REAL SPACE TRANSFER DEVICES IN SOI

Serge Luryi

ECE Department and NY State Center for Advanced Sensor Technology
State University of New York at Stony Brook
Stony Brook, NY 11794-2350, USA

Abstract. Three-terminal real-space transfer (RST) devices employ charge injection of hot electrons over a potential barrier into an independently contacted second conducting layer. The first conducting layer is the usual transistor channel, where electrons are heated by the source to drain field. The high RST injection is accompanied by a strong negative differential resistance (NDR) in the source-drain circuit. Because of the NDR property, RST transistors can be used, e.g., as voltage-controlled oscillators for wireless applications. Efficiency of the RST has been amply demonstrated in III-V heterostructures, where bandwidths in excess of 100 GHz have been demonstrated. Silicon implementations of RST have been so far limited to the Si/GeSi heterosystem. In this paper I discuss possible implementation of RST transistors in Si, using low-barrier dielectrics, such as zirconium silicates, which are currently under intensive investigation in the context of the high-κ dielectrics program. SOI technology offers special advantages both for the implementation of RST devices and their likely application.

INTRODUCTION

The Charge Injection Transistor (CHINT) is a three terminal semiconductor device based on real-space transfer of hot electrons. The typical device structure is illustrated in Fig. 1.

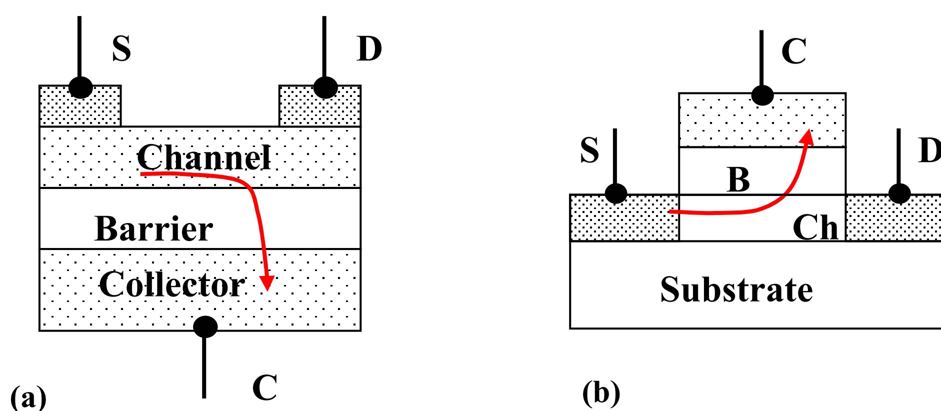

Figure 1. Schematic diagram of real space transfer transistors. (a) Bottom collector configuration, used in many heterojunction implementations of the CHINT; (b) top collector configuration, readily amenable to SOI implementation.

Figure 1a shows the conventional "bottom-collector" structure. Heated by the electric field applied between electrodes S and D, channel electrons become hot and undergo emission over the barrier into the second, independently contacted, conducting layer. The contact C to the second conducting layer is often referred to as collector contact and the layer itself as collector layer. The contacts S and D to the hot-electron channel are similar to source/drain transistor contacts. Figure 2a illustrates a top-collector configuration, which has certain advantages in terms of the speed of operation but has been rarely used, being more difficult to implement in semiconductor heterojunctions.

Since the first disclosure of the CHINT idea by Kastalsky and Luryi [1] and the first demonstration of the device in GaAs/AlGaAs heterostructures [2], CHINT has been implemented in a variety of semiconductor heterostructure systems, including InGaAs/InAlAs, InGaAs/InP, strained-layer InGaAs/AlGaAs/GaAs, as well as GeSi/Si material systems. Powerful logic devices have been proposed and demonstrated based on special symmetries of charge injection [3]. Light emitting version of the CHINT, which results when the collector layer is implemented in a complementary material, has also been demonstrated and extensively studied [4]. These devices show powerful logic functions in the dependence of output light on electron-heating voltages regarded as input signals. References to the history and the development of CHINT can be found in a recent encyclopedia article [5].

