A high-speed silicon-based few-electron memory with metal–oxide–semiconductor field-effect transistor gain element

Andrew C. Irvine, Zahid A. K. Durrani, and Haroon Ahmed[a)]
Microelectronics Research Centre, University of Cambridge, Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, United Kingdom

(Received 2 August 1999; accepted for publication 7 March 2000)

The design and operation of a single-electron transistor-controlled memory cell with gain provided by an integrated metal–oxide–semiconductor field-effect transistor is described. The field-effect transistor has a split gate, the central section of which is addressed by the single-electron transistor. This design effectively reduces the size of the cell's memory node such that the memory states are represented by a difference of only a few hundred electrons, while obtaining output currents in the microamp range. The hysteresis loop in the field-effect transistor current shows clear step discontinuities which arise from Coulomb oscillations in the single-electron transistor's drain-to-source current. The cell operates with write times as short as 10 ns and voltages of less than 5 V at 4.2 K, and operation persists to 45 K. The cell design is compact and the memory is fabricated in silicon-on-insulator material by processes that are compatible with current silicon fabrication methods. © *2000 American Institute of Physics.* [S0021-8979(00)00912-9]

I. INTRODUCTION

The progressive miniaturization of classical silicon dynamic random-access memory (DRAM) cells is ultimately limited by the need for the stored charge on the memory node capacitor to provide an output current above the noise margin during the cell's ''read'' cycle. The reduction of device dimensions requires the reduction of stored electron number, and for this purpose it is necessary to develop an alternative memory cell design with gain as well as scalability. The Coulomb blockade principle has the potential to provide suitable carrier control down to the single-electron limit, and Coulomb blockade devices imply small device dimensions, making such devices inherently suitable for high-density integration. However, a SET typically has subunity gain, and for practical purposes it is necessary to develop an integrated high-gain cell for the sense architecture. Performance levels must be at least comparable with classical devices, and compatibility with well-established fabrication procedures is imperative. Several silicon-based memory cells utilizing single-electron charging effects have been reported.[1–4] The device of Yano *et al.*[1] uses a gate electrode to induce charge onto isolated grains in a granular silicon thin film, thus modifying the current along a percolation pathway through the film. Tiwari *et al.*[2] employ a control gate to induce tunneling from the channel of a field-effect transistor onto nanocrystalline islands in the gate oxide, while Nakajima *et al.*[3] use similar principles, having instead a lithographically defined island as the floating gate. In all cases, room-temperature, long-retention-time operation was achieved, but as a consequence of high tunnel resistances, the write and/or erase times in each case are long, making the devices potentially more suitable for flash memory rather than high speed random-access memory applications. In addition, the latter two devices operate by using carrier tunneling through the gate oxide of the field-effect transistor, leading to degradation of the oxide and to reduced durability. A high-speed crystalline silicon nanowire-based memory has been reported by Stone and Ahmed,[4] using single-electron transistor sensing, but this has low gain. A similar cell incorporating a metal–oxide–semiconductor field-effect transistor (MOSFET) as a gain element was subsequently demonstrated by Durrani *et al.*[5] However in that device the memory states were represented by a difference of $\sim 10^6$ electrons and the full advantages of the SET were not realized.

In this work, we investigate the operation of a silicon-based, single-electron-transistor-controlled memory cell with 10 ns write/erase time and high gain. The memory cell uses a highly doped crystalline silicon nanowire integrated with a MOSFET. A split gate on the MOSFET minimizes the memory node size, enabling operation with memory states defined by a difference of a few hundred electrons on the memory node. The cell may be operated at voltages of less than 5 V. Hysteresis in the MOSFET current confirms memory operation and steps observed in this current are attributed to Coulomb oscillations in the single-electron transistor. The memory operation persists to 45 K.

II. CELL DESIGN AND FABRICATION

We describe the operating principle of a silicon-on-insulator (SOI) memory cell of the type described previously[5] with reference to the schematic drawing in Fig. 1(a). The device consists of a silicon nanowire single-electron transistor[6] integrated with a MOSFET, the latter having its channel immediately below the buried oxide layer. The memory node acts both as an electrode of the SET nanowire and as the central gate of the MOSFET, and is thus capacitively coupled to the MOSFET channel. In this way,

[a)]Author to whom correspondence should be addressed. Electronic mail: ha10@phy.cam.ac.uk

 © 2000 American Institute of Physics

Downloaded 24 Oct 2005 to 129.169.10.56. Redistribution subject to AIP license or copyright, see http://jap.aip.org/jap/copyright.jsp

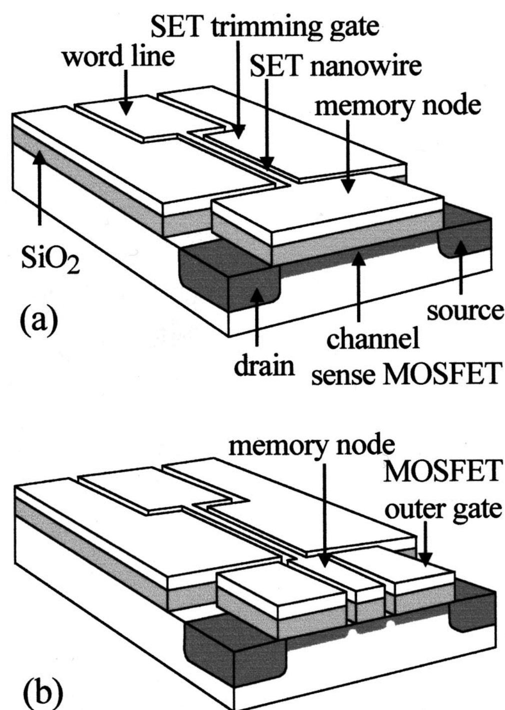

FIG. 1. Memory cells (schematic, not to scale) with integrated silicon-on-insulator nanowire single-electron transistor and metal–oxide–semiconductor field-effect transistor: (a) large-area single-gate structure (Ref. 5). (b) split-gate structure with small memory node.

the SET nanowire can be biased using voltages applied to the word line at one end and/or the MOSFET at the other end, and charge can flow onto the memory node only via the SET nanowire. Under Coulomb blockade conditions, an idealized nanowire has a zero-current gap of width $2V_c$ centered at zero voltage across the SET ($V_{\mathrm{SET}}=0$). Thus the potential drop across the nanowire may be maintained at a stable value between $-V_c$ and $+V_c$. The precise value of this potential difference depends on the cell's biasing history. If at any time the voltage at the word line and/or the MOSFET is changed such that the nanowire is biased outside the Coulomb gap, charge will flow to the memory node until the blockade condition is restored. The steady-state memory-node potential will then differ from the word-line potential by $\pm V_c$, and the sign of this difference will depend on the polarity of the change applied to the word-line voltage and/or MOSFET. Hence for a given biasing condition there are two memory states corresponding to the two states of memory-node potential. These states may be detected as different levels of drain–source current in the MOSFET. Figure 1(b) shows how the memory-node size, and therefore the stored charge, may be reduced for a given drain–source separation by the use of a split MOSFET gate. The outer-gate electrodes are used to ensure that the MOSFET channel is extended in the regions between the memory node or central gate and the drain and source implants. During cell operation, the MOSFET conductance is controlled by the gating effect of the memory node. The separation between outer gates and the memory node must be sufficiently small to ensure adequate continuity of the channel from drain to source. The outer gates can also be employed to close the channel, allowing the cell to be deselected in memory array applications.

