Graphene Base Transistors: A Simulation Study of DC and Small-Signal Operation

Valerio Di Lecce, Roberto Grassi, Antonio Gnudi, *Member, IEEE*, Elena Gnani, *Member, IEEE*, Susanna Reggiani, *Member, IEEE*, and Giorgio Baccarani, *Fellow, IEEE*

Abstract—A simulation study aimed at investigating the main features in dc and small-signal operating conditions of the hot-electron graphene base transistor (GBT) for analog terahertz operation is presented. Intrinsic silicon is used as reference material. The numerical model is based on a self-consistent Schrödinger–Poisson solution, using a 1-D transport approximation and accounting for multiple-valley and nonparabolicity band effects. Some limitations in the extension of the saturation region and in the output conductance related to the finite quantum capacitance of graphene and to space charge effects are discussed. A small-signal model is developed that catches the essential physics behind the voltage gain and the cutoff frequency, which shows that the graphene quantum capacitance severely limits the former but not the latter. According to simulations carried out within the ballistic transport approximation, a 20-nm-long GBT can achieve at the same time a voltage gain larger than 10 and a cutoff frequency largely above 1 THz within a reasonably wide bias range.

Index Terms—Graphene, graphene base transistor (GBT), hot-electron transistor (HET), terahertz operation.

I. INTRODUCTION

IN THE last decade, graphene has gained the scientific community's attention as a potential candidate for improving the performance of electronic devices operating at RF frequencies. Its 2-D nature [1] with high-speed massless-like carriers, group velocity around 10^8 cm/s [2]–[4], and long enough carrier mean-free path at room temperature to ensure quasi-ballistic transport at sizes compatible with CMOS devices, are indeed quite attractive. In particular, after the first demonstration of graphene field effect transistors (GFETs) [4], remarkable RF performance has been achieved in terms of cutoff frequency (in the range of hundreds of gigahertz [5], [6]), and various RF applications have been investigated [7], [8]. The absence of a bandgap is, however, responsible for GFETs suffering from the lack of drain current saturation, ultimately limiting their potential use in RF circuits.

Recently, a novel approach to exploit the properties of graphene in RF devices was proposed in [9], leading to the graphene base transistor (GBT) concept. In this new perspective, graphene is not used as a semiconductor channel material with current flowing on its plane, but rather as a highly-conductive monolayer semimetallic film forming the transistor base, almost transparent to the current flowing in the normal direction. In [9], the GBT is reported to be capable of addressing the issues left unsolved by the GFET, i.e., high on–off current ratio, drain-current saturation, and power amplification.

In this paper, we investigate the behavior of the GBT by means of numerical simulations based on a full-quantum transport model coupled with Poisson's equation, going beyond the analysis carried out in [9] for several aspects. Namely, space charge effects due to traveling electrons as well as band structure effects in the semiconducting layers, such as multiple ellipsoidal valleys including nonparabolicity corrections, are fully accounted for. Some peculiar features of the device characteristics are highlighted and explained. Moreover, a small-signal model is presented, which catches the essential device physics in the active region and clarifies the dependence of the cutoff frequency f_T and of the intrinsic voltage gain A_{v0} on the device parameters. From the above analysis, optimization guidelines for terahertz operation are suggested.

This paper is organized as follows. In Section II, details on the device structure and on the model used in the simulations are provided. In Section III, the device $I-V$ characteristics are presented and the peculiar features are illustrated. In Section IV, the small-signal model including capacitive effects, valid in the saturation region, is discussed. Finally, conclusions are drawn in Section V.

II. DEVICE STRUCTURE AND NUMERICAL MODEL

A 1-D cross-section of the intrinsic part of the GBT is shown in Fig. 1 along with a qualitative band diagram in the active operating region. The device is obtained by stacking a metallic region (emitter), a semiconducting or insulating layer [emitter–base insulator (EBI)], a graphene monolayer (base), another semiconducting or insulating layer [base-collector insulator (BCI)], and a second metallic region (collector). The graphene layer is supposed to be contacted as well. The device behavior is similar to a conventional n-p-n bipolar junction transistor: in the active operating region, electrons are injected into the device from the emitter, a fraction of them reaches the collector, while a—hopefully much smaller—fraction is captured by the base or backscattered. The EBI and BCI regions act as potential barriers for the electrons: current flows by Fowler–Nordheim tunneling through

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

2   IEEE TRANSACTIONS ON ELECTRON DEVICES
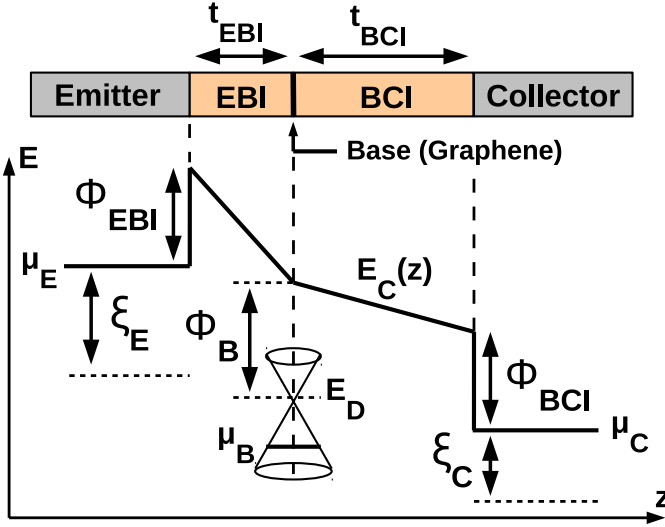

Fig. 1.   1-D cross-section of the GBT and its conduction band diagram $E_C(z)$ in the active operating region: $\mu_{E/B/C}$ is the emitter/base/collector Fermi level and E_D is the graphene Dirac energy. The other parameters are defined in Section II.

the EBI barrier, the shape of which is modulated by the V_{BE} voltage, resulting in the transistor effect. As for the EBI and BCI material, a semiconductor typically provides a lower energy barrier Φ_{EBI} (Fig. 1) and consequently larger operating currents. Besides, it allows for a more efficient carrier transport, because of its crystalline nature and higher mobility.

