Ultra-low power circuits using graphene p-n junctions and adiabatic computing

Original
Ultra-low power circuits using graphene p-n junctions and adiabatic computing / Miryala, Sandeep; Tenace, Valerio; Calimera, Andrea; Macii, Enrico; Poncino, Massimo. - In: MICROPROCESSORS AND MICROSYSTEMS. - ISSN 0141-9331. - ELETTRONICO. - 39:8(2015), pp. 962-972. [10.1016/j.micpro.2015.05.018]

Availability:
This version is available at: 11583/2629471 since: 2016-01-29T11:47:39Z

Publisher:
Elsevier

Published
DOI:10.1016/j.micpro.2015.05.018

Terms of use:
openAccess
This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)
Ultra-Low Power Circuits Using Graphene P-N Junctions and Adiabatic Computing

Sandeep Miryala, Valerio Tenace,
Andrea Calimera, Enrico Macii and Massimo Poncino

Dipartimento di Automatica e Informatica, Politecnico di Torino, 10129, Torino, Italy

Abstract
Recent works have proven the functionality of electrostatically controlled graphene p-n junctions that can serve as basic primitive for the implementation of a new class of compact graphene-based reconfigurable multiplexer logic gates. Those gates, referred as RG-MUXes, while having higher expressive power and better performance w.r.t. standard CMOS gates, they also have the drawback of being intrinsically less power/energy efficient.

In this work we address this problem from a circuit perspective, namely, we revisit RG-MUXes as devices that can operate adiabatically and hence with ultra-low (ideally, almost zero) power consumption. More specifically, we show how to build basic logic gates and, eventually, more complex logic functions, by appropriately interconnecting graphene-based p-n junctions as to implement the adiabatic charging principle.

We provide a comparison in terms of power and performance against both adiabatic CMOS and their non-adiabatic graphene-based counterparts; characterization results collected from SPICE simulations on a set of representative functions show that the proposed ultra-low power graphene circuits can operate with 1.5 to 4 orders of magnitude less average power w.r.t. adiabatic CMOS and non-adiabatic graphene counterparts respectively. When it comes to performance, adiabatic graphene shows 1.3 (w.r.t. adiabatic CMOS) to 4.5 orders of magnitude (w.r.t. non-adiabatic technologies) better power-delay product.

Keywords: Adiabatic Circuits, Graphene p-n junction, Low Power Circuits, Graphene Nanoelectronics

1. Introduction
In the last decade, graphene rapidly emerged as a potential candidate to replace silicon in the next generation of electronic circuits due to its astounding electro-mechanical properties [1].

Graphene is a stretchable and transparent electrical conductor with carrier mobility and saturation velocity far larger than standard silicon-based semiconductors. Those features, combined with the possibility of arranging graphene with other materials to form new composites, represent a perfect match for the growing market of flexible and wearable mobile applications.

Besides these superlatives, however, graphene also shows an indisputable limit, that is, the lack of an energy band gap; conduction and valence band touch each other at zero-energy, where the Fermi Energy \( E_F \) passes, thereby preventing the material to implement the OFF state. In other words, the ON/OFF current ratio in graphene is quite below the value reached in silicon, resulting in a weak separation between logic 0’s and 1’s. Needless to say, this characteristic has been initially used as an argument to support the inadequacy of graphene in the implementation of electronic devices. However, recent studies have proven possible ways to overcome this drawback.

On one hand people have addressed the problem by following a “semiconductor-like” strategy focusing on possible fabrication techniques that physically open an energy gap in the material. Most of them belong to industries that do not want to waste the huge investments done for silicon and would like to replicate as much as possible the successful story of silicon semiconductors with minimal efforts. Graphene Nanoribbons [2] (GNRs) are the most popular embodiment of this class of approaches. GNRs consist of narrow stripes of graphene that show an energy band gap inversely proportional to their width; like standard semiconductors, GNRs can be used to implement Field Effects Transistors (GNR-FETs), e.g., those presented in [3] and [4].

Although even very narrow GNRs exhibit energy gaps sufficiently large for use as a semiconductor to implement graphene-FETs [5], it is quite difficult to fabricate samples of graphene with perfect edges. Edge roughness alters the level of disorder in the material and results in significant degradation of device characteristics and its electrical properties [6].

Another approach exploits the semi-metallic nature of graphene and tries to accommodate it by means of alternative solutions that do not require any physical distortion of the lattice structure. Electrostatic doping [7] is the most representative strategy belonging to this class. It allows a fine-tuning of the Fermi Energy \( E_F \) that can be shifted down into the valence band (to obtain p-type graphene) or up into the conduction band (to obtain n-type graphene) using external electrical filed applied through metal gates. Face-to-face regions with opposite doping profiles form an equivalent p-n junction [8], the key component behind any electronic circuit. It is worth emphasizing that since p-n junctions are built on a pristine sheet of graphene, they preserve the main characteristics of the material.

At this preliminary stage it is hard to predict which of these strategies will prevail in the electronics market, and in how
much time; we embrace the basic principle that an efficient use of graphene should inevitably exploit its intrinsic properties rather than trying to change them. That brought our attention to the second class of methods, i.e., the implementation of digital circuits based on electrostatically controlled graphene p-n junctions.

