energy spectrum is correlated to the photon statistics and number of photon exchanges that took place during the interaction. We give an overview of DLA and PINEM physics with a focus on electron phase-space manipulation.

https://doi.org/10.1364/AOP.461142

1. Introduction . . . 864
1.1. Nanostructure-Mediated Electron–Photon Interaction . . . 865
1.2. Silicon Dual-Pillar Accelerator Ddevices . . . 866
1.3. On-Chip Integration and Prospects of a Teraelectronvolt-Range Energy Collider . . . 868
1.4. Recent Experimental DLA Literature: a Comparison . . . 871
2. Classical Theory of DLA . . . 874
2.1. Dual-Pillar Optical Modes . . . 876
3. Phase-Space Control of DLA . . . 878
3.1. Beam Transport Using APF . . . 879
3.2. Ballistic Bunching Down Toward Zeptosecond Bunches . . . 883
3.3. Phase Manipulation for Low-Energy-Spread Microbunching . . . 887
3.4. Pulse Compression with Optical Beat Note . . . 889
3.5. Soft Tuning of Phase-Space Dynamics . . . 889
3.5a. PFT Illumination . . . 891
3.5b. Trading Energy Gain Versus Transverse Acceptance . . . 893
4. Quantum Nature of DLA . . . 894
4.1. The Quantum Electron Wave Packet in a DLA . . . 895
4.1a. Time Scales in DLA Experiments . . . 895
4.1b. PINEM . . . 897
4.1c. The Quantum-Classical Correspondence in DLA . . . 901
4.1d. Measuring Quantum Features in DLA Experiments . . . 902
4.2. Reading Coherent Information from a Quantum System . . . 904
4.3. SP and the Electron Wave Function in Periodic Structures . . . 912
4.3a. Spontaneous Quantum Interactions . . . 913
4.3b. Stimulated Quantum Interactions . . . 914
Funding . . . 916
Disclosures . . . 916
Data availability . . . 916
References . . . 916

Miniature light-driven nanophotonic electron acceleration and control

Roy Shiloh, Norbert Schönenberger, Yuval Adiv, Ron Ruimy, Aviv Karnieli, Tyler Hughes, R. Joel England, Kenneth James Leedle, Dylan S. Black, Zhexin Zhao, Pietro Musumeci, Robert L. Byer, Ady Arie, Ido Kaminer, and Peter Hommelhoff

1. INTRODUCTION

Dielectric laser accelerators (DLAs) are a highly promising technology that could be used to miniaturize particle accelerators down to the chip-scale. The usage of dielectric structures allows the application of at least an order of magnitude larger optical electric fields onto the photonic nanostructures as compared with the traditional, metallic, radio frequency (RF) acceleration cavities. Further, modern rugged and power-efficient ultrafast lasers are poised to induce these high fields, and have additional advantages over RF technology, including high repetition rates, femtosecond temporal period, and minimal jitter, because the electron emission and subsequent manipulation can be governed by the same laser. Recently, an on-chip electron accelerating structure completely integrated with laser grating couplers was demonstrated [1]. The structure, a product of the inverse design approach in nanophotonics [2], is the first major demonstration using guided- rather than free-space laser illumination.

In 2018, a record 1.8 GeV/m peak acceleration gradient was demonstrated in a DLA structure [3]. At 850 MeV/m average gradient, this is more than 30 times larger than the gradients used today in conventional RF accelerating cavities. However, until recently, DLA structures have been limited in acceleration length, for a fundamentally critical reason: even the best accelerator requires collimation forces. This is because Earnshaw's theorem forbids constant focusing forces in all three spatial directions from being exerted at the same time [4], and as a consequence of the Lorentz force, when accelerating particles, these would also diverge (defocus) spatially, only to eventually crash into the physical boundaries of the structure [5]. To circumvent this effect, the laser-induced acceleration forces can be periodically modulated to achieve collimation and acceleration simultaneously, as discussed in Section 3.1. Such manipulations can be nicely illustrated in phase space using classical mechanics, which is at the core of accelerator physics theory. Only recently, the complex phase-space manipulation of electrons with the help of *optical* fields was experimentally demonstrated using specially designed alternating-phase focusing (APF) nanophotonic structures [6]. There, a sub-relativistic electron beam was actively confined in a channel just 225 nm wide but roughly 80 μm long: on-chip, demonstrating the potential to extend DLAs into millimeter-length structures, and beyond.

When considering DLAs for extended on-chip accelerators, the concept of controlled phase-space manipulation is indispensable. As explained in detail in Section 3, the electron pulse's phase space can then be shaped by (1) designing the photonic nanostructure or (2) engineering the incident laser pulse.

In the case of structure design, attosecond micro-pulses, or *bunch trains*, were generated simply by adding a drift space following an accelerating stage [7,8]. In these examples, expanded on in Sections 3.2–3.4, the shortest electron bunch length of 270 as was experimentally shown.

When using two lasers to drive the DLA structure, the relative phase between them can be tuned to switch between inducing primarily energy modulation (acceleration) and additionally skewing the beam transversely (deflection) [9]. In a more rigorous formulation, different modes of operation were classified in phase space for this case [10], as is derived in Section 2. This kind of phase-space control is envisaged not only to compress electron pulses but also to introduce more complex manipulations [11]. Furthermore, the laser pulse itself can also be engineered to manipulate the electron beam in phase space, as described in Section 3.5, using both (linear) DLA schemes as well as ponderomotive schemes (quadratic in the field).

In the future, these tiny accelerators are projected to serve as compact radiation sources and localized medical irradiation devices, among other applications [12–14]. In particular, we consider here the prospects of using DLA for teraelectronvolt-scale colliders [15] in Section 1.3. We point the interested reader to previously extensive review papers on DLAs in Refs. [13,14].

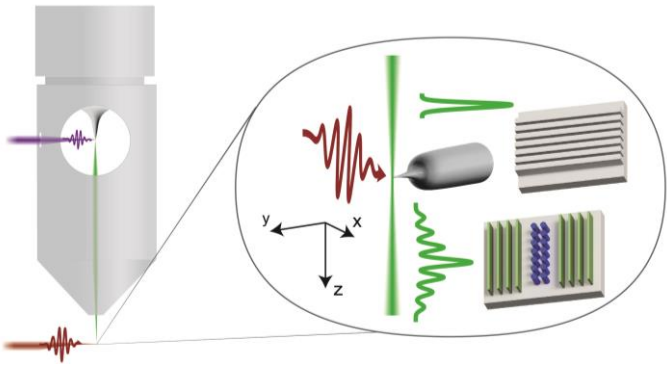
Section 4 is devoted to a comprehensive tutorial and review of recent research in the *quantum nature* of DLAs. A revolution in light and free electron experiments originated from the work of Zewail and Thomas [16], who combined the outstanding electron beam quality of a transmission electron microscope (TEM), its dedicated high-resolution energy spectrometer, and ultrashort laser pulses. Exciting a sample with laser pulses, this kind of laser pump–electron probe experimental scheme enabled the unprecedented ability to observe ultrashort processes in matter with sub-nanometer resolution [17]. It was only natural to extend control of the experimental system and create on-demand ultrashort electron pulses by exciting the electron emitter with an ultraviolet ultrashort pulse. This, in effect, is identical to the DLA scheme: using lasers to manipulate a free electron wave packet, and its energy spectrum, in particular.

From the viewpoint of quantum science, this, in turn, promoted the emergence of photon-induced near-field electron microscopy (PINEM) in 2009 [18], which is essentially the demonstration of quantized energy transitions of an electron wave packet by absorption and emission of photons. The relevant theory and its relation to DLA is elaborated on in Section 4.1. Although seminal experiments were mostly done in TEMs [19–24], PINEM physics was recently also demonstrated in a scanning electron microscope (SEM) [25], which is expected to begin a new trend of research into cascaded photon–electron multi-site interaction. Tens, hundreds, and even thousands of independent interaction sites would be the basis for manipulation of free-electron quantum bits and quantum computing [26], for example, which is derived in Section 4.2. The idea of using the energy spectrum of free electrons as quantum information (bits) carriers is intriguing, because quantum bits are traditionally considered in bulk materials or gases, where the energy levels of bound electrons are practically limited. Conversely, free electrons have practically unlimited equally spaced discrete energy levels, as was recently observed in a quantum DLA experiment with several thousand discrete energy peaks [27], which are accessible in ultrashort time scales.

1.1. Nanostructure-Mediated Electron–Photon Interaction

SEMs and TEMs are popular instruments for electron–photon experiments in academic institution. In the context of this review, these machines are adapted to have optical viewports with access to the electron emitter, usually a sharp metallic tip, see Fig. 1. An ultrafast laser system is then configured to emit two phase-locked short pulses: an

Figure 1



General schematic of the experimental setup for exploring nanostructure-mediated electron–photon interaction. An ultraviolet pulse excites an electron pulse that, following free space propagation, interacts with the evanescent fields generated by a second IR pulse and the nanostructure. The spectrum of the electron pulse can then be analyzed using an electron spectrometer.

ultraviolet pulse (purple), which excites the electron emitter and generates an electron pulse, and an infrared (IR) pulse (red), which is timed to intercept the electron pulse as it traverses in or very close to a nanostructure (see the inset), such that the generated nearfields can mediate the energy exchange between the electrons and photons. In the inset of Fig. 1, three nanostructures are shown, from left to right: a metallic needle tip, which has often been used in PINEM experiments (Sections 4.1 and 4.2), a simple grating as usually described for Smith–Purcell (SP; see Section 4.3), and a dual-pillar design for a DLA (Sections 2 and 3). The initial electron wave packet prior to interaction (green Gaussian shape, top) is then modulated (green peaks, bottom), and measured with an electron spectrometer downstream.

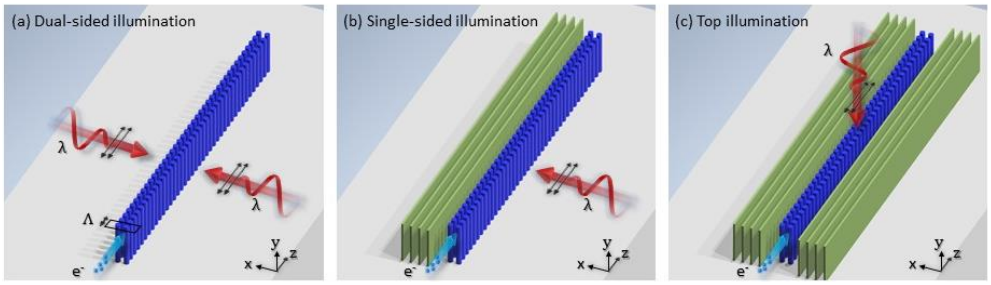
In the configuration shown in Fig. 1, the ultraviolet pulse would have its polarization aligned with the axis of the electron tip emitter, such that the electrical field can efficiently emit electrons from the emitter and toward the first electrostatic acceleration stage inside the electron microscope column. The benefits of using a commercial column are in its industry standard, optimized electron-optical components: lenses, stigmators, deflectors, and in advanced machines also aberration correctors and monochromators, all very stable and easy to operate. For more details on ultrafast electron microscopes, the reader may resort to Refs. [28,29].

At the nanostructure and to impart energy onto the electrons along the propagation direction, the IR laser polarization is set along the electron propagation direction ("z" in Fig. 1), such that the component of the electrical field is parallel. For an efficient interaction, phase-matched structures are preferred. In a simple periodic structure, significant (over ~5%) energy gain is not expected, unlike in tapered structures [30]. Still, acceleration can be measured: the design of a periodic subwavelength grating structure would have a period that, upon spectral decomposition, matches the phase velocity of its first order with the velocity of the electron. In recent years, most of the effort toward demonstrating electron acceleration has been made by using the dual-pillar accelerator structures, which are covered in the next section.

1.2. Silicon Dual-Pillar Accelerator Ddevices

Dual-pillar structures are pairs of pillars, arranged in a generally periodic colonnade where the electron beam passes through its center, as shown in Fig. 2. They were first proposed by Palmer [31] for laser-driven accelerators in the 1980's. They have a

Figure 2



Three popular options of driving a dual-pillar acceleration structure. Blue, dual-pillar colonnade. Green, distributed Bragg mirror. (a) Dual-sided illumination, (b) single-sided driving, and (c) top illumination. Electron and laser propagation direction as well as laser polarization are indicated in each image.

number of advantages for building tunable accelerator structures because they can be driven by multiple lasers both in-plane [32] and out-of-plane [33], to tailor the optical mode inside the device and to account for fabrication imperfections. This enables a single device to operate as an accelerating, focusing, or bunching stage, or be used for beam steering and streaking [8,9,34]. In particular, Fig. 2(a) depicts the classical dual-sided illumination configuration, with the dual-pillar colonnade colored in blue. The electron pulse is directed through the colonnade and between the pairs of pillars. The laser polarization and hence the direction of the electric field component is along the "z" longitudinal direction (parallel double-sided black arrows) [32]. This setup has the advantage of potentially perfect symmetry, both of the lasers and geometry, and the ability to dynamically control the electron's phase space, as shown in Section 2.1. Practically, the spatiotemporal alignment and calibration of the system is challenging. Figure 2(b) includes a distributed Bragg mirror (green plates),which imitates dual-sided illumination. This geometry ameliorates the difficulty in the dual-sided laser illumination setup, where only one laser is required to drive the structure [35]. However, including the mirror as part of the subwavelength nanostructure means long designs are more sensitive to fabrication errors. Figure 2(c) shows a dual-pillar structure illuminated from the top. Although experimentally this structure has shown a slightly lower acceleration gradient than the other two, the great advantage of it is in the ease of the experimental procedure, and, it allows multiple colonnades to be driven equally and concurrently [33].

Silicon is an attractive material for fabricating dual-pillar structures because of the relative ease of device fabrication, its reasonable laser damage threshold, and electrical and thermal conductivity. In addition, due to the high refractive index of silicon at IR wavelengths, dual-pillar devices can work effectively from electron velocities ranging from β<0.25 all the way to relativistic energies in a single architecture [36].

For many of the early demonstration experiments, in-plane illumination with lasers from each side was used as shown in Fig. 2(a). By using dual-pillar gratings with a relatively high reflectivity, the optical mode from each side could be tailored for tunability. Similarly, if a dual-pillar design was intended for single sided drive, the drive side and opposite side of the device could be tailored appropriately for a specific mode profile [32,37]. A typical device designed for subrelativistic electrons at 96 keV or β = 0.54 may have a pillar width of 825 nm in the z (grating periodic) direction, a thickness of 600 nm in the y direction, a periodicity of 1060 nm, an electron channel gap of 400 nm, and a height of 2.8 μm for a drive wavelength of 1960 nm. These devices can produce hyperbolic cosine accelerator modes, hyperbolic sine deflection

modes, and skew modes with correlated acceleration and deflection, as described in Section 2.1, depending on the relative phase of the drive laser from each side.

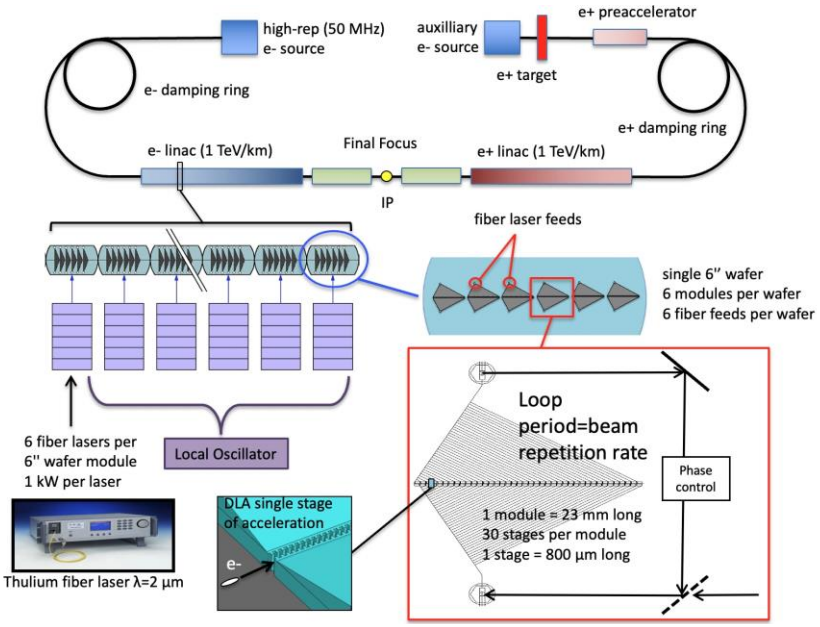
Dual-pillar silicon devices have been demonstrated at wavelengths ranging from the near-infrared (NIR) at 870 nm [32] to the mid-IR at 1980 nm [9]. Longer-wavelength drive lasers allow the dual-pillar device to scale accordingly, resulting in larger structures that are easier to inject into and have larger acceleration buckets. Owing to the indirect bandgap of silicon, NIR wavelengths only experience a small amount of absorption in the dual-pillar structures and hence do not see a large laser damage threshold reduction compared to longer wavelengths. Most recent experiments have been carried out at Tm:fiber compatible wavelengths of 1940–1980 nm, which enables device channels nearly twice the size of NIR wavelengths, and enables better fabrication accuracy as well. Moving forward, telecom wavelengths in the 1550 nm band offer a number of advantages, especially as devices become more integrated with on-chip waveguide power delivery. It is advantageous to use laser wavelengths below 2050 nm to facilitate laser alignment in free-space via one- or two-photon absorption in silicon-based microscopes. At laser repetition rates of 100 kHz and up, there is a loose correlation between having thicker pillars and a higher laser damage threshold in the 2 μm range. Typical silicon dual-pillar devices damage at laser fluences of 10 mJ/cm^2, which corresponds to incident laser fields of 600 MeV/m for 220 fs pulses. At 96keV beam energy, a silicon dual-pillar device with a channel width of $\lambda_0/5$, λ_0 being the laser's central wavelength, will be able to provide maximum channel center gradients of up to 180 MeV/m with 220 fs drive pulses. This maximum gradient increases with the electron energy and with smaller channel widths.

Various experiments were performed to increase the robustness of dual-pillar devices through the use of hydrogen annealing to improve surface roughness and the use of additional surface coatings of Silicon Nitride and other materials to improve the device performance at its laser damage threshold [38]. These treatments were able to yield small but consistent improvements in device performance. Additional photonic elements such as Bragg reflectors can also be incorporated into dual-pillar devices for specific purposes, such as symmetrizing the device fields from one-sided drive laser [37]. In addition, dual pillars can be arrayed into parallel accelerator channels [39] for charge scaling.

1.3. On-Chip Integration and Prospects of a Teraelectronvolt-Range Energy Collider

A future DLA-based linear collider, schematically illustrated in Fig. 3, will require the development of high-gradient accelerator structures as well as suitable diagnostics and beam manipulation techniques, including compatible small-footprint deflectors, focusing elements, and beam position monitors. Key developments in these areas have been made within the last 5 years, including the demonstration of high average gradients (300–850 MeV/m), axial fields up to 1.8 GV/m, speed-of-light synchronous acceleration in laser-driven dielectric microstructures [3,40,41], non-relativistic acceleration with gradients up to 350 MV/m [32,42], and development of preliminary design concepts for compatible photonic components and power distribution networks [30,43]. The power distribution scheme is then envisioned as a fiber-to-chip coupler that brings a pulse from an external fiber laser onto the integrated chip, distributes it between multiple structures via on-chip waveguide power splitters, and then recombines the spent laser pulse and extracts it from the chip via a mirror-image fiber output coupler [44], after which the power is either dumped, or for optimal efficiency, recycled [45]. For long structures, multiple fiber-to-chip couplers will be required. Maintaining phase synchronicity of the laser pulse and the accelerated electrons between many separately fed structures could be accomplished by fabricating the requisite phase delays into the lengths of the waveguide feeds and employing the use of active feedback systems.

Figure 3

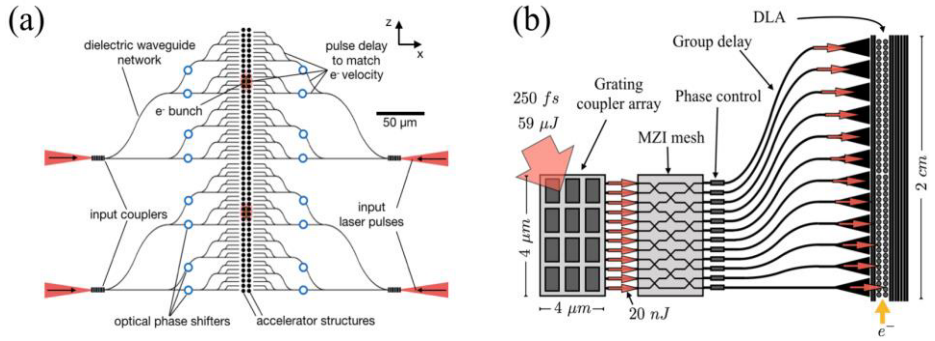


Conceptual schematic of a 30 TeV DLA e^+e^- collider driven by a phase-locked network of energy-efficient solid-state fiber lasers at 20 MHz repetition rate. Laser power is distributed by photonic waveguides to a sequence of dielectric accelerating, focusing, and steering elements co-fabricated on 6-inch wafers which are aligned and stabilized using mechanical and thermal active feedback systems. England *et al.*, "Considerations for a TeV collider based on dielectric laser accelerators," J. Instrum. **17**, P05012 (2022), doi:10.1088/1748-0221/17/05/P05012 [15].

A major challenge of DLA is scaling up the interaction length between the driving laser and the electron beam, which is limited by both the beam dynamics and the laser delivery system. A promising solution is to use integrated optics platforms, built with precise nanofabrication, to provide controlled laser power delivery to the DLA; this would further eliminate many free-space optical components, which are bulky, expensive, and sensitive to alignment. The laser control mechanisms may additionally be implemented on-chip, which will add to the compactness and robustness of the device and allow for precise implementation of laser-driven focusing schemes.

A system for laser coupling to DLA was recently proposed in Ref. [43], in which the laser beam is first coupled into a single dielectric waveguide on the chip and then split several times to spread over the accelerator structure. Here, waveguide bends are designed to implement an on-chip pulse-front tilt (PFT), which delays the incident laser energy to arrive at the accelerator structure at the same time as the moving electron beam [46–48]. Although this work provides a way to achieving interaction lengths on the 100 µm to 1 mm scale, the power splitting approach has the disadvantage of concentrating the optical power at a single-input facet. For longer length structures, requiring more splits, the input facet becomes a bottleneck for damage and nonlinear effects, and future versions of DLA using integrated optical power delivery systems would ideally have "one-to-many" coupling mechanisms where a single laser beam is directly coupled into several waveguides, eliminating this bottleneck. One approach to this kind of coupling uses a large array of grating couplers on the surface of the chip, each supplying power to an individual waveguide [49], as shown in Fig. 4(b). Theoretical studies of grating couplers, combined with inverse design optimization

Figure 4



(a) DLA power distribution network concept proposed in Ref. [43] using dielectric waveguides to split and delay a single input pulse to the accelerator structure. Figure 1 reprinted with permission from Hughes *et al.*, Phys. Rev. Appl. **9**, 054017 (2018), Ref. [43]. Copyright 2018 by the American Physical Society. (b) Schematic of a proposed waveguide-fed DLA designed for long interaction length (not to scale). Figure 1 reprinted with permission from Hughes *et al.*, Phys. Rev. Appl. **11**, 64014 (2019), Ref. [49]. Copyright 2019 by the American Physical Society.

have shown that coupling efficiencies close to 100% may be possible [50]. With areas of several square micrometers having been demonstrated for grating couplers, several thousand may fit on a square millimeter area, which may easily be aligned with a free-space laser source.

To increase the robustness of the DLA coupling, an integrated mesh of Mach-Zehnder interferometers (MZIs) could be fabricated onto the chip. These MZIs have been experimentally demonstrated and act as tunable beam splitters that may share power between waveguides as controlled by integrated optical phase shifters [51]. Initial simulations suggest that for a 250 fs pulse, up to 10 MZIs can be accommodated, which would roughly correspond to being able to share power between 10 adjacent waveguides and should be sufficient for the purposes of DLA. Phase control may be accomplished by thermal or electro-optic phase shifters integrated into the final waveguide sections. Either the electron beam signal or the light scattered out-of-plane may be used as a diagnostic tool for sequentially optimizing the phase shifters. These phase shifters may also be used to implement laser-driven focusing schemes, such as ponderomotive focusing [52] or APF [53], which have shown significant promise for DLA in recent simulation studies [5,54]. Thus, integrated optical phase control gives a path forward for combined acceleration and focusing of the electron beam.

The group delay necessary for matching the arrival of each pulse to the moving electron bunch can be implemented by designing the fixed waveguide geometry, such as the bends described in Ref. [43], in combination with subwavelength gratings [55] embedded in the waveguides. With bend radii as low as 50 μm, it is possible to get close to 100% transmission through the bends [43]. Additional stages may be necessary for compensating dispersion encountered in the waveguides. This will be especially important for longer structures. Previous simulations [43] have shown that these effects will occur at around 1 cm waveguide lengths when using weakly-guided silicon nitride waveguides. An attractive option is to engineer this dispersion to avoid damage and nonlinearities, by sending in an initially chirped and broadened pulse, providing recompression closer to the accelerator.

The coupling from waveguide to several DLA periods may be accomplished with an inverse taper on the waveguide. Alternatively, the DLA structures may be etched

directly into the waveguide, such as in a buried grating [56]. This method is currently being tested experimentally. It was shown in Ref. [43] that a moderate amount of resonance may be beneficial for enhancing the electric fields in the accelerator gap and avoiding the damage and nonlinear constraints in the waveguides. Thus, a quality factor of about 10 may be useful to design into the DLA structures either by inverse design using the adjoint method [57] or by defining dielectric mirrors surrounding the DLA structures. The entire structure may either be driven symmetrically on each side or a dielectric mirror may be used to reflect the incoming light from one side of the device, as discussed in Section 1.2.

As the acceleration process of DLA is optical field driven, the optical phase must be well controlled. Poor synchronization would result in either a decrease of efficiency, electron energy spreading, or defocusing. Frequency comb technologies can detect and control both the phase and repetition rate of the delivered pulses. Because the electrons in a DLA will be optically compressed to form microbunches less than one optical cycle in duration and separated by a single laser wavelength, this requires sub-cycle stabilization of the absolute frequency and relative phase of each pulse.

1.4. Recent Experimental DLA Literature: a Comparison

In 1987, Mizuno *et al.* [58] used a 496 μm sub-millimeter wave laser to demonstrate the inverse SP effect on a metallic grating, by measuring a 5 eV increase in the full width at half maximum (FWHM) of a 80 keV electron beam. However, only in 2013, with the first experimental demonstrations of acceleration using dielectric structures, did the practical potential of using grating accelerators become apparent. Table 1 includes DLA experiments published in peer-reviewed journals, performed since then. When looking at the entries on this table, it is important to note that we quote the average gradient over the interaction length. As the relative phase between the electrons and the laser can vary over this distance, this can be significantly smaller than the peak accelerating gradient (defined as the maximum energy gain over a period of the structure). For example, in the UCLA experiments [3], 850 MV/m was the gradient observed over 21 μm effective interaction length as limited by the non-linear phase chirp induced by Kerr effects, but the peak accelerating gradients in the structure were larger than 1.8 GeV/m. It is worthwhile to point out for the sake of comparison here that in the literature of most accelerating schemes (RF, plasma-based, inverse free-electron laser (IFEL), and others), the peak accelerating gradient is typically quoted.