The samples were fabricated from lightly doped p-type bonded SOI material which has a thermally grown buried oxide layer ~40 nm thick and a top silicon layer ~40 nm thick. The top silicon was implanted uniformly with 35 kV arsenic to a dose of $8\times10^{13}\,\mathrm{cm}^{-2}$, giving a mean n-type doping level of $1-2\times10^{19}\,\mathrm{cm}^{-3}$, well above the metal–insulator transition threshold at cryogenic temperatures. The MOSFET characteristics were defined by subsequent implantation of dopants into the substrate through the top silicon and buried oxide layers. The MOSFET channel doping was obtained by a uniform double implant of boron ($2\times10^{12}\,\mathrm{cm}^{-2}$ at 55 kV and $5\times10^{12}\,\mathrm{cm}^{-2}$ at 75 kV). Patterned optical resist masking was then used to define the heavily doped n-type source and drain contacts (75 kV phosphorus, $2\times10^{15}\,\mathrm{cm}^{-2}$) and the heavily doped p-type substrate contact (50 kV boron, $1\times10^{15}\,\mathrm{cm}^{-2}$). The drain–source separation in all cases was ~8 μm.

The fabrication scheme was similar to that described elsewhere in detail.[5] The SET and MOSFET gate patterns were written in poly-methyl-methacrylate electron-beam resist, and a $CF_4/SiCl_4$ reactive ion plasma was used to etch the top silicon as far as the buried oxide. The SET nanowires defined in this way were 1 μm in length and typically 50 nm wide, and the separation between the memory node and the adjacent outer MOSFET gates was ~70 nm. The samples were then thermally oxidized at atmospheric pressure under dry oxygen flow at 1000 °C for 15 min in order to thin the silicon and to passivate the nanowires electrically. The oxidation process consumed approximately 10 nm of silicon at free surfaces. Contacts were made by using a standard complementary MOS metallization process. A 500-nm-thick SiO_2 dielectric was sputter deposited and subsequently H_2 passivated at atmospheric pressure by annealing for 15 min at 400 °C in an ambient of forming gas (5% H_2 in N_2). Sloping-profile contact holes were selectively wet etched through the SiO_2 dielectric, and a $Ti/Al_{0.99}Si_{0.01}/Ti$ blanket metallization (80 nm/500 nm/80 nm) was sputter deposited. The excess metal was removed by $SiCl_4$ reactive-ion etching, after which the contacts were sintered for 15 min at 200 °C in forming gas.

Electrical measurements were carried out using an HP4156 parameter analyzer. An HP8110A pulse generator was also employed for pulsed measurements. Measurements at 4.2 K were principally made with the samples immersed in liquid helium, while measurements at higher temperatures were made in an Oxford Instruments continuous-flow cryostat.

Selected device structures prior to oxidation and dielectric deposition are shown in the scanning-electron micrographs of Figs. 2–4. Figure 2 shows the implementation of the split-gate memory cell configuration, using a MOSFET of ~8 μm drain–source separation, and Fig. 3 shows the memory node and nanowire of the same device in more detail. In this case, the memory node has a 1 μm$\times$1 μm area, with a 70 nm separation between the memory node and the outer gates on either side. Other electrodes, such as the SET trimming gates, were isolated from the principal MOSFET

Downloaded 24 Oct 2005 to 129.169.10.56. Redistribution subject to AIP license or copyright, see http://jap.aip.org/jap/copyright.jsp

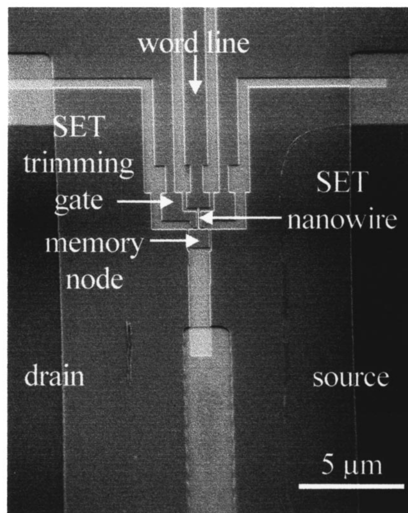

FIG. 2. Scanning-electron micrograph of memory cell prior to oxidation.

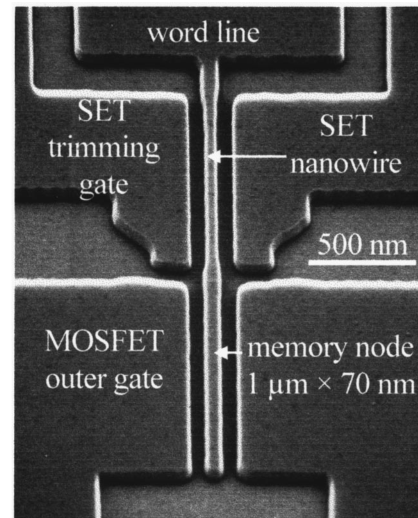

FIG. 4. Micrograph of an optimized memory cell with memory node area 1 μm$\times$70 nm.

channel by a trench of width 200–500 nm where practicable, to optimize the controllability of the MOSFET current by the memory node. Finally, Fig. 4 shows a memory cell of memory node size 1 μm$\times$70 nm. The length-to-width ratio of the memory node in this structure permits the combination of few-electron operation with high MOSFET transconductance. This device has a memory node area smaller by a factor of ~3000 than that of the single-gate prototype cell reported earlier,[5] for which a typical voltage variation on the memory node represents a change in charge equivalent to $\sim 10^6$ electrons. We have used a two-dimensional electrostatic finite-element analysis method in a section across the memory node, outer gates, and MOSFET channel to calculate the memory-node capacitance. The resulting number of electrons per volt of memory node potential change is plotted as a function of split-gate memory node width in Fig. 5. The parameters of the calculation are chosen to represent the specific device structures discussed in this work, although for simplicity purely metallic behavior in the top silicon layer is assumed. As a best approximation, the memory node and outer gates are taken to be rectangular in cross section, with 10 nm of silicon removed by oxidation at free surfaces. The figure shows that the electron population on the ultrasmall split-gate memory node (1 μm long, and effectively 50 nm in width when oxidation is taken into account) changes by approximately 1000 electrons/V change in potential. The memory states in this device will therefore differ by roughly 200 electrons for a typical Coulomb gap of 0.2 V, a figure considerably lower than in classical DRAM devices.[7]

