Ballistic Deflection Transistors and the Emerging Nanoscale Era

David Wolpert, Hiroshi Irie, Roman Sobolewski, and Paul Ampadu
Department of Electrical and Computer Engineering
University of Rochester
Rochester, NY 14627

Quentin Diduck and Martin Margala
Department of Electrical and Computer Engineering
University of Massachusetts Lowell
Lowell, MA 01854

Abstract—This paper presents a brief survey of the state of the art in nanoscale electronics, with special emphasis on room-temperature nanoscale ballistic deflection transistors (BDTs) and T-branch junctions (TBJs). Both devices are planar structures etched into a two-dimensional electron gas (2DEG). Extremely low capacitances (~0.2 fF) in the 2DEG system and low switching voltages (~0.15 V) predict THz performance and ultra-low power consumption, making BDTs and TBJs among the most promising and versatile of ballistic nanoelectronic devices. Obstacles in circuit and logic design using the BDT are presented along with potential solutions. I-V characteristics from a fabricated BDT and simulation results from a two-input BDT NAND gate are provided. Future plans to facilitate large-scale integration are discussed.

I. INTRODUCTION

As CMOS scaling nears its inevitable limit, researchers are searching for novel devices and reliable architectures to replace or extend conventional silicon electronics. A wide variety of alternative devices have been proposed as potential successors to conventional Si CMOS. These devices can be categorized into “traditional,” having properties similar to conventional FETs, and “non-traditional” devices with fundamentally different properties and/or novel materials.

This paper provides an overview of emerging devices, circuits, and architectures. Section II provides a brief survey of nanoscale devices and architectures. Section III describes ballistic transport in 2-dimensional electron gas (2DEG) heterostructures, and examines operation of planar T-branch junctions (TBJs) and ballistic deflection transistors (BDTs); measured and simulated results are further presented. Conclusions and future trends are presented in Section IV.

II. NANOSCALE DEVICES AND ARCHITECTURES

A. Nanoscale Architectures

As devices are scaled into the nanometer regime, they become more susceptible to process, voltage, and temperature (PVT) variations and noise. Functional must be maintained in the presence of these conditions. One reliability improvement is the use of regular structures, which have been shown to improve yield by 7% while reducing area by over 15% [1]. On-chip sensors can detect voltage and temperature variations, allowing supply voltage or operating frequency to be adapted to guarantee functionality [2]. The Razor architecture [3] uses a delayed latch to check for delay errors at runtime. Error control methods are needed to handle single- and multi-wire errors, which are prevalent in nanoscale systems [4]. In [5], Hamming product codes are used to address multi-wire errors in an energy efficient manner. Adaptive error control schemes have been proposed [6] to provide average case energy dissipation while still handling worst case noise conditions.

B. Traditional Nanoscale Devices

Traditional device improvements include new material structures, fabrication processes, or integration methods. Surveys of nanoscale devices emphasizing these traditional improvements are available [7][8]. III-V compounds, such as GaAs, have been used in electronic and optical applications, resulting in operating frequencies of over 250 GHz in a GaAs HBT [9]. Channel properties can be adjusted through the use of Ge, SiGe, or strained Si. An excellent discussion of these techniques is available in [10], with mobility improvements of up to 10X reported. Gate design has been re-engineered to reduce leakage current and increase the I_on/I_off ratio. These techniques include metal gates with high-k dielectrics, dual-gate FETs, and finFETs. Silicon-on-insulator (SOI) techniques result in large performance improvements by allowing tighter control of noise, coupling, and substrate leakage [7]. Creation of ultra-thin substrates has facilitated the adoption of 3D integrated circuits [11]. Using through-Si vias to connect multiple stacked device layers, this allows for large reductions in wire length, improving latency and system performance.

C. Non-Traditional Devices

Molecular structures have been used to create both passive and active devices [12]. Nanowires can be used in conjunction with active molecular switches to create logic using crossbar structures, such as memristor crossbar latches [13]. Graphene sheets, nanoribbons, and carbon nanotubes can exhibit either metallic or semiconducting properties depending on their chirality. They exhibit some of the highest carrier mobilities of any known device, and should result in operating frequencies well into the THz range [14]. Single-electron transistors [15] are deep nanoscale devices that make use of Coulomb barrier tunneling to provide logic. This "one-electron-at-a-time" transport can be regarded as the ultimate limit in nanoelectronics. Spintronics [16][17] uses the spin of an electron or group of electrons to encode logic information. Ferromagnetic semiconductors are increasingly implemented as spin-valve devices [16] and offer promise for large, nonvolatile random-access memories. The above topics are compared in a survey of non-traditional devices in [18].
Research is also ongoing in the areas of DNA computing [19], peptide computing [20], and, most notably, quantum computing [21]. These technologies have yet to provide general computing solutions, though they promise immense computational power for parallel computing problems such as searching databases or factoring very large numbers. Finally, nonlinear optical devices such as semiconductor all-optical switches and amplifiers have recently been used to create Boolean logic functions [22].

