

## Efficient coupling between free electrons and the supermode of a silicon slot waveguide: supplement

**YANNICK D'MELLO,<sup>1,\*</sup>  RAPHAEL DAHAN,<sup>2</sup> SANTIAGO BERNAL,<sup>1</sup> XIHANG SHI,<sup>2</sup> IDO KAMINER,<sup>2</sup>  AND DAVID V. PLANT<sup>1</sup>**

<sup>1</sup>*Department of Electrical Engineering, Faculty of Engineering, McGill University, Montreal, Canada*

<sup>2</sup>*Solid State Institute, Technion-Israel Institute of Technology, Haifa 32000, Israel*

\**Yannick.DMello@mail.McGill.ca*

---

This supplement published with Optica Publishing Group on 25 May 2023 by The Authors under the terms of the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/) in the format provided by the authors and unedited. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Supplement DOI: <https://doi.org/10.6084/m9.figshare.22811660>

Parent Article DOI: <https://doi.org/10.1364/OE.491452>

# Efficient coupling between free electrons and the supermode of a silicon slot waveguide: Supplementary Material

YANNICK D'MELLO,<sup>1,\*</sup> RAPHAEL DAHAN,<sup>2</sup> SANTIAGO BERNAL,<sup>1</sup> XIHANG SHI,<sup>2</sup> IDO KAMINER,<sup>2</sup> AND DAVID V. PLANT<sup>1</sup>

<sup>1</sup> Department of Electrical Engineering, Faculty of Engineering, McGill University, Montreal, Canada.

<sup>2</sup> Solid State Institute, Technion-Israel Institute of Technology, Haifa 32000, Israel.

\* [Yannick.Dmello@mail.McGill.ca](mailto:Yannick.Dmello@mail.McGill.ca)

## 1. Experimental Setup

The electron beam (eBeam) emitted by the ultrafast transmission electron microscope (UTEM) can be selected within a kinetic energy range of 80 keV to 200 keV, which corresponds to a range of sub-relativistic velocities,  $v_e$ , of  $0.5c_0$  to  $0.7c_0$  respectively, where  $c_0$  is the speed of light in vacuum. The UTEM can be operated in one of the following two modes. The thermionic emission mode delivers a flow of single electrons as random current once the temperature reaches the work function of the material, which is useful to study quantum interactions [1-6]. For 200 keV electrons in this mode, the emittance is 6 nm-mrad at, implying that the diameter of the spot size,  $d_e$ , can be 2-20 nm with a corresponding convergence semi-angle,  $\theta_e$ , of 3-0.3 mrad, respectively. Alternatively, the photoemission mode [7] delivers electron pulses, which are useful for testing dielectric laser accelerators (DLAs) [8-14]. In this mode, the emittance is 400 nm-mrad at 200 keV, which implies that  $d_e$  can be 20-200 nm with a corresponding  $\theta_e$  of 20-2 mrad, respectively. At a common  $d_e = 20$  nm, the thermionic emission mode was selected because its lower emittance allowed for a better overlap between the electron and optical supermode. The dimensions of the eBeam along its trajectory were calculated from these parameters using Rayleigh criterion. We assume a symmetric eBeam profile in the transverse dimension [7] such that the emittance in the X dimension (see Fig. 1 of the main manuscript for reference),  $\epsilon_x = \sigma_x \sigma_{ax} = \epsilon_y = \epsilon$  where  $\sigma_x$  and  $\sigma_{ax}$  are the width and convergence angle of the eBeam, respectively. Emittance remains constant for a specific eBeam energy. Within the emittance, the convergence angle and spot size can be modified via the numerical aperture of the magnetic lensing system. To adjust for different eBeam energies, we normalize the emittance [15],

$$\epsilon_n = \beta\gamma\epsilon = \beta\gamma\sigma_x\sigma_{ax} \quad (S1)$$

Where  $\beta = v_e/c_0$  and  $\gamma = 1/\sqrt{1-\beta^2}$  is the Lorentz factor. The electron energy (with relativistic correction) is,

$$\epsilon_e = (\gamma - 1)m_e c_0^2 \quad (S2)$$

Where  $m_e$  is the rest mass of the electron. The width of the eBeam is adjusted by the de Broglie wavelength of the electron,

$$\lambda_e = \frac{h_j}{\gamma m_e v_e} \quad (S3)$$

Where  $h_j$  is Planck's constant in Joules. Note that lower electron energies have a larger de Broglie wavelength which enlarges the spot size of the eBeam.

The eBeam undergoes magnetic lensing in order to steer and focus it onto the sample as well as separate the electron energies after the interaction. Since the electrons travel in a helical path, the on-chip device is surrounded by trenches so that the substrate can act as a mesa to ensure beam clearance. The protrusion of the mesa must be 50  $\mu\text{m}$  higher than the substrate, which is satisfied by the 675  $\mu\text{m}$  thickness of the Si substrate. Additionally, appropriate clearance is also required beyond the device due to the divergence and the helical trajectory of the eBeam. These

constraints limit the maximum effective interaction length,  $L_{\text{eff}}$ , to 500  $\mu\text{m}$ . After interaction, the electrons would be filtered via a magnetic prism and aperture to measure their energy distribution. The relative arrival time of the electron with respect to the laser pulse, and therefore the location of their interaction, can be determined by the relative delay between the pump and probe pulses.

## 2. Circuit Layout

We design a photonic integrated circuit to be compatible with the constraints of the UTEM. The layout is shown in Fig. 1 of the main text. Inside the UTEM, light can be focused onto the chip from a lens with a focal length of 50 mm. This implies that a collimated beam of 5 mm diameter can be focused down to a spot size of 30  $\mu\text{m}$  with a convergence angle of  $0.4394^\circ$ . Incident light is therefore coupled from free-space into the chip by a surface, vertical grating coupler (GC) [16]. We employed only one GC in order to reduce any coupling loss or phase mismatch between the arms of a dual-GC setup [17]. The GC is connected to a 90:10 splitter [18, 19] which routes 90% of the light to the device and uses 10% for feedback. In the 90% arm, light is sent to a mode evolution region (MER) in which the fundamental transverse magnetic ( $\text{TM}_0$ ) mode of the strip waveguide is converted into the desired supermode of the slot waveguide [20, 21]. A waveguide bend ensures the proximity of the MER to the edge of the device. A deep trench is carved into the chip at both ends of the device to act as a mesa and ensure that the diverging eBeam is not obstructed by the underlying BOX layer. The 10% arm of the splitter is connected to a Bragg grating (BG) reflector to provide feedback during the optical alignment of the chip. A secondary GC and BG are also placed next to the device to enable a coarse alignment. Once the incident angle is coarsely aligned using this secondary GC, the chip could be moved along the Z-axis to couple light into the primary GC and its alignment could be fine tuned based on feedback from the BG in the 10% arm.