The usefulness of CHINT derives from (i) efficient control of the injection current I_C by the heating voltage V_D (relative to source S) which enables ultra high-speed operation of the device as an amplifier, (ii) symmetry of charge injection with respect to the polarity of voltage applied to the heating electrodes, which enables powerful logic functions, and (iii) the very strong negative differential resistance observed in the channel circuit, i.e. in the current voltage characteristic $I_D (V_D)$ at a fixed V_C. Property (iii) allows one to use CHINT as a voltage controlled oscillator of ultra-high bandwidth. In this embodiment, the device is sometimes referred to as the Negative Resistance FET (NERFET) and both names are used interchangeably.

Of particular interest for the present discussion is the demonstration [6] of ultra-high speed operation top-collector InGaAs/InP CHINT. This device, implemented with conventional optical lithography and wet etching at 1 μm design rules, showed a much higher bandwidth (f_T of 115 GHz) than those which had been previously achieved in bottom-collector configurations.

As evident from Fig. 1b, the structure of a top-collector CHINT is similar to that of a field-effect transistor. A top collector device can, in principle, be operated in both the CHINT mode and the FET mode. In the FET mode, the collector of CHINT plays the role of a controlling (gate) electrode. Maezawa and Mizutani [7] have studied experimentally the two modes of operation of the same device — the FET (gate controlled channel conduction) mode and the CHINT mode (drain controlled hot-electron injection into the collector) — and concluded that in a device with the channel length of 1 μm the CHINT mode was approximately three times faster. This conclusion was also supported by Monte Carlo simulations by Akeyoshi et al [7].

Compound semiconductor heterostructures, because of their typically low barrier height at the heterojunction, can naturally be operated in both the CHINT and the FET modes. Silicon MOS transistors, in contrast, can only be operated in the FET mode, since hot-electron injection into the gate occurs only at a very low level owing to the large band offset between Si and SiO_2. The barrier height for hot-electron emission at

the Si/SiO_2 interface is above 3 eV and hot-electron currents rarely exceed a few pA. This is, of course, fortunate for the FET operation, because no Si circuit designer likes leaky gates. Hot electron injection is only a nuisance for the conventional silicon MOSFET, but it may become a current to be reckoned with in the insulated gate devices with lower-barrier dielectrics.

SUITABLE DIELECTRIC MATERIAL

In recent years, a large worldwide effort has been devoted to the search of suitable high-permittivity dielectrics that could replace the trusted SiO_2. It is certainly an extremely challenging task, because one can say that the primary reason for the undisputed triumph of silicon MOS technology over all other transistor technologies in the 20th Century lies in a long list of fortunate properties of the thermally grown oxide on Si (100) surface. One property lacking in this list, however, is a high dielectric permittivity. Higher-permittivity gate dielectrics would lead to higher-speed circuits by enhancing the ratio of the useful (gate-to-channel) capacitance to most of the parasitic capacitances. Whether or not the 21st Century will find a replacement for SiO_2 in Si CMOS circuits is very much an open question at this time, but smart money is not ignoring this possibility.

Some of the candidate dielectrics that have been considered for the high-permittivity insulated gates demand a troublesome penalty for the enhancement of permittivity, which occurs at the expense of a significant narrowing of the dielectric bandgap. What is particularly disturbing for insulated-gate transistor applications is the situation when most of the narrowing occurs in the conduction band. Thus, according to recent calculations [8], the band offset for Si electrons at $Zr_{\{x\}}Si_{\{1-x\}}O_2$ interface becomes lower than 0.5 eV for x=0.18. While disagreeing with the calculations [8] for zirconium silicates, Roberts [9] gives his own table of calculated band offsets of oxides on Si, where one finds $\Delta E_C = 0.3\,\text{eV}$ for Ta_2O_5, $\Delta E_C = 0.8\,\text{eV}$ for $BaZrO_3$ and a negative offset for $BaTiO_3$. The low-offset zirconium silicate (or a similar material) would be an ideal barrier layer for Si CHINT.