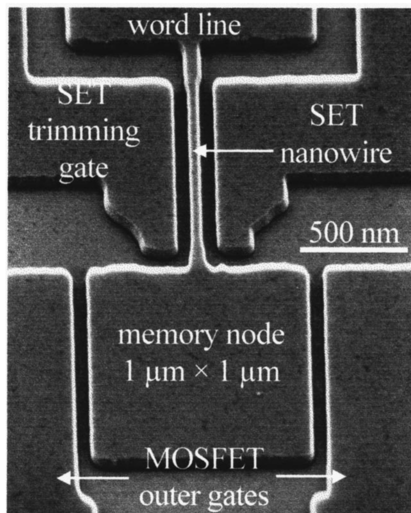

FIG. 3. Detail of the device displayed in Fig. 2. The memory node area is 1 μm$\times$1 μm.

III. SINGLE-ELECTRON TRANSISTOR CHARACTERIZATION

The operation of SOI nanowire single-electron transistors has been described elsewhere in detail.[6] The SET consists of a highly doped silicon nanowire with adjacent trimming-gate electrodes. Although the nanowire is nominally uniform, it is sufficiently small that statistical fluctuations in doping density can create significant inhomogeneity in the width of the space-charge depletion region at the

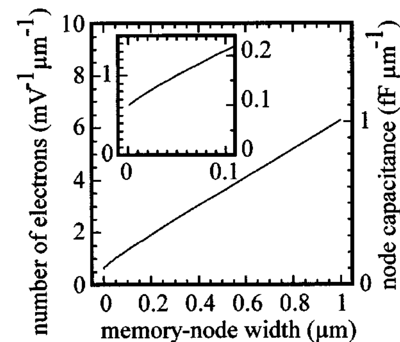

FIG. 5. Electron number per volt on the memory node and memory node capacitance for a split-gate memory cell, as obtained from a two-dimensional capacitance model. The model assumes that the gates and inversion layer channel are metallic in behavior. The gate-to-memory node separation was taken to be 90 nm and the silicon layer thickness 30 nm, as is expected for a real device after oxidation.

Downloaded 24 Oct 2005 to 129.169.10.56. Redistribution subject to AIP license or copyright, see http://jap.aip.org/jap/copyright.jsp

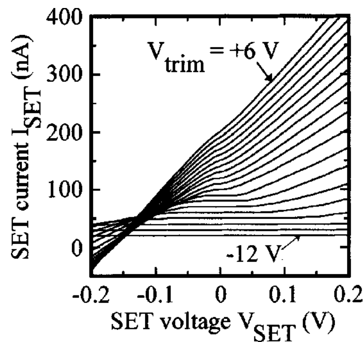

FIG. 6. Current–voltage characteristics as a function of trimming-gate voltage V_{trim} for a silicon nanowire single electron transistor at 4.2 K. The device varies between conducting (positive V_{trim}) and pinched off (negative V_{trim}). At intermediate V_{trim} the conducting pathway is segregated into islands and the device exhibits a zero-current Coulomb gap.

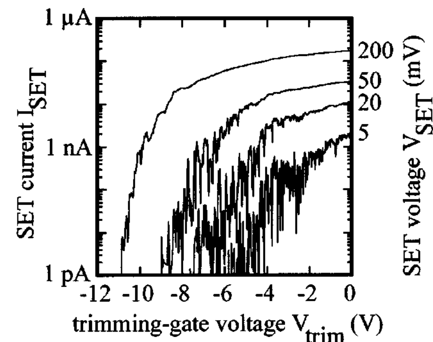

FIG. 8. SET current I_{SET} as a function of trimming-gate voltage V_{trim} for selected SET voltages in the silicon nanowire SET of Fig. 6 at 4.2 K. The plot shows complex, reproducible Coulomb oscillations.

wire's surfaces. Thus the conducting pathway along the nanowire is also of variable width. Nonuniformities in the wire's surface structure may contribute further to this inhomogeneity. An applied voltage V_{trim} on the SET trimming gates is used to vary the depletion width in the wire, such that increasingly negative trimming-gate voltage may lead to localized necking in the wire's conductive pathway where the local dopant density is lowest and/or where the wire is narrowest. In this case, the wire may be likened to a chain of essentially metallic islands separated by depletion barriers of varying width and tunnel resistance. Provided that one or more of these islands is sufficiently small that the energy required to add a single electron to that island is at least comparable with kT, Coulomb blockade may be observed in the nanowire's current–voltage characteristics. Figures 6–9 show typical electrical characteristics of a SET at 4.2 K. Note that this device is fabricated without other memory cell components, and the nanowire is contacted at both ends to permit electrical characterization. In all other respects, the SET is nominally identical to those used in memory cells described later. In Fig. 6, the nanowire's current–voltage ($I_{SET}-V_{SET}$) characteristic is plotted for a series of trimming-gate voltages, and an image map of nanowire differential conductance G for the same device is displayed in Fig. 7. It can be seen that the conduction characteristics vary greatly as the trimming-gate voltage is changed. At positive trimming-gate voltages, the wire's current varies almost linearly with applied voltage. This corresponds to the situation where the nanowire channel is approximately continuous. As the gating potential is decreased, the nonlinearity becomes more pronounced, and a zero-current voltage gap is observed. Subsequent trimming-gate voltage changes widen this gap until the wire is shut off completely. In Fig. 8 the SET current is plotted against the trimming-gate voltage at the fixed SET voltages shown. Complex, reproducible Coulomb oscillations are seen in the current, and are evident also in the voltage gap variations in Fig. 7. Finally, the nanowire voltage applied in constant-current mode as a function of trimming-gate voltage is plotted in Fig. 9 for a range of nanowire currents. Oscillations are clearly evident in the measured voltage, indicative of variations in the voltage gap. However, it should be remembered that stray leakage currents related to interconnects and external circuitry are non-negligible—in the picoamp range here—and therefore the ultralow-current blockade regime cannot be characterized directly. The complexity of the Coulomb oscillations indicates that the SET cannot be described as a single-island system, which would give rise to a single oscillation period. Therefore the voltage gap observed in the SET characteristics represents not a single Coulomb gap, but rather the sum of a

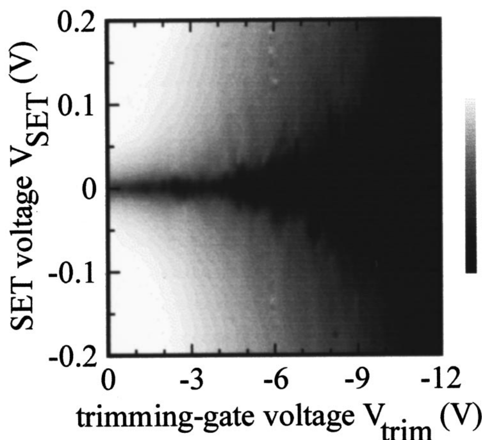

FIG. 7. Image plot of drain–source differential conductance G as a function of V_{trim} and SET voltage V_{SET} for the device of Fig. 6 at 4.2 K. The black regions represent Coulomb blockade or pinch-off.