An approximate 1-D device model is adopted: uniformity is assumed in the x and y directions, so that the transverse wave vector $\vec{k}_t \equiv (k_x, k_y)$ is conserved, while transport occurs in the z direction. Transport through the valence band of the semiconducting layers is ignored. This approximation is appropriate when: 1) the largest applied voltages do not or only slightly exceed the bandgap voltage E_G/q, such that electron-hole pair generation by band-to-band tunneling and impact ionization are unimportant and 2) the value of Φ_{EBI} is smaller than $E_G/2$, so that emission of electrons is favored over emission of holes.

Electron transport is solved within the ballistic nonequilibrium Green's function (NEGF) formalism [10] using an effective mass (EM) Hamiltonian. The neglect of any type of scattering mechanism certainly leads to an overestimation of the device performance, but can be roughly justified for the nearly deca-nanometer devices considered in this paper, providing moreover a sort of performance upper limit.

To correctly estimate the group velocity of high energy electrons, in particular of those in the BCI region near the collector, nonparabolic corrections to the EM model are introduced, adopting the following energy-dispersion relation [11] for a conduction-band valley having the principal axes aligned with the reference axes as follows:

$$\epsilon(1 + \alpha\epsilon) = \frac{\hbar^2}{2} \left(\frac{k_x^2}{m_x} + \frac{k_y^2}{m_y} + \frac{k_z^2}{m_z} \right) \tag{1}$$

where $\epsilon = E - E_C$ is the electron kinetic energy, E_C is the conduction band edge, E is the total electron energy, α is the nonparabolicity factor, and $\hbar$ is the reduced Planck constant. Having defined the normalized transverse wavevector

$$\vec{\zeta}_t \equiv \left(\frac{\hbar k_x}{\sqrt{2m_x}}, \frac{\hbar k_y}{\sqrt{2m_y}} \right) \tag{2}$$

the following expression for the 1-D energy-dispersion relation for a given $\vec{\zeta}_t$ (subband) is obtained

$$\epsilon_z(1 + \alpha_{\zeta_t}\epsilon_z) = \frac{\hbar^2 k_z^2}{2m_{\zeta_t}} \tag{3}$$

where ϵ_z is the longitudinal kinetic energy, i.e., $\epsilon_z = \epsilon - \epsilon_t$, $\epsilon_t \equiv \epsilon_t(\zeta_t)$ is the energy of the sub-band minimum measured from E_C given by the following:

$$\epsilon_t(\zeta_t) = -\frac{1}{2\alpha} + \sqrt{\left(\frac{1}{2\alpha}\right)^2 + \frac{\zeta_t^2}{\alpha}} \tag{4}$$

$\zeta_t = ||\vec{\zeta}_t||$, and

$$\alpha_{\zeta_t} = \frac{\alpha}{1 + 2\alpha\epsilon_t}, \quad m_{\zeta_t} = m_z (1 + 2\alpha\epsilon_t). \tag{5}$$

Equation (3) can be used to define an energy-dependent EM $\widetilde{m}_{\zeta_t} \equiv \widetilde{m}_{\zeta_t}(\epsilon_z)$

$$\widetilde{m}_{\zeta_t}(\epsilon_z) = m_{\zeta_t}(1 + \alpha_{\zeta_t}\epsilon_z) \tag{6}$$

such that

$$\epsilon_z = \frac{\hbar^2 k_z^2}{2\widetilde{m}_{\zeta_t}}. \tag{7}$$

The expression (6) is used for $\epsilon_z > 0$. To avoid an unrealistic description of the dispersion relation inside the energy gap, the parabolic expression $\widetilde{m}_{\zeta_t} = m_{\zeta_t}$ is used for $\epsilon_z < 0$. From the above considerations, the nonparabolic 1-D Hamiltonian for electrons in the EBI and BCI regions at a given $\vec{\zeta}_t$ in the considered valley is

$$\hat{H}_{\vec{\zeta}_t} = E_C + \epsilon_t(\zeta_t) - q\varphi(z) - \frac{\hbar^2}{2} \frac{\partial}{\partial z} \left(\frac{1}{\widetilde{m}_{\zeta_t}(z, E)} \frac{\partial}{\partial z} \right) \tag{8}$$

where q is the electron charge, $\varphi(z)$ is the electrostatic potential, and the EM $\widetilde{m}_{\zeta_t}$ is dependent through ϵ_z on both the longitudinal position z and the total energy E.

In this paper, silicon is assumed as the material for the EBI and BCI regions, with the z-axis oriented along the $\langle 100 \rangle$ direction. Such orientation gives rise to two families of sub-bands (ladders) with different EMs and degeneracy factors g_v: the unprimed ladder with $m_x = m_z = m_t$, $m_y = m_l$ or $m_y = m_z = m_t$, $m_x = m_l$, and $g_v = 4$; the primed ladder with $m_x = m_y = m_t$, $m_z = m_l$, and $g_v = 2$. The longitudinal and transverse masses are $m_l = 0.92\,m_0$ and $m_t = 0.19\,m_0$, respectively, with m_0 the free electron mass. Moreover, $\alpha = 0.5$ eV^{-1} [12].

The Hamiltonian (8) is inserted into the NEGF formalism in order to calculate the charge and the current (the latter through the Landauer formula) for the sub-band corresponding to the given $\vec{\zeta}_t$ and ladder. The total charge and current are computed by means of numerical integrations in $\vec{\zeta}_t$, eventually

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

DI LECCE *et al.*: GRAPHENE BASE TRANSISTORS

3

adding the results for the two ladders and accounting for the degeneracy factors. Exploiting the cylindrical symmetry of the Hamiltonian (8) with respect to ζ_t, the integrals are conveniently calculated in polar coordinates using a Gaussian quadrature scheme in ζ_t with typically 20 grid points.

It could be objected that, due to the current injection being controlled by the electron dynamics within an extremely thin layer next to the emitter-EBI interface, an atomistic Hamiltonian should be more appropriate. However, consistently with the scope of this paper, aiming at investigating the main device features, the simpler and faster EM model (with nonparabolic corrections) is preferred here, which is anyhow known to provide good results even when dealing with phenomena occurring within extremely thin material regions [13]–[15].

The simulation domain is composed of EBI, base, and BCI regions. Open-boundary conditions are imposed at the emitter and collector contacts by defining appropriate self-energies. In particular, we assume that the emitter/collector contact can be described by a Hamiltonian similar to (8) with flat $\varphi(z)$ and with E_C replaced with $\mu_{E/C} - \xi_{E/C}$, where μ_E and $\mu_C = \mu_E - qV_{CE}$ are the emitter and collector Fermi levels, respectively, and $\xi_{E/C}$ is a parameter that describes the degeneracy of the emitter/collector region (Fig. 1). Values $\xi_E = 0.8$ eV and $\xi_C = 0.06$ eV are used (Section III). In addition, for simplicity, the emitter/collector Hamiltonian is assumed strictly parabolic, with constant EMs equal to the values of bulk silicon for the two ladders. The closed-form expression of the self-energy in [10, Ch. 8], valid for a semi-infinite lead, is used.