Since the seminal 2013 publications, four record average acceleration gradients have been measured: In the subrelativistic, SEM energies (up to 30 keV) regime, Kozák *et al.* [59] used a short single-side grating (only three periods) and an ultrafast few-cycle laser pulse of 20 fs to reach an average acceleration gradient of 170 MeV/m using a 30.7 keV electron beam. The combination of a short number of periods avoided dephasing, whereas the few-cycle laser pulse allowed for a much higher damage threshold. In the subrelativistic, TEM-energies (up to 300 keV), Leedle *et al.* [32] used a dual-pillar structure and reached an average acceleration gradient of 376 MeV/m using a 96.3 keV electron beam. The dual-pillar structure has the advantage of enabling a symmetric field profile in the channel at no cost to the fabrication process, assuming it is driven symmetrically from both sides. In the mildly relativistic regime, Cesar *et al.* [3] used a double-grating structure and measured an average acceleration gradient of 850 MeV/m, as mentioned previously. Although double-gratings are challenging to mechanically align one to the other and require compensation to the nonlinear optical effects of the lasers propagating through the bulk, this work boasts the largest gradients measured so far to the best of the authors' knowledge. Lastly, in the relativistic regime, Peralta *et al.* [40] measured an average acceleration gradient of 310 MeV/m using a

Table 1. Comparison of DLA Structures in the Literature

Reference	Type of Structure	Length of Structure or Effective Interaction	Material	Electron Energy	Laser Wavelength	Average Acceleration Gradient
Peralta *et al.* (2013) [40]	Double-sided grating	550 μm (laser spot size: 30 × 300 μm r.m.s.)	SiO_2	60 MeV	800 nm	309.8 ± 20.7 MeV/m
Breuer *et al.* (2013) [42], Breuer *et al.* (2014) [60]	Single-sided grating	24.63 μm (laser spot size: 18 ± 0.8 μm)	SiO_2	27.9 keV	787 nm	25 MeV/m
Leedle *et al.* (2015) [61]	Single-sided grating	15 μm (5.6 ± 0.5 μm interaction distance)	Si	96.3 keV	907 nm	218 ± 20 MeV/m
Leedle *et al.* (2015) [32]	Dual pillar (symmetric)	10.5 μm (5.6 ± 0.5 μm interaction length)	Si	86.5 keV	907 nm	370 ± 40 MeV/m max 196 ± 20 MeV/m channel center
	Dual pillar (asymmetric)	10.5 μm (5.6 ± 0.5 μm interaction length)	Si	96.3 keV	870 nm	376 ± 40 MeV/m max
McNeur *et al.* (2016) [35]	Single-sided grating	8.6 μm (effective interaction length)	SiO_2	28.4 keV 15.2 keV 9.6 keV	787 nm	27 MeV/m 15 MeV/m 7.5 MeV/m
Wootton *et al.* (2016) [41]	Double-sided grating (illuminated from one side)	1 mm (16.3 μm interaction length)	SiO_2	60 MeV	800 nm	690 ± 100 MV/m
Kozák *et al.* (2017) [62]	Planar boundary	23.5 μm (effective interaction length)	Ge Si	28.4 keV	1.93 μm	19 ± 2 MeV/m
Kozák *et al.* (2017) [63]	Single-sided grating	(Not given) Two laser spots with 7 ± 1 μm ($1/e^2$) spot size and spatial separation of 18 ± 1 μm	Si	28.1 keV	1930 nm	(Not given)
Kozák *et al.* (2017) [59]	Single-sided few-period grating	1.86 μm (beam encompasses whole structure)	Si	30.7 keV	1800 nm	170 MeV/m
Leedle *et al.* (2018) [64]	Dual-pillar (symmetric, illuminated from both sides)	15 μm (beam encompasses whole structure)	Si	99.0 keV	1940 nm	200 ± 8 MeV/m max 133 ± 8 MeV/m channel center
Cesar *et al.* (2018) [3]	Double-sided grating	0.5 mm, 1 mm (21.5 μm effective interaction length)	SiO_2	8 MeV	800 nm	850 MeV/m
Cesar *et al.* (2018) [46]	Double-sided grating	0.5 mm	SiO_2	6.5 MeV	800 nm	560 MeV/m
Yousefi *et al.* (2019) [37]	Dual pillar (symmetric, Bragg mirror on one side)	12.0 ± 0.1 μm (beam encompasses whole structure)	Si	28.4 keV	1930 nm	133 ± 9 MeV/m
Black *et al.* (2019) [8]	Dual pillar (two stages of dual-pillar structures separated by drift space, illumination from both sides)	7.88 μm (first stage) 35 μm (drift space) 12.25 μm (second stage) 55.1 μm (total length) (Stages illuminated separately)	Si	57 keV	1980 nm	58 ± 5 MeV/m (measured in second stage, assumed to be the same in first stage)

(Table continued)

Table 1. Continued

Reference	Type of structure	Length of structure or effective interaction	Material	Electron energy	Laser wavelength	Average acceleration gradient
Black *et al.* (2019) [34]	Dual pillar (one stage of dual-pillar structure followed by a drift space followed by an aperture)	14.18 μm (pillars) 39.6 μm (drift space) (Beam encompasses whole structure)	Si	89.4 ± 0.1 keV	1.950 μm	111 ± 6 MeV/m
Schönenberger *et al.* (2019) [7]	Dual pillar (two identical stages of dual-pillar structures, modulator and analyzer, separated by drift space)	13.2 μm (modulator) 30 μm (drift space) 13.2 μm (analyzer) 56.4 μm (total length) (Beam encompasses whole structure)	Si	28.4 keV	1932 nm	34-60 MeV/m
Sapra *et al.* (2020) [1]	Inverse designed structure (not directly illuminated but through a waveguide)	30 μm (40 μm beam diameter on input grating coupler)	Si	83.4 keV	1.94 μm	30.5 MeV/m
Deng *et al.* (2020) [65]	Single-sided grating	18 μm (effective interaction length) 23 μm (effective interaction length)	Sapphire Ga_2O_3	96.3 keV	2000 nm	56 ± 6 MeV/m 104 ± 4 MeV/m
Adiv *et al.* (2020) [66]	Single-sided grating	89 μm (total length) 30 μm (effective interaction length)	Si	195 keV	950 nm	67 MeV/m
Miao *et al.* (2020) [38]	Symmetric dual pillar, illuminated from both sides	15 μm (beam encompasses whole structure)	Annealed and nitride coated Si	97 keV	1960 nm	178 ± 1.8 MeV/m
Niedermayer *et al.* (2021) [67]	Dual pillar (two stages with drift section in between, illuminated from both sides)	31 μm (first stage) 35 μm (drift section) 28 μm (second stage) 94 μm (total length) (Stages illuminated separately)	Si	57 keV	1980 nm	53.6 MeV/m (second stage only)
Shiloh *et al.* (2021) [33]	Dual pillar (symmetric, Bragg mirror on both sides, illuminated from top)	25 μm (beam encompasses whole structure)	Si	28.4 keV	1.93 μm	49.2 ± 3.1 MeV/m
	Dual pillar (period shift of π, Bragg mirror on one side, illuminated from side)	15 μm (beam encompasses whole structure)	Si	28.4 keV	1.93 μm	68.7 ± 5.6 MeV/m
Shiloh *et al.* (2021) [6], Shiloh *et al.* (2022) [68]	APF transport structure with Bragg mirror	77.7 μm (beam encompasses whole structure)	Si	28.4 keV	1933 nm	(Irrelevant)
Chlouba *et al.* (2022) [69]	Single-sided grating	15 μm (effective interaction length)	Diamond	26.8 keV	1933 nm	59.3 ± 6 MeV/m

60 MeV beam. This experiment paved the way to dielectric laser acceleration in the relativistic regime, surpassing the RF-based cavities' highest average gradients of by 6-10 times.

We note that, in a recent publication [6], the authors state that "To our knowledge, the longest DLA structure hitherto used for subrelativistic electrons was 13.2 μm long." This rather narrow claim is inaccurate at best. It is our hope that the following information in Table 1 is a more complete list of existing experimental literature.

2. CLASSICAL THEORY OF DLA

The formulation of DLA interactions begins with the electromagnetic fields in the nanophotonic structure. The fields can be calculated by assuming an incident plane wave impinging on a thin phase-only sub-wavelength diffraction grating, which is periodic along the electron propagation axis. They can then be expanded in a Fourier series, where the first coefficient or spatial harmonic is designed to predominantly contribute to the electron's energy modulation. As the interaction relies on evanescent fields, and as such is only meaningful in a transverse aperture smaller than the decay of these fields, the aperture is usually designed to be about one-tenth of the incident wavelength. For a 2 μm laser, this amounts to roughly 200 nm, and consequently any transverse, off-axis deviation in position or velocity, and hence contribution of transverse forces to the energy gain of the particles can be neglected.

Although it may be intuitively reasonable to begin analyzing the force on a charged point-particle propagating in the channel with the Lorentz force, utilizing instead the rich toolset of accelerator physics theory, the usual treatment starts with the energy gained by the particle over Λ, the length of one unit cell [70], as shown in Fig. 2(a). It is also convenient to convert time to position by assuming a reference particle propagating with constant velocity βc in the lab frame, such that its position along z is $z = \beta ct$, with β the velocity normalized to the speed of light c. Then, the energy gain of any particle that is a distance s behind the reference particle is mainly affected by the z-component of the electric field:

$$\Delta W(x, y; s) = q\ \mathrm{Re}\left\{\int_{-\Lambda/2}^{\Lambda/2} \underline{E}_z(x, y, z) e^{i\omega(z+s)/\beta c} dz\right\}, \tag{1}$$

where $\underline{E}_z(x, y, z)$ is the complex amplitude of the electric field, separating the harmonic carrier dependence $e^{i\omega t}$ of frequency ω at time $t = (z + s)/\beta c$. Owing to the periodicity of the grating Λ along z, we can express the field in a Fourier series,

$$\underline{E}_z(x, y, z) = \sum_{m=-\infty}^{\infty} \underline{\mathrm{e}}_m(x, y) e^{-im2\pi z/\Lambda}, \tag{2}$$

where the series field coefficients are found to be,

$$\underline{\mathrm{e}}_m(x, y) = \frac{1}{\Lambda}\int_{-\Lambda/2}^{\Lambda/2} \underline{E}_z(x, y, z) e^{im2\pi z/\Lambda} dz. \tag{3}$$

We can now calculate the energy gain in terms of the spatial harmonics by inserting Eq. (2) into Eq. (1) and rearranging:

$$\Delta W(x, y; s) = q\ \mathrm{Re}\left\{e^{i\omega s/\beta c} \sum_{m=-\infty}^{\infty} \underline{\mathrm{e}}_m(x, y) \int_{-\Lambda/2}^{\Lambda/2} e^{-iz\left(\frac{\omega}{\beta c} - \frac{2\pi m}{\Lambda}\right)} dz\right\}. \tag{4}$$

Relating the frequency to wavelength, $\omega = 2\pi c/\lambda$, we find that the oscillatory integral is non-zero only when the *synchronicity condition*, which resembles the Widerøe condition in accelerator physics, is satisfied:

$$\Lambda = m\beta\lambda. \tag{5}$$

The energy gain per particle in one structure period due to spatial harmonic m is then given by

$$\Delta W(x, y; s) = \sum_{m=-\infty}^{\infty} q\Lambda |\underline{e}_m(x, y)| \cos\left(\frac{2\pi}{\beta\lambda} s + \varphi_m(x, y)\right), \tag{6}$$

in which we have expressed the field coefficients in phasor form, $\underline{e}_m = |\underline{e}_m| e^{+i\varphi_m}$. It is interesting to note that the spatial harmonic m would be equal to the number of *buckets* per grating period, or areas of stable dynamics in phase space, which can be used for acceleration [70].

As stated previously, we most usually aim to use the first spatial harmonic, where all others are generally negligible in comparison. In such a case, we arrive at the final approximation using $m = 1$,

$$\Delta W(x, y; s) \approx q\Lambda |\underline{e}_1| \cos\left(\frac{2\pi}{\beta\lambda} s + \varphi_1\right). \tag{7}$$

We remind the reader that the particle discussed so far was defined to be at a distance s behind the (arbitrary) reference particle. We now designate this particle as the *synchronous particle*, its phase the argument of the cosine, $\varphi_s = 2\pi s/\beta\lambda + \varphi_1$, which can be chosen arbitrarily in each period by engineering the structure (φ_1). From this definition, and considering the phase interval $[0, 2\pi]$, the maximum energy gain and loss would occur at $\varphi_s = 0$ and $\varphi_s = \pi$, respectively, whereas no energy is gained at $\varphi_s = \pi/2$ and $\varphi_s = 3\pi/2$. The latter, in particular, provides two stable buckets suitable for transport of particles [5,6].

Knowing the energy gain, we can now proceed to determine the momentum differential Δp_s along the longitudinal direction, or the particle trajectory s. This assumes that in one structure period Λ, the velocity change of the particle is small and the rigid beam approximation, meaning that the particle does not change its trajectory, is made. Then, we can use the Panofsky–Wenzel theorem to relate the longitudinal momentum to the transverse one [71], which would give us the prerequisites to describe the particle dynamics.

The Panofsky–Wenzel theorem was published in 1953 in the context of wake-fields in accelerators. However, it is quite general, and assumes only the rigid beam approximation and that the impulse of the particle in question is of key interest, with constant velocity. The theorem states that when considering the Lorentz force $\vec{F} = q(\vec{E} + c\vec{\beta} \times \vec{B})$ and the associated impulse along a rigid path

$$\Delta\vec{p}(x, y; s) = \int_{-\infty}^{\infty} dt\, \vec{F}(x, y, s = -z + \beta ct; t), \tag{8}$$

the curl of this impulse is zero, assuming the magnetic field decays to zero at the boundaries. When considering periodic conditions, then in one such DLA period, the

magnetic field is equal at the boundaries, and we can also conclude that

$$\overrightarrow{\nabla'} \times \Delta\vec{p}(x,y;s) = \int_{-T/2}^{T/2} dt\, [\vec{\nabla} \times \vec{F}(x,y,-z+\beta ct;t)]$$

$$= -q \int_{-T/2}^{T/2} dt \frac{d}{dt}\vec{B}(x,y,-z+\beta ct;t)$$

$$= -q\vec{B}(x,y,-z+\beta ct;t)|_{-T/2}^{T/2} = 0. \tag{9}$$

In the latter, $\vec{\nabla} \equiv (\partial x, \partial y, \partial z)$, and following the transformation $z = -s + \beta ct$, $\overrightarrow{\nabla'} \equiv (\partial x, \partial y, -\partial s)$, and T is the time the electron requires to traverse a unit cell of period Λ. A proof can be shown to use a couple of vector identities and the Maxwell equations [72,73]. In one period of our structure, we call $\Delta\vec{p}$ the *kick* such that when taking its cross product $\hat{s} \times (\overrightarrow{\nabla'} \times \Delta\vec{p}(x,y;s)) = 0$, we find a relation between the longitudinal and transverse kicks, namely that $\vec{\nabla}_{xy}\Delta p_s = -\partial\Delta\vec{p}_{xy}/\partial s$. From the fundamental definition of work, we can write $\Delta W = \beta c\, \Delta p_s$ and assuming the synchronicity condition is satisfied (Eq. (5)), we can use Eq. (4) and conclude that,

$$\Delta\vec{p}_{xy}(x,y;s) = -\int \vec{\nabla}_{xy}\Delta p_s ds = \frac{q\Lambda}{\omega} \sum_{m=-\infty}^{\infty} \mathrm{Im}\{e^{i\omega s/\beta c}\, \vec{\nabla}_{xy}\underline{e}_m(x,y)\}. \tag{10}$$

With this result, we now have the full information on both transverse and longitudinal momenta, and energy change, imparted onto the particle using a laser field modulated by a periodic nanostructure. One can use these to devise a quasi-analytic particle tracking and beam dynamics code, which can be utilized to optimize DLA structures [54,70].

2.1. Dual-Pillar Optical Modes

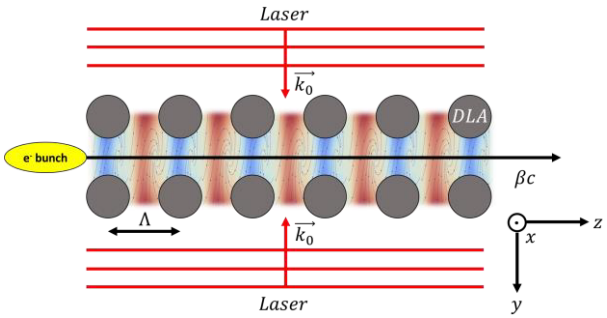
In this section, we follow the derivation from Ref. [10]; for consistency, the constants in Table 2 and definitions are kept.

Figure 5 shows a typical operating mode for a dual-pillar device driven by two normally incident lasers in the y–z device plane. Following Ref. [10], we consider the case of a device that is invariant in the x-dimension, and thus only analyze the 2D case described by a transverse magnetic mode H_x and transverse electric mode E_x. Three-dimensional effects are ignored for the current discussion, but are important for practical accelerator structures, especially as the accelerator length increases. We examine the case of the

Table 2. Constants and Their Definitions for This Section

q	Electron charge
ω	Laser angular frequency
k_0	Laser wave vector in vacuum
λ	Laser wavelength
Λ	Grating periodicity
β	Normalized electron speed v/c
k_z	Bloch wave vector associated with grating phase advance per period
k_n	nth Fourier wave vector of periodic field in grating
Γ_n	Grating fields transverse wave vector
ϵ_n	Complex field strength of nth Fourier harmonic
r_n	Complex geometry factor related to transmissivity of grating in nth harmonic
α_n	Phase of nth harmonic relative to electron
θ	Relative phase of counter-propagating drive lasers

Figure 5



Electron bunch traveling along z through a dual-grating DLA structure. The grating has periodicity Λ in the propagation direction and is assumed to be invariant in x. Two lasers of equal amplitude with relative phase θ are normally incident on the grating structure, with wave vectors $\pm k_0 \hat{y}$. A reference particle enters the structure traveling with speed βc at some injection phase α relative to the laser field and experiences a force given by Eq. (22). Figure 1 reprinted with permission from Black *et al.*, Phys. Rev. Accel. Beams **23**, 114001 (2020), Ref. [10]. Copyright 2020 by the American Physical Society.

transverse magnetic mode for an accelerator device with time harmonic fields at angular frequency ω:

$$\mathbf{H}(\mathbf{r}, t) = \mathbf{H}(\mathbf{r}) e^{-i\omega t}. \tag{11}$$

The transverse magnetic mode must satisfy the wave equation with periodic boundary conditions according to the Floquet–Bloch theorem because the DLA device has periodicity Λ in the z direction,

$$\nabla \times \left(\frac{1}{\epsilon(\mathbf{r})} \nabla \times \mathbf{H} \right) = \mu_0 \omega^2 \mathbf{H}, \tag{12}$$

$$\mathbf{H}(\mathbf{r}) = \mathbf{H}_p(\mathbf{r}) e^{ik_z z}, \quad \mathbf{H}_p(\mathbf{r} + \Lambda \hat{\mathbf{z}}) = \mathbf{H}_p(\mathbf{r}), \tag{13}$$

where the periodic mode $\mathbf{H}_p(\mathbf{r})$ can then be expanded into a Fourier series with $k_n = 2\pi n/\Lambda$

$$\mathbf{H}_p(\mathbf{r}) = \sum_{n=-\infty}^{\infty} \mathbf{h}_n(x, y) e^{ik_n z}. \tag{14}$$

Ignoring the x dependence of the magnetic field yields the transverse magnetic mode of the DLA device:

$$H_x = \sum_n H_x^{(n)} = \sum_{n=-\infty}^{\infty} h_n(y) e^{i(k_n + k_z)z}. \tag{15}$$

Next, we set $k_z = 0$ for normal laser incidence onto the structure. We can describe the transverse dependence of the nth harmonic of the fields in the accelerator channel by the transverse complex wave vector Γ_n, resulting in a hyperbolic cosine and sine component of the device fields. Assuming a rigid beam and the paraxial approximation, the Lorentz force on the electron beam propagating as $z \approx \beta c t + z_0$, z_0 related to the injection phase, in the device can be described by

$$\mathbf{F} = \Re\{[\mathbf{E}(\mathbf{r}) + \beta c \hat{\mathbf{z}} \times \mathbf{B}(\mathbf{r})] e^{-i\omega t}\}, \tag{16}$$

$$\mathbf{F} = \Re\{\tilde{\mathbf{F}}\}. \tag{17}$$

Averaging the Lorentz force the electron experiences over a propagation distance L results in

$$\bar{\mathbf{F}}_n = q\epsilon_n e^{i\frac{k_0 z_0}{\beta}} \begin{bmatrix} 0 \\ \xi_n(a_n^- \cosh(\Gamma_n y) + a_n^+ \sinh(\Gamma_n y)) \\ i(a_n^+ \cosh(\Gamma_n y) + a_n^- \sinh(\Gamma_n y)) \end{bmatrix} \times \frac{1}{L} \int_{-\frac{L}{2}}^{\frac{L}{2}} e^{i\left(k_n - \frac{k_0}{\beta}\right)z} dz \tag{18}$$

with

$$a_n^{\pm} = (1 \mp r_n)\frac{e^{i\theta} \pm 1}{2}, \tag{19}$$

$$\xi_n = (k_n - k_0\beta)/\Gamma_n, \tag{20}$$

$$k_n^2 - \Gamma_n^2 = k_0^2. \tag{21}$$

The geometry factor r_n in Eqs. (14) and (15) describes the transmissivity of the specific periodic nanophotonic architecture and any Fabry–Pérot effects into a single complex amplitude and phase. Here r_n classifies the possible DLA modes between hyperbolic cosine, hyperbolic sine, skew modes, or elliptical modes with relative phase θ between the incident plane wave from each side.

In these terms, the synchronicity condition (Eq. (5)) may also be written as $\beta = k_0/k_n$, and for a particle propagating at $z = \beta ct + z_0$ that satisfies it, the particle is "stationary" with respect to the synchronized wave. We then can define the "electron injection phase" $\alpha_n = k_0 z_0/\beta + \phi_{\epsilon_n}$ for $\epsilon_n = |\epsilon_n| e^{i\phi_{\epsilon_n}}$ and rewrite the average Lorentz force as follows:

$$\bar{\mathbf{F}}_n = q|\epsilon_n| e^{i\alpha_n} \begin{bmatrix} 0 \\ \frac{1}{\gamma}(a_n^- \cosh(\Gamma_n y) + a_n^+ \sinh(\Gamma_n y)) \\ i(a_n^+ \cosh(\Gamma_n y) + a_n^- \sinh(\Gamma_n y)) \end{bmatrix}. \tag{22}$$

Finally, the momentum kick over a single DLA period due to the nth harmonic on a rigid, phase-matched beam can be computed as

$$\Delta\vec{p}_n \approx \int_0^{\Lambda} \frac{dz}{\beta c} \Re\left\{\bar{\mathbf{F}}_n\right\} = \frac{\Lambda}{\beta c} \Re\left\{\bar{\mathbf{F}}_n\right\} \tag{23}$$

or

$$\Delta p_{n,\perp} = \Re\left\{\frac{q|\epsilon_n|\Lambda}{\gamma\beta c} e^{i\alpha_n}(a_n^- \cosh(\Gamma_n y) + a_n^+ \sinh(\Gamma_n y))\right\} \tag{24}$$

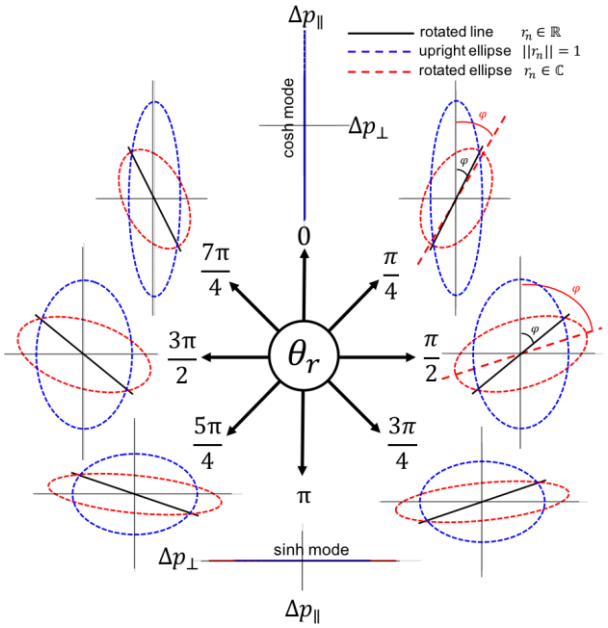
$$\Delta p_{n,\parallel} = \Re\left\{\frac{q|\epsilon_n|\Lambda}{\beta c}\, i e^{i\alpha_n}(a_n^+ \cosh(\Gamma_n y) + a_n^- \sinh(\Gamma_n y))\right\}. \tag{25}$$

This defines the phase-space mode in the dual-pillar device, which can be cosh, sinh, skewed, or elliptical in nature classified by r_n, as illustrated in Fig. 6 and exemplified in experimental form in Fig. 7.

3. PHASE-SPACE CONTROL OF DLA

A systematic study of accelerator physics is based on Hamiltonian mechanics, where the concept of phase space is principal to predict, describe, and optimize the electron interaction with the accelerator device. Also in DLA, this approach was embraced and is the basis of different methods to control and manipulate the electron beam, and is in fact essential to successfully accelerate electrons to high energy with ever more complex schemes. Some required building blocks to achieve this are discussed: in Section 3.1 we explain how to transport electron beams through the nanophotonic structures; in Section 3.2 we describe how to transform an electron pulse into a train of

Figure 6



Plot of the phase space ($\Delta p_\perp$, $\Delta p_\parallel$) for the three classes of modes. The (normalized) momentum kicks plotted here are for the $n = 1$ component of the momentum kicks for an on-axis particle ($y = 0$). γ is set to 1 for convenience. As α varies from 0 to 2π, it traces out ellipses in ($\Delta p_\perp$, $\Delta p_\parallel$) space. Values of r_n corresponding to the three mode classes are then chosen, and θ is varied from 0 to 2π for each mode. Linear skew modes are represented by the rotated line (solid black) with $r_1 = 0.1$. The upright ellipse (dashed blue) has $r_1 = e^{2i}$. The general ellipse (dashed red) has $r_1 = 0.7e^{i}$. The principal cosh and sinh modes occur when $\theta = 0$ or π, respectively. All other θ values produce skew modes. Figure 2 reprinted with permission from Black *et al.*, Phys. Rev. Accel. Beams **23**, 114001 (2020), Ref. [10]. Copyright 2020 by the American Physical Society.

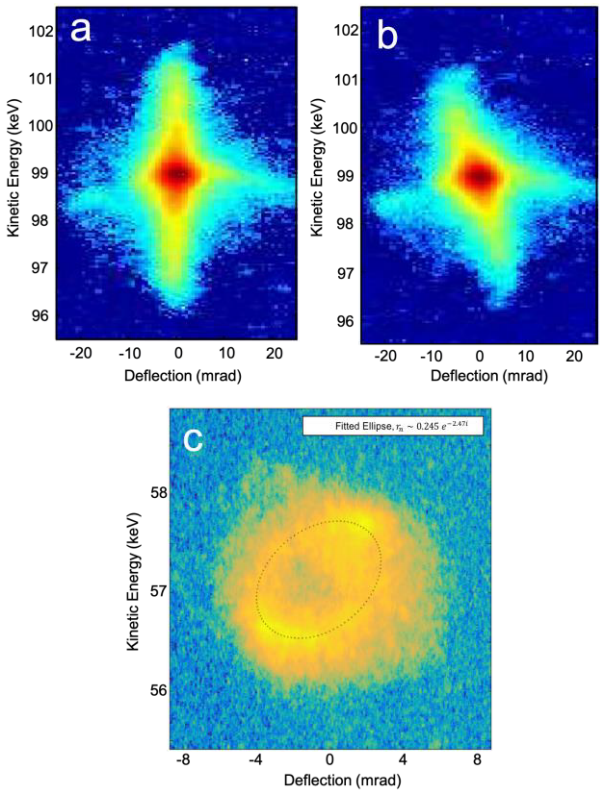
attosecond bunches; Section 3.3 augments on the latter and describes a more complex structure able to provide an attosecond and down toward zeptosecond bunch train with much lower energy spread; a theoretical proposal on the compression of a large fraction of an electron pulse is discussed in Section 3.4; finally, the soft-tuning of phase-space dynamics, using laser pulse modulation, is elaborated on in Section 3.5.

3.1. Beam Transport Using APF

So far, we have treated electron beams in periodic (but finite-length) dual-pillar structures with the natural goal of high acceleration gradients. However, we have also learned from the Lorentz force and the Panofsky–Wenzel theorem that when accelerating the beam (longitudinal force), deflection (a transverse force) is inevitable. These deflections are related to the strong optical fields, usually between 1 and 10 GV/m, and as such cannot be compensated for by conventional electron-optical elements, such as solenoids, quadrupoles, einzel lenses, and similar devices. It is clear that to accelerate over long distances, a scheme must be devised to confine the beam inside the nanostructure, otherwise, sooner rather than later, all electrons would be lost by impact onto the structure.

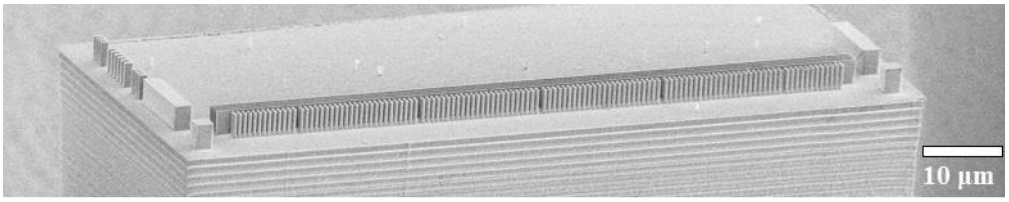
In fact, the most plausible solution is to involve the laser itself, by modulating the laser pulse, for example, using ponderomotive schemes as proposed by Naranjo *et al.* [52] and discussed in Section 3.5. Alternatively, one could modulate the nanostructure,

Figure 7



(a) Symmetric cosh mode produced with dual drive. (b) Linear skew mode produced with dual drive. (c) General elliptical mode produced with single-side drive. Adapted from Ref. [10].

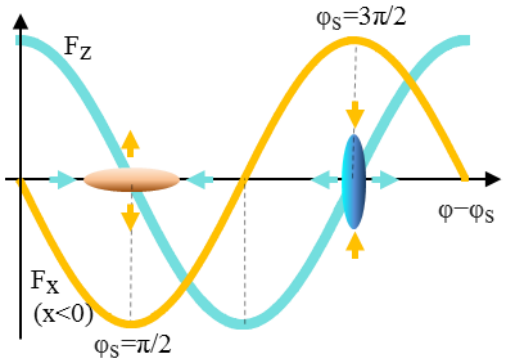
Figure 8



Example of an APF electron beam transport structure. The pillar colonnade is separated into periodic sections, or macro-cells, which alternate between the effective focusing and defocusing forces applied by the nanostructure to electrons traveling inside. The five gaps separating the macro-cells are easily discernible. When optimized, this ensures maximal transmission of electrons to long distances. In this example, a Bragg mirror is situated behind the pillar colonnade so that single-sided illumination suffices, and the whole structure is placed on a pedestal (mesa) to provide clearance for both the electron and laser beams.

for example, using the APF scheme [5,74]. In the latter, when implemented using the dual-pillar design, a section of several longitudinally adjacent dual pillars is defined as a "macro-cell" and these macro-cells are quasi-periodically separated by a gap so as to superimpose two quasi-periodicities: the dual-pillar periodicity (tapered throughout the structure to accommodate the increasing electron velocity); and the APF macro-cell periodicity (also tapered). The gaps between each macro-cell and the next (see

Figure 9



Longitudinal (F_z) and transverse (F_x) forces acting on an a particle of phase φ relative to the (locally) synchronous phase φ_s, in an APF transport structure, for a particle slightly off the electron optical axis toward negative x (meaning that the particle is even driven further away from the axis in the left part of the figure, and toward the axis in the right part of the figure, as indicated). Two stable points, where no energy is gained (see Section 2), exist at $\pi/2$ and $3\pi/2$, where the forces act in a complementary fashion, either focusing-debunching or defocusing-bunching the electrons around each point. In the experimental structure of Fig. 8 and by virtue of the gaps, at every half of an APF period, the synchronous phase if flipped and the actions are reversed to provide stable transport along the structure.