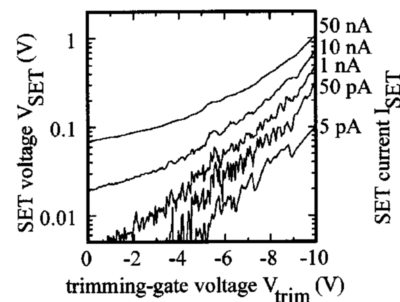

FIG. 9. SET voltage V_{SET} as a function of trimming-gate V_{trim} for selected constant SET currents I_{SET} in the silicon nanowire SET of Fig. 6 at 4.2 K. The fine structure on the curves is reproducible and represents complex Coulomb oscillations consistent with a highly disordered multiple-tunnel junction.

Downloaded 24 Oct 2005 to 129.169.10.56. Redistribution subject to AIP license or copyright, see http://jap.aip.org/jap/copyright.jsp

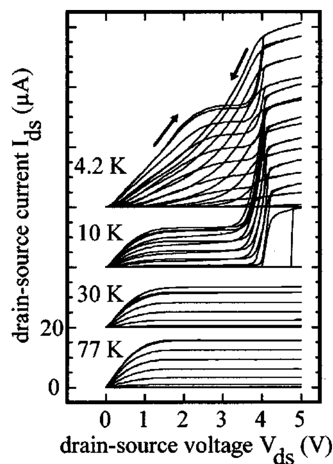

FIG. 10. Double-sweep transistor characteristics for the MOSFET central gate at the temperatures indicated, showing the effects of carrier freezeout below 30 K. The central gate voltage V_{central} is stepped between 0 and 5 V in 0.5 V intervals, and the outer gate voltage V_{out} is maintained at 5 V.

series of Coulomb gaps within a chain of conducting islands in the nanowire. In addition, depletion in the nanowire causes an increase in the number of islands with more negative trimming-gate voltage. The apparent Coulomb gap $2V_c$ obtainable from Figs. 6 and 7 therefore cannot yield a realistic island capacitance, and the estimation of island size in our devices is not practical.

IV. FIELD-EFFECT TRANSISTOR CHARACTERIZATION

The temperature dependence of the operation of the split-gate MOSFET in a memory cell is shown in Figs. 10–12. The memory node, which is the central gate of the MOSFET, is 1 μm×1 μm in area. In each case, the SET trimming-gate voltage is 5 V, under which conditions the SET is effectively a linear resistance and the central MOSFET gate voltage V_{central} is therefore only insignificantly different from the word-line voltage V_{ws}. Figure 10 shows drain–source current I_{ds} as a function of drain–source voltage V_{ds} for word-line voltage steps between 0 and 5 V at the temperatures shown. In Fig. 11. I_{ds} is plotted against V_{central}. In both of these figures, the voltage V_{out} applied to the outer gates of the MOSFET is 5 V, with the intention of creating a continuous inversion layer channel between the

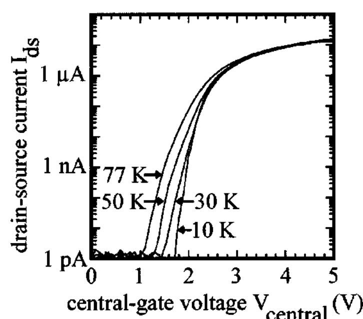

FIG. 11. Transfer characteristic for the central gate of the device of Fig. 10 at the temperatures indicated: $V_{\mathrm{out}}=5$ V.

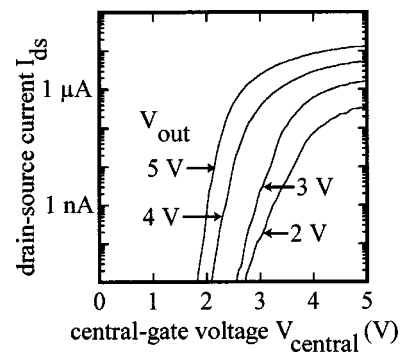

FIG. 12. Transfer characteristic at 10 K for the central gate of the device of Figs. 10 and 11 at the values of outer-gate voltage V_{out} indicated. The plot illustrates the cell selection property of the outer gates.

implanted source and drain regions and the central gate of the MOSFET. As can be seen in the figures, the central gate is effective in controlling the MOSFET channel conductance, indicating that ~70 nm gaps in the split gate are sufficiently small that fringing fields permit the MOSFET channel to be gated without discontinuity when the gate oxide is ~40 nm thick. It can also be seen from Fig. 11 that a small variation in memory-node voltage can yield a change in I_{ds} by several orders of magnitude for a suitably selected operating range, and that this current change is enhanced at lower temperatures.

A feature of Fig. 10 is the kink effect exhibited at temperatures below 30 K. This characteristic is commonly seen in MOSFETs at low temperatures,[8] and is due to carrier freezeout in the bulk region. Under high drain–source voltage conditions, the high electric field leads to impact ionization, and the holes generated in this way diffuse to the bulk. At low temperatures, the high bulk contact resistance means that these holes may accumulate in the bulk region, acting effectively as a parasitic back gate and reducing the MOSFET's threshold voltage. The result is a greater drain–source current at high drain–source voltage than would otherwise be the case. Low temperatures also lead to long carrier trapping and emission times at donor sites, causing space-charge boundaries, and therefore also the inversion channel width, to respond slowly to voltage changes. This results in a hysteresis of the type observed here when the drain–source voltage is swept cyclically. It should be emphasized that the fabrication of an optimized low-temperature MOSFET is beyond the scope of this work. However in the memory data presented here the memory operation may be readily distinguished from any anomalies in the MOSFET behavior.

