Tuning of ungated plasmons by a gate in the field-effect transistor with two-dimensional electron channel

V. V. Popov' , A. N. Koudymov, M. Shur, and O. V. Polischuk

Citation: Journal of Applied Physics **104**, 024508 (2008); doi: 10.1063/1.2955731
View online: http://dx.doi.org/10.1063/1.2955731
View Table of Contents: http://aip.scitation.org/toc/jap/104/2
Published by the American Institute of Physics

Tuning of ungated plasmons by a gate in the field-effect transistor with two-dimensional electron channel

V. V. Popov,[1,a] A. N. Koudymov,[1] M. Shur,[1] and O. V. Polischuk[2]
[1]*Department of Electrical, Computer, and System Engineering and Center for Integrated Electronics, CII 9015, Rensselaer Polytechnic Institute, Troy, New York 12180, USA*
[2]*Institute of Radio Engineering and Electronics (Saratov Branch), Russian Academy of Sciences, 410019 Saratov, Russia*

(Received 9 April 2008; accepted 2 May 2008; published online 23 July 2008)

We show that voltage variations at a short gate can effectively tune higher-order ungated plasmon resonances in field-effect transistors (FETs). These higher-order ungated plasmon resonances may be excited by incoming terahertz radiation with much greater efficiency than the gated plasmon resonances. We calculate the spectra of the terahertz plasmon absorption in the frame of a rigorous electromagnetic approach, which allows us to estimate the radiation resistance. Based on the calculation results, we explain the behavior of different plasmon resonances in terms of the alternating-current FET equivalent circuit. The results may help design high performance plasmonic devices operating in the terahertz frequency range. © *2008 American Institute of Physics*.
[DOI: 10.1063/1.2955731]

I. INTRODUCTION

The high-frequency response of field-effect transistors (FETs) with two-dimensional (2D) electron channels is strongly affected by plasma oscillations excited in the FET structure. This phenomenon in its various manifestations can be used for the detection, frequency multiplication, and generation of terahertz radiation.[1–5] Two different types of plasma oscillations exist in a FET structure. They are plasma oscillations excited in gated or ungated regions of the electron channel and termed as gated and ungated plasmons, respectively. The gated plasmons are considered to be more attractive for electronic applications because their frequencies can be effectively tuned by varying the gate voltage. For identical boundary conditions at the source and drain sides of the gated section of the channel, the frequencies of plasma oscillations excited under the FET gate can be estimated as[1]

$$\bar{\omega}_n = \sqrt{\frac{e^2 N_s^{(g)} d}{m^* \varepsilon_0 \varepsilon_b}} \frac{\pi}{l_{\text{eff}}}(2n-1) \quad (n=1,2,3,\ \ldots), \tag{1}$$

where $N_s^{(g)}$ is the sheet electron density in the channel under the gate electrode, l_{eff} is the effective gate length related to the geometrical gate length l by $l_{\text{eff}}=l+2d$ with d being the gate-to-channel distance, and e and m^* are the electronic charge and effective mass, respectively. [Equation (1) applies when $\bar{\omega}_n \tau \gg 1$, where τ is the electron scattering time in 2D electron layer.] The sheet electron density in the channel under the gate electrode, $N_s^{(g)}$, can be approximately estimated using the parallel plate capacitor model as

$$N_s^{(g)} = \frac{\varepsilon_0 \varepsilon_b}{ed} U_0, \tag{2}$$

where ε_b is the dielectric constant of the barrier layer separating the gate contact from the 2D electron channel, ε_0 is the permittivity of free space, and U_0 is the effective gate-to-channel voltage swing, which is the difference between the gate voltage U_g and the channel depletion threshold voltage U_{th}: $U_0=U_g-U_{\text{th}}$. For the typical parameters of InGaAs high-electron-mobility transistor with a submicron gate, $l=400$ nm, $d=27$ nm, $\varepsilon_b=13.88$, $N_s^{(g)}=10^{12}$ cm^{-2}, and $m^*=0.042m_0$, where m_0 is the free-electron mass, Eq. (1) yields the fundamental plasmon frequency $(n=1)$ about 1.34 THz. However, the gated plasmons are weakly coupled to the terahertz radiation because (i) they are strongly screened by the metal gate electrode, (ii) they have a vanishingly small net dipole moment due to their acoustic nature, and (iii) they strongly leak into the ungated plasmons in access regions of the channel.[6] That is why the gated plasmons are difficult to excite by impinging terahertz radiation upon the FET without using special antenna elements coupling the gated plasmons to the terahertz radiation.

The effect of the ungated portions of the channel on the resonant frequencies of the gated plasmons was studied in Refs. 4 and 7, where the ungated portions of the channel were considered as resistive regions. Such a model is only applicable when the frequency of the ungated plasmons is much higher than the frequency of the gated plasmons.