Graphene p-n junctions can serve as basic switch for a complex logic gate, called RG-MUX (Reconfigurable Gate MultipleXer) because it implements the functionality of a multiplexer. RG-MUXes use a wide graphene sheet (around 190nm) due to which the material defects are within the limits [9, 10]. The authors of [11] propose the realization of various logic gates using RG-MUXes and show that these gates have superior performance and smaller area than traditional CMOS-based ones. RG-MUXes have been characterized for timing and power [12, 13], and various synthesis and design styles that exploit graphene p-n junction [14] and MUX-based [15] have also been investigated [16, 17, 18]. In spite of its great advantage in terms of speed and area, however, RG-MUXes have the drawback of being less energy-efficient than equivalent CMOS gates [19]. This is due to a larger gate capacitance, a consequence of the larger gate area. Energy benefits can be obtained only indirectly thanks to the smaller size of such devices, which allow shorter interconnects [19].

In this work, we revisit RG-MUXes with the objective of making them also energy-efficient, while preserving the highly desirable characteristics of graphene. We do this by recognizing that RG-MUXes, for both structural and functional reasons, naturally lend themselves to implement logic elements that can operate adiabatically. Adiabatic logic [20], aims at mimicking an adiabatic, i.e., without energy exchange, computation process in digital logic. The basic idea of reducing energy dissipation during the switching process relies on the use of a variable (ramp) power supply to recycle a portion of the energy from the load capacitance. Although regarded as a mostly theoretical and somehow exotic computational style, research on the topic has been constantly active over the years, providing several demonstrations of working implementations [21, 22, 23]. The basic building block of any adiabatic circuit is the adiabatic amplifier, a buffer that uses adiabatic charging to drive a capacitive load. In its traditional embodiment, the adiabatic amplifier is implemented using transmission gates (TGs) [20].

Extending a previous work [24], this paper provides three main new contributions: (i) we show that an RG-MUX, by properly assigning voltage values at its terminals, can naturally operate as an adiabatic amplifier, and that each of its two p-n junctions can be regarded as TGs; (ii) we also characterize the figures of merit of these gates and compare them against both adiabatic and non-adiabatic CMOS implementations; (iii) we show how to build fast and extremely low-power adiabatic logic gates based on graphene p-n junctions. Simulation results obtained from SPICE simulations using a dedicated Verilog-A model for the p-n junction, show that the proposed adiabatic gates are 1.5X more power efficient than equivalent adiabatic CMOS-based implementations, still showing more than one order of magnitude better power-delay product. We also demonstrate that such devices are able to operate with about 4X less power and about 4.5X improvement in power-delay product with respect to non-adiabatic counterparts.

2. Background

2.1. Graphene physics

Theorized since 1947 [25], but isolated for the first time only sixty years later [26, 27, 28], graphene is the most surprising allotrope of carbon. It consists of a one-atom-thick sheet of graphite where all the carbon atoms form covalent bonds in a single plane. The resulting two-dimensional (2D) structure is made up of atoms packed in a hexagonal crystal lattice with a carbon-carbon bond length $a_{C-C}$ of 1.42Å. This gives graphene special electrical properties [29] that are a direct expression of its unique energy band structure.

As for any solid material, it is possible to derive an analytical expression of the electron dispersion by solving the time-independent Schrödinger’s equation through the periodic potential of the lattice [30, 31]. An exact solution is computationally challenging, but reasonable and accurate approximations are possible. We refer here to the solution proposed by [32], in which the equation describing the energy dispersion $E$ of graphene is given in (1) and plotted in Figure 1-a. Interested readers can refer to [32, 33] for a formal derivation, which is out of the scope of this work.

$$E^\pm(k) = \pm \gamma \sqrt{1 + 4 \cos \left( \frac{\sqrt{3}a}{2} k_x \right) \cos \left( \frac{a}{2} k_y \right) + 4 \cos^2 \left( \frac{a}{2} k_y \right)} \quad (1)$$

The positive and negative energy branches are the conduction bands ($E^+$) and valence bands ($E^-$), respectively; the vector $\mathbf{k} = \{k_x, k_y\}$ represents the 2D wavevector; $\gamma$ is a fitting parameter whose value can range from 2.7eV to 3.3eV; $a = \sqrt{3}a_{C-C}$ is the side length of the parallelogram representing the primitive cell in the Bravais lattice [32].

From the plot, we notice that the conduction and valence curves touch each other near the edges of the Brillouin Zone, i.e., at
zero energy, where the Fermi energy $E_F$ passes (Figure 1-b). This gapless spectrum provides graphene with semi-metallic properties, different from metals (where $E_F$ is in the conduction band) and semiconductors (where $E_F$ falls in the bandgap). As a consequence, graphenene can only implement a weak “OFF” state. Recent works have shown, however, the possibility to implement equivalent p-n junctions by means of electrostatic doping. Those p-n junctions can be thereby used to implement logic gates.