Regarding the appropriate treatment of the graphene layer, the issue is far from being straightforward. The in-plane transport properties of graphene have been intensively investigated both theoretically and experimentally. However, the investigation of transport for electrons impinging onto the graphene layer in the normal direction is still in an early stage: studies can be found stating that graphene acts like a barrier [9], together with other studies where it is assimilated to a potential well [16]. In this paper, graphene is treated in the following manner. As far as transport is concerned, merely an imaginary retarded self-energy $\Sigma_B^R = -j\Delta_B\delta(z - z_B)$ independent of E is added to the semiconductor Hamiltonian (8) at the single grid point $z = z_B$ corresponding to the location of the graphene sheet, with $\delta(z)$ the Dirac delta. Such self-energy phenomenologically mimics the presence of a base contact region, through which electrons can flow in and out, thus allowing for a nonzero base current. The base contact is assumed at thermodynamic equilibrium with Fermi level $\mu_B = \mu_E - qV_{BE}$. The value of Δ_B, which represents the coupling constant between the contact and the rest of the device, is considered as a fitting parameter, to be adjusted on the basis of future detailed transport studies of electrons across the graphene layer, or more simply on measured base currents. It should be noticed that setting Δ_B equal to zero is equivalent to ignoring the base current. In this paper, the value $\Delta_B = 10^{-13}$ eV·cm has been used, which provides $\beta_F = I_C/I_B \simeq 10^5$ in the active region.

In addition, the effect of graphene on the device electrostatics is fully considered. The net electron density within the device is calculated as follows:

$$n(z) = n_{SC}(z) + \delta(z - z_B)n_{GR} \qquad (9)$$

where $n_{SC}(z)$ is the contribution arising from the semiconductor and calculated with the NEGF formalism in the standard way, while n_{GR} is the net electron sheet density on the graphene layer. The latter is computed from the graphene Dirac band model, assuming an equilibrium distribution with Fermi level $\mu_B = \mu_E - qV_{BE}$, as follows:

$$n_{GR} = \frac{2(k_B T)^2}{\pi(\hbar v_F)^2}\big[F_1(\eta) - F_1(-\eta)\big] \qquad (10)$$

where k_B is Boltzmann's constant, $T = 300$ K is the temperature, $v_F = 10^8$ cm/s is the graphene Fermi velocity, F_1 is the Fermi–Dirac integral of order one, and $\eta = [\mu_B - E_D + q\varphi(z_B)]/(k_B T)$ the normalized difference between Fermi level and Dirac energy, $E_D = E_C - \Phi_B$ being the Dirac energy for $\varphi(z_B) = 0$ and Φ_B a constant, which depends on the EBI/BCI semiconductor (Fig. 1). For silicon $\Phi_B = 0.5$ eV, calculated as the difference between the graphene workfunction and the silicon electron affinity. It should be noticed that n_{GR} is positive when $\eta > 0$, corresponding to an effective electron density in the graphene sheet, n_{GR} is negative when $\eta < 0$, corresponding to an effective hole density.

The NEGF equations together with (10) are self-consistently solved with the 1-D Poisson equation, which provides the electrostatic potential $\varphi(z)$. Dirichlet boundary conditions for the potential are imposed at the emitter and collector contacts, corresponding to the barrier heights Φ_{EBI} and Φ_{BCI} (Fig. 1). Schottky barrier lowering due to image forces is neglected.

III. DEVICE DC OPERATION

In this section, making use of the model described above, the fundamental aspects of the device behavior in dc are investigated. Two reference devices have been simulated as case studies, both having the EBI/BCI regions made of undoped silicon with $t_{EBI} = 3$ nm: in the first one, hereafter called GBT1, $t_{BCI} = 20$ nm, in the second, called GBT2, $t_{BCI} = 10$ nm. The maximum total device length (23 nm for GBT1) should roughly ensure the validity of the ballistic transport assumption. For both devices, $\Phi_{EBI} = \Phi_{BCI} = 0.2$ eV. The choice of the barrier height and of t_{EBI} has been driven by the consideration that a relatively low Φ_{EBI} and a thin EBI layer are key factors to increase the current injection and hence the device speed [9], [17]. As far as the barrier height is concerned, the search for materials providing a low Schottky barrier with silicon has led to the exploration of a wide range of solutions: titanium is one of the metals with the lowest Schottky barrier for electrons on n-type silicon (0.5 eV) [18], while rare-earth silicides allow for barriers lower than 0.4 eV [19], [20]. Considering that barrier lowering is not explicitly considered in our model, the choice made here of $\Phi_{EBI} = 0.2$ eV appears to be reasonable. Concerning the EBI layer thickness, the chosen value $t_{EBI} = 3$ nm is low but still within reach of the present fabrication capabilities.
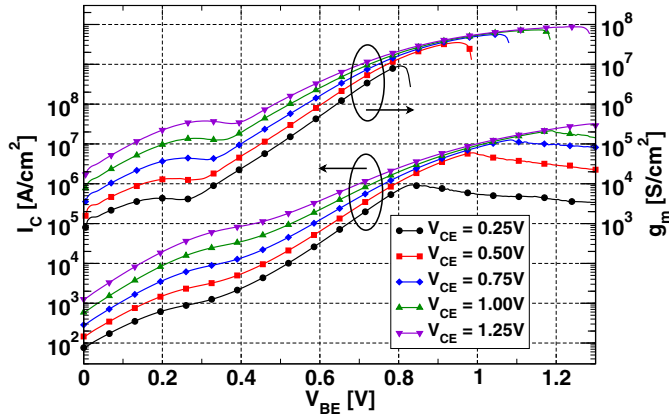
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