The frequencies of ungated plasmon modes in a finite 2D electron channel bridging the gap between the source and drain contacts in a slot diode can be estimated as[8]

$$\omega_n^2 = \frac{e^2 N_s^{(\text{un})}}{m^* \varepsilon_0 (\varepsilon_1+\varepsilon_2) L_{\text{eff}}} \pi(2n-1) \quad (n=1,2,3,\ \ldots), \tag{3}$$

where $N_s^{(\text{un})}$ is the sheet electron density in the ungated channel, L_{eff} is the effective length of the ungated channel in the

[a]On leave from Institute of Radio Engineering and Electronics, Saratov, Russia. Electronic mail: popov@soire.renet.ru.

slot diode ($L_{\text{eff}} > L$, where L is the geometrical length of the channel due to the effect of image charges induced in the source and drain metal contacts[8]), and ε_1 and ε_2 are the effective dielectric constants of media under and above the 2D electron channel. [Once again, we assume that $\bar{\omega}_n \tau \gg 1$.] In contrast to the gated plasmons, the ungated plasmons have optical nature and, hence, can be effectively excited by terahertz radiation due to a large dipole moment inherent in this optically active plasmon mode.[8] In fact, the dispersion relation for the ungated plasmons given by Eq. (3) is the well known square-root dispersion law for 2D plasmons derived in Ref. 9 with the plasmon wavevector quantized due to the spatial confinement of the channel by the side metal contacts.

In this paper, we show that the frequencies of higher-order resonances of the ungated plasmon modes can be effectively tuned by the gate-voltage variation at a short gate (the length of the gate is much shorter than the source-to-drain separation), which mimics conventional gated plasmon resonances but with a much stronger coupling to the terahertz radiation. In Sec. II, we describe an electromagnetic approach that we use to calculate the plasmonic properties of the FET. Calculation results are discussed in Sec. III in terms of the alternating-current FET equivalent circuit. The conclusions are summarized in Sec. IV.

II. THEORETICAL FORMALISM

We consider a 2D electron channel of length L located in the $y=0$ plane on the surface of a substrate with a dielectric constant ε_s (see Fig. 1) bridging two perfectly conductive semiplanes modeling the source and drain contacts. A barrier slab of an insulator with dielectric constant ε_b of thickness d separates the 2D electron channel from the gate contact. The gate electrode is assumed to be perfectly conductive and infinitely long along the z-axis strip of width l (which is the gate length of the FET according to a conventional definition) and is of zero thickness. The gate electrode is located in the $y=d$ plane on the surface of the barrier insulator slab. Above the gate ($y>d$) is the ambient medium with dielectric constant $\varepsilon_a=1$ (vacuum). A plane electromagnetic wave impinges on the FET structure normally from the top. We assume that the electric field of the incident wave $E^{(0)} \exp[-ik_y^{(0)} y - i\omega t]$ is polarized across the gate-contact strip along the x axis. Here ω is the angular frequency and $k_y^{(0)} = k_0$, where $k_0 = \omega\sqrt{\varepsilon_0 \mu_0}$ with ε_0 and μ_0 being the permittivity and permeability of free space, respectively.

We write the sought-for fields and currents in the structure in terms of the Fourier transforms. For instance, for the x-component of the electric field in any of the three media we have

$$E_x(x,y) = \int_{-\infty}^{+\infty} E_x(k_x,y)\exp(ik_x x)dk_x, \tag{4}$$

where

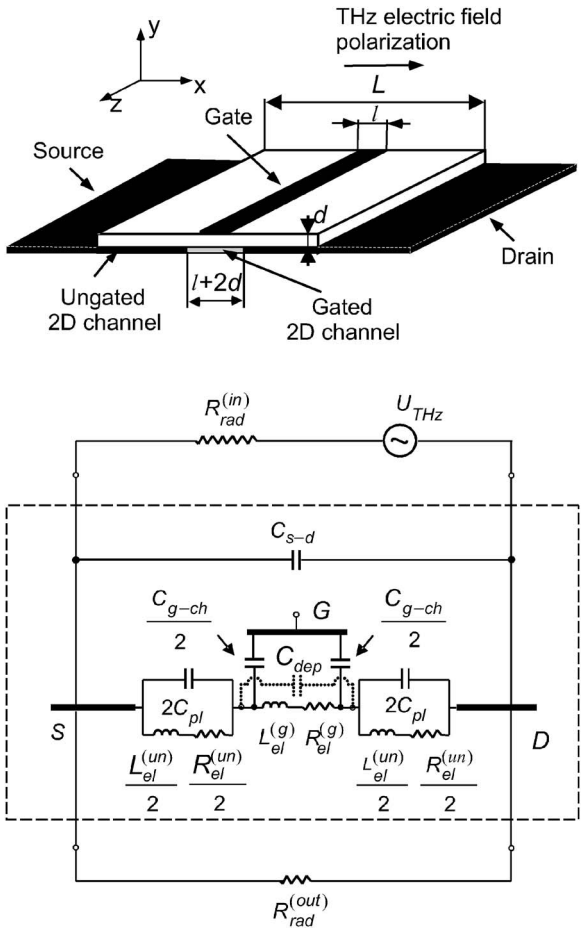

FIG. 1. Schematic of the transistor structure under consideration and coordinate system. The external terahertz radiation is incident normally from the top. An equivalent oscillating-current circuit is shown below (intrinsic transistor is in a dashed box).