2.2. Graphene p-n junction

Figure 2 shows the basic structure of a graphene p-n junction that uses electrostatic doping. The device is composed of four layers, namely: (i) the bottom layer, that includes two split gates (referred as back gates in the figure) made of conductive material separated by a distance $D$; (ii) an insulating layer of oxide, which is placed on top of the split gates; (iii) a wide graphene sheet; and, (iv) two electrodes (front metal contacts in the figure) placed on top of the graphene sheet, which serve to supply a reference current to the device.

The application of a negative voltage on a back-gate shifts $E_F$ towards the valence band resulting in p-type graphene in the above region. On the other hand, a positive voltage shifts $E_F$ towards the conduction band leading to n-type graphene [34]. In this way, by applying asymmetric voltages to the two back gates a p-n junction is formed [35]. The front metal contacts represent the conceptual source (left) and the probe (right) that are emitting and receiving carriers.

As demonstrated in [36], carriers injected in the p-region through the left front contact cross the potential barrier at the p-n junction with a transmission probability $T(\theta)$ which depends on two parameters: the angle $\theta$ between the electron’s wave vector $\mathbf{k}$ and the normal of the junction, and the width $D$ of the p-n transition region. The analytical expression of $T(\theta)$ is:

$$T(\theta) = \cos^2(\theta)e^{-\frac{2\pi D}{\hbar} \sin^2 \theta}$$

(2)

The transmission probability $T(\theta)$ is thus 1 for carriers that travel orthogonally with respect to the junction ($\theta = 0$), regardless of $D$, and decreases exponentially for larger values of $\theta$, e.g., $T(\theta) = 0$ for $\theta = \pi/2$. Notice that when a symmetric control voltage is simultaneously applied to two adjacent back-gates ($(+V, +V)$ or $(-V, -V)$), the graphene layer is entirely of p- or n-type, respectively; the device is thus transparent to the charge flow.

2.3. Electrical model of graphene p-n junction

Fig. 3 shows a schematic of the electrical model implemented borrowing the works of [11, 19]. Terminals A and Z denote the left and right front contacts respectively, whereas S and U denote the back gate potentials on the two gates. The resistor $R_{AZ}$ represents the resistive equivalent of a graphene p-n junction between input A and output Z. Its value ranges from $R_{ON} = 300 \Omega$, under p-p or n-n configurations, to $R_{OFF} = 10^3 \Omega$, under p-n or n-p configurations. The analytical expression for the junction resistance is expressed as:

$$R_{AZ} = \frac{R_0}{N_0kT(\theta)}$$

(3)

In (3), the transmission probability of the carriers across the junction is given by (2), $R_0 = \frac{1}{2e^2}$ is the quantum resistance per propagation mode, and $N_0$ is the number of excited propagation modes [37]. The electrical model also includes parasitics of the metal contacts. The resistors $R_A$ at the front pins A and $Z$ model the resistance of the metal-to-graphene contacts [38]. The lumped capacitance $C_g$ at the back-gates, i.e., $C_{gS}$ at S and $C_{gU}$ at U, consists of the series of the oxide capacitance $C_{ox}$ and the quantum capacitance of the graphene sheet $C_q$, namely $C_g = 1/(C_{ox}^{-1} + C_q^{-1})$. $C_{gS}$ and $C_{gU}$ denote the gate capacitances at the back gates S and U respectively. This electrical model was implemented in Verilog-A and included in our SPICE simulations.

Fig 4 shows the variation of the junction resistance obtained through SPICE simulations; there are two curves: one with the potential of back-gate $U$ fixed at either a positive potential ($+V_{dd}/2$, cross mark) and at a negative potential ($-V_{dd}/2$, plus mark). Let us consider the curve with the potential of $U$ fixed at $+V_{dd}/2$ and let us vary the potential of the other back-gate $S$ from $-V_{dd}/2$ to $+V_{dd}/2$. As it can be seen, at $-V_{dd}/2$, the junction resistance is very large, and as it approaches $+V_{dd}/2$ the resistance settles at few hundred Qs. In practice when the two back-gates have the same polarity the junction resistance is small.
When on the voltage assignments at its back-gate inputs. The resulting p-n-n doping profile of the graphene sheet creates a low-resistive path between the front contacts A-Z and a high-resistive R_{pp} path between the front contacts B-Z (Figure 6-a). This forces the output Z to follow the input signal associated with the lowest resistance, i.e., \( Z = A \).

Similarly, when \( S = "1" \) the central graphene region is n-doped, leading to a p-n-n doping profile of the graphene layer. Therefore a low-resistive path between the contacts B-Z makes the output to follow B, i.e., \( Z = B \) (Figure 6-b).

From the above analysis, the RG-MUX can be seen as a 3-input, 1-output device, where A and B are the data input, S the control input, and Z the output. Depending on the polarity of S the resistances on the two input-to-output paths are properly set (refer to Figure 6-c). From a functional point of view, the RG-MUX implements a multiplexer, where S is the selection input, i.e., \( Z = S \cdot B + \overline{S} \cdot A \).