4

IEEE TRANSACTIONS ON ELECTRON DEVICES

Fig. 2. Turn-on characteristics (left axis) and transconductance (right axis) of the device GBT1 defined in the text for different V_{CE} voltages (color online).
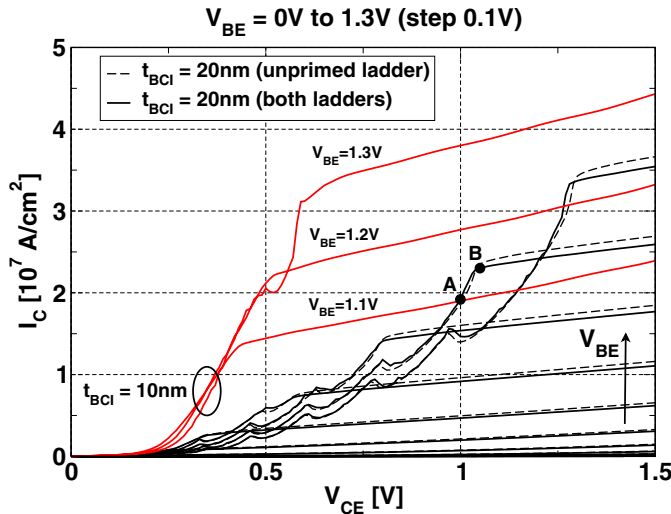

Fig. 3. Output current characteristics of the device GBT1 defined in the text with $t_{BCI} = 20$ nm calculated with both ladders (solid black lines) and with the unprimed ladder only (dashed black lines) for V_{BE} ranging from 0 to 1.3 V in steps of 0.1 V. Red lines: output characteristics of the device GBT2 with $t_{BCI} = 10$ nm calculated with both ladders for $V_{BE} = 1.1$, 1.2, and 1.3 V. The negative differential resistance in the unsaturated region (low V_{CE}) is most likely related to quantum resonance effects in the potential well (Fig. 5). The phenomenon is still under investigation (color online).
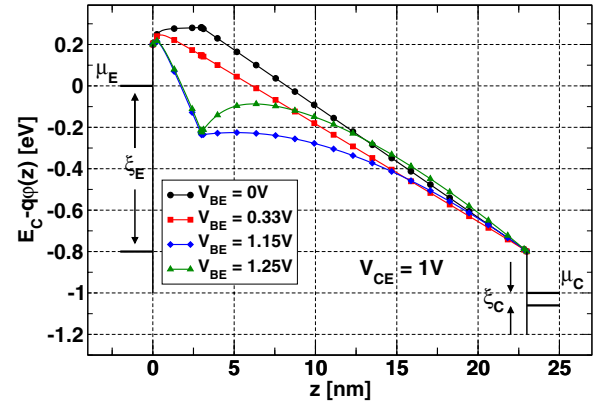

Fig. 4. Conduction-band diagram of the device GBT1 at $V_{CE} = 1$ V and different V_{BE} values. The graphene layer is located at $z = z_B = 3$ nm. The emitter/collector degeneracy parameters ξ_E and ξ_C are also shown (color online).
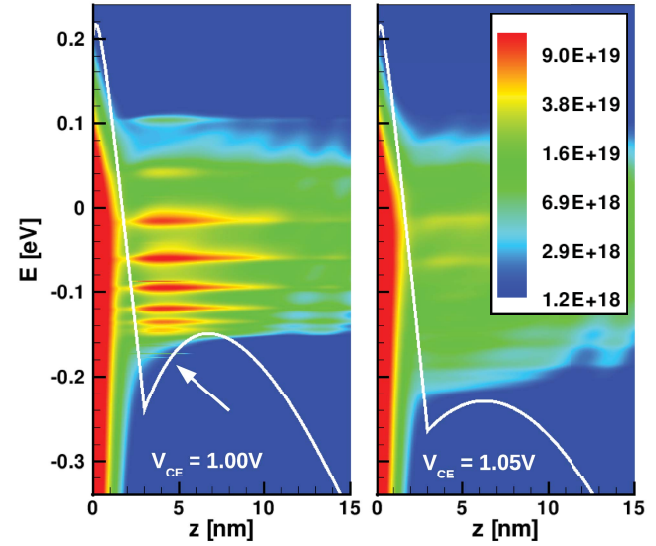

Fig. 5. Conduction-band diagrams (white lines) and electron density spectra in $eV^{-1}cm^{-3}$ (colored plots) of the device GBT1 for the two bias points marked as A (left part: $V_{BE} = 1.2$ V, $V_{CE} = 1$ V) and B (right part: $V_{BE} = 1.2$ V, $V_{CE} = 1.05$ V) in Fig. 3. The arrow points at the lowest quasi-bound state within the potential well. The high density near $z = 0$ is due to the fast-decaying electron states injected from the emitter into the silicon gap (color online).

The simulated turn-on and output characteristics of the GBT1 device are reported in Figs. 2 and 3, respectively. The maximum applied voltages are $V_{BE,max} = 1.3$ V and $V_{CE,max} = 1.5$ V, corresponding approximately to the limits within which impact ionization and valence band transport can be neglected, as discussed in Section II. It should be noted that such limitations on the operating voltages make the device inadequate for use as a power amplifier: the design aimed at here is oriented to analog high-frequency rather than high-power operation.

From the two figures, distinct operating regions emerge. For approximately $V_{BE} < 0.8$ V, the current behaves roughly exponentially with respect to both V_{BE} and V_{CE} (Fig. 2). To gain physical insight, Fig. 4 shows the conduction-band diagram at $V_{CE} = 1$ V for different V_{BE} values. The EBI tunneling barrier gets thinner and thinner with V_{BE} increasing above 0 V.

The discontinuity of the conduction-band slope occurring at $z = z_B = 3$ nm for $V_{BE} = 0$ V (Fig. 4, black line) is due to the graphene sheet charge, which turns out to be negative (i.e., $n_{GR} > 0$) with $\mu_B > E_D - q\varphi(z_B)$. At $V_{BE} = 0.33$ V (Fig. 4, red line), the discontinuity disappears, corresponding to $n_{GR} = 0$. At this bias point, $\mu_B = E_D - q\varphi(z_B)$ and the graphene quantum capacitance reaches its minimum: a change in V_{BE} leads to only a small change of $\varphi(z_B)$, and this in turn leads to an inflection point in the turn-on curve, which is clearly visible around $V_{BE} = 0.3$ V. Accordingly, the device transconductance (also reported in Fig. 2) is subject to a small drop around this bias point. For $V_{BE} > 0.3$ V, the graphene charge becomes positive and the transconductance starts increasing again.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