$$E_x(k_x,y) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} E_x(x,y)\exp(-ik_x x)dx \tag{5}$$

is the amplitude of the Fourier transforms. Using Eq. (4) and similar corresponding formulas for every sought-for field component and electric currents in the 2D electron channel and gate electrode, we rewrite the system of the Maxwell equations in each medium in the Fourier representation. The boundary conditions for the tangential components of the electric field E_x and magnetic field H_z are

$$E_x^{(b)}(k_x,0) = E_x^{(s)}(k_x,0),$$

$$H_z^{(b)}(k_x,0) - H_z^{(s)}(k_x,0) = I_x(k_x,0) \tag{6}$$

at $y=0$ and

$$E_x^{(a)}(k_x,d) + E_x^{(0)}\exp[-ik_0 d]\delta(k_x) = E_x^{(b)}(k_x,d),$$

$$H_z^{(a)}(k_x,d) + H_z^{(0)}\exp[-ik_0 d]\delta(k_x) - H_z^{(b)}(k_x,d) = I_x(k_x,d) \tag{7}$$

at $y=d$, where $I_x(k_x,0)$ and $I_x(k_x,d)$ are the amplitudes of the Fourier transforms of the sheet current density in the planes of the 2D electron layer and in the gate strip, respectively, and $\delta(k_x)$ is the Dirac delta function. Hereafter, superscripts

(a), (b), and (s) refer to the ambient medium (vacuum in the case under consideration), barrier slab, and substrate, respectively. The current in the 2D electron layer is related to the in-plane electric field by the Ohm law $I_x(k_x,0)=\sigma(\omega)E_x(k_x,0)$, where $\sigma(\omega)$ is the sheet conductivity of the 2D plasma. We describe the response of electrons in the 2D channel by the sheet conductivity $\sigma(x,\omega)=N_s(x)e^2/m(\nu-i\omega)$, where $\nu=1/\tau$ is the electron scattering rate. The equilibrium sheet electron density $N_s(x)=N_s^{(\mathrm{un})}$ is constant in the ungated regions of the channel. Equation (2) describes the variation of $N_s(x)=N_s^{(g)}$ as a function of the gate voltage in the gated region of the 2D electron channel. Thus, in our theoretical model we assume that the electron sheet density in the 2D channel depends on the x-coordinate,

$$\sigma(\omega,x)=\begin{cases}\sigma^{(\mathrm{un})} & \text{for} \quad -\dfrac{L}{2}<x<\dfrac{l_{\mathrm{eff}}}{2} \quad \text{and} \quad \dfrac{l_{\mathrm{eff}}}{2}<x<\dfrac{L}{2}\\ \sigma^{(g)} & \text{for} \quad -\dfrac{l_{\mathrm{eff}}}{2}<x<\dfrac{l_{\mathrm{eff}}}{2}\\ 0 & \text{for} \quad |x|>\dfrac{L}{2},\end{cases}$$

where $l_{\mathrm{eff}}=l+2d$. Then we can expand the sheet impedance of the 2D electron channel, $\rho(x)=1/\sigma(x)$, into the Fourier series within the interval $-L/2<x<L/2$ as

$$\rho(x)=\sum_{p=-\infty}^{\infty}\rho_p\exp\left[ip\frac{2\pi}{L}x\right],$$

where

$$\rho_p=\frac{1}{\pi p}\sin\left(\pi p\frac{l_{\mathrm{eff}}}{L}\right)(\rho^{(g)}-\rho^{(\mathrm{un})}) \quad \text{for} \quad p\neq 0$$

and

$$\rho_0=\rho^{(\mathrm{un})}+(\rho^{(g)}-\rho^{(\mathrm{un})})\frac{l_{\mathrm{eff}}}{L}.$$

In the geometry under consideration, both the magnetic field of the incident wave and the induced magnetic field have only the z-components (TM polarization). The amplitudes of the Fourier transforms of the induced magnetic field in different media can be written in the forms

$$H_z^{(a)}(k_x,y)=A(k_x)\exp[-\alpha^{(a)}y] \quad \text{for} \quad y>d,$$

$$H_z^{(b)}(k_x,y)=B(k_x)\exp[-\alpha^{(b)}y]+C(k_x)\exp[-\alpha^{(b)}(d-y)] \quad \text{for} \quad 0<y<d,$$

$$H_z^{(s)}(k_x,y)=D(k_x)\exp[\alpha^{(s)}y] \quad \text{for} \quad y<0. \tag{8}$$