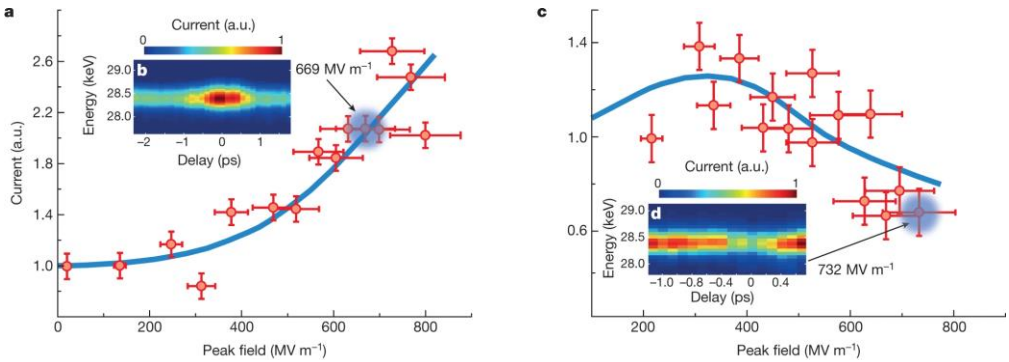
Fig. 8) are designed to apply a phase shift to the effective near-field forces acting on the propagating electron pulse, such that the transverse defocusing forces become focusing, and vice versa. The link between transverse and longitudinal momentum changes described in the introduction to Section 2 is also relevant here: when transverse focusing forces are acting on the particles, longitudinal defocusing (debunching) forces lengthen the electron bunch. The inverse analog is also true: transverse defocusing forces are coupled with longitudinal focusing (bunching) forces. Owing to the periodicity of the electric field of the narrowband laser (treated as a sine wave), this behavior happens concurrently and in a complementary fashion around two points (or "synchronous phases") in one APF period: if a continuous stream of electrons is injected into an APF structure, such that in one laser cycle all phases are sampled, around one such synchronous phase the electrons would be focusing transversely and debunching longitudinally, whereas electrons near the complementary synchronous phase would defocus transversely and bunch longitudinally (see Fig. 9). Denoting the synchronous phase as φ_s, the effect of the gap between macro-cells manifests as the switching of the effective forces that an electron of some phase φ experiences; denoting the transverse force as F_x and the longitudinal force as F_z, we can write an expression for these forces [6]:

$$\begin{aligned} F_x &= \frac{qc}{\beta\gamma} C_c \frac{1}{\gamma} \sinh(k_x x) \sin(\varphi - \varphi_s), \\ F_z &= \frac{qc}{\beta\gamma} C_c \cosh(k_x x) \cos(\varphi - \varphi_s), \end{aligned} \tag{26}$$

where C_c is a constant dependent on the structure design and laser and k_x is the laser wave number of the DLA structure.

Prior to the rather complex, full integration of APF with an accelerating structure, a milestone demonstration of electron transport on chip, using APF but without acceleration, was measured in experiment [6,68]. In this case, the two stable points are at the nodes of the longitudinal force F_z in Fig. 9, and as such provide two complementary

Figure 10



Normalized current or contrast through APF nanophotonic structures as function of the peak optical field. Red points, experimental data; blue curve, particle tracking simulation results. (a) Current increases from 1 (laser off) with increasing field strength up to a maximum value of 2.67 ± 0.05. (b) Time delay scan between electron and laser pulses of the spectrally resolved current at 669 MV $m^{-1} \pm 70$ MV m^{-1}. Clearly, the largest current is observed for maximum temporal overlap (zero time delay). (c) As in (a) but now for the over-focusing structure. Over-focusing sets in for fields larger than 300 MV m^{-1}. The current drops to below 1 for fields larger than 550 MV m^{-1}. (d) Same as in (b) for the over-focusing structure at 732 MV $m^{-1} \pm 76$ MV m^{-1}. Now, the largest electron loss happens at maximum overlap. Adapted from Ref. [6].

regions, or buckets, for transport: in each region, electrons orbit these points in when described in phase space. Although the average electron energy remains unchanged, the prolonged APF action does result in slight bunching of the electron pulse, which is advantageous as discussed in Section 3.2.

In experiment, the behavior of the electrons inside the nanostructure cannot be probed, simply owing to the structure dimensions and inability to access the electrons while they are transported within. However, the transmitted current through the structure can be measured with different field excitations, by tuning the laser power. Without laser irradiation, the natural divergence of the electron beam implies some amount of current will be lost, depending on the structure length. This "laser off" measurement is normalized to unity. The current is measured relative to this laser-off current, defining the contrast. Then, increasing the laser peak field, we expect the contrast to increase because the optical fields can confine the electrons inside the structure. After a certain optimal peak field, where maximum transmission is achieved, the contrast would begin dropping because the optical forces become overly strong and deflect some of the electrons to impact with the pillars. Eventually, if the peak field is set to be too strong, electrons will be actively lost. Owing to the damage threshold of the silicon structures used in that experiment [6], the full behavior could not be measured with the optimized structure (Fig. 10(a)). Therefore, a second "over-focusing" structure was fabricated and measured (Fig. 10(c)). There, the structure did not provide optimal transmission (contrast less than 1.4), but the remainder of the expected transmission curve is readily measurable.

It is important to point out that due to the finiteness of the APF structure, this effect of increased transmission and then a drop due to too-strong optical forces can also occur simply due to variable focusing of the electron beam: when increasing the laser power, electrons focus similar to the action of a lens, and following an aperture (which the long structure effectively creates), a similar transmission curve may be measured. However, in such a case, a definite signature in the energy spectrum would be visible,

as discussed thoroughly in Ref. [68]: although in the APF structure the average energy remains constant and the energy distribution remains small, in a structure where the APF effect is not acting ideally, higher-energy electrons would survive and the measured energy width can increase by a large amount.

In the previous discussion, we treated the vertical (out-of-plane) direction as invariant. Indeed, as DLA structures become longer, the variability of the laser field in this direction becomes important, because the electrons still diverge in this direction and begin to experience different field strengths. One proposed solution would be to use a complex silicon-on-insulator dual-pillar structure, which could provide confinement in the vertical direction, so an extension of the APF principle to three dimensions [75]. Based on this idea, a scalable design has also been proposed theoretically [76]. Alternatively, confinement is also sought after using a top-illumination scheme [33], which may simplify the route to 3D confinement by keeping the same fabrication recipes that are followed today.

3.2. Ballistic Bunching Down Toward Zeptosecond Bunches

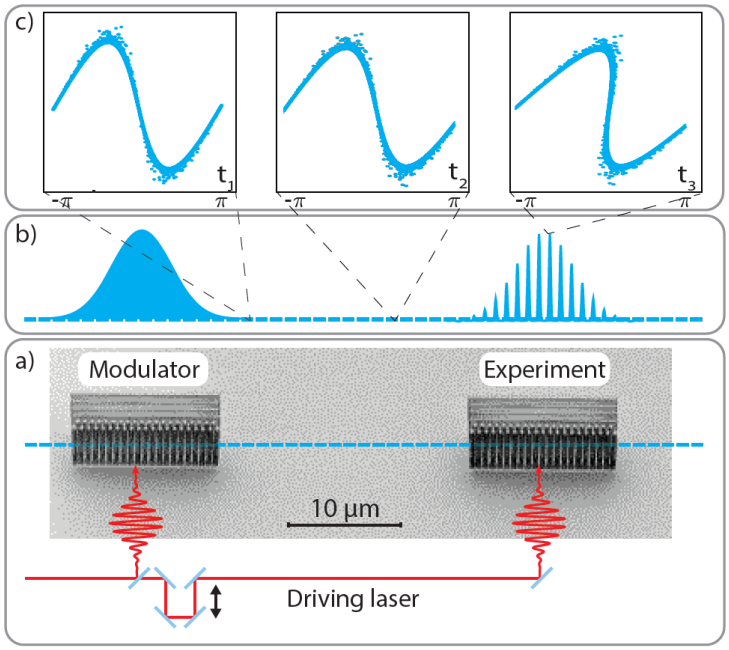
Bunching, i.e., the process of longitudinal compression of electrons, is an important technique for running an efficient accelerator, because only a certain portion of phase space can be accelerated. Furthermore, it can boost the signal-to-noise ratios and improve the temporal resolution in experiments and enable new experiments in ultrafast diffraction [77–79] and ultrafast electron microscopy [29,80–82]. The temporal resolution of laser-triggered electron sources is usually limited by the temporal duration of the electron-releasing laser pulses and subsequent dispersive broadening of the electron pulses. Typical electron pulse durations at the sample have been in the range of 30 fs to 1 ps in experiments not working with actively bunched electron beams [28,29,77–83].

Different schemes can decrease the electron pulse duration via *bunching*. Here we discuss *ballistic bunching* [7,8]. As opposed to, for example, adiabatic bunching, that can happen in a suitably designed accelerator structure when the electrons settle into a bucket, ballistic bunching uses the dispersion of the electrons, either from speed differences in subrelativistic electrons or path length differences in a magnetic chicane for relativistic electrons, to achieve the longitudinal density modulation. Here we only examine the subrelativistic case in detail, and follow Ref. [7].

Figure 11 shows an experiment that has demonstrated ballistic *microbunching* with DLAs. An energy modulation ΔE is imprinted on the electrons in a sinusoidal fashion, due to the nature of the driving field, in a dual-pillar DLA structure. As electron sources typically deliver electron pulses of several tens to hundreds of femtoseconds, which is equivalent to at least a few optical cycles of the driving laser field, the bunching happens for each optical cycle individually, hence the term microbunching. The overall bunch of electrons is microbunched, however, the interaction is the same, and therefore we primarily discuss one individual bunching event.

As depicted in Fig. 12, the initial sinusoidal energy modulation leads to a shearing motion in phase space during the drift in vacuum that follows the initial modulation. This happens, as mentioned previously, due to the different velocities of the electrons with different energies. The shearing reaches a point, where the highest longitudinal electron density is reached (Fig. 12(c)). We call this point the longitudinal focus. Any experimental interaction relying on very short electron pulses has to be placed precisely at this position because the shearing in phase space continues after the focal point, resulting in an increase of pulse duration.

Figure 11



(a) Sketch of the experimental setup with modulator and an experiment. (b) Sketch of the evolution of the electron pulse duration. A relatively broad electron pulse is energy modulated in the first DLA structure. (c) Sketch of the phase-space evolution during the electron drift. The vertical axis denotes the energy of the electrons plotted over one cycle. Following a drift section, the faster, higher-energy electrons catch up with the slower electrons, forming the micro-bunched pulse train. Adapted from Ref. [7].

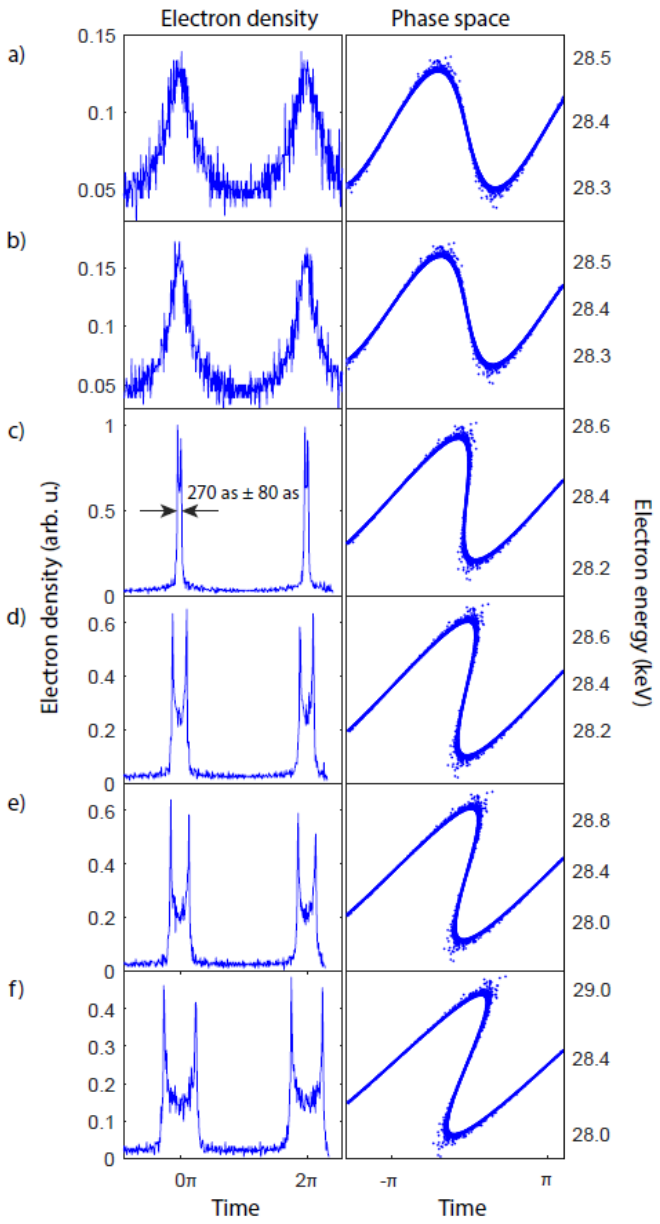
This is visualized in Fig. 12 where simulations show the electron density profile in front (Figs. 12(a) and 12(b)), close to (Fig. 12(c)), and behind the longitudinal focal point (Figs. 12(d)–12(f)). The same effect can be seen in Fig. 13, where by increasing the optical power supplied to the modulator structure, the focal length is decreased, analogous to an optical lens where the radius of curvature is increased.

To design and optimize a DLA bunching stage, various parameters must be considered, and for the best-case scenario the system has to be simulated with methods such as finite difference time domain FDTD to calculate the electromagnetic field distribution inside the DLA structure and particle-tracking software to account for the electron motion through those fields. However, a simplified calculation in phase space can give quite a few insights into a ballistic buncher. For simplicity we assume that the interaction is linear, without higher-order effects, so only one spatial mode of the driving field contributes to the energy modulation of the electrons, and that the electrons are not dephasing. This guarantees that the energy modulation is purely sinusoidal. The resulting energy spread

$$\Delta E = -qe_1\lambda_g N \sin\left(\frac{2\pi}{T}t\right), \tag{27}$$

dependent on the injection time t, is then only dependent on e_1, the peak electric field of the accelerating mode, λ_g the structure period, the number of periods N, and the temporal period of the sinusoidal modulation T (see Ref. [7] and the supplementary material thereof). As here we describe ballistic bunching in the subrelativistic regime, it is advantageous to rewrite the equation in terms of β, the speed of the particles

Figure 12

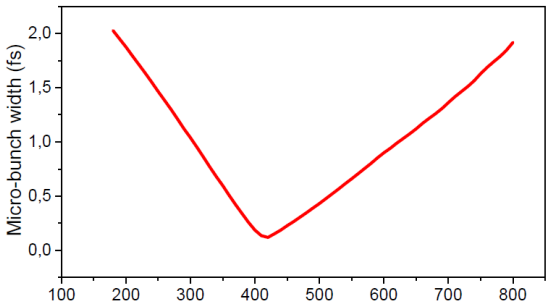


Electron density and phase space at the analyzer structure. Underbunching is visible in the first two rows (a) and (b), where the temporal focus is not yet reached. Shortest micropulses are shown in (c). Various degrees of over-bunching are displayed in the last three rows (d)–(f). The second column shows the phase-space distribution of the microbunches: the vertical axes represent the electrons' energy. When the maximum and minimum of the modulation coincide in time, the temporal focus is reached. After that, the characteristic over-bunching shape is formed, when the high-energy electrons have passed the slow electrons. Adapted from Ref. [7].

normalized to the speed of light, with $\Delta\beta(t) = \Delta E(t)/mc^2\beta\gamma^3$. To further simplify, we only consider the linear region around $t = 0$, resulting in

$$\Delta\beta(t) = -qe_1\frac{2\pi}{T}t\,\frac{\lambda_g N}{mc^2\beta\gamma^3}. \tag{28}$$

Figure 13



Microbunch duration versus incident field strength. The curve shows simulation results, with a minimum of 125 as. Adapted from Ref. [7].

As a first characteristic milestone, we calculate the longitudinal focal length L_0. To do so we first need the time it takes the electrons to reach the temporal focus. By utilizing the distance of each electron to the temporal focus, $\Delta s(t) = {}^{t}/_{T}\lambda_g$, we arrive at

$$\tau = \frac{\Delta s}{\Delta\beta c} = \frac{\lambda_g \beta\gamma^3}{2\pi}\frac{qe_1\lambda_g Nc}{mc^2}. \tag{29}$$

Finally, multiplying this time by the electron velocity yields the longitudinal focal length

$$L_0 = \beta c\tau = \frac{\lambda_g}{2\pi}\frac{\beta^2\gamma^3 mc^2}{\Delta E}, \tag{30}$$

with $\Delta E = e_1\lambda_g N$ the width of the energy modulation.

From this, we can see that the position of the temporal focus is only dependent on the initial electron velocity encoded in β and γ and the induced energy spread. Higher energy spread means the temporal focus is reached earlier, because the velocity differences are larger. Conversely, the temporal focal length increases for higher initial particle velocities, because the relative velocity differences diminish.

The duration of the bunched electrons is now simply the difference in arrival time at the temporal focus between the slowest and the fastest electron. The arrival time can simply be stated as $\tau = \frac{L_0}{v_e}$. Therefore, the time difference is

$$\Delta\tau = \frac{L_0}{v_{min}} - \frac{L_0}{v_{max}} = L_0\frac{\delta v}{v^2} = L_0\frac{\delta\beta c}{\beta^2 c^2} = \frac{L_0}{c\beta^3\gamma^3}\frac{\delta E}{mc^2}, \tag{31}$$

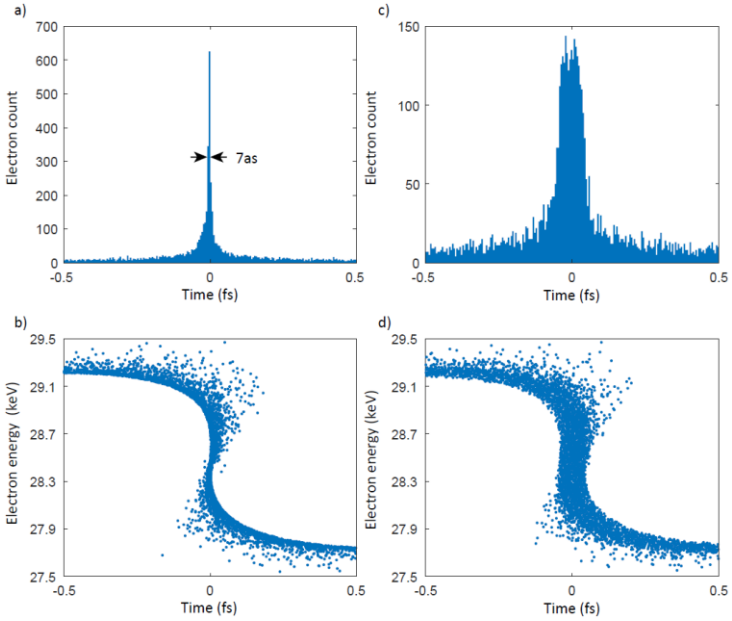
where $\delta\beta = \delta E/mc^2\beta\gamma^3$ is the velocity spread due to the initial energy distribution. Substituting L_0 into this equation, we finally arrive at

$$\frac{\Delta\tau}{T} = \frac{1}{2\pi}\frac{\delta E}{\Delta E}, \tag{32}$$

where T is the laser period, $\Delta\tau$ is the FWHM of the microbunch, δE is the FWHM of the initial energy spread, and ΔE is the peak value of the energy modulation.

We see that the bunch length decreases for increasing energy modulation and decreasing period of the driving field. This is similar to an optical lens, where a more strongly diffracting lens, i.e., a lens with a lower focal length or higher index of refraction, focuses more tightly. We also deduce form this result that an initial energy spread is detrimental for the bunch length. To visualize this phenomenon, we plot the phase space and resulting longitudinal electron density for two different initial energy spreads and otherwise identical conditions in Fig. 14.

Figure 14



(a) Histogram of a 7 as bunch. (b) Corresponding phase space with initial energy spread of 0.5 eV. (c) Bunch histogram for an initial energy spread of 50 eV leading to a bunch duration of 100 as. This exaggerated energy spread was chosen to more easily visualize the dependence of the final bunchlength on the initial energy spread. (d) Corresponding phase space shows the increased energy spread, which contributes significantly to the increased microbunch duration after the phase-space rotation. Adapted from Ref. [7].

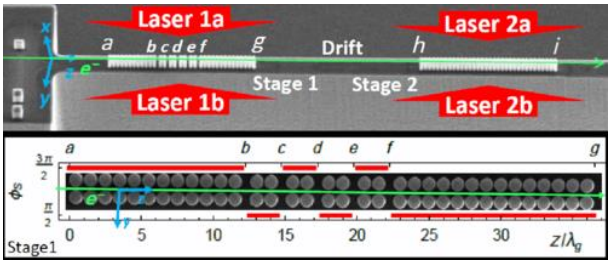
The final phase-space area of the bunch is $\Delta E \Delta \tau$, which is equal to the initial phase-space area $\delta E T$ divided by 2π. This means that the fraction of electrons that end up in the bunch is $1/2\pi = 16\%$, which is the fraction of electrons that are linearly modulated by the sinusoidal fields. A full in-depth discussion of two recent DLA microbunching experiments can be found in Refs. [7,8].

3.3. Phase Manipulation for Low-Energy-Spread Microbunching

APF, as discussed previously in Section 3.1, can also be used when performing microbunching experiments. This essentially allows sections of the DLA device to function as a modulator for microbunching, and a demodulator, which reduces the energy spread of the beam while preserving the microbunch characteristics and transversely focusing the electron beam. Such a modulation–demodulation scheme forms the critical component of a DLA injector, allowing for a net shift of the beam energy in a subsequent accelerator stage.

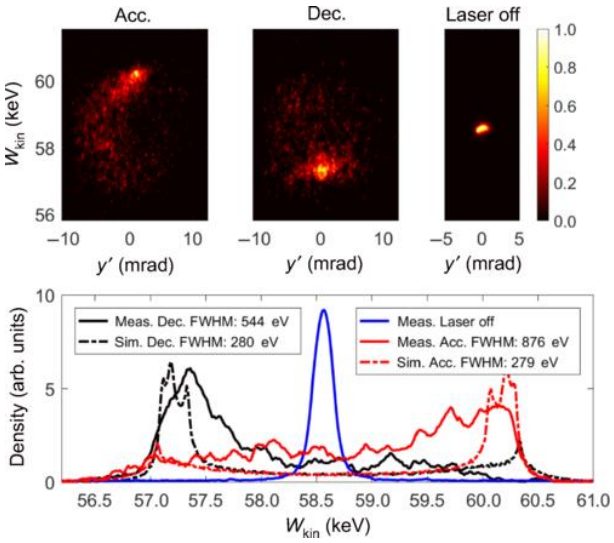
Multi-stage modulation–demodulation and net acceleration was demonstrated in Ref. [67] using the device shown in Fig. 15. This experiment injected a 57 keV electron beam into the hybrid two stage structure, illuminated by four drive laser beams similar to Ref. [8]. In this case, however, the first stage consists of a modulation segment (Figs. 15(a) and 15(b)), transport segment (Figs. 15(b)–15(f)), and demodulation section (Figs. 15(f) and 15(g)). This scheme reduces the net energy spread coming out of stage 1 from approximately 1 keV to 240 eV for a 35 μm bunching drift distance to the second stage. The reduced energy spread improves the subsequent DLA stage performance significantly, because the dispersive action on the electron microbunches

Figure 15



SEM image of the low-energy spread microbuncher, showing the drive laser configuration. In contrast to the case of ballistic bunching discussed in Section 3.2, here the APF section in stage 1 (denoted b–f) reduces the overall energy modulation. Figure XX reprinted with permission from Niedermayer *et al*., Phys. Rev. Appl. **15**, L021002 (2021) Ref. [67]. Copyright 2021 by the American Physical Society.

Figure 16

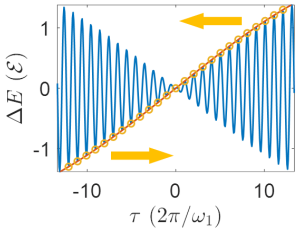


Top: Spectral data as recorded on a microchannel plate detector for accelerating and decelerating phases, as well as laser off conditions. Bottom: Integrated spectrum showing a maximum energy gain of 1.5 ± 0.1 keV with a FWHM energy spread of 0.88 keV. Figure 3 reprinted with permission from Niedermayer *et al*., Phys. Rev. Appl. **15**, L021002 (2021), Ref. [67]. Copyright 2021 by the American Physical Society.

is smaller, and so they will propagate much farther before elongating, which in turn makes it more suitable as a DLA injector stage.

Figure 16 shows the experimental results of the low energy spread microbuching experiment; as can be seen, by controlling the injection phase into the 28 μm long second stage, the overall beam energy can be shifted much more uniformly than in the case of a simple buncher and accelerator [8]. The net energy gain in this experiment was 1.5 ± 0.1 keV with a FWHM energy spread of 0.88 keV for an average gradient of 56 MeV/m in the second stage. Note that the acceleration gradients in the microbunching experiments are lower than in acceleration focused experiments, because the FWHM laser pulse length was increased to 605 fs from 220 fs for more uniform laser fields over the entire 730 fs electron pulse length (note also that the laser electric field FWHM is $\sqrt{2}$ longer than the intensity FWHM).

Figure 17



Energy modulation of the optical beat note node. Drift trajectories of the orange circled phases are indicated by the arrows. Pulse compression of the picosecond bunch envelope occurs after a free space drift length. Adapted from Ref. [11].

3.4. Pulse Compression with Optical Beat Note

In addition to being able to microbunch electrons, optical waveforms in laser accelerators can also be used for manipulation or compression on the picosecond timescale as an alternative to using terahertz or RF radiation. RF compressors are difficult to synchronize on the few femtosecond time scale, and generating terahertz waves typically has very low efficiency and can be difficult to integrate into existing electron microscopes. Using an all optical technique with readily available IR wavelengths could make the few-femtosecond time scale much more accessible for electron diffraction and other experiments.

Following Ref. [11], consider the case of using two optical pulses of different central frequencies ω_1 and ω_2, $\omega_1 > \omega_2$, where the energy modulation for phase-matched electrons becomes $\Delta E(\tau) = \mathcal{E}_1 A_1(\tau)\cos(\omega_1\tau - \phi_1) + \mathcal{E}_2 A_2(\tau)\cos(\omega_2\tau - \phi_2)$. If we set the Gaussian pulse envelope $A_1(\tau) = A_2(\tau) = A(\tau)$ and set the phase difference $\phi_1 - \phi_2 = \pi$, the energy modulation can be simplified to

$$\Delta E(\tau) = 2\mathcal{E}A(\tau)\sin\left(\frac{\omega_1 - \omega_2}{2}\tau\right)\cos\left(\frac{\omega_1 + \omega_2}{2}\tau\right), \tag{33}$$

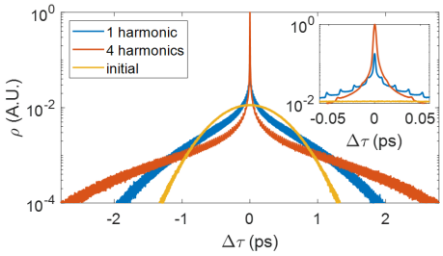
with a fast oscillating component and a slow oscillating component that leads to compression near $\tau = 0$.This is illustrated in Fig. 17. After modulation, those electron phases circled in orange migrate toward the field node as it drifts to a second stage, just like in other bunching schemes. After the proper drift distance, the (faster) phase lagging microbunches have caught up with the (slower) phase leading microbunches, leading to a bunched electron beam. All other phases wash out.

Figure 18 shows an example simulated macrobunching experiment starting with a 1 ps 57 keV electron bunch modulated with $\mathcal{E} = 400$ eV using 20 ps pulses at frequency-corresponding wavelengths $\lambda_1 = 2240$ nm and $\lambda_2 = 2245$ nm. After a 45 mm drift length, velocity bunching yields a 3 fs bunch containing about 3% of the electrons. Adding additional driving harmonics can improve the macrobunching performance of this type of scheme. Using four harmonics in the optical beat note waveform can increase the bunched fraction from 3% to 21% [11]. Note that it is possible to achieve similar results with other means of forming a node in an optical drive waveform. The main disadvantage of this scheme is that it does not bunch the entire electron bunch, but it does capture a larger fraction of electrons and provides shorter bunches than single cycle mid-IR deflection schemes [84].

3.5. Soft Tuning of Phase-Space Dynamics

One of the main challenges in DLA is the relatively small size of the stable region in phase space which strongly limits the amount of electron current/charge which can be accelerated with this technology.