### 3. Adiabatic Computing

The development of adiabatic logic in CMOS technology has been widely investigated in the past decades [39]. However, several of the proposed techniques show some critical downsides, mainly regarding: (i) increase in logic design complexity; (ii) the need of multiple supplies for proper interfacing between stages; and (iii) a self-charging effect at the output nodes, which is due to leakage currents. In the proposed graphene-based implementation the first two concerns are removed by construction, whereas only the third issue remains open. In this section, we employ a single-supply adiabatic strategy that allows the realization of compact logic circuits operating at low power.

#### 3.1. Adiabatic switching

Consider the application of a step voltage source (from 0 to \( V_{dd} \)) to an RC circuit (Figure 7-a). The total energy supplied by the voltage source is given by \( E_{supplied} = Q \cdot V_{DD} = CV_{DD}^2 \).

The energy stored in a capacitor when the output has reached the final value is given by \( E_c = \frac{1}{2}CV_{DD}^2 \), i.e., only half of the supplied energy. The other half is dissipated across the resistor in the form of heat (\( E_R \)). This scenario represents the traditional "computation" model, in which energy waste is maximum.
In contrast, adiabatic logic uses a slow-rising ramp supply signal with transition time \( T_r \) to change from 0 to \( V_{DD} \) (Figure 7-b). Assuming that \( T_r \) is such that the capacitor is able to charge instantaneously to the input supply voltage, the current is given by:

\[
i(t) = C \frac{dv(t)}{dt} = \frac{CV_{DD}}{T_r}
\]

(4)

The energy dissipated across the resistor is obtained by integrating the power across the resistor over time \( T_r \):

\[
E = \int_0^T v(t) i(t) dt = \int_0^T \frac{C^2 V_{DD}^2}{T_r^2} dt = \frac{RC^2 V_{DD}^2}{T_r} V_{DD} (5)
\]

where \( v(t) \) is the voltage drop across the resistor.

The minimum transition time for which the step and ramp source produce the same energy dissipation can be computed from (5) and the expression of \( E_{\text{supplied}} \). For \( T_r > 2RC \), adiabatic circuits are more energy efficient than regular circuits with step input [40].

3.2. The adiabatic amplifier

The adiabatic amplifier [20] is the basic building block used in adiabatic circuits; from a functional point of view it is a simple buffer that uses adiabatic charging to drive a capacitive load. In traditional CMOS implementations, the typical realization uses transmission gates (TGs) as shown in Figure 8. The choice of TGs is because of their fully-restoring feature, i.e., the NMOS transistor is used to pass logic “0” whereas the PMOS is used to pass logic “1”, so the output is always strongly driven and the levels are never degraded.

The basic operations of the adiabatic amplifier are quite intuitive. Depending on the values of \( X \) and \( \overline{X} \), one of the output capacitances will be adiabatically (through \( V_A \)) charged. Clearly, the two sides are mutually exclusive and only one side drives the supply voltage to the output.

The circuit uses dual-rail encoding for both inputs (because driving the TGs require the double polarity of \( X \)) and outputs; in the latter case this is required for interconnecting the amplifier to other adiabatic circuit elements. Operations occur in three phases, corresponding to the three “regions” of the ramp input. The input is first set to a stable value: then the “charge” phase starts, in which the supply \( V_A \) is ramped and the load capacitor is adiabatically charged. In the second “evaluate” phase, when \( V_A \) has reached the final value, output voltages are stable and can be used by next logic stages. In the third “recovery” phase, the load capacitor discharges back into \( V_A \) as it ramps down to 0V.

4. Adiabatic through RG-MUXes

4.1. RG-MUX as an adiabatic amplifier

If we compare the RG-MUX of Figure 5 and its model in Figure 6 with the adiabatic amplifier of Figure 8 we can immediately spot many similarities. Both devices operate as selectors with variable resistance from inputs to outputs: in the RG-MUX, signal \( S \) determines the low- and high-resistance paths, whereas in the adiabatic amplifier it is represented by the signal \( X \). There is, however, a significant difference between the two devices, namely the flow of “computation”, which occurs in opposite directions. In the RG-MUX the central node \( Z \) is an output, whereas in the adiabatic amplifier the central node is connected to the ramped supply voltage. The opposite occurs for the two side nodes \( (A \) and \( B) \) which are inputs in the RG-MUX and outputs in the adiabatic amplifier.

Figure 9 shows the input and output signals assignment of an RG-MUX so that it can be used to implement the operations of an adiabatic amplifier.

Specifically, we notice how signal \( Z \) (an output in a RG-MUX) is now connected to the variable supply voltage \( V_a \), whereas pins \( A \) and \( B \) are now the dual-rail encoded values of the output. The role of the back-gate input \( S \) (together with the specific encoding of \( U \) and \( \overline{U} \), which are normally fixed) is similar to the role of \( X \) and \( \overline{X} \) signals in the amplifier: \( S \) = “0” (“1”) is equivalent to \( X = 0 \) \( (X = 1) \), that is, output \( F \) (\( \overline{F} \)) is charged by \( V_a \).