The operation of the outer gates of the MOSFET is demonstrated further in Fig. 12, which shows the memory node's transfer characteristic at a set of values of outer-gate voltage. We note that the equivalent curve for $V_{\mathrm{out}}=1$ V falls below the I_{ds} noise limit. For a fixed memory node voltage, the MOSFET drain–source current I_{ds} is governed by the outer-gate potential, allowing the cell to be operated in either an ''off'' or an ''on'' state. However, it is important to note that there is a capacitive coupling between the memory node and the outer gates, and when the device is operated as a memory cell as described in Sec. V, it is in principle possible to

Downloaded 24 Oct 2005 to 129.169.10.56. Redistribution subject to AIP license or copyright, see http://jap.aip.org/jap/copyright.jsp

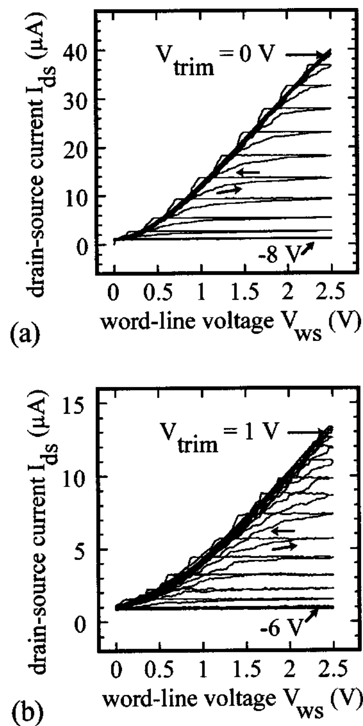

FIG. 13. Memory action in split-gate memory cells at 4.2 K. Memory node dimensions prior to oxidation are: (a) 1 μm×1 μm and (b) 1 μm ×70 nm, the outer-gate voltage is 5 V and the drain–source voltage is 5 V. At intermediate values of SET trimming-gate voltage V_{trim}, the SET nanowire exhibits a zero-current Coulomb gap of width $2V_c$ and the memory node voltage lags the word-line voltage V_{ws} by an amount sufficiently in excess of $2V_c$ to permit a suitable memory node charging current, creating a hysteresis in the MOSFET output current I_{ds}. The variation of the width of the hysteresis loop arises from changes in the size of the Coulomb gap.

produce an unwanted change in the cell's memory state by altering the outer-gate voltage.

V. MEMORY CELL OPERATION

The memory effect in split-gate cells at 4.2 K is demonstrated in Figs. 13(a) and 13(b), for two cells with memory node dimensions 1 μm×1 μm and 1 μm×70 nm, respectively. The figures show the current I_{ds} through the sense MOSFET at a fixed drain–source voltage V_{ds} as the word-line voltage V_{ws} is swept cyclically. The SET trimming-gate voltage V_{trim} is stepped negatively as shown to vary the nanowire's conduction characteristic. The MOSFET outer-gate voltage V_{out} is fixed at 5 V in these figures to ensure that there is a continuous, open drain–source channel in the MOSFET. At the most positive values of SET trimming-gate voltage, the nanowire is ohmic at all times, and the word line acts as an orthodox gate contact. At the most negative SET trimming-gate voltage, the wire is completely pinched off, isolating the memory node from the word line and preventing any change in memory-node potential. At intermediate voltages, the MOSFET current exhibits a clear hysteresis, with a separation between the sweep-up and sweep-down curves dependent principally on the extent of the Coulomb gap $2V_c$ in the nanowire for each specific biasing condition (or more correctly on the SET nanowire voltage at the SET current level required to charge the memory node at the required rate, as outlined below). However, the precise nature of the hysteresis cannot be described entirely in terms of a unique SET voltage, since the hysteresis width varies with word-line potential for any given sweep rate. This may be explained by considering the potential of the nanowire relative to its trimming gates, and relative to any other conducting component to which it is capacitively coupled. As the word line is made more positive, the net effect is to gate the nanowire more negatively. By comparison with Figs. 6–9, dependent on the range of the word-line voltage sweep, the nanowire's current–voltage characteristic will vary continuously, with a trend towards increased Coulomb gap $2V_c$ with increasing V_{ws}, and a correspondingly wider hysteresis loop.

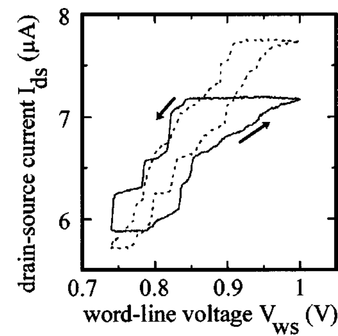

FIG. 14. Hysteresis loop at 4.2 K in a cell of memory node dimensions 1 μm×1 μm, with cycle period 60 s (solid line) and 320 s (dashed line). V_{trim}= −5 V. The plateau regions in I_{ds} indicate where the SET nanowire is biased within the Coulomb gap. The steps in the hystersis loop are produced by Coulomb oscillations in the SET. Faster cycle speeds require the SET to be biased such that the charging current is high, hence the wider hysteresis loop.

Figure 14 shows the hysteresis at 4.2 K in a device of memory node area 1 μm×1 μm during cyclic operation in a reduced range of word-line voltage V_{ws}. For this device, variation in the SET characteristics between samples leads to a different V_{trim} operating range than in results shown elsewhere, though the results are qualitatively sample independent. The hysteresis is plotted for two different cycle speeds. Both curves exhibit constant-current regimes at the upper and lower current extremes, representing the parts of the cycle where the SET nanowire is in blockade. It can be seen that the hysteresis has a sweep-period dependence, although the shape is qualitatively unchanged. This dependence results from the fact that the charging or discharging rate of the memory node is a function of the potential drop across the SET nanowire. The memory-node potential lags the word-line potential by an amount which depends on the charging rate; faster sweep rates lead to a greater potential drop across the nanowire and a wider and flatter hysteresis loop. The potential difference across the SET V_{SET} is given directly by the deviation (in terms of V_{ws}) of the hysteresis loop from the MOSFET's standard transfer characteristic under nonhysteretic conditions (as obtained using a positive SET trimming-gate voltage). In Fig. 14, the magnitude of V_{SET} is typically ~40 and ~20 mV for the faster and slower sweeps, respectively. We may also calculate the mean current flowing in the SET nanowire during the sweeps using the MOSFET transconductance (~11 μS) and the calculated memory-node ca-

Downloaded 24 Oct 2005 to 129.169.10.56. Redistribution subject to AIP license or copyright, see http://jap.aip.org/jap/copyright.jsp

pacitance (~1 fF). The mean charging currents are in the low 10^{-18} A range for the slower sweep and close to 10^{-17} A for the faster sweep, although the current is sharply peaked at the ''rise'' part of the hysteresis steps. These current levels are well below our current measurement sensitivity limits ($\sim 10^{-12}$ A), and therefore indicate an apparent Coulomb gap smaller than would be obtained if the nanowire current could be measured directly. The Coulomb gap therefore can only be defined in the memory cells in terms of a threshold current value, and the effective Coulomb gap is a function of the measurement procedure.