As shown in Fig. 1(a) of the main manuscript, the input GC produces a  $\text{TM}_0$  mode which must then be coupled into the slot waveguide from a single sidewall (to prevent obstruction of the eBeam path), and with a specified electric field distribution. These requirements are addressed by the MER, which we design using supermode analysis [22, 23] for directional couplers and slot waveguides. In this design, a strip waveguide is connected from the 90% arm of the splitter to one of the sidewalls of the slot waveguide whereas the other sidewall was tapered and bent outward. The connected sidewall gradually evolves the effective index of the  $\text{TM}_0$  mode in the strip waveguide to become the desired supermode in the slot waveguide. The conversion efficiency to the target supermode is measured as the insertion loss (IL) from the power in the input  $\text{TM}_0$  mode to the desired supermode. We assume negligible energy leakage into the other modes which implies that they do not significantly interfere with the desired supermode. So, the phase relationship between the modes is not considered. Since the MER depends on the cross-section of the slot waveguide and its supermode, it must be redesigned for each supermode and waveguide cross-section. In cases where a strip waveguide is used for the interaction, that is,  $w_{\text{gap}} = 0$ , then the IL can be set as 0 dB since a MER is not necessary. In a future iteration of the design, the IL of the MER can be optimized by also injecting the supermode in the output to maximize the power for both forward and backward propagation. However, note that although the IL of the MER affects the overall interaction strength, it does not affect the coupling strength per photon because this value is normalized to the power of the optical pulse.

For example, an incident beam can be coupled into the GC with an IL of -9 dB [18] and guided through the variable intensity splitter and MER with ILs of -0.45 dB [19] and -10 dB, respectively. The optical pulse parameters in the slot are determined from the material constraints of the slot waveguide in the following section.

### 3. Optical Pulse Parameters

The parameters of the input optical pulse are constrained by the accelerating structure. Under specific conditions, confined photons in the pulse get absorbed by bound electrons in the structure. This photoionization pushes electrons into the conduction band [24, 25]. The conduction electrons get accelerated in the laser field and generate an avalanche from impact ionization [26]. This plasma is heated by the laser pulse, transferring the electron energy to the lattice. If its electron density increases to the point where the plasma frequency matches the laser frequency, then resonance occurs and the material gets damaged by ablation [25]. Although dielectrics can withstand orders of magnitude higher electric fields than metals, their damage threshold still often limits the acceleration gradient,  $G$ . It scales as [24],  $G \propto \sqrt{F_{\text{th}}}$ , where  $F_{\text{th}}$  is the damage threshold fluence of the material. The fluence is given by [26],  $F_{\text{th}} = \frac{2U_{\text{th}}}{\pi w_x w_y}$ , where  $U_{\text{th}}$  is the threshold pulse energy and  $w_x, w_y$  are the spot size dimensions. The threshold is further lowered by local field enhancements due to light being confined in the relatively small dimensions of the accelerating structure. For example, a silica grating has a damage threshold fluence of 1.85 J/cm<sup>2</sup> at 1500 nm, which is 0.53 times that of bulk silica (but still allows for a  $G$  of 300 MV/m) [25].

The optical damage threshold of the host material depends on the wavelength and pulse duration of the laser. Since it increases with the laser pulse width [27], we set it at 1 ps for our Si waveguide. The damage threshold of Si for a 1 ps pulse reaches a local maximum within the wavelength range of 1200-1600 nm [24, 25] as compared to 1064 nm [27] or 2000 nm [17] (all of which encompass the one- and two-photon ionization regimes). We therefore select two operating wavelengths of 1310 nm and 1550 nm, which also benefit from the availability of components and laser sources as well as the maturity of fabrication techniques.

For 1 ps pulses at a wavelength of 1550 nm, breakdown occurs around 0.2 J/cm<sup>2</sup> [26]. In the case of our slot waveguide, the Si core thickness is 220 nm with an average total width  $2w_{\text{Si}}$  of 500 nm, which occupies a cross-sectional area of  $1.1 \times 10^{-9}$  cm<sup>2</sup>. This allows for a maximum pulse energy of 0.22 nJ corresponding to a peak pulse power of 220 W in the slot. Considering an average loss of -19.4576 dB from the GC, splitter, and MER together (as calculated in the previous section), the energy of the free-space optical pulse that is incident on the chip must be 19.4169 nJ corresponding to 19.4169 kW.