DI LECCE *et al.*: GRAPHENE BASE TRANSISTORS

5

For approximately $V_{BE} > 0.8$ V, two different behaviors can be distinguished from Figs. 2 and 3 for any given V_{BE} value: an unsaturated region at low V_{CE}, and a saturation region at higher V_{CE}, where the current is almost independent of V_{CE}. Besides, for a given V_{CE}, the current first increases with V_{BE} and then decreases. The transition between unsaturated and saturated behavior is shown better in Fig. 5, which shows the conduction-band diagrams and the electron spectral densities relative to the two bias points highlighted with the marks A and B in Fig. 3. At $V_{CE} = 1$ V (left plot) the device is still in the unsaturated condition. This regime is characterized by the presence of a potential well near $z = z_B$, due to the large positive charge present on the graphene layer, where the Fermi level $\mu_B = -1.2$ eV is far below the Dirac point energy $E_D - q\varphi(z_B)$. The current is limited also by the potential barrier, which is formed within the BCI layer close to the graphene and is sustained by the negative space charge due to the traveling electrons. The height of such barrier is controlled by both the base and the collector voltages, thus explaining the V_{CE} dependence of the current. Within the potential well, quasi-bound electron states are formed, which in Fig. 5-left are clearly identified by the red density peaks corresponding to the contributions given by the simulated sub-bands (ζ_t values) and used for the integration process in the transverse momentum space. Such quasi-bound electron charge induces a large positive charge on the graphene layer. Fig. 5-left corresponds almost to the edge of the unsaturated region: a small increase of V_{CE} (Fig. 5-right) reduces the BCI barrier height and suppresses the quasi-bound states in the well. As a consequence, the positive graphene charge decreases, the Dirac point energy decreases as well, and the barrier height is further lowered. At $V_{CE} = 1.05$ V the device is already in the saturated regime: the current is almost independent of V_{CE}, since the large majority of the electrons are injected through the EBI tunneling barrier at energies above the BCI barrier, as can be seen from the electron density in Fig. 5-right.

The already mentioned nonmonotonic current behavior with V_{BE} for a given V_{CE} (Fig. 2) can be explained in a similar way, through the presence of a BCI barrier. Fig. 4 reports the band diagrams for $V_{BE} = 1.15$ V (blue curve) and $V_{BE} = 1.25$ V (green curve) for the same $V_{CE} = 1$ V. It is seen that in this operating condition the increase of V_{BE} does not cause a decrease of the Dirac energy, but rather a significant increment of the positive graphene charge, which in turn develops the BCI barrier, which eventually leads to a current decrease.

The considerations above explain the limited extension of the saturation region. An additional concern is the origin of the slope of the saturation current with V_{CE} (output conductance). This can be explained observing that, even if in saturation the BCI barrier is low or absent, the collector voltage still indirectly affects the current through the electrostatic influence on the graphene charge and, hence, the Dirac potential, due to the limited quantum capacitance of graphene. For comparison, Fig. 3 shows the output characteristics also of the GBT2 device ($t_{BCI} = 10$ nm). The effect of the smaller BCI layer thickness is twofold. First, a significant improvement is observed concerning the extension of the saturation region:

e.g., at $V_{BE} = 1.2$ V, a V_{CE} slightly above 0.5 V is sufficient to suppress the BCI barrier (onset of saturation), as opposed to the $V_{CE} \simeq 1$ V of the GBT1 device. On the other hand, the output conductance is clearly degraded, representing a possible limitation for the maximum achievable voltage gain. This problem will be further investigated in the next section.

To understand the relative importance of the two sub-band ladders on the total current, the complete GBT1 characteristics (both ladders) are compared with those computed with only the unprimed ladder in Fig. 3. Understandably, the primed ladder only slightly modifies the current, since it is characterized by the transport EM equal to the large longitudinal EM, which considerably limits the tunneling current. However, it is interesting to observe that its effect is to decrease, rather than to increase, the current. This can be explained with the electrostatic feedback: the slowly moving primed electrons give a contribution to the total electron charge, which acts in the sense of slightly increasing the EBI tunneling barrier, thus reducing the injected unprimed current.

A final remark concerning the choice of the parameters ξ_E and ξ_C. As can be seen from Fig. 4, the value of $\xi_E = 0.8$ eV guarantees a sufficiently large energy range for carrier injection, including the energies in the potential well near $z = z_B$, and it is therefore appropriate to mimic a metal contact. On the other hand, injection from the collector is negligible, therefore the chosen value $\xi_C = 0.06$ eV has little impact on the results and allows to save energy grid points.

IV. SMALL-SIGNAL MODEL

In this section, a small-signal model is developed, which is valid in the saturation regime and includes capacitive effects. The following notation will be used to identify the charges per unit area in the device: Q_{GR} indicates the charge on the graphene layer, given by $Q_{GR} = -qn_{GR}$, with n_{GR} defined in (10); Q_N is the total transport electron charge, which results from the integration of the electron density in the EBI and BCI semiconductor layers given by the NEGF procedure; Q_E and Q_C are the sheet charges at the emitter and collector metal contacts, respectively, obtained by applying Gauss theorem at the contact interfaces and assuming zero electric field within the external contact regions.

Instead of considering the terminal currents and charges as functions of the applied voltages V_{BE} and V_{CE}, it is convenient to work with the voltages $V_{DE} = V_{DB} + V_{BE}$ and $V_{DC} = V_{DB} + V_{BC}$, where $V_{DB} = \eta k_B T/q$, η being defined in (10), is related to the energy difference between the Dirac point and the Fermi level in the graphene layer. From (10) it follows that $Q_{GR} \equiv Q_{GR}(V_{DB}) \equiv Q_{GR}(V_{DE} - V_{BE})$. Besides, according to the analysis of the previous section and considering that the electron injection in saturation regime is controlled essentially by the potential within the EBI layer, it is reasonable to assume that $I_C \equiv I_C(V_{DE})$ and $I_B \equiv I_B(V_{DE})$. The collector voltage affects the currents through its influence on V_{DE}. This is confirmed by Fig. 6, where I_C and I_B are drawn versus V_{DE} for two different values of the collector voltage within the saturation region. It is seen that the approximation works well in particular for large currents. The picture is
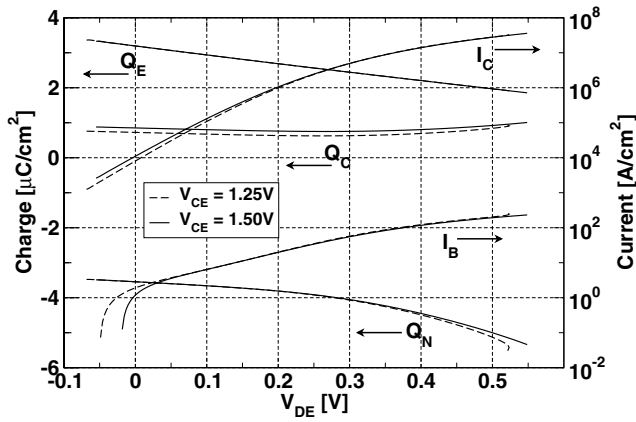
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

