Calculating drain delay in high electron mobility transistors

CrossMark

R. Coffie *

RLC Solutions, 6520 Tearose Dr., Plano, TX 75074, USA

A B S T R A C T

An expression for the signal delay (drain delay) associated with electrons traveling through the gate–drain depletion region has been obtained for nonuniform electron velocity. Due to the presence of the gate metal, the signal delay through the gate–drain depletion region was shown to be larger than the signal delay in the base–collector depletion region of a bipolar transistor when equal depletion lengths and velocity profiles were assumed. Drain delay is also shown to be larger in transistors with field plates (independent of field plate connection) compared to transistors without field plates when equal depletion lengths and velocity profiles were assumed. For the case of constant velocity, two expressions for the proportionality constant relating drain delay and electron transit time across the depletion were obtained.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

The signal delay, referred to as drain delay or τ_D, through the non-gradual channel region (gate–drain depletion region) of the intrinsic portion in field effect transistors (FETs) is well known to impact the short-circuit unity-current gain frequency f_T [1]. Drain delay can be a significant portion of the total delay in short gate length devices [2,3]. Drain delay also plays a dominant role in high voltage devices due to the extended gate–drain depletion regions required to support large drain voltages [4,5]. Proper calculation and understanding of drain delay is therefore important for designing high speed and high voltage FET amplifiers.

Similar to the collector signal delay in bipolar transistors [6], drain delay is associated with the phase delay in the drain transport factor (ζ) defined as the ratio of the total current collected by the drain to the conduction current entering the gate–drain depletion region. In [7], drain delay was properly treated for a constant electron velocity in the depletion region. This paper will present an expression for drain delay valid for nonuniform carrier velocities in the gate–drain depletion region. Analysis will show the electron velocity near the start of the gate–drain depletion region is weighted more than the velocity near the end of the depletion region (similar to bipolar transistors [8]), but drain delay is larger than collector delay for the same velocity profile and depletion length. Analysis will also show independent of the field plate connection (typically gate connected or source connected), FETs with field plates will have larger drain delay than FETs

without field plates if the same velocity profile and depletion length are assumed. To prevent potential future confusion, discrepancy between the results of this paper and results from [9] for drain delay are discussed.

2. Theory

Currents induced by moving charges in vacuum tubes were first analyzed by Shockley [10] and Ramo [11]. Their work is now referred to as the Shockley–Ramo Theorem and has been generalized for inhomogeneous linear media with an arbitrary fixed charge distribution under time-varying potentials [12] and to non-linear media [13]. The Shockley–Ramo Theorem states in a system of N conductors, a charge q_0 located at position $\vec{r}_0$ traveling with velocity $\vec{v}$ will induce a current in the ith conductor of

$$I_i = q_0 \frac{d\Phi_i[\vec{r}_0]}{d\vec{r}_0} \cdot \frac{d\vec{r}_0}{dt} = q_0 \vec{\nabla} \Phi_i \cdot \vec{v} \tag{1}$$

where Φ_i is often referred to as the normalized electrostatic potential of the system but should not be confused with the true electrostatic potential of the system under normal operating conditions. A general definition for Φ_i valid for non-linear and linear media is [13]

$$\Phi_i = \left. \frac{\partial \phi_{el}}{\partial V_i} \right|_{V_{op}} \tag{2}$$

where V_{op} is the operating bias point, V_i is the potential of the ith conductor and ϕ_{el} is the electrostatic potential produced with all the electrodes being biased at their operating potentials, the fixed

* Tel.: +1 972 767 9388.
E-mail address: rcoffie@rlcsolutions.com

space charge and mobile charge of the system removed, but with the material's dielectric properties left in place.

Current induced in the drain due to electrons traveling in the gate–drain depletion region is a special case of the Shockley–Ramo Theorem [7]. Before calculating the induced drain current, several simplifying assumptions similar to [8] will be made: (1) electron confinement in the channel is maintained throughout the gate–drain depletion region; (2) the variation of the gate–drain depletion region boundaries due to carrier modulation is small compared to the depletion region length allowing the DC values to be used; (3) the electron density entering the depletion region is density modulated, but not velocity modulated (the local velocity is always $v[x]$); and (4) all electrons entering the depletion region have the same velocity $v[0]$.

As shown in Fig. 1, the channel of the transistor is taken along the x-axis, with the start and end of the gate–drain depletion located at $x = 0$ and $x = x_m$, respectively. With the above assumptions, the AC conduction current entering the depletion region is

$$i_{ch}[t] = i_{ch0} \exp[j\omega t] = q n_s[0] W_g v[0] \exp[j\omega t] \tag{3}$$

where q is the elementary charge, $n_s[0]$ is the sheet density of electrons at the start of the gate–drain depletion region, W_g is the gate width, and ω is the angular frequency of modulation. Between t and $t + dt$, we have mobile charge q_{inj} injected into the depletion region of

$$q_{inj}[0, t] = q n_s[0] W_g v[0] \exp[j\omega t] dt \tag{4}$$

At a time $(t + t')$ later where

$$t' = \int_0^x (dx'/v[x']) \tag{5}$$

this same charge is now located at x. Therefore, the injected charge in the depletion region has the form of a traveling wave in the $+x$ direction given by

$$q_{inj}[x, t] = q n_s[0] W_g v[0] \exp[j\omega(t - t')] dt \tag{6}$$

The drain transport factor is obtained by substituting (6) into (1) for q_0, dividing by $i_{ch}[t]$, and then integrating over all the injected mobile charge in the depletion region