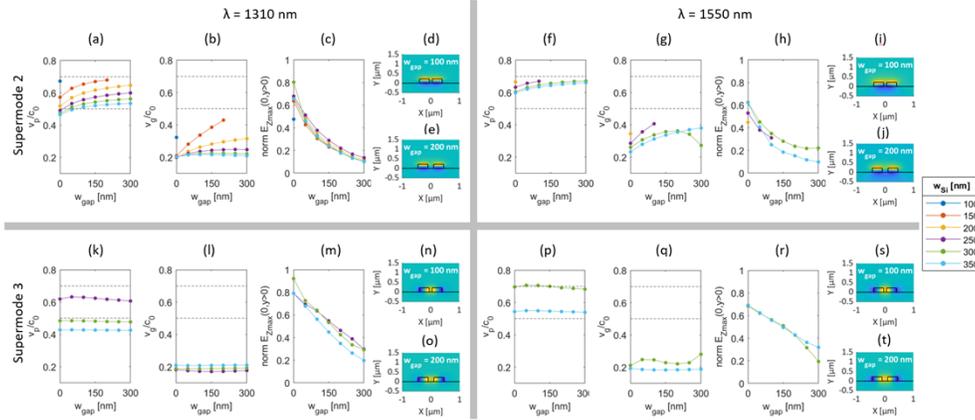
Such a high peak pulse power can induce strong nonlinear effects. In fact, these effects are even caused at lower powers in Si waveguides because the high refractive index of Si increases the confinement [28] and reduces the group velocity [29] of the optical mode, which enhances optical interactions with the Si core. Below 5 mW, intra-bandgap states induce single-photon absorption and as the optical power increases, the 1.12 eV bandgap of Si induces two-photon absorption which dominates beyond 300 mW [29]. Absorption of the optical energy consequently generates electron-hole pairs, which cause free-carrier dispersion (FCD) and free-carrier absorption (FCA). FCA is further enhanced by the fact that the free-carriers cannot diffuse out of the modal region in the waveguide [29]. It is determined by the effective modal area (calculated from the Poynting vector or the electric field intensity) [29] which corresponds to the surface-to-volume ratio of the waveguide [28]. This nonlinearity enhancement increases the TPA coefficient to 1.5 cm/GW [29] as compared to 0.4-1.2 cm/GW for bulk structures [28, 29], and the SPA coefficient to 1.9 m<sup>-1</sup> [29]. In this case, the recombination rate of the free-carriers limits the rate at which optical pulses can be generated. Carriers are found to have shorter lifetimes of ~800 ps initially, with longer lifetimes up to ~300 ns at later stages, assuming no initial carrier density [28]. In our case, since the IL of the GC already reduces the power by -9 dB when coupling into the chip, followed by a -10 dB loss at the MER, we assume these effects to be negligible in the interaction region. Future analysis could quantify these effects and even exploit FCA and FCD to control light within the Si structure [29].

## 4. Supermodes

The limits of  $v_e$  that are imposed by the UTEM also constrain the range of phase velocities  $v_p = \omega/k$  of the supermode and consequently, its effective refractive index,  $n_{\text{eff}}$ . For example, an eBeam of 80 keV or  $0.5c$  requires a  $n_{\text{eff}}$  of 1.9904 for the supermode. Similarly, an eBeam of 200 keV requires a  $n_{\text{eff}}$  of 1.4382 which is lower than the index of the buried oxide (BOX) layer. It is therefore unachievable using the standard SOI platform because the mode gets coupled into the BOX layer. The amplitude,  $E_0$ , of the electric field of the supermode,  $E_0 \cos(\omega t - kz)$ , is obtained from its individual field components as  $E_0 = \sqrt{E_X^2 + E_Y^2 + E_Z^2}$ . To account for a non-uniform electric field distribution (as opposed to a Gaussian beam profile), we simulate the electric field distribution and obtain its amplitude at each  $(x,y)$  point around the waveguide cross-section,  $E_0(x,y)$ . It is multiplied by the sign of the phase to retain the direction of oscillation of the field,

$$E_{0,Z} = |E_Z| \cdot \text{sgn}(\angle E_Z) \quad (\text{S4})$$

We then characterize the properties of the three lowest order modes for each variation of the waveguide geometry. These properties are the effective index,  $n_{\text{eff}}$ , group index,  $n_g$ , and the maximum amplitude of the longitudinal electric field component in the center of the gap,  $E_{0,Z}(0,y > 0)$ . Both  $n_{\text{eff}}$  and  $n_g$  were converted into the corresponding phase and group velocities of the supermode,  $v_p$  and  $v_g$ , respectively, to compare them directly with the electron velocity,  $v_e$ . We discard the first supermode because its  $E_Z$  field exhibits an antisymmetric distribution about the  $Y$  axis which results in  $E_{0,Z}(0,y > 0) = 0$  as well as opposite phases on either side, which renders it useless for coupling. Fig. S1 shows the change in  $v_p$ ,  $v_g$ , and  $E_{0,Z}(0,y > 0)$  for the remaining two supermodes as a result of varying the cross-sectional dimensions of the waveguide,  $w_{\text{Si}}$  and  $w_{\text{gap}}$ . The limits of  $v_e$  imposed by the UTEM are shown as black, dashed, horizontal lines at  $0.5c$  and  $0.7c$  in Fig. S1(a,f,k,p).



**Fig S1.** Supermode characteristics for each variation of the waveguide cross-section. (a)-(j) Variation of supermode 2 properties with slot parameters at wavelengths of (a)-(e) 1310 nm and (f)-(j) 1550 nm. (k)-(t) Variation of supermode 3 properties with slot parameters at wavelengths of (k)-(o) 1310 nm and (p)-(t) 1550 nm. In each set, the first two columns show  $v_p$  and  $v_g$  along with horizontal, black, dashed lines indicating the limits of  $v_e$ . The third column of each set shows the maximum  $E_z$  field at the center of the waveguide.

In addition to the upper limit on  $v_p$  imposed by the UTEM, the refractive index of the BOX layer also limits the minimum  $n_{\text{eff}}$  beyond which confinement of the supermodes is weakened. This corresponds to a  $v_p$  of  $0.6944c$ , which is approximately equal to the UTEM limit and therefore not shown. Additionally, we find that just beyond this limit, the supermodes still remain confined over short propagation lengths since majority of the mode field is distributed

in either Si or the external vacuum rather than in the BOX layer, that is, due to the vertical asymmetry of the waveguide cross-section. For both supermodes, the positive phase of their  $E_Z$  field component is above the waveguide due to the lower refractive index of vacuum compared to the underlying BOX layer. The distribution of  $E_Z$  around the waveguide depends on the cross-sectional dimensions of the slot. Narrower sidewall widths are unable to support supermode 3 especially at the longer wavelength of 1550 nm. Additionally, longer wavelengths are less confined, which generally decreases  $E_{0,Z}$  in the gap. This behaviour is evident in the comparison of field distributions between Fig. S1(c) and (h) for supermode 2, and equivalently Fig. S1(m) and (r) for supermode 3. The field amplitude also monotonically decreases as  $w_{\text{gap}}$  is increased as shown in Fig. S1(c,h,m,r). This is because narrowing  $w_{\text{gap}}$  squeezes the  $E_Z$  field between the sidewalls [30]. Although a higher  $E_{0,Z}$  is beneficial to the interaction, note that a wider  $w_{\text{gap}}$  accommodates a larger eBeam diameter and therefore a longer interaction length. As the confinement is weakened for a wide  $w_{\text{gap}}$  and thin  $w_{\text{Si}}$ , this lowers  $n_{\text{eff}}$  allowing  $v_p$  to reach the upper limit of the system. A higher  $v_p$  is more suitable for applications in which the final electron velocity, or kinetic energy, is important.

