Wigner wave packets: Transmission, reflection, and tunneling

Kevin L. Jensen ●[*]
Naval Research Laboratory, Washington, DC 20375, USA

Joel L. Lebowitz
Departments of Mathematics and Physics, Rutgers University, New Jersey 08854, USA

Jeanne M. Riga and Donald A. Shiffler ●
Directed Energy Directorate, Air Force Research Laboratory, Albuquerque, New Mexico 87117, USA

Rebecca Seviour ●
School of Computing and Engineering, University of Huddersfield, United Kingdom

A numerical solution to the time evolution equation of the Wigner distribution function (WDF) with an accuracy necessary to simulate the passage of a wave packet past a barrier is developed, where quantum effects require high accuracy and fine discretization. A wave packet incident on a barrier, a portion of which tunnels through, demonstrates behavior that can define various characteristic transmission and reflection delay (TARD) times useful in the simulation of electron emission. A model for the TARD times is proposed that relies only on the asymptotic maxima of the position $\rho(x, t)$ and wave number $\rho(k, t)$ densities given by the WDF and applied to a ballistic trajectory model for the (faster) transmitted and (slower) reflected parts. The dependence of the TARD times on barrier width, symmetry, and abruptness is analyzed. For symmetrical barriers with characteristics similar to field emission barrier heights and widths, TARD times are on the order of a fraction of a femtosecond. The TARD times for when tunneling predominates are contrasted to tunneling times in the literature. Use of the TARD times in simulations of field emission in nanogap devices or to model ultrashort pulses generated under rapidly changing conditions for electron sources are proposed.

I. INTRODUCTION

Processes associated with the field emission of electrons operate on disparate length scales that span many orders of magnitude [1–3]. Tunneling is exquisitely sensitive to the barrier shape [4–6], which in turn depends on microscale surface curvature [7–12], emitter shape [13–15], surface roughness [16–22], and nearest-neighbor (shielding) effects in arrays [23–25]. Simulating emission using, for example, particle-in-cell or molecular dynamics codes is already challenging because of the difficulty in reconciling particle transport with tunneling (wave) behavior [17,26–28]. Recent developments in nanogap emitters [29–32], thin insulators [33,34], and ultrafast emission [32,35–40] introduce further significant complications because the emission processes may be characterized by timescales [41] approaching tunneling time estimates [42–47]. However, a consensus on what the tunneling time is has not converged [47–54], making the assignment of a characteristic emission delay time difficult for simulation. Transmission and reflection delay (TARD) and tunneling time estimates are necessary if tunneling estimates [55–62] are to be simulated when space charge and transit time across the

anode-cathode gap with the barrier is present. In such a scenario, for example, an electron deemed eligible for emission at time t determined from a Poisson process [41] may not be released until time $t + \tau$, where τ accounts for the TARD time with the barrier considered herein. The TARD time is related to, but not identical with, what is conventionally referred to as a tunneling time, with various candidates proposed in the literature [52,53,63,64]).

Importantly for simulation purposes, Zimmerman *et al.* [53] concluded that "...*tunneling is unlikely to be an* instantaneous *process*" (emphasis added) which speaks to a simulation need and supports measurements suggesting a finite tunneling time [47,52,65]. The atomistic processes (tunneling from nanoscale emission sites over timescales comparable to the ratio of Planck's constant with a Rydberg energy, or $4\pi\hbar/m(\alpha_{fs}c)^2 = 0.304\,\text{fs}$) and phenomena affect the overall macroscale processes such as space charge [30,41,66], and understanding their relation is important for being able to predict overall cathode performance in the ultrafast and ultrasmall limits. Such conditions are not accurately modeled using conventional one-dimensional emission models such as the static Fowler-Nordheim equation [67–69] for predicting emission dynamics for well-understood reasons [9,10,25,41,70]. Simulation of thermal-field emission, in which tunneling is accompanied by over-the-barrier processes, introduces

*kjensen@mailaps.org

additional complications, but ones that must be considered to simulate emitters that become hot during their operation [3,71]. Atomistic phenomena, then, render purely steady-state approaches to modeling current emission inaccurate in nanogap devices and ultrafast conditions.

Initial efforts by us to model quantum mechanical tunneling dynamics on the atomistic level was begun by exploring a trajectory-based Wigner distribution function (WDF) approach [61,62], which was intended to provide a means for allowing microscale processes to be captured in simulations of electron emission in nanodevices. The WDF is a quantum distribution [72,73] but, by virtue of its analogy with a classical transport equation, it has analogs to a classical distribution of particles, and therefore is potentially useful to describing purely quantum mechanical behaviors such as resonant tunneling [56,74–79] or providing a link to Monte Carlo [80–82]. The exact models provided both analytic trajectories and time-dependent behavior in addition to describing intriguing quantum carpets [83] for a weighting of states modeled after a thermal distribution. Changes in the eigenstates associated with a step potential lead to trajectories whose tunneling behavior was examined. An exploration of the trajectory interpretation, however, is made difficult because there are few analytically solvable cases on which to test the methods. In our previous work, exact models for closed boundary conditions exhibited rapidly oscillating behavior that complicated a trajectory analysis. Although such rapid oscillations can be alleviated by, for example, invoking a Husimi distribution [73,84], doing so does not provide correct estimates of charge and current density.

Another trajectory approach for mitigating the problem of rapidly oscillating behavior was presented by Donoso and Martens [85]. This approach involves entangled classical trajectories for cases of smoothly varying potentials in which the state of the system evolves adiabatically. It is based on the hydrodynamic formulation of quantum mechanics originally proposed by Bohm [86,87] as a way of enabling a more efficient solution of the tunneling problem without sacrificing realism in the calculation [88–90]. The full Wigner function is represented by a smoothed Gaussian distribution function in this approach, which averages over the negative regions of the Wigner function, thereby eliminating problematic interference oscillations. Using this approximate methodology, Donoso and Martens were able to obtain good, qualitative agreement with the full quantum mechanical calculation for both the tunneling probability and the tunneling time as a function of initial mean ensemble energy. While both of these quantities were systematically overestimated by a few percent, the calculation was shown to be numerically stable, correctly reproduced the long-time growth of the tunneling probability, and resulted in an increase in computational speed by as much as an order of magnitude for the system considered. Lastly, Heim *et al.* [91], examined analogous phase space trajectories using a Wigner function for a parabolic barrier (for which the integral containing $V(x, k − k')$ simplifies as discussed in Sec. II D) for energy eigenstates and show how the Kemble form of the transmission probability [Eq. (C1)] naturally follows.

In contrast to the exact closed boundary models of our prior work, open boundary conditions are required to con-

sider current flow for electron emission models but require numerical approaches to the time evolution behavior for even simple barriers (compare to exact solutions to Schrödinger's equation (SE) [46] or alternate numerical methods [54,92]). Present methods are intended to allow addressing emission processes with short spatial and temporal dependencies [34,70,93] in a manner commensurate with the development of the moments of a distribution model uniting the theoretical description of the commonly used equations of electron emission, particularly thermal-field processes [94]. In contrast to the closed boundary condition studies with a thermal distribution, the present open boundary study focuses on Gaussian wave packets: As shown by Kluksdahl *et al.* [75], although thermal boundary conditions [56,74,95,96] and wave packets differ, predictions of processes associated with tunneling (particularly resonant tunneling) are similar for both [95] and therefore the wave packet studies may be profitably used to introduce a computationally useful time parameter that would enable including delays into the simulation of tunneling and field emission even in the absence of a trajectory model. Even so, numerical considerations limited the number of discretization points in phase space in prior treatments to values far too coarse for the demands of simulating TARD times here.

The organization of the paper is therefore as follows. First, a far more capable methodology for time evolving the WDF is developed that allows for accurate simulations capable of resolving the fine features associated with WDF evolution, in particular, the highly oscillatory features associated with quantum interference that are problematic for trajectory approaches. Second, a method for uniquely specifying a computationally useful TARD time that will aid in simulations is given: Although static (and symmetric) barriers are presently considered exclusively, the method in principle can be employed to treat oscillating field emission barriers [37] and interface barriers [97,98] (triangular or triangularlike) or transient conditions [99] of great interest and which are briefly discussed. Moreover, the capabilities enable the examination of the speeding up of the tunneling wave packet, and that examination is related to a simple δ-function barrier model. Third, the properties of various potentials are considered, from which Gaussian and parabolic barriers are focused upon for their desirable characteristics. Lastly, simulations are undertaken to describe the interference behavior and relate it to analytical models, and then characterize the TARD times for Gaussian and parabolic barriers and briefly compare them to the semiclassical time of Büttiker and Landauer, and the dwell time τ_d of Winful's analysis [63] (Appendix D).

II. WIGNER WAVE PACKETS

A. Time-dependent formulation

The single particle one-dimensional WDF is obtained from the density matrix $\hat{\rho}$ by [72,100]

$$f(x, k) \equiv \frac{1}{\pi} \int_{-\infty}^{\infty} dy \, e^{2iky} \langle x − y | \hat{\rho} | x + y \rangle, \qquad (1)$$

where x is the position coordinate and $\hbar k$ is the conjugate momentum coordinate, with notation and conventions following Refs. [61,69] but, briefly, are (i) to use units of [eV, fs, nm],

(ii) to absorb the unit charge q (the electron charge is $-q$) term into $V = q\varphi$ and $F = q|\mathcal{E}|$, where φ and $\mathcal{E}$ are the potential and electric field, respectively, (iii) the units of V are [eV] and of F are [eV/nm], so V and F are a potential energy and force, respectively, and (iv) all other terms have their usual representations and values (e.g., m for electron mass, etc.). Quantum nonlocal features are made manifest in the kets $|x \pm y\rangle$ on which $\hat{\rho}$ operates. The time evolution of the density matrix $\hat{\rho}$ is known to satisfy the relation [5,101,102]

$$i\hbar\partial_t\hat{\rho} = [\hat{H}, \hat{\rho}], \tag{2}$$

where $\hat{H} = \hbar^2\hat{k}^2/2m + V(\hat{x})$ is the Hamiltonian and the hat notation is used to reinforce that it is an operator. Combining Eqs. (1) and (2) shows that f satisfies [69]

$$\partial_t f(x, k, t) = -\frac{\hbar k}{m}\partial_x f(x, k, t)$$
$$+ \int_{-\infty}^{\infty} V(x, k - k')f(x, k', t)dk', \tag{3}$$

$$V(x, k - k') = -\frac{i}{\pi\hbar}\int_{-\infty}^{\infty} e^{2i(k-k')y}$$
$$\times \{V(x + y) - V(x - y)\}dy. \tag{4}$$

Observe that Eq. (3) can be written in the Liouville form $i\hbar\partial_t f = \hat{\mathcal{L}}f$, where $\hat{\mathcal{L}}$ is determined by that relation. Integration over all k recovers the continuity equation relating number density $\rho(x, t)$ to current density $J(x, t)$ given by

$$\partial_t\rho(x, t) + \partial_x J(x, t) = 0, \tag{5}$$

where $\rho(x, t)$ and $J(x, t)$ are defined by

$$\rho(x, t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x, k, t)\, dk, \tag{6}$$

$$J(x, t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\hbar k}{m} f(x, k, t)\, dk, \tag{7}$$

which resemble their classical counterparts. As given, ρ and J are number density and number current density. To make them charge and conventional current density, a factor of q must be included to account for electron charge ($-q$), as will be done below when explicitly needed. Similarly, a momentum density $\rho(k, t)$ is obtained by integration over x rather than k, resulting in

$$\rho(k, t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x, k, t)\, dx. \tag{8}$$

It is seen that number density is related to the zeroth moment in k, and current density to the first moment, of the distribution function. In terms of the Liouville operator $\hat{\mathcal{L}}$, Eq. (3) is [61]

$$f(x, k, t + \Delta t) = \exp(-i\hat{\mathcal{L}}\Delta t/\hbar)f(x, k, t). \tag{9}$$

Using the approximation $e^{-2\hat{X}} \approx (1 + \hat{X})^{-1}(1 - \hat{X})$, then the Cayley form [56]

$$[\hat{I} + ir\hat{\mathcal{L}}]f(x, k, t + \Delta t) = [\hat{I} - ir\hat{\mathcal{L}}]f(x, k, t) \tag{10}$$

is useful for implementing a time-evolution matrix formulation [56], and where $\hat{I}$ is the identity operator and $r = \Delta t/2\hbar$. The form of Eq. (10) is similar to the Crank-Nicolson implicit method for solving differential equations numerically [103] and encountered in numerical solutions of SE [54].

B. Free Gaussian wave packet

The minimum uncertainty Gaussian wave packet at $t = 0$ is (up to an arbitrary phase factor $e^{i\varphi}$) given by [5,73,104,105]

$$\langle x|\psi\rangle = \left(\frac{2}{\pi a^2}\right)^{1/2} e^{ik_o x - (x - x_o)^2/2a^2}. \tag{11}$$

When inserted into Eq. (1), then

$$\langle x + y|\psi\rangle\langle\psi|x - y\rangle = \frac{e^{-2ik_o y}}{\pi a^2}\exp\left[-\frac{(x - x_o)^2 + y^2}{a^2}\right].$$

Inserting $\hat{\rho} = |\psi\rangle\langle\psi|$ into the integration in Eq. (1) results in

$$f(x, k, 0) = \frac{2}{\sqrt{\pi}\hbar a}e^{-[(x - x_o)/a]^2 - [a(k - k_o)]^2}. \tag{12}$$

The initial state is $f(x, k, 0) \equiv f_o(x, k)$. For a free wave packet, then $V(x, k - k') = 0$, and so Eq. (3) entails [100] (compare Fig. 1 of Ref. [106])

$$f(x, k, t) = f_o\left(x - \frac{\hbar k t}{m}, k\right). \tag{13}$$

Another measure is the rise in density and current density at a particular point, chosen to be the origin ($x = 0$). For the parameters chosen, the shape of $J(0, t)$ closely resembles $\rho(0, t)$, and so consider instead the moment $\langle k - k_o\rangle$ evaluated for the distribution is $f(0, k, t)$, for which

$$\frac{\langle k - k_o\rangle}{\langle k_o\rangle} \equiv \frac{\int_{-\infty}^{\infty}(k - k_o)f(x, k, t)dk}{\int_{-\infty}^{\infty}k_o f(x, k, t)dk}$$
$$= \frac{J(x, t) - qv_o\rho(x, t)}{qv_o\rho(x, t)}, \tag{14}$$

where $v_o = \hbar k_o/m$ and the factor q for unit charge has been explicitly included to make J a charge current density. It follows $J(0, t) > q(\hbar k_o/m)\rho(0, t)$ for $t < t_o$, expected because the head of the wave packet contains a greater proportion of the higher momentum states. The rotation of the ellipse in phase space corresponds to the spreading of the wave packet's wave function determined from SE [5,100] and shall be revisited in Sec. IV A.

C. Interference

A simple model of free wave packet interference provides an analytic model anticipating behavior shown by a wave packet interacting with a potential barrier below, and is therefore treated explicitly. Let the wave function of Eq. (11) be modified to describe two approaching wave packets, both characterized by a center wave number k_o. Introduce the term $\varsigma = \pm 1$, and define

$$\langle x|\psi_\varsigma\rangle = \left(\frac{2}{\pi a^2}\right)^{1/2} e^{i\varsigma k_o x - (x - \varsigma x_o)^2/2a^2} \tag{15}$$

It is seen that Eq. (11) corresponds to $\langle x|\psi_+\rangle$. The mixed state for two Gaussian wave packets is then, following Weinbub and Ferry [107],

$$|\psi\rangle = \frac{1}{\sqrt{2}}\{|\psi_+\rangle + |\psi_-\rangle\}. \tag{16}$$

Now, $|\psi\rangle\langle\psi|$ is explicitly

$$2|\psi\rangle\langle\psi| = \{|\psi_+\rangle\langle\psi_+|\} + \{|\psi_-\rangle\langle\psi_-|\}$$
$$+ \{|\psi_+\rangle\langle\psi_-|\} + \{|\psi_-\rangle\langle\psi_+|\}, \qquad (17)$$

with the first two (top) are direct terms and the last two (bottom) are cross terms. The direct contribution to $f(x, k)$ is easily anticipated from Eq. (12) and is

$$f_d(x, k) = f_+(x, k) + f_-(x, k),$$
$$f_\pm(x, k) = \frac{1}{\sqrt{\pi}\hbar a} e^{-[(x\mp x_o)/a]^2 - [a(k\mp k_o)]^2}, \qquad (18)$$

and visualized as two noninteracting Gaussian wave packets, one centered at $(+x_o)$ and moving to the right, and the second at $(-x_o)$ and moving to the left (note that x_o itself can be separately positive or negative). The cross term can be shown to be [108]

$$f_c(x, k) = \frac{2}{\sqrt{\pi}\hbar a} e^{-(x/a)^2 - (ak)^2} \cos(k_o x - k x_o). \qquad (19)$$

It is centered at the origin of phase space and oscillates rapidly when k_o is large: Weinbub and Ferry [107] designated it as the *central entanglement term* but in the present paper it will instead be termed the *central interference term* because of how entanglement differs from coherence, even though both involve an interference term. The orientation of the ridges is seen to be dictated by the argument of the cos term in Eq. (19) when the time evolution of Eq. (13) is invoked: the argument of that term is constant when $k/x = k_o/x_o$. In contrast, the direct terms of Eqs. (18) are seen to correspond to the wave packets as they move away from the origin: They do not give negative contributions, in contrast to the cross term as a consequence of the cosine term. The Wigner function $f(x, k) = f_d(x, k) + f_c(x, k)$ for various x_o and k_o, scaled by a, is shown in Fig. 1. The clearly visible central interference term concentrated near the origin anticipates features that appear when a wave packet incident on a barrier separates into transmitted and reflected portions.

Significantly, the central interference term is such that both $\rho(k)$ and $\rho(x)$ are not appreciable near the origin because the rapid oscillations of the central interference term average to small values. For example, along $x = 0$, the contributions to $\rho(x)$ from the central interference term from Eq. (6) go as

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} f_c(0, k) \, dk = \frac{1}{\pi\hbar} e^{-(k_o a/2)^2}, \qquad (20)$$

and therefore are exponentially small as $k_o a$ increases, with an analogous evaluation for $\rho(k)$. Although the central interference term will alter in the presence of a barrier, nevertheless, it is expected that $\rho(x)$ and $\rho(k)$ will be dominated by the transmitted and reflected portion of a wave packet incident on a barrier.