6

IEEE TRANSACTIONS ON ELECTRON DEVICES

Fig. 6. Left vertical axis: charge densities per unit device area—emitter interface charge *[etc.]* Q_E, collector interface charge Q_C, electron density integrated on the 1-D domain Q_N. Right vertical axis: collector and base currents. Horizontal axis: potential difference V_{DE} between graphene Dirac energy and emitter Fermi level. All curves are drawn for $V_{CE} = 1.25$ V and $V_{CE} = 1.5$ V.

more complex for the charges Q_N, Q_E, and Q_C, which are also shown in Fig. 6. While $Q_E \equiv Q_E(V_{DE})$ is very well satisfied, a small dependence on V_{CE} is clearly visible for Q_N and Q_C, for the former mainly at high voltages. This can be simply explained considering a more intuitive semi-classical (and again ballistic) transport picture within the BCI layer, as follows. Since, as stated above, the electron current density is controlled by the graphene Dirac potential [i.e., $I_C \equiv I_C(V_{DE})$], a variation of V_{CE} does not affect the product of the concentration and the mean velocity of those high kinetic energy electrons, which are traveling within the BCI next to the collector contact and have been injected from the emitter. The variation of V_{CE}, however, does affect their mean velocity and concentration separately, due to the shift of the conduction band edge and hence of their kinetic energy. This means that Q_N must somehow depend on V_{CE}. The V_{CE} dependence is more evident for Q_C, since it represents the sheet charge on the collector contact, which is electrostatically dependent on the traveling electron charge close to the collector. Therefore, to maintain the discussion general, no simplification is assumed for Q_N and Q_C, i.e., $Q_N \equiv Q_N(V_{DE}, V_{DC})$ and $Q_C \equiv Q_C(V_{DE}, V_{DC})$. In addition, the displacement terminal currents $\tilde{I}_E$, $\tilde{I}_C$, and $\tilde{I}_B$ are assumed to be related to the charges by $\tilde{I}_E = \dot{Q}_E + \dot{Q}_N$, $\tilde{I}_C = \dot{Q}_C$, and $\tilde{I}_B = \dot{Q}_{GR}$, where the dots indicate time derivatives, as usual.

With the above assumptions, the standard linearization procedure around the bias point leads to the small-signal equivalent circuit shown in Fig. 7, where the differential parameters are defined as follows:

$$\hat{g}_{BE} = \frac{dI_B}{dV_{DE}}, \quad \hat{g}_m = \frac{dI_C}{dV_{DE}} \tag{11}$$

$$C_Q = \frac{dQ_{GR}}{dV_{BD}}, \quad C_{DE} = -\frac{\partial(Q_E + Q_N)}{\partial V_{DE}} \tag{12}$$

$$C_{DC} = -\frac{\partial Q_C}{\partial V_{DC}}, \quad C_m = \frac{\partial Q_C}{\partial V_{DE}}, \quad C_n = \frac{\partial Q_N}{\partial V_{DC}} \tag{13}$$

where all derivatives are computed at the bias point. For reference, the differential capacitances are plotted versus V_{BE}
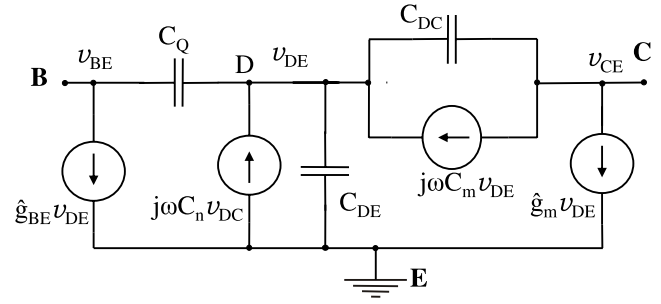

Fig. 7. Small-signal equivalent circuit in saturation regime. The differential parameters are defined in (11)–(13).
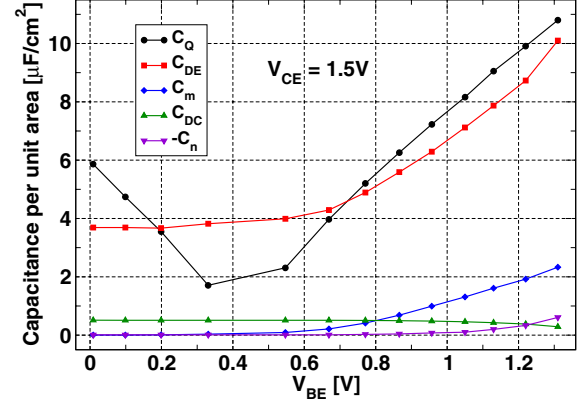

Fig. 8. Differential capacitances defined in (12)–(13) versus V_{BE} at $V_{CE} = 1.5$ V for the GBT1 device. Notice that for convenience $-C_n$ is reported (color online).

in Fig. 8 at $V_{CE} = 1.5$ V for the device GBT1. From the small-signal circuit, the following expression can be derived for the variation of the Dirac potential

$$v_{DE} = \frac{v_{BE} C_Q + v_{CE}(C_{DC} - C_n)}{C_Q + C_{DE} + C_{DC} - C_m - C_n} \tag{14}$$

from which the device transconductance g_m and output conductance g_{CE} are calculated as follows:

$$g_m \equiv \left(\frac{i_C}{v_{BE}}\right)_{v_{CE} = 0} = \frac{\hat{g}_m C_Q}{C_Q + C_{DE} + C_{DC} - C_m - C_n} \tag{15}$$

$$g_{CE} \equiv \left(\frac{i_C}{v_{CE}}\right)_{v_{BE} = 0} = \frac{\hat{g}_m (C_{DC} - C_n)}{C_Q + C_{DE} + C_{DC} - C_m - C_n}. \tag{16}$$