To obtain the total pulse power, we integrate the Poynting vector over the cross-sectional area of the simulation region since it is a local property of the optical field. This was done by summing the discretized power over the pixels of the simulation region, such that,

$$P = 0.5 \int \text{Re}(P_Z) dA = 0.5 \sum_{i=1}^{n_{\text{px}}} P_{Z,i} \cdot \Delta a_{\text{px}} \quad (\text{S5})$$

Where  $\Delta a_{\text{px}}$  is the area of each pixel. We evaluate only the  $P_Z$  component because the other two components are null since the Poynting vector points in the direction of propagation. The simulated, normalized E field was therefore scaled to match the input pulse power by,

$$E_{0,\text{exp}} = \sqrt{\frac{P_{\text{exp}}}{P_{\text{sim}}}} E_{0,\text{sim}} \quad (\text{S6})$$

## 5. Interaction Length

In a quantum interpretation of the electron, the spatiotemporal distribution of its wavefunction is larger than a classical point particle. Hence, we estimate the duration as 500 fs based on previous measurements in the UTEM (as indicated in the Supplementary Material of [1]). In these calculations, we assume the probability distribution of the electron to be uniform within the spot size of the eBeam since energy can be exchanged at any point in the cross-section. We also approximate the helical trajectory of the eBeam to be straight along  $L_{\text{eff}}$ . In the X-Y transverse dimension,  $L_{\text{eff}}$  is limited by the divergence of the eBeam within the slot. For example, since a larger  $d_e$  corresponds to a lower  $\theta_e$ , it allows a longer slot and therefore, longer interaction length,  $L_{\text{eff}}$ . The maximum  $L_{\text{eff}}$  that can be accommodated by the slot is calculated for every combination of eBeam diameter and height,

$$L_{\text{eff}} = \frac{2 \sqrt{(\min(h_e, t_{\text{Si}}) - t_{\text{Si}})^2 + \left(\frac{w_{\text{gap}}}{2}\right)^2}}{\tan \theta_e} \quad (\text{S7})$$

However, maximizing  $L_{\text{eff}}$  does not necessarily maximize the overall interaction. The interaction could benefit from either a small, focused spot size in a narrow slot width, or a low convergence angle over a longer interaction length, which indicates a trade-off. Along the Z propagation direction,  $L_{\text{eff}}$  is limited by dephasing as  $v_e$  is accelerated beyond  $v_p$ . It is also limited by the mismatch with  $v_g$  depending on the limited durations of the optical pulse and electron wavefunction, but this limitation is not as severe. This is why  $L_{\text{eff}}$  is a crucial variable in maximizing the coupling efficiency. In this process, the  $E_Z$  field distribution overlapping with the electron must also be maximized. As the eBeam converges and diverges through its focal

point, the hourglass shape is determined by the convergence angle, width, and depth of focus [31]. To account for these gradually changing conditions along the length of the interaction, we discretize  $L_{\text{eff}}$  into 10 segments and evaluate coupling using the average cross-sectional area of the eBeam in each segment. The sum of these values determines the overall coupling. Regarding the evolution of the supermode over the interaction region, we assume the pulse profile and peak power to undergo negligible dispersion or modulation throughout propagation [17].

## 6. Coupling Strength

When an optical pulse co-propagates with a free electron, its properties are imprinted onto the wavefunction of the electron by the phase-front of its longitudinal electrical field. Maximum energy transfer occurs when both the photon and electron are traveling in the same direction at the same speed, that is, when  $v_e$  is matched to  $v_p$ . The electron can then surf the plasma wake-field of the optical wavefront. This stimulated interaction modulates the kinetic energy spectrum of the electron. Adding or removing kinetic energy is equivalent to accelerating or decelerating the electron, respectively. In our case, we assume that the interaction is localized in the near field of the device where diffraction effects are insignificant. We therefore analyze it using PINEM theory [32, 33].

Exciting the charged tip inside the UTEM emits an electron traveling in the  $+Z$  direction.

Its wavefunction can be written as,  $\psi(\mathbf{r}, t) = \phi(\mathbf{r}, t)e^{i(k_0 z - E_0 t/\hbar)}$  with its behaviour described by the reduced Schrödinger equation,

$$\left(v \frac{\partial}{\partial z} - \frac{\partial}{\partial t}\right)\phi = -\frac{q_e v_e}{\hbar \omega} (e^{-i\omega t} E_Z - e^{i\omega t} E_Z^*)\phi \quad (\text{S8})$$

Here, we replace  $\nabla$  by  $ik_0 \hat{z}$ , and ignore any wavefunction gradients that do not contribute to the direction of propagation  $e^{ik_0 \hat{z}}$ . Hence, in the absence of any interaction, the electric field contributing to the electron energy is  $E_Z = 0$  resulting in an incident wavefunction of the form,

$$\phi(\mathbf{r}, t) = \phi_0(\mathbf{r} - v_e t \hat{z}) \quad (\text{S9})$$

Which satisfies the reduced Schrödinger equation above and accounts for  $v_e$  (or equivalently  $\varepsilon_e$  as per equation S2). We ignore transverse and mismatched components of the field due to their negligible effect from a co-propagating pulse. During the interaction, the electron wavefunction exchanges quanta of  $\hbar\omega$  energy to co-propagating photons. As per the PINEM effect [33], it has an equal probability of absorbing or emitting a photon and therefore gaining or losing energy, which causes it to accelerate or decelerate, respectively. Its spectrum therefore gets redistributed around a periodically spaced set of energies  $\varepsilon_0 \pm \ell \hbar\omega$ , which modifies equation S9 to,

$$\phi(\mathbf{r}, t) = \phi_0(\mathbf{r} - v_e t \hat{z}) \sum_{\ell} e^{i\ell\omega\left(\frac{z}{v_e} - t\right)} f_{\ell}(\mathbf{r}) \quad (\text{S10})$$