Figure 18



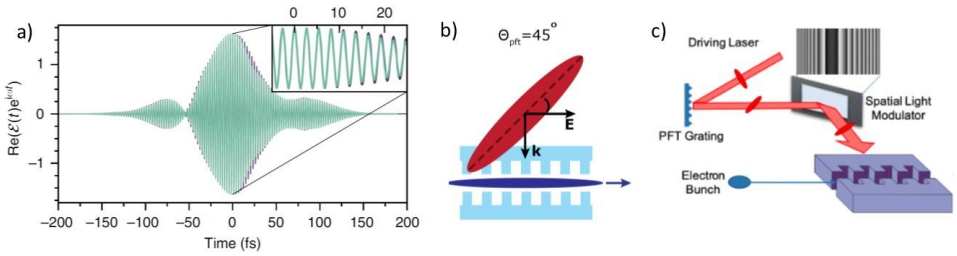
Electron density distribution before (yellow) and after the compression with only fundamental frequencies (blue) or the first four harmonics (red). Here $\Delta\tau$ is the time difference between an electron and the electron pulse center passing an observation point. For normalization, the time integrated densities are kept the same. The inset shows an enlarged view of the peak. Figure 3 reprinted with permission from Zhao *et al.*, Phys. Rev. Lett. **127**, 164802 (2021), Ref. [11]. Copyright 2021 by the American Physical Society.

In the transverse direction, shrinking the accelerator dimensions implies that the aperture to be cleared by the electrons is significantly smaller than in conventional RF accelerators. Strong focusing forces must be applied to keep the particles confined in micrometer-size gaps. The issue can be somewhat mitigated by considering slab-symmetric structures, where the clearance in one of the transverse dimensions can be kept much larger than the laser wavelength. Slab geometry allows for the use of flat beams (having emittance ratio in the two planes much different than one). Flat beams fit well the requirements of some DLA applications (i.e., linear colliders) [85]. At the same time, many applications require round beams, and some structures cannot be fabricated with very large aspect ratios. As discussed in Section 3.1, conventional magnetic field-based systems are typically insufficient to provide the required focusing strengths and a great deal of research has gone into figuring out how to take advantage of the laser fields not just for acceleration, but also for controlling the transverse beam dynamics [5,6].

In the longitudinal dimension, it is important to note that in a DLA structure essentially only particles of a given energy can interact with the electromagnetic field. In fact, prolonged interaction can only occur when the resonant or synchronicity condition (Eq. (5)) is met, that is, when the particle longitudinal velocity is equal to the phase velocity of the wave in the channel enabling phase-synchronous energy-exchange between the laser and the charged particle. The relevant parameter setting the resonant energy bandwidth of the accelerator (i.e., the maximum energy deviation from the resonant condition for which there can still be significant DLA interaction), is the normalized wave amplitude $\alpha = eE_0/mc\omega$ [4]. The relative energy acceptance (i.e., energy bandwidth over resonant energy) is proportional to the square root of this quantity, which is typically very small (10^{-5}) for DLA, especially compared with $\alpha \approx 1$ in conventional RF accelerators. Consequently, one has to carefully design DLA structures to include focusing properties while at the same time be resonant with (and efficiently accelerate) the target energy electron beam. Relatively sophisticated structure engineering has been proposed to ensure that the field seen by the electrons (phase and amplitude) optimizes the DLA acceleration process.

Notably, due to the low Q-factor of the dielectric structures used for DLA, the fields in the electron beam channel are actually a faithful reproduction of the illuminating laser pulses (see Fig. 19(a)), so that controlling phase and amplitude of the drive laser actually offers an interesting alternative to soften the tight tolerance requirements on

Figure 19



(a) Lumerical simulation comparison of the electric field profile of the illuminating laser and the field in the channel demonstrating faithful reproduction (green and purple curves). (b) PFT illumination scheme. (c) Principle of PFT-enabled phase-space control.

structure fabrication and enable tuning of the accelerator characteristics without the need to modify/refabricate the actual dielectric structure [86]. Thanks to the progress in laser pulse manipulation and the development of technologies such as liquid crystal masks and digital micromirror arrays, this tuning can occur online, on demand, and at very high repetition rates (up to kilohertz), by simply modifying the mask applied to the driving laser. We therefore term it *soft-tuning*, to contrast with the hardware-based tuning of the accelerator which does require structure modification. The importance of this is not to be underestimated and can be easily gauged by looking at the continuous developments in low-level RF control systems which are routinely used in conventional accelerator technology to manipulate quickly and nearly arbitrarily the characteristics of the output beams (energy, current, size, etc.) [87].

3.5a. PFT Illumination

Although it is likely that in the future phase and amplitude control of the drive pulses will be implemented using on-chip laser manipulation [1,43,88], the first exploratory research can be carried out using free space coupling. In fact, PFT illumination in combination with modern techniques for spatial control of the phase and amplitude of a laser pulse allow for nearly arbitrary shaping of the field experienced by the electrons in a DLA.

PFT is a standard laser manipulation technique which is used to increase the interaction length in side-illuminated DLA geometries [28,46,89]. Effectively tilting the pulse front of the drive laser ensures that as the electrons travel into the structure they always see an active region of the accelerator. Experimental demonstration of the concept enabled the achievement of 300 keV record high energy gain in DLA extending the interaction region to ~700 μm. For relativistic beams, the PFT angle is nearly exactly 45° (see Fig. 19(b)), which can be easily obtained using a grating and imaging optics with a proper magnification [90]. For non-relativistic beams, changing velocity in the accelerator, the PFT geometry is more complicated as the angle has to vary along the interaction giving rise to the requirement for a PFT banana shape still to be demonstrated in practice.

PFT illumination can be easily coupled with standard methods for spatial light manipulation such as digital micromirror devices or liquid crystal masks [91]. These can be used to apply arbitrary phase/amplitude masks to the transverse profile of the laser which the PFT converts into temporal modulation seen by the electrons. As masks can be changed essentially online, such a system will allow to test and tune the DLA phase spaces using a variety of beam dynamics control methods, including APF, ponderomotive focusing or anything else in between. With a full control over the amplitude

and phase seen by the particles, the question becomes how to best optimize the drive pulse.

We had discussed in Section 3.1 the principle of APF and its implementation by inserting small drifts at proper locations along the DLA structure. In the soft-tuning scheme using the PFT– spatial light modulator combination, there is a limit in implementing APF due to how sharp each transition can be made, ultimately set by the pixel size on the spatial light modulator and resolution of the imaging system that transports the laser to the DLA structure. Depending on the optical magnification between the mask and the accelerator (typically below 2 in order to satisfy the PFT angle requirements), and considering a pixel pitch of 5 μm for the spatial light modulator, even for a nearly perfect imaging system, we can estimate 10 μm as the typical length over which the phase can be controlled. This will likely be too large for non-relativistic electron beam manipulation, but might allow testing of this concept on relativistic beamlines [70].

An alternative scheme has been proposed by Naranjo *et al.* [52] and borrows from the concept of second-order focusing in conventional RF structures [92]. If a superperiodicity δ_k is added to the accelerating wave (the original plan proposed to hard-wire this into the structure, but the PFT–LCM (liquid crystal modulator) combination can be used to modulate the drive laser phase), using Floquet's theorem the field can be decomposed into spatial harmonics. Owing to the different phase velocities of these harmonics, only one of these can be resonant (i.e., maintain a nearly constant phase) with the electrons, whereas all the others will wash out over the beam as it propagates in the DLA. The alternation of focusing and defocusing phases has a net focusing effect which can be used to counteract the resonant defocusing force and maintain the beam confined in the accelerating channel.

To simplify the discussion one can imagine the electrons to be moving in the field of two waves, one with velocity β_r, phase synchronous with the beam (which provides stable longitudinal acceleration, but with a net defocusing effect), and the other at a slower velocity $\beta_1 \cong \beta_r - \delta_k/k$ where k is the wavenumber of the DLA structure. At first order, the equations of motion for the electron deviation from the resonant energy $\delta\gamma$ and for its normalized transverse velocity y' can be written as [52,93]

$$\frac{\partial \delta\gamma}{\partial z} = -\frac{E_0}{mc^2}\sin\varphi - \frac{E_1}{mc^2}\sin(\varphi + \delta_k z)$$

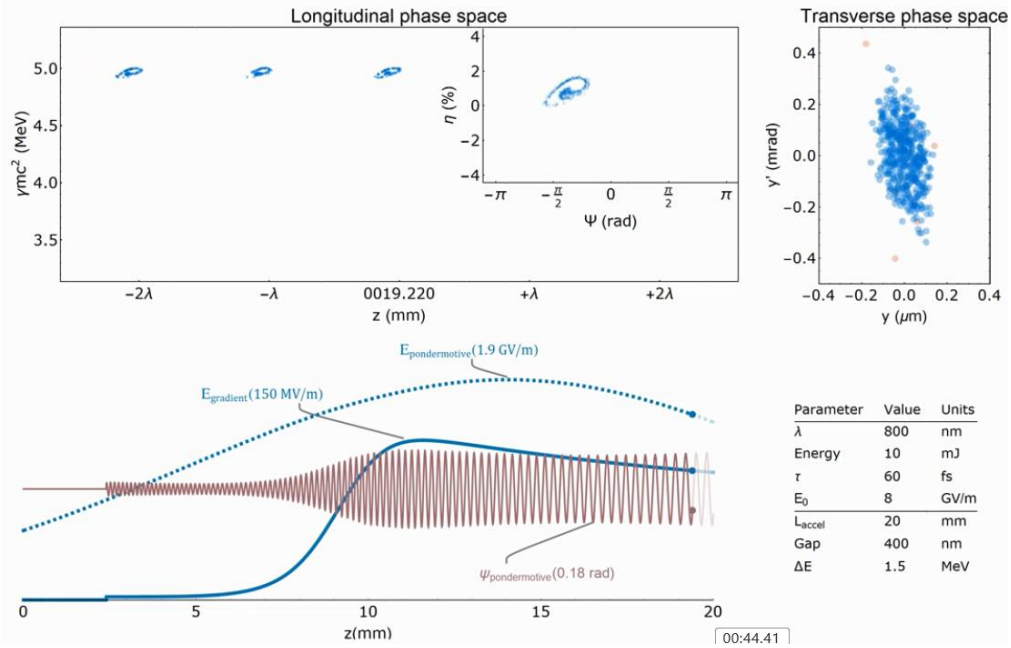
$$\frac{\partial y'}{\partial z} = \frac{E_0}{mc^2}\frac{ky}{\gamma^3\beta^2}\cos\varphi - \frac{E_1}{mc^2}\frac{ky}{\gamma\beta}(1-\beta\beta_1)\cos(\varphi + \delta_k z). \tag{34}$$

Analyzing the solutions of the equation for transverse motion into a slowly varying secular component and a fast oscillation, we can rewrite it for the slow drift motion [94]:

$$\frac{\partial y'}{\partial z} = \frac{E_0}{mc^2}\frac{ky}{\gamma^3\beta^2}\cos\varphi - \left[\frac{E_1\cos\varphi}{mc^2}\frac{k}{\gamma\beta}(1-\beta\beta_1)\right]^2\frac{y}{2\delta_k{}^2} \tag{35}$$

and noting that the coefficient of the second term is negative for all phases, retrieve the ponderomotive focusing effect [95]. The main drawback of the ponderomotive focusing scheme (compared with the APF scheme discussed previously) is the significant need for power in the non-resonant harmonic E_1 to compensate the strong resonant defocusing, so that the laser is not efficiently used to accelerate the particles (i.e., E_0 is relatively small). An interesting advantage is due to the energy scaling of the ponderomotive focusing term (i.e. $1/\gamma^2$), which eventually, for ultrahigh electron energies, dominates over the resonant defocusing term ($1/\gamma^3$). This also results in a matched beta function proportional to the beam energy, perfectly compensating the adiabatic geometric emittance decrease to provide a constant spot size along the accelerator.

Figure 20



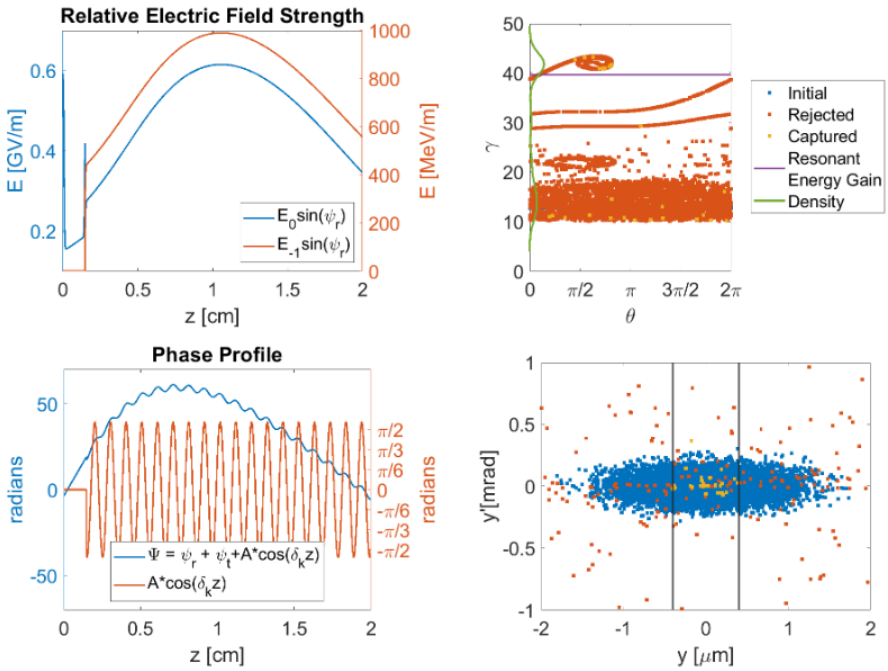
Simulation of the planned ponderomotive focusing and acceleration for 3.5 MeV injection energy at the UCLA Pegasus Laboratory. In this example, most of the laser power (peak field 2 GV/m) is in the non-resonant harmonic $E_{ponderomotive}$ and the accelerating gradient ($E_{gradient}$) is less than 150 MV/m. The super-periodicity ($2\pi/\delta_k$= 300 μm) is added using a sinusoidal modulation (amplitude 0.18 rads) to the drive laser phase (bottom left). The initial section with small modulation is used to bunch the beam and optimize the fraction of particles trapped. Longitudinal and transverse phase spaces at the end of 2 cm interaction are shown (top row). Adapted from Ref. [97].

In addition, looking at the first line in Eq. (34), one can note that the non-resonant spatial harmonics also contribute to the longitudinal phase-space evolution, but because they do not maintain a constant phase relationship with the electrons they will simply induce small-amplitude quivering in the beam's longitudinal velocity, hopefully small enough not to perturb the stability of the trapped region in the longitudinal phase space. To ensure that the acceleration dynamics is not strongly affected by the other spatial harmonics, one can use the Chirikov criterion [96], which states that in order to avoid chaotic motion, the maximum height of the stable bucket in the un-normalized longitudinal phase space $\frac{\delta\gamma}{\gamma} = \sqrt{4\beta^3\gamma\alpha}$ needs to be smaller than the separation of the other spatial harmonics' accelerating buckets $\gamma^2\beta^2\delta_k/k$. Thus, a compromise must be found between the requirements of transverse stability (toward small δ_k/k) and those of longitudinal stability (large δ_k/k). An example of the application of the ponderomotive phase focusing principle is shown in Fig. 20. Here, the amplitude of the resonant and non-resonant harmonics are optimized to transverse confine an electron beam which is accelerated from 3.5 to 5 MeV in a 2-cm-long DLA structure.

3.5b. Trading Energy Gain Versus Transverse Acceptance

Another possibility opened by the use of a 2D liquid crystal mask is amplitude modulation. Essentially, this is achieved by adding a tunable cylindrical lens in the non-PFT axis (i.e., the axis orthogonal to the PFT manipulation), which can be used to control the drive laser intensity along the direction of propagation of the electrons

Figure 21



Simulation using SHARD [98] of mixed modality where amplitude shaping is used to bunch the beam at the beginning of a high gradient acceleration section. The ponderomotive and resonant harmonic have about the same power, and 14 MeV over 2 cm can be obtained in this scheme. The number of particles trapped transversely is relatively small (< 1% trapped fraction).

in the DLA. Amplitude modulation is particularly useful to achieve strong bunching, for example using a high-intensity region followed by a drift section. In combination with phase modulation, it could also be exploited to fine tune the characteristics of the output phase spaces, for example, allowing to trade energy gain and transverse acceptance by simply adjusting the phase mask.

In the example shown in Fig. 21, a short bunching section is followed by a strongly tapered DLA section where the resonant energy is increased from an input level of 6 MeV to a final output energy of nearly 20 MeV in just 2 cm. The phase modulation is used to provide some transverse focusing, but the fraction of trapped particles remains below 1%. As the technology has advanced to the level that 2D mask tuning can happen online, it is envisioned that artificial intelligence-based numerical optimization might be utilized to fine-tune all the accelerator parameters.

In conclusion, soft-tuning for DLAs is a very attractive technique to control the output phase spaces of the beam from DLAs. In particular, until the time on-chip amplitude and phase control will be available at the level required for high gradient acceleration, a PFT-based approach will allow to explore the various options and study the beam dynamics in detail.

4. QUANTUM NATURE OF DLA

In this section, we investigate the quantum aspects in the process of free-electron acceleration and modulation that arise due to the quantum wave nature of the electron. For many years, electron acceleration was fully described by treating the electrons classically, as point charges, despite it being well-known that an electron is not, in fact, a point charge but a wave packet of finite size. Intriguing phenomena can be

observed when modulating the electron wave packet by interaction with nanostructures, with or without light. Accordingly, in Section 4.1 we discuss when is this the wave packet description important in DLA experiments and the relation of DLA to PINEM, along with experimental evidence of quantum features in these regimes. Then, in Section 4.2 we discuss applications of these quantum phenomena, such as the encoding of information onto free electron qubits, and the different quantum effects that are involved. Finally, while the absorption of light that drives electron acceleration in DLA is described as the inverse SP effect, Section 4.3 is dedicated to the SP effect: spontaneous and stimulated emission of light, and the role of the electron wave function in periodic structures.

4.1. The Quantum Electron Wave Packet in a DLA

Our goal for here is to identify the conditions under which a quantum picture can do better to describe the DLA process. These conditions have to do with the time scale of the electron wave packet and the cycle of the laser field.

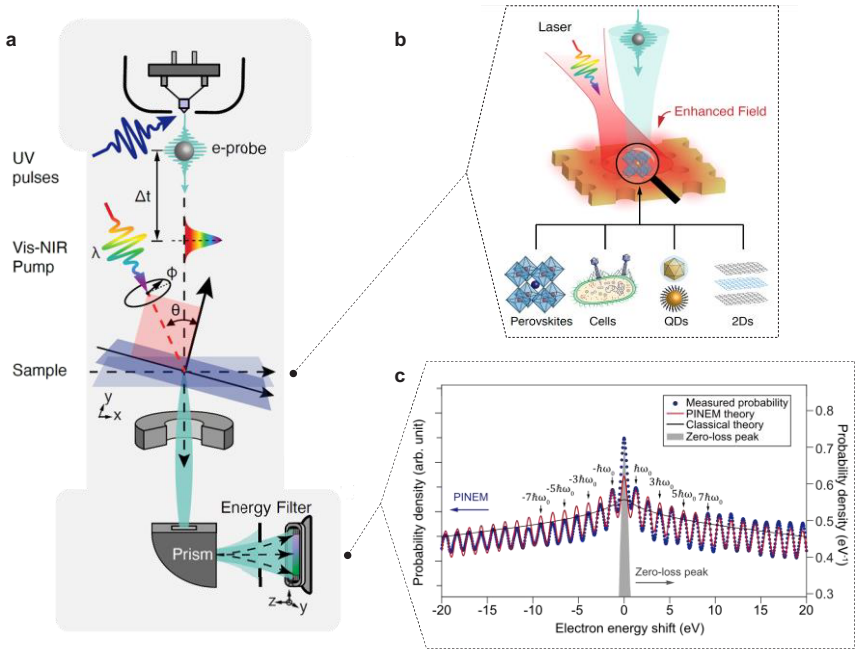
4.1a. Time Scales in DLA Experiments

Using pulsed electron beams as the source for the dielectric laser acceleration process, it is easy to experiment in the regime of a single electron per pulse. Surprisingly, even the spectrum of a single electron can be stretched by such an interaction, namely part of the electron is accelerated and part of it is decelerated, simultaneously. To understand this phenomenon and accurately describe the interaction of a single electron with the near-fields in a DLA structure, it is required to consider the underlying wave nature of the electron, or in other words, we must consider the quantum description of the electron as it is reflected in its wave function. Moreover, the measurement of electron pulses with a high-resolution electron spectrometer unveils the quantum nature of the interaction: the resulted electron spectrum consists of multiple peaks, which correspond in a quantum picture to the exchange of even thousands of photons with a single electron. Such observations were first demonstrated only recently in Ref. [27], showing that the DLA interaction can be described with the same theory as PINEM.

The PINEM approach first emerged as an experimental technique in 2009 [18] in a study of the interaction of free-electron wave packets with optically excited nanotubes. It exploits a modified version of a TEM, which was upgraded to have two optical ports through which femtosecond pulses are coupled into the TEM, in what is now known as ultrafast TEM (UTEM). In the UTEM, an electron pulse is optically excited by an ultraviolet laser pulse and interacts with an optical excitation that is driven by a second laser pulse, in a pump-probe experiment, impinging on the studied sample (e.g., nanowires, membranes, and DLA); see Fig. 22. In 2010, following the first experimental demonstration of PINEM, two papers came out presenting its underlying theory, by accounting for the electron wave function using the Schrödinger equation [99,100].

To probe the regime where PINEM effects can be seen in pump-probe type experiments (with DLA among them), we must consider the time scales in such experiments and how they affect the results. There are four important time scales: (1) the total electron pulse duration (τ^e_{tot}), which is composed of the coherent and incoherent durations; (2) the coherent electron pulse duration (τ^e_{coh}), which is closely related to the coherence duration of the excitation pulse responsible for the photo-emission; (3) the cycle of the laser pulse (τ^l_{cyc}), which is equal to λ/c, where λ is the central wavelength of the pulse and c is the speed of light; and (4) the laser pulse duration (τ^l_p). The relations between these four time scales directly influence the characteristics of the electron

Figure 22



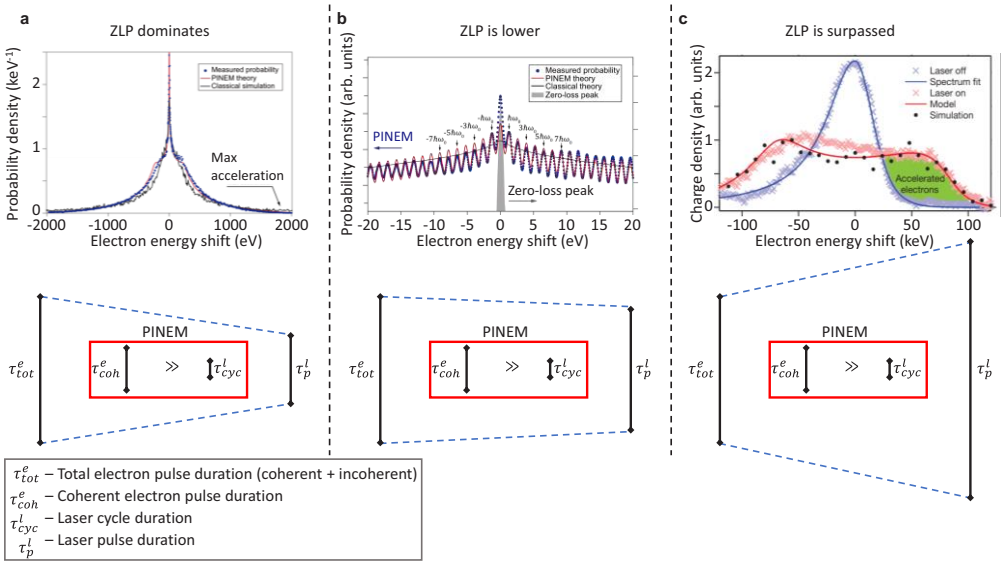
Experimental setup used for PINEM experiments. (a) Schematic of the ultrafast transmission electron microscope (UTEM) used to conduct pump-probe experiments in the PINEM technique. (b) Illustration of the wealth of possible investigated samples, from biology to material engineering and physics. (c) Example of a recorded electron energy spectrum after its interaction with the sample's near-field, showing quantized features that can be explained only by accounting for the electron wave nature. Panels (a) and (b) are reprinted by permission from Nature: Wang *et al*., Nature **582**, 50–54 (2020) [22], Copyright 2020 Nature. Panel (c) is Fig. 4 reprinted with permission from Adiv *et al*., Phys. Rev. X **11**, 041042 (2021), Ref. [27]. Copyright 2021 by the American Physical Society.

energy spectrum after the interaction and also determine whether the quantum nature of the interaction can be seen (see also Fig. 23).

If the coherent electron pulse duration is much longer than the laser cycle, i.e., $\tau^e_{coh} \gg \tau^l_{cyc}$, then the electron wave function extends over several cycles of the laser field and can be simultaneously accelerated and decelerated in what appears as having individual photons exchanged with the electron. Therefore, $\tau^e_{coh} \gg \tau^l_{cyc}$ is a necessary condition for the PINEM interaction to occur. However, it is still not enough for it to be observed in an experiment as the measurement device, for example an electron spectrometer, needs to be sensitive enough to detect energy change as small as the single-photon energy quanta.

The relation between the two remaining time scales, the total electron pulse duration τ^e_{tot} and the laser pulse duration τ^l_p, determines the achievable temporal overlap between the two, and directly influences the shape of the resulting electron spectrum. If $\tau^e_{tot} > \tau^l_p$, then only part of the electron wave function interacts with the near-field, and different parts of the same electron experience varying interaction strengths. The result would be a significant peak at the initial energy of the electron coming from the fact that most of the electron did not interact with the near-field and, hence, did not change its energy (Fig. 23(a)). If $\tau^e_{tot} \approx \tau^l_p$ than most of the electron wave function interacts with the field and the nominal, central energy peak would be suppressed (Fig. 23(b)).

Figure 23



Relevant time scales in the interaction of free electrons with near-fields and their implications on the observed spectrum. (a) $\tau^e_{tot} > \tau^l_p$, (b) $\tau^e_{tot} \approx \tau^l_p$, (c) $\tau^e_{tot} < \tau^l_p$. Four time scales (three regimes of detected spectra) should be considered for predicting the resulting electron energy spectrum. The necessary condition to observe PINEM phenomena and probe the interaction's quantized nature is to have the electron coherence duration τ^e_{coh} much longer than the laser cycle duration τ^l_{cyc}. The characteristic shape of the electron spectrum after the interaction is strongly dependent on the relation between the total electron pulse duration and the laser pulse duration, as can be seen by the three panels. In panels (a) and (b) the discrete peaks in the spectrum are not observed due to insufficient detection resolution, but the condition of $\tau^e_{coh} \gg \tau^l_{cyc}$ is satisfied and thus (unmeasurable) discrete peaks exist. See the text for more details. ZLP, zero-loss peak. Top panels in (a) and (b) adapted from [27] and top panel in (c) adapted from [40].

Lastly, if $\tau^e_{tot} < \tau^l_p$, then the entire electron wave function can interact with a comparable accelerator field amplitude, which give raise to a spectrum that is characterized by a double-humped shape at the edges (Fig. 23(c)).

To put the discussion in a wider context, let us consider the above time scales for electron interactions with RF and microwave fields: For such cases, $\tau^e_{tot} \ll \tau^l_{cyc}$, and thus the entire electron is contained in a fraction of a single cycle of the field, and can be considered as a point particle. This is why the classical picture is sufficient for most purposes in traditional accelerator physics, and why in laser–electron interactions it was necessary to extend the electron description into the quantum regime.

4.1b. PINEM

Prior to the introduction of PINEM as was discussed in the previous section, few studies dealt with the quantum description of free electron interaction with electromagnetic fields [101]. A quantum theory of non-relativistic electron interaction with an electromagnetic wave in an infinite planar surface was introduced in 1971, discussing a stimulated exchange of integer quanta of the wave energy between the electron and the electromagnetic field [102]. In 1977, multiphoton emission and absorption by electrons scattered from an optically excited gas of argon atoms was observed [103].

Other theoretical studies use the correspondence principle to explore the similarities and differences between the classical and quantum descriptions of free electron radiation in the context of inverse Bremsstrahlung radiation [104] and free-electron lasers (FELs) [105,106]. Those studies can be considered as preliminary, unique cases of what is conceived today as PINEM interaction.

Since the emergence of PINEM into the electron microscopy field [18,99,100], several PINEM experiments were carried out, revealing the power of free electrons as probes for different optical phenomena. Using the UTEM platform and the PINEM technique, researchers were able to capture both the spatial distribution of standing plasmon waves on nanowires and their quantized nature in the energy exchange with free electrons [107]. PINEM was also used to demonstrate the "quantum walk" of free electrons in momentum space as they interact with the near-field of a metallic tip [19].

PINEM experiments made continuous advances and entered different areas of free-electron science, allowing the reconstruction and manipulation of the electron wave function. Among them: quantum state tomography of free electrons [108], generation of free electron attosecond pulses [79], attosecond coherent manipulation of the free electron wave function [20], and the creation of electron vortex beams [109]. PINEM physics was also recently demonstrated for the first time in a SEM [25].

In 2020, PINEM experiments were conducted for the first time in the scope of cavity quantum electrodynamics, exploring the interaction of free electrons with a photonic cavity mode. Wang *et al.* looked into the spatial distribution of light in a photonic crystal, deduced the photonic mode lifetime and reconstructed the cavity band structure [22]. Kfir *et al.* observed the interaction of free electrons with whispering gallery modes generated in sphere-shaped micro-resonators, and measured the photonic mode lifetime [23].