Here $\alpha^{(i)}=\pm\sqrt{k_x^2-k_0^2\varepsilon^{(i)}}$ with $i=a,b,s$. The sign before the radical in the last expression is chosen to satisfy the conditions at $y\to\pm\infty$, which are the zero field condition for the evanescent fields $(k_x>k_0)$ and the scattering condition for the radiative fields $(k_x<k_0)$. The latter condition requires that only outgoing waves exist at $y\to\pm\infty$. By making use of Eqs. (6)–(8) and the Maxwell equations in each medium in the Fourier representation, after some obvious but rather cumbersome algebra, we obtain the following relations between the in-plane electric field and currents in the planes $y=0$ and $y=d$:

$$E_x(k_x,d)=G^{(1,1)}(k_x)E_x(k_x,0)+G^{(1,2)}(k_x)I_x(k_x,d)+G^{(1,0)}(k_x)\times E^{(0)}\delta(k_x),$$

$$I_x(k_x,0)=G^{(2,1)}(k_x)E_x(k_x,0)+G^{(2,2)}(k_x)I_x(k_x,d)+G^{(2,0)}(k_x)\times E^{(0)}\delta(k_x),$$

where

$$G^{(1,1)}(k_x)=2\frac{\chi\mathrm{s}_2}{\mathrm{s}_2(\chi^2+1)-\mathrm{s}_1(\chi^2-1)},$$

$$G^{(1,2)}(k_x)=iZ_0\frac{\chi^2-1}{\mathrm{s}_2(\chi^2+1)-\mathrm{s}_1(\chi^2-1)},$$

$$G^{(2,1)}(k_x)=\frac{i}{Z_0}\left[\frac{\mathrm{s}_2(\chi^2+1)-\mathrm{s}_3(\chi^2-1)}{\chi^2-1}+\frac{4\chi^2\mathrm{s}_2^2}{\chi^2-1}\frac{1}{\mathrm{s}_1(\chi^2-1)-\mathrm{s}_2(\chi^2+1)}\right],$$

$$G^{(2,2)}(k_x)=-G^{(1,1)}(k_x),$$

and

$$G^{(1,0)}(k_x)=2\mathrm{s}_1\frac{\chi^2-1}{\mathrm{s}_2(\chi^2+1)-\mathrm{s}_1(\chi^2-1)}\exp(-ik_0d),$$

$$G^{(2,0)}(k_x)=-\frac{4i}{Z_0}\frac{\chi\mathrm{s}_1\mathrm{s}_2}{\mathrm{s}_2(\chi^2+1)-\mathrm{s}_1(\chi^2-1)}\exp(-ik_0d),$$

with $\chi=\exp(-\alpha^{(b)}d)$, $\mathrm{s}_i=k_0\varepsilon^{(i)}/\alpha^{(i)}$, and $Z_0=120\pi\ \Omega$ being the free space impedance.

Using the condition $E_x(x,d)=0$ at the perfectly conductive gate strip and the Ohm law in the 2D electron channel, we construct the following system of two coupled integral equations for the sheet current density $I_x(x,d)$ at the gate strip and for the in-plane electric field $E_x(x,0)$ at the 2D electron channel:

$$\int_{-l/2}^{l/2}[\bar{G}^{(1,1)}(x,x')E_x(x',0)+\bar{G}^{(1,2)}(x,x')I_x(x',d)]dx'=\bar{G}^{(1,0)}E^{(0)},$$

$$E_x(x,0)+\int_{-L/2}^{L/2}[\bar{G}^{(2,1)}(x,x')E_x(x',0)+\bar{G}^{(2,2)}(x,x')\times I_x(x',d)]dx'=\bar{G}^{(2,0)}E^{(0)}, \tag{9}$$

with the kernels

$$\bar{G}^{(1,1)}(x,x')=\frac{L}{2\pi}\int_{-\infty}^{\infty}G^{(1,1)}(k_x)\exp\left[ik_x\frac{1}{L}(xl-x'L)\right]dk_x,$$

$$\bar{G}^{(1,2)}(x,x')=\frac{l}{2\pi}\int_{-\infty}^{\infty}G^{(1,2)}(k_x)\exp\left[ik_x\frac{1}{L}(xl-x'l_{\mathrm{eff}})\right]dk_x,$$

$$\bar{G}^{(2,1)}(x,x') = \frac{L}{2\pi}\int_{-\infty}^{\infty} G^{(2,1)}(k_x)\rho(x)\exp[ik_x(x-x')]dk_x,$$

$$\bar{G}^{(2,2)}(x,x') = \frac{l_{\text{eff}}}{2\pi}\int_{-\infty}^{\infty} G^{(2,2)}(k_x)\rho(x)\exp\left[ik_x\frac{1}{L}(xL - x'l_{\text{eff}})\right]dk_x,$$

and $\bar{G}^{(1,0)}=-G^{(1,0)}(0)$, $\bar{G}^{(2,0)}=-\rho(x)G^{(2,0)}(0)$. The first and second equations in Eq. (9) are the Fredholm integral equations of the first and second kinds, respectively.