A further notable feature in Figs. 13 and 14 is the series of abrupt current changes or steps in the hysteresis curve. The trend of the hysteresis loop to broaden as V_{ws} increases was attributed earlier to the progressive increase in the SET nanowire's Coulomb gap, so it seems reasonable to suggest that the steps in the hysteresis may be connected with fine oscillations in the Coulomb gap, of the type seen in Fig. 8. In parts of the sweep where the Coulomb gap is progressively decreasing, the memory node voltage would vary more rapidly than V_{ws}, giving a sharp rise in output current. Where it is increasing, a plateau would be observed. If the observed phenomena are due to Coulomb oscillations, it would be expected that the SET trimming-gate voltage should have a very strong influence on the position of the hysteresis steps, and specifically that the steps should change position in terms of V_{ws} if the SET trimming-gate voltage is varied. Image maps of the transconductance G of the sweep-up part of a series of hysteresis loops are plotted as a function of word-line voltage V_{ws} and SET trimming-gate voltage V_{trim} in Figs. 15(a) and 15(b), for $1\ \mu m \times 1\ \mu m$ and $1\ \mu m \times 70$ nm memory node devices, respectively. The lighter lines represent sharp increases in the MOSFET current, and therefore in the memory node potential, and the darker regions in between indicate relatively slow charging ranges. The black areas at the lower left show the bias conditions under which the SET nanowire is in blockade, while those at the lower right represent nanowire pinch-off. The relatively uniform gray areas at the top of the maps indicate roughly constant transconductance in the regime where the nanowires tend toward linear conduction. The hysteresis loop steps are represented by the light and dark bands, and as expected, their position in V_{ws} varies approximately linearly with the SET trimming-gate voltage V_{trim}. Since we assume that each band represents an isogram along which the nanowire characteristics are identical, we may assume also that the potential drop across the nanowire is constant and therefore that the slopes of the plots indicate the change in SET trimming-gate bias required to maintain a constant potential difference between the conducting island or islands in the wire and the wire electrodes. The gradient of the bands is therefore a direct measure of the capacitive coupling between the SET trimming gate and the island. In Figs. 15(a) and 15(b), the gradient dV_{trim}/dV_{ws} is similar for all of the hysteresis steps and is equal to ~2.0 and ~1.8 for the small and large memory node devices, respectively. Furthermore, the corresponding data for the sweep-down part of the hysteresis cycles give the same results. By contrast, the hysteresis loop steps depend only very weakly on the MOSFET terminal voltages; dV_{ds}/dV_{ws} and dV_{out}/dV_{ws} are ~100 or more for both memory node geometries. The high capacitive link between the steps and the SET trimming gates is consistent with a link between the steps and the SET, and gives direct evidence that the hysteresis loop steps arise from Coulomb oscillations.

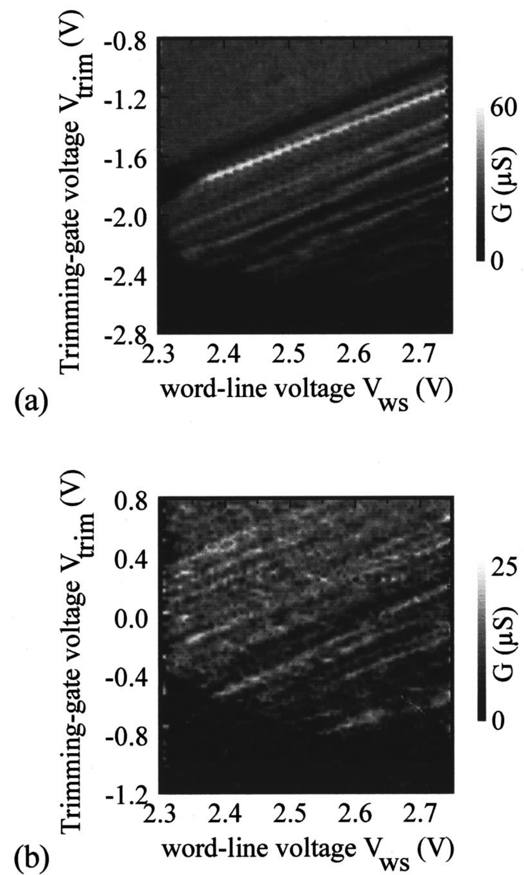

FIG. 15. Image maps of differential cell transconductance G at 4.2 K obtained from a series of hysteresis loops in devices of preoxidation memory node dimensions: (a) $1\ \mu m \times 1\ \mu m$ and (b) $1\ \mu m \times 70$ nm, as a function of SET trimming-gate voltage V_{trim}. G is obtained from the slope of the sweepup (increasing V_{ws}) part of the I_{ds}/V_{ws} hystereses only, and the SET trimming-gate voltage V_{trim} is varied between the limits indicated in increments of −25 mV. The bright and dark bands indicate steps in the hysteresis loop, and their position in terms of V_{ws} varies strongly and linearly with V_{trim}.

The speed of operation is a critical parameter in a practical memory cell. Figure 16 shows pulsed operation of a cell with a $1\ \mu m \times 1\ \mu m$ memory node for a range of write pulse widths from 1 μs to 10 ns. In this measurement, the word line is held at a steady-state ''read'' voltage V_{read} upon which is superimposed alternating ''write-zero'' and ''write-one'' pulses of voltage level V_{write0} and V_{write1}, respectively, where $V_{write0} < V_{read} < V_{write1}$, and of duration t_{write0} and t_{write1}, respectively. The MOSFET drain–source voltage V_{ds} is held at a constant 5 V. The write pulse widths are shorter than the MOSFET current sampling interval, and the plots represent the MOSFET current I_{ds} during the successive 25-ms-long ''read'' intervals. The cell operation is essentially the same as a hysteresis sweep in the interval $V_{write0} < V_{ws} < V_{write1}$, and consequently for any given intermediate volt-

Downloaded 24 Oct 2005 to 129.169.10.56. Redistribution subject to AIP license or copyright, see http://jap.aip.org/jap/copyright.jsp

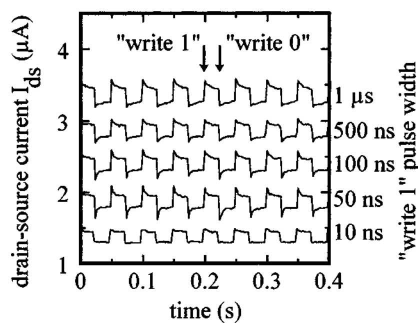

FIG. 16. Pulsed operation speed for a cell of 1 μm×1 μm memory node size at 4.2 K. The MOSFET drain–source current I_{ds} is measured over time at a constant drain–source voltage V_{ds} of 5 V. The steady-state word-line ''read'' voltage is 1.35 V, upon which is superimposed alternating ''write 1'' and ''write 0'' square pulses of voltage level $V_{write1}=1.95$ V and $V_{write0}=0.85$ V, respectively. The ''write 0'' pulse width t_{write0} was maintained at 1 μs, while the ''write 1'' pulse width t_{write1} was varied as shown. $V_{trim}=0$ V, $V_{out}=5$ V.