Roberts [9] further gives an excellent exposition of the physics that controls the conduction band offset between silicon and a high-κ dielectric. Most of these dielectrics are metal oxides. Their dielectric gap is usually well known, while the asymmetry of the band alignment with silicon is often not. It can be measured by internal photoemission involving transitions between the valence bands of silicon and the oxide conduction band. Such measurements are not very precise and one often relies on calculated values. It is believed that the band alignment is controlled by charge transfer between the partially filled dangling bonds of the semiconductor and the oxide. The highest filled energy level of the dangling bonds in each material, called the charge neutrality level or CNL, plays the role of the chemical potential. In equilibrium, charge transfer at the interface aligns the CNLs of adjacent materials. Since the CNL of Si is low in the gap, only about 0.2 eV from the valence band, whereas the CNL of most oxides is high in the gap, the alignment of these materials tends to give low conduction-band offsets. One man's worry is another man's joy!

PROPOSED DEVICE STRUCTURE

The proposed silicon CHINT structure is illustrated in Fig. 2. The device looks much like a MOSFET, but the principle of operation is entirely different. Channel electrons, accelerated by the electric field (directed into the page in Fig. 2) and equilibrating their energy via electron-electron collisions, become "hot" in the sense that their energy distribution, while remaining approximately maxwellian, is characterized by an elevated effective temperature $T_e > T$. Hot electrons undergo thermionic emission (RST) over the low-barrier dielectric into the independently contacted collector, here assumed implemented as a heavily doped polysilicon layer.

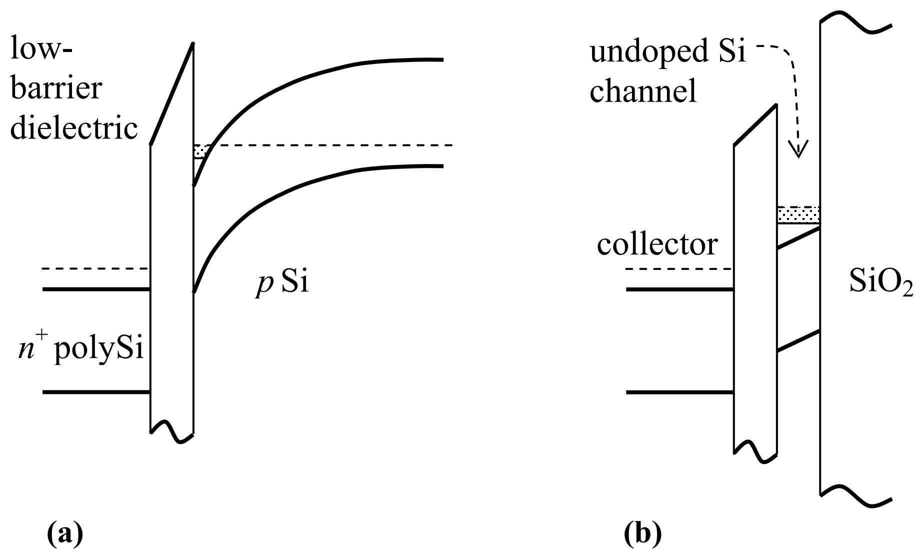

Figure 2. Schematic cross-section of the proposed silicon CHINT with a low-barrier (high-κ) dielectric. (a) Implemented on a Si substrate; (b) SOI implementation.

The hot electron ensemble employed in CHINT has quasi-equilibrium energy and momentum distributions. This makes it very much different from and much more robust than the so-called ballistic electron ensembles [10] often employed in other hot-electron devices. The electron temperature T_e can be determined, at least qualitatively, from a simple energy balance equation. Analysis of the electron heating suggests an important advantage of SOI structure arising due to the higher mobility in undoped silicon channels [11].

WHY SOI ?