4.2. Implementing adiabatic logic gates

The adiabatic amplifier is the simplest of many possible adiabatic logic gates and functionally corresponds to a buffer. In order to implement complex logic functions, we adopt the architectural template of a generic logic function of [20], in which a

---

2 The original implementation of the adiabatic amplifier includes two NMOS clamp at the two outputs driven by \( X \) and \( \overline{X} \), which are not shown here.
parallel of the structure of classical CMOS gates is maintained. In practice, the complementary pull-up and pull-down networks of CMOS are replaced by two equivalent networks of TGs. The new pull-up network is used to charge the output \( F \), whereas the pull-down one will charge the complemented output \( \overline{F} \) (Figure 10-a). Figure 10-b shows an example of a logic gate which realizes an AND/NAND function.

The implementation of arbitrarily complex logic functions requires the availability of the graphene counterpart of a TG. Since the RG-MUX is essentially a side-by-side combination of two junctions (in the p-n/n-n or p-p/p-p pattern), we can use a single graphene p-n junction ("half" RG-MUX) to implement a TG. Figure 11 illustrates the basic concept of this equivalence.

\[ \overline{U}/U \] pins, which are assumed to be fixed at "0" and "1" respectively. The symbol on the left (dark red back-gate) denotes \( \overline{U} = "0", \) i.e., \(-Vdd/2\), and is therefore equivalent of the TG with signal transmitted when \( X = 0 \); the symbol on the right (light green back-gate) denotes \( U = "1", \) i.e., \( Vdd/2\), and corresponds to the TG transmitting signal with \( X = 1 \). In other words, by splitting the RG-MUX, one can obtain two TGs with positive and negative polarity, with no need of complementing the input signal \( X \).

Using the general architecture of Figure 10, the basic logic gates (INV/BUF, AND/NAND and OR/NOR) based on these TGs can be implemented as shown in Figure 12. \( V_a \) denotes the ramped-supply voltage that enables adiabatic operations. Two p-n junctions are needed to realize an INV/BUF, whereas four junctions are needed for AND/NAND and OR/NOR.

For the sake of clarity, we depict in Figure 13 the signal waveforms of a 2-input AND (left plot) and of a 2-input OR (right plot). Signal \( V(a) \) represents the power supply, \( V(X) \) and \( V(Y) \) represent the two input signals, and signals \( V(F) \) and \( V(Fbar) \) represent the true and complemented outputs of the circuits. Let us consider the plot on the left. When the two input signals are both at logic 1, i.e., the AND operator evaluates to 1, the supply signal \( V(a) \) is propagated to the output; whereas, in the other three cases, namely when the AND operator evaluates to logic 0, \( V(a) \) is not allowed to propagate. In other words, a graphene p-n junction based circuit propagates the supply signal to the output node if, and only if, the logic function evaluates to logic 1. This is an example of the intrinsic pass-through characteristics of graphene p-n junctions. As the plot suggests, an important aspect to be considered is the leakage-induced charging of the output nodes, referred as the Self-Charging effect. Consider the situation in which the input \( X/\overline{X} \) is kept constant for several cycles. The \( I_{eff} \) current drained from \( V_a \) charges the output capacitance. The effect is plotted in Figure 13, where \( \overline{F} \) (plot \( v(Fbar) \) in the picture) slowly charges to intermediate voltage levels. After some cycles, this will cause a bit-flip at the output node. A quantitative analysis of the output voltage refresh rate is presented in the Section 5.

5. Simulation results

5.1. Power dissipation across RC-circuits

The first experiment compares the energy benefits obtained by adiabatic operations on graphene-based devices against those achieved with CMOS. More precisely, we measured the power dissipation for the RC circuit shown in Figure 7 for both technologies. The CMOS configuration consists of a TG designed using minimum size PMOS and NMOS available in a 40nm technology library provided by STMicroelectronics, driving a capacitance corresponding to the input pin/gate capacitance of a minimum sized inverter, i.e., \( 0.6fF \). The equivalent resistance is \( 2.56k\Omega \), obtained by the parallel of the resistances of the p- and n- networks in the linear region of operation. The graphene configuration consists of a graphene p-n junction driving a capacitance equal to the back-gate capacitance of an RG-MUX, i.e., \( 0.62fF \). The resistance is that of the junction (including perpendicular...
Graphene starts consuming less power. Notice that, although the contact resistance), i.e., 1.38kΩ. In both cases, power dissipation after application of the ramp input signal is estimated through the power command available in HSPICE over a signal transition at the output node. Experimental results are plotted in Fig 14.

We notice that for very small values of $T_r$ (< 10 ps) the equivalent CMOS network is more energy-efficient than the graphene one. However, such small values of $T_r$ actually correspond to non-adiabatic operations. As a matter of fact, even the steepest transition in a conventional non-adiabatic circuit takes a few picoseconds of rise/fall time. Such superiority corresponds to the intrinsic energy efficiency of CMOS over graphene-based devices mentioned in Section 1. For $T_r > 10 ps$, graphene and CMOS initially exhibit similar power consumption, before graphene starts consuming less power. Notice that, although the truly adiabatic region for graphene starts at $T_r = 2RC ≈ 1.7 ps$, power consumption is already lower for smaller $T_r$ values. On average, there are about two orders of magnitude consistently over the entire range of $T_r$ values greater than 10 ps.