We solve the system of integral equations [Eq. (9)] numerically, approximating the functions $E_x(x,0)$ and $I_x(x,d)$ by the expansions

$$E_x(x,0) = \sum_{n=0}^{\infty} a_n P_n(\xi),$$

$$I_x(x,d) = \sum_{n=0}^{\infty} b_n P_n(\eta), \tag{10}$$

where $P_n(\xi)$ and $P_n(\eta)$ are the Legendre polynomials of the nth degree with $\xi=2x/L$ and $\eta=2x/l$, and a_n and b_n are unknown coefficients. Substituting Eq. (10) into the system of integral equations [Eq. (9)] and using the Galerkin procedure,[10] with the Legendre polynomials $P_n(\xi)$ as the orthogonal basis functions, we transform the system of integral equations [Eq. (9)] for the functions $E_x(x,0)$ and $I_x(x,d)$ into an infinite system of the linear algebraic equations for the coefficients a_n and b_n,

$$\sum_{n=0}^{\infty} G_{m,n}^{(1,1)} a_n + \sum_{n=0}^{\infty} G_{m,n}^{(1,2)} b_n = G^{(1,0)}E^{(0)},$$

$$a_m + \sum_{n=0}^{\infty} G_{m,n}^{(2,1)} a_n + \sum_{n=0}^{\infty} G_{m,n}^{(2,2)} b_n = G^{(2,0)}E^{(0)}, \tag{11}$$

where

$$G_{m,n}^{(1,1)} = (2m+1)i^{(2n+m)}(-1)^n\frac{L}{\pi}\int_{-\infty}^{\infty} G^{(1,1)}(k_x) \times J_n\left(k_x\frac{L}{2}\right)J_m\left(k_x\frac{l_{\text{eff}}}{2}\right)dk_x,$$

$$G_{m,n}^{(1,2)} = -(2m+1)\frac{n}{2}i^{(2n+m-1)}(-1)^n\frac{l}{l_{\text{eff}}}\int_{-\infty}^{\infty} G^{(1,2)}(k_x) \times \frac{1}{k_x}J_n\left(k_x\frac{l_{\text{eff}}}{2}\right)J_m\left(k_x\frac{l_{\text{eff}}}{2}\right)dk_x,$$

$$G_{m,n}^{(2,1)} = (2m+1)i^{(2n+m)}(-1)^n\frac{L}{\pi}\int_{-\infty}^{\infty} G^{(2,1)}(k_x) \times J_n\left(k_x\frac{L}{2}\right)\sum_{p=-\infty}^{\infty}\rho_p J_m\left(p\pi + k_x\frac{L}{2}\right)dk_x,$$

$$G_{m,n}^{(2,2)} = -(2m+1)\frac{n}{2}i^{(2n+m-1)}(-1)^n\int_{-\infty}^{\infty} G^{(2,2)}(k_x) \times \frac{1}{k_x}J_n\left(k_x\frac{l_{\text{eff}}}{2}\right)\sum_{p=-\infty}^{\infty}\rho_p J_m\left(p\pi + k_x\frac{L}{2}\right)dk_x,$$

and

$$G_{m,0}^{(1,0)} = \bar{G}^{(1,0)}\delta_{m,0},$$

$$G_{m,0}^{(2,0)} = \bar{G}^{(2,0)}i^m(2m+1)\sum_{p=-\infty}^{\infty}\rho_p J_m(p\pi).$$

Here $J_m(\zeta)$ is the spherical Bessel function of the first kind of the mth order. The infinite system of the linear algebraic equations [Eq. (11)] can then be truncated and solved numerically to achieve a desired level of convergence. As a result, we calculate the in-plane electric field in the 2D electron channel plane $E_x(x,0)$ and the current density in the gate contact $I_x(x,d)$ and, finally, calculate the absorption length $L^{(\text{ab})}=Q/P_0$, where P_0 is the energy flux density in the incident wave and

$$Q = \frac{1}{2}\text{Re}\{\sigma^{(\text{un})}\int_{-L/2}^{L/2}|E_x(x)|^2dx + (\sigma^{(g)}-\sigma^{(\text{un})})\int_{-l_{\text{eff}}/2}^{l_{\text{eff}}/2}|E_x(x)|^2dx\}$$

is the absorption power (per unit width of the channel).

It should be noted that in many previous papers (e.g., see Refs. 1, 2, 4, and 7) the quasielectrostatic approximation was used for describing the plasma oscillations in FET structures. This approximation appears to be quite sufficient for calculating the plasmon frequencies in such short (much shorter than the electromagnetic wavelength) devices. However, the quasielectrostatic approximation might fail to correctly describe the plasmon resonance line shape, even if the dimensions of the device are much shorter than the electromagnetic wavelength. Without considering the electromagnetic effects, one is not able to obtain the radiative part of the plasmon resonance, which may provide the dominant contribution to the total plasmon resonance linewidth of the ungated plasmons due to the optical nature of this plasmon mode. (In Sec. III we show that the radiative contribution to the total linewidth of the fundamental ungated plasmon resonance is as much as seven times greater than the dissipative contribution.)