For the higher RST efficiency it is important that hot electrons are compressed in a narrow layer in the vicinity of the barrier. Therefore, when contemplating bulk silicon transistor structures (Fig. 2a) for the CHINT application, one must assume a relatively heavy substrate doping, $N_A \cong 10^{18}$ cm^{-3} or higher. This does not only introduce a large parasitic input capacitance (remember that drain is the input electrode in CHINT) but it

may also adversely affect the onset of RST process. Indeed, consider the energy balance equation assuming that the dominant energy loss mechanism is associated with the emission of optical phonons (this is indeed true only at relatively low T_e) and is characterized by a relaxation time constant τ_E,

$$k\,(T_e - T) \;=\; e\mu F^2 \tau_E\,. \qquad (1)$$

For a given electric field in the channel, $F \cong V_{SD} / L_{CH}$, the excess carrier temperature is proportional to channel mobility. Since in the confined undoped layers the value of μ can be much higher than in doped bulk transistor channels [11], it should be much easier to establish the hot-electron RST with thin SOI channel layers. Optical phonon emission is no longer important at higher T_e when the dominant loss term in the energy balance is due to the energy flux carried away by hot electrons leaving the channel over the barrier [12]. Once the RST process sets in, the electron temperature rises very high, $T_e > 1{,}000$ K and the electrostatic configuration in the channel becomes unstable with respect to formation of hot-electron domains [13]. In this regime, the carrier mobility is no longer relevant, since the velocity becomes saturated.

CHARACTERISTICS

Lacking experimental data for silicon CHINT with an oxide barrier, we can learn from the earlier experience with compound semiconductor heterostructures. The most impressive characteristics have been obtained in InGaAs/InAlAs heterostructures lattice matched to InP. In these materials, the barrier height is determined by the conduction band discontinuity between $In_{0.53}Ga_{0.47}As$ and $In_{0.52}Al_{0.48}As$, which is 0.5 eV.

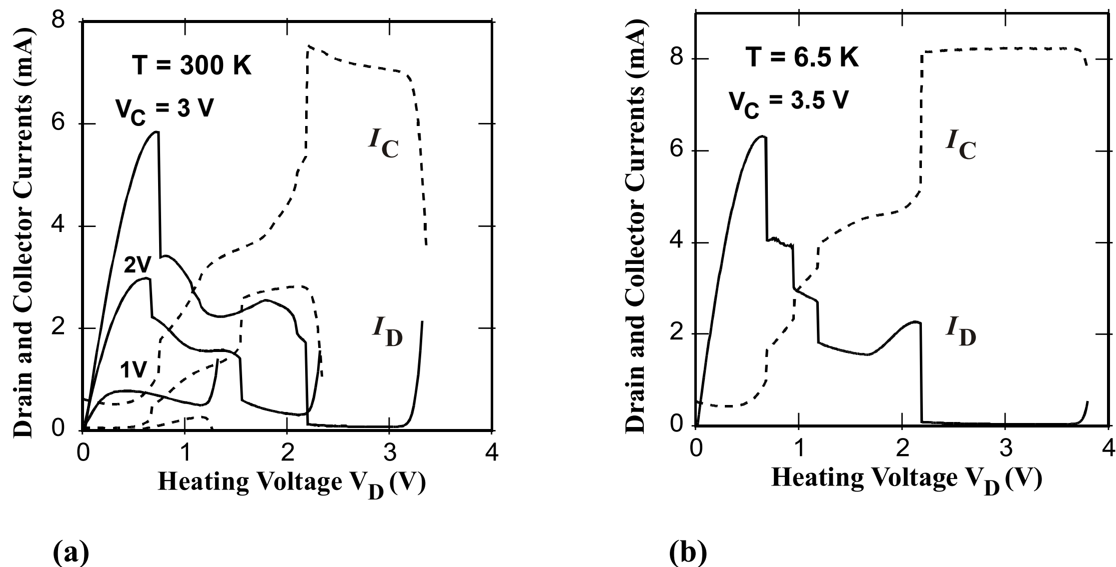

Figure 3. Family of current-voltage characteristics at room (a) and cryogenic (b) temperatures of an InGaAs/InAlAs heterostructure CHINT/NERFET device with a 25-μm wide channel [14]. Solid lines show current in the channel circuit while dashed lines show the collector current.