D. Barriers

If $V(x)$ is linear, that is, $V(x) = -Fx$, then $V(x + y) - V(x - y) = -2Fy$. Similarly, if $V(x)$ is quadratic, that is, $V(x) = (\gamma/2)x^2$, then $V(x + y) - V(x - y) = 2\gamma xy$. In either
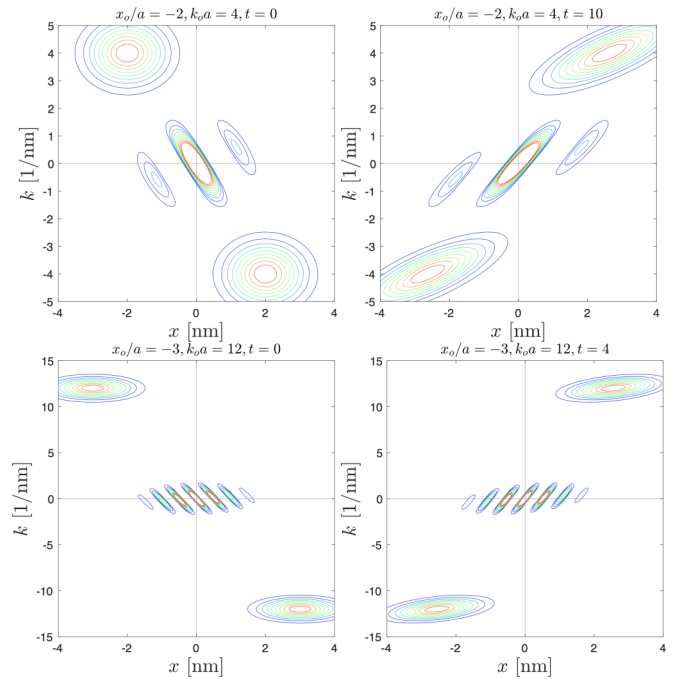

FIG. 1. $f(x, k) = f_d(x, k) + f_c(x, k)$ for a wave function of two Gaussian wave packets traveling in opposite directions, evaluated using the analytic Eqs. (18) and (19) for various x_o and k_o, evaluated at $t = 0$ (left) and t (right). Only contours for $f(x, k) > 0$ are shown with the jth contour of 10 at $(j/10)f_+(x_o, k_o)$.

case, the integral in $V(x, k)$ of Eq. (4) is proportional to

$$-\frac{i}{2\pi\hbar} \int_{-\infty}^{\infty} e^{iks} s \, ds = -\frac{1}{\hbar} \partial_k \delta(k), \qquad (21)$$

where $\delta(k)$ is the Dirac delta function. As a result, the integral term in Eq. (3) becomes, for the linear case, $(F/\hbar)\partial_k f(x, k, t)$, and for the quadratic case, $F \to -\gamma x$. It is also seen that Eq (4) can be written

$$\partial_t f(x, k, t) = -\dot{x} \, \partial_x f - \dot{k} \, \partial_k f, \qquad (22)$$

where $\dot{x} = \hbar k/m$ and $\dot{k} = -\partial_x V(x)$, which can then be used to construct the classical trajectories for linear fields and harmonic oscillators. In the linear case, it follows [69,100]

$$f(x, k, t) = f_o\left(x - \frac{\hbar k}{m}t + \frac{F}{2m}t^2, k - \frac{F}{\hbar}t\right), \qquad (23)$$

for which the $F = 0$ case corresponds to a free wave packet. Figure 1 evidences this behavior by sheering the ellipsoid because the phase space points with higher k move faster to the right (in the $+\hat{x}$ direction): Were the initial condition ellipsoids to straddle the $k = 0$ boundary, all phase space points for which $k < 0$ would likewise move leftward, unaffected by the evolution of the $k > 0$ phase space points. This casual observation is behind the form of the optimal numerical scheme leading to the time evolution of the wave packet even when a barrier is present. The behavior of $V(x, k)$, then, governs in what manner quantum mechanics alters those trajectories in a phase space approach, but the behavior of the free wave packet suggests a useful scheme to solve Eq. (3).

By far, the simplest barrier to consider is the δ-function barrier of Eq. (A1), for which

$$V_d(x, k - k') = -\frac{4\lambda V_o}{\pi \hbar} \sin[2(k - k')x] \qquad (24)$$

[compare Eq. (A7)], a form which allows an analytic solution to $f(x, k, t)$ for closed boundary conditions [61,62]. Open boundary conditions [48,76,109], however, require additional considerations that make the time evolution simulations difficult, and therefore, numerical means are advantageous. For an abrupt potential like $V_d(x)$, however, ripples appear in $f(x, k, t)$ for large values of k, particularly near regions where $V(x)$ undergoes abrupt changes in magnitude or slope that cannot be ignored (e.g., Figs. 2 and 5 of Ref. [61]), and which undermine finite difference methods applied to Eq. (3). Abruptness is manifested as the extent of $V(x, k)$ along the k dimension, and for $V_d(x)$, it does not diminish as do the other potentials of Eqs. (A2)–(A5), thereby complicating its use in numerical methods that consider only a portion of k space. The next simplest form of a rectangular barrier also results in highly oscillatory behavior for large k of $f(x, k, t)$ near the boundaries of the barrier (e.g., Figs. 8 and 9 of Ref. [62]). Of the simple candidates available, considered in Appendix A, those potentials that do not exhibit large or discontinuous changes in $V(x)$ and $\partial_x V(x)$ (that is, they are not abrupt) and are symmetrical most strongly reduce the behavior of $V(x, k - k')$ that lead to large k behavior in $f(x, k, t)$ and therefore complicate numerical accuracy. Consequently, attention here is confined to the Gaussian $V_g(x)$ and parabolic $V_p(x)$ potentials of Eqs. (A3) and (A4), as they are sufficiently smooth in the sense of Fig. 24.

III. WIGNER TIME EVOLUTION

A. Explicit scheme

The numerical solution of the time-dependent WDF equation of Eq. (3) with open boundary conditions [56,76] using a direct solution [110] proceeds by defining $f(x, k, t)$ at equispaced points in position ($x \rightarrow x_i$), wave number ($k \rightarrow k_j$), and time ($t \rightarrow t_n$) via

$$x_i = (2i - N - 1)\Delta x/2, \quad k_j = (2j - N_k - 1)\Delta k/2,$$
$$t_n = n\,\Delta t, \qquad (25)$$

using standard finite-difference methods [103]. The time taken to evaluate $f(x, k, t)$ increases substantially as the terms (N_x, N_k, N_t) increase, as the size of the resulting $\mathbf{L}$ operator (the discrete form of the Liouville operator $\hat{\mathcal{L}}$) rapidly increases, and the matrix solution slows. Moreover, the boundaries must be sufficiently far away that the incoming distribution into the simulation region can be approximated, usually by a thermal supply function derived from a Fermi-Dirac distribution when simulating a resonant tunneling diode (RTD) [56,61,111]. The problem is revisited here with the intent to develop a far more rapid numerical solution capable of sufficient accuracy to treat the wave-packet problem, enabling the values of (N_x, N_k) to be enlarged by an order of magnitude but an iteration performed in a fraction of the time of previous methods on which the present approach is based, e.g., a single time step took 3.2 s for $(N_x, N_k) = (86, 72)$ on a mainframe

whereas the present approach takes 0.042 s per iteration for $(N_x, N_k) = (512, 128)$ on a desktop [112].

The values of Δx and Δk are related by requirements on how to treat the $\sin[2(k_j - k_{j'})x_i]$ term in $V(x_i, k_j - k_{j'})$. Substitution of Eqs. (25) into the $\sin[2(k - k')x]$ part gives

$$\sin[2(k_j - k_{j'})x_i] \equiv \sin\left[2\pi\frac{(j - j')(i - 1)}{N_x - 1}\right], \qquad (26)$$

if $\Delta x = L/(N_x - 1)$, where L is the width of the simulation region, and $\Delta k = 2\pi/L$. This choice insures that $V(\pm L/2, k_j - k_{j'}) \equiv 0$, that is, $V(x, k - k')$ vanishes on the boundaries $x = \pm L/2$ and the choice of Δk enables the numerically advantageous form of Eq. (26).

Next, define $f(x, k, t)$ and $V(x, k - k')$ on those points. This shall be done by introducing indices on f and V such that

$$f(x, k, t) \rightarrow f(x_i, k_j, t_n) \equiv \left[\mathbf{f}_i^n\right]_j, \qquad (27)$$

$$V(x, k - k') \rightarrow [\mathbf{V}_i]_{j,j'}. \qquad (28)$$

Integrations over k' become summations over j' for the multiplication of the $\mathbf{V}_i$ matrix with the $\mathbf{f}_i^n$ vector. The derivatives become, in the finite difference approach,

$$\partial_t f(x, k, t) \rightarrow \frac{\mathbf{f}_i^{n+1} - \mathbf{f}_i^n}{\Delta t},$$
$$\partial_x f(x, k, t) \rightarrow \mp\frac{1}{2\Delta x}\left[3\mathbf{f}_i^n - 4\mathbf{f}_{i\pm1}^n + \mathbf{f}_{i\pm2}^n\right], \qquad (29)$$

where the spatial derivative uses an upwind/downwind second-order differencing scheme (SDS) [113]. The notation will be that bold face quantities (e.g., $\mathbf{M}$, $\mathbf{f}_i^n$) are matrices or vectors, and terms such as $\mathbb{D}$ are operators. In describing the finite differencing scheme below, the introduction of a difference operator that acts on the i index by

$$\mathbb{D}(f_{i,j}^n) = \tfrac{1}{2}\left[3f_{i,j}^n - 4f_{i\pm1,j}^n + f_{i\pm2,j}^n\right], \qquad (30)$$

where ($+$) is used for $k_j < 0$ and ($-$) is used for $k_j > 0$. Introduce the matrix $\boldsymbol{\alpha}$ defined by

$$[\boldsymbol{\alpha}]_{jj'} = \frac{\hbar|k_j|\Delta t}{4m\Delta x}\delta_{j,j'} \equiv \alpha_j \delta_{j,j'}, \qquad (31)$$

where $\delta_{jj'}$ is the Kronecker delta function. Using a simple *explicit* Euler scheme for the time derivative, then the vectors $\mathbf{f}_i^{n+1}$ are straightforwardly deduced from $\mathbf{f}_i^n$ by the equation

$$\mathbf{f}_i^{n+1} - \mathbf{f}_i^n = -\boldsymbol{\alpha} \cdot \left[3\mathbf{f}_i^n - 4\mathbf{f}_{i\pm1}^n + \mathbf{f}_{i\pm2}^n\right] + \mathbf{V}_i \cdot \mathbf{f}_i^n.$$

This equation can be rewritten as

$$\mathbf{f}_i^{n+1} = [\mathbf{I} - 2\boldsymbol{\alpha}\mathbb{D} + \mathbf{V}_i] \cdot \mathbf{f}_i^n, \qquad (32)$$

where $\mathbf{I}$ is the identity matrix. Such an *explicit* scheme is accurate to order $O(\Delta t)$.

B. Implicit scheme

To improve the accuracy to order $O(\Delta t^2)$, all instances of $\mathbf{f}_i^n$ on the right-hand side of Eq. (32) can be replaced by the average $(\mathbf{f}_i^{n+1} + \mathbf{f}_i^n)/2$, resulting in an *implicit* matrix equation cast in the form of Eq. (10) and solvable using matrix solution techniques based on LU decomposition [56]. The presence of $\mathbf{V}_i$ requires the inversion of matrices of size $N_k \times N_k$ for

$N_x \times N_t$ times. The form of the SDS differential operator, however, makes possible an alternate approach that spares the requirements for such an investment of computational time and resources. It makes use of the formulation used to solve for $\mathbf{f}_i^{n+1}$ using an implicit scheme for a free wave packet.

For the free case, the implicit matrix equation when $V_o = 0$ is

$$[\mathbf{I} + \boldsymbol{\alpha}\mathbb{D}] \cdot \mathbf{f}_i^{n+1} = [\mathbf{I} - \boldsymbol{\alpha}\mathbb{D}] \cdot \mathbf{f}_i^n \equiv \mathbf{b}_i, \qquad (33)$$

because $\mathbb{D}$ acts in a manner dependent on the jth element within $\mathbf{f}_i^n$, it is profitable to represent $\mathbf{f}_i^n$ as the combination of two vectors of size $N_k/2$ corresponding to regions where $k_j < 0$ (leading to terms with an l subscript) and where $k_j > 0$ (leading to terms with a u subscript) so

$$\mathbf{b}_i \equiv \begin{bmatrix} \mathbf{l}_i \\ \mathbf{u}_i \end{bmatrix}. \qquad (34)$$

Next, because $\mathbf{M} = \mathbf{I} + (3/2)\boldsymbol{\alpha}$ only has diagonal entries, then it, too, can be separated into parts as can $\boldsymbol{\alpha}/2$, or

$$\mathbf{I} + \frac{3}{2}\boldsymbol{\alpha} = \mathbf{M} = \begin{bmatrix} \mathbf{M}_l & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_u \end{bmatrix}; \quad \frac{\boldsymbol{\alpha}}{2} = \begin{bmatrix} \mathbf{L} & \mathbf{0} \\ \mathbf{0} & \mathbf{U} \end{bmatrix}, \qquad (35)$$

where $\mathbf{0}$ is the $(N_k/2) \times (N_k/2)$ zero matrix, and $\mathbf{L}$ and $\mathbf{U}$ also only have diagonal entries [114]. As a result, Eq. (33) becomes two separate matrix equations

$$\mathbf{M}_l \cdot \mathbf{l}_i^* = \mathbf{l}_i + \mathbf{L} \cdot [4\mathbf{l}_{i-1}^* - \mathbf{l}_{i-2}^*],$$
$$\mathbf{M}_u \cdot \mathbf{u}_i^* = \mathbf{u}_i + \mathbf{U} \cdot [4\mathbf{u}_{i+1}^* - \mathbf{u}_{i+2}^*], \qquad (36)$$

where the action of $\mathbb{D}$ is explicitly shown. Importantly, both $\mathbf{l}_i$ and $\mathbf{u}_i$ are evaluated sequentially and rapidly: Because $[\mathbf{M}^{-1}]_{j,j} = 1/[\mathbf{M}]_{j,j'} = \delta_{j,j'}/(1 + 3\alpha_j/2)$, no costly matrix inversions are required. All $\mathbf{l}_i^*$ and $\mathbf{u}_i^*$ in Eqs. (36) which have indices $i < 1$ or $i > N_x$, respectively, are set to zero vectors because there are no incident contributions from the boundaries (in contrast to simulations of RTDs [113], where they are conventionally set to a thermal supply function, although doing so may not be accurate [61]). Finally,

$$\mathbf{f}_i^{n+1} = \begin{bmatrix} \mathbf{l}_i^* \\ \mathbf{u}_i^* \end{bmatrix} \qquad (37)$$

in the free wave packet ($V_o = 0$) case. By virtue of the chosen structure of the $(\mp \Delta x \, \partial_x \rightarrow \mathbb{D})$ differential operator of Eq. (30), the solution represented by Eq. (37) for the free propagation case ($V_o = 0$) can be computed with the speed of the explicit scheme associated with Eq. (32), and *the solution of $\mathbf{l}_i^*$ is independent of $\mathbf{u}_i^*$* (they can be evaluated separately). This is a consequence of the ability to separate the larger matrix equation into noninteracting upward and downward elimination schemes (as shown in Fig. 2) enabled by the upwind/downwind SDS differential operator, and reflects that the positive and negative k regions of $f(x, k, t)$ in Eq. (13) evolve without any interaction when $V_o = 0$. The diagonal nature of $\mathbf{M}$ insures that no time-consuming matrix inversions are required. Retaining that speed advantage requires that a means to bypass those matrix inversions when $V_o \neq 0$, which links $\mathbf{l}_i^*$ and $\mathbf{u}_i^*$, must be developed.
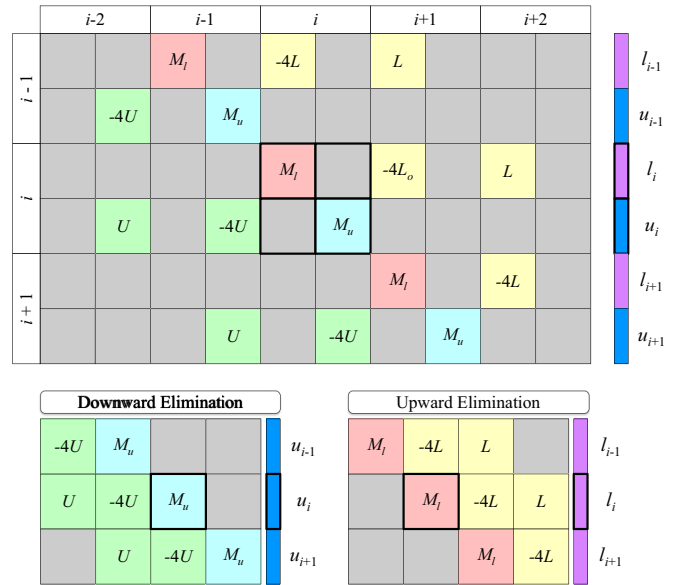

FIG. 2. Reorganization of the matrix equations for $\mathbf{f}_i^n$ into two smaller matrix equations with entries entirely above or below the diagonal. Size of each box is $(N_k/2) \times (N_k/2)$. Center diagonal boxes for ith row and $(j = i)$ column designated by thicker black borders. An l subscript and L matrix refer to $k_j < 0$; a u subscript and U matrix to $k_j > 0$. $\mathbf{L}$, $\mathbf{U}$ and $\mathbf{M}$ have only nonzero diagonal elements.

The inclusion of a barrier ($V_o > 0$) is accomplished by first solving Eq. (32) for the predictor $\mathbf{f}_i^*$. The correction then incorporates $\mathbf{f}_i^*$ in a modified potential term that is now included with the $\mathbf{b}_i$ vector. That is,

$$\mathbf{f}_i^* = [\mathbf{I} - 2\boldsymbol{\alpha}\mathbb{D} + \mathbf{V}_i] \cdot \mathbf{f}_i^n,$$
$$\mathbf{b}_i = [\mathbf{I} - \boldsymbol{\alpha}\mathbb{D}] \cdot \mathbf{f}_i^n + \tfrac{1}{2}\mathbf{V}_i \cdot (\mathbf{f}_i^* + \mathbf{f}_i^n), \qquad (38)$$

where the solution of $\mathbf{f}_i^{n+1}$ unfolds according to Eqs. (36) and (37). Although iterations along the lines of higher order Runge-Kutta methods [115] could be used to improve the corrector (thereby enabling larger time steps), here, one prediction and one correction are employed and smaller time steps are used as a matter of simplicity.