III. RESULTS AND DISCUSSION

The left panel in Fig. 2 shows the calculated terahertz absorption spectra of a FET with a short gate for several different values of the gate length l. If the gate length is much shorter than the total length of the channel L (the bottom spectrum), the gate contact does not noticeably influence the resonances of the ungated plasmon modes excited in the source-to-drain spacing. As a result, the terahertz absorption spectrum of FET with an ultrashort gate essentially replicates the fundamental and second-order plasmon resonances in a

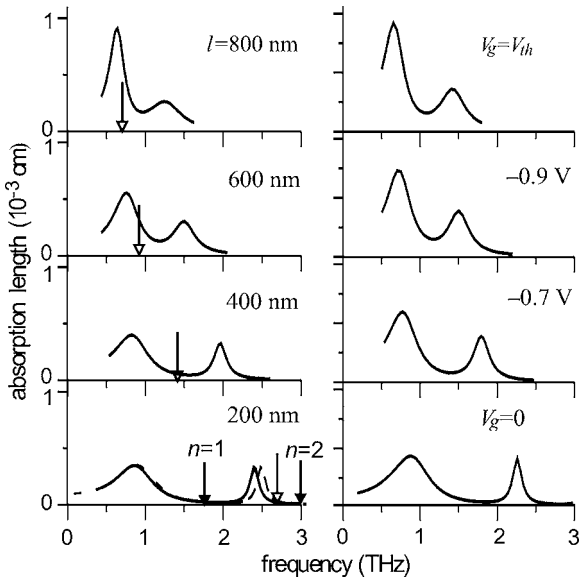

FIG. 2. Absorption length of FET as a function of frequency for (left panel) different lengths of the gate for fixed electron densities $N_s^{(g)}=10^{12}$ cm^{-2}, ($V_g=-0.7$ V), $N_s^{(un)}=3\times10^{12}$ cm^{-2}, and (right panel) for different gate-voltage values for a fixed gate length $l=400$ nm. Filled arrows in the left bottom panel show the frequencies of the fundamental ($n=1$) and second-order ($n=2$) ungated plasmon modes calculated using Eq. (3) for $L_{\text{eff}}=L$. Unfilled arrows in the left panel show the frequencies of the fundamental gated plasmon mode calculated using Eq. (1). The FET parameters are $L=4$ μm, $d=27$ nm, and $\tau=2.3\times10^{-12}$ s. The sheet electron density in the gated portion of the channel was calculated using Eq. (2) with $V_{\text{th}}=-1.0$ V.

slot diode with 2D electron channel.[8] (The dashed line in the left bottom panel in Fig. 2 shows the absorption-length spectrum of the FET structure with no gate.) In this case, fitting the resonance frequencies by Eq. (3) with $\varepsilon_1=\varepsilon_s$ and $\varepsilon_2=\varepsilon_a=1$ yields the relative effective channel lengths $L_{\text{eff}}/L=4.5$ and $L_{\text{eff}}/L=1.45$ for the fundamental and the second-order ungated plasmon resonances, respectively.

A considerable excess of the effective length of the channel, L_{eff}, over the geometrical source-to-drain distance L for the fundamental plasmon mode can be explained in terms of equivalent circuit parameters (see the equivalent circuit in Fig. 1) by the effect of the planar intercontact source-to-drain capacitance $C_{\text{s-d}}$. (The equivalent circuit approach for describing the gated plasmons was developed in a recent paper in Ref. 11.) Of course, there are also the source-to-gate and gate-to-drain intercontact capacitances of the FET. However, those capacitances (not shown in the equivalent circuit in Fig. 1) are small in a FET with a short gate, so they do not affect the fundamental plasmon resonance and only slightly shift the second-order plasmon resonance down from its position in the device with no gate at all (the left bottom panel in Fig. 2). The characteristic capacitance (per unit width of the channel of the length L) involved in an ungated plasmon mode can be estimated as $C_{\text{pl}}^{(n)}=(\varepsilon_s+1)\varepsilon_0/k_nL$, where k_n is the plasmon wavevector estimated as $k_n=(2n-1)\pi/L$, with n being an integer. Multiplied by the kinetic electron inductance (per unit width of the channel of the length L), $L_{\text{el}}=m^*L/e^2N_s^{(un)}$, the characteristic plasmon mode capacitance yields the ungated plasmon frequency, $\omega_n=\sqrt{1/C_{\text{pl}}^{(n)}L_{\text{el}}}$, given by Eq. (2), which neglects the effect of the side contacts. For the fundamental plasmon mode ($k=\pi/L$), one obtains $C_{\text{pl}}\approx0.04$ fF/μm. The source-to-drain capacitance $C_{\text{s-d}}$ shunts the plasmon capacitance (see Fig. 1) with the total capacitance of their parallel connection $C_{\text{tot}}=C_{\text{s-d}}+C_{\text{pl}}$. This leads to decreasing the fundamental plasmon resonance frequency by a factor of $L_{\text{eff}}/L=4.5$, which yields $C_{\text{s-d}}\approx0.14$ fF/μm. This value of $C_{\text{s-d}}$ is in agreement with the electrostatic calculations[12–14] for the length of the side contacts of 40 μm each, which gives an estimated size of the side contact area supporting the image charge oscillations. (This length is close to the half of the terahertz radiation wavelength in the semiconductor.)