A family of CHINT/NERFET characteristics is shown in Fig. 3. The RST diverts a fraction of channel current into the collector. For a sufficiently high collector bias V_C this fraction tends to nearly 100%, resulting in a strong negative differential resistance (NDR) in the drain circuit. When biased into the NDR regime, the device is unstable and can be used as a generator of oscillations. With an *LC* load the single device can be used as a voltage-controlled oscillator; the oscillations can be turned on and off by V_C . In principle, with the parasitic capacitances eliminated by SOI, the CHINT bandwidth is limited only by hot-electron equilibration, which is characterized by the energy relaxation time τ_E . In silicon, τ_E is in the subpicosecond range.

Because of the very high mobility of electrons in InGaAs, the mobility is no longer an important parameter. The NDR curves in Fig.3 are quite similar for both 300K and 6.5K. This happens because the electron velocity is saturated before the onset of the NDR even at room temperature. The peak-to-value (PTV) ratio of the NDR seems to improve at lower *T* but this is only due to higher V_C bias. The 6.5K data in Fig.3b correspond to a PTV of 210 but still higher values of PTV ≅ 8,000 were subsequently observed at 300K [15]. The quest for highest PTV turned out rather unexpectedly when practically infinite ratios and then even negative values of the valley I_D were observed. Explanation of this strange behavior was given by Luryi and Pinto [13] in terms of the formation of a hot-electron domain in the channel, with its electrostatic potential exceeding that of the drain. Simulations [16] revealed the existence of multiply connected current-voltage characteristics $(I_D - V_D)$ with many loops and folds – which qualitatively accounted for the observed sharp non-linearities. These features arise due to interplay between the fast RST process and the relatively slower dielectric relaxation in the channel. An example of the simulated behavior is shown in Figs. 4a,b.

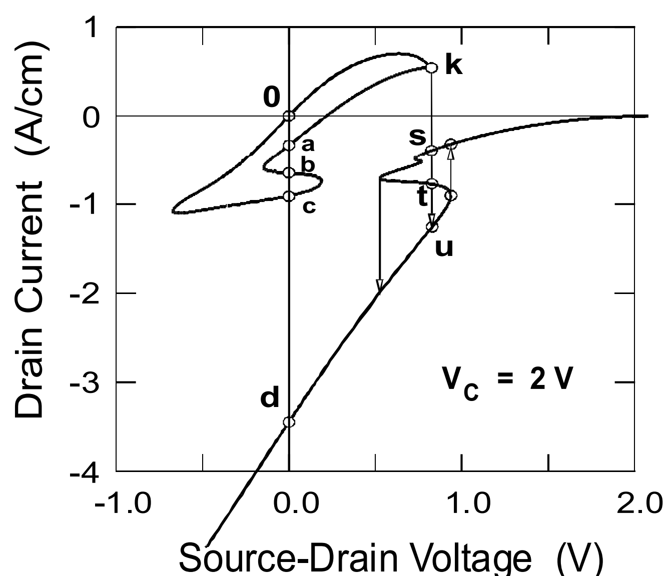

Figure 4a. Multiply connected current voltage characteristics of a CHINT, obtained by simulations [13,16]. The device was assumed to possess silicon-like velocity-field characteristics with $v_{sat} = 10^7$ cm/s. Channel length 5 μm, barrier height 0.5 eV, barrier thickness 0.2 μm.