age, such as V_{read} in this instance, there are two possible values of the cell output I_{ds}, provided that the charging time of the memory node is not small compared with the ''write'' pulse widths. With t_{write0} fixed at 1 μs, t_{write1} was varied, and it can be seen in Fig. 16 that even at $t_{write1}=10$ ns the memory node charging rate is sufficient for the memory states to be clearly resolved in the two values of I_{ds}. The biasing conditions in Fig. 16 differ from those in the hysteretic regime of Fig. 13(a) principally because it is a different sample, and the localized variations in pinch-off voltage in the nanowire which produce the Coulomb blockade effect are determined by random and/or uncontrolled variations in the wire's structure, and blockade characteristics may be observed in these devices even for $V_{trim}\geqslant 0$ if localized sections of the wire are depleted. As a consequence, the sample of Fig. 16 requires a more positive trimming-gate voltage V_{trim} than that of Fig. 13(a) to produce a comparable effect. Furthermore, in a real memory cell at any nonzero SET voltage, the memory-node leakage current will not be strictly zero, and excess memory-node charge would be expected to decay with time. As a result, for a given V_{trim}, a memory effect may be observable only for faster cycle speeds. Thus, even if two measurements were from the same device, the fast, pulsed cycle of Fig. 16 would show a memory effect in a more positive range of V_{trim} (or alternatively in a more negative range of V_{ws}) than would the slow hysteresis of Fig. 13(a).

The performance shown in Fig. 16 at $t_{write1}=10$ ns is consistent with a simple time constant estimate given a memory node capacitance of order 1 fF and a nanowire resistance of order 1 M Ω. It is seen in Fig. 16 that the separation between the two memory states after any initial residual transient is more or less independent of the ''write-one'' pulse width, in contrast to the clear sweep-rate dependence of the hysteresis curves in Fig. 14. This is because there is a fundamental difference between the two measurements. During a swept measurement such as those in Fig. 14, the SET nanowire maintains an equilibrium, drawing just enough charging current to maintain the necessary voltage drop between the word line and the memory node. Hence the SET current and the hysteresis width are both defined by the word-line voltage sweep rate and the memory-node capacitance. During a pulsed measurement, the SET is switched discontinuously between high- and low-current voltage ranges (here $\geqslant 0.1$ μA and $\leqslant 0.1$ pA, respectively) in a nonequilibrium manner. To ''write'' to the memory cell, the SET is pulsed well out of ''blockade.'' This may lead very rapidly to the storage of an amount of charge significantly higher than the memory node can retain under the subsequent ''read'' conditions. The start of the read period will then be accompanied by a discharging transient, as seen in Fig. 16, before reaching a quasisteady state. Since the rate of change of memory-node voltage with time is proportional to the SET current, and therefore varies sharply with SET voltage, the initial rate of memory-node discharge at the start of the read period may be rapid. Assuming that the nanowire current varies continuously with voltage across many decades, the read period is likely to be dominated by discharging with a time constant of order of the read duration, irrespective of whether the write-period charging has reached saturation. As measured, the form of the read-period MOSFET current output will therefore be effectively independent of write pulse width. The absence of a strong dependence of memory state separation on write time in Fig. 16 demonstrates that for the pulse widths used here, the pulsing operation satisfies this condition, although it may be marginal in the case of the 10-ns-wide write1 pulse.

It is significant to note that the ratio of retention time to write time determined from Fig. 16 is greater than 10^6. This value arises directly from the strong nonlinearity between current and voltage in the silicon nanowire. Nanowire SETs in isolation exhibit Coulomb blockade behavior, including Coulomb oscillations at nanowire current levels in the nA range (Fig. 8), more than 3 orders of magnitude above the sub-pA range in which the nanowire is biased during the read period of Fig. 16. In addition, steps in memory-node hysteresis measurements have been attributed to Coulomb oscillation effects. It is therefore reasonable to propose that the high-speed memory effect shown in Fig. 16 is predominantly due to Coulomb blockade.

The dependence of the memory effect on operating temperature is illustrated in Fig. 17, in which the hysteresis in a cell of memory node size 1 μm×1 μm is plotted as a function of temperature. As the temperature increases, the hysteretic part of the curve shifts to higher V_{ws}, a dependence which may be explained by a temperature-dependent increase in the free carrier concentration in the material and/or an increase in thermally activated conduction across depletion barriers in the wire. The other notable feature is the disappearance of the current steps in the hysteresis loop at temperatures above 45 K. It seems likely that this temperature is the threshold above which Coulomb blockade effects cease to be significant in this particular device. Hysteresis at higher temperatures would therefore be due only to slow charging and discharging of the memory node. This is a reasonable proposition since the nanowire may still be effectively pinched off by the electric field of the SET trimming gates at the higher temperatures, even in the absence of Coulomb blockade effects. This conclusion is supported by the

Downloaded 24 Oct 2005 to 129.169.10.56. Redistribution subject to AIP license or copyright, see http://jap.aip.org/jap/copyright.jsp

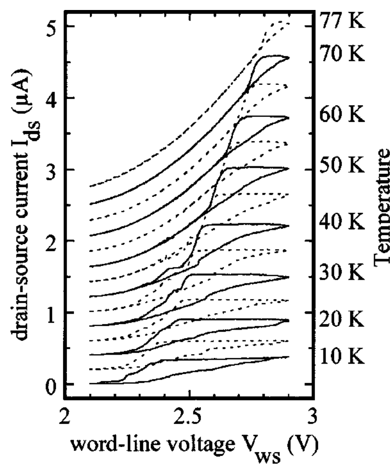

FIG. 17. Temperature dependence of a memory hysteresis loop for a cell of 1 μm×1 μm memory node size, showing the disappearance of the step features above approximately 45 K. $V_{ds}=2$ V, $V_{trim}=0$ V, $V_{out}=5$ V.

observation that above 45 K it becomes impossible to observe a constant-I_{ds} regime at the low-V_{ws} end of short-range hystereses such as those in Fig. 14. It should be noted that in the case of a linearly conducting, high-resistance nanowire, the memory node charging time would be comparable to the retention time. Coulomb blockade is therefore a necessary requirement for a useful memory cell of the type we describe here.