III. BALLISTIC TRANSPORT DEVICES

Various III-V compound heterostructures, such as AlGaAs/InGaAs or InGaAs/InAlAs systems, can be used to create a 2DEG layer with very high electron mobilities. In 2DEG structures with dimensions smaller than the electron mean free path $l_e$, electrons can travel without any scattering events. This is referred to as ballistic transport and can be observed even at room temperature over small distances; an excellent overview of ballistic transport is presented in [23]. One-dimensional electron gas transistor channels are used to create high-electron mobility transistors (HEMTs) [24], which can exhibit up to THz performance.

Ballistic transport can be explained using Fig. 1, which is similar to an experiment by Hirayama and Tarucha [25]. The black lines are boundaries, electrons are ejected from terminal 1, and terminal 2 is positively charged. In the ordinary drift-diffusive transport regime, the initial electron momentum is lost because of scattering before electrons reach the central region. This is defined by the momentum relaxation time

$$\tau_m = \frac{m \mu}{e}$$  \hspace{1cm} (1)

where $m^*$ is the electron effective mass, $\mu$ is the electron mobility, and $e$ is an elementary charge. Scattered electrons drift towards the positively charged terminal 2, resulting in the electron flow shown in Fig. 1(a). As transport becomes more ballistic, $\tau_m$ increases, and electrons maintain their momentum for a longer period, resulting in a larger percentage reaching terminal 3 instead of terminal 2. In the ballistic limit (where $\tau_m$ is much larger than the travel time between terminal 1 and 3), electron paths follow a billiard model [26], and all electrons travel from terminal 1 to 3, shown in Fig. 1(b). In the region between ballistic and drift-diffusive transport, electrons behave quasi-ballistically and can be guided by applied potentials and geometrical boundaries. In semiconducting 2DEG systems, $\tau_m$ and $l_e$ strongly depend on temperature, so that devices that are ballistic at liquid nitrogen or liquid helium temperatures can become drift-diffusive at room temperature. A number of novel devices have been proposed using 2DEG ballistic transport [23][27][28]. Our ballistic nanoelectronics research concentrates on the T-Branch Junction (TBJ) and the Ballistic Deflection Transistor (BDT).

A. Ballistic T-Branch Junction (TBJ) Devices

TBJs are three-terminal devices patterned in a 2DEG [31]. An SEM image of a TBJ is shown in Fig. 2. TBJs exhibit a nonlinear input-output transfer function [31] with efficiency characterized by a ballisticity factor $\alpha/\alpha_0$, where $\alpha_0 = e^2/2\mu_e$ at $T = 0$ K and $\mu_e$ is the Fermi energy of the 2DEG [32].

We recently studied TBJs in a 2DEG InGaAs/InAlAs heterostructure by varying the device channel length $L$ and operating temperature $T$ [29]. Fig. 3 summarizes the experimental results in terms of $\alpha/\alpha_0$, which increases with decreasing $T$ and decreases as the $U/I$ ratio increases. Another interesting feature is that $\alpha/\alpha_0$ saturates at a nonzero value when $L >> l_e$; thus, the nonlinear ballistic effect is still observable even for long TBJs at high temperatures. The robustness of the TBJ’s nonlinear response is a unique advantage in terms of realizing room-temperature circuits.

B. Ballistic Deflection Transistors

The BDT [30] is a six-terminal coplanar structure etched into a 2DEG. The dimensions and materials allow electrons to travel quasi-ballistically at room temperature, guided by the central deflector and lateral gate potentials as shown in Fig. 4. The steering voltage is much smaller than that required for gate pinch-off. This low voltage, combined with the low capacitance of the 2DEG features, results in an estimated $f_T$ in the THz range [23]. An SEM of a BDT is shown in Fig. 5.