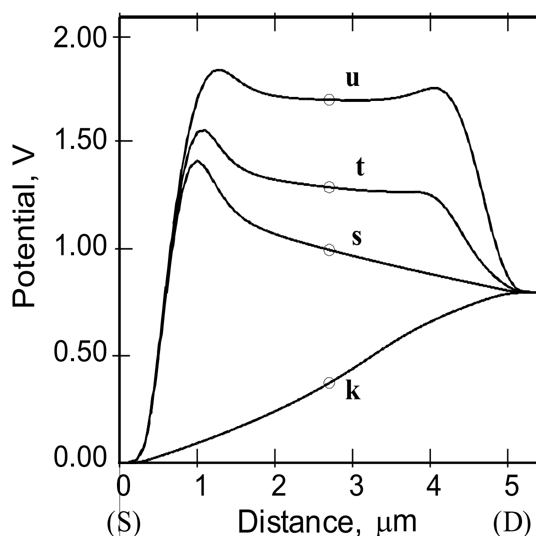

Figure 4b. The channel potential profile *V(x)* in states, u, t, s, k of the current-voltage diagram. All four states correspond to exactly the same applied bias, V_C =2V, V_D≈0.82V.

The diagram in Fig.4a is not an observable current-voltage characteristic; it rather represents a phase portrait of all stationary states of the device in the $V_D - I_D$ plane. However, it does allow tracing the observable characteristic. Thus, as we increase the drain bias from 0 we follow the "normal" characteristic until point *k* when a global redistribution of charge and field in the device is forced. Such redistribution is reminiscent of a phase transition, and the device switches to point *u*, corresponding to the highest collector current. Two hot electron domains are formed, one near the source, the other near the drain, where the channel is depleted of electrons. The collector field thus remains unscreened and the local channel potential dips below that at the drain, resulting in a negative drain current.

The electrostatic potential in the channel in the direction from the source to the drain is plotted in Fig. 4b for the four states corresponding to exactly identical external bias configuration. Three of these states are stable and, in principle, observable; state *t* is unstable with respect to small perturbations.

Hot electrons domains form when the finite supply rate of electrons to a "hot spot" is exceeded by the RST flux from that spot, the supply being limited by the electronic transport along the channel. Domains can form even at zero source-drain bias, $V_D = 0$, cf. Fig.4a. Depending on the bias history, the device can be in either of two stable states, *o* and *d*. State *o* is "FET-like" with the collector field screened by channel electrons, and state *d* is "anomalous" (similar to *u*) with the channel depleted. Three other states, *a, b* and *c,* are unstable with respect to small perturbations and if perturbed rapidly evolve into one of the two stable states. Time-dependent simulations [17] revealed that the characteristic switching time roughly corresponds to the time of electron travel from the S and D contacts to the middle of the channel. Besides the geometry of the device and the barrier height, the saturated velocity v_{sat} is the most relevant parameter that governs switching processes in RST transistors.

The simulations [13,16,17] were in some sense more relevant to an implementation in silicon (which was at the time purely hypothetical) than in GaAs, since they ignored the intervalley transfer process that dominates high-field transport in GaAs and related compounds. It also slows down the speed of RST transistors in these materials.

SPEED

Silicon RST transistors are expected to be faster than those implemented in GaAs, especially in SOI implementations that eliminate much of the parasitic capacitance. Monte-Carlo simulations by Kizillyalli and Hess [18] indicate that the fast RST process in GaAs is considerably slowed down by the much slower momentum-space transfer. The record speed in RST transistors (extrapolated current-gain cut-off of 115 GHz) was demonstrated [6] in InGaAs/InP heterojunctions with InP barriers, where the momentum transfer is suppressed because the barrier height for RST ($\Delta E_C \approx$ 0.25 eV) is less than half of the intervalley separation in the InGaAs channel ($E_{L\Gamma} = 0.55$ eV). We expect that the small-signal bandwidth of an SOI CHINT should be limited only by the energy relaxation time in Si. Large-signal switching process should be somewhat slower (FET-like) as it is still limited by electron transport along the channel.