A close consideration of Eqs. (38) reveals that an additional approximation may be made that substantially increases the speed of simulations. The presence of $\sin[2(k_j - k_{j'})x_i]$ in $\mathbf{V}_i$ means that far enough away from the potential barrier, the contribution of $\mathbf{V}_i \cdot (\mathbf{f}_i^* + \mathbf{f}_i^n)$ to $\mathbf{b}_i$ becomes negligible, reflecting that the rapid oscillations of $V(x, k - k')$ caused by $\sin[2(k - k')x]$ makes $\int V(x, k - k')f(x, k', t)dk'$ negligible. Because the $\mathbf{V}_i$ matrix is the only term connecting the $k > 0$ region of phase space to the $k < 0$ region, the structure shown in Fig. 2 entails that the iterative solutions of Eqs. (36) need not be extended all the way to the boundaries when L is large, as is required to ensure that Δk is small. As a result, until the presence of the barrier impacts $f(x, k, t)$, the free propagation equations of Eq. (33) can be used. However, because the numerical solution of Eq. (38) already provides a substantial computational savings, utilization of this approximation is not done herein.

C. Density and current density

The evaluation of density proceeds straightforwardly from Eq. (6) and is

$$\rho(x,t) \to \rho_i^n = \frac{\Delta k}{2\pi} \sum_{j=1}^{N_k} f_{i,j}^n, \tag{39}$$

with an analogous equation holding for $\rho(k,t) \to \rho_j^n$, for which the summation is over i and Δk is replaced by Δx. By virtue of the j-dependent operator $\mathbb{D}$ of Eq. (30), the current density $J(x,t)$ is different. For steady-state conditions, the continuity equation $\partial_t \rho + \partial_x J = 0$ becomes, in discrete form,

$$\rho_i^{n+1} - \rho_i^n = 0 = -\frac{\Delta t}{\Delta x} \sum_{j=1}^{N_k} \frac{\hbar k_j}{m} \mathbb{D}(f_{i,j}^n). \tag{40}$$

Because $\mathbb{D}$ acts on $f_{i,j}^n$ differently depending on whether $k_j \lesseqgtr 0$ (which is the same as $2j \lesseqgtr N_k + 1$), the summation over j separates into two summations depending on the sign of k_j. In those summations, reordering Eq. (30) gives [56,77]

$$2\mathbb{D}(f_{i,j}^n) = (3f_{i,j}^n - f_{i\pm1,j}^n) - (3f_{i\pm1,j}^n - f_{i\pm2,j}^n). \tag{41}$$

Defining $\Delta x \partial_x J_i = J_{i+\frac{1}{2}} - J_{i-\frac{1}{2}}$ therefore identifies [116]

$$J_p^{\mp} = \sum_{k_j \lesseqgtr 0} \mp 2\alpha_j (3f_{p\pm\frac{1}{2},j}^n - f_{p\pm\frac{3}{2},j}^n),$$

$$J_p^n = J_p^+ + J_p^-, \tag{42}$$

where $p = i \pm \frac{1}{2}$, the form of which insures that $\partial_x J(x) = 0$ under steady-state conditions. Importantly, because $\alpha_j > 0$ for all j, a ($\mp$) is attached to it to account for the behavior of k_j. Such a construction is demanded by how the upwind/downwind SDS scheme for $\mathbb{D}$ is defined in Eq. (30).

IV. WAVE-PACKET TUNNELING

A. Free propagation

To form a baseline comparison, the numerical time evolution scheme of Eq. (33) is first applied to the free wave packet so as to demonstrate stability and accuracy, but more importantly to demonstrate how the simulations are to be represented. The parameters of the simulation are given in Table I but for $V_o = 0$. Significantly, observe that the present formulation is not beholden to requirements imposed by fast Fourier transform (FFT) methods because $V(x, k - k')$ is rendered analytically. This has two consequences: (i) the incentive to make N_x comparable in size to N_k is no longer necessary, and so N_x is increased substantially to allow the boundaries to be pushed further out, to increase L and thereby decrease Δk, and to improve computational speed and accuracy of the SDS approach, and (ii) using FFTs on $V(x, k - k')$ led to Eq. (33) being considered for *every value of i from 1 to N_x* in past studies, but the present method can begin and end at any value of i for which $f(x, k, t)$ is appreciably nonzero, and so in principle much of the numerical evaluations near $x \approx \pm L/2$ need not be undertaken even when a barrier is present, resulting in a further substantial reduction in the numerical investment. The advantages of the second consequence, however, are not employed here.

TABLE I. Default parameters for the simulation of free wave packets and wave packets incident on barriers, grouped according to phase space, wave packet, and barrier parameters.

Term	Symbol	Value	Unit
Number of time increments	N_t	64	–
Time increment	Δt	0.05	fs
Simulation region	L	14	nm
Number of length units	N_x	512	–
Length unit	Δx	$L/(N_x - 1)$	nm
Wave-number units	N_k	128	–
Wave-number increment	Δk	$2\pi/L$	1/nm
Wave-packet parameter	a	0.6	nm
Initial center position	x_o	−1.6	nm
Initial center wave number	k_o	13.55	1/nm
Initial center energy	$\hbar^2 k_o^2/2m$	7	eV
Barrier height	V_o	8	eV
Barrier wave number	k_v	14.491	1/nm
$V(x)$ parameter	λ	0.5	nm

The results of a simulation of a free Gaussian wave packet, cast in a manner that will be used for subsequent simulations, are shown in Fig. 3. Small departures from oval contour behavior are a consequence of the graphical interpolation procedure applied to a relatively coarse grid in k space, which is governed by the size of L. It is nevertheless seen that the simulation is sufficiently accurate. At each time step, Eqs. (39) and (42) are evaluated for $\rho(x,t)$ and $J(x,t)$, respectively. A conventional plot of the former at various times would show a spreading of the wave packet, but here, $\rho(x,t)$ is presented as a contour plot such that the spreading of the wave packet is shown by a widening of the contours for a given value of the t axis, as in Fig. 4. The behavior of $J(x,t)$ is shown in Fig. 5. Both figures show that the dashed white line along which $x(t) = x_o + \hbar k_o/m$ also tracks the maximum of J. Below, these lines shall be called *ballistic* in the sense of free propagation not subject to a barrier. As a result, the motion of the maximum locations can serve to develop a TARD time estimate.
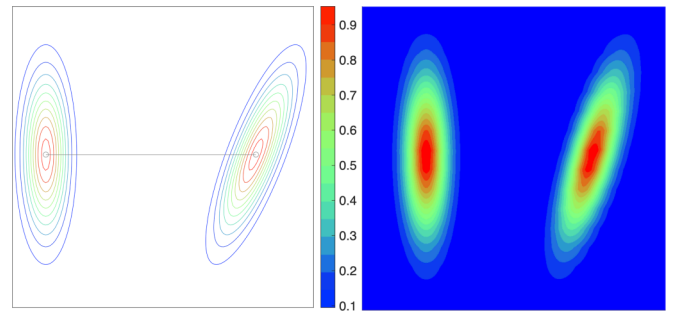

FIG. 3. Simulation of the evolution of WDF for a Gaussian wave packet described by the initial state Eq. (12) for Table I parameters, except for $V_o = 0$. Initial ($t = 0$) is to the left of each figure, final ($t = N_t \Delta t$) is to the right. Left: Analytic contours of Eq. (13). Right: Numerical solution of time-dependent WDF. Range of x: -3.5 nm $< x < 5.5$ nm. Range of k: 10 nm$^{-1} < k < 17$ nm^{-1}.
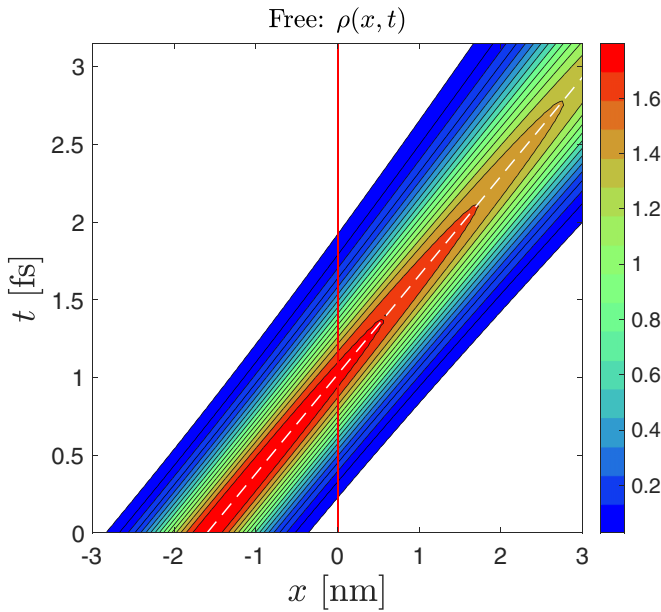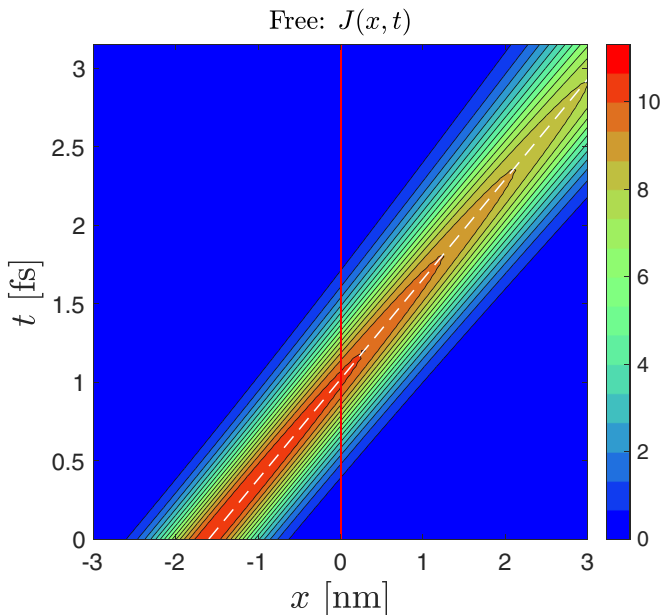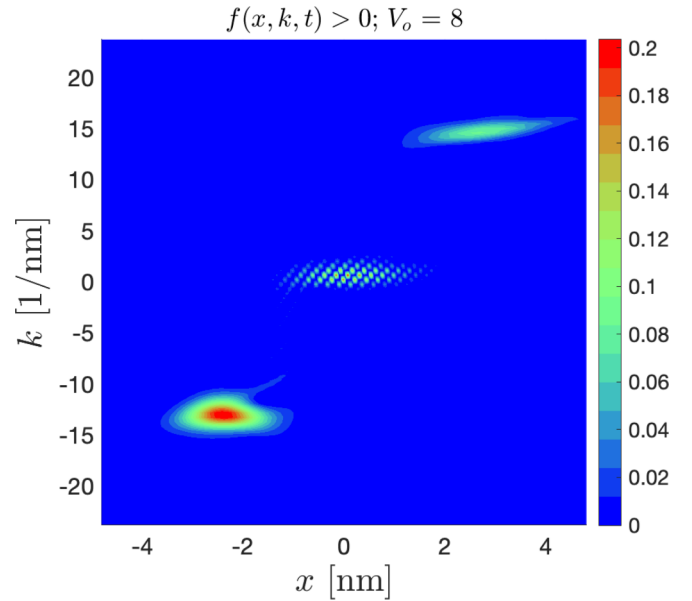
FIG. 4. Time evolution of density $\rho(x,t)$ for a free wave packet using Table I parameters. Vertical red line is at $x = 0$. Dashed white (ballistic) line corresponds to $x(t) = x_o + \hbar k_o t/m$.

B. Gaussian barrier

A Gaussian barrier specified by [see Eq. (A3)] $V_g(x) = V_o \exp[-(x/\lambda)^2]$ with the parameters of Table I is now included: V_o is set to be 1 eV larger than $\hbar^2 k_o^2/2m$ so what emerges past the barrier contains contributions due to tunneling. A small portion of the wave packet does pass through after 3.2 fs, as shown in Fig. 6: Compared to Fig. 3, now the $k < 0$ region of phase space must be included to show the reflected portion. After collision with the barrier, the

FIG. 6. Simulation of the evolution of WDF at $t = 3.15$ fs for a wave packet described by the initial state Eq. (12) for Table I parameters, but now with a Gaussian barrier described by Eq. (A3). The ellipse on the upper right is the transmitted portion of the packet; the deformed ellipse on the lower left is the reflected portion. The center constitutes interference analogous to Fig. 1.

wave packet separates into transmitted and reflected portions, mimicking the separating wave packets of Fig. 1 and the persistent oscillations visible at the origin associated with the central interference term $f_c(x,k)$ (compare Fig. 3(a) of Ref. [84], although it is for a double barrier). At the resolution considered herein, the oscillations appear to be hatching, but when examined in closer detail in Fig. 7, they are seen to replicate the behavior of the cos term in $f_c(x,t)$ in Eq. (19).

FIG. 5. Simulation of the evolution of WDF for a Gaussian wave packet described by the initial state Eq. (12) for Table I parameters. Dashed white (ballistic) line corresponds to $x(t) = x_o + \hbar k_o t/m$.
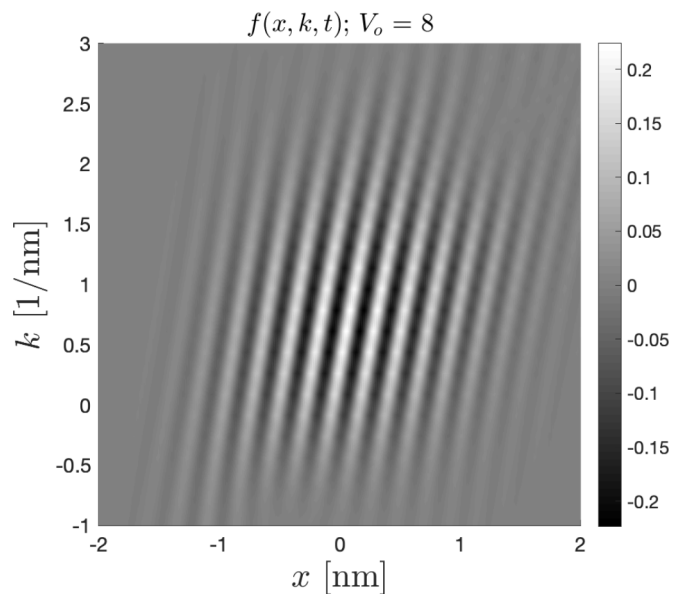

FIG. 7. Close-up of the interference region of Fig. 6, but shown in a gray color scheme to enhance the visibility of the ridges. Both positive and negative regions of $f(x,k,t)$ are shown.
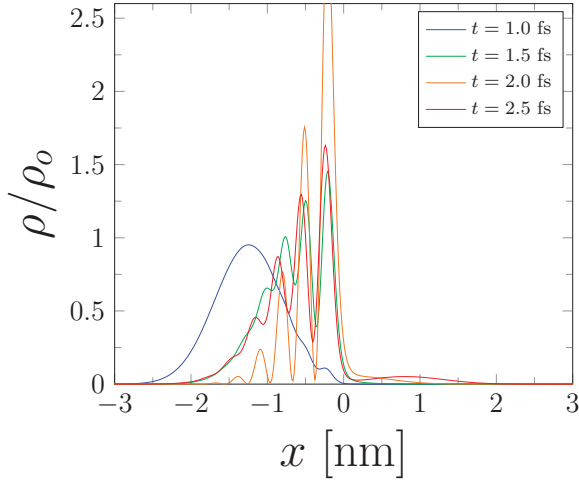
FIG. 8. Density $\rho(x,t)$ normalized to $\rho_o = \rho(x_o, 0)$ at four representative times identified in the legend, using the parameters of Table I.

Increasing $N_x \to 2000$, $N_k \to 240$, and $L \to 36$ nm (parameters that would have defied consideration in the absence of methods developed herein) allows them to be better resolved. The orientation of the oscillations along diagonals in phase space average out when evaluating either $\rho(x,t)$ or $\rho(k,t)$, just as occurred in Sec. II C for the central interference term $f_c(x,k)$. Significantly, the interference is seen to arise when portions of the wave packet begin to separate into reflected and transmitted portions, which occurs at the onset of the barrier. These oscillations, analogs of which appear in exact closed boundary Wigner function studies, complicate the definition of a trajectory and associated tunneling time, but values suggested in Ref. [62] are comparable to the present estimate.