All of the above-mentioned experiments show great agreement with the conventional theory of PINEM, as it first proposed in 2010 [99,100]. This theory is suitable for interactions that are localized in space (up to a few micrometers) and, hence, relatively short in time (few femtoseconds). However, this is not the case for DLA interactions with typical length scales of dozens of micrometers. The first PINEM experiment to sustain resonant, phase-matched long interaction of free electrons and near-fields was presented in 2020 by Dahan *et al.* [21]: they maintained a resonant phase-matching between the wave function of a free-electron pulse and an evanescent optical mode over hundreds of micrometers. This phase-matching was shown to be the quantum analog of the classical electron-light phase-matching used in DLAs, or synchronicity condition in accelerator physics. To theoretically describe their results, an extended PINEM theory was developed to capture the effects associated with interaction of an extended length and prolonged duration. This theory then successfully described the first DLA–PINEM experiment [27].

The purpose of the PINEM theory is to describe the interaction of a free-electron wave function with a classical electromagnetic field. Hence, its starting point is the time-dependent Schrödinger equation, which takes the form

$$\left[\frac{\left(\hat{\boldsymbol{P}}+q_e\boldsymbol{A}\right)^2}{2m_e}-q_eV\right]\psi=i\hbar\frac{\partial\psi}{\partial t}, \tag{36}$$

where $\hat{\boldsymbol{P}}=-i\hbar\nabla$ is the momentum operator, q_e and m_e are the electron's charge and mass, respectively, $\boldsymbol{A}$ and V are the electromagnetic field vector and scalar potentials, respectively, and $\psi=\psi(\boldsymbol{r},t)$ is the electron wave function which depends on the spatial position, $\boldsymbol{r}$, and time, t.

PINEM has the following assumptions: (1) electron paraxiality neglecting transverse dynamics of the wave function (i.e., the electron trajectory is constrained to a linear axis which we denote by z); (2) negligible recoil by photon emission (i.e., the electron momentum is much larger than that of the interacting photon); and (3) no ponderomotive corrections (i.e., the field amplitude is small enough so that forces quadratic in the field are smaller than the linear field effects). Under these assumptions, Eq. (36) for an electron with a momentum of $\hbar k_0 \hat{z}$ reads

$$\left[U_0 - \hbar v_e \left(i\frac{\partial}{\partial z} + k_0\right) - i\frac{q_e v_e}{\omega}(\tilde{E}_z e^{-i\omega t} - \tilde{E}_z^* e^{i\omega t})\right]\psi = i\hbar\frac{\partial \psi}{\partial t}, \tag{37}$$

where v_e is the electron velocity and U_0 is the initial electron energy. The electric field's projection along z in complex representation is written as $E_z(\boldsymbol{r},t) = \tilde{E}_z(\boldsymbol{r},t)e^{-i\omega t}$, whereas ω is the laser's angular frequency and $\tilde{E}_z$ is the complex field phasor. For completeness, the physical field is equal to $2\Re[E_z(\boldsymbol{r},t)]$. Here we assume that $e^{-i\omega t}$ is the time harmonic of the field, and any other time component within $\tilde{E}_z(\boldsymbol{r},t)$ is much slower than that.

At this point, we split the electron wave function using the slowly varying envelope approximation into a product, $\psi(\boldsymbol{r},t) = exp\left[i\left(k_0 z - \frac{U_0}{\hbar}t\right)\right]\phi(\boldsymbol{r},t)$, where the first term is a fast-oscillating carrier containing the electron's initial energy and momentum, and the second term describes slower dynamics, seen as the wave function's envelope, which evolves during the interaction. Plugging this expression into Eq. (37) reduces it to

$$v_e\frac{\partial \phi}{\partial z} + \frac{\partial \phi}{\partial t} = -\frac{q_e v_e}{\hbar\omega}(\tilde{E}_z e^{-i\omega t} - \tilde{E}_z^* e^{i\omega t})\phi, \tag{38}$$

which is solved by

$$\phi(r,t) = \phi_0(x,y,T)e^{-2i\Im[g(r,t)e^{-i\omega T}]}. \tag{39}$$

In Eq. (39), ϕ_0 represents the coherent electron wave function before the interaction with the field, and $T = t - z/v_e$ is the time variable in the electron's frame of reference. The function $g(\boldsymbol{r},t)$ is known as the PINEM interaction coupling strength, and it quantifies the overall coupling efficiency between the electron and the electric field. It is also a dimensionless complex parameter that is given by

$$\begin{aligned} g(r,t) &= \frac{q_e}{\hbar\omega}\int_{-\infty}^{z}\tilde{E}_z\left(x,y,z';T+\frac{z'}{v_e}\right)e^{-i\omega\frac{z'}{v_e}}dz' \\ &= \frac{q_e e^{i\omega T}}{\hbar\omega}\int_{-\infty}^{z}E_z\left(x,y,z';T+\frac{z'}{v_e}\right)dz', \end{aligned} \tag{40}$$

where we exploit the relation $E_z(\boldsymbol{r},t) = \tilde{E}_z(\boldsymbol{r},t)e^{-i\omega t}$ in the derivation of Eq. (40).

In the conventional PINEM theory, because the interaction has a limited extent, the field phasor is approximated to be constant in time in the definition of the PINEM coupling strength [99,100]. In contrast, when the interaction maintained over dozens of micrometers, then both the spatial and temporal overlap between the electromagnetic field and the electron wave function should be considered, hence the time-dependence of the electromagnetic field was kept in Eq. (40).

At this point we note that the integration is carried out until the observation point in the z axis. In practice, the electrons are being measured far from the interaction region, and hence we can take the upper integration limit to $z \to \infty$, where g drops its explicit dependence on z. In addition, we see that the transverse (x and y) dependence of g comes

solely from the distribution of E_z. Another important remark is that by multiplying g with the fundamental photon energy of the incident laser ($\hbar\omega$), and dividing it with the interaction length, it becomes equivalent to the acceleration gradient that is commonly discussed when describing the results of a typical DLA experiment, as discussed in Section 4.1.3.

To complete the derivation to a point that is beneficial to describe experimental results, we must calculate the electron energy probability density after the interaction. In general, g can depend on T in a complicated form. To avoid that, we write the Fourier transform of the electron wave function with respect to T as follows: $\phi(x,y,T) = \int_{-\infty}^{\infty} f(x,y,u) exp(iuT/\hbar) du$, where u is the energy variable and $f(x,y,u)$ is a coherent probability density for the electron to be at a certain energy state u. Therefore, the expression $P(U \le u \le U + \Delta U) = |f(x,y,u)|^2 \Delta U$ gives the probability P that the electron energy is within the interval $u \in [U, U + \Delta U]$.

Now, we can apply the Jacobi–Anger expansion, $exp(iq\sin\varphi) = \sum_{\ell=-\infty}^{\infty} J_\ell(q) exp(i\ell\varphi)$, where J_ℓ is the Bessel function of order ℓ, to write Eq. (39) as

$$\phi(x,y,T) = \phi_0(x,y,T) \sum_{\ell=-\infty}^{\infty} f_\ell(x,y,T) e^{i\ell\omega T}$$

$$f_\ell(x,y,T) = J_\ell(2|g(x,y,T)|) e^{i\ell \arg[-g(x,y,T)]}. \tag{41}$$

Equation (41) represents the most general electron wave function modulation that is triggered by the interaction with a classical electric field under the PINEM framework. To derive the electron energy spectrum after the interaction is, however, a delicate step. We note that the coefficients of the coherent energy amplitudes f_ℓ depend on the time coordinate T, hence they no longer represent the formal (independent of T) Fourier coefficients of the electron wave function. Nevertheless, the difference in time scales makes this description accurate for experiments in the UTEM and other platforms for DLAs, because the components $f_\ell(x,y,T)$ in Eq. (41) are slow-varying with respect to the time harmonic $e^{i\ell\omega T}$.

The coherent energy probability density ρ_{coh} of the electron can now be written as

$$\rho_{coh}(x,y,u,\Delta t) = \sum_{\ell=-\infty}^{\infty} P_\ell^{coh}(x,y,\Delta t)\delta(u - \ell\hbar\omega),$$

$$P_\ell^{coh}(x,y,\Delta t) = J_\ell^2(2|g(x,y,\Delta t)|). \tag{42}$$

In the above equation, Δt is the relative time delay between the electron and the incident laser pulse.

To complete the derivation and apply it to a realistic scenario, we must consider the incoherent contributions to the interaction. These contributions are the initial electron (incoherent) energy width (also known as the zero-loss peak (ZLP)), and the transverse distribution (in x and y) of the electron pulse. Both can be easily measured in the UTEM during the experiment and from that we can construct their function form which we denote by ρ_0. Then the overall electron probability density is given by convolving the incoherent part with the result of Eq. (42) with respect to space and time:

$$\rho(x,y,u,\Delta t) = \int\int du dt\ \rho_0(x,y,u,t)\rho_{coh}(x,y,u-u,\Delta t - t). \tag{43}$$

Equation (43) is the most general expression for the electron probability density in space, time, and energy in a typical PINEM experiment within the UTEM. In experiment, conducting an electron energy loss spectroscopy (EELS) measurement,

we measure the quantity $\rho(u,\Delta t_0)=\int\int\rho(x,y,u,\Delta t_0)dxdy$ which is the electron energy probability density for a given electron-laser delay Δt_0.

It is worth noting that Eqs. (42) and (43) contain a set of delta functions that samples the spectrum in multiples of the fundamental photon energy, i.e., $u_\ell=\ell\hbar\omega$. This is the origin of observing the quantum nature of the interaction in PINEM experiment, as each delta function corresponds to the exchange (either emission or absorption) of a photon with a single electron. The full PINEM formula for the electron probability density that is shown in Eq. (43) captures a realistic scenario where around each delta function component there is a broadening due to incoherent contributions. Further, in order to reveal the individual exchange of photons with a single electron, and probe the quantum nature of the observed interaction, it is also needed to use highly coherent electrons (with low energy uncertainty) and a sensitive enough detector (see Ref. [27] for more details).

4.1c. The Quantum-Classical Correspondence in DLA

The above derivation for the extended PINEM theory, and in particular Eq. (43), captures well the experimental results of certain DLA interactions, specifically those in which the electron dispersion is negligible throughout the interaction with the field, and electron–electron interactions are neglected (so self-bunching effects are negligible). Its main advantage is that it reveals the quantum nature of the interaction, which is absent from the classical model so far used to describe DLA experiments. Furthermore, it can be used even for scenarios where the incoherent broadening smeared the quantum feature in the electron spectrum. In this section, we compare the classical formula for the electron spectrum to that of PINEM.

In the classical case, the electron is considered as a point particle subjected to the Lorentz force. Hence, neglecting relativistic effects, its velocity changes via

$$m_e\frac{dv}{dt}=q_e\Re[2\tilde{E}_z(z)e^{-i\omega_0 t}\ ], \tag{44}$$

where the z axis is parallel to the electron trajectory and we assume no deflections. Here $\tilde{E}_z(z)$ is the electric field phasor's z component, ω is the laser central frequency, and v is the electron velocity. To find how the electron velocity changes with time, we integrate Eq. (44) with respect to time to get

$$m_e\Delta v=q_e\int\Re[2\tilde{E}_z(z)e^{-i\omega_0 t}\ ]dt. \tag{45}$$

In most DLA experiments, it is reasonable to assume that the electron energy does not change much with respect to its initial energy, i.e., $v=\ v_e+\Delta v$ and $\Delta v\ll v_e$. In that case, the time step is approximately given by $t\approx(z-z_0)/v_e$, where z_0 is the starting point of the electron–DLA interaction. By changing variables, the change in the electron's kinetic energy is

$$\Delta E_k=q_e\int\Re\left[2\tilde{E}_z(z)e^{-i\frac{\omega_0}{v_e}(z-z_0)}\right]dz. \tag{46}$$

To show the correspondence of this classical derivation and PINEM, we define the quantity $W_g\overset{\Delta}{=}q_e\int\tilde{E}_z(z)exp(-i\omega_0 z/v_e)dz$ and note that $W_g=g\hbar\omega_0$, where g is the coupling strength of the conventional PINEM theory which is the time independent version of Eq. (40). With this definition we rewrite Eq. (46) as $\Delta E_k=\Re\left[W_g exp\left(i\frac{\omega}{v_e}z_0\right)\right]$.

We assume that the electron starts its interaction with the electromagnetic field in an arbitrary point in space, and this point is uniformly distributed over one optical cycle

of the laser field, i.e., $f_{\text{int.s.p.}}(z_0) \overset{A}{=} \text{Uniform}\left(0, \frac{2\pi v_e}{\omega_0}\right)$. Then, the probability density of the electron energy change for the classical case then reads

$$P_{\text{classical}}(\Delta E_k) = \frac{f_{\text{int.s.p.}}(z_0)}{\left|\frac{d\Delta E_k}{dz_0(\Delta E_k)}\right|} = \frac{1}{2\pi|W_g|\sqrt{1-\left(\frac{\Delta E_k}{2|W_g|}\right)^2}}. \tag{47}$$

To compare this result with the PINEM theory, we introduce a dimensionless *continuous* parameter ℓ that quantifies the electron energy change by $\Delta E_k = \ell\hbar\omega_0$, and write Eq. (47) as

$$P_{\text{classical}}(\ell) = \frac{f_{\text{int.s.p.}}(z_0)}{\left|\frac{d\Delta E_k}{dz_0(\Delta E_k)}\right|} = \frac{1}{2\pi|g|\hbar\omega_0\sqrt{1-\left(\frac{\ell}{2|g|}\right)^2}}. \tag{48}$$

This is the classical analog to the coherent probability that derived in Eq. (42).

In the PINEM framework, the coupling strength g appears as the argument of a series of Bessel functions which gives the electron spectrum its oscillatory profile. Furthermore, for the quantum case, the dimensionless parameter ℓ counts the number of photons that the electron exchanges with the field. Thus, for the quantum case ℓ is an integer, while for the classical case ℓ is a continuous variable.

4.1d. Measuring Quantum Features in DLA Experiments

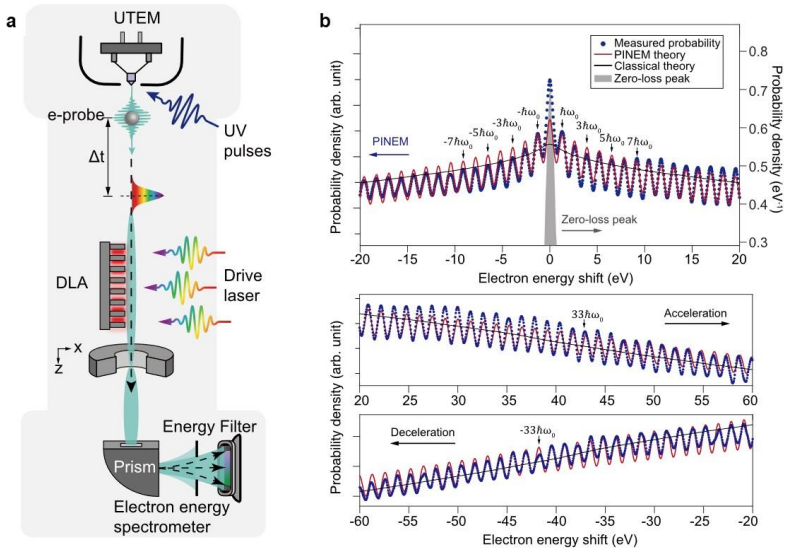
The theoretical discussion in Section 4.1.3 was demonstrated recently in a DLA experiment conducted within a UTEM. This experiment uncovered the multi-peak shape of the electron spectrum after its interaction with a DLA near-field, and boasts excellent agreement between the extended PINEM theory and the experimental results (even for a case where the quantum features are not apparent, see Fig. 24).

The full electron bandwidth after the DLA interaction is shown in Fig. 25. There, owing to the incoherent broadening of the electron and the lower spectrometer resolution needed to capture a wider energy range, the quantization of the electron spectrum is indistinguishable. Yet, the extended PINEM theory is useful even for a case where the quantum features are not apparent.

Interestingly, the interaction with the DLA caused the single electron wave function to split into several thousand distinct peaks of different energy in a coherent manner. In the energy domain, it means that the single electron occupies thousands of energy levels, simultaneously, and in the time domain it constitutes a comb of sub-attosecond pulses. This effect can be explained with the extended PINEM description (Eq. (43), which reproduce the observed spectrum with great agreement, however, the classical description (Eq. (48) can only follow the trend of the spectrum envelope. The quantum description of the electron provides additional degrees of freedom to the acceleration process, which is still yet to be explored.

Importantly, the above result shows that the electron–DLA interaction maintains quantum coherence for timescales of hundred laser cycles. This happens despite the presence of different interaction mechanisms that may lead to loss of coherency (e.g., bremsstrahlung, bulk excitations, and phonons). In addition, for strong interactions, the physics of quantum walk is modified by a gradual broadening of the electron energy levels due to the coherent uncertainty of the driving laser. This leads to an open question regarding whether quantum coherence is preserved also for strong interactions. Although the above result does not solve this issue, it shows that the quantum coherence remains even for the exchange of thousands of photons.

Figure 24

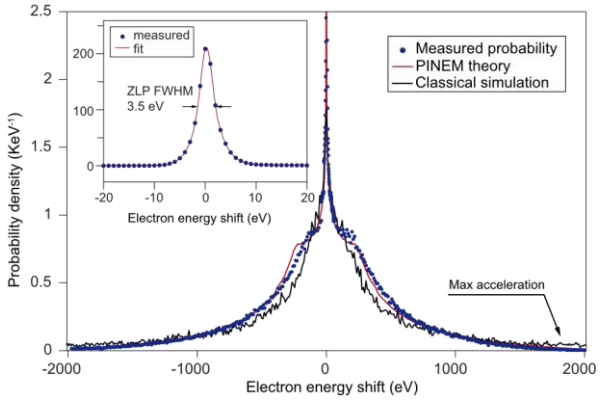


First DLA experiment in a UTEM. (a) Scheme of the experimental setup: a DLA formed by a single-sided grating is placed inside the UTEM and is pumped by a 950 nm femtosecond pulse in perpendicular incidence. A femtosecond electron pulse is than graze the DLA across its total length of 89 μm experienced acceleration and deceleration due to the interaction with the DLA near field. (b) Electron energy spectrum after DLA interaction. In (a) and (b) the oscillatory nature of the electron spectrum unravels when all the conditions to observe a PINEM-type interaction are fulfilled. The multi-peak shape corresponds to individual photons being emitted or absorbed by a single electron wave function. The PINEM theory (Eq. (43)) shows great agreement with the data, whereas the classical approach (Eq. (48)) can only follow the spectrum's envelope trend. Figure 1 (part (a)) reprinted with permission from Adiv *et al.*, Phys. Rev. X **11**, 041042 (2021), Ref. [27]. Copyright 2021 by the American Physical Society. Part (b) reprinted from [27] under a Creative Commons license.

Contemporary DLA schemes contain a pre-bunching stage for the electron beam [7,8]. Squeezing the electrons into attosecond bunches is highly valuable, because it can lead to net acceleration (or deceleration) and it opens the door to time-resolved atomic-scale dynamics [108,110,111]. In the quantum regime, bunching can also happen for a single electron pulse. As described previously, the single-electron wave function is altered by the DLA interaction where parts of it are accelerated while others are decelerated. This means that by letting the single-electron wave function propagate in free space there is a special distance at which the higher energetic parts of it catch up with the lower energetic components, resulting in a single free electron bunching. The quantum uncertainty plays a crucial role here as it applies a lower bound on the achievable duration of the electron bunch, which was studied theoretically in Ref. [112].

The developments in DLA research can be harnessed to serve as platforms for exploring other quantum phenomena in electron–light interactions. For instance, DLA was proved to sustain an efficient quasi-phase-matching interaction between the electron wave function and a light wave. This suggests that optimized structures may be able to mediate interactions between a single electron and a single photon [66]. Such an achievement would require the quantization of the electromagnetic field as well [113,114], and will open a way to conduct quantum optical experiments. In this

Figure 25



The electron wave function after the DLA interaction within the UTEM has a 4 keV bandwidth. The quantum features of the spectrum are hidden due to coarser resolution of the spectrometer, but the extended PINEM theory, when considering the incoherent broadening of the electron and the energy resolution, still gives a great predication of the measured data. Figure 3 reprinted with permission from Adiv *et al.*, Phys. Rev. X **11**, 041042 (2021) Ref. [27]. Copyright 2021 by the American Physical Society.

regard, it has been shown recently how to use advanced DLA designs to imprint photon statistics on a free electron and deduce the photon state by measuring the electrons [115,116].

4.2. Reading Coherent Information from a Quantum System

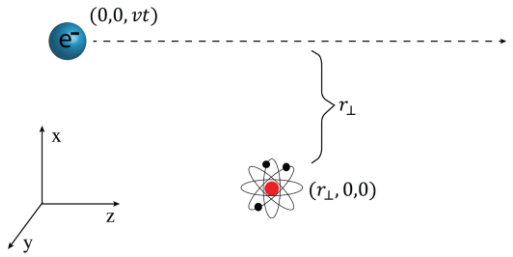
This section focuses on a recently proposed application that utilizes the quantum wave nature of the electrons in the DLA. The DLA serves to modulate the electron wave function, and this modulated electron can then be used to read and write quantum information in other quantum systems. The idea to use shaped electrons to control quantum state of matter was proposed first using a semiclassical model by Gover and Yariv [117], receiving the name "free-electron bound-electron resonant interaction" (FEBERI). The idea is that the temporal shape of the free electron's wave function can alter its coherent interaction with quantum two-level systems and induce Rabi oscillations, enabling control over the quantum state using free electrons. In that sense, the electron interaction is similar to how classical light can be used for coherent control [118]. Later, the idea was explored and a full quantum theory of the interaction was presented [119–121]. In addition, free electrons were proposed to measure the quantum state of qubits, measure decoherence times and to generate qubit-qubit and photon-photon entanglement [119,120,122]. These advances reveal the potential of free-electrons as novel and promising tools for manipulating quantum information, combining the high coherence of lasers and the high spatial and temporal resolution provided by free electrons.

We can model the electron–qubit interaction by the Hamiltonian [120]

$$H = -i\hbar v\partial_z + \frac{\hbar\omega_0}{2}\sigma_z + \frac{e}{4\pi\varepsilon_0}\cdot\frac{(\boldsymbol{d}_\perp\cdot\boldsymbol{r}_\perp + d_\parallel z)\sigma_+ + (\boldsymbol{d}_\perp^*\cdot\boldsymbol{r}_\perp + d_\parallel^* z)\sigma_-}{(r_\perp^2 + z^2)^{3/2}} \tag{49}$$

where the first two terms describe the Hamiltonians of the electron and the qubit. The average velocity of the electron wave packet is v. The qubit has an energy separation of $\hbar\omega_0$; $\sigma_\pm$, σ_z are Pauli matrices. The third term in Eq. (49) describes the interaction with the transition dipole moment $\boldsymbol{d} = \mathrm{g}|e\boldsymbol{r}|\mathrm{e}$, with components $\boldsymbol{d}_\parallel = \hat{\boldsymbol{z}}d_z$ and $\boldsymbol{d}_\perp = \hat{\boldsymbol{x}}d_x + \hat{\boldsymbol{y}}d_y$. The distance between the center of the electron wave packet and the center of the qubit

Figure 26



Schematic of the interaction between a free electron and an atom. The electron has coordinates $x = 0,\ \ y = 0$ and moves along the z axis with speed v so that $z = vt$. The atom is taken as stationary and situated at the origin $(r_\perp, 0, 0)$. Adapted from [120].

(the impact parameter) is $r_\perp$ (Fig. 26); the vacuum permittivity is ε_0. The two main approximations behind Eq. (49) are: (1) the paraxial approximation for the electron; it is valid because the energy of the electron is much larger than that of the excitation; (2) external decoherence channels by other material excitations such as Bremsstrahlung radiation [123] and characteristic x ray [124] occur at probabilities much smaller than unity. The latter condition can be made less strict by post-selecting electrons that lost energy $\pm\hbar\omega_0$, because some other transitions might also destroy the purity regardless of post-selection, depending on the specifics of the sample. Thus, the reduced density matrix of the electron and the qubit (for the excitation of interest) will be the same as if the external channels are not considered at all.

Assuming weak interaction, the electron–qubit scattering matrix can be found from the Magnus expansion [125] as

$$S = e^{-i(g\sigma^+ b + g^* b^+ \sigma^-)} = \cos|g| - i\cdot\sin|g|(e^{i\phi_g}\sigma_+ b + e^{-i\phi_g}\sigma_- b^+). \tag{50}$$

Here, the operators b and b^+ are momentum translation operators for the electron, defined as $b = e^{i\omega_0 z/v}$ [126] (for sufficiently fast electrons, they equivalently describe energy translation). The PINEM interaction coupling strength, $g = |g|e^{i\phi_g}$, can be approximated in the weak interaction limit as [120]

$$g = \frac{ed_x\omega_0 K_1\left(\frac{\omega_0 r_\perp}{\gamma v}\right)}{2\pi\varepsilon_0\gamma\hbar v^2} + i\frac{ed_z\omega_0 K_0\left(\frac{\omega_0 r_\perp}{\gamma v}\right)}{2\pi\varepsilon_0\gamma^2\hbar v^2}, \tag{51}$$

where $K_0(x)$ and $K_1(x)$ are the modified Bessel functions of the second kind and $\gamma = 1/\sqrt{1 - v^2/c^2}$ is the Lorentz factor. A more detailed derivation including higher-order corrections to g and the complete form of the scattering matrix can be found at [120].

After the scattering matrix has been established, the calculation of the scattering of any modulated electron from a qubit in any general state can be calculated. The quantum nature of the interaction can now be utilized to perform several tasks, such as controlling the state of the qubit [117], measuring coherence times and the full quantum state of the qubit [120], and even inducing entanglement between separated non-interacting qubits [119]. The ability to perform these tasks arises from the ability to use shaped electron wave functions. The latter was initially quite surprising, as it was debated whether the coherent shape of the free electron's wave function affects the interaction [117,121,127,128]. However, this turned out to be the case only when the quantum coherence vanishes, for example as happens in thermodynamical equilibrium. When the system maintains coherence (meaning, non-zero off diagonal elements in the system's density matrix exist) different components of the free electron's wave

function can interfere during the interaction, and this interference can lead to enhanced interaction. When exploited correctly, it can reveal desired information about the quantum system.

As an example, consider a qubit with energy separation $\hbar\omega_0$ described by a general density matrix ρ, and a free electron in the initial state $|\psi_e\rangle$. We can calculate the expectation value in the change of the free electron's energy after the interaction and get

$$\langle \Delta E \rangle = \hbar\omega_0 \sin^2|g|(\rho_{ee} - \rho_{gg}) - 2\hbar\omega_0 \cos|g| \sin|g| \mathrm{Im}\{\rho_{ge} e^{i\phi_g} \psi_e|b|\psi_e\}, \tag{52}$$

where ρ is the qubit's density matrix defined as $\rho_{ij} = \langle i|\psi\rangle\langle\psi|j\rangle$ for a pure state. The parameter $s_e \equiv \langle\psi_e|b|\psi_e\rangle$ will be called from now on *the bunching coefficient* of the electron. This parameter is expressed by the Fourier component of the electron's charge density, $b = \int \exp(i\omega z/v)\ |\psi(z)|^2 dz$, and is analogs to the classical first bunching coefficient in FELs [129,13]. At the typical limit where $|g| \ll 1$, Eq. (52) can be rewritten as

$$\frac{\Delta E}{\hbar\omega_0} = |g|^2(\rho_{ee} - \rho_{gg}) - 2|g|\mathrm{Im}\{\rho_{ge} e^{i\phi_g} s_e\}. \tag{53}$$

The first term is the incoherent one, and so does not depend on the shape of the initial free electron's wave function. We can see that if the system is in the ground state $(\rho_{gg} = 1, \rho_{ie} = 0)$, the electron experiences energy loss. If it is in the excited state $(\rho_{ee} = 1, \rho_{gi} = 0)$, the electron experiences energy gain. When the system is in superposition, its coherence, given by the second term, alters the final energy spectrum. If coherence is ignored, both processes are summed up in a weighted average for the final energy spectrum. If many qubits are present and there is no coherence between them, the first term will give us the population statistics.

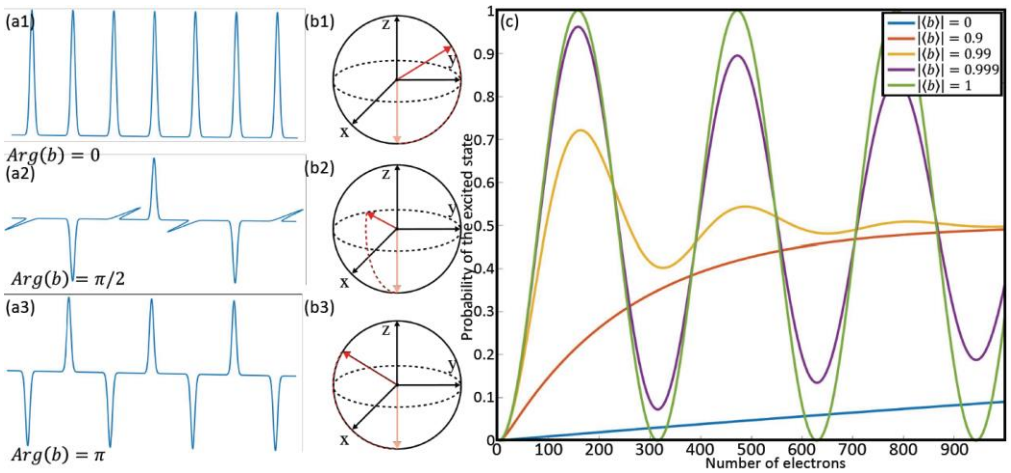
The second term contains the coherent information. We see that if the electron is shaped in such a way that $s_e \neq 0$, the off-diagonal elements can alter the average change in the electron's energy. This change is proportional to $|g|$, as opposed to $|g|^2$, which is the standard case for electrons that are not coherently modulated. This enhancement was first predicted by Ref. [117] using a semiclassical analysis. If our ability to control the free electron wave function were absolute, we could imagine engineering an electron with any s_e we want (under the constraint $|s_e| \leq 1$ resulting from the unitarity of b), allowing us to extract the value of ρ_{ge}, completely measuring the qubit's state and decoherence times [120].