VI. DISCUSSION

We have demonstrated the operation of an integrated SET-MOSFET memory gain cell in which the memory states are represented by MOSFET current output levels corresponding to a difference of a few hundred electrons on the memory node. Single-electron charging effects are visible in the output hysteresis. The fabrication scheme is fully compatible with current silicon-based memory chip technologies. A suitable scheme for cell selection in an array of cells of this type has been published elsewhere.[9] The Coulomb blockade memory effect persists to 45 K and the cell can be operated with voltages of less than 5 V. The principal advantage of Coulomb blockade in few-electron memory applications is that it gives the potential to control few-electron populations with single-electron precision, provided that the charging energy per electron of the memory node (given by e^2/C, where C is the memory node capacitance) is large compared with the thermal energy kT. Under such conditions, each tunneling event onto the memory node will strongly inhibit the probability of further carrier transfer. However in the present devices, which operate using differences of hundreds of electrons and have a memory-node charging energy smaller than kT, single-electron precision is neither achieved nor necessary. Memory operation in this case is dependent on the SET current–voltage nonlinearity, and in an idealized Coulomb blockade device, the ''off'' current can be zero. SETs are therefore uniquely suitable for the storage of small charges.

The node capacitance, and therefore its electron population, is determined by the geometry of the cell, and the cell's write time to retention time ratio is essentially determined by the SET's current–voltage characteristic. Hence the memory node capacitance is the principal quantity defining the performance of the memory cell. A small node capacitance implies a rapid write speed, and the 10 ns pulsed operation and long retention times demonstrated in Fig. 16 are comparable with practical DRAMs.

Several cell design considerations are apparent. The excess electron number on the memory node and the cell's current output both depend approximately linearly on the memory node length, and this length may be minimized as far as can be tolerated by the external sense circuitry. The electron number also depends on the memory node capacitance, and this is reduced for a thicker gate oxide, which would also permit a wider separation between the memory node and the MOSFET outer gates. However, the MOSFET transconductance is greatly enhanced by decreasing the oxide thickness, and there should be a tradeoff in performance between few-electron operation and high output sensitivity. To permit high sensitivity while avoiding unnecessarily high lithography demands, the design may be modified such that the memory node forms an orthodox MOSFET with a short gate length, eliminating the need for a split gate. In such a cell, the ''select'' function of the outer gates in the present design may be performed by a pass transistor in series with the memory node MOSFET, such that the gap between the memory node and the pass transistor is bridged by a noncontacted MOSFET drain/source implant, at minimal cost in terms of cell size.

Since this design of memory cell utilizes Coulomb blockade as a means of charge retention, it is potentially suitable for scaling to the single electron limit. However, to be practical, it must have the potential to perform at least as well as standard DRAM at room temperature or at least 77 K. A theoretical analysis of the performance of the present type of memory structure has been published elsewhere.[10] To be competitive in terms of power dissipation, the cell requires a memory refresh rate similar to DRAM, and therefore a retention time of at least ~1 s is needed. The memory cell studied here satisfies this requirement at 4.2 K, but retention time decreases very sharply with increasing temperature. However, the write time, memory retention time and the threshold voltage difference between memory states are all determined by the SET's current–voltage characteristics. The cell may therefore be improved by the use of an enhanced SET. The operating temperature depends principally on the SET island size, which in the present design may perhaps be improved by reducing the nanowire's width and increasing its doping level correspondingly. Alternatively, a different type of SET may be used. Coulomb blockade effects have been observed at room temperature in granular silicon SETs,[11] and the present limitations of nanoscale lithography are technological rather than fundamental.

The complexity of our cell can be reduced if the SET characteristics can be controlled with sufficient reproducibly that the SET trimming gates may be omitted, allowing the device to operate with only four terminals. Given that no storage capacitor is necessary, the cell is in principle significantly smaller than classical DRAM. One further issue to

Downloaded 24 Oct 2005 to 129.169.10.56. Redistribution subject to AIP license or copyright, see http://jap.aip.org/jap/copyright.jsp

consider is the operating voltages, which in the results presented here are within ±5 V, somewhat higher than in conventional DRAM, and therefore indicative of power dissipation and compatibility problems. However, these voltages are necessary as a result of SET variations (which may in principle be overcome) and the kink effect in our MOSFETs. With a suitable low-temperature MOSFET, the drain–source voltage may be small, and the word-line voltage may be varied in and around the MOSFET subthreshold regime, permitting orders of magnitude difference between ''1'' and ''0'' state output current levels.[10] Thus the maximum voltages applied to the cell may be little more than the Coulomb gap width. Assuming that the SET operating temperature issue can be solved, the present cell design has the potential for write times, retention times, and power dissipation at least comparable with classical DRAM cells, and with scalability well beyond the present limits.

ACKNOWLEDGMENTS

The authors acknowledge with thanks the advice and encouragement of K. Katayama, H. Mizuta, D. A. Williams, H.-O. Müller, and K. Nakazato of Hitachi Cambridge Laboratory, and technical discussions with S. Biesemans and M. Rosmeulen of IMEC, Leuven, Belgium. This work was supported by the European Union's Microelectronics Advanced Research Initiative (MEL-ARI), as part of the Fabrication and Architecture of Single-Electron Memories (FASEM) project.

[1] K. Yano, T. Ishii, T. Hashimoto, T. Kobayashi, F. Murai, and K. Seki, Tech. Dig. Int. Electron Devices Meet. **93**, 541 (1993).
[2] S. Tiwari, F. Rana, H. Hanafi, A. Hartstein, E. F. Crabbé, and K. Chan, Appl. Phys. Lett. **68**, 1377 (1996).
[3] A. Nakajima, T. Futatsugi, K. Kosemura, T. Fukano, and N. Yokoyama, Appl. Phys. Lett. **70**, 1742 (1997).
[4] N. J. Stone and H. Ahmed, Appl. Phys. Lett. **73**, 2134 (1998).
[5] Z. A. K. Durrani, A. C. Irvine, H. Ahmed, and K. Nakazato, Appl. Phys. Lett. **74**, 1293 (1999).
[6] R. A. Smith and H. Ahmed, J. Appl. Phys. **81**, 2699 (1997).
[7] K. Itoh, Y. Nakagome, S. Kimura, and T. Watanabe, IEEE J. Solid-State Circuits **32**, 624 (1997).
[8] F. Balestra and G. Ghibaudo, Solid-State Electron. **37**, 1967 (1994).
[9] H. Mizuta, K. Katayama, H.-O. Müller, and D. A. Williams, Extended Abstracts IWCE-6, Osaka University, 1998, p. 17
[10] K. Katayama, H. Mizuta, H.-O. Müller, D. A. Williams, and K. Nakazato, IEEE Trans. Electron Devices **46**, 2210 (1999).
[11] B. H. Choi, S. W. Hwang, I. G. Kim, H. C. Shin, Y. Kim, and E. K. Kim, Appl. Phys. Lett. **73**, 3129 (1998).

Downloaded 24 Oct 2005 to 129.169.10.56. Redistribution subject to AIP license or copyright, see http://jap.aip.org/jap/copyright.jsp