As before, consider the time evolution of $\rho(x,t)$: its behavior at select times is shown in Fig. 8, and compare to solutions based on solving Schrödinger's Eq. [(45), [117]]: The pile-up prior to the barrier is the particle being said by Elberfeld and Kleber to "wait" before tunneling. If the time slices are more densely represented, then a top-down view of these lines can be mapped as contours along axes of x and t. Figures 9–11 show $\rho(x,t)$, $\rho(k,t)$, and $J(x,t)$, respectively, as contours. The minor wiggles along the lowest (blue) contour lines near $(x,t) = (0, 2.7)$ in Fig. 9 are a consequence of the highly oscillatory WDF associated with the central interference region concentrated near $(x,k) = (0,0)$ in Fig. 6: these wiggles are a numerical artifact and disappear as Δx reduces. Reflected and transmitted wave packets are visible in $\rho(x,t)$ and $J(x,t)$ by elements which move away from the red centerline at $(x = 0)$, and $\rho(k,t)$ by the convergence to transmitted (k_t) and reflected (k_r) center momenta. The white dashed (ballistic) lines in Figs. 9 and 11 are evaluated analogously to Fig. 4 but, for the transmitted and reflected lines, the procedure is modified by finding the local maximum of the transmitted and reflected peaks in Fig. 10, called k_t and k_r, respectively, and then defining

$$x_i(t) = x_o + \hbar k_o t/m, \quad x_t(t) = x_t - \hbar k_t(t_{max} - t)/m,$$
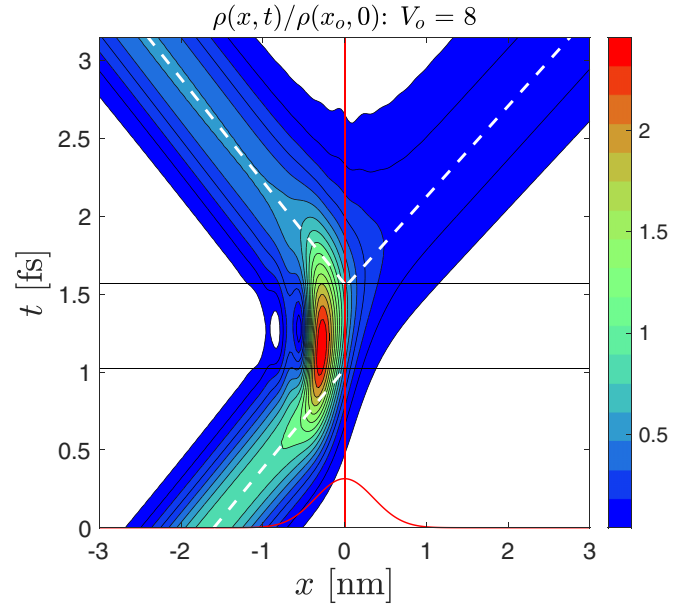$$x_r(t) = x_r - \hbar k_r(t_{max} - t)/m \tag{43}$$

FIG. 9. Time evolution of position density $\rho(x,t)$ as evaluated using Eq. (39). The two horizontal black lines signify where $x_i(t)$ crosses ($x = 0$) (lower at $t_o = 1.0196$ fs) and where the vertex from which $x_t(t)$ and $x_r(t)$ diverge (higher at $t_r = 1.5657$ fs), a separation of 0.5461 fs for Table I parameters.

for the incident, transmitted, and reflected trajectories, respectively, where $t_{max} = (N_t - 1)\Delta t$, (x_t, x_r) are the locations of the transmitted and reflected maxima of $\rho(x, t_{max})$. We find $(k_o, k_r, k_t) = (13.5546, -13.1140, 14.8152)$ are the locations of the incident, reflected, and transmitted maxima of
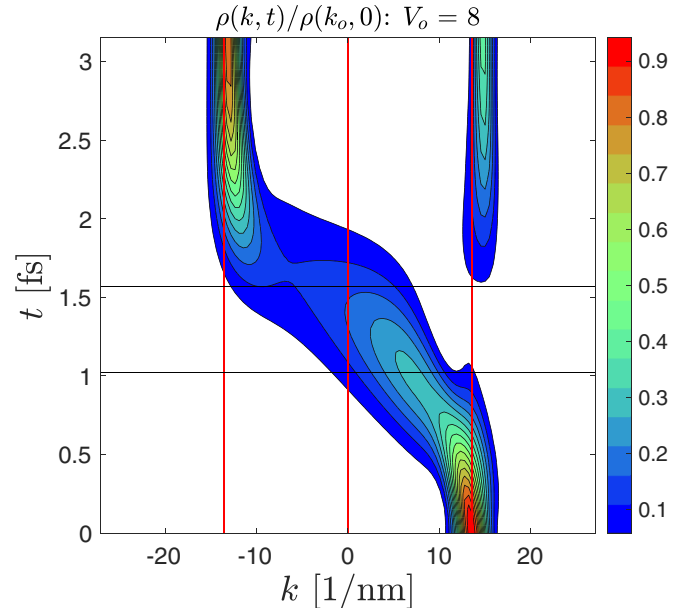

FIG. 10. Time evolution of momentum density $\rho(k,t)$ as evaluated using Eq. (39), but for summations over i. The locations of the peaks of $\rho(k, t_{max})$ define $k_r = -13.1140$ nm^{-1} and $k_t = 14.8152$ nm^{-1}. The vertical red lines are at $k = \pm k_o = \pm 13.5546$ nm^{-1} and at $k = 0$.
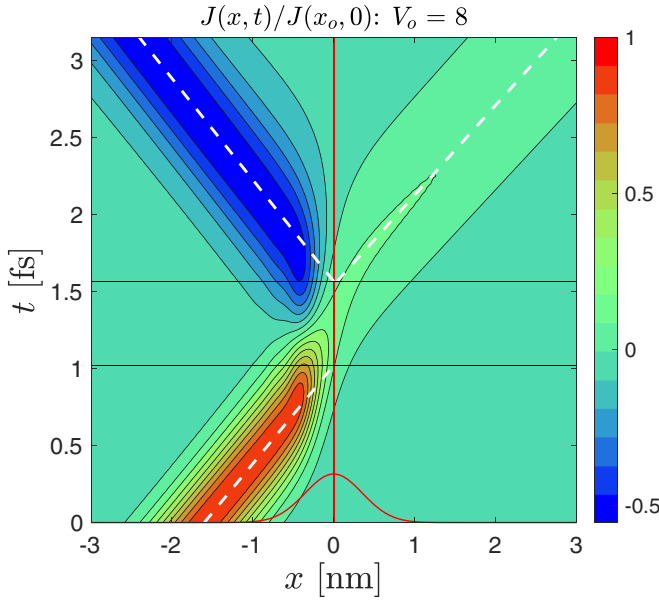
FIG. 11. Time evolution of current density $J(x,t) = J_+(x,t) + J_-(x,t)$ as evaluated using Eqs. (42). White dashed lines are as in Fig. 9. Negative current (blue) moves to the left, positive (red) to the right.

$\rho(k, t_{\max})$, respectively, in units of [nm^{-1}]. It is seen that (k_r, k_t) are slight departures from $(-k_o, k_o)$. Further, $x_t(t)$ and $x_r(t)$ diverge from approximately the same vertex: Letting t_s be such that $x_s(t_s) = 0$ for the subscripts $s \in (i, r, t)$, then $t_r = 1.5657$ fs is at a later time than $x_i(t)$ crosses the origin ($x = 0$) at $t_i = 1.0196$ fs. The behaviors of $f(x, k, t)$ at these times are shown in Figs. 12 and 13. The peaks of $\rho(k, t) \geqslant 0$ as $t \to \pm\infty$ are chosen because they are well-defined and their locations asymptotically fixed. For the purposes of modeling emission, a definition of the time between where $x_i(t)$ crosses
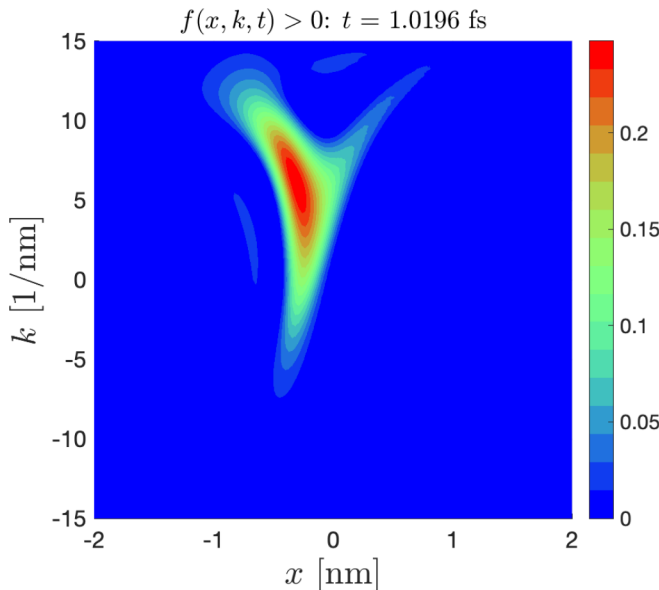

FIG. 12. The distribution $f(x, k, t)$ at $t = t_o$ (positive range only), corresponding to the lower horizontal line of Fig. 9.
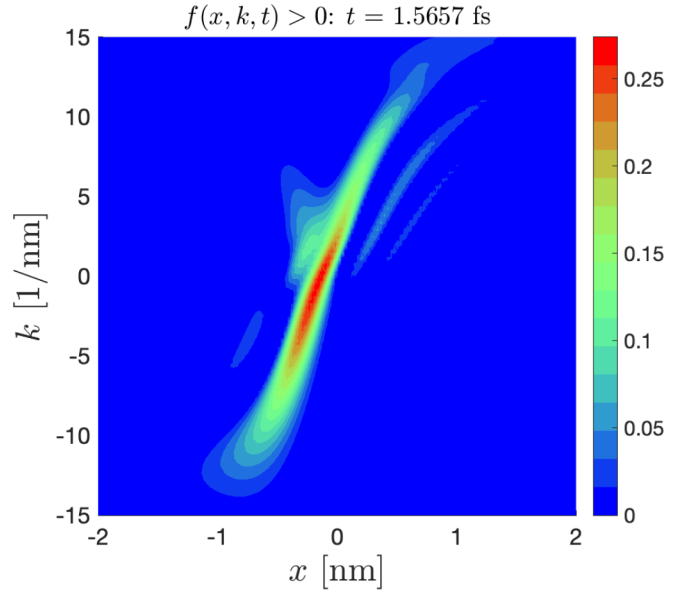
the origin and where $x_r(t)$ and $x_t(t)$ diverges from it, is found by setting the left-hand side of Eq. (43) to 0 and solving for t, giving rise to the TARD times

$$\tau_r = t_r - t_i, \quad \tau_t = t_t - t_i, \tag{44}$$

where t_j is such that $x_j(t_j) = 0$ for $j \in (i, r, t)$.

That $k_t > k_o$ is a consequence of tunneling is numerically evident in the WDF approach, but it can be demonstrated using methods of the Schrödinger approach. Tunneling through a barrier is associated with a tunneling probability $D(k)$. In the simple case of a δ function barrier [Eq. (A1)], for which all incident waves experience tunneling,

$$D(k) = \frac{4k^2}{4k^2 + \gamma^2}, \tag{45}$$

where $\gamma = 4\lambda m V_o / \hbar^2$, as is known [5] or can be shown through a limiting procedure [61]. As a result, the transmitted wave packet resulting from an incident Gaussian wave packet is no longer strictly Gaussian itself, introducing complexity in the identification of the dynamics of its peak with k_t. The current density [compare Eq. (7)]

$$J(x, t) = \frac{q}{2\pi} \int_{-\infty}^{\infty} (\hbar k/m) D(k) f_s(k) \tag{46}$$

in a conventional formulation using a supply function $f_s(k)$ [94] does show that the transmitted wave packet has sped up (corresponding to $k_t > k_o$): If $f_s(k) = (a/\sqrt{\pi}) \exp[-a^2(k - k_o)^2]$ (intended to mimic Eq (11)), then for $\gamma \gg k_o$ [compare Eq. (14)],

$$\langle k \rangle = \frac{\int k D(k) f_s(k) dk}{\int D(k) f_s(k) dk} \approx k_o \left(\frac{2a^2 k_o^2 + 3}{2a^2 k_o^2 + 1} \right) \tag{47}$$

compared to the incident Gaussian ($\langle k \rangle_{\text{inc}} = k_o$ with $D(k) = 1$). For Table I parameters, $\langle k \rangle_{\text{trans}} / \langle k \rangle_{\text{inc}} = 1.015$. Such a result is expected because states with larger k have greater

FIG. 13. The distribution $f(x, k, t)$ at $t = t_r$ (positive range only), corresponding to the upper horizontal line of Fig. 9.

transmission [$D(k)$ is monotonically increasing for the δ-function barrier], an expectation that persists for barriers of finite thickness, and which is clearly evident in SE simulations (see, for example, Fig. 4 of Ref. [54]). This seeming speedup of the transmitted wave packet in a tunneling barrier, first suggested by Hartman [118], is behind the superluminal claims in both tunneling calculations and experiments using electromagnetic analogs of quantum tunneling [54].

Analogous to the Wigner approach studied here, time-dependent simulations of wave packets composed of states for which $\hbar^2 k^2/2m < V_o$ for rectangular barriers [54,92,119] such as Eq. (A2) show that the transmitted/reflected packets produce non-Gaussian but peaked packets [even though the less abrupt potentials of Eqs. (A4) and (A5) would soften features in the transmitted and reflected features]. As shown by Smith and Blaylock [54], FFT methods can then be used to obtain $\rho(k,t)$ from $\rho(x,t)$, from which the momentum of the peaks of the reflected and transmitted wave packets can be found. This would allow the numerical solution of SE to provide (k_i, k_r, k_t) to the TARD time approach of Eq. (43).

The present WDF method differs from the SE approach in important respects: (i) Most importantly, finite difference methods to solve SE do not unambiguously resolve the potential barrier $V(x)$ [albeit a much greater issue for abrupt potentials such as $V_r(x)$ and $V_t(x)$ of Appendix A than for Gaussian potentials such as $V_g(x)$] compared to the exact methods behind Eqs. (A7) and (A8). (ii) Boundary conditions in the SE approach are comparatively difficult because the finite difference methods must apply to the full wave function $\psi(x,t)$ and so its boundaries at each end must be known, whereas the entering boundaries of the WDF are specified and the exiting boundaries open (for example, to minimize systematic error, the simulation region considered by Smith and Blaylock [54] was large enough that the wave packet remained "very nearly zero at the boundaries for the entirety of the simulation," meaning that the range of the simulation was well beyond where the wave packet propagated—by requiring boundary conditions $\psi(\pm L, t) = 0$, they are, in effect, closed). (iii) The relation $f(x, k, t + \delta t) = f(x - \hbar k \delta t/m, k, t)$ [compare Eq. (23)] can be used in WDF simulations when the influence of $V(x, k - k')$ subsides, allowing broad wave packets with a narrow momentum spread to be considered, whereas (for computational reasons) narrow wave packets with wide momentum spread are more advantageous in the SE approach. (iv) Identifying the peaks of the reflected and transmitted packets requires additional computational complexity (e.g., FFT methods) compared to the straightforward WDF approach and do not necessarily identify the velocity of the peaks of $\rho(x,t)$ with the values of (k_i, k_r, k_t) from $\rho(k,t)$; and (v) the trajectory interpretation for the Schrödinger approach (Bohm trajectories) are of a very different kind than the Wigner trajectories that are well-suited to the emission models relying on packets *but more importantly on distributions* [61,62]. Some of these distinctions address constraints placed on the numerical evaluation of $\psi(x,t)$ in the SE approach, e.g., the six constraints on simulation parameters discussed by Goldberg *et al.* [92].

Final considerations with respect to peak identification, as well as how well finite difference methods in the SE approach (with its approximations behind generating the ini-
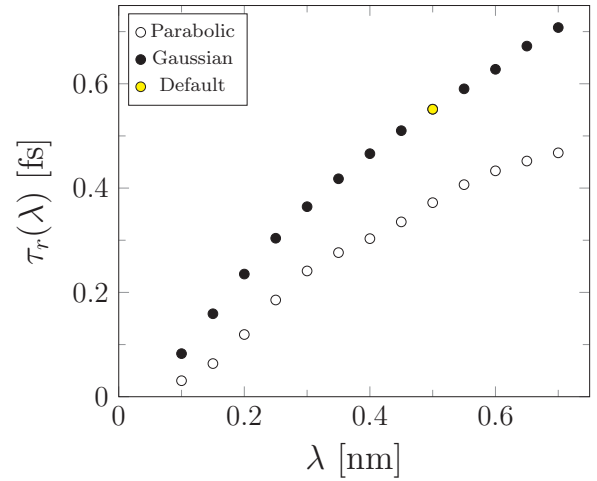

FIG. 14. Evaluation of reflection TARD time $\tau_r(\lambda)$ for Table I parameters but with different λ (Table I value for λ is a yellow dot labeled default for a Gaussian barrier) for both parabolic and Gaussian barriers.

tial wave packet) perform, are in progress and will be reported separately [120]. Therefore, for computational, physical, and utilization reasons, the WDF approach identifies k_t and k_r unambiguously and is therefore better suited to return (k_i, k_r, k_t) in Eq. (43) on which the TARD time model here is based. When the energy and spread of the packet is sufficiently low, the TARD time acquires a tunneling time characteristic that is examined.

C. TARD time

The magnitude of the reflection delay time τ_r of Eqs. (44) will equal the transmission delay time τ_t if the ballistic reflected and transmitted lines converge to a common vertex at $x = 0$. τ_r is the separation of the two horizontal black lines in Figs. 9–11, and numerically equal to $\tau_r = 0.5461$ fs for the parameters of the simulation given in Table I. It includes both transport over the barrier ($\hbar^2 k^2/2m > V_o$, or fly over) and transport under the barrier ($\hbar^2 k^2/2m \leqslant V_o$, or tunneling). Such a definition is reliant on utilizing the probabilistic nature of $\rho(x,t)$ needed when deciding if an electron has passed a barrier or has been reflected by it, and therefore hearkens to the Bohmian trajectory approach [53]. For all parameters held the same except for the barrier now specified by $0.1 \leqslant \lambda \leqslant 0.7$ nm, then τ_r varies as shown in Fig. 14, where the default case is the yellow dot. As the barrier thins (λ decreases), more of the transmitted current is due to tunneling. Estimates of τ_r depend on both the choice of x_o (because of the extent of the initial Gaussian packet) and the abruptness of the barrier governed by $\partial_x V(x)$ (the interference effects extend further from the origin).

We note that the apparent discontinuity (white space) along the transmitted ridge between the black horizontal lines of Fig. 10 is only a consequence of how the 16 contour lines are selected (equispaced in height with the smallest at $1/17 = 0.0588$): A ridge is present but is too small in height to be discerned for the lowest contour level, but is revealed for contours at heights less than 0.035. Because $k_r < -k_o$ and

$k_t > k_o$, it is seen that the transmitted wave packet is traveling slightly faster, and the reflected wave packet slightly slower, than the initial wave packet, a circumstance that must be accommodated by how the delay time is defined. This is not unexpected [119]: The transmitted wave packet is composed of contributions that have a higher transmission probability, which increases with k. Significantly, observe that the horizontal black lines approximately demarcate half the separation where the peak of $\rho(k, t)$ begins to first decline and then separate into two distinct peaks in Fig. 10.

Accounting for how long an electron takes to pass an emission barrier is difficult for theory, simulations, and experiments on ultrafast processes associated with emission and transmission [32,37,43,46,56,121]. A method to incorporate delays associated with electrons passing a barrier (particularly if tunneling), as occur in field and photoemission studies using quantum mechanical, Monte Carlo, or particle-based methods [17,20,70,93,122–125], may profitably use the notion of TARD time introduced here, as it accounts for changes in overall number and current density in a manner that respects their time evolution for a wave packet but which does not run afoul of interference oscillations associated with Wigner trajectories.