To exemplify the importance of the electron's bunching parameter even further, another interesting question is to ask how the electron alters the qubit's state. It was already proposed that modulated electrons can be used to perform Rabi oscillations on the qubit state [117], with each interaction rotating the qubit's state around a vector on the x–y plane in the Bloch sphere by angle $2|g|$. However, the semi-classical analysis misses the fact that during the interaction, the electron and qubit get entangled. Hence, the electron carries some of the coherence of the qubit outside the system, effectively decohering the qubit state and damping the Rabi oscillations. One way to measure this is to look at the purity of the qubit after the interaction,

$$\mathrm{Tr}[\rho^2] = 1 + \frac{1}{2}(|s_e|^2 - 1)\sin^2 2|g|. \tag{54}$$

We can see that when the bunching parameter approaches unity, $s_e \to 1$, the post-interaction qubit's state stays pure, indicating that the electron and the qubit do not get entangled. In Fig. 27, we exemplify how different free-electron wave functions affect the qubit's state and how the bunching parameter changes the resulting Rabi

Figure 27



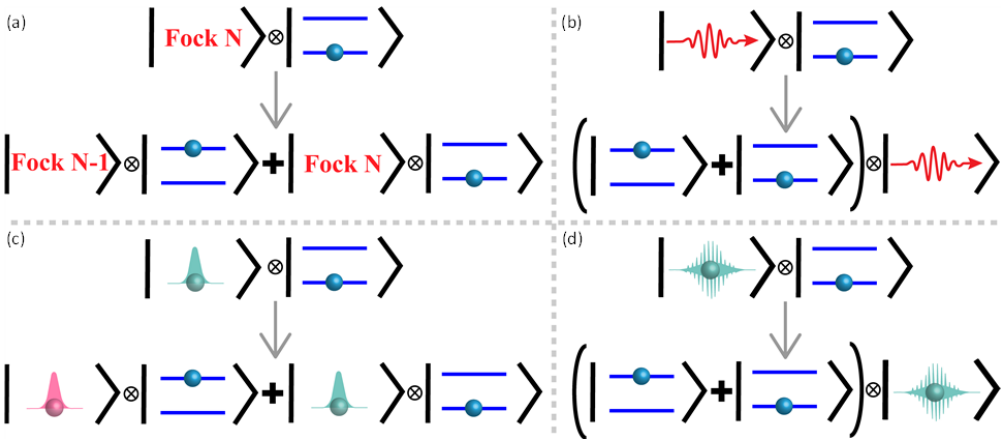
Coherent control using coherently shaped free electrons. (a) Electron's wave function in energy. The electron is modulated such that the different energy components have linear phase relation between them. The phase of the expectation value of the free-electron energy displacement operator b indicates the phase relation between the different free electron energy components. (b) When a series of such electrons interact with a two-level system, the state starts to rotate around the Bloch sphere in an angle dictated by the free-electron energy components' phase relation. Each electron in (a) relates to the adjacent rotation on the Bloch sphere in (b), where we assumed that the coupling constant g is real. (c) Probability of the two-level system to be excited as a function of number of electrons. It can be seen that when the comb is not perfect (meaning $|b|<1$), the oscillations decay until the state of the two-level system is completely mixed.

oscillations. We see that in order to perform coherent control on the qubit's state, we need an electron with a bunching parameter approaching unity.

It is already known that to perform quantum coherent tasks (such as control over qubit states and to perform quantum measurements), a coherent light (i.e., Glauber state) is needed [118,130]. This can be understood from the fact that coherent light is an eigenstate of the photonic annihilation operator (a). This implies that the light does not change significantly when single photons are emitted or absorbed, or in a more mathematical way, up to a constant, $|\psi_{coherent} \approx a|\psi_{\text{coherent}}$. In the context of coherent control, the state of the material qubits is altered due to exchange of energy between the light and the material's excitation. If the light would change significantly when a photon is absorbed by the qubit, the coherence of the qubit will be lost, as the excited qubit state will be entangled to a very different light state than the ground qubit state (Figs. 28(a) and 28(b)). An analogous thing happens with free electrons (Figs. 28(c) and 28(d)), however, instead of photon creation and annihilation operators, the free electron can alter its state by exchanging quanta's of energy with the material and translating up and down the energy ladder. The best analog for coherent light will be an electron wave function which does not change significantly when such translation is acted on it. Such an electron will be a coherent energy comb of the form:

$$|\psi_{comb} = \sum_n e^{in\phi}|E_0 - n\hbar\omega,\ b|\psi_{comb} = e^{i\phi}|\psi_{comb}. \tag{55}$$

Figure 28



Quantum description of qubit interactions. When a quantum probe such as (a) photons or (c) monoenergetic electrons interact with a qubit, they entangle and exchange quantum information. As a result, different qubit states get entangled into different probe states, washing out the coherent information of the qubit. More complicated probes, such as (b) coherent light and (d) shaped electrons, can potentially alter the state of the qubit without entangling to it, enabling us to read and write the quantum state of the qubit. To understand what quantum information is exchanged and how the interaction can be exploited to read/write quantum states, a fully quantum mechanical treatment is required. For photons, this description is based on the well-known Jaynes–Cumming model and its extensions [109]. For electrons, the analogous description has only been developed recently [119,120], and many of its aspects are still being investigated.

This kind of electron cannot be truly realized, but a good way to quantify how close the electron is to be fully bunched is by the moments of the ladder operators (i.e., $|b^n|$), which will approach unity as we get close to the ideal case.

Mathematically, this point can be realized by acting with our qubit–electron or photon–electron scattering matrix on the comb-electron state. As $|\psi_{comb}\rangle$ is an eigenstate of b and b^+ (with eigenvalues $e^{i\phi_e}$ and $e^{-i\phi_e}$ accordingly), it is also an eigenstate of the scattering matrices, and so the electron–photon [113] and the electron–qubit scattering matrices can be written alternatively as

$$S_{qubit} = e^{-i|g|(e^{i(\phi_g+\phi_e)}\sigma^+ + e^{-i(\phi_g+\phi_e)}\sigma^-)} = R_{(\cos(\phi_g+\phi_e),\sin(\phi_g+\phi_e),0)}(2|g|),$$

$$S_{photon} = e^{|g_Q|(e^{i(\phi_{g_Q}+\phi_e)}a^+ - e^{-i(\phi_{g_Q}+\phi_e)}a)} = D(|g_Q|e^{i(\phi_{g_Q}+\phi_e)}). \quad (56)$$

We see that when the electron is a comb, the scattering between the electron and qubit translates into a rotation around the x–y plane in the Bloch sphere (R is the rotation matrix) with an angle $2|g|$, and the scattering between the electron and photons translate into a displacement of the light state by $|g_Q|$ (D is the displacement operator [118]). Therefore, the comb electron allows us to manipulate light states and qubits in a coherent and controllable way, revealing why we expect to see Rabi oscillations on the qubit and showing us that the radiation emitted by such electron is expected to be coherent [131].

If the electron's energy width is smaller than the qubit's energy separation, s_e will clearly be zero. This can be understood classically because the electron's temporal duration will be longer than the phase cycle of the quantum two level system, and

so all the coherent phase information will be washed out during the interaction. One might suggest to change s_e using electron energy shaping with PINEM interactions [18,19,99,110], however, this is irrelevant since the PINEM scattering matrix commute with the energy ladder operators. To increase the value of s_e, a temporal shaping must be performed on the electron, compressing it in time into sub-femtosecond duration.

One way to temporally shape the free electron's wave function (see Fig. 29) is to combine energy shaping (using methods such as PINEM) with free space propagation (FSP), i.e., "drift" [26]. The dispersion relation of the electron's energy will result in different energy components accumulating phases at different rates, changing the temporal shape of the free electron's wave function. The simplest way to generate such a "comb" electron is by using a single PINEM interaction with frequency ω_0 and some interaction strength g, and then letting the electron propagate freely for some distance. It is shown that the value of $|s_e|$ achieved using this technique is [119]

$$|s_e| = \left| J_1 \left[4|g| \sin\left(2\pi \frac{z}{z_T} \right) \right] \right|, \tag{57}$$

where z is the propagation distance, z_T is the Talbot distance [132] and J_1 is the first Bessel function of the first kind. Consequently, the optimal value of $|s_e|$ achieved by this technique will be the maximal value achieved by J_1, which is $|s_e|_{optimal} \approx 0.58$. Optimal bunching coefficient using one interaction site has been already realized in classical DLA experimentally [7,8]. Interestingly, in the classical case, the optimal drift length in order to maximize the bunching coefficient is different for small values of g (Fig. 30(d)).

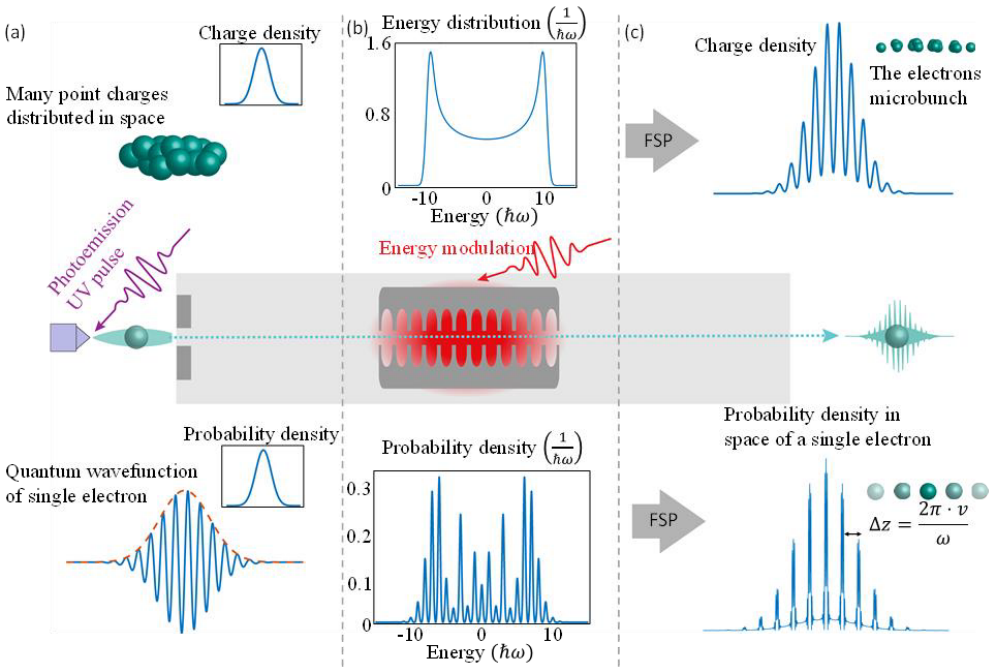
It is important to note that the optimal bunching distance from a classical point of view is related to, but different than, the resulting one from a quantum point of view, as discussed and compared in the following.

The scheme presented previously (achieving $|s_e|_{optimal} \approx 0.58$) is still not sufficient for useful, coherent control using free electrons. Stronger attosecond compression of the free electron's wave function is needed. Such attosecond compression of the electrons has been first proposed using an optimized multi-harmonic modulating laser pulse (before the FSP) [134]. More recent works [11,133] proposed a way to compress electrons by using two PINEM interactions separated by separate stages of FSP. This scheme was further optimized with three points of interaction and an additional stage of FSP to reach $|s_e|_{optimal} \approx 0.99$ (see Ref. [133]), thus generating approximated electron combs. In another recent approach described in Section 3.4, Zhao *et al.* [11] suggested to compress electrons using the beat-note of two optical frequencies. Overall, more work is needed to resolve the quantum properties of such generated electrons and quantify their $|s_e|_{optimal}$. The resulting high temporal compression of the free electron's wave function translates into a large spread in the electron's energy spectra, which is expected to push $|s_e|$ closer to 1.

Next, we focus on Ref. [133], which, so far, best exemplifies the ability to bunch the quantum free electron's wave function. This paper proposes the use of weak preparation pulses before the main compression stage. The technique requires using optimization to select the drift distances and the interaction parameter for each PINEM interaction. Using sufficient preparation pulses, one might be able to push both $|s_e|$ and even higher moments of $|b^n|$ to be arbitrarily close to unity, enabling the vision of using electrons to generate and measure quantum states of material's excitations.

To exemplify the quality of the electrons generated by such a technique, we can look at the radiation emitted by them. As shown previously, perfect combs interacting with an empty cavity will generate coherent states of light with an average of $|g_Q|^2$ photons,

Figure 29

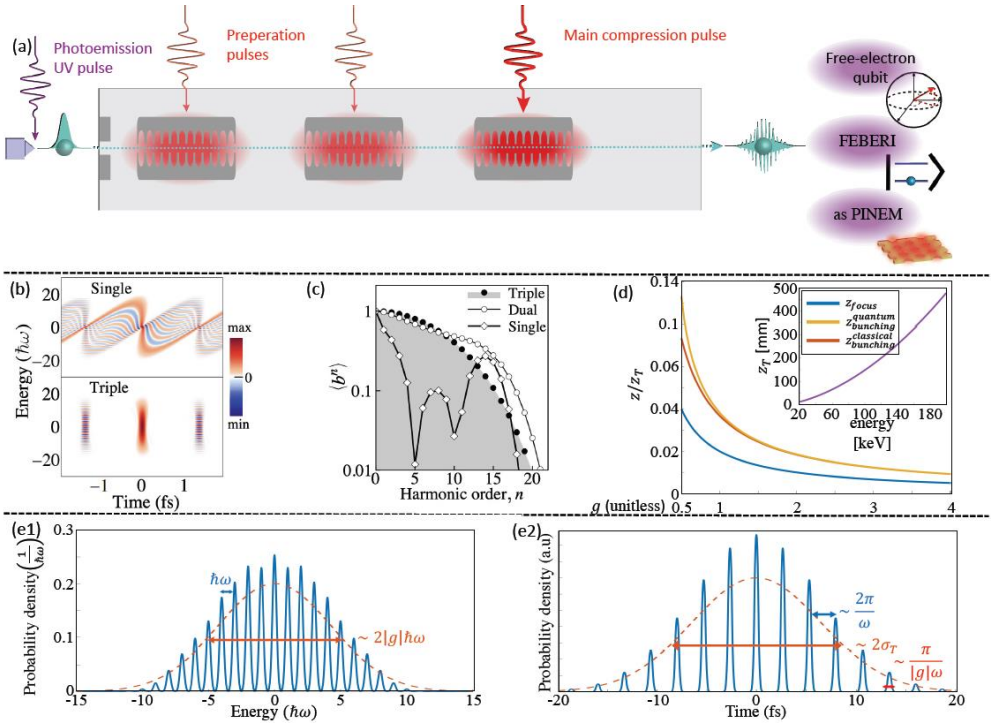


Classical (top) versus quantum (bottom) description of the electron's temporal shaping. (a) Electrons are emitted from an electron gun via a photoemission process. Classically the electrons are point particles distributed in space according to some charge distribution, which is typically much wider than optical wavelengths. Quantum mechanically, each electron is smeared in space according to its wave function. (b) Energy modulation is performed on the electrons via PINEM-like processes. Classically, each electron sees a different phase of the modulating laser, while quantum mechanically each electron sees many periods of the laser and interferes with it. Classically, each electron experiences a different Lorenz force and therefore gains (or loses) a different amount of energy, while quantum mechanically each electron spreads in energy. The resulting quantum energy distribution constitutes of discrete peaks, resulting from the wave interference between the electron and the light. For this figure, $g = 4$ was chosen for the interaction strength. (c) Electrons propagate and eventually get micro bunched to bunches separated in time by one cycle of the modulating field. In the quantum picture each electron gets bunched, resulting in a probability density with very narrow features.

where g_Q is the interaction strength with the cavity [113]. When the combs are not perfect, one can still calculate the resulting electron-light state. Then the light state after tracing out the electron, so the state of light if the electron is not measured and information regarding the electron's state is "lost," is generally a mixed state. This state can be compared with the desired coherent state using measures such as the fidelity between the states [135]. We assume a general electron state $|\psi_e\rangle$ interacting with an empty cavity $|0\rangle$ with coupling constant g_Q. We then trace out the electron and stay with a partially mixed light state described by a density matrix σ. The fidelity between the light state and the desired coherent state is then given by $F(\sigma) = g_Q|\sigma|g_Q$. We define F_n as the fidelity of the resulting light state with $g_Q = 1$ when the electron is created using n PINEM interactions with the parameters taken from [136]. The results are striking:

$$F_0 \sim 0.3,\ F_1 \sim 0.66,\ F_2 \sim 0.95,\ F_3 \sim 0.98. \tag{58}$$

Figure 30

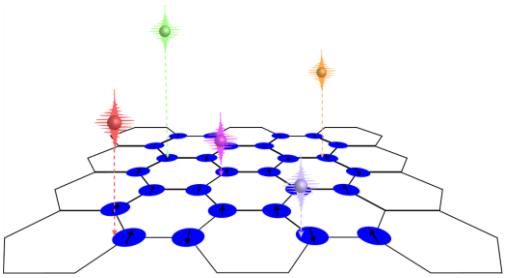


Attosecond-scale compression of the electron using multiple interaction points with a modulating laser. (a) The electron is shaped using multiple-stage DLA, where the strength of the interactions and the distance between them is tuned to result in optimal bunching. The resulting bunched electron is then fit to perform FEBERI interactions. In this figure the interaction strength for the main compression is chosen to be $g = 4$, and the preparation pulses are optimized accordingly to result in optimal bunching. (b) Wigner distributions in the focus for single point of interaction (top) and three points of interaction (bottom), showing a localization of electrons in phase space. (c) Values of $\langle b^n \rangle$ obtained at the optimal bunching distance. (d) Important distances for electron shaping: z_{focus} is the distance in which the wave packet become maximally narrow [112], and $z_{bunching}$ is the distance in which the bunching coefficient becomes maximal; for small values of g this distance differs between the classical [13,129] and the quantum [119] case. (e) Resulting probability distribution of the electron in (e1) energy and (e2) time, at the optimal bunching distance after three points of interactions. (b) and (c) Reprinted from [133] under a Creative Commons license.

It is clear that already one preparation pulse (2 PINEM interactions, F_2), generates a good approximation of coherent light. Two-stage DLAs are already in use today [7,8] and can be exploited to demonstrate such abilities.

Concluding this section, DLA contains all the building blocks to read and write the quantum state of coherent systems. To do that, we combine the quantum theory for the interaction between quantum systems and modulated electrons with the ability to create attosecond compressed electron pulses. Such capabilities, especially if achieved at deep subwavelength and potentially atomic resolutions, are attractive for creating new types of quantum simulators (Fig. 31). The vision for this pursuit, as presented in Ref. [120], is to develop quantum simulators in which shaped electrons enable the depiction of the initial state of each element and allow to read the final (or intermediate) states using femtosecond (and eventually attosecond [20–22,108,136]) time resolution.

Figure 31



Vision for an electron-based quantum simulator. The electron is used as the probe to both read and write the state of individual qubits, enabling to control the state of the simulator on spatial scales much smaller than these achievable using laser-based probes.

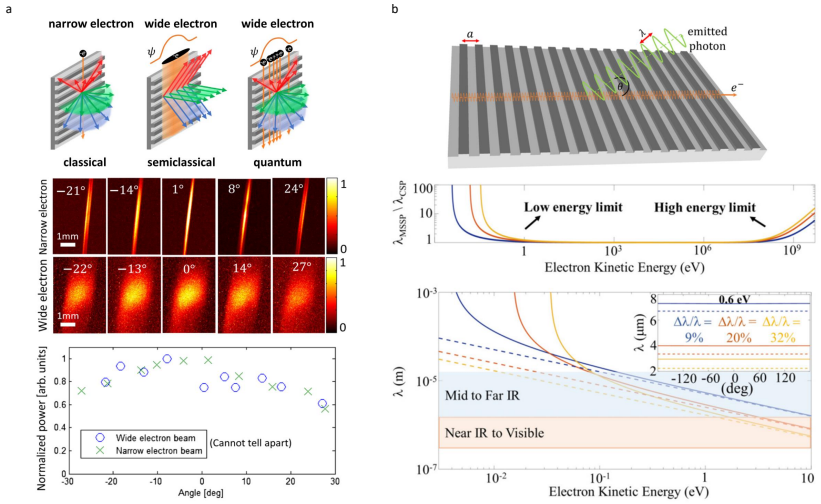
4.3. SP and the Electron Wave Function in Periodic Structures

The onset of the quantum regime in spontaneous and stimulated electron–photon interactions could be mainly attributed to three different phenomena: (1) quantum decoherence and quantum interference of light emitted by single electron wave functions [122–144]; (2) sensitivity of the electron wave function to the incident quantum photon statistics [113,114,116,133,135,145]; and (3) quantum corrections to the emitted spectrum associated with the recoil exerted on the free electron by the emission of photon quanta [146–148]. To observe the first type of effect, the electron wave function ought to have large spatial (transversal or longitudinal) coherence on the scale of the optical wavelength. The second family of effects requires strong quantum coupling between light and free electrons, while ensuring single-mode operation [116], thus ensuring that the nanostructure effectively supports one spatial mode at the relevant spectral regime. Lastly, the third effect, which has only recently been reported experimentally [149], suggests using either slow electrons or periodic structures with very small periodicities [147].

Periodic structures driven by free electrons in electron microscopes offer a fertile ground for the observation and tests of quantum interactions, thanks to the versatility of the SP effect [150] for controlling the spectrum, emission directionality, and even polarization [151–162], dictated by the structure periodicity. Quasi-phase-matching further allows mode selectivity, which is found important for probing approximately single-mode quantum effects [116]. In addition, periodic structures could be amended with effective cavities that enhance the near field, based on inverse design [1,27,116], which can increase the quantum coupling strength between light and free electrons.

In TEMs, the size of the electron's transverse wave function could be increased far beyond the scale of a single optical wavelength, while maintaining high spatial coherence [163,164]. Further, ultrafast TEMs allow tens of millielectronvolts photoemitted electron coherent energy spreads to be measured, determined by the laser linewidth [165], whereas modern monochromators allow energy resolutions down to the single millielectronvolt scale [166]. These properties allow for large longitudinal coherence of the electron wave function, extending far beyond a single radiation wavelength. Finally, EELS, frequently used in TEMs [167] and recently also realized in SEMs for PINEM experiments [25], can be employed as a detection technique for resolving spontaneous and stimulated quantum effects [168]. Stimulated interactions can be cascaded with spontaneous emission, wherein the electron wave function is pre-formed by the incident laser light and then allowed to emit spontaneously, testing its effect on the emitted light [137,143,147].

Figure 32



Quantum regime of spontaneous free-electron light emission in periodic structures. (a) Testing spatial quantum decoherence of SP radiation from a transversely wide electron, comparing the classical, semiclassical, and quantum approaches. The experimental results (diverging far-field SP radiation irrespective of the wave function shape) are in favor of the quantum approach. Figure 1 reprinted with permission from Remez *et al.*, Phys. Rev. Lett. **123**, 060401 (2019), Ref. [139] ©. 2019 American Physical Society. (b) Quantum recoil corrections to SP radiation; shown on the right is the divergence from the classical model for slow electrons. Figure 3 reprinted with permission from Tsesses *et al.*, Phys. Rev. A **95**, 013832 (2017), Ref. [147]. Copyright 2017 by the American Physical Society.

4.3a. Spontaneous Quantum Interactions

The interpretations of the wave function of quantum particles have been addressed by many researchers over the history of quantum mechanics. When Schrödinger first tried to make sense of the solutions to his equations for the hydrogen atom, he interpreted the modulus squared of the wave function $|\psi|^2$ as a classical charge density [169,170]. Owing to the inconsistencies arising from this conjecture when applied to *bound* electron systems, this idea was quickly dismissed in favor of the probabilistic approach, where $|\psi|^2$ is interpreted instead as a probability density [170,171].

Surprisingly, for *free* electron systems the interpretation of $|\psi|^2$ as a charge density does not lead to such obvious inconsistencies, and was thus favored by many researchers as it greatly simplifies the semiclassical models of free-electron radiation [158,172–176] (Fig. 32(b)). Further, it is relatively accurate for the treatment of stimulated processes [172] and many-electron superradiant emission [156,177,178]. The latter generally holds except for cases in which free-electron entanglement is present [179], or for spontaneous emission from electrons with a small energy uncertainty compared with the photonic recoil, as in the quantum FEL regime [122,180–183].

Going beyond the semiclassical regime of spontaneous emission from a quantum free electron necessitates a careful quantum treatment. For example, to examine the effect of the wave function shape (longitudinal and/or transverse) on the emitted radiation, a quantum electrodynamical derivation is usually utilized. Ritchie and Howie [141] analyzed this scenario in EELS detection, showing no wave function dependence when all electron momenta are collected. Friedman *et al.* [140] showed that, for electrons

spontaneously emitting light in periodic structures, no dependence of the emitted power spectrum on the longitudinal wave function is to be expected.

More recent works investigated the breakdown of the semiclassical interpretation of the free-electron wave function. Pan and Gover [137,138,176] analyzed the problem of spontaneous (as well as stimulated) SP emission from pre-shaped electrons in both the semiclassical and quantum regimes, resulting in different predictions regarding the dependence on the wave function. Remez *et al.* [139] have shown that these different predictions could be told apart experimentally by observing the transverse far-field divergence of SP radiation, and observed results that are in favor of the quantum probability distribution interpretation (Fig. 32(a)). Their results were also consistent with the expectation of no wave function dependence (under the paraxial approximation and assuming a transversely uniform grating). While no dependence on the wave function is expected when observing the radiated *power spectrum*, recent works by Karnieli *et al.* [143] and Kfir *et al.* [147] have shown that the *spectral coherence* of the spontaneously emitted light does depend on the wave function. Further, Wong *et al.* [184] have calculated further corrections beyond the paraxial approximation and for non-uniform media, that allow wave function dependence and interference effects. Di Giulio *et al.* [122,185] investigated the interface between the spontaneous and stimulated regimes of free-electron radiation, showing the emergence of wave-function-dependent interactions whenever an external field is involved.

Finally, the quantum corrections associated with photonic recoil, first investigated by Ginzburg for Cherenkov radiation [186], and famously dismissed as unimportant by his PhD advisor Landau [146], can also be potentially observed using quantum spontaneous emission by free electrons in periodic structures, as investigated by Tsesses *et al.* [147]. In their paper, Tsesses *et al.* [147] calculated quantum corrections to SP radiation from electrons traversing a nanometric grating, and from electrons traversing two-dimensional periodic gratings, and analyzed the regimes at which the corrections could potentially be resolved (Fig. 32(b)). This effect was recently reported to have been demonstrated experimentally for free-electron x ray emission from van der Waals materials [149].

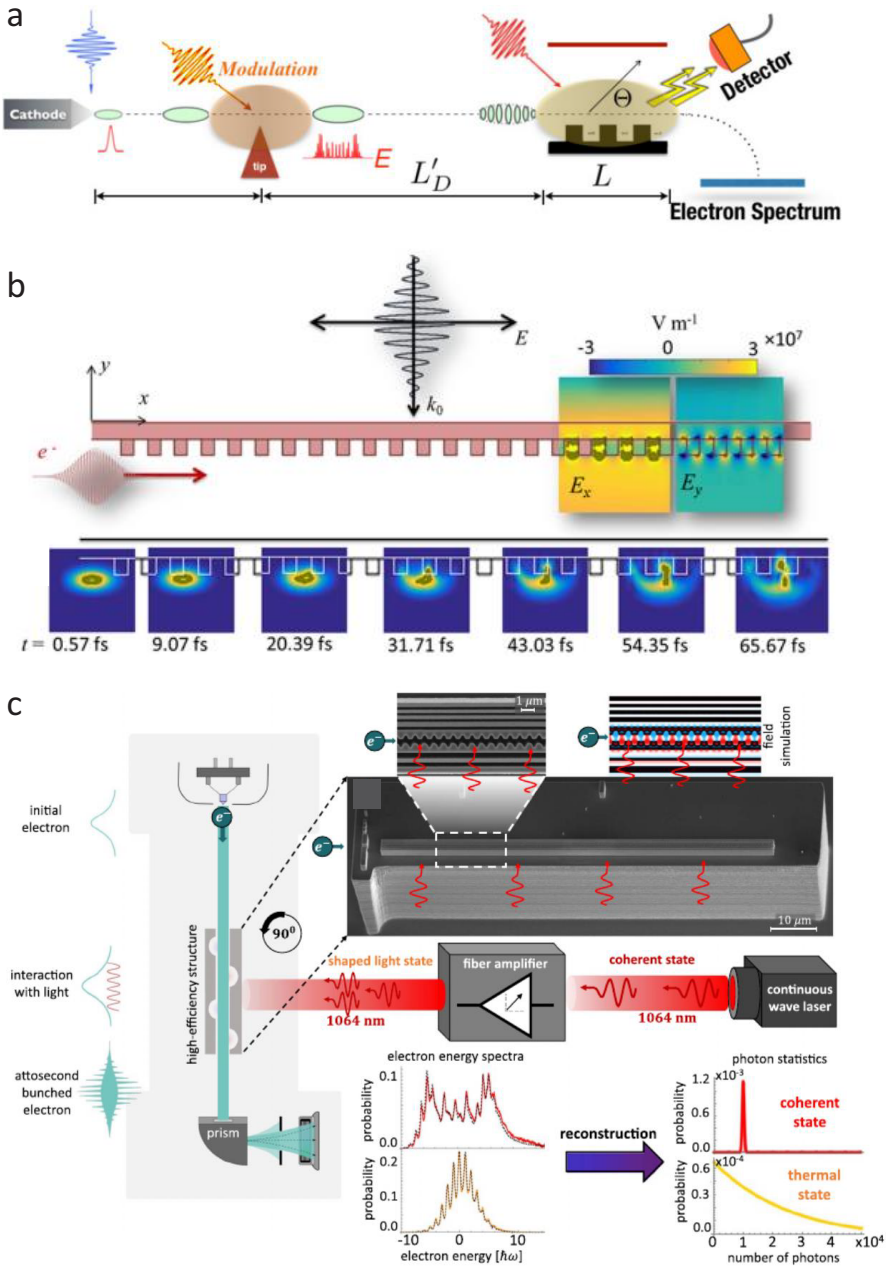
4.3b. Stimulated Quantum Interactions

Much interest has been devoted to the study of the quantum nature of stimulated free-electron–light interactions. In periodic structures, many authors focused on the inverse SP effect [58], the quantum FEL [183,184,187] and, recently, the quantum DLA [27], motivated by corrections to the stimulated emission and absorption processes owing to the wave nature of the electron. Such examples are the studies by Pan and Gover [137] and Talebi [172] (Fig. 33(a),(b)). Here onward, however, we primarily focus on the fully quantum interactions of free electrons and quantum light, which can be experimentally probed using periodic structures [116].