Please note that we reported a large range of $T_r$ values although some of the larger values in the ps range are clearly impractical but for some specific low-throughput applications.

5.2. Power characterization of the adiabatic amplifier

We characterized the power consumption of the adiabatic amplifier using HSPICE, by assessing in particular the breakdown of power between ON and OFF states in the two junctions. Figure 15 shows these two components as a function of $T_r$. The plot reports the average power $P_{on}$ and $P_{off}$ of the two p-n junctions of the amplifier. Subscripts “on” and “off” denote the fact that the two junctions have complementary states.

The figure clearly shows that both components of power dissipation decreases as $T_r$ increases, yet with different scales. While total power (sum of $P_{on}$ and $P_{off}$) exhibits about seven orders of magnitude power reduction over the range of $T_r$ values.
reported, for smaller $T_r$, power is determined by active power $P_{on}$, whereas around $T_r = 100\,\text{ps}$ the two components have similar weight, and for larger $T_r$ values the off current (junction leakage) becomes dominant. Experiments were replicated for different device width ($W$), i.e., the size of the p-n junction along the $k$, axis depicted in Figure 2. In particular, we characterized devices assuming $W = 189\,\text{nm}$ and $W = 100\,\text{nm}$. As can be seen from the plot, for lower values of $T_r$, increasing the device width results in increased off-state power. Also, increasing device width decreases the junction resistance, as per (3), and increases the average current across the branch. However, for large values of $T_r$, the signal at $F$ charges towards logic “1” irrespective of the junction resistance making constant off-state current. Similarly, at low $T_r$ the power in the on-state increases with device width due to the increase of the average current across the junction and remains constant for larger values of $T_r$.

5.3. Relation with operating frequency

An important concern to be considered is the relation of $T_r$ with the operating frequency of the circuit. Assuming a 50% duty cycle for the variable power supply. The load capacitor has to completely charge to its final value in the given duty cycle and remain in that state for a significant time. Therefore, although power savings are maximum for large $T_r$’s, there exists an optimal $T_r$ for each frequency, calculated as the value that allows the capacitor to charge to its final value in the middle of the duty cycle. As an intuitive rule of thumb, higher frequencies imply smaller $T_r$ values. Calculation of such optimal value yields the following expression of $T_r$ (in seconds) vs. frequency (in Hz): $T_{r,\text{opt}}(f) = \frac{1}{2f}$.

5.4. Leakage currents and self-charging effect

To avoid the self-charging effect (see Section 4), adiabatic circuits must periodically refresh the output voltages. Since $I_{off}$ increases for increasing $T_r$, the lower the operating frequency the faster the charging of the capacitor and thus the more frequent the need of refreshing the voltage. The actual refresh frequency depends on what level of the output voltage should be considered as critical, e.g., some percentage of $V_{dd}$.

The time taken to reach that level, expressed in cycles at the desired frequency, will determine the feasibility of the operations. Since we can refresh at most once per cycle, this time should be greater than 1 cycle. Tab 1 reports the results of an exploration of feasibility by spanning different frequency values ($1 \div 1000\,\text{MHz}$), and different voltage thresholds (up to 50% of $V_{dd}$ in steps of 10%). Entries in bold are the feasible ones. We can observe that very low frequencies are not feasible even for large voltage degradation at the output, whereas frequencies in the order of a few hundred of MHz are mostly feasible. Using the relation between $T_r$ and $f$ of Section 5.3, this corresponds to $T_r$’s in the order of $1\,\text{ns}$, in the region in which there is an adiabatic benefit over CMOS (please refer to Figure 14).

5.5. Characterization of simple logic functions

We realized 46 simple logic functions using the adiabatic approach presented in this paper and a classical standard cell approach which is non-adiabatic. The characterization was done for a range of transition time starting from $1\,\text{ps}$ to $1\,\text{ns}$. We compare the characterization results of graphene technology with that of conventional CMOS technology. Adopted logic functions are summarized in Table 2. We limited our experiments to functions having no more than six devices in series. We then compared the following four implementations:

**Adiabatic graphene**: logic functions are realized according to the methodology described in this paper, i.e., each p-n junction acting as a transmission gate. A ramp signal is fed to the common intersection point of the dual logic design style.

**Adiabatic CMOS**: logic functions are implemented using MOSFET based transmission gates, i.e., parallel connection of NMOS/PMOS. Also in this case a ramp signal is used to exploit the adiabatic charging principle.

**Non-adiabatic graphene**: logic functions are first synthesized using the ABC synthesis tool using a subset of all possible gates, namely INV, BUF, AND, OR. Then, each gate is replaced acting as a transmission gate. A ramp signal is used to refresh the voltage. The actual refresh frequency depends on what level of the output voltage should be considered as critical, e.g., some percentage of $V_{dd}$.

**Non-adiabatic CMOS**: logic functions are created using standard complementary MOS architectures, namely, by means of pull-up/down networks connected to $V_{dd}$/GND; the pull-up network is implemented using PMOS whereas the pull-down network is realized using NMOS with proper size.