A feature that was well-hidden for the Gaussian barrier (and will only be slightly better revealed for the parabolic barrier of Sec. IV D) is the nature of the intersection of the reflected and transmitted dashed white (ballistic) lines. When these lines cross at the origin ($x = 0$) (that is, appear to originate together from a vertex), the TARD times are equal in magnitude, or $|\tau_r| = |\tau_t|$. The barriers simulated herein are symmetrical and reveal finite TARD times, in contrast to arguments using different definitions of tunneling time or suggesting vanishing tunneling times for symmetrical barriers when delay is entirely due to tunneling (considered below). The present distinction will therefore capture probabilistic emission times in a manner usable by particle simulation codes in a way that the other definitions will not.

The contribution of leakage to the transmitted packet is indicated by what portion of the initial packet has $k > k_v \equiv \sqrt{2mV_o}/\hbar$ (see Fig. 3), and successfully passes the barrier. From Eq. (12), it is seen that $f(x, k, 0)$ separates into a product of an x-dependent term with a k-dependent term, the latter of which behaves as $f_s(k) = \exp[-a^2(k - k_o)^2]$. If $D(k)$ is the transmission probability [Eq. (C1)] as a function of k, then the portion of $D(k)f_s(k)$ for $k > k_v$ is a measure of the nontunneling contribution [126]. For $k_o = 13.5546$ (nm^{-1}) ($E_o = 7$ eV) as in Table I, then the portion of $D(k)f_s(k)$ for $k > k_v$, or leakage, is shown as the shaded gray region in Fig. 15(a), and is clearly significant: It is found that the leakage contribution dominates ($>50\%$) for $E_o > 6$ eV. When the barrier is thinned to $\lambda = 0.16$ nm, and k_o is reduced to 10.868 (nm^{-1}) ($E_o = 4.5$ eV), as in Fig. 15(b), the leakage is so small as to not be visible ($\lesssim 1\%$)

The analysis of TARD times when tunneling dominates can then be investigated using the smaller k_o and λ parameters. For resolution of details, the tunneling calculations require much larger values of N_t, N_x: Below, they shall be set equal to $N_x = 2^{10} = 1024$ and N_t comparable. A close-up of the wave packet evolution for $E_o = 4.5$ eV is shown in Fig. 16, with the contours adjusted so as to starkly bring out the tunneling
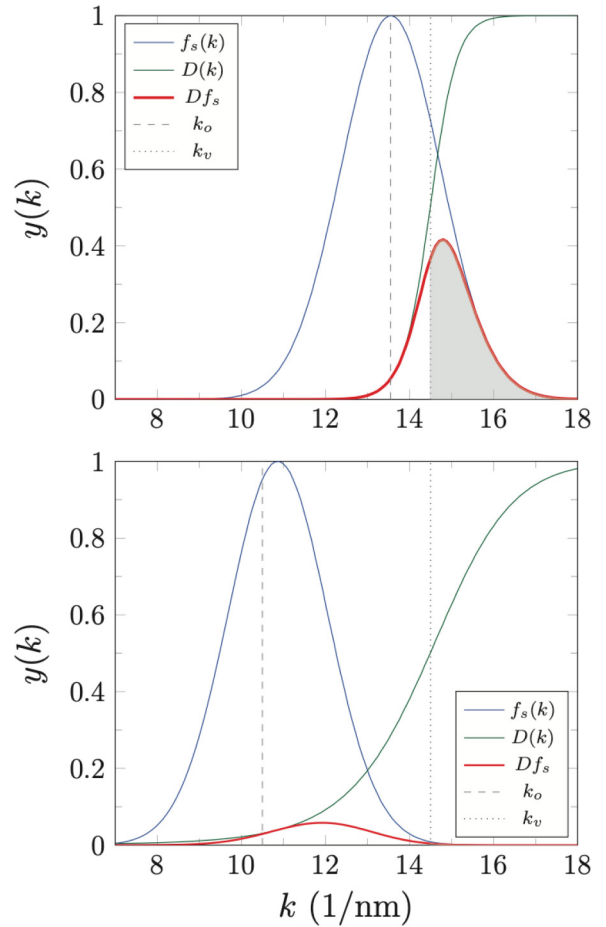

FIG. 15. Transmission probability for a parabolic barrier $D(E)$ (green), Ratio $f_s(E) = f(x_o, k, 0)/f(x_o, k_o, 0)$ (blue), and product $D(E)f_s(E)/D(E_o)f_s(E_o)$ (red) for the cases $V_o = 8$ eV (or $k_v = 14.49$ nm^{-1}) and (top) $\lambda = 0.5$ nm and $E_o = \hbar^2 k_o^2/2m = 7$ eV (or $k_o = 13.55$ nm^{-1}) and (bottom) $\lambda = 0.16$ nm and $E_o = 4.5$ eV (or $k_o = 10.87$ nm^{-1}). The gray shaded region represents leakage, for which $E(k) = \hbar^2 k^2/2m$ is larger than $V_o = \hbar^2 k_v^2/2m$ but is not easily visible for the lower figure.

portion. Because the incident energy is lower, the wave packet is started further away from the center of the Gaussian barrier at $x_o = 2.5$ nm, so a free wave packet would pass the center ($x = 0$) line at approximately 2 fs later. Several features are immediately noticeable. First, the gap between the horizontal lines representing where $[x_i(t), x_r(t), x_t(t)]$ cross the origin $x = 0$ have become much closer. Second, $x_r(t)$ and $x_t(t)$ no longer appear to originate from a single vertex: the transmitted ballistic line initiates slightly to the right of the origin. To explore the smaller TARD times associated with tunneling, evaluations were done for E_o spanning a range from tunneling dominated ($E_o < 6$ eV) to leakage dominated ($E_o \gtrsim 6$ eV), and shown in Fig. 17. It is seen that τ_r decreases as the tunneling contribution shown in Fig. 15 increases.

Observe that lower energy wave packets are reflected sooner than higher energy packets precisely because their penetration of the barrier is less, leading to a *reflection delay* that occurs because particles with higher energy penetrate a greater distance into the barrier before being turned back com-
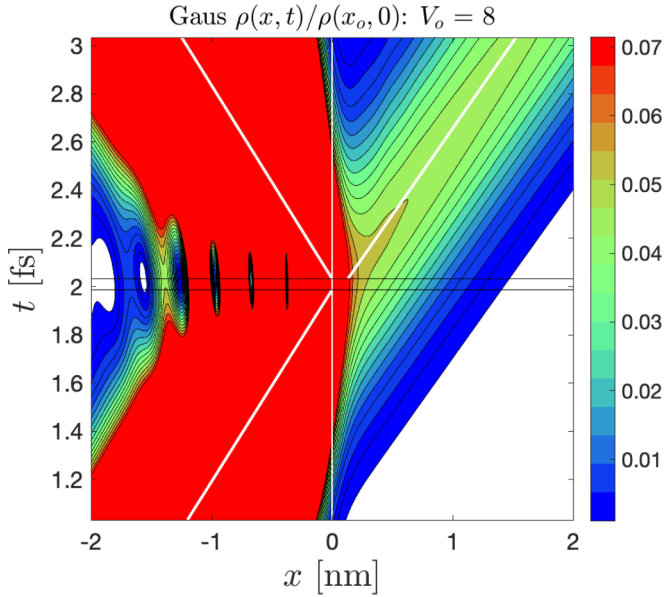
FIG. 16. Time evolution of position density $\rho(x,t)$ as evaluated using Eq. (39) but now for $k_o = 10.8679\,\mathrm{nm}^{-1}$ ($E_o = 4.5\,\mathrm{eV}$), $x_o = -2.5\,\mathrm{nm}$, and $\lambda = 0.16\,\mathrm{nm}$ for a Gaussian barrier. The contour lines are 2.5% the levels shown in Fig. 9 and a smaller field of view is chosen. Dashed white lines replaced by solid white lines for visibility. The two horizontal black lines are now separated by 0.046 fs. Thin white vertical line marks the center of $V(x)$.

pared to lower energy particles, and such a delay presumably contributes to the separation between where the incident and reflected ballistic lines pass the origin ($x = 0$) in Figs. 9 and 10. By comparison, a classical particle would reflect instantaneously, such that there would be no separation between the ballistic incident and reflected lines, both crossing where the particle strikes the barrier. Because lower energy incident
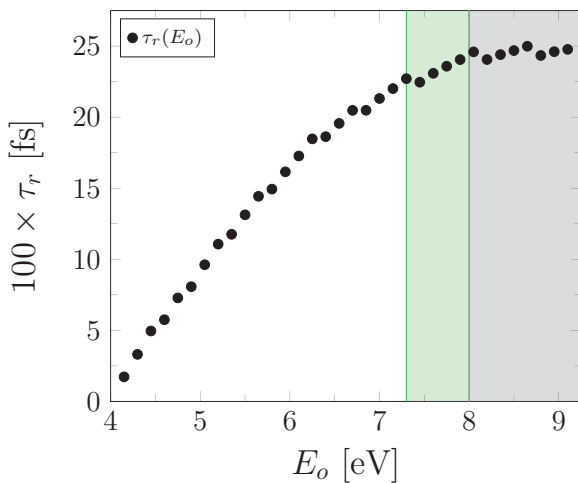

FIG. 17. TARD time τ_r as the initial center energy of the wave packet $E_o = \hbar^2 k_o^2/2m$ changes for a Gaussian barrier with $V_o = 8\,\mathrm{eV}$ and $\lambda = 0.16\,\mathrm{nm}$. All simulations used $N_t = 512$ and $N_x = 1024$. Numerical noise is a consequence of using fitting procedures to find (k_o, k_r, k_t) and $[x_r(t_{max}), x_t(t_{max})]$. $E_o > V_o$ in the gray region, and leakage dominates in the green region ($7.3\,\mathrm{eV} \leqslant E_o \leqslant V_o$).

wave packets are more completely reflected, the expectation then becomes that the ballistic incident and reflected lines form a vertex, with a separation now appearing between $x_r(t)$ and $x_t(t)$ (they no longer form a vertex at $x = 0$).

The $x_o(t)$ line, if extended, would be above the $x_t(t)$ line in Fig. 16. That circumstance mimics Fig. 4 of Ref. [54] and, just as they discuss, it is not an indication of "superluminal" transport. Whereas $\tau_r = t_r - t_i \to 0$ as E_o decreases, the separation $\tau_t \equiv t_t - t_i$, where $x_i(t_i) = 0$ and $x_t(t_t) = 0$, remains finite and is the TARD time due to tunneling, to borrow the parlance of Büttiker and Landauer [43]. From Eq. (43) and the simulation leading to Fig. 16, for $x_o = -2.5\,\mathrm{nm}$, $x_t = 2.8196\,\mathrm{nm}$, $k_o = 10.8679\,\mathrm{nm}^{-1}$, $k_t = 11.578\,\mathrm{nm}^{-1}$, $t_{max} = (N_t - 1)(4\,\mathrm{fs})/N_t = 3.9961\,\mathrm{fs}$ for $N_t = N_x = 2^{10}$, then

$$\tau_t \equiv t_t - t_i = -\frac{mx_t}{\hbar k_t} + \frac{mx_o}{\hbar k_o} + t_{max} = -0.094567\,\mathrm{fs}. \quad (48)$$

Finding the behavior of τ_t in the tunneling-only limit requires greater accuracy because of interference of the reflected components of the wave packet associated with more abrupt barriers, and therefore significantly larger N_t and N_x for the simulation. Comparisons to simulations based on solving SE for when tunneling is exponentially small shall be reported separately [120]. Observe that the negative value associated with Eq. (48) is an artifact of the arbitrary identification of τ_t with trajectories that cross the t axis (vertical): Were the crossing along the x axis evaluated instead, and a time difference created by the ratio of that separation with the incident velocity, then a positive time metric would result that may result in a preferred time parameter. The interpretation will be taken up separately. Here, it is important to emphasize that the sign should not be associated with superluminal or noncausal behavior; rather, the emitted wave packet is constructed of higher velocity components analogous to the discussion surrounding Eq. (47) or as in Fig. 4 of Ref. [54].

The timescales associated with Fig. 17 in the evolution of $\rho(x,t)$ determined herein for Gaussian wave packets from the Wigner function approach (compare also Ref. [127]) closely match those determined by Petersen and Pollak [128,129] for Gaussian wave packets similar to Eq. (11) (e.g., compare Fig. 9 here with Fig. 2(a) of Ref. [129]), but $\tau(\lambda) > 0$ here. In their study, Petersen and Pollak argue that the tunneling time for both symmetric barriers (such as the Gaussian barrier here) and asymmetric barriers vanishes. They use a definition of τ_a reliant on the size of $1/a^2$ in Eq. (11), which is Γ in their notation. It is based on a definition of τ which compares the time it takes the Gaussian wave function centered at $t = 0$ at a position $-x_i$ to the left of the barrier to reach positions $\pm L$, $L \gg x_i$. Extrapolating to large a, they find that it takes longer to reach $-L$ than $+L$, with a time difference τ_a which depends on the width of the wave packet. They explain this by noting that the transmitted wave contains more of the higher components of the momentum in the original wave packet (similar to the arguments of Ref. [54]). They find that $\tau_a \to 0$ as $a \to \infty$. The time τ_a is clearly not the time the transmitted particle spends inside the barrier.

In contrast, here, the TARD times (τ_r, τ_t) are explicitly related to the center momenta of the incident, reflected, and transmitted packets. Such a formulation is closer to the view discussed by Landsman and Keller [52] when they say "…*the*
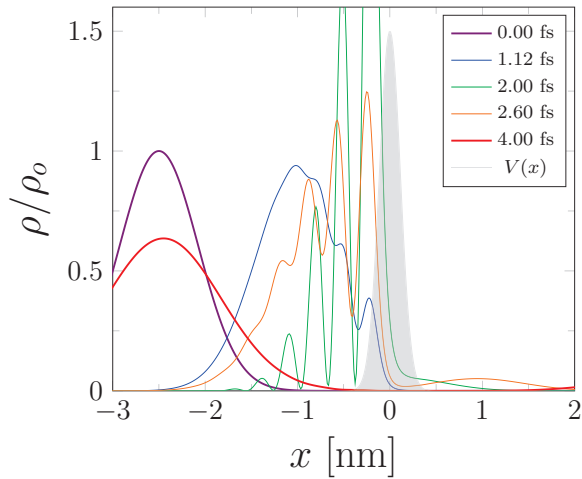
FIG. 18. Density $\rho(x,t)$ normalized to $\rho_o = \rho(x_o, 0)$ at four representative times identified in the legend, using the parameters of Table I, but now for $k_o = 10.8679$ ($E_o = 4.5\,\mathrm{eV}$) and $\lambda = 0.16\,\mathrm{nm}$ so the transmitted portion is predominantly due to tunneling. Shaded region is the barrier.

time it takes an electron to tunnel is a probabilistic, *rather than a deterministic process*" (emphasis added). Because $\rho(x,t)$ represents the probability of a measurement finding an electron at x, delays associated with its behavior therefore account for time-dependent behavior associated with tunneling. They will therefore be useful in simulations that rely on particlelike behavior (e.g., beam optics codes, molecular dynamics, or Monte Carlo simulations [17,27,28]) for the description of nanogap and/or ultrashort conditions. That is, in field emission simulations, electrons are localized as they emerge from the barrier, and are most often treated as point particles; therefore, a particle code is better served by the ballistic trajectories and the TARD times they give rise to (although a phase space trajectory representation is desirable [62]), particularly if the applied field rapidly oscillates [35,37,130] or internal processes that affect emission result in a delay in the arrival time of the electron to the surface [93].