The relative values of the effective lengths (normalized to the geometrical length of the channel) remain essentially (within 10%) unchanged with a variation of the geometrical channel length L because such a short device (the channel length L being much shorter than the terahertz radiation wavelength) has to obey the electrostatic scaling. [The 10% drop in the effective channel length for the fundamental plasmon resonance takes place by varying the electron densities in the channel within the ranges$(1-5)\times10^{12}$ cm^{-2}.] For higher-order plasmon resonances, the effective length of the ungated channel approaches a geometrical value of the source-to-drain separation because short-wavelength higher-order plasmon modes induce the image charges only in a small area at the edges of the side contacts.

The electron channel resistance (per unit width of the channel of the length L) $R_{\text{el}}=m^*\nu L/e^2N_s^{(un)}$ contributes, along with the radiation resistance R_{rad}, to the plasmon resonance linewidth (see relevant arguments, e.g., in Refs. 15 and 16). If the source and drain contacts are longer than the terahertz wavelength (they are infinitely long in our electrodynamic model), the input radiation resistance $R_{\text{rad}}^{(\text{in})}$ is much larger than the FET impedance so that a fixed terahertz drain current induced by such a bow-tie-like antenna flows in the FET.[8] In this case, the output radiation resistance $R_{\text{rad}}^{(\text{out})}$ dominates in the plasmon resonance radiative linewidth as $\gamma_{\text{rad}}=1/R_{\text{rad}}^{(\text{out})}C_{\text{tot}}$ so that the total full width of the plasmon resonance line at its half maximum (FWHM) is $\Delta\omega=\nu+\gamma_{\text{rad}}$. Subtracting the dissipative contribution $\nu\approx0.07$ THz from the FWHM of the fundamental ungated plasmon resonance line in the FET with 200 nm gate, one can estimate $\gamma_{\text{rad}}\approx0.49$ THz yielding $R_{\text{rad}}^{(\text{out})}\approx1800$ Ω μm. It is remarkable that the radiative part of the plasmon resonance completely dominates in the total plasmon resonance linewidth of the ungated plasmons ($\gamma_{\text{rad}}\gg\nu$) due to the optical nature of this type of plasmons. (Note that the radiative resistance cannot be evaluated using a simpler quasistatic approach to solving the problem considered in this paper.)

When the length of the gate contact becomes comparable with the oscillating loop of a higher-order ungated (intercontact) plasmon mode (Fig. 3), the gate effectively screens this mode and, as a consequence, the frequency of the corresponding resonance goes down (see the left panel in Fig. 2). However, a short gate weakly screens the fundamental plasmon mode and, hence, its frequency only slightly decreases with an increase in the gate length (which corresponds to a slight increase in the source-to-gate and gate-to-drain capacitances). Obviously, the screening effect of a short gate on the

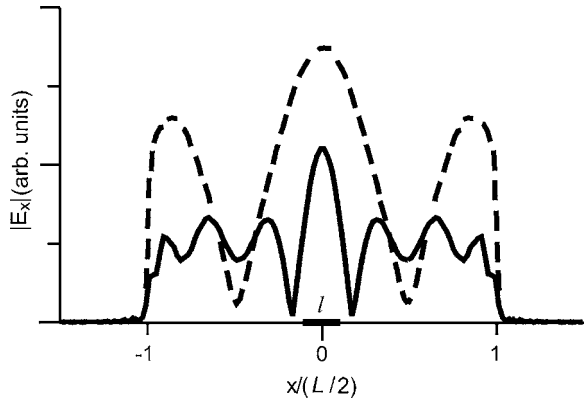

FIG. 3. Distributions of the in-plane component of the electric field in the second-order plasmon mode in the FET channel for the gate length $l=400$ nm (solid line) and in the same FET structure with no gate (dashed line). The gate area is shown by a short thick bar of length l located around $x=0$.

higher-order ungated plasmon mode should depend on the gate placement. In order to achieve an effective screening, a short gate should be placed above the antinode of the in-plane electric field of the ungated plasmon mode, e.g., as shown in Fig. 3. However, if a short gate was placed above the node of the in-plane electric field of the ungated plasmon mode, it would not affect the ungated plasmon mode frequency.