PINEM, as discussed in detail in Section 4.1.2, has seen dramatic advance. Recent publications include achievements in both the coupling efficiency [21,188], diversity of near-field structures [22,23], and the quantum coherence of the electron wave function [19,20,108,189,190]. The classical PINEM regime is restricted to interactions between quantized electrons with classical light, namely, a Glauber coherent state, which renders the modeling of such processes compatible with a time-dependent Schrödinger equation, where one plugs in the classical field as a time-varying potential term that drives the interaction [134].

Very recently, a surge of fruitful research emerged in the regime of quantum PINEM (QPINEM), wherein the quantum state of light becomes important for the correct description of the PINEM interaction. Kfir [113] theoretically showed that whenever

Figure 33



Quantum regime of stimulated free-electron light emission in periodic structures. (a) Proposal for testing the effect of the wave function shape (modulated first by the inverse-SP effect) on a successive spontaneous and stimulated SP radiation. Figure 1 reprinted with permission from Pan and Gover, Phys. Rev. A **99**, 052107 (2019), Ref. [137]. Copyright 2019 by the American Physical Society. (b) Semiclassical Maxwell–Schrödinger simulations for the inverse SP effect. Reprinted from [172] under a Creative Commons license. (c) Quantum PINEM experiment employing an inverse-designed periodic structure. The incident CW light changes its photon statistics from thermal to Poissonian, and the EELS signal shows a direct dependence on the quantum photon statistics, as predicted by the QPINEM theory. From Dahan *et al.*, Science **373**, eabj7128 (2021) [116]. Reprinted with permission from AAAS.

a free electron interacts with a single-mode cavity, it gets entangled with the photons, resulting in non-trivial correlations. Soon after, Di Giulio *et al.* [114,135] calculated the EELS signal observed following a QPINEM interaction, showing its dependence on the exact photon statistics hosted in the cavity. Gorlach *et al.* [145] then proposed a scheme for quantum optical detection using this principle, showing that the entire photonic quantum state can be extracted from a novel free-electron homodyne-type interaction. Ben Hayun *et al.* [133] further proposed to use free electrons to generate nontrivial quantum states of light using the QPINEM interaction.

To observe quantum photon statistics effects, one must significantly increase the value of the dimensionless quantum coupling constant g_Q, close to unity. This has been recently achieved experimentally using a 2D plasmonic Cherenkov scheme [188], agreeing well with the spontaneous QPINEM theory. A similar result was demonstrated for spontaneous emission of single photons into an integrated optical cavity [191]. Given a large enough g_Q (even if it is still two orders of magnitude below unity), one can create interactions with continuous-wave (CW) light, which has a much smaller incident field amplitude than a femtosecond pulse. Such an interaction enables probing light with non-trivial quantum statistics, i.e., states which differ from a classical coherent Glauber state [192] of laser light.

Recent experiments used a DLA structure [116] and a high-Q micro-resonator [24] to implement CW PINEM [193–196]. In this context, the use of periodic structures, and in particular inverse-designed structures [1,27,116], can boost the quantum coupling. Lastly, because the QPINEM effect requires mainly a single-mode operation [113], restrictions are posed on the optical structure (waveguides or microcavities) to ensure this. The use of highly selective quasi-phase-matching in periodic structures can relatively easily ensure a single-mode operation in a stimulated process [116].

The first and currently only experimental demonstration of stimulated QPINEM was reported by Dahan *et al.* [116] (Fig. 33(c)), using the inverse SP effect in a DLA, stimulated by amplified CW light, continuously shifting its quantum photon statistics from a Glauber coherent state to a thermal state. The 80 μm inverse-designed periodic structure allowed enough enhancement for low-intensity CW operation with quantum light, as well as nearly single-mode operation thanks to quasi-phase-matching selectivity. In their work, Dahan *et al.* [116] observed the continuous transition of quantum to classical random walk of the free electron on the discrete energy ladder, owing solely to the change in photon statistics, and in agreement with the stimulated QPINEM prediction.

FUNDING

Israel Science Foundation (830/19); HORIZON EUROPE European Research Council (AccelOnChip); Gordon and Betty Moore Foundation (4744).

DISCLOSURES

The authors declare that there are no conflicts of interest related to this article.

DATA AVAILABILITY

No data were generated or analyzed in the presented research.

REFERENCES

1. N. V. Sapra, K. Y. Yang, D. Vercruysse, K. J. Leedle, D. S. Black, R. J. England, L. Su, R. Trivedi, Y. Miao, O. Solgaard, R. L. Byer, and J. Vučković, "On-chip integrated laser-driven particle accelerator," Science **367**, 79–83 (2020).

2. S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, and A. W. Rodriguez, "Inverse design in nanophotonics," Nat. Photonics **12**, 659–670 (2018).
3. D. Cesar, S. Custodio, J. Maxson, P. Musumeci, X. Shen, E. Threlkeld, R. J. England, A. Hanuka, I. V. Makasyuk, E. A. Peralta, K. P. Wootton, and Z. Wu, "High-field nonlinear optical response and phase control in a dielectric laser accelerator," Commun. Phys. **1**, 46 (2018).
4. T. P. Wangler, *RF Linear Accelerators* (Wiley, 2008).
5. U. Niedermayer, T. Egenolf, O. Boine-Frankenheim, and P. Hommelhoff, "Alternating-phase focusing for dielectric-laser acceleration," Phys. Rev. Lett. **121**, 214801 (2018).
6. R. Shiloh, J. Illmer, T. Chlouba, P. Yousefi, N. Schönenberger, U. Niedermayer, A. Mittelbach, and P. Hommelhoff, "Electron phase space control in photonic chip-based particle acceleration," Nature **597**, 498–502 (2021).
7. N. Schönenberger, A. Mittelbach, P. Yousefi, J. McNeur, U. Niedermayer, and P. Hommelhoff, "Generation and characterization of attosecond microbunched electron pulse trains via dielectric laser acceleration," Phys. Rev. Lett. **123**, 264803 (2019).
8. D. S. Black, U. Niedermayer, Y. Miao, Z. Zhao, O. Solgaard, R. L. Byer, and K. J. Leedle, "Net acceleration and direct measurement of attosecond electron pulses in a silicon dielectric laser accelerator," Phys. Rev. Lett. **123**, 264802 (2019).
9. K. J. Leedle, D. S. Black, Y. Miao, K. E. Urbanek, A. Ceballos, H. Deng, J. S. Harris, O. Solgaard, and R. L. Byer, "Phase-dependent dielectric laser acceleration of 99 keV electrons with symmetrically driven silicon dual pillar gratings," *2018 Conf. Lasers Electro-Optics, CLEO 2018 - Proc.* 43, 2181–2184 (2018).
10. D. S. Black, Z. Zhao, K. J. Leedle, Y. Miao, R. L. Byer, S. Fan, and O. Solgaard, "Operating modes of dual-grating dielectric laser accelerators," Phys. Rev. Accel. Beams **23**, 114001 (2020).
11. Z. Zhao, K. J. Leedle, D. S. Black, O. Solgaard, R. L. Byer, and S. Fan, "Electron pulse compression with optical beat note," Phys. Rev. Lett. **127**, 164802 (2021).
12. T. Plettner and R. L. Byer, "Proposed dielectric-based microstructure laser-driven undulator," Phys. Rev. Spec. Top.–Accel. Beams **11**, 030704 (2008).
13. R. J. England, R. J. Noble, and K. Bane, *et al.*, "Dielectric laser accelerators," Rev. Mod. Phys. **86**, 1337–1389 (2014).
14. K. P. Wootton, J. McNeur, and K. J. Leedle, "Dielectric laser accelerators: designs, experiments, and applications," Rev. Accel. Sci. Technol. **09**, 105–126 (2016).
15. R. J. England, U. Niedermayer, L. Schächter, T. Hughes, P. Musumeci, R. K. Li, and W. D. Kimura, "Considerations for a TeV collider based on dielectric laser accelerators," J. Instrum. **17**, P05012 (2022).
16. A. H. Zewail and J. M. Thomas, *4D Electron Microscopy* (Imperial College Press, 2009).
17. A. H. Zewail, "Four-dimensional electron microscopy," Science **328**, 187–193 (2010).
18. B. Barwick, D. J. Flannigan, and A. H. Zewail, "Photon-induced near-field electron microscopy," Nature **462**, 902–906 (2009).
19. A. Feist, K. E. Echternkamp, J. Schauss, S. V. Yalunin, S. Schäfer, and C. Ropers, "Quantum coherent optical phase modulation in an ultrafast transmission electron microscope," Nature **521**, 200–203 (2015).
20. G. M. Vanacore, I. Madan, G. Berruto, K. Wang, E. Pomarico, R. J. Lamb, D. McGrouther, I. Kaminer, B. Barwick, F. Javier García de Abajo, and F. Carbone, "Attosecond coherent control of free-electron wave functions using semi-infinite light fields," Nat. Commun. **9**, 2694 (2018).

21. R. Dahan, S. Nehemia, M. Shentcis, O. Reinhardt, Y. Adiv, X. Shi, O. Be'er, M. H. Lynch, Y. Kurman, K. Wang, and I. Kaminer, "Resonant phase-matching between a light wave and a free-electron wavefunction," Nat. Phys. **16**, 1123–1131 (2020).
22. K. Wang, R. Dahan, M. Shentcis, Y. Kauffmann, A. Ben Hayun, O. Reinhardt, S. Tsesses, I. Kaminer, A. Ben Hayun, O. Reinhardt, S. Tsesses, and I. Kaminer, "Coherent interaction between free electrons and a photonic cavity," Nature **582**, 50–54 (2020).
23. O. Kfir, H. Lourenço-Martins, G. Storeck, M. Sivis, T. R. Harvey, T. J. Kippenberg, A. Feist, and C. Ropers, "Controlling free electrons with optical whispering-gallery modes," Nature **582**, 46–49 (2020).
24. J.-W. Henke, A. S. Raja, A. Feist, G. Huang, G. Arend, Y. Yang, J. Kappert, R. N. Wang, M. Möller, J. Pan, J. Liu, O. Kfir, C. Ropers, and T. J. Kippenberg, "Integrated photonics enables continuous-beam electron phase modulation," Nature **600**, 653–658 (2021).
25. R. Shiloh, T. Chlouba, and P. Hommelhoff, "Quantum-coherent light-electron interaction in a scanning electron microscope," Phys. Rev. Lett. **128**, 235301 (2022).
26. O. Reinhardt, C. Mechel, M. Lynch, and I. Kaminer, "Free-electron qubits," Ann. Phys. **533**, 2000254 (2021).
27. Y. Adiv, K. Wang, R. Dahan, P. Broaddus, Y. Miao, D. Black, K. Leedle, R. L. Byer, O. Solgaard, R. J. England, I. Kaminer, J. England, and I. Kaminer, "Quantum nature of dielectric laser accelerators," Phys. Rev. X **11**, 041042 (2021).
28. M. Kozák, J. McNeur, N. Schönenberger, J. Illmer, A. Li, A. Tafel, P. Yousefi, T. Eckstein, and P. Hommelhoff, "Ultrafast scanning electron microscope applied for studying the interaction between free electrons and optical near-fields of periodic nanostructures," J. Appl. Phys. **124**, 023104 (2018).
29. A. Feist, N. Bach, N. Rubiano da Silva, T. Danz, M. Möller, K. E. Priebe, T. Dömrose, J. G. Gatzmann, S. Rost, J. Schauss, S. Strauch, R. Bormann, M. Sivis, S. Schäfer, and C. Ropers, "Ultrafast transmission electron microscopy using a laser-driven field emitter: femtosecond resolution with a high coherence electron beam," Ultramicroscopy **176**, 63–73 (2017).
30. J. McNeur, M. Kozák, N. Schönenberger, K. J. Leedle, H. Deng, A. Ceballos, H. Hoogland, A. Ruehl, I. Hartl, R. Holzwarth, O. Solgaard, J. S. Harris, R. L. Byer, and P. Hommelhoff, "Elements of a dielectric laser accelerator," Optica **5**, 687–690 (2018).
31. R. B. Palmer, "Laser-driven grating linac," Part. Accel. **11**, 81–90 (1980).
32. K. J. Leedle, A. Ceballos, H. Deng, O. Solgaard, R. F. Pease, R. L. Byer, J. S. Harris, R. Fabian Pease, R. L. Byer, and J. S. Harris, "Dielectric laser acceleration of sub-100 keV electrons with silicon dual-pillar grating structures," Opt. Lett. **40**, 4344–4347 (2015).
33. R. Shiloh, T. Chlouba, P. Yousefi, and P. Hommelhoff, "Particle acceleration using top-illuminated nanophotonic dielectric structures," Opt. Express **29**, 14403–14411 (2021).
34. D. S. Black, K. J. Leedle, Y. Miao, U. Niedermayer, R. L. Byer, and O. Solgaard, "Laser-driven electron lensing in silicon microstructures," Phys. Rev. Lett. **122**, 104801 (2019).
35. P. Yousefi, J. McNeur, M. Kozák, U. Niedermayer, F. Gannott, O. Lohse, O. Boine-Frankenheim, and P. Hommelhoff, "Silicon dual pillar structure with a distributed Bragg reflector for dielectric laser accelerators: design and fabrication," Nucl. Instrum. Methods Phys. Res., Sect. A **909**, 221–223 (2018).

36. J. McNeur, M. Kozak, D. Ehberger, N. Schönenberger, A. Tafel, A. Li, and P. Hommelhoff, "A miniaturized electron source based on dielectric laser accelerator operation at higher spatial harmonics and a nanotip photoemitter," J. Phys. B: At., Mol. Opt. Phys. **49**, 034006 (2016).
37. P. Yousefi, N. Schönenberger, J. Mcneur, M. Kozák, U. Niedermayer, and P. Hommelhoff, "Dielectric laser electron acceleration in a dual pillar grating with a distributed Bragg reflector," Opt. Lett. **44**, 1520 (2019).
38. Y. Miao, D. S. Black, K. J. Leedle, Z. Zhao, H. Deng, A. Ceballos, R. L. Byer, J. S. Harris, and O. Solgaard, "Surface treatments of dielectric laser accelerators for increased laser-induced damage threshold," Opt. Lett. **45**, 391 (2020).
39. Z. Zhao, D. S. Black, R. J. England, T. W. Hughes, Y. Miao, O. Solgaard, R. L. Byer, S. Fan, R. Joel England, T. W. Hughes, Y. Miao, O. Solgaard, R. L. Byer, and S. Fan, "Design of a multichannel photonic crystal dielectric laser accelerator," Photonics Res. **8**, 1586 (2020).
40. E. A. Peralta, K. Soong, R. J. England, E. R. Colby, Z. Wu, B. Montazeri, C. McGuinness, J. McNeur, K. J. Leedle, D. Walz, E. B. Sozer, B. Cowan, B. Schwartz, G. Travish, and R. L. Byer, "Demonstration of electron acceleration in a laser-driven dielectric microstructure," Nature **503**, 91–94 (2013).
41. K. P. Wootton, Z. Wu, B. M. Cowan, A. Hanuka, I. V. Makasyuk, E. A. Peralta, K. Soong, R. L. Byer, and R. Joel England, "Demonstration of acceleration of relativistic electrons at a dielectric microstructure using femtosecond laser pulses," Opt. Lett. **41**, 2696–2699 (2016).
42. J. Breuer and P. Hommelhoff, "Laser-based acceleration of nonrelativistic electrons at a dielectric structure," Phys. Rev. Lett. **111**, 134803 (2013).
43. T. W. Hughes, S. Tan, Z. Zhao, N. V. Sapra, K. J. Leedle, H. Deng, Y. Miao, D. S. Black, O. Solgaard, J. S. Harris, J. Vuckovic, R. L. Byer, S. Fan, R. J. England, Y. J. Lee, and M. Qi, "On-chip laser-power delivery system for dielectric laser accelerators," Phys. Rev. Appl. **9**, 054017 (2018).
44. E. Colby, "A laser-driven linear collider: sample machine parameters and configuration," in *Proceedings of 2011 Particle Accelerator Conference*, New York, NY, USA (2011).
45. R. H. Siemann, "Energy efficiency of laser driven, structure based accelerators," Phys. Rev. Spec. Top.–Accel. Beams **7**, 061303 (2004).
46. D. Cesar, J. Maxson, X. Shen, K. P. Wootton, S. Tan, R. J. England, and P. Musumeci, "Enhanced energy gain in a dielectric laser accelerator using a tilted pulse front laser," Opt. Express **26**, 29216 (2018).
47. T. Plettner, P. P. Lu, and R. L. Byer, "Proposed few-optical cycle laser-driven particle accelerator structure," Phys. Rev. Spec. Top.–Accel. Beams **9**, 111301 (2006).
48. Y. Wei, G. Xia, J. D. A. Smith, and C. P. Welsch, "Dual-gratings with a Bragg reflector for dielectric laser-driven accelerators," Phys. Plasmas **24**, 073115 (2017).
49. T. W. Hughes, R. J. England, and S. Fan, "Reconfigurable photonic circuit for controlled power delivery to laser-driven accelerators on a chip," Phys. Rev. Appl. **11**, 064014 (2019).
50. L. Su, R. Trivedi, N. V. Sapra, A. Y. Piggott, D. Vercruysse, and J. Vučković, "Fully-automated optimization of grating couplers," Opt. Express **26**, 4023–4034 (2018).
51. D. A. B. Miller, "Perfect optics with imperfect components," Optica **2**, 747–750 (2015).
52. B. Naranjo, A. Valloni, S. Putterman, and J. B. Rosenzweig, "Stable charged-particle acceleration and focusing in a laser accelerator using spatial harmonics," Phys. Rev. Lett. **109**, 164803 (2012).

53. D. A. Swenson, "Alternating phase focused linacs," Part. Accel. **7**, 61–67 (1976).
54. U. Niedermayer, A. Adelmann, and S. Bettoni, *et al.*, "Challenges in simulating beam dynamics of dielectric laser acceleration," Int. J. Mod. Phys. A **34**, 1942031 (2019).
55. J. Wang, *Grating and Ring Based Devices on SOI Platform* (McGill University, 2016).
56. C.-M. Chang and O. Solgaard, "Silicon buried gratings for dielectric laser electron accelerators," Appl. Phys. Lett. **104**, 184102 (2014).
57. T. Hughes, G. Veronis, K. P. Wootton, R. J. England, and S. Fan, "Method for computationally efficient design of dielectric laser accelerator structures," Opt. Express **25**, 15414–15727 (2017).
58. K. Mizuno, J. Pae, T. Nozokido, and K. Furuya, "Experimental evidence of the inverse Smith–Purcell effect," Nature **328**, 45–47 (1987).
59. M. Kozák, M. Förster, J. McNeur, N. Schönenberger, K. Leedle, H. Deng, J. S. Harris, R. L. Byer, and P. Hommelhoff, "Dielectric laser acceleration of sub-relativistic electrons by few-cycle laser pulses," Nucl. Instrum. Methods Phys. Res., Sect. A **865**, 84–86 (2017).
60. J. Breuer, R. Graf, A. Apolonski, and P. Hommelhoff, "Dielectric laser acceleration of nonrelativistic electrons at a single fused silica grating structure: Experimental part," Phys. Rev. Spec. Top.–Accel. Beams **17**, 021301 (2014).
61. K. J. Leedle, R. Fabian Pease, R. L. Byer, and J. S. Harris, "Laser acceleration and deflection of 96.3 keV electrons with a silicon dielectric structure," Optica **2**, 158–161 (2015).
62. M. Kozák, P. Beck, H. Deng, J. McNeur, N. Schönenberger, C. Gaida, F. Stutzki, M. Gebhardt, J. Limpert, A. Ruehl, I. Hartl, O. Solgaard, J. S. Harris, R. L. Byer, and P. Hommelhoff, "Acceleration of sub-relativistic electrons with an evanescent optical wave at a planar interface," Opt. Express **25**, 19195–19204 (2017).
63. M. Kozák, J. McNeur, K. J. Leedle, H. Deng, N. Schönenberger, A. Ruehl, I. Hartl, J. S. Harris, R. L. Byer, and P. Hommelhoff, "Optical gating and streaking of free electrons with sub-optical cycle precision," Nat. Commun. **8**, 14342 (2017).
64. K. J. Leedle, D. S. Black, Y. Miao, K. E. Urbanek, A. Ceballos, H. Deng, J. S. Harris, O. Solgaard, and R. L. Byer, "Phase-dependent laser acceleration of electrons with symmetrically driven silicon dual pillar gratings," Opt. Lett. **43**, 2181–2184 (2018).
65. H. Deng, K. J. Leedle, Y. Miao, D. S. Black, K. E. Urbanek, J. McNeur, M. Kozák, A. Ceballos, P. Hommelhoff, O. Solgaard, R. L. Byer, and J. S. Harris, "Gallium oxide for high-power optical applications," Adv. Opt. Mater. **8**, 1901522 (2020).
66. Y. Adiv, K. Wang, R. Dahan, P. Broaddus, Y. Miao, D. Black, K. Leedle, O. Solgaard, J. England, and I. Kaminer, "Observation of the quantum nature of laser-driven particle acceleration," *Conference on Lasers and Electro-Optics* (2020).
67. U. Niedermayer, D. S. Black, K. J. Leedle, Y. Miao, R. L. Byer, and O. Solgaard, "Low-energy-spread attosecond bunching and coherent electron acceleration in dielectric nanostructures," Phys. Rev. Appl. **15**, L021002 (2021).
68. R. Shiloh, T. Chlouba, and P. Hommelhoff, "Experimental considerations in electron beam transport on a nanophotonic chip using alternating phase focusing," J. Vac. Sci. Technol., B: Nanotechnol. Microelectron.: Mater., Process., Meas., Phenom. **40**, 010602 (2022).
69. T. Chlouba, R. Shiloh, P. Forsberg, M. Hamberg, M. Karlsson, M. Kozák, and P. Hommelhoff, "Diamond-based dielectric laser acceleration," Opt. Express **30**, 505–510 (2022).

70. U. Niedermayer, T. Egenolf, and O. Boine-Frankenheim, "Beam dynamics analysis of dielectric laser acceleration using a fast 6D tracking scheme," Phys. Rev. Accel. Beams **20**, 111302 (2017).
71. W. K. H. Panofsky and W. A. Wenzel, "Some considerations concerning the transverse deflection of charged particles in radio-frequency fields," Rev. Sci. Instrum. **27**, 967 (1956).
72. M. J. Browman, "Using the Panofsky–Wenzel theorem in the analysis of radio-frequency deflectors," in *Proceedings of International Conference on Particle Accelerators* (IEEE, 1993), pp. 800–802.
73. A. W. Chao, *Lecture Notes on Topics in Accelerator Physics* (2002).
74. I. B. Fainberg, "Alternating Phase Focusing," in *CERN Symposium On High-Energy Accelerators And Pion Phyiscs* (1956), pp. 91–100.
75. U. Niedermayer, T. Egenolf, and O. Boine-Frankenheim, "Three dimensional alternating-phase focusing for dielectric-laser electron accelerators," Phys. Rev. Lett. **125**, 164801 (2020).
76. U. Niedermayer, J. Lautenschläger, T. Egenolf, and O. Boine-Frankenheim, "Design of a scalable integrated nanophotonic electron accelerator on a chip," Phys. Rev. Appl. **16**, 024022 (2021).
77. B. J. Siwick, J. R. Dwyer, R. E. Jordan, and R. J. D. Miller, "An atomic-level view of melting using femtosecond electron diffraction," Science **302**, 1382–1385 (2003).
78. P. Baum, D.-S. Yang, and A. H. Zewail, "4D Visualization of Transitional Structures in Phase Transformations by Electron Diffraction," Science **318**, 788–792 (2007).
79. Y. Morimoto and P. Baum, "Diffraction and microscopy with attosecond electron pulse trains," Nat. Phys. **14**, 252–256 (2018).
80. A. H. Zewail and J. M. Thomas, *4D Electron Microscopy Imaging in Space and Time* (Imperial College Press, 2010).
81. K. B. Schliep, P. Quarterman, J. P. Wang, and D. J. Flannigan, "Picosecond Fresnel transmission electron microscopy," Appl. Phys. Lett. **110**, 222404 (2017).
82. G. Berruto, I. Madan, Y. Murooka, G. M. Vanacore, E. Pomarico, J. Rajeswari, R. Lamb, P. Huang, A. J. Kruchkov, Y. Togawa, T. LaGrange, D. McGrouther, H. M. Rønnow, and F. Carbone, "Laser-induced skyrmion writing and erasing in an ultrafast cryo-Lorentz transmission electron microscope," Phys. Rev. Lett. **120**, 117201 (2018).
83. S. Ackermann, A. Azima, and S. Bajt, *et al.*, "Generation of coherent 19- and 38-nm radiation at a free-electron laser directly seeded at 38 nm," Phys. Rev. Lett. **111**, 114801 (2013).
84. Y. Morimoto and P. Baum, "Single-cycle optical control of beam electrons," Phys. Rev. Lett. **125**, 193202 (2020).
85. A. Ody, P. Musumeci, J. Maxson, D. Cesar, R. J. England, and K. P. Wootton, "Flat electron beam sources for DLA accelerators," Nucl. Instrum. Methods Phys. Res., Sect. A **865**, 75–83 (2017).
86. S. Crisp, A. Ody, P. Musumeci, and R. J. England, "Resonant phase matching by oblique illumination of a dielectric laser accelerator," Phys. Rev. Accel. Beams **24**, 121305 (2021).
87. L. R. Doolittle, H. Ma, and M. Champion, "Digital low-level RF control using non-IQ sampling," in *Proceedings of LINAC 2006, Knoxville, Tennessee USA* (2006).
88. L. Su, A. Y. Piggott, N. V. Sapra, J. Petykiewicz, and J. Vučković, "Inverse design and demonstration of a compact on-chip narrowband three-channel wavelength demultiplexer," ACS Photonics **5**, 301–305 (2018).