D. Parabolic barrier

The TARD times change as a consequence of barrier shape. A rapid demonstration is afforded by considering the parabolic barrier $V_p(x) = V_o[1 - (x/\lambda)^2]$ (Eq. (A4)) but otherwise using the parameters of Table I. For brevity, only the evaluations of $\rho(x,t)$ and $J(x,t)$ are reproduced, and shown in Fig. 19. Compared to the delay time associated with Fig. 10, the time separation between the horizontal lines in Fig. 19 is now 0.3693 fs and is a consequence of the narrower extent of the parabolic barrier compared to the Gaussian barrier. Conversely, where the figures are similar is a consequence of the parabolic barrier parameters being chosen to closely match the Gaussian barrier at its apex. Although the parabolic barrier has a more abrupt onset than a Gaussian barrier and therefore causes $V_p(x, k - k')$ to span further in k space, the sensitivity to the location of x_o for the evaluation of τ_r is reduced. It was found, however, that by choosing parameters for which tunneling dominates, as in Figs. 16 and 18 for a parabolic
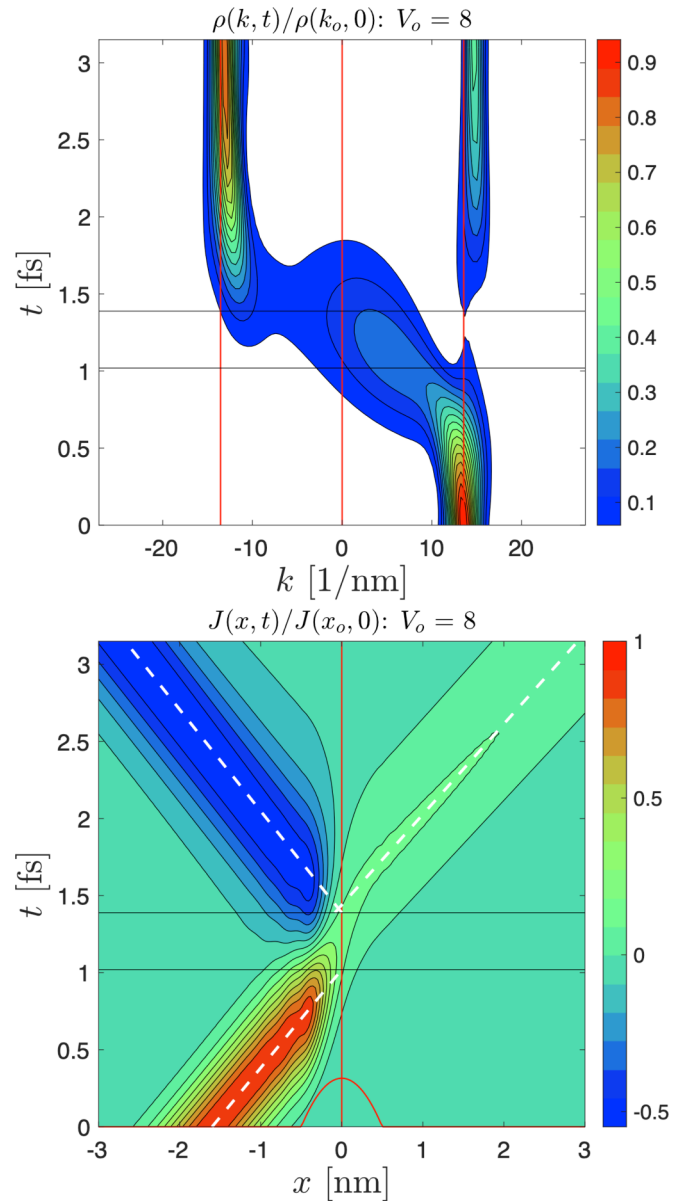

FIG. 19. The simulation for (top) density $\rho(k,t)$ and (bottom) current density $J(x,t)$ for Table I parameters (compare Figs. 10 and 11, respectively), for the parabolic barrier (represented as the red curve in the lower figure). The delay time has been shortened (closer spacing of the horizontal lines), which is now 0.3693 fs.

barrier, Fig. 20 results. Lastly, when $E_o \gtrsim 7.25\,\mathrm{eV}$ in Fig. 17, both τ_r and τ_t are nonzero, so none of the ballistic (white) lines diverge from a common vertex along $x = 0$ for those conditions. Both τ_r and τ_t are required to model where the reflected and transmitted wave packets are at later times. In simulations of electron emission, the probability of whether an electron is launched, when it is launched, and with what velocity it departs are therefore derived from $\rho(x,t)$, τ_t and k_t. The effects of abruptness on the transmitted and incident packets, and the development and application of that emission model, shall be examined separately [120].
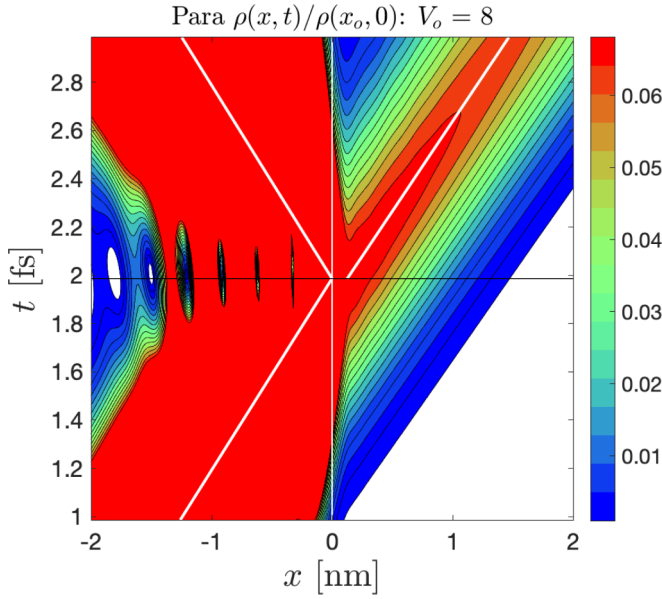
FIG. 20. Time evolution of position density $\rho(x,t)$ as evaluated using Eq. (39) but now for $k_o = 10.8679\,\text{nm}^{-1}$ ($E_o = 4.5\,\text{eV}$), $x_o = -2.5\,\text{nm}$, and $\lambda = 0.16\,\text{nm}$ for a parabolic barrier (compare Fig. 16). The two horizontal black lines are now overlapping. The thin white vertical line marks the center of the Gaussian $V(x)$.

E. Comparisons to other times

The Büttiker-Landauer semiclassical time τ_{sc} [43] is commonly invoked to describe tunneling (See Refs. [52] and [131] for its relation to Bohmian and Larmour times). τ_{sc} is defined by [43]

$$\tau_{sc} = \int_{x_-}^{x_+} \frac{dx}{v(x)}. \tag{49}$$

The terms $x_\pm$ are defined by $V(x_\pm) - E = 0$, $V(x)$ is the barrier profile, E is the energy, and $v(x) = \sqrt{2|V(x) - E|/m}$ for a given E (the absolute value is used by Büttiker and Landauer to make τ_{sc} real-valued for E_o both below and above V_o: here, only $E_o < V_o$ is considered). As is evident from its definition, τ_{sc} is defined for a single energy (for which, as Büttiker and Landauer confirm, the incident current density is simply $j = q\hbar k/m$ as follows from $\psi_{\text{inc}}(x) \propto e^{ikx}$), which already renders it different from the TARD times for wave packets composed of many wave-number components. Using shape factor methods of Appendix C [132] for evaluating the Gamow factor, where $L(E) = x_+(E) - x_-(E)$ and $\kappa(E) = \sqrt{2m(V_o - E)}/\hbar$, and V_o is the maximum height of the barrier, the semiclassical time becomes

$$\tau_{sc}(E) = u(E)\frac{mL(E)}{\hbar\kappa(E)}, \tag{50}$$

where $\kappa(E) = \sqrt{k_v^2 - k^2}$, $E = \hbar^2 k^2/2m$, and $V_o = \hbar^2 k_v^2/2m$. The rectangular, triangular, and parabolic barriers are special cases of a class of potentials that have constant shape factor terms [6,132], e.g., for the rectangular barrier of Eq. (A2), $\sigma(E) = u(E) = 1$, whereas for the parabolic barrier of Eq. (A4), $2\sigma(E) = u(E) = \pi/2$ [compare Eq. (C6)]. More complex potentials, such as Gaussian and image charge barriers, have energy-dependent shape factors for which the energy

dependence can be taken as weak [70]. Restricting attention for now to the parabolic barrier $V_p(x)$ of Eq. (A4), it is found

$$x_\pm[E(k)] = \pm\frac{\lambda}{k_v}\sqrt{k_v^2 - k^2},$$

$$L[E(k)] = \frac{2\lambda}{k_v}\sqrt{k_v^2 - k^2}. \tag{51}$$

Therefore, for the parabolic barrier $V_p(x)$, the semiclassical time is

$$\tau_{sc}(\text{parabolic}) = \frac{\pi m\lambda}{\hbar k_v} = \pi\lambda\sqrt{\frac{m}{2V_o}} \tag{52}$$

and is therefore independent of $E(k)$: In contrast, barriers for which $u(E)$ is *not* constant (such as Gaussian barriers) exhibit an energy dependent $\tau_{sc}(E)$. For the parameters of Fig. 1, where $V_o = 8\,\text{eV}$ and $\lambda = 0.16\,\text{nm}$, $\tau_{sc}(\text{parabolic}) \equiv \tau_p = 0.29964\,\text{fs}$. The ratio of τ_p with τ_t is therefore $\tau_t/\tau_p = 0.31560$, or the TARD time due to tunneling is about a third of the semiclassical time.

The dwell time (Appendix D) is the difference between the time a wave packet spends in the region of a barrier with the time it would have spent in the same region when the barrier is absent. In the treatment by Winful [63,64], the wave function in the barrier is evanescent, not propagating, and therefore exponentially declines with a decay constant of $\kappa(E)$ that behaves as a skin depth. By analogy, then, energy in the form of $|\psi(x)|^2$ is stored in the barrier governed by that skin depth, and so a substantial increase in L will not appreciably alter how much is stored. In his analogy, transmitted electrons correspond to leaking of the stored energy in the forward direction, with reflection being the leaking in the back direction. The time associated with that leaking is τ_d.

Although there is not an equivalence between the TARD times (τ_r, τ_t) and the dwell time (τ_d) of Winful [63], it is seen that they bear some relation (e.g., $\tau_r \to 0$ when the ballistic incident trajectory overlaps the transmitted trajectory, which is equivalent to Fig. 4, as does τ_d by definition). The rectangular barrier of Winful is such that $L = 2\lambda$ here, for which the Gamow factor is $\theta = 2\kappa L = 4\kappa\lambda$. Consequently, Eq. (20) of Winful for τ_d is generalized to Eq. (D1) here, where $\tau_v \equiv \hbar/V_o = 0.0823\,\text{fs}$ for $V_o = 8\,\text{eV}$. For sufficiently large θ, then $\text{sech}(\theta/2) \ll 1$, and so Eq. (D1) with $N_b \to 0$ becomes

$$\tau_d \approx \tau_v\left(\frac{k}{\kappa}\right)\tanh\left(\frac{\theta}{2}\right) \tag{53}$$

and is shown in Fig. 21. In this limit, $\tau_i(\theta)$ is simply related to $\tau_d(\theta)$ by $k^2\tau_i \approx \kappa^2\tau_d$ as per Eq. (D2). The equation is extended to parabolic barriers here through the replacement $\theta \to 4\sigma\kappa\lambda$, where σ is the shape factor of the parabolic barrier (Appendix C). Now, where the barriers are purposely thin to encourage a visible transmitted wave packet, the largeness of θ is governed by E_o rather than λ (as in Appendix D).

The behavior of τ_r from Figs. 14 and 17 are now compared to the dwell time τ_d for barrier widths specified by the parameter λ. A direct comparison as a function of λ is shown in Fig. 22 for Gaussian and parabolic $V(x)$, where the asymptotic lines use θ_p only [Eq. (C7)] for simplicity, because the difference to lines using θ_g is small, as demonstrated in
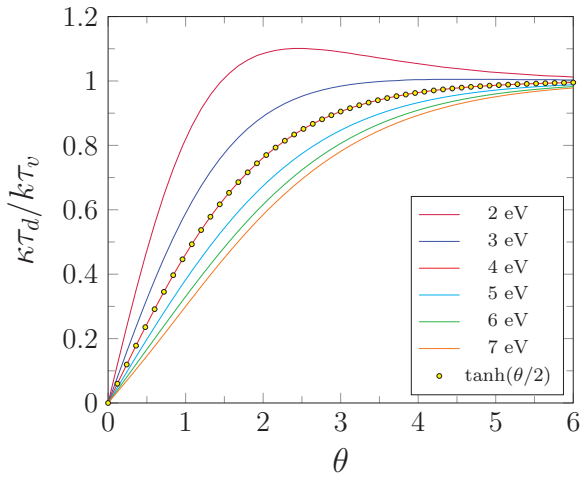
FIG. 21. The dwell time $\tau_d(\theta)$ [Eq. (D1)] in the ratio $\kappa\tau_d/k\tau_v$, for the parameters $V_o = 8\,\mathrm{eV}$ and E as shown (lines). Circles ($\circ$) correspond to Eq. (53) and overlap the $E/V_o = 1/2$ line for which $k = \kappa$.

FIG. 23. The TARD time τ_r of Fig. 17 compared to the dwell time $\tau_d(\theta)$ [Eq. (D1)], both in the ratio $\kappa\tau/k\tau_v$, as a function of the Gamow factor $\theta_g(k)$ [Eq. (C10)] for $V_o = 8\,\mathrm{eV}$ and $\lambda = 0.16\,\mathrm{nm}$. Leakage dominates in the green range.

the discussion following Eq. (C10). For Fig. 23 [Gaussian $V(x)$ only], the Gaussian Gamow factor θ_g is approximated by Eq. (C10) over the range of interest for $V_o = 8\,\mathrm{eV}$ and $\lambda = 0.16\,\mathrm{nm}$. The comparison is only to show shared trends, as the reflection TARD time τ_r differs from the dwell time τ_d. The green region of Fig. 17 corresponds to the green region of Fig. 23 for which $\theta_g(E) < 0.6373$, and that region contains a significant contribution to emission due to leakage. It is in keeping with the observation by Winful that "...*propagating above-barrier components begin to dominate....*"

V. SUMMARY

Approaches to introducing quantum effects into simulations of electron emission that otherwise treat electrons as
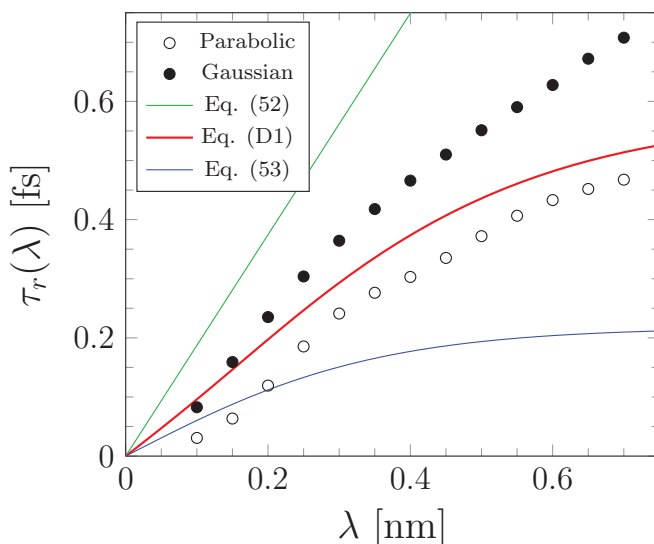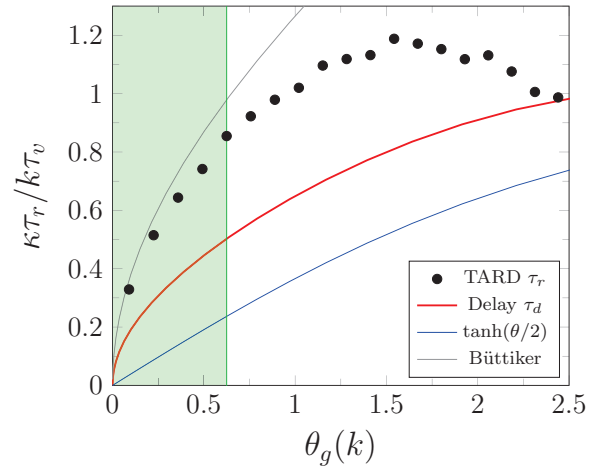

FIG. 22. The reflection TARD time $\tau_r(\lambda)$ of Fig. 14 for parabolic and Gaussian barriers. Lines use only $\theta = \theta_p$ of Eq. (C7) as σ_g is within 2.4% of $\sigma_p = \pi/4$.

pointlike particles have sought to use trajectory concepts made possible by the WDF, which acts like a phase space probability function even though it exhibits negative values as a consequence of quantum mechanical effects. For wave packets that separate into transmitted and reflected components, the negative regions are associated with oscillatory regions and correspond to effects of interference. Past studies of WDFs for closed boundary conditions show similar rapid oscillations near barriers, the magnitudes of which depend on the abruptness and height of the barrier. Such oscillations are problematic for a trajectory interpretation but are an unavoidable artifact of quantum effects, interference, and wave reflection for static conditions, with their behavior for dynamic conditions being presently unaddressed. Time-dependent Wigner wave-packet studies were therefore investigated herein to develop a tunneling time model that would allow such processes to be modeled even in dynamic circumstances.

In the present study, an accurate and fast means of numerically evaluating the time evolution of a WDF was developed that, in addition to being substantially more rapid than prior versions, allows using finite difference methods with a far greater number of points characterizing f_{ij}^n. The large number of phase space points allowed for the determination of $f(x, k, t)$ using second-order accurate schemes in both position x and time t (the discretization of wave number k is fixed by the length of the simulation region). The high accuracy is required to numerically resolve the central interference region (so named to correspond to the central entanglement region in the work of Weinbub and Ferry [107]) after the wave packet interacts with the barrier and separates into a transmitted and reflected portion. The central interference region is shown to be similar to the central interference term that arises analytically from two Gaussian wave packets traveling in opposite directions. Although the focus was on Gaussian and parabolic barriers, analytic representations for other barriers (delta function, rectangular, triangular, double barrier for resonant tunneling) were developed, for which the numerical

techniques developed herein are applicable, although only the parabolic and Gaussian barriers were considered explicitly because they exhibit a desirable smoothness.

Simple analytic barriers allowed for similarly simple $V_c(x, k - k')$ functions in the WDF time-evolution equation [Eq. (4)] to be considered, a development that allowed for a substantial increase in execution speed in addition to the hybrid predictor-corrector and implicit methods used. Nevertheless, as also shown, more complex potentials (e.g., biased RTDs, Fowler-Nordheim, or triangularlike asymmetric barriers) can be built up through the addition of the simple components, thereby retaining the advantage of the rapid analytical methods without having to resort to fast Fourier transform (FFT) techniques characterizing earlier studies. Although not pursued herein, the structure of the time evolution approach in Eqs. (38) may allow for only a subsection of the simulation region to be considered, thereby enabling further significant reductions in execution time.

Based on the behavior of the transmitted and reflected wave packets and and their dependence on barrier height, width, symmetry, and abruptness, TARD times were introduced that relied only on knowing the asymptotic locations of peaks (x_o, x_r, x_t) of the position density $\rho(x, t)$ and (k_o, k_r, k_t) of the momentum density $\rho(k, t)$, which serve to define the times (t_o, t_r) by setting the left-hand side of Eq. (43) to 0 and then defining the tunneling time via $\tau_r = t_r - t_o$ and $\tau_t = t_t - t_o$ although we observe that in the case when tunneling dominates, the definition of τ_t may merit revisiting to address its sign. When developed for force-free motion away from the barrier, it is suggested that the method can be applied to when the wave packets are being accelerated in a linear field because for up to constant fields, the WDF satisfies Eq. (23), although such an assertion remains to be demonstrated.

The TARD times provide a natural approach to introduce delays associated with quantum mechanical effects for either static or dynamically changing barriers. As a result, a means to include these effects in simulations of nanogaps and/or changes in the barrier occurring over fs (or smaller) timescales is possible without relying on an instantaneous application of a current density equation (e.g., the Fowler-Nordheim equation) that presupposes static barriers under the assumption that the tunneling is instantaneous. In trajectory simulations where spurious oscillations at the barrier become problematic, we may explore the entangled classical trajectory approach of Donoso and Martens to obtain the average behavior of the ensemble, including approximate tunneling probability and tunneling time.

Future work will be in several directions. First, complications due to the influence of resonances [see Eq. (B1)] will be considered. Second, comparisons to and modifications by Schrödinger-based approaches will enable the examination of deep tunneling contributions where the transmitted packet is exponentially small. Third, demonstrating the utility of the tunneling time definition in emission simulations that are generally reliant on a particle model (due to the large number of emission events characteristic of operation) will be considered. Lastly, the present analysis will be extended to build on the long-established analogy between quantum mechanics and electromagnetic wave propagation [64] to devise a series of microwave experiments for the purpose of investigating

extensions of the present study. The cm wavelength scale in the microwave regime enables effects presented in this paper to be examined at the macroscopic scale.