This represents a Fourier decomposition of the wave function in terms of its energy, which is quantized, or discretized, by the index  $\ell$ . In other words, the main Fourier coefficients  $f_{\ell}$  are spaced from each other by the laser frequency  $\hbar\omega$ , which allows them to be expressed as a summation instead of an integral. These  $f_{\ell}$  coefficients can be written in terms of Bessel functions as  $f_{\ell}(\beta) = e^{i\ell \arg(-\beta)} J_{\ell}(2|\beta|)$  and integrated along the  $+Z$  interaction length [34]. The probability of gaining or losing  $\ell$  quanta of photon energy is therefore  $P_{\ell} = |J_{\ell}(2|\beta|)|^2$ . The coupling strength over the total interaction length  $(-\infty, \infty)$  can then be obtained from the probability of absorbing or emitting a photon,

$$\beta(z) = \frac{q_e}{\hbar\omega} \int_{-\infty}^{\infty} E_Z(z) e^{-i\omega \frac{z}{v}} dz \quad (\text{S11})$$

Note that since  $\omega$  is in the denominator, this implies that lower frequencies (or equivalently, longer wavelengths) have a higher probability of interaction. However, since they exchange less  $\hbar\omega$  energy with the electron, the resulting coupling strength could still be lower. This could

explain why our design produces comparable results at both wavelengths of 1550 nm and 1310 nm alongside previously reported calculations at 2000 nm [17].

After the inelastic interaction with the laser field, the spectral distribution of  $\varepsilon_e$  can be quantified by the coupling or Rabi parameter [35],  $g$ , which represents the transition amplitude between adjacent electron energy states. This coupling constant,  $g$ , links the classical acceleration of an electron to the field of the optical mode. If  $g \gg 1$ , the electron energy width (or bandwidth) scales as  $4|g|$  times the optical photon energy [36]. Taking into account the spatial distribution of  $E_z$ , the interaction at each spatial co-ordinate  $(x_0, y_0)$  overlapping with the spot size of the eBeam can then be written in terms of this dimensionless coupling constant [33, 37, 38],

$$g(x_0, y_0) = \frac{q_e}{\hbar\omega} \int_{-\infty}^{\infty} E_z^{(\omega)}(x_0, y_0, z, t(z)) dz \quad (\text{S12})$$

The superscript  $(\omega)$  denotes the frequency dependence of the electric field and  $t(z)$  indicates the temporal delay with respect to the electron wavefunction. This integral represents the electron energy gain/loss normalized by the photon energy. It suggests that a maximum of  $2|g|$  photons are exchanged during the interaction, which implies both gain and loss resulting in a  $2|g|\hbar\omega$  energy spread of the electron spectrum [1].

For optimal coupling,  $E_z^{(\omega)}$  must be constant so that the interaction with the electron beam is maintained. In contrast, a DLA configuration would require the phase velocity of  $E_z$  to be increased along the interaction length to maximize the energy gain of the electron. The mode field distribution of the electric field  $E_z(x, y)$  was obtained from 3D FDTD simulations. The main contribution to the interaction is at the maximal field amplitude near the waveguide surface. At this proximity, the field travels approximately as a phasor so that the expression for  $E_z$  can be separated as  $E_{0,z}(x_0, y_0)e^{ik_z z}$  and the phase mismatch  $(k_z - \omega/v_e)$  can be separated in equation S12 as,

$$g(x_0, y_0) = \frac{q_e}{\hbar\omega} \int_{-\infty}^{\infty} E_{0,z}(x_0, y_0) e^{i(k_z - \omega/v_e)z} dz \quad (\text{S13})$$

Since energy can be exchanged at any  $(x, y)$  point in the region of overlap, we average the electric field over this region. Additionally, assuming  $v_e = v_p$ , the coupling strength along a fixed length simplifies to,

$$g = \frac{q_e}{\hbar\omega} E_{z,\text{avg}} L_{\text{eff}} \quad (\text{S14})$$

Note that this is only valid for length scales in which the eBeam spot size does not significantly change. Hence, this simplified form of  $g$  must be calculated separately for each discretized segment along  $L_{\text{eff}}$  depending on the convergence angle of the eBeam, as mentioned in the previous section. Additionally, as in equation S10, PINEM theory uses the approximation that the electron velocities remain constant during the interaction. However, when the electrons are accelerated, they no longer match  $v_p$ . This implies that  $L_{\text{eff}}$  is limited by the phase-matching condition,  $\frac{1}{v_e} \int (v_e(z) - v_p) dz < \frac{\pi v_p}{\omega}$  where  $\frac{\pi v_p}{\omega}$  is the half wavelength of the mode. Note that in this equation,  $v_e$ ,  $g$ , and  $L_{\text{eff}}$  are inseparable variables and the rigorous electron velocity  $v_e(z)$  is not known along its trajectory. Hence, we modify this condition to,

$$\frac{L_{\text{eff}}}{v_{e,\text{avg}}(g(L_{\text{eff}}))} (v_{e,\text{avg}}(g(L_{\text{eff}})) - v_p) \leq \frac{\pi v_p}{\omega} \quad (\text{S15})$$

Here,  $v_{e,\text{avg}}(g(L_{\text{eff}})) = \frac{(v_e(z=0) + v_e(z=L_{\text{eff}}))}{2}$  assuming linear acceleration and  $\frac{L_{\text{eff}}}{v_{e,\text{avg}}(g(L_{\text{eff}}))}$  is the interaction time. Moreover, since equation S14 does not account for the dependence of the pulse envelope, we effectively assume that the optical pulse and electron do not incur significant drift along a particular segment of  $L_{\text{eff}}$ . The validity of this assumption depends on whether the segment is short enough for the interaction to not be affected by a group velocity

mismatch or if the duration of the pulse is long enough for the driving-field amplitude to be relatively constant over the segment. Although the interaction is maximized when both conditions are satisfied, the phase modulation of the electron wavefunction is still expected to be uniform and deterministic for nonideal cases [35, 36]. The energy gained (or lost) by the electron can then be evaluated as [1],