The frequency of the second-order (and higher-order) intercontact plasmon resonances can be effectively tuned like the frequency of the conventional gated plasmon resonance by applying the gate voltage that induces the electron density variations in the short gated region of the channel (see the right panel in Fig. 2). However, in distinction from the conventional gated plasmon resonance, where the resonant frequency is directly proportional to the square root of the gate-voltage swing, the frequency of the "gated intercontact resonance" exhibits saturation when the gate voltage approaches the threshold voltage. It is remarkable that the intensity of the gated intercontact resonance is much stronger (by two orders of magnitude) compared to that of a conventional gated plasmon resonance due to the optical nature of the ungated plasmons. (The gated plasmon resonances are not seen in the scale of Fig. 2 at their predicted positions marked by unfilled arrows in the left panel of Fig. 2.) The fundamental plasmon resonance is only slightly affected by the depletion of a short region of the channel under the gate contact. This fact can be also understood by considering the effective capacitance of the depletion region (shown by dotted lines in the equivalent circuit in Fig. 1), which is about $C_{\mathrm{dep}} \approx 0.23$ fF/μm for the fundamental plasmon resonance (estimated for $l=400$ nm employing the approach in Refs. 12–14). This capacitance is much greater than the characteristic plasmon capacitance, $C_{\mathrm{pl}} \approx 0.04$ fF/μm. Being connected in series with C_{pl}, capacitance C_{dep} effectively shunts the depletion region of the channel and, hence, the position and resonance line contour of the fundamental ungated plasmon resonance do not essentially change. [Due to a small gate-to-channel separation in the structure ($\sim$27 nm), gate-to-channel fringing capacitances, C_{fr}, at the edges of the depletion region play an important role and also shunt the depletion region at terahertz (plasmon) frequencies (the fringing capacitances are included in the gate-to-channel capacitance in the equivalent circuit in Fig. 1).] A slight decrease in the fundamental resonance frequency with the channel depletion under the gate may be explained by a small decrease in the electronic inductance with a decrease in the average electron density in the channel.

IV. CONCLUSIONS

We have shown that the frequencies of higher-order resonances of the ungated plasmon modes in the FET with a short gate can be effectively tuned by the gate-voltage variation. Such pseudogated plasmon resonances exhibit a strong tunability similar to the conventional gated plasmon resonances but can be excited by the external terahertz radiation with a much stronger intensity.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research and by the National Science Foundation under the auspices of the I/UCRC "Connection One." The work at IRE RAS was supported by the Russian Foundation for Basic Research through Grant No 06-02-16155 and by the Russian Academy of Sciences Program "Quantum Nanostructures."

[1] M. Dyakonov and M. Shur, IEEE Trans. Electron Devices **43**, 380 (1996).
[2] M. S. Shur and J.-Q. L. Lü, IEEE Trans. Microwave Theory Tech. **48**, 750 (2000).
[3] W. Knap, J. Lusakowski, T. Parenty, S. Bollaert, A. Cappy, V. V. Popov, and M. S. Shur, Appl. Phys. Lett. **84**, 2331 (2004).
[4] A. Satou, V. Ryzhii, I. Khmyrova, M. Ryzhii, and M. S. Shur, J. Appl. Phys. **95**, 2084 (2004).
[5] F. Teppe, D. Veksler, A. P. Dmitriev, V. Yu. Kachorovskii, S. Rumyantsev, W. Knap, and M. S. Shur, Appl. Phys. Lett. **87**, 052107 (2005).
[6] V. V. Popov, O. V. Polischuk, and M. S. Shur, J. Appl. Phys. **98**, 033510 (2005).
[7] V. Ryzhii, A. Satou, W. Knap, and M. S. Shur, J. Appl. Phys. **99**, 084507 (2006).
[8] V. V. Popov, G. M. Tsymbalov, M. S. Shur, and W. Knap, Semiconductors **39**, 142 (2005).
[9] F. Stern, Phys. Rev. Lett. **18**, 546 (1967).
[10] G. Korn and T. Korn, *Mathematical Handbook for Scientists and Engineers*, 2nd ed. (McGraw-Hill, New York, 1968).
[11] I. Khmyrova and Yu. Seijyou, Appl. Phys. Lett. **91**, 143515 (2007).
[12] J. S. Wei, IEEE J. Quantum Electron. **13**, 152 (1977).
[13] M. Ito and O. Wada, IEEE J. Quantum Electron. **22**, 1073 (1986).
[14] A. Koudymov, X. Hu, K. Simin, G. Simin, M. Ali, J. Yang, and M. Asif Khan, IEEE Electron Device Lett. **23**, 449 (2002).
[15] V. V. Popov, O. V. Polischuk, T. V. Teperik, X. G. Peralta, S. J. Allen, N. J. M. Horing, and M. C. Wanke, J. Appl. Phys. **94**, 3556 (2003).
[16] T. V. Teperik, V. V. Popov, and F. J. García de Abajo, Phys. Rev. B **71**, 085408 (2005).