The simulations for estimating power/performance was carried out using Synopsys HSPICE. For graphene technology we use the electrical model presented in Section 2.3 written in Verilog-A, whereas for CMOS technology we use the models from the ST-Microelectronics at 40nm technology node. The load capacitance was fixed at 0.6/$f$, which is the minimum gate capacitance in both technologies.

5.5.1 Power characterization

Power dissipation results for the logic functions in Table 2 is plotted in Figure 16. We report the average power dissipation
over the 46 logic functions versus the input signal transition time. First key observation is that, irrespective of the technology, non-adiabatic implementations exhibit higher power dissipation compared to adiabatic ones. However, non-adiabatic graphene technology has the least power dissipation due to the smaller junction resistance compared to that of a MOSFET. For the adiabatic implementations, graphene technology has smaller power dissipation over CMOS technology over the entire range of \( T_r \), with a maximum gain of about 1.5 orders of magnitude at \( T_r = 1 \text{ns} \).

5.5.2. Power-delay product characterization

Figure 17 depicts the power delay product (PDP) of adiabatic and non-adiabatic implementations in graphene and CMOS technologies. The PDP plotted is averaged over the 46 logic functions for a given transition time. The objective of this plot is to show that power savings in adiabatic implementations do not come at the expense of performance loss. As can be seen from the plot, the PDP of graphene-based adiabatic implementations is more than an order of magnitude smaller than that of adiabatic CMOS. However, in the non-adiabatic approach it is the CMOS technology that has better PDP than a graphene-based implementation.

5.6. A CAD tool for adiabatic graphene circuits

The logic functions described in the previous section were built manually. In order to allow an automated synthesis usable for realistically-sized designs, we implemented a CAD tool that maps a generic Boolean function onto a graphene-based p-n junction technology. We adopted Binary Decision Diagrams (BDDs) \([41, 42]\) as a common data structure to represent the functions of the circuits. A generic node of a BDD (Figure 18-a) has two in-going edges \( f \) and \( f' \), representing the positive and negative co-factors of the function, respectively. They propagate the input ramp-signals generated by the two terminal nodes 0 and 1 to the out-going edge, labeled \( f \). Such evaluation is based on the value assumed by a generic primary input \( x \) associated to the decision node.

![Figure 16: Average Power of logic functions in Table 2.](image)

![Figure 17: Average Power Delay Product (PDP) of logic functions in Table 2.](image)

![Figure 18: Decision node in BDD. (a) Logic structure, (b) PN-BDD realization, and (c) TG-BDD realization.](image)
Figure 19: Design flow.

- **TG-BDD**: each internal BDD node is mapped with MOSFET transmission gates, as depicted in Figure 18-c.

In both cases we exploit the adiabatic charging principle to achieve the minimum power consumption. The adopted design flow is illustrated in Figure 19. Each benchmark is first processed with the open-source CUDD library [43] in order to generate the corresponding BDD structure; then, a TCL script maps each decision node in the corresponding technology, i.e., PN-BDD or TG-BDD. Node descriptions are stored in a technology library which is fed to the TCL mapping script. The output is a SPICE-compliant netlist.

Table 3 summarizes the results we obtained on a set of open-source benchmarks. Columns **PI** and **PO** show the total number of primary inputs and outputs of the design and the total number of nodes in the BDD structure. Columns **PN-BDD** and **TG-BDD** report, for the two implementation styles, the number of p-n junctions (for the PN-BDD) and transistors (for the TG-BDD) as well as the resulting area.

For the PN-BDD structure, the equivalent area of a single BDD node is given by the sum of the two p-n junctions (refer to Figure 18), which corresponds to 0.382$\mu$m² [11]. The total area occupation for a given circuit is simply obtained by multiplying this number by the number of BDD nodes. The transmission gate implementation conversely consists of six MOSFET transistors, i.e., a PMOS/NMOS for each transmission gate, and one PMOS/NMOS for the inverter stage. In our case, we adopted minimum-sized transistors defined in a commercial technology library from STMicroelectronics at 40nm node, where each PMOS has an equivalent area of 0.205$\mu$m², and each NMOS has an equivalent area of 0.147$\mu$m². Therefore, the total area of a single TG-BDD node is equal to 1.056$\mu$m².

It is therefore clear that the first advantage of the PN-BDD structure is in the number of devices needed to implement a BDD node. Infact, a TG-BDD node requires an additional inverter stage that is not needed in a PN-BDD node. Another key-point is signal degradation: a stack of MOSFET transmission gates connected in series shows a higher signal degradation w.r.t. a series of p-n junctions. For this reason, in our mapping script, we added a signal-restoration stage for each stack of ten MOSFET transmission gates, by employing a minimum-sized buffer from the same library with an equivalent area of 1.038$\mu$m². Results reported in Table 3 shows that the proposed PN-BDD mapping strategy requires, on average, almost 70% less devices w.r.t. the TG-BDD implementation, which translates to a roughly 39% area saving.