89. Y. Wei, M. Ibison, G. Xia, J. D. A. Smith, and C. P. Welsch, "Dual-grating dielectric accelerators driven by a pulse-front-tilted laser," Appl. Opt. **56**, 8201–8206 (2017).
90. J. Hebling, "Derivation of the pulse front tilt caused by angular dispersion," Opt. Quantum Electron. **28**, 1759–1763 (1996).
91. D. Cesar, J. Maxson, P. Musumeci, X. Shen, R. J. England, and K. P. Wootton, "Optical design for increased interaction length in a high gradient dielectric laser accelerator," Nucl. Instrum. Methods Phys. Res., Sect. A **909**, 252–256 (2018).
92. S. Reiche, J. B. Rosenzweig, S. Anderson, P. Frigola, M. Hogan, A. Murokh, C. Pellegrini, L. Serafini, G. Travish, and A. Tremaine, "Experimental confirmation of transverse focusing and adiabatic damping in a standing wave linear accelerator," Phys. Rev. E **56**, 3572–3577 (1997).
93. D. B. Cesar, "Probing ultrafast dynamics with relativistic electrons," Ph.D dissertation (University of CaliforniaLos Angeles, 2019).
94. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics* (Elsevier Science, 2013).
95. A. Szczepkowicz, "Application of transfer matrix and transfer function analysis to grating-type dielectric laser accelerators: ponderomotive focusing of electrons," Phys. Rev. Accel. Beams **20**, 081302 (2017).
96. B. V. Chirikov, "Resonance processes in magnetic traps," J. Nucl. Energy, Part C **1**, 253–260 (1960).
97. D. Cesar, P. Musumeci, and J. England, "All optical control of beam dynamics in a dla," in *2018 IEEE Advanced Accelerator Concepts Workshop (AAC)* (2018), pp. 1–5.
98. A. Ody, S. Crisp, P. Musumeci, D. Cesar, and R. J. England, "SHarD: A beam dynamics simulation code for dielectric laser accelerators based on spatial harmonic field expansion," Nucl. Instrum. Methods Phys. Res., Sect. A **1013**, 165635 (2021).
99. S. T. Park, M. Lin, and A. H. Zewail, "Photon-induced near-field electron microscopy (PINEM): theoretical and experimental," New J. Phys. **12**, 123028 (2010).
100. F. J. G. de Abajo, A. Asenjo-Garcia, and M. Kociak, "Multiphoton absorption and emission by interaction of swift electrons with evanescent light fields," Nano Lett. **10**, 1859–1863 (2010).
101. F. H. M. Faisal, *Theory of Multiphoton Processes* (Springer Science+ Business Media, 1987).
102. D. A. Varshalovich and M. I. D'yakonov, "Quantum theory of the modulation of the electron beam at optical frequencies," Sov. Phys. JETP **33**, 51–57 (1971).
103. A. Weingartshofer, J. K. Holmes, G. Caudle, E. M. Clarke, and H. Krüger, "Direct observation of multiphoton processes in laser-induced free-free transitions," Phys. Rev. Lett. **39**, 269–270 (1977).
104. L. Friedland, "Correspondence principle in multiphoton inverse bremsstrahlung," J. Phys. B: At. Mol. Phys. **12**, 409–418 (1979).
105. L. Friedland, "Correspondence principle in free-electron lasers," Phys. Rev. A **29**, 1310–1314 (1984).
106. A. Fruchtman and L. Friedland, "Simplified small signal gain calculations in free electron lasers," Int. J. Infrared Millimeter Waves **5**, 683–690 (1984).
107. L. Piazza, T. T. A. Lummen, E. Quiñonez, Y. Murooka, B. W. Reed, B. Barwick, and F. Carbone, "Simultaneous observation of the quantization and the interference pattern of a plasmonic near-field," Nat. Commun. **6**, 6407 (2015).
108. K. E. Priebe, C. Rathje, S. V. Yalunin, T. Hohage, A. Feist, S. Schäfer, and C. Ropers, "Attosecond electron pulse trains and quantum state reconstruction in ultrafast transmission electron microscopy," Nat. Photonics **11**, 793–797 (2017).

109. G. M. Vanacore, G. Berruto, I. Madan, E. Pomarico, P. Biagioni, R. J. Lamb, D. McGrouther, O. Reinhardt, I. Kaminer, B. Barwick, H. Larocque, V. Grillo, E. Karimi, F. Javier García de Abajo, and F. Carbone, "Ultrafast generation and control of an electron vortex beam via chiral plasmonic near fields," Nat. Mater. **18**, 573–579 (2019).
110. P. B. Corkum and F. Krausz, "Attosecond science," Nat. Phys. **3**, 381–387 (2007).
111. P. Baum and A. H. Zewail, "Attosecond electron pulses for 4D diffraction and microscopy," Proc. Natl. Acad. Sci. U. S. A. **104**, 18409–18414 (2007).
112. P. Baum, "Quantum dynamics of attosecond electron pulse compression," J. Appl. Phys. **122**, 223105 (2017).
113. O. Kfir, "Entanglements of electrons and cavity photons in the strong-coupling regime," Phys. Rev. Lett. **123**, 103602 (2019).
114. V. Di Giulio, M. Kociak, and F. J. G. de Abajo, "Probing quantum optical excitations with fast electrons," Optica **6**, 1524–1534 (2019).
115. A. Gorlach, A. Karnieli, R. Dahan, E. Cohen, A. Pe'er, and I. Kaminer, "Ultrafast non-destructive measurement of the quantum state of light using free electrons," in *Conference on Lasers and Electro-Optics,* OSA Technical Digest (Optica Publishing Group, 2021), paper FF2I.4.
116. R. Dahan, A. Gorlach, U. Haeusler, A. Karnieli, O. Eyal, P. Yousefi, M. Segev, A. Arie, G. Eisenstein, P. Hommelhoff, and I. Kaminer, "Imprinting the quantum statistics of photons on free electrons," Science **373**, eabj7128 (2021).
117. A. Gover and A. Yariv, "Free-electron–bound-electron resonant interaction," Phys. Rev. Lett. **124**, 064801 (2020).
118. M. O. Scully and M. S. Zubairy, *Quantum Optics* (1999).
119. Z. Zhao, X.-Q. Sun, and S. Fan, "Quantum entanglement and modulation enhancement of free-electron–bound-electron interaction," Phys. Rev. Lett. **126**, 233402 (2021).
120. R. Ruimy, A. Gorlach, C. Mechel, N. Rivera, and I. Kaminer, "Toward atomic-resolution quantum measurements with coherently shaped free electrons," Phys. Rev. Lett. **126**, 233403 (2021).
121. F. J. G. de Abajo and V. Di Giulio, "Optical excitations with electron beams: challenges and opportunities," ACS Photonics **8**, 945–974 (2021).
122. G. Baranes, R. Ruimy, A. Gorlach, and I. Kaminer, "Free electrons can induce quantum correlations between two separate photonic cavities," in *Conference on Lasers and Electro-Optics*, J. Kang, S. Tomasulo, I. Ilev, D. Müller, N. Litchinitser, S. Polyakov, V. Podolskiy, J. Nunn, C. Dorrer, T. Fortier, Q. Gan, and C. Saraceno, eds., OSA Technical Digest (Optical Society of America, 2021), p. FTh1N.7.
123. M. E. Peskin, *An Introduction to Quantum Field Theory* (CRC Press, 2018).
124. D. B. Williams and C. B. Carter, "The transmission electron microscope," in *Transmission Electron Microscopy* (Springer, 1996), pp. 3–17.
125. W. Magnus, "On the exponential solution of differential equations for a linear operator," Commun. Pure and Appl. Maths. **7**, 649–673 (1954).
126. N. Rivera and I. Kaminer, "Light–matter interactions with photonic quasiparticles," Nat. Rev. Phys. **2**, 538–561 (2020).
127. F. J. G. De Abajo, "Comment on 'Free-electron-bound-electron resonant interaction'," Phys. Rev. Lett. **126**, 019501 (2021).
128. A. Gover and A. Yariv, "Gover and Yariv reply," Phys. Rev. Lett. **126**, 019502 (2021).
129. D. L. Webster, "Cathode-ray bunching," J. Appl. Phys. **10**, 501–508 (1939).
130. R. J. Glauber, "Coherent and incoherent states of the radiation field," Phys. Rev. **131**, 2766–2788 (1963).

131. A. Ben Hayun, O. Reinhardt, J. Nemirovsky, A. Karnieli, N. Rivera, and I. Kaminer, "Shaping quantum photonic states using free electrons," Sci. Adv. **7**, eabe4270 (2021).
132. V. Di Giulio and F. J. G. de Abajo, "Free-electron shaping using quantum light," Optica **7**, 1820 (2020).
133. S. V. Yalunin, A. Feist, and C. Ropers, "Tailored high-contrast attosecond electron pulses for coherent excitation and scattering," Phys. Rev. Res. **3**, L032036 (2021).
134. O. Reinhardt and I. Kaminer, "Theory of shaping electron wavepackets with light," ACS Photonics **7**, 2859–2870 (2020).
135. R. Jozsa, "Fidelity for mixed quantum states," J. Mod. Opt. **41**, 2315–2323 (1994).
136. H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press on Demand, 2002).
137. Y. Pan and A. Gover, "Spontaneous and stimulated emissions of a preformed quantum free-electron wave function," Phys. Rev. A **99**, 052107 (2019).
138. A. Gover and Y. Pan, "Dimension-dependent stimulated radiative interaction of a single electron quantum wavepacket," Phys. Lett. A **382**, 1550–1555 (2018).
139. R. Remez, A. Karnieli, S. Trajtenberg-Mills, N. Shapira, I. Kaminer, Y. Lereah, and A. Arie, "Observing the quantum wave nature of free electrons through spontaneous emission," Phys. Rev. Lett. **123**, 060401 (2019).
140. A. Friedman, A. Gover, G. Kurizki, S. Ruschin, and A. Yariv, "Spontaneous and stimulated emission from quasifree electrons," Rev. Mod. Phys. **60**, 471–535 (1988).
141. R. H. Ritchie and A. Howie, "Inelastic scattering probabilities in scanning transmission electron microscopy," Philos. Mag. A **58**, 753–767 (1988).
142. F. J. G. de Abajo, "Optical excitations in electron microscopy," Rev. Mod. Phys. **82**, 209–275 (2010).
143. A. Karnieli, N. Rivera, A. Arie, and I. Kaminer, "The coherence of light is fundamentally tied to the quantum coherence of the emitting particle," Sci. Adv. **7**, eabf8096 (2021).
144. O. Kfir, V. Di Giulio, F. J. G. de Abajo, and C. Ropers, "Optical coherence transfer mediated by free electrons," Sci. Adv. **7**, 6380 (2021).
145. A. Gorlach, A. Karnieli, R. Dahan, E. Cohen, A. Pe'er, and I. Kaminer, "Ultrafast non-destructive measurement of the quantum state of light using free electrons," (2020).
146. V. L. Ginzburg, "Radiation by uniformly moving sources (Vavilov–Cherenkov effect, transition radiation, and other phenomena)," Phys.-Usp. **39**, 973–982 (1996).
147. S. Tsesses, G. Bartal, and I. Kaminer, "Light generation via quantum interaction of electrons with periodic nanostructures," Phys. Rev. A **95**, 013832 (2017).
148. I. Kaminer, M. Mutzafi, A. Levy, G. Harari, H. H. Sheinfux, S. Skirlo, J. Nemirovsky, J. D. Joannopoulos, M. Segev, and M. Soljacic, "Quantum Čerenkov radiation: spectral cutoffs and the role of spin and orbital angular momentum," Phys. Rev. X **6**, 011006 (2016).
149. S. Huang, R. Duan, N. Pramanik, J. S. Herrin, C. Boothroyd, Z. Liu, and L. J. Wong, "Quantum recoil in free electron-driven spontaneous emission from van der Waals crystals," in *Conference on Lasers and Electro-Optics*, Technical Digest Series (Optica Publishing Group, 2022), p. FF4C.4.
150. S. J. Smith and E. M. Purcell, "Visible light from localized surface charges moving across a grating," Phys. Rev. **92**, 1069 (1953).
151. R. Remez, N. Shapira, C. Roques-Carmes, R. Tirole, Y. Yang, Y. Lereah, M. Soljačić, I. Kaminer, and A. Arie, "Spectral and spatial shaping of Smith–Purcell radiation," Phys. Rev. A **96**, 061801 (2017).

152. M. J. Moran, "X-ray generation by the Smith–Purcell effect," Phys. Rev. Lett. **69**, 2523–2526 (1992).
153. A. Massuda, C. Roques-Carmes, Y. Yang, S. E. Kooi, Y. Yang, C. Murdia, K. K. Berggren, I. Kaminer, and M. Soljačić, "Smith–Purcell radiation from low-energy electrons," ACS Photonics **5**, 3513–3518 (2018).
154. Z. Su, F. Cheng, L. Li, and Y. Liu, "Complete control of Smith–Purcell radiation by graphene metasurfaces," ACS Photonics **6**, 1947–1954 (2019).
155. L. Jing, X. Lin, Z. Wang, I. Kaminer, H. Hu, E. Li, Y. Liu, M. Chen, B. Zhang, and H. Chen, "Polarization shaping of free-electron radiation by gradient bianisotropic metasurfaces," Laser Photonics Rev. **15**, 2000426 (2021).
156. S. E. Korbly, A. S. Kesar, J. R. Sirigiri, and R. J. Temkin, "Observation of frequency-locked coherent terahertz Smith–Purcell radiation," Phys. Rev. Lett. **94**, 054803 (2005).
157. I. Kaminer, S. E. Kooi, R. Shiloh, B. Zhen, Y. Shen, J. J. López, R. Remez, S. A. Skirlo, Y. Yang, J. D. Joannopoulos, A. Arie, and M. Soljacic, "Spectrally and spatially resolved Smith–Purcell radiation in plasmonic crystals with short-range disorder," Phys. Rev. X **7**, 011003 (2017).
158. Z. Wang, K. Yao, M. Chen, H. Chen, and Y. Liu, "Manipulating Smith–Purcell emission with babinet metasurfaces," Phys. Rev. Lett. **117**, 157401 (2016).
159. L. Liang, W. Liu, Y. Liu, Q. Jia, L. Wang, and Y. Lu, "Multi-color and multidirectional-steerable Smith–Purcell radiation from 2D sub-wavelength hole arrays," Appl. Phys. Lett. **113**, 013501 (2018).
160. T. Fu, D. Wang, Z. Yang, Z. Deng, and W. Liu, "Steering Smith–Purcell radiation angle in a fixed frequency by the Fano-resonant metasurface," Opt. Express **29**, 26983–26994 (2021).
161. Y.-C. Lai, T. C. Kuang, B. H. Cheng, Y.-C. Lan, and D. P. Tsai, "Generation of convergent light beams by using surface plasmon locked Smith–Purcell radiation," Sci. Rep. **7**, 11096 (2017).
162. Y. Yang, C. Roques-Carmes, I. Kaminer, A. Zaidi, A. Massuda, Y. Yang, S. E. Kooi, K. K. Berggren, and M. Soljacic, "Manipulating Smith–Purcell radiation polarization with metasurfaces," in *Conference on Lasers and Electro-Optics* (2018), Paper FW4H.1.
163. T. Latychevskaia, "Spatial coherence of electron beams from field emitters and its effect on the resolution of imaged objects," Ultramicroscopy **175**, 121–129 (2017).
164. B. Cho, T. Ichimura, R. Shimizu, and C. Oshima, "Quantitative evaluation of spatial coherence of the electron beam from low temperature field emitters," Phys. Rev. Lett. **92**, 246103 (2004).
165. E. Pomarico, I. Madan, G. Berruto, G. M. Vanacore, K. Wang, I. Kaminer, F. J. García De Abajo, and F. Carbone, "meV resolution in laser-assisted energy-filtered transmission electron microscopy," ACS Photonics **5**, 759–764 (2018).
166. O. L. Krivanek, N. Dellby, J. A. Hachtel, J. C. Idrobo, M. T. Hotz, B. Plotkin-Swing, N. J. Bacon, A. L. Bleloch, G. J. Corbin, M. V. Hoffman, C. E. Meyer, and T. C. Lovejoy, "Progress in ultrahigh energy resolution EELS," Ultramicroscopy **203**, 60–67 (2019).
167. L. Reimer and H. Kohl, *Transmission Electron Microscopy: Physics of Image Formation* (Springer, 2008).
168. M. Liebtrau, M. Sivis, A. Feist, H. Lourenço-Martins, N. Pazos-Pérez, R. A. Alvarez-Puebla, F. J. G. de Abajo, A. Polman, and C. Ropers, "Spontaneous and stimulated electron–photon interactions in nanoscale plasmonic near fields," Light: Sci. Appl. **10**, 82 (2021).
169. E. Schrödinger, "An undulatory theory of the mechanics of atoms and molecules," Phys. Rev. **28**, 1049–1070 (1926).

170. R. P. Feynman, R. B. Leighton, and M. L. Sands, *The Feynman Lectures on Physics*.
171. M. Born, "Born's Nobel lecture on the statistical interpretation of quantum mechanics," *Nobel Lect. Tech. Report* (1954).
172. N. Talebi, "Schrödinger electrons interacting with optical gratings: quantum mechanical study of the inverse Smith–Purcell effect," New J. Phys. **18**, 123006 (2016).
173. E. Lorin, S. Chelkowski, and A. Bandrauk, "A numerical Maxwell–Schrödinger model for intense laser–matter interaction and propagation," Comput. Phys. Commun. **177**, 908–932 (2007).
174. I. P. Ivanov, V. G. Serbo, and V. A. Zaytsev, "Quantum calculation of the Vavilov-Cherenkov radiation by twisted electrons," Phys. Rev. A **93**, 053825 (2016).
175. D. V. Karlovets and A. M. Pupasov-Maksimov, "Nonlinear quantum effects in electromagnetic radiation of a vortex electron," Phys. Rev. A **103**, 012214 (2021).
176. Y. Pan and A. Gover, "Spontaneous and stimulated radiative emission of modulated free-electron quantum wavepackets—semiclassical analysis," J. Phys. Commun. **2**, 115026 (2018).
177. A. Gover, R. Ianconescu, A. Friedman, C. Emma, N. Sudar, P. Musumeci, and C. Pellegrini, "Superradiant and stimulated-superradiant emission of bunched electron beams," Rev. Mod. Phys. **91**, 035003 (2019).
178. C. Pellegrini, A. Marinelli, and S. Reiche, "The physics of x-ray free-electron lasers," Rev. Mod. Phys. **88**, 015006 (2016).
179. A. Karnieli, N. Rivera, A. Arie, and I. Kaminer, "Superradiance and subradiance due to quantum interference of entangled free electrons," Phys. Rev. Lett. **127**, 060403 (2021).
180. A. Angioi and A. Di Piazza, "Quantum limitation to the coherent emission of accelerated charges," Phys. Rev. Lett. **121**, 010402 (2018).
181. P. Kling, E. Giese, C. M. Carmesin, R. Sauerbrey, and W. P. Schleich, "High-gain quantum free-electron laser: emergence and exponential gain," Phys. Rev. A **99**, 053823 (2019).
182. W. P. S. Peter Kling, E. Giese, R. Endrich, P. Preiss, and R. Sauerbrey, "What defines the quantum regime of the free-electron laser?" New J. Phys. **17**, 123019 (2015).
183. G. R. M. Robb and R. Bonifacio, "Coherent and spontaneous emission in the quantum free electron laser," Phys. Plasmas **19**, 073101 (2012).
184. L. J. Wong, N. Rivera, C. Murdia, T. Christensen, J. D. Joannopoulos, M. Soljačić, and I. Kaminer, "Control of quantum electrodynamical processes by shaping electron wavepackets," Nat. Commun. **12**, 1700 (2021).
185. V. Di Giulio, O. Kfir, C. Ropers, and F. Javier García de Abajo, "Modulation of cathodoluminescence emission by interference with external light," ACS Nano **15**, 7290–7304 (2021).
186. V. L. Ginzburg, "Quantum theory of radiation of electron uniformly moving in medium," Zh. Eksp. Teor. Fiz. **10**, 589 (1940).
187. C. M. Carmesin, P. Kling, E. Giese, R. Sauerbrey, and W. P. Schleich, "Quantum and classical phase-space dynamics of a free-electron laser," Phys. Rev. Res. **2**, 023027 (2020).
188. Y. Adiv, H. Hu, S. Tsesses, R. Dahan, K. Wang, Y. Kurman, H. Chen, X. Lin, G. Bartal, and I. Kaminer, "Observation of 2D Cherenkov radiation and its quantized photonic nature using free-electrons," *Conf. Lasers Electro-Optics* (2021), Pap. FM1L.6.
189. K. E. Echternkamp, A. Feist, S. Schäfer, and C. Ropers, "Ramsey-type phase control of free-electron beams," Nat. Phys. **12**, 1000–1004 (2016).

190. M. Krüger, M. Schenk, and P. Hommelhoff, "Attosecond control of electrons emitted from a nanoscale metal tip," Nature **475**, 78–81 (2011).
191. A. Feist, G. Huang, G. Arend, Y. Yang, J.-W. Henke, A. S. Raja, F. J. Kappert, R. N. Wang, H. Lourenço-Martins, Z. Qiu, J. Liu, O. Kfir, T. J. Kippenberg, and C. Ropers, "Cavity-mediated electron–photon pairs," Science **377**, 777–780 (2022).
192. R. J. Glauber, "Quantum theory of optical coherence," Phys. Rev. **130**, 2529–2539 (1963).
193. O. Schwartz, J. J. Axelrod, S. L. Campbell, C. Turnbaugh, R. M. Glaeser, and H. Müller, "Laser phase plate for transmission electron microscopy," Nat. Methods **16**, 1016–1020 (2019).
194. P. Das, J. D. Blazit, M. Tencé, L. F. Zagonel, Y. Auad, Y. H. Lee, X. Y. Ling, A. Losquin, C. Colliex, O. Stéphan, M. García de Abajo, and Kociak, "Stimulated electron energy loss and gain in an electron microscope without a pulsed electron gun," Ultramicroscopy **203**, 44–51 (2019).
195. C. Liu, Y. Wu, Z. Hu, J. A. Busche, E. K. Beutler, N. P. Montoni, T. M. Moore, G. A. Magel, J. P. Camden, D. J. Masiello, G. Duscher, and P. D. Rack, "Continuous wave resonant photon stimulated electron energy-gain and electron energy-loss spectroscopy of individual plasmonic nanoparticles," ACS Photonics **6**, 2499–2508 (2019).
196. A. Ryabov, J. W. Thurner, D. Nabben, M. V. Tsarev, and P. Baum, "Attosecond metrology in a continuous-beam transmission electron microscope," Sci. Adv. **6**, abb1393 (2020).

Roy Shiloh studied Physics and Electrical Engineering and received his PhD in 2017 from Tel Aviv University under the supervision of Prof. Ady Arie, where he pioneered the field of computer-generated holograms in electron optics using thin membranes as amplitude and phase masks. Since then he has been researching methods for nanophotonic control of electron beams, mainly in the contexts of the on-chip electron accelerator and the PINEM effect. He was the co-leader of the Injector group under the Accelerator on a Chip International Program (ACHIP). In 2023 he will join the faculty of the Applied Physics Department of the Hebrew University of Jerusalem.

Norbert Schönenberger was instrumental in the Accelerator on a Chip International Program (ACHIP) efforts in Erlangen, culminating in the experimental demonstration of the record-breaking 270 as electron bunch train generation using DLA.

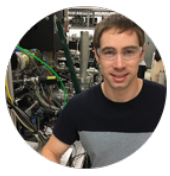
Yuval Adiv is a PhD candidate at the Technion – Israel Institute of Technology. He received his Bachelor of Science in Electrical Engineering and Physics in 2019 from the Technion. He is currently investigating quantum interactions between free electrons and light in the AdQuanta group led by Professor Ido Kaminer.

Ron Ruimy graduated *summa cum laude* in Electrical Engineering and Physics double major BSc at the Technion, Israel. Currently in the direct PhD track in the faculty of electrical engineering in the Technion under the supervision of Professor Ido Kaminer. During his graduate studies, his research focus was theoretical work concerning the quantum aspects of electron–matter interaction, trying to explore how the quantum properties of the free electrons could be utilized as a unique probe for quantum technologies.

Aviv Karnieli received his BSc degrees (*summa cum laude*) in Physics and Electrical Engineering from Tel-Aviv University in 2017. He then enrolled to the direct PhD track in Physics, in the Sackler School of Physics and Astronomy in Tel-Aviv University, receiving the prestigious Adams Fellowship by the Israel Academy of Sciences and Humanities in 2019. In his PhD research, under the joint supervision of Prof. Ady Arie and Prof. Ido Kaminer from the Technion, Israel, he studies quantum optical effects in nonlinear optics and free-electron–light interactions.

Tyler Hughes is a research scientist at Flexcompute, Inc. He received his PhD and Master's degree in Applied Physics from Stanford University. His research interests include computational electromagnetics and integrated photonics.

Dr. Joel England obtained his PhD in accelerator physics at the University of California Los Angeles, where he developed a method of generating tailored drive beams for more efficient excitation of plasma-based particle accelerators. In 2008, he joined the Advanced Accelerator Research Department at SLAC as a postdoc working on optical-scale particle accelerators powered by solid-state lasers. As a recipient of the Panofsky Fellowship, he led the Dielectric Laser Acceleration Group at SLAC from 2010 to 2015. This led to the first demonstrations of high-gradient acceleration in laser-driven dielectric structures conducted in collaboration with Prof. Robert Byer at Stanford. He is currently a lead scientist at SLAC and the Head of Accelerator Operations for the Ultrafast Electron Diffraction Group. He serves as co-leader and executive committee member for the Accelerator on a Chip International Program (ACHIP), which aims to develop a laser-driven microchip particle accelerator.



Ken Leedle was the main experimentalist of the Accelerator on a Chip International Program (ACHIP) in Stanford for many years. He was the co-leader of the Injector group and made substantial contributions toward the program's goals on nanostructures and integration into a stand-alone compact device.

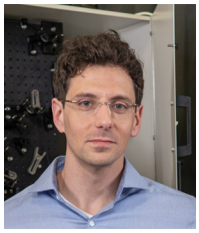
Dr. Dylan Black received his PhD from Stanford University, and his Bachelor's degree from UCLA. His thesis work demonstrates the first direct measurement of attosecond-scale electron bunching in dielectric laser accelerators and was chosen as Stanford University's entry to the Council of Graduate Schools' Distinguished Dissertation award. His research on laser-powered electron lenses has also been featured in *Science* magazine. He is currently employed as a Research Scientist at PsiQuantum, Inc. His interests include accelerator physics, nonlinear optics, photonic quantum computing, and Middle Egyptian hieroglyphics.

Zhexin Zhao received the Bachelor's degree (with honors) in Electronic Engineering from Tsinghua University, Beijing, China, in 2015, and the MS and PhD degrees in Electrical Engineering (with PhD minor in physics) from Stanford University in 2018 and 2021, respectively. Her research interests include electromagnetic theory, photonic design and optimization, quantum physics, and accelerator physics.

Prof. Pietro Musumeci is the director of the UCLA Pegasus Photoinjector Lab. He joined the Department of Physics and Astronomy at the University of California Los Angeles in 2006. Before that, he was one of the leaders of the experimental effort on the Italian free-electron laser project SPARC at the Laboratori Nazionali di Frascati, Italy. He was elected APS Fellow in 2017 for his contribution on laser acceleration, photoinjector and high-brightness ultrashort electron beam production. His interests lie in the application of new ultrafast high-power laser technologies to the field of accelerators and beam physics. He has authored over 200 publications and serves on various committees for funding agencies and international conferences. He is active in the field of high-brightness electron sources, ultrafast electron scattering techniques, high-efficiency free-electron lasers, and high-gradient laser accelerators.

Professor Robert L. Byer is the William R. Kenan, Jr. Professor of Applied Physics at Stanford University. He has conducted research and taught classes in lasers and nonlinear optics at Stanford University since 1969. He has made numerous contributions to laser science and technology, including the demonstration of the first tunable visible parametric oscillator, the development of the Q-switched unstable resonator Nd:YAG laser, remote sensing and coherent laser radar, and precision spectroscopy using Coherent Anti Stokes Raman Scattering (CARS). Current research includes lasers and suspensions for LIGO, gravitational wave detection, and laser particle acceleration; accelerators on a chip. Professor Byer has published more than 500 scientific papers and holds 54 patents in the fields of lasers and nonlinear optics. His Google Scholar lifetime citations are more than 93,000 (h ~140). Professor Byer was elected to the National Academy of Engineering in 1987 and to the National Academy of Science in 2000 and as a charter member of the National Academy of Inventors in 2012.

Prof. Ady Arie received his BSc degree in Mathematics and Physics from the Hebrew University of Jerusalem in 1983. In 1986 and 1992 he received his MSc degree in Physics and PhD degree in Engineering from Tel-Aviv University. Since 2006 he is a Professor of Electrical Engineering in the Department of Physical Electronics, School of Electrical Engineering in the Fleischman Faculty of Engineering, Tel-Aviv University, and holds the Marko and Lucie Chaoul Chair in Nano-Photonics. His research in recent years is in the areas of classical and quantum nonlinear optics, hydrodynamics, electron optics, and light–electron interaction. He is a Fellow of the Optical Society. In the years 2008–2014 he served as a Topical Editor of *Optics Letters* and since 2018 he is an Associate Editor of Optica. In 2016 he won the Kadar Foundation Award for Excellence in Research.



Ido Kaminer received his PhD under Prof. Mordechai Segev, where he discovered new classes of accelerating optical beams, for which he received the 2014 American Physical Society (APS) Award for Outstanding Doctoral Dissertation in Laser Science. He later did his postdoc at MIT with Prof. Marin Soljacic and Prof. John Joannopoulos, where he established the foundations of macroscopic quantum electrodynamics (MQED) for photonic quasiparticles. He is now an associate professor at the Technion. His group connected the field of free-electron radiation to quantum optics. He was recently elected to the Israeli Young Academy, which includes 32 young Israeli faculty members below the age of 45. He has won multiple awards and grants, including the ERC Starting Grant, the Krill Prize, and the 2022 Schmidt Science Polymath Award. He is also the laureate of the 2021 Blavatnik Award in Physical Sciences and Engineering in Israel, and the recipient of the 2022 Adolph Lomb Medal, the top international award for a young scientist (age 35 or younger) in the field of optics.

Peter Hommelhoff studied physics at TU Berlin and ETH Zurich, received his PhD under Theodor Hänsch at LMU Munich (2002), went to Stanford for a 4-year postdoctoral stay with Mark Kasevich, and ran a Max Planck Research Group in Garching from 2007 to 2012 before becoming a professor of physics at Friedrich-Alexander-Universtität (FAU) in Erlangen. His current research interests include laser-based electron acceleration in nanophotonic structures, attosecond physics at the surface of and inside of solids, quantum coherent coupling of free electrons and light, and quantum-mechanically enhanced electron microscopy. He is a Fellow of the Max Planck Institute for the Science of Light and has received the Leibniz Prize of the German Research Foundation (DFG), the Leibinger Innovation Award, two ERC grants, and a Lynen Fellowship, amongst others. Jointly with Robert L. Byer of Stanford, he was Principal Investigator of the Moore Foundation-funded Accelerator on a Chip International Program (ACHIP).