It is seen that the model implicitly includes the effect of the nonzero output conductance, which is given by the variation of the Dirac potential with the collector voltage, which in turn modulates the current. It should be noticed that since the graphene quantum capacitance C_Q is not infinite, as opposed to the case of an ideal metal, the variation of v_{DE} does not coincide with the variation of the base voltage v_{BE}. In particular, when C_Q is near its minimum, which occurs around $V_{BE} \simeq 0.3 \div 0.4$ V according to Fig. 8, v_{BE} has a reduced capability of controlling v_{DE}, from which the already observed drop of g_m originates (Fig. 2). From (15) and (16) the intrinsic voltage gain is written as follows:

$$A_{v0} = \frac{g_m}{g_{ce}} = \frac{C_Q}{C_{DC} - C_n}. \tag{17}$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.
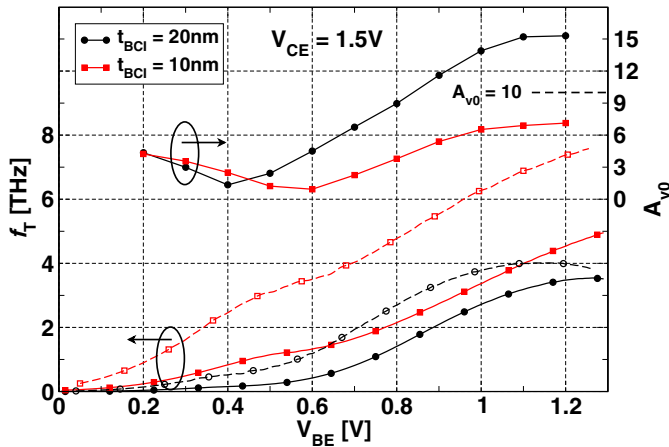
DI LECCE *et al.*: GRAPHENE BASE TRANSISTORS

7

Fig. 9. Cutoff frequency (left axis) and intrinsic voltage gain (right axis) versus V_{BE} at $V_{CE} = 1.5$ V for the GBT1 (black lines) and GBT2 (red lines) devices. The dashed-line f_T curves are calculated using (18) with only the $C_{DE} \simeq -\partial Q_N / \partial V_{DE}$ term in the denominator (color online).

At not too large V_{BE} (low injection) the dependence of Q_N on the collector voltage can be neglected. Hence, C_n can be neglected and C_{DC} tends to the capacitance per unit area of the BCI layer, i.e., $C_{DC} - C_n \simeq \varepsilon_{Si}/t_{BCI} = 0.52 \ \mu F/cm^2$, as confirmed by Fig. 8. For large V_{BE} the situation is more complicated, however the figure shows that $C_{DC} - C_n$ remains comparable with ε_{Si}/t_{BCI}. This indicates that t_{BCI} plays a major role in determining the output conductance and the voltage gain: the small-t_{BCI} device GBT2 has a larger output conductance in saturation than GBT1 (Fig. 3) and a lower gain, as shown in Fig. 9. Indeed, from Fig. 9 it appears that the maximum A_{v0}, which occurs at the largest V_{BE}, drops from nearly 15 to nearly 7 just by halving t_{BCI}, consistently with the previous considerations. Obviously, the limited quantum capacitance of graphene is ultimately responsible for the relative low A_{v0} values. This represents a serious issue for the scalability of the device length.

From the small-signal equivalent circuit of Fig. 7 the cutoff frequency f_T is derived by setting the short-circuit current gain equal to one. It results

$$f_T = \frac{1}{2\pi} \frac{\hat{g}_m}{C_{DE} + C_{DC} - C_m - C_n} \equiv \frac{1}{2\pi} \frac{dI_C}{dQ_{GR}} \qquad (18)$$

where the last identity descends from the definitions (11)–(13) and from the property that the sum of all charge contributions amounts to zero. Interestingly, the cutoff frequency does not suffer from the limitations introduced by the graphene quantum capacitance. Its behavior versus V_{BE} is shown in Fig. 9 for the two devices and it is seen to exceed the 1 THz threshold for a large bias range. The shorter device exhibits a higher f_T, as expected.

To gain a better insight of what ultimately limits f_T, Fig. 9 also shows (dashed lines) for the two devices the cutoff frequencies calculated using again (18), but accounting only for the $-\partial Q_N / \partial V_{DE}$ term in the denominator, which is part of C_{DE}, and neglecting the other capacitances. This represents a sort of upper limit for f_T, since it amounts to considering only the capacitive effects due to the mobile electrons. It is

seen that the GBT1 device is not far from the upper limit at high V_{BE}, while the additive capacitances play a more important role for GBT2. Any further improvement of f_T for GBT1 should therefore go through the use of materials with lower EM and higher electron velocities. Even with silicon, however, the device remains quite promising, since from Fig. 9 a margin seems to exist for achieving at the same time a voltage gain larger than 10 and a cutoff frequency larger than 1 THz.

V. CONCLUSION

The DC characteristics of silicon-based GBT devices have been investigated by means of simulations based on a full-quantum mechanical transport model accounting for nonparabolicity and multiple-valley band effects in ballistic conditions. The analysis has highlighted some peculiarities in the device behavior, mainly related to the space charge effects in the BCI layer, which are responsible for the formation of an internal potential barrier and ultimately limit the extension of the saturation region. The problem can be alleviated by engineering the BCI layer (making it shorter), but a tradeoff exists with the voltage gain, due to the limited quantum capacitance of graphene.

A small-signal model has been developed, which catches the essential physics behind the voltage gain and the cutoff frequency in the saturation regime. It has been shown that, in case a voltage gain of at least 10 is desired, the device length cannot be scaled below 15–20 nm. Nonetheless, the simulations show that a bias window exists within which it is possible to achieve at the same time both a voltage gain better than 10 and a cutoff frequency larger than 1 THz, up to at least 3 THz. Even considering that electron scattering has been totally neglected, this represents a quite promising result, which stimulates new research on this type of devices.

We believe that the presented investigation represents a solid and useful starting point which can guide future optimization studies, devoted to explore systematically the parameter space and the material choice.