SYMMETRY AND LOGIC

A fundamental property of RST transistors is the symmetry equivalence [17] between the internal states of the device $\mathbf{S}[\boldsymbol{V_D}, \boldsymbol{V_C}]$ at different external bias configurations:

$$\mathbf{S}[\boldsymbol{V_D}, \boldsymbol{V_C}] \quad \Leftrightarrow \quad \mathbf{S}[-\boldsymbol{V_D}, (\boldsymbol{V_C} - \boldsymbol{V_D})]. \tag{2}$$

This correspondence follows from the reflection symmetry in the plane normal to the source-drain direction that cuts the channel in the middle. Thus the output current is invariant under an interchange of the input voltages $\boldsymbol{V_S} \leftrightarrow \boldsymbol{V_D}$ and the device exhibits an exclusive OR dependence of the collector current on the input voltages, regarded as binary logic signals. More powerful logic functionality is obtained in a RST device [3] with three input terminals. In such a device, called the ORNAND gate, the output current has either OR or NAND dependence on the voltage applied to any two of the three electrodes, depending on the voltage applied to the third electrode.

Of particular interest are optoelectronic versions of the RST logic device, which obtain when the collector is implemented as a complementary direct-gap semiconductor layer. Hot electrons injected in a separately contacted *p*-type collector produce light of intensity proportional to the RST current. Mastrapasqua et al [4] demonstrated an ORNAND with excellent optical logic characteristics. It would be very exciting if such a device could also be implemented in silicon. Can one possibly deposit over the low-barrier-oxide dielectric layer in Fig. 2 a crystalline collector layer comprising a *p*-type direct-gap semiconductor?

CONCLUSION

I have argued that the currently fashionable program aimed at the development of high-κ gate dielectrics for VLSI applications, may have an unexpected by-product that does not suit the originally intended application but is still important in its own right. Low-barrier dielectrics may provide the material base for silicon implementation of real space transfer transistors and open a variety of useful circuit applications. I have discussed the expected characteristics and limitations of such devices that naturally lend themselves to SOI implementation. Silicon based RST transistors employing low-barrier oxide in the collector structure should perform in many respects similar to and in some respects (notably, speed) better than previously explored compound semiconductor devices. The SOI technology may offer the best route to achieve the ultimate performance of Si RST transistors.

ACKNOWLEDGEMENTS

I am grateful to Sorin Cristoloveanu and Marco Mastrapasqua for useful discussions. This work was partially supported by the NY Center for Advanced Technology in Sensor Systems, funded by the NY State Office of Science, Technology and Academic Research (NYSTAR).

REFERENCES

[1] A.Kastalsky and S.Luryi, "Novel real-space transfer hot-electron transfer devices", *IEEE Electron Device Lett*. **EDL-4**, pp. 334-336 (1983).
US Pat. **4,903,092** "Real space electron transfer device using hot electron injection" (filed 1983, issued 1990).

[2] S.Luryi, A.Kastalsky, A.C.Gossard, R.Hendel, "Charge injection transistor based on real-space hot-electron transfer", *IEEE Trans. Electron Devices* **ED-31**, pp. 832-839 (1984); "A field-effect transistor with a negative differential resistance", *IEEE Electron Device Lett.* **EDL-5**, pp. 57-60 (1984).

[3] S.Luryi and M.R.Pinto, US Pat. **4,999,687** (filed 1990, issued 1991).
S.Luryi, P.M.Mensz, M.R.Pinto, P.A.Garbinski, A.Y.Cho, D.L.Sivco, "Charge injection logic", *Appl. Phys. Lett.* **57**, 1787-1789 (1990).
M.Mastrapasqua, C.A.King, P.R.Smith, M.R.Pinto, "Functional devices based on real space transfer in {Si/SiGe} structure", *IEEE Trans. Electron Devices* **ED-43**, pp. 1671--1677, 1996.

[4] S.Luryi, "Light emitting devices based on the real-space-transfer of hot electrons", *Appl. Phys. Lett.* **58**, pp. 1727-1729 (1991); US Pat. **5,223,723** (filed 1990, refiled 1991; issued 1993).
M.Mastrapasqua, S.Luryi, G.L.Belenky, P.A.Garbinski, A.Y.Cho, D.L.Sivco, "Multi-terminal light emitting logic device electrically reprogrammable between OR and NAND functions", *IEEE Trans. Electron Devices* **ED-40**, pp. 1371-1377 (1993).