APPENDIX A: DELTA-FUNCTION SEQUENCES

The $V(x)$ barriers (and by extension, wells) give analytical $V(x, k - k')$ from Eq. (4), for which ones of relevance here are δ function (V_d), rectangular (V_r), parabolic (V_p), Gaussian (V_g), and triangular (V_t), defined by

$$V_d(x) = 2V_o\lambda\delta(x), \tag{A1}$$

$$V_r(x) = V_o \Theta(\lambda^2 - x^2), \tag{A2}$$

$$V_g(x) = V_o \exp[-(x/\lambda)^2], \tag{A3}$$

$$V_p(x) = V_o[1 - (x/\lambda)^2]\Theta(\lambda^2 - x^2), \tag{A4}$$

$$V_t(x) = V_o[1 - (x/\lambda)]\Theta(x)\Theta(\lambda - x), \tag{A5}$$

where $\Theta(x)$ is the Heaviside step function and $\delta(x)$ is the Dirac δ function (compare Eq. (10) of Ref. [61]). In the case of Eq. (A1), it is quickly found that

$$V_d(x, k) = -\frac{4\lambda V_o}{\pi \hbar} \sin(2kx). \tag{A6}$$

The other cases are chosen to emphasize distinctions associated with continuous behavior such that $V(x + \epsilon) = V(x - \epsilon)$ as $\epsilon \to 0$, abrupt behavior such that $\partial_x V(x + \epsilon) \neq \partial_x V(x - \epsilon)$, and symmetric behavior such that $V(+|x|) = V(-|x|)$. Specifically, (i) cases (r, t) are discontinuous whereas cases (p, g) are continuous, (ii) cases (r, p, t) are abrupt whereas case (g) is smooth, and (iii) cases (r, p, g) are symmetric whereas case (t) is asymmetric. It can be shown that the symmetric cases result in $V_c(x, k)$ given by [for $c \in (d, r, p, g)$]

$$\frac{V_s(x, k)}{V_o N_c} = -\eta_c(2k\lambda) \sin(2kx), \tag{A7}$$

where $k - k' \to k$ for notational simplicity, whereas the antisymmetric case contains an additional term such that

$$\frac{V_t(x, k)}{V_o N_t} = \eta_t(2k\lambda) \sin(2kx) - \eta_t'(2k\lambda) \cos(2kx), \tag{A8}$$

where the (t) subscript is used directly as it is the only asymmetric case considered, although the right triangle barrier can be made symmetrical by the replacement
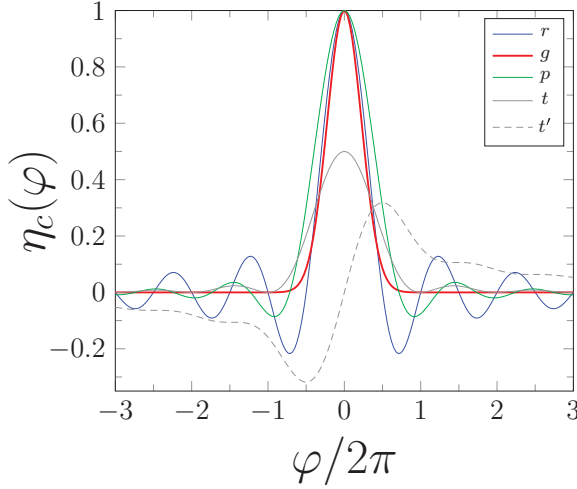
FIG. 24. The $\eta_c(\varphi)$ functions of Eqs. (A7) and (A8). Discontinuity and abruptness result in a greater range over which variation is visible. Legend refers to the cases of Eq. (A10).

FIG. 25. $V(x, k)$ for a single barrier ($V_r(x)$ of Eq. (A2)) compared to a double barrier separated by a well of width $w = W - 2\lambda$ ($V_{rtd}(x)$ of Eq. (B1)) with $2\lambda = 3W/8$.

$V_t(x) \to [V_t(x) + V_t(-x)]/2$, or isosceles barrier. The right triangle barrier is, however, of greater importance given its relation to the Fowler-Nordheim barrier of field emission [67,69]. Explicit evaluation of Eq. (4) gives

$$N_r = \frac{4\lambda}{\pi\hbar}; \quad N_p = \frac{8\lambda}{3\pi\hbar}; \quad N_g = \frac{2\lambda}{\sqrt{\pi}\hbar}; \quad N_t = \frac{2\lambda}{\pi\hbar}, \quad \text{(A9)}$$

$$\eta_r(\varphi) = \frac{\sin\varphi}{\varphi}; \quad \eta_g = e^{-\varphi^2/4},$$

$$\eta_p(\varphi) = \frac{3}{\varphi^2}(\sin\varphi - \varphi\cos\varphi),$$

$$\eta_t(\varphi) = \frac{1 - \cos\varphi}{\varphi^2}; \quad \eta_t'(\varphi) = \frac{\varphi - \sin\varphi}{\varphi^2}. \quad \text{(A10)}$$

The behavior is shown in Fig. 24. Observe that for the symmetric cases, the relation

$$\frac{N_c\hbar}{4\lambda}\int_{-\infty}^{\infty}\eta_c(\varphi)d\varphi = 1 \quad \text{(A11)}$$

holds, so the $N_c\hbar\eta_c(2k\lambda)/(4\lambda)$ are seen to be δ-function sequences (analytic representations of the Dirac δ function) [133] in the limit that $\lambda \to 0$. The rapidity with which $V_c(x, k)$ decreases with increasing k greatly affects oscillations in $f(x, k, t)$ for large $\pm|k|$, which also depends on N_c, e.g., $V_r(x, k)/V_o N_r$ diminishes with increasing $\varphi = 2k\lambda$ but $V_d(x, k)/V_o N_r$ does not, showing starkly the effects of the most abrupt barrier. Therefore, the Gaussian barrier is the most advantageous as it most rapidly diminishes with φ, followed by the parabolic barrier, whereas the triangular and rectangular barriers are more susceptible to the incursion of numerical noise for large $|k|$ in time evolution simulations using present methods.

APPENDIX B: RESONANT TUNNELING BARRIER

The associated $V(x, k)$ to successively more complex barriers can be constructed from the components of Eq. (A2). An example of importance to past WDF simulations is the super-
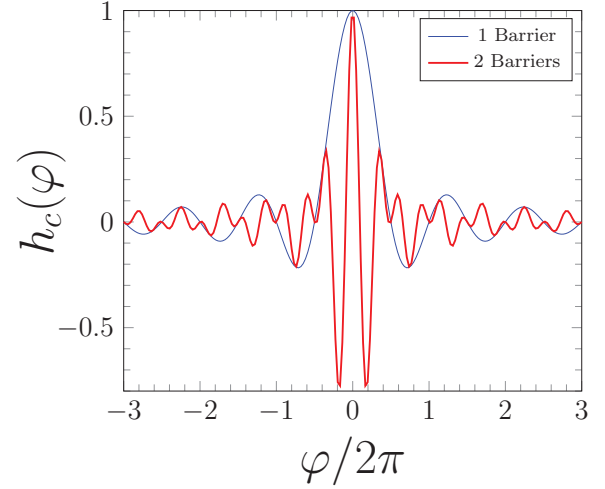
lattice barriers examined by Tsu and Esaki [134] for which the double barrier instance figures prominently in the treatment of RTDs [56,74,95,96]. The single barrier is replaced by two barriers centered at $\pm W/2$, resulting in

$$V_{\text{RTD}}(x) = V_r\left(x - \frac{W}{2}\right) + V_r\left(x + \frac{W}{2}\right), \quad \text{(B1)}$$

where $W - 2\lambda$ is the width of the well region, to which a bias potential may be added, although that complication is not considered here. Explicit evaluation shows that

$$V_{\text{RTD}}(x, k) = -4V_o\frac{\cos(kW)\sin(2k\lambda)\sin(2kx)}{\pi\hbar k}, \quad \text{(B2)}$$

which differs from $V_r(x, k)$ by a factor of $2\cos(kW)$. The effects of the additional factor are shown in Fig. 25.

Finally, RTDs are often subject to a potential drop that can be represented as the inclusion of an additional bias term $V_b(x)$ that linearly varies between $(-W/2 - \lambda < x < W/2 + \lambda)$ and constant at 0 for $x < -W/2 - \lambda$ but at $-\Delta V$ for $x > W/2 + \lambda$. The resulting $V_b(x, k - k')$ is then also analytic and can be evaluated using the techniques introduced in the consideration of Eqs. (21) and (A5).

APPENDIX C: GAUSSIAN SHAPE FACTORS

A commonly used approximation to the tunneling probability $D[E(k)]$ for thermal and field emission [135] uses the Kemble approximation,

$$D[E(k)] = \frac{1}{1 + e^{\theta[E(k)]}}, \quad \text{(C1)}$$

and is exact for a parabolic barrier [136] and very good for barriers such as $V_p(x)$ [6]. The Gamow factor in it is given by

$$\theta(E) \equiv \frac{2\sqrt{2m}}{\hbar}\int_{x_-(E)}^{x_+(E)}\sqrt{V(x) - E}\,dx, \quad \text{(C2)}$$

where $x_\pm$ are such that $V(x_\pm) - E = 0$. The shape factor method [70,132] expresses the Gamow factor as

$$\theta(E) \equiv 2\sigma(E)\kappa(E)L(E), \tag{C3}$$

where $\kappa(E) \equiv \sqrt{2m(V_o - E)}/\hbar = \sqrt{k_v^2 - k^2}$ is the maximum height of the integrand, $L(E) = x_+(E) - x_-(E)$ is the width of the integration region, and $\sigma(E)$ accounts for the proportion of the integration area $\kappa(E)L(E)$ that is occupied by the integrand. For the rectangular barrier, the shape factor $\sigma_r(E) \equiv 1$, such that $\theta_o(E) \equiv 2\kappa(E)L$ and L is the width of the barrier because a rectangular barrier is of constant height and width. For general barriers, $\sigma(E)$ and the companion $u(E)$ are [132]

$$\sigma(E) = \frac{1}{L(E)} \int_{x_-(E)}^{x_+(E)} \left[\frac{V(x) - E}{V_o - E} \right]^{1/2} dx, \tag{C4}$$

$$u(E) = \frac{1}{L(E)} \int_{x_-(E)}^{x_+(E)} \left[\frac{V_o - E}{V(x) - E} \right]^{1/2} dx. \tag{C5}$$

As a result, for $V(x) = V_p(x)$ of Eq. (A4), $x_\pm(E) = \pm\lambda\sqrt{1 - (E/V_o)}$ and so

$$\sigma_p = \int_0^1 \sqrt{1 - s^2}\, ds = \frac{\pi}{4},$$

$$u_p = \frac{1}{2} \int_0^1 \frac{ds}{\sqrt{s(1 - s)}} = \frac{\pi}{2}, \tag{C6}$$

$$\theta_p(k) = \left(\frac{\pi\lambda}{k_v} \right)(k_v^2 - k^2) = \frac{\pi}{4}\theta_o[E(k)] \tag{C7}$$

for the parabolic barrier [132], both of which are constant. Observe that no restriction appears on E to keep it below V_o: in fact, extending E past V_o provides a good account of $D(E > V_o)$ for parabolic barriers [6].

The Gaussian barrier for $V(x) = V_g(x)$ of Eq. (A3) proceeds analogously. A form more amenable to numerical integration uses $x_\pm(E) = \lambda\sqrt{\eta(E)}$ where $\eta(E) \equiv \ln(V_o/E) = 2\ln(k_v/k)$ results in

$$\sigma_g[\eta(E)] = \int_0^{\pi/2} \left[\frac{e^{\eta\cos^2 s} - 1}{e^\eta - 1} \right]^{1/2} \cos s\, ds. \tag{C8}$$

Limiting cases are $\sigma_g(0) = \pi/4$ and $\sigma_g(\eta \gg 1) \approx \sqrt{\pi/2\eta}$. The variation of $\sigma_g(\eta)$ with η is well fit by a quadratic equation with coefficients anticipated from finite difference approximations to derivatives, giving

$$\sigma_g(\eta) \approx A + B\eta + C\eta^2, \tag{C9}$$

where

$$A = \sigma_g(0) = 0.78540,$$

$$B = \tfrac{1}{2}[-3\sigma_g(0) + 4\sigma_g(1) - \sigma_g(2)] = -0.050396,$$

$$C = \tfrac{1}{2}[\sigma_g(0) - 2\sigma_g(1) + \sigma_g(2)] = -0.0011621,$$

and $\sigma_g(1) = 0.73384$ and $\sigma_g(2) = 0.67996$, an approximation for which the error is $<0.04\%$ for $E > V_o/8$. Consequently, a useful approximation is

$$\theta_g(E) = 4\lambda\sqrt{\eta}\,\sigma_g(\eta)\kappa(E), \tag{C10}$$

where $\sigma_g(\eta)$ is approximated by Eq. (C9). It follows $\theta_g/\theta_p < [\eta e^\eta/(e^\eta - 1)]^{1/2} \approx 1 + (\eta/4)[1 + (\eta/24)]$

APPENDIX D: DWELL TIME AND HARTMAN EFFECT

Hartman [118] considered the tunneling of a wave packet through a metal-insulator-metal (MIM) structure, modeled as a rectangular barrier equivalent to Eq. (A2). He found that that for sufficiently thick barriers ($k_v\lambda \gg 1$ in the present notation), tunneling is dominated by contributions smaller than but near to the top of the barrier ($D[E(k)]$ sharply peaked near $k \lesssim k_v$ in the present notation). He further found that the "transmission time" (his δt_3) converges to $(k_v^2/k\kappa)\tau_v$ (in the present notation where $\tau_v \equiv \hbar/V_o$) in the same thick barrier limit—that is, it is independent of the width of the barrier, an effect subsequently designated the *Hartman effect* [54,63,64,137].

Winful compactly summarizes the Hartman effect in Ref. [63] as the saturation of τ_g (the group delay) with increasing λ, where he disentangles the role of self-interference of the reflected wave so as to find

$$\tau_g = \tau_d + \tau_i,$$

with τ_d being the dwell time and τ_i being the self-interference delay with the latter given by $\tau_i = -(\hbar/k)\Im(r)\partial_E k$, where $\Im(r)$ is the imaginary part of the reflection coefficient $r(k)$ (cf. Eqs. (9) and (10) of Ref. [63]), e.g., for an incident plane wave of e^{ikx}, the reflected plane wave is of the form $r(k)e^{-ikx}$. Following Ref. [6], introduce $\theta \equiv 2\kappa L = 4\kappa\lambda$. In terms of θ, Winful's Eqs. (16) and (17) of Ref. [63] are

$$\frac{\tau_d}{\tau_v} = \left(\frac{kk_v^2}{2\kappa} \right) \frac{N_a + N_b}{D_o}, \tag{D1}$$

$$\frac{\tau_i}{\tau_v} = \left(\frac{\kappa k_v^2}{2k} \right) \frac{N_a}{D_o}, \tag{D2}$$

where the numerator (N_a, N_b) and denominator (D_o) terms are given by

$$N_a = 2k_v^2 \tanh(\theta/2), \quad N_b = (\kappa^2 - k^2)\theta\,\mathrm{sech}^2(\theta/2),$$

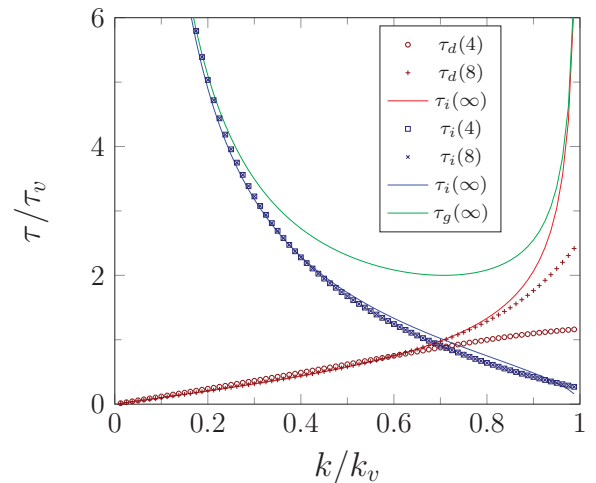$$D_o = k_v^4 - (\kappa^2 - k^2)^2 \tanh^2(\theta/2). \tag{D3}$$

FIG. 26. Delay $\tau_d(\theta)$ [Eq. (D1)] and self-interference $\tau_i(\theta)$ [Eq. (D2)] for a rectangular barrier with $V_o = 8$ eV and $\theta = 4$ ($\circ$, $\square$) and 8 ($\times$, $+$). The asymptotic forms of Eq. (D4) are the red, blue, and green lines for [$\tau_d(\infty)$, $\tau_i(\infty)$, $\tau_g(\infty)$], respectively.

The behavior of $\tanh(\theta/2)$ and $\mathrm{sech}(\theta/2)$ alone govern the width dependence of τ_d and τ_i. The $\lambda \to \infty$ limit corresponds to the $\theta \to \infty$ limit, for which $N_a \to 2k_v^2$, $N_b \to 0$, and $D_o \to k_v^4$, resulting in

$$\frac{\tau_d}{\tau_v} \to \frac{k}{\kappa}, \quad \frac{\tau_i}{\tau_v} \to \frac{\kappa}{k}, \quad \frac{\tau_g}{\tau_v} \to \frac{k}{\kappa} + \frac{\kappa}{k} = \frac{k_v^2}{k\kappa}, \quad \text{(D4)}$$

and is equivalent to Eq. (19) of Winful, although a better approximation for large but finite θ is to retain N_a/D_o as in Eq. (53). If saturation is said to occur for $\tanh(\theta/2) > 1 - \delta$, then $\theta \gtrsim \ln[(2/\delta) - 1]$, or $\theta \gtrsim 6$ for $\delta = 0.005$. The behavior of Eqs. (D1) and (D2) is compared to the asymptotic limit of Eq. (D4) in Fig. 26 for values of θ that bracket $\theta = 6$.

[1] J. R. Harris, K. L. Jensen, and D. A. Shiffler, J. Phys. D: Appl. Phys. **48**, 385203 (2015).

[2] J. R. Harris, K. L. Jensen, J. J. Petillo, S. Maestas, W. Tang, and D. A. Shiffler, J. Appl. Phys. **121**, 203303 (2017).