REFERENCES

[1] Y. Zhang, Y.-W. Tan, H. L. Stormer, and P. Kim, "Experimental observation of the quantum hall effect and Berry's phase in graphene," *Nature*, vol. 438, no. 7065, pp. 201–204, Nov. 2005.
[2] W. A. de Heer, C. Berger, X. Wu, P. N. First, E. H. Conrad, X. Li, T. Li, M. Sprinkle, J. Hass, M. L. Sadowski, M. Potemski, and G. Martinez, "Epitaxial graphene," *Solid State Commun.*, vol. 143, nos. 1–2, pp. 92–100, 2007.
[3] A. K. Geim and K. S. Novoselov, "The rise of graphene," *Nature Mater.*, vol. 6, no. 3, pp. 183–191, 2007.
[4] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, "Electric field effect in atomically thin carbon films," *Science*, vol. 306, no. 5696, pp. 666–669, 2004.
[5] Y. Wu, Y.-M. Lin, A. A. Bol, K. A. Jenkins, F. Xia, D. B. Farmer, Y. Zhu, and P. Avouris, "High-frequency, scaled graphene transistors on diamond-like carbon," *Nature*, vol. 472, no. 7341, pp. 74–78, 2011.
[6] L. Liao, Y.-C. Lin, M. Bao, R. Cheng, J. Bai, Y. Liu, Y. Qu, K. L. Wang, Y. Huang, and X. Duan, "High-speed graphene transistors with a self-aligned nanowire gate," *Nature*, vol. 467, no. 7313, pp. 305–308, 2010.
[7] H. Wang, A. Hsu, J. Wu, J. Kong, and T. Palacios, "Graphene-based ambipolar RF mixers," *IEEE Electron Device Lett.*, vol. 31, no. 9, pp. 906–908, Sep. 2010.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

8   IEEE TRANSACTIONS ON ELECTRON DEVICES

[8] J. Moon, D. Curtis, D. Zehnder, S. Kim, D. Gaskill, G. Jernigan, R. Myers-Ward, C. Eddy, P. Campbell, K.-M. Lee, and P. Asbeck, "Low-phase-noise graphene FETs in ambipolar RF applications," *IEEE Electron Device Lett.*, vol. 32, no. 3, pp. 270–272, Mar. 2011.

[9] W. Mehr, J. Dabrowski, J. C. Scheytt, G. Lippert, Y.-H. Xie, M. C. Lemme, M. Östling, and G. Lupina, "Vertical graphene base transistor," *IEEE Electron Device Lett.*, vol. 33, no. 5, pp. 691–693, May 2012.

[10] S. Datta, *Quantum Transport: Atom to Transistor*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[11] H. Flietner, "The E(k) relation for a two-band scheme of semiconductors and the application to the metal-semiconductor contact," *Phys. Status Solidi B*, vol. 54, no. 1, pp. 201–208, 1972.

[12] V. Sverdlov, *Strain-Induced Effects in Advanced MOSFETs*. New York, NY, USA: Springer-Verlag, 2011.

[13] M. Luisier and G. Klimeck, "Simulation of nanowire tunneling transistors: From the Wentzel–Kramers–Brillouin approximation to full-band phonon-assisted tunneling," *J. Appl. Phys.*, vol. 107, no. 8, pp. 084507-1–084507-6, 2010.

[14] M. Khayer and R. Lake, "Drive currents and leakage currents in InSb and InAs nanowire and carbon nanotube band-to-band tunneling FETs," *IEEE Electron Device Lett.*, vol. 30, no. 12, pp. 1257–1259, Dec. 2009.

[15] S. Poli, S. Reggiani, A. Gnudi, E. Gnani, and G. Baccarani, "Computational study of the ultimate scaling limits of CNT tunneling devices," *IEEE Trans. Electron Devices*, vol. 55, no. 1, pp. 313–321, Jan. 2008.

[16] H. Yang, J. Heo, S. Park, H. J. Song, D. H. Seo, K.-E. Byun, P. Kim, I. Yoo, H.-J. Chung, and K. Kim, "Graphene barristor, a triode device with a gate-controlled Schottky barrier," *Science*, vol. 336, no. 6085, pp. 1140–1143, 2012.

[17] S. Vaziri, G. Lupina, C. Henkel, A. D. Smith, M. Östling, J. Dabrowski, G. Lippert, W. Mehr, and M. C. Lemme, "A graphene-based hot electron transistor," *Nano Lett.*, vol. 13, no. 4, pp. 1435–1439, 2013.

[18] M. O. Aboelfotoh, "Electrical characteristics of Ti/Si(100) interfaces," *J. Appl. Phys.*, vol. 64, no. 8, pp. 4046–4055, 1988.

[19] K. N. Tu, R. D. Thompson, and B. Y. Tsaur, "Low Schottky barrier of rare-earth silicide on n-Si," *Appl. Phys. Lett.*, vol. 38, no. 8, pp. 626–628, 1981.

[20] S. Zhu, J. Chen, M.-F. Li, S. Lee, J. Singh, C. Zhu, A. Du, C. Tung, A. Chin, and D. Kwong, "N-type Schottky barrier source/drain MOSFET using ytterbium silicide," *IEEE Electron Device Lett.*, vol. 25, no. 8, pp. 565–567, Aug. 2004.

Roberto Grassi received the Ph.D. degree in electrical engineering from the University of Bologna, Bologna, Italy, in 2011.

He is currently a Post-Doctoral Researcher with E. De Castro Advanced Research Center on Electronic Systems, University of Bologna, Bologna, Italy.

Antonio Gnudi (M'98) received the Ph.D. degree in electrical engineering from the University of Bologna, Bologna, Italy, in 1989.

He is currently an Associate Professor of electronics with the University of Bologna.

Elena Gnani (M'98) received the Ph.D. degree from the University of Bologna, Bologna, Italy.

She is currently with the E. De Castro Advanced Research Center on Electronic Systems, University of Bologna.

Susanna Reggiani (M'07) received the Ph.D. degree in electrical engineering from the University of Bologna, Bologna, Italy, in 2001.

She is currently an Associate Professor with the University of Bologna.

Valerio Di Lecce received the Ph.D. degree in electronics and telecommunications from the University of Modena and Reggio Emilia, Modena, Italy, in 2012.

He is currently a Post-Doctoral Researcher with E. De Castro Advanced Research Center on Electronic Systems, University of Bologna, Bologna, Italy.

Giorgio Baccarani (S'68–M'80–SM'92–F'00) received the M.S. degree in electrical engineering and physics from the University of Bologna, Bologna, Italy, in 1967 and 1969, respectively.

He is currently a Full Professor of electronics with the University of Bologna.