[5] S.Luryi and M.Mastrapasqua, "Charge Injection Devices", Wiley Encyclopedia of Electrical and Electronic Engineering, John G. Webster, Editor (John Wiley @ Sons, Publishers, ISBN: 0-471-13946-7) Volume 3, pp. 258-263 (1999).

[6] G.L.Belenky, P.A.Garbinski, S.Luryi, P.R.Smith, A.Y.Cho, R.A.Hamm, and D.L.Sivco, "Microwave studies of self-aligned top-collector charge injection transistors", *1993-IEDM Tech. Digest*, pp. 423-426 (1993); *Semicond. Sci. Tech*. **9**, pp. 1215-1219 (1994).

[7] K.Maezawa and T.Mizutani, "High-frequency characteristics of charge-injection transistor-mode operation in AlGaAs/InGaAs/GaAs metal-insulator-semiconductor field-effect transistors", *Jpn. J. Appl. Phys*. **30**, No.6, pp 1190-1193 (1991); T.Akeyoshi, K.Maezawa, M.Tomizawa, and T.Mizutani, "Monte Carlo study of charge injection transistors (CHINTs)", *Jpn. J. Appl. Phys*. **32**, pp.26-30 (1993).

[8] A.Kawamoto, K.Cho, P.Griffin, and R.Dutton, "First principles investigation of scaling trends of zirconium silicate interface band offsets", *J. Appl. Phys.* **90**, pp. 1333-1341 (2001).

[9] J.Robertson, "Electronic structure and band offsets of high-dielectric-constant gate oxides", MRS Bulletin, http://www.mrs.org/publications/bulletin (March 2002).

[10] S.Luryi, "Hot electrons in semiconductor devices", in *Hot Electrons in Semiconductors*, ed. by N.Balkan, (Oxford University Press, ISBN 0-19-850058-0) pp. 385-427 (1998).

[11] D.Esseni, C.Fiegna, and M.Mastrapasqua, "Ultra-thin single- and double-gate MOSFETS for future ULSI applications: Measurements, simulations, and open issues", in *Future Trends in Microelectronics: The Nano Millennium*, ed. by S.Luryi, J.M.Xu, and A.Zaslavsky, Wiley Interscience, New York (2002) pp. 63-78.

[12] A.A.Grinberg, A.Kastalsky, and S.Luryi, "Theory of hot electron injection in CHINT/NERFET devices", *IEEE Trans. Electron Devices* **ED-34**, pp. 409-419 (1987).

[13] S.Luryi and M.R.Pinto, "Broken symmetry and the formation of hot-electron domains in real-space transfer transistors", *Phys. Rev. Lett*. **67**, pp. 2351-2354 (1991).

[14] P.M.Mensz, S.Luryi, A.Y.Cho, D.L.Sivco, F. Ren, "Real space transfer in three-terminal InGaAs/InAlAs/InGaAs heterostructure devices", *Appl. Phys. Lett*. **56**, pp. 2563-2565 (1990).

[15] P.M.Mensz, P.A.Garbinski, A.Y.Cho, D.L.Sivco, and S.Luryi, "High transconductance and large peak-to-valley ratio of negative differential conductance in three-terminal InGaAs/InAlAs real-space transfer devices", *Appl. Phys. Lett*. **57**, pp. 2558-2560 (1990).

[16] M.R.Pinto and S.Luryi, "Simulation of multiply connected current-voltage characteristics in charge injection transistors", *1991-IEDM Tech. Digest*, pp. 507-510 (1991).

[17] S. Luryi and M.R. Pinto, "Symmetry of the real-space transfer and collector-controlled states in charge injection transistors", *Semicond. Sci. Tech*. 7, No. 3B, pp. B520-B526 (1992).

[18] I.C. Kizilyalli and K. Hess, "Physics of real-space transfer transistors," *J. Appl. Phys*. **65**, pp. 2005-2013 (1989).