$$
\begin{aligned}
\zeta &= \int_0^{x_m} \exp[-j\omega t'] \frac{\partial \Phi_D[x, 0]}{\partial x} dx \\
\zeta &= \int_0^{\tau_1} v[x[t']] \exp[-j\omega t'] \frac{\partial \Phi_D[x, 0]}{\partial x} dt'
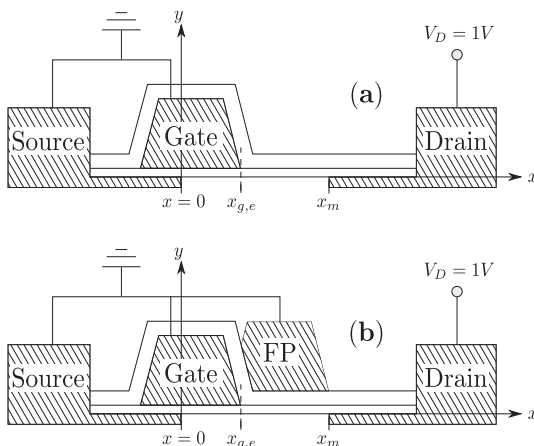\end{aligned}
\tag{7}
$$

Fig. 1. Configurations used to simulate Φ_D. The gate–drain depletion region starts at $x = 0$ and ends at $x = x_m$. The end of the gate is located at $x = x_{g,e}$. (a) FET without field plates. (b) FET with a field plate.

where $\tau_1 = \int_0^{x_m} (dx/v[x])$ is the electron transit time across the depletion region. In (7), Φ_D is the normalized electrostatic potential along the channel calculated as outlined in Fig. 1. For this calculation the media is assumed linear. Φ_D is determined by removing all mobile and depletion charges from the system, leaving the material's dielectric properties in place, grounding the gate and field plate (if present), the source and channel regions are approximated by a grounded metal up to the start of the gate–drain depletion region and the channel from the end of the gate–drain depletion region to the drain contact is set to unit potential. Although the bias point used in calculating Φ_D is different from the operating bias point, the same result is obtained since the media is assumed linear [13]. Determining Φ_D is easily done with any 2D electrostatic simulator. The freely available electrostatic simulator FEMM 4.2 [14] is used for calculating Φ_D in this paper. The drain delay is then obtained as

$$\tau_D = -\text{phase}[\zeta]/\omega \tag{8}$$

Comparison to collector delay in bipolar transistors is more easily done by considering angular frequencies $\omega \ll \tau_1^{-1}$. The low-frequency limit is obtained by expanding (7) to first order in $\omega t'$

$$
\begin{aligned}
\zeta_{LF} &= \int_0^{x_m} (1 - j\omega t') \frac{\partial \Phi_D}{\partial x} dx \\
\zeta_{LF} &= 1 - j\omega \int_0^{x_m} (1 - \Phi_D)(v[x])^{-1} dx
\end{aligned}
\tag{9}
$$

The second form of (9) is obtained through integration by parts. From (9), the low-frequency drain delay is

$$\tau_D = \int_0^{x_m} (1 - \Phi_D)(v[x])^{-1} dx \tag{10}$$

As a check, evaluation of Φ_D for a 1D device (such as the bipolar transistor) can be done analytically. In the 1D case, Φ_D is simply the electrostatic potential between two capacitor plates with unit voltage between plates ($\Phi_D = x/x_m$) and (10) reduces to the well-known collector delay expression given in [8] for bipolar transistors. Similar to the low frequency collector delay in bipolar transistors, the drain delay has a weighting function that is unity at the start of the gate–drain depletion region ($\Phi_D \approx 0$ near $x = 0$) and becomes zero at the end of the depletion region ($\Phi_D \approx 1$ near $x = x_m$). For all cases investigated, the FET weighting function is a monotonic decreasing function from one to zero (see Fig. 3). Therefore, the drain delay is more sensitive to velocities near $x = 0$ compared to velocities near $x = x_m$. Unlike the linear bipolar weighting function $(1 - x/x_m)$, the FET weighting function is nonlinear.

Incorporating current induced velocity modulation in the low-frequency drain delay expression requires determining the change in velocity with change in injected current. The current dependent velocity expressed as a first order Taylor series is

$$\frac{dx}{dt} = v_x[x, i_{ch0} + di_{ch0}] = v_x[x, i_{ch0}] + \frac{dv_x[x, i_{ch0}]}{di_{ch0}} di_{ch0} \tag{11}$$

Inserting (11) into (10) for $v[x]$ and using $(1 + y)^{-1} \approx (1 - y)$ for $y \ll 1$ results in

$$\tau'_D = \int_0^{x_m} (1 - \Phi_D)\left(1 - \frac{dv_x[x, i_{ch0}]}{di_{ch0}} \frac{di_{ch0}}{v_x[x, i_{ch0}]}\right) \frac{dx}{v_x[x, i_{ch0}]} \tag{12}$$

where τ'_D is the drain delay including velocity modulation. In 1D, (12) is consistent with the collector delay expression in the presence of velocity modulation for bipolar transistors derived in [15].

3. Results and discussion

FET weighting functions are calculated for a FET without a field plate and a FET with a field plate (see Fig. 1). Short (60 nm) and