$$\Delta\varepsilon_e = 2g\hbar\omega \quad (\text{S16})$$

With the acceleration gradient given by,

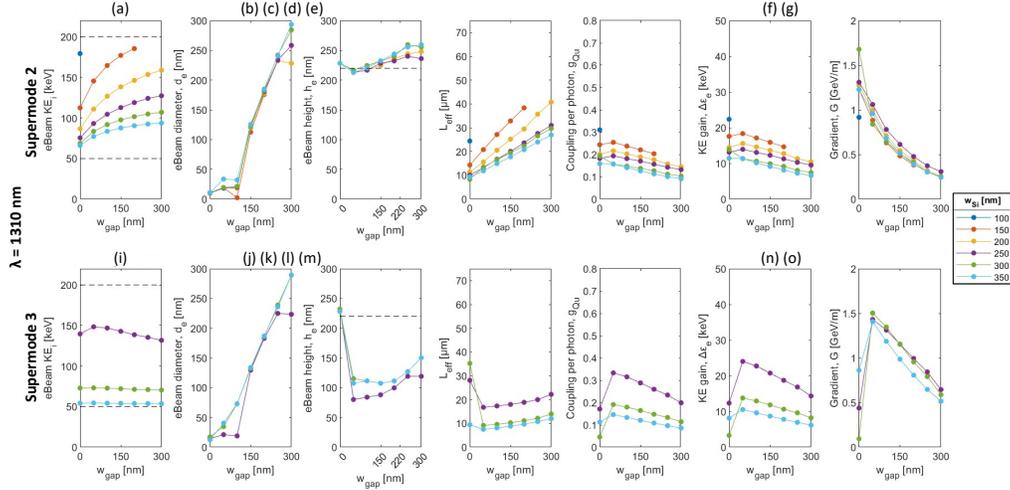
$$G = \Delta\varepsilon_e/L_{\text{eff}} \quad (\text{S17})$$

Note that this value is dependent on  $g$ , which is dependent on the power (energy and duration) of the optical pulse [2]. In a quantum-optics description [39], we consider PINEM driven by a coherent state  $|\alpha\rangle$  with an average number of  $|\alpha|^2$  photons. In the classical field of a mode with total energy per  $\mu\text{m}$ ,  $U$ , the average number of photons is  $\langle n \rangle = U/\hbar\omega$ . Therefore, we interpret the coupling strength per photon as [39],

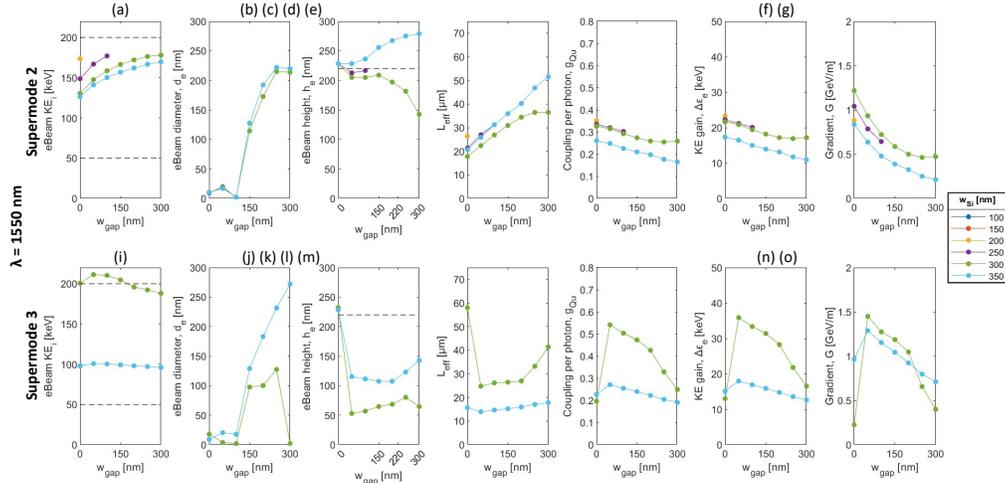
$$g_{\text{Qu}} = \frac{g}{\sqrt{n}} = \frac{g}{\alpha} = \frac{\Delta\varepsilon_e}{2\sqrt{P_{\text{slot}}T_{\text{pulse}}\hbar\omega}}, \quad (\text{S18})$$

which is a dimensionless parameter [35, 39] representing the proportion of the electron energy gain from one photon in the supermode. The regime of strong coupling is accessed by  $g_{\text{Qu}} \geq 1$  [39, 40]. In this manner,  $g_{\text{Qu}}$  is related to both the output kinetic energy gain of the electron and the input optical energy of the laser pulse, which therefore makes it a suitable indicator of coupling efficiency.

We confirm the validity of our reasoning and assumptions by applying them to the use case presented in this work. Specifically, we use these equations to find the maximum coupling strength per photon or equivalently, the coupling efficiency of the device design. It is done by characterizing the parameter space of the design in the following manner. Each set of slot dimensions,  $w_{\text{gap}}$  and  $w_{\text{Si}}$ , host specific supermode profiles with a corresponding effective index,  $n_{\text{eff}}$ , and phase velocity,  $v_p$ . This sets the electron velocity,  $v_e$ , which constrains the eBeam emittance, that is, its diameter,  $d_e$ , and convergence angle,  $\theta_e$ . The eBeam height from the BOX layer,  $h_e$ , is restricted to less than 500 nm since the interaction strength decayed exponentially with increasing separation [36]. Hence, the maximum  $L_{\text{eff}}$  is geometrically constrained by  $w_{\text{gap}}$ ,  $h_e$ , and  $d_e$ . These parameters are used to calculate the coupling constant,  $g$ , which is divided by the square root of the number of photons in the pulse,  $\alpha$ , [39] to produce the coupling constant per photon,  $g_{\text{Qu}}$ . In this manner,  $w_{\text{gap}}$ ,  $w_{\text{Si}}$ ,  $d_e \leftrightarrow \theta_e$ , and  $h_e$  are the independent variables in the parameter space;  $E_{0,z}$ ,  $v_e = v_p$ ,  $L_{\text{eff}}$ , and  $g$  are the dependent variables; and  $g_{\text{Qu}}$  is the figure of merit. We calculate  $g_{\text{Qu}}$  for both supermodes and both wavelengths over all variations of slot geometry to determine the optimal combination. The characterization of this parameter space is shown in Fig. S2 and Fig. S3.



**Fig S2.** Coupling results at  $\lambda = 1310$  nm for (a-g, top) supermode 2 and (i-o, bottom) supermode 3 hosted by a waveguide with thickness of 220 nm and varying wall and gap widths. Optimal eBeam parameters are shown as (a,i) initial KE,  $KE_i$ , (b,j) diameter,  $d_e$ , and (c,k) height,  $h_e$ . They overlap over (d,l) the effective interaction length,  $L_{\text{eff}}$ , to maximize (e,m) the coupling per photon,  $g_{\text{Qu}}$ , to produce the calculated (f,n) energy gain,  $\Delta\epsilon_e$ , and (g,o) acceleration gradient,  $G$ .