Concerning the efficiency of the PN-BDD over the TG-BDD strategy, we can refer to Figure 20, where we depict the averaged power consumption of the two techniques in function of the transition time ($T_r$) of the input signal. Both structures are considered in adiabatic configuration. As can be seen from the plot, in non-adiabatic region, i.e., $T_r = 1$ps, TG-BDD (circle mark) is slightly more efficient (about 3.5%), whereas, as $T_r$ increases PN-BDD (square mark) results in less power consumption. The best case is recorded at $T_r = 1$ns with a difference of 1.8 orders of magnitude, which is mainly due to the lower power requirements of an adiabatic p-n junction w.r.t. a MOSFET transmission gate.

In Figure 21, we also address the performances of the proposed mapping solutions by means of the power-delay product (PDP). As can be seen from the plot, the PN-BDD structure shows lower PDP w.r.t. the TG-BDD counterpart over the entire $T_r$ range. The motivation lies in two fundamental differences of the structures, namely the presence of an inverter stage in each TG-BDD node, which increases the configuration time, and the additional signal-restoring stages that increase the total propagation delay.

<table>
<thead>
<tr>
<th>Benchmark</th>
<th>PI</th>
<th>PO</th>
<th>Nodes</th>
<th>PN-BDD</th>
<th>Area [$\mu$m²]</th>
<th>Transistors</th>
<th>TG-BDD</th>
<th>Area [$\mu$m²]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sram</td>
<td>9</td>
<td>1</td>
<td>25</td>
<td>500</td>
<td>19.14</td>
<td>150</td>
<td>26.4</td>
<td></td>
</tr>
<tr>
<td>fbsp</td>
<td>7</td>
<td>10</td>
<td>42</td>
<td>84</td>
<td>32.08</td>
<td>252</td>
<td>44.35</td>
<td></td>
</tr>
<tr>
<td>ric</td>
<td>8</td>
<td>31</td>
<td>79</td>
<td>158</td>
<td>60.35</td>
<td>474</td>
<td>81.82</td>
<td></td>
</tr>
<tr>
<td>max1024</td>
<td>10</td>
<td>6</td>
<td>250</td>
<td>500</td>
<td>191</td>
<td>1500</td>
<td>240.11</td>
<td></td>
</tr>
<tr>
<td>eg5t</td>
<td>8</td>
<td>6</td>
<td>250</td>
<td>500</td>
<td>191</td>
<td>1500</td>
<td>240.11</td>
<td></td>
</tr>
<tr>
<td>apex2</td>
<td>38</td>
<td>3</td>
<td>329</td>
<td>658</td>
<td>251.35</td>
<td>2182</td>
<td>400.33</td>
<td></td>
</tr>
<tr>
<td>ala4</td>
<td>14</td>
<td>8</td>
<td>351</td>
<td>702</td>
<td>288.10</td>
<td>2308</td>
<td>431.90</td>
<td></td>
</tr>
<tr>
<td>e52</td>
<td>35</td>
<td>7</td>
<td>1106</td>
<td>2352</td>
<td>1090.32</td>
<td>2900</td>
<td>1469.55</td>
<td></td>
</tr>
<tr>
<td>apex1</td>
<td>45</td>
<td>45</td>
<td>1246</td>
<td>2492</td>
<td>932.34</td>
<td>7894</td>
<td>1402.36</td>
<td></td>
</tr>
<tr>
<td>lps</td>
<td>24</td>
<td>109</td>
<td>1693</td>
<td>3586</td>
<td>1237.45</td>
<td>11120</td>
<td>2091.93</td>
<td></td>
</tr>
<tr>
<td>square</td>
<td>31</td>
<td>71</td>
<td>2268</td>
<td>4536</td>
<td>1522.75</td>
<td>3385</td>
<td>5986.11</td>
<td></td>
</tr>
<tr>
<td>c880</td>
<td>60</td>
<td>26</td>
<td>9010</td>
<td>18000</td>
<td>6883.60</td>
<td>41320</td>
<td>11431.2</td>
<td></td>
</tr>
<tr>
<td>sqe</td>
<td>41</td>
<td>42</td>
<td>9159</td>
<td>18318</td>
<td>6997.42</td>
<td>56250</td>
<td>10522.04</td>
<td></td>
</tr>
<tr>
<td>c1908</td>
<td>33</td>
<td>25</td>
<td>9714</td>
<td>19482</td>
<td>6134.84</td>
<td>36609</td>
<td>12998.30</td>
<td></td>
</tr>
</tbody>
</table>

**avg.**

Total area:

- **PO**
- **Nodes**
- **PN-BDD**
- **Area**
- **Transistors**
- **Area**

6. Conclusions

This paper explores the potential of graphene p-n junctions for implementing logic devices that operate according to the principle of adiabatic computation.

Results show that these gates can operate with about 2 orders of magnitude lower power than their CMOS counterpart, overcoming a limitation of these graphene-based elements observed in previous works. Moreover, these power benefits do not require very low frequencies as in typical adiabatic operations.
and can be achieved with operating frequencies in the order of the hundred of MHz.

[33] T. Yan, Q. Ma, S. Chilstedt, M. D. Wong, D. Chen, Routing with graphene


URL http://vlsi.colorado.edu/~fabio/CUDD/