[3] M. Cahay, W. Zhu, J. Ludwick, K. Jensen, R. Forbes, S. Fairchild, T. Back, P. Murray, J. Harris, and D. Shiffler, in *Nanotube Superfiber Materials* (Elsevier, Amsterdam, Netherland, 2019), pp. 511–539.

[4] P. J. Price, Semicond. Sci. Technol. **19**, S241 (2004).

[5] M. Cahay and S. Bandyopadhyay, *Problem Solving in Quantum Mechanics* (John Wiley & Sons, Hoboken, New Jersey, 2017), p. 368.

[6] K. L. Jensen, D. Finkenstadt, D. A. Shiffler, A. Shabaev, S. G. Lambrakos, N. A. Moody, and J. J. Petillo, J. Appl. Phys. **123**, 065301 (2018).

[7] P. Kapur and R. Peierls, Proc. R. Soc. A **163**, 606 (1937).

[8] C. J. Edgcombe and N. d. Jonge, J. Phys. D: Appl. Phys. **40**, 4123 (2007).

[9] A. Kyritsakis, G. Kokkorakis, J. Xanthakis, T. Kirk, and D. Pescia, Appl. Phys. Lett. **97**, 023104 (2010).

[10] C. J. Edgcombe, in *Advances in Imaging and Electron Physics*, edited by P. Hawkes (Elsevier, Amsterdam, Netherlands, 2010), Vol. 162, pp. 77–127.

[11] K. L. Jensen, D. A. Shiffler, J. R. Harris, I. M. Rittersdorf, and J. J. Petillo, J. Vac. Sci. Technol. B **35**, 02C101 (2017).

[12] V. Filip, L. D. Filip, and H. Wong, Solid-State Electron. **138**, 3 (2017).

[13] P. R. Schwoebel, C. A. Spindt, and C. Holland, J. Vac. Sci. Technol. B **21**, 433 (2003).

[14] K. L. Jensen, in *Wiley Encyclopedia of Electrical and Electronics Engineering*, edited by J. G. Webster (John Wiley & Sons, Inc., New York, 2014), pp. 1–29.

[15] R. G. Forbes, J. Appl. Phys. **120**, 054302 (2016).

[16] W. Li and D. Y. Li, J. Chem. Phys. **122**, 064708 (2005).

[17] D. A. Dimitrov, G. I. Bell, J. Smedley, I. Ben-Zvi, J. Feng, S. Karkare, and H. A. Padmore, J. Appl. Phys. **122**, 165303 (2017).

[18] G. S. Gevorkyan, S. Karkare, S. Emamian, I. V. Bazarov, and H. A. Padmore, Phys. Rev. Accel. Beams **21**, 093401 (2018).

[19] R. S. Brayfield, A. J. Fairbanks, A. M. Loveless, S. Gao, A. Dhanabal, W. Li, C. Darr, W. Wu, and A. L. Garner, J. Appl. Phys. **125**, 203302 (2019).

[20] K. L. Jensen, M. McDonald, O. Chubenko, J. R. Harris, D. A. Shiffler, N. A. Moody, J. J. Petillo, and A. J. Jensen, J. Appl. Phys. **125**, 234303 (2019).

[21] K. L. Jensen, M. McDonald, J. R. Harris, D. A. Shiffler, M. Cahay, and J. J. Petillo, J. Appl. Phys. **126**, 245301 (2019).

[22] J. R. Malayter and A. L. Garner, AIP Adv. **10**, 095110 (2020).

[23] L. Nilsson, O. Groening, C. Emmenegger, O. Kuettel, E. Schaller, L. Schlapbach, H. Kind, J.-M. Bonard, and K. Kern, Appl. Phys. Lett. **76**, 2071 (2000).

[24] J. R. Harris, K. L. Jensen, D. A. Shiffler, and J. J. Petillo, Appl. Phys. Lett. **106**, 201603 (2015).

[25] F. Andreas, S. M. Marwan, and G. F. Richard, J. Vac. Sci. Technol. B **31**, 032201 (2013).

[26] Y. Feng and J. Verboncoeur, Phys. Plasmas **13**, 073105 (2006).

[27] J. J. Petillo, D. N. Panagos, and K. L. Jensen, in *IEEE International Vacuum Electronics Conference, Monterey, CA, USA* (IEEE, 2014), pp. 79–80.

[28] K. Torfason, A. Valfells, and A. Manolescu, Phys. Plasmas **22**, 033109 (2015).

[29] S. Bhattacharjee, A. Vartak, and V. Mukherjee, Appl. Phys. Lett. **92**, 191503 (2008).

[30] W. Koh and L. Ang, Appl. Phys. Lett. **89**, 183107 (2006).

[31] W. S. Koh and L. K. Ang, Nanotechnology **19**, 235402 (2008).

[32] M. Krüger, C. Lemell, G. Wachter, J. Burgdörfer, and P. Hommelhoff, J. Phys. B: At. Mol. Opt. Phys. **51**, 172001 (2018).

[33] S. Banerjee and P. Zhang, AIP Adv. **9**, 085302 (2019).

[34] K. L. Jensen, A. Shabaev, S. G. Lambrakos, D. Finkenstadt, N. A. Moody, A. J. Neukirch, S. Tretiak, D. A. Shiffler, and J. J. Petillo, J. Appl. Phys. **127**, 235301 (2020).

[35] P. Hommelhoff, Y. Sortais, A. Aghajani-Talesh, and M. A. Kasevich, Phys. Rev. Lett. **96**, 077401 (2006).

[36] C. Kealhofer, S. M. Foreman, P. Hommelhoff, and M. A. Kasevich, in *Ultrafast Phenomena XVI*, edited by P. Corkum, S. DeSilvestri, K. Nelson, E. Riedle, and R. Schoenlein (Springer, Berlin, 2009), pp. 702–704.

[37] P. Zhang and Y. Y. Lau, Sci. Rep. **6**, 19894 (2016).

[38] P. Zhang and Y. Y. Lau, J. Plasma Phys. **82**, 595820505 (2016).

[39] J. K. Bae, I. Bazarov, P. Musumeci, S. Karkare, H. Padmore, and J. Maxson, J. Appl. Phys. **124**, 244903 (2018).

[40] A. Tafel, S. Meier, J. Ristein, and P. Hommelhoff, Phys. Rev. Lett. **123**, 146802 (2019).

[41] K. L. Jensen, D. A. Shiffler, I. M. Rittersdorf, J. L. Lebowitz, J. R. Harris, Y. Y. Lau, J. J. Petillo, W. Tang, and J. W. Luginsland, J. Appl. Phys. **117**, 194902 (2015).

[42] L. V. Keldysh, Sov. Phys. JETP **20**, 1307 (1965).

[43] M. Büttiker and R. Landauer, Phys. Rev. Lett. **49**, 1739 (1982).

[44] P. J. Price, Phys. Rev. B **38**, 1994 (1988).

[45] E. H. Hauge and J. A. Støvneng, Rev. Mod. Phys. **61**, 917 (1989).

[46] O. Costin, R. Costin, I. Jauslin, and J. L. Lebowitz, J. Appl. Phys. **124**, 213104 (2018).

[47] R. Ramos, D. Spierings, I. Racicot, and A. Steinberg, Nature **583**, 529 (2020).

[48] F. A. Buot and A. K. Rajagopal, Superlattices Microstruct. **23**, 641 (1998).

[49] P. C. W. Davies, Am. J. Phys. **73**, 23 (2005).

[50] H. R. Reiss, Phys. Rev. Lett. **101**, 043002 (2008).

[51] H. R. Reiss, Phys. Rev. A **82**, 023418 (2010).

[52] A. S. Landsman and U. Keller, Phys. Rep. **547**, 1 (2015).

[53] T. Zimmermann, S. Mishra, B. R. Doran, D. F. Gordon, and A. S. Landsman, Phys. Rev. Lett. **116**, 233603 (2016).

[54] K. Smith and G. Blaylock, Am. J. Phys. **85**, 763 (2017).

[55] C. Dewdney and B. Hiley, Found. Phys. **12**, 27 (1982).

[56] F. A. Buot and K. L. Jensen, Phys. Rev. B **42**, 9429 (1990).

[57] M. Razavy, Phys. Lett. A **212**, 119 (1996).

[58] K. Na and R. E. Wyatt, Int. J. Quantum Chem. **81**, 206 (2001).

[59] D. Sels, F. Brosens, and W. Magnus, Physica A **391**, 78 (2012).

[60] E. Yakaboylu, M. Klaiber, and K. Z. Hatsagortsyan, Phys. Rev. A **90**, 012116 (2014).

[61] K. L. Jensen, D. A. Shiffler, J. L. Lebowitz, M. Cahay, and J. J. Petillo, J. Appl. Phys. **125**, 114303 (2019).

[62] K. L. Jensen, D. A. Shiffler, J. M. Riga, J. R. Harris, J. L. Lebowitz, M. Cahay, and J. J. Petillo, J. Appl. Phys. **126**, 144301 (2019).

[63] H. G. Winful, Phys. Rev. Lett. **91**, 260401 (2003).

[64] H. G. Winful, Phys. Rep. **436**, 1 (2006).

[65] M. J. Hagmann, J. Vac. Sci. Technol. B **13**, 403 (1995).

[66] Y. Y. Lau, Y. Liu, and R. Parker, Phys. Plasmas **1**, 2082 (1994).

[67] R. H. Fowler and L. Nordheim, Proc. R. Soc. A **119**, 173 (1928).

[68] R. G. Forbes and J. H. B. Deane, Proc. R. Soc. London A **463**, 2907 (2007).

[69] K. L. Jensen, *Introduction to the Physics of Electron Emission* (John Wiley & Sons, Inc., NJ, 2017).

[70] K. L. Jensen, D. A. Shiffler, M. Peckerar, J. R. Harris, and J. J. Petillo, J. Appl. Phys. **122**, 064501 (2017).

[71] K. L. Jensen, Y. Y. Lau, D. W. Feldman, and P. G. O'Shea, Phys. Rev. ST Accel. Beams **11**, 081001 (2008).

[72] M. Hillery, R. Oconnell, M. Scully, and E. Wigner, Phys. Rep. **106**, 121 (1984).

[73] H.-W. Lee, Phys. Rep. **259**, 147 (1995).

[74] W. R. Frensley, Phys. Rev. B **36**, 1570 (1987).

[75] N. Kluksdahl, A. Kriman, and D. Ferry, Superlattices Microstruct. **4**, 127 (1988).

[76] W. R. Frensley, Rev. Mod. Phys. **62**, 745 (1990).

[77] B. A. Biegel and J. D. Plummer, Phys. Rev. B **54**, 8070 (1996).

[78] P. Zhao, H. Cui, D. Woolard, K. L. Jensen, and F. A. Buot, Int. J. Mod. Phys. B **14**, 411 (2000).

[79] K.-Y. Kim, J. Kim, and S. Kim, AIP Adv. **6**, 065314 (2016).

[80] R. E. Salvino and F. A. Buot, J. Appl. Phys. **72**, 5975 (1992).

[81] X. Oriols, F. Martin, and J. Sune, Solid State Commun. **99**, 123 (1996).

[82] F. Martin, J. Garcia-Garcia, X. Oriols, and J. Sune, Solid-State Electron. **43**, 315 (1999).

[83] M. J. W. Hall, M. S. Reineker, and W. P. Schleich, J. Phys. A: Math. Gen. **32**, 8275 (1999).

[84] E. Colomés, Z. Zhan, and X. Oriols, J. Comput. Electron. **14**, 894 (2015).

[85] A. Donoso and C. C. Martens, Phys. Rev. Lett. **87**, 223202 (2001).

[86] D. Bohm, Phys. Rev. **85**, 166 (1952).

[87] D. Bohm, Phys. Rev. **85**, 180 (1952).

[88] L. de Broglie, C. R. Acad. Sci. Paris **183** (1926).

[89] E. Madelung, Z. Phys. **40**, 322 (1927).

[90] *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton University Press, Princeton, NJ, 1983).

[91] D. Heim, W. Schleich, P. Alsing, J. Dahl, and S. Varro, Phys. Lett. A **377**, 1822 (2013).

[92] A. Goldberg, H. M. Schey, and J. L. Schwartz, Am. J. Phys. **35**, 177 (1967).

[93] K. L. Jensen, J. J. Petillo, S. Ovtchinnikov, N. Panagos, D. N. A. Moody, and S. G. Lambrakos, J. Appl. Phys. **122**, 164501 (2017).

[94] K. L. Jensen, IEEE Trans. Plasma Sci. **46**, 1881 (2018).

[95] N. C. Kluksdahl, A. M. Kriman, D. K. Ferry, and C. Ringhofer, Phys. Rev. B **39**, 7720 (1989).

[96] P. Zhao, H. Cui, D. Woolard, K. Jensen, and F. Buot, J. Appl. Phys. **87**, 1337 (2000).

[97] K. L. Jensen, D. Finkenstadt, A. Shabaev, S. G. Lambrakos, N. A. Moody, J. J. Petillo, H. Yamaguchi, and F. Liu, J. Appl. Phys. **123**, 045301 (2018).

[98] L. Cultrera, A. Galdi, J. K. Bae, F. Ikponmwen, J. Maxson, and I. Bazarov, Phys. Rev. Accel. Beams **23**, 023401 (2020).

[99] M. Krüger, M. Schenk, M. Förster, and P. Hommelhoff, J. Phys. B: At., Mol. Opt. Phys. **45**, 074006 (2012).

[100] Y. S. Kim and M. E. Noz, *Phase Space Picture of Quantum Mechanics: Group Theoretical Approach* (World Scientific, Singapore, 1991).

[101] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1968).

[102] G. J. Iafrate, V. N. Sokolov, and J. B. Krieger, Phys. Rev. B **96**, 144303 (2017).

[103] G. D. Smith, *Numerical Solution of Partial Differential Equations: Finite Difference Methods*, Oxford Applied Mathematics and Computing Science Series (Clarendon Press Oxford University Press, Oxford, 1985).

[104] S. M. Blinder, Am. J. Phys. **36**, 525 (1968).

[105] C.-T. Claude, D. Bernard, and L. Franck, *Quantum Mechanics* (Wiley-Interscience, New York, 1977), pp. 69–78.

[106] C. Jacoboni and P. Bordone, Rep. Prog. Phys. **67**, 1033 (2004).

[107] J. Weinbub and D. K. Ferry, Appl. Phys. Rev. **5**, 041104 (2018).

[108] Figure 1 of Ref. [107] has as the cosine term the factor $\cos(2k_o x)$ instead, which has been corrected here.

[109] V. Sverdlov, E. Ungersboeck, H. Kosina, and S. Selberherr, Mater. Sci. Eng., R **58**, 228 (2008).

[110] D. Querlioz and P. Dollfus, *The Wigner Monte-Carlo Method for Nanoelectronic Devices: Particle Description of Quantum Transport and Decoherence* (Wiley, London, 2010), p. 243.

[111] K. L. Jensen and F. A. Buot, Phys. Rev. Lett. **66**, 1078 (1991).

[112] The calculations of Refs. [56,111] were on a Cray X-MP (circa 1991) making use of IMSL matrix solution subroutines and FFT algorithms. The platform for present calculations used an iMac desktop with a 3 GHz Intel Core i5 processor running MATLAB R2019b interactively, no matrix inversions, and no specialized numerical packages other than vectorization. Execution would be even faster were the present MATLAB scripts optimized and rendered into a compiled language like Fortran. The speed-up comparison is therefore illustrative, not quantitative.

[113] K. L. Jensen and F. A. Buot, IEEE Trans. Electron Devices **38**, 2337 (1991).

[114] Observe that in standard LU decomposition, L would refer to the entries below the diagonal and, conversely, for U: Here, instead, L and U refer to the negative and positive regions of $f(x, k, t)$, respectively.

[115] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in Fortran: The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1992).

[116] K. L. Jensen and F. A. Buot, J. Appl. Phys. **67**, 2153 (1990).

[117] W. Elberfeld and M. Kleber, Am. J. Phys. **56**, 154 (1988).

[118] T. E. Hartman, J. Appl. Phys. **33**, 3427 (1962).

[119] M. H. Bramhall and B. M. Casper, Am. J. Phys. **38**, 1136 (1970).

[120] K. L. Jensen, J. M. Riga, J. L. Lebowitz, R. Seviour, and D. A. Shiffler (unpublished).

[121] L. K. Ang and P. Zhang, Phys. Rev. Lett. **98**, 164802 (2007).

[122] S. Karkare, D. Dimitrov, W. Schaff, L. Cultrera, A. Bartnik, X. Liu, E. Sawyer, T. Esposito, and I. V. Bazarov, J. Appl. Phys. **113**, 104904 (2013).

[123] Y. Ang, S.-J. Liang, and L. Ang, MRS Bull. **42**, 505 (2017).

[124] J. K. Bae, L. Cultrera, P. DiGiacomo, and I. Bazarov, Appl. Phys. Lett. **112**, 154101 (2018).

[125] O. Chubenko, S. S. Baturin, and S. V. Baryshev, J. Appl. Phys. **125**, 205303 (2019).

[126] Note that $kD(k)f_s(k)$ would better mimic current density but, for present purposes, the difference is small.

[127] Y. E. Lozovik and A. V. Filinov, J. Exp. Theor. Phys. **88**, 1026 (1999).

[128] J. Petersen and E. Pollak, J. Phys. Chem. Lett. **8**, 4017 (2017).

[129] J. Petersen and E. Pollak, J. Phys. Chem. A **122**, 3563 (2018).

[130] G. Herink, L. Wimmer, and C. Ropers, New J. Phys. **16**, 123005 (2014).

[131] Z. Shu, X. Hao, W. Li, and J. Chen, Chin. Phys. B **28**, 050301 (2019).

[132] K. L. Jensen, J. Appl. Phys. **111**, 054916 (2012).

[133] E. Butkov, *Mathematical Physics* (Addison-Wesley Pub. Co, Reading, MA, 1968).

[134] R. Tsu and L. Esaki, Appl. Phys. Lett. **22**, 562 (1973).

[135] E. L. Murphy and R. H. Good, Phys. Rev. **102**, 1464 (1956).

[136] N. Fröman and P. O. Fröman, *JWKB Approximation: Contributions to the Theory* (North-Holland Pub. Co., New York, 1965).

[137] V. Petrillo and V. Olkhovsky, Open Phys. **3**, 339 (2005).