**Fig S3.** Coupling results at  $\lambda = 1550$  nm for (a-g, top) supermode 2 and (i-o, bottom) supermode 3 hosted by a waveguide with thickness of 220 nm and varying wall and gap widths. Optimal eBeam parameters are shown as (a,i) initial KE,  $KE_i$ , (b,j) diameter,  $d_e$ , and (c,k) height,  $h_e$ . They overlap over (d,l) the effective interaction length,  $L_{\text{eff}}$ , to maximize (e,m) the coupling per photon,  $g_{\text{Qu}}$ , to produce the calculated (f,n) energy gain,  $\Delta\epsilon_e$ , and (g,o) acceleration gradient,  $G$ .

For a strip waveguide,  $w_{\text{gap}} = 0$ , which implies that the MER is not needed and therefore a larger amount of optical energy can reach the interaction region, resulting in a higher  $g$  although this effect is normalized by  $g_{\text{Qu}}$ . For cases of a wide  $w_{\text{gap}}$ , since  $d_e$  is determined by  $v_e$ , it cannot be increased enough to maximize overlap with the regions where the amplitude of  $E_z$  is the highest, such as near the sidewalls. For a narrower  $w_{\text{gap}}$ , the eBeam is aligned directly above the waveguide whereas for a wider  $w_{\text{gap}}$ , the optimal  $h_e$  is inside the slot. An increase in  $w_{\text{gap}}$

accommodates an extended  $L_{\text{eff}}$  since larger eBeam dimensions can fit into a wider gap. It also generally increases  $v_g$  towards  $v_p$  (as seen in Fig. S1), which improves the temporal overlap. A longer  $L_{\text{eff}}$  improves the overlap and therefore the coupling strength. The highest  $E_Z$  amplitudes are found in supermode 3, but this higher order supermode can only be confined if  $w_{\text{Si}} > 250 \mu\text{m}$ . At a wavelength of 1550 nm, the weaker confinement allows for generally a higher  $g_{\text{Qu}}$ , especially for supermode 3. However, we must ignore the first 4 results for supermode 3 at 1550 nm, as shown in Fig. S3(i), because their corresponding  $n_{\text{eff}}$  and  $v_p$  are beyond the limits of the BOX and UTEM respectively. Despite this, supermode 3 still achieves the highest  $g_{\text{Qu}}$  of 0.4266 with a sidewall width of 300 nm and gap width of 200 nm, which induces the largest energy gain of 28.27 keV. This corresponds to an acceleration gradient of 1.05 GeV/m, which is one of the highest, but not the highest gradient achieved in this design scheme. The highest gradient of 1.68 GeV/m is achieved for supermode 2 at 1310 nm hosted by a sidewall width of 300 nm and a gap width of 0 nm as seen in Fig. S2(g) resulting in a  $g_{\text{Qu}}$  of only 0.1928 corresponding to an energy gain of 13.9 keV. Since both the gradient and gain are important performance metrics for such a device, our results confirm that maximizing  $g_{\text{Qu}}$  produces the best overall coupling efficiency.

On a chip, the interaction length might need to be shortened due to the space occupied by the waveguide bend and the MER. As a result, we expect a measured  $\Delta\varepsilon_e$  to be lower than the calculated value. However, note that before reaching the interaction region, the eBeam must anyway converge over a certain finite length between the surface of the BOX layer and the region above the slot waveguide. Hence, for future designs, this space might be sufficient to fit a compact, dispersion engineered [41] MER without significant loss of coupling strength.

## 7. Methods

The normalized  $E_{0,z}$  field was simulated in 2D using Ansys Lumerical Mode. Its calculated properties were then refined using a 3D simulation in Ansys Lumerical FDTD. The MER was also simulated and optimized in 3D using Ansys Lumerical FDTD. The mesh size of the simulation region was set at 8 nm, which determined the resolution of the simulated electric field. This imposed an incremental area  $dA$  of  $6.4 \times 10^{-17} \text{ m}^2$  over which the E-field was discretized. The simulated E-field was distributed over a specific number of pixels as determined by the resolution of the mesh and was normalized to a maximum amplitude of 1 V/m. These results were exported to MATLAB where the interaction with the electron was calculated and optimized.

## References

1. 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," *Nature Physics* **16**, 1123-1131 (2020).
2. J.-W. Henke, A. S. Raja, A. Feist, G. Huang, G. Arend, Y. Yang, F. 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).
3. R. Ruimy, A. Gorlach, C. Mechel, N. Rivera, and I. Kaminer, "Toward Atomic-Resolution Quantum Measurements with Coherently Shaped Free Electrons," *Physical Review Letters* **126**, 233403 (2021).
4. Y. Adiv, H. Hu, S. Tsesses, R. Dahan, K. Wang, Y. Kurman, A. Gorlach, H. Chen, X. Lin, G. Bartal, and I. Kaminer, "Observation of 2D Cherenkov radiation," (2022), p. arXiv:2203.01698.
5. R. Dahan, G. Baranes, A. Gorlach, R. Ruimy, N. Rivera, and I. Kaminer, "Creation of Optical Cat and GKP States Using Shaped Free Electrons," (2022), p. arXiv:2206.08828.

