Single electron tunneling and suppression of short-channel effects in submicron silicon transistors

M. G. Peters, S. G. den Hartog and J. I. Dijkhuis

Faculty of Physics and Astronomy, Debye Institute, University of Utrecht, 3508 TA Utrecht, The Netherlands

O. J. A. Buyk
Philips Research Laboratories, 5656 AA Eindhoven, The Netherlands

L. W. Molenkamp
2. Physikalisches Institut, RWTH-Aachen, D-52056 Aachen, Germany

(Received 10 March 1998; accepted for publication 18 July 1998)

We report measurements on submicron metal–oxide–semiconductor field effect transistors equipped with a gate on three sides of the channel. At room temperature, a strong suppression of short-channel effects has been achieved for the narrowest channels. At liquid helium temperatures, the same devices exhibit clear conductance oscillations in the subthreshold regime, indicating that a quantum dot has formed in the disordered channel. © 1998 American Institute of Physics.

I. INTRODUCTION

In the last decades, there has been a strong effort to reduce the dimensions of silicon devices, in order to achieve a closer packing of transistors. However, at a given design rule, the scaling of the physical processes breaks down and new phenomena that are absent in larger structures dominate the device behavior. For example, on decreasing the size of a metal–oxide–semiconductor field effect transistor (MOSFET), at some point the channel length approaches the depletion layer widths of the source and drain. This results, among other things, in a degradation of the subthreshold characteristics of the device and a failure to achieve current saturation. One can suppress these and other short-channel effects by high doping in the channel, at the expense of a reduced mobility, lower operating speed, and an increased risk for avalanches at the drain. An alternative is to employ silicon-on-insulator (SOI) devices, where the silicon channel can be fully depleted. Moreover, SOI allows for the fabrication of three-dimensional structures with the gate surrounding the channel. The specific gate geometry may even, under favorable conditions, invert the entire channel, and this volume inversion is thought to be advantageous for the electrical properties.

At the same time, SOI devices have become increasingly important for the entirely different field of single electron tunneling. Although most of the fundamental research was focused on quantum dots in III–V semiconductors at ultralow temperatures (millikelvins), quite recently single electron tunneling has been reported at elevated temperatures in silicon devices. This has opened the door to future applications and silicon has since then gained renewed interest.

Beside the fact that many of these structures were based on SOI, the suppression of short-channel effects, volume inversion, and single electron tunneling were not investigated earlier in one and the same device. A combined study is important since the suppression of short-channel effects is a prerequisite to achieving single electron tunneling at room temperature. On the other hand, as the dimensions of these devices continue to decrease, it becomes quite conceivable that the inversion is not restricted to the interface and volume inversion effects can no longer be ignored.

We present a novel submicron SOI MOSFET equipped with a gate on three sides of the channel. For the narrowest devices, short-channel effects are strongly suppressed while maintaining a low doping concentration. We have examined to what extent volume inversion occurs in these devices by analyzing the subthreshold swing and the transconductance. Simulations are performed to support our analysis. For the devices exhibiting suppression of short-channel effects, we find at liquid helium temperatures, clear conductance oscillations in the threshold regime, indicating that the devices act as single electron transistors. In contrast to most other devices, our single-gated devices are the result of standard lithographic techniques that can be fabricated on large-scale wafers and yet show clear single electron tunneling phenomena.

II. DEVICE FABRICATION AND EXPERIMENTAL TECHNIQUE

The layout of our MOSFET is shown in Fig. 1. The SIMOX SOI wafer consists of a lightly p-doped silicon substrate, a 400 nm thick layer of silicon oxide, 140 nm of p-doped (1 × 10^{15} cm^{-3}) silicon, and a 40 nm thick silicon oxide. Electron-beam lithography and anisotropic dry etching down to the 400 nm thick oxide layer were then used to create the silicon channel. The remaining oxide was removed by wet etching and a 25 nm thick thermal oxide was grown.
around the channel at 1000 °C. As a result the height of the channel is reduced to approximately 100 nm. Subsequently, a 300 nm thick gate contact of poly-silicon was deposited, implanted with phosphorous ions \((1 \times 10^{16} \text{ cm}^{