The room temperature measured response of the BDT across a range of differential gate voltages is shown in Fig. 6. The x-axis is the left gate voltage. The asymmetry between the left and right outputs is caused by process variation and slight unevenness in contact placement. A positive left gate and negative right gate voltage results in current gain through the left output branch; current gain is seen through the right output branch when the voltages are reversed. The overlapping response in Fig. 6 is a result of electrons being scattered into the incorrect output. Electron scattering into the pull-up channel accounts for less than 10% of the total current. Gate leakage, where electrons tunnel under the minimum width etching between the gate and channel, accounts for 20% of the total current. These leakage values will decrease as geometries and fabrication techniques are further refined.
One of the major challenges of creating logic with BDTs is the method of converting the output current of one device into the gate voltage of the next device. In the BDT, accumulated gate charge is not dissipated by switching the driving gate input; instead, the lack of driving current creates a high resistance path to VDD and ground through the driving device. One solution is to use a string of resistors between VDD and VSS, shown outside the dashed box in Fig. 7. The output current from the BDT affects the voltage division between VDD and VSS, and the node $V_{out}$ can then be applied to the next gate. More electrons on a node results in a lower voltage; thus, a flow of electrons represents logic ‘0’.

The empirical model in Fig. 7 uses two voltage-controlled current sources (VCCS) to recreate the output behavior from Fig. 4, with a leakage path between ground and the supply voltage represented by a resistor. To simplify the model, output symmetry is forced by inverting the right channel response about 0 V to create the left channel response.

The 2-input NAND gate can be used to construct any arbitrarily complex logic function. It is thus of immense importance for novel devices to achieve this function if they are to be used for general purpose computation. The NAND gate design shown in Fig. 8 functions as follows: the source is shown as the arrow entering the bottom channel of the left-hand BDT. The differential gate input $A$ guides electrons into the channel labeled $A = 1$ when gate $A$ is high and gate $\overline{A}$ is low, and into channel $A = 0$ in the opposite case. The differential gate input $B$ guides electrons in the central region to the channels labeled $B = 0$ or $B = 1$ similar to gate $A$. This results in a flow of electrons (representing a ‘0’) at the output $F$ only when $A$ and $B$ are high. For each other input combination, electrons are driven either to the left output channel or the top right output channel. This behavior results in the logic function $A$ NAND $B$. To maintain the differential output required to drive the next stage of gates, the left output channel and top right output channel must be connected to create the logic function $A$ AND $B$.

The empirical model in Fig. 9 is similar to Fig. 8, with a current-controlled voltage source (CCVS) allowing the VCCS instances in the right-hand BDT to react to the current exiting the left-hand BDT. This results in the waveform in Fig. 10, with the outputs delayed by 1 ns. As shown, the output $v_f$ switches between 0.15 V and -0.15 V, matching the input voltages. The -0.15 V state, logic low, is achieved only when...
both inputs are ‘1’. When both inputs are ‘0’, the 0.15 V state is reached. For the ‘01’ and ‘10’ cases, a slightly reduced state of 0.1 V is reached. Thus, the model represents a NAND gate.

Additional experiments planned with fabricated BDT's include testing the current-to-voltage converters to cascade devices, testing other Boolean logic gates, and using cross-coupled BDTs to create memory elements.

IV. CONCLUSIONS AND FUTURE TRENDS

The presented survey of devices reveals a number of exciting options with the potential of replacing conventional CMOS when further scaling becomes too costly. This wide variety of options may in fact introduce unexpected side effects, as there may not be a single “predominant” device in which resources are pooled, as was the case when CMOS emerged as the electronics standard of sorts. A number of traditional improvements may be combined, such as strained-Si channels with fiFET gates, to extend the reach of traditional devices; however, these improvements do not offer the same potential performance enhancement as non- conventional ideas such as molecular or optical computing. The ballistic deflection transistor provides an interesting middle ground, pairing semiconductor technology with non- traditional device and circuit design, and is shown to offer a number of properties which could make it an excellent choice for the next era of computing. Continued work on improving the device and circuit design should help further differentiate it from other alternatives.

ACKNOWLEDGMENT

The authors would like to acknowledge contributions of the BDT Research Group at the University of Rochester (UR) and University of Massachusetts at Lowell. This work is sponsored by NSF Grant ECCS0609140, ONR Grant N00014-08-1-0093, AFSOR Grant FA9550-07-1-0032, and a NYSTAR Grant to the UR CAT-NEIS. TBJ and BDT fabrication was performed at the Cornell Nanofabrication Facility (CNF), a member of the NSF sponsored NNI network.

REFERENCES