6. 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," (2022), p. arXiv:2202.12821.
7. A. Feist, N. Bach, N. Rubiano da Silva, T. Danz, M. Möller, K. E. Priebe, T. Domröse, J. G. Gatzmann, S. Rost, J. Schauss, S. Strauch, R. Bormann, M. Siviş, 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).
8. J. Breuer and P. Hommelhoff, "Laser-Based Acceleration of Nonrelativistic Electrons at a Dielectric Structure," *Physical Review Letters* **111**, 134803 (2013).
9. 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).
10. 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).
11. E. A. Nanni, W. R. Huang, K.-H. Hong, K. Ravi, A. Fallahi, G. Moriena, R. J. Dwayne Miller, and F. X. Kärtner, "Terahertz-driven linear electron acceleration," *Nature Communications* **6**, 8486 (2015).
12. J. McNeur, M. Kozák, N. Schönerberger, 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).
13. N. V. Sapra, K. Y. Yang, D. Verduyck, 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).
14. H. Tang, L. Zhao, P. Zhu, X. Zou, J. Qi, Y. Cheng, J. Qiu, X. Hu, W. Song, D. Xiang, and J. Zhang, "Stable and Scalable Multistage Terahertz-Driven Particle Accelerator," *Physical Review Letters* **127**, 074801 (2021).
15. "Beam Optics and Focusing Systems without Space Charge: Sections 3.1 - 3.5," in *Theory and Design of Charged Particle Beams* (2008), pp. 51-103.
16. Y. Wang, X. Wang, J. Flueckiger, H. Yun, W. Shi, R. Bojko, N. A. F. Jaeger, and L. Chrostowski, "Focusing sub-wavelength grating couplers with low back reflections for rapid prototyping of silicon photonic circuits," *Optics Express* **22**, 20652-20662 (2014).
17. Z. Zhao, T. W. Hughes, S. Tan, H. Deng, N. Sapra, R. J. England, J. Vuckovic, J. S. Harris, R. L. Byer, and S. Fan, "Design of a tapered slot waveguide dielectric laser accelerator for sub-relativistic electrons," *Optics Express* **26**, 22801-22815 (2018).
18. A. Zanzi, A. Brimont, A. Griol, P. Sanchis, and J. Marti, "Compact and low-loss asymmetrical multimode interference splitter for power monitoring applications," *Optics letters* **41**, 227-229 (2016).
19. E. El-Fiky, Y. D'Mello, Y. Wang, J. Skoric, M. G. Saber, A. Kumar, A. Samani, L. Xu, R. Li, D. Patel, and D. V. Plant, "Ultra-Broadband and Compact Asymmetrical Beam Splitter Enabled by Angled Sub-Wavelength Grating MMI," in *Conference on Lasers and Electro-Optics*, OSA Technical Digest (online) (Optical Society of America, 2018), STh4A.7.
20. N.-N. Feng, R. Sun, L. C. Kimerling, and J. Michel, "Lossless strip-to-slot waveguide transformer," *Optics Letters* **32**, 1250-1252 (2007).
21. R. Palmer, L. Alloatti, D. Korn, W. Heni, P. C. Schindler, J. Bolten, M. Karl, M. Waldow, T. Wahlbrink, W. Freude, C. Koos, and J. Leuthold, "Low-Loss Silicon Strip-to-Slot Mode Converters," *IEEE Photonics Journal* **5**, 2200409-2200409 (2013).
22. L. Chrostowski and M. Hochberg, *Silicon photonics design: from devices to systems* (Cambridge University Press, 2015).
23. K. Okamoto, *Fundamentals of optical waveguides* (Academic press, 2006).

24. K. Soong, R. L. Byer, E. R. Colby, R. J. England, and E. A. Peralta, "Laser damage threshold measurements of optical materials for direct laser accelerators," AIP Conference Proceedings **1507**, 511-515 (2012).
25. K. Soong, R. Byer, C. McGuinness, E. Peralta, E. Colby, and R. England, "Experimental determination of damage threshold characteristics of IR compatible optical materials," in *2011 Particle Accelerator Conference Proceedings*, (2011),
26. B. Cowan, *Optical damage threshold of silicon for ultrafast infrared pulses*, Boulder Damage Symposium XXXIX: Annual Symposium on Optical Materials for High Power Lasers (SPIE, 2007), Vol. 6720.
27. X. Wang, Z. H. Shen, J. Lu, and X. W. Ni, "Laser-induced damage threshold of silicon in millisecond, nanosecond, and picosecond regimes," Journal of Applied Physics **108**, 033103 (2010).
28. I. Aldaya, A. Gil-Molina, J. L. Pita, L. H. Gabrielli, H. L. Fragnito, and P. Dainese, "Nonlinear carrier dynamics in silicon nano-waveguides," Optica **4**, 1219-1227 (2017).
29. A. Gil-Molina, I. Aldaya, J. L. Pita, L. H. Gabrielli, H. L. Fragnito, and P. Dainese, "Optical free-carrier generation in silicon nano-waveguides at 1550 nm," Applied Physics Letters **112**, 251104 (2018).
30. V. R. Almeida, Q. Xu, C. A. Barrios, and M. Lipson, "Guiding and confining light in void nanostructure," Optics Letters **29**, 1209-1211 (2004).
31. D. Negi, L. Jones, J.-C. Idrobo, and J. Ruzs, "Proposal for a three-dimensional magnetic measurement method with nanometer-scale depth resolution," Physical Review B **98**, 174409 (2018).
32. B. Barwick, D. J. Flannigan, and A. H. Zewail, "Photon-induced near-field electron microscopy," Nature **462**, 902-906 (2009).
33. S. T. Park, M. Lin, and A. H. Zewail, "Photon-induced near-field electron microscopy (PINEM): theoretical and experimental," New Journal of Physics **12**, 123028 (2010).
34. F. J. García de Abajo, B. Barwick, and F. Carbone, "Electron diffraction by plasmon waves," Physical Review B **94**, 041404 (2016).
35. 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).
36. 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).
37. S. Nehemia, R. Dahan, M. Shentcis, O. Reinhardt, Y. Adiv, K. Wang, O. Beer, Y. Kurman, X. Shi, and M. H. Lynch, "Observation of the Stimulated Quantum Cherenkov Effect," arXiv preprint arXiv:1909.00757 (2019).
38. F. J. García de Abajo, A. Asenjo-Garcia, and M. Kociak, "Multiphoton Absorption and Emission by Interaction of Swift Electrons with Evanescent Light Fields," Nano Letters **10**, 1859-1863 (2010).
39. O. Kfir, "Entanglements of Electrons and Cavity Photons in the Strong-Coupling Regime," Physical Review Letters **123**, 103602 (2019).
40. V. Di Giulio, M. Kociak, and F. J. G. de Abajo, "Probing quantum optical excitations with fast electrons," Optica **6**, 1524-1534 (2019).
41. Y. D'Mello, O. Reshef, S. Bernal, E. El-fiky, Y. Wang, M. Jacques, and D. V. Plant, "Integration of periodic, sub-wavelength structures in silicon-on-insulator photonic device design," IET Optoelectronics **14**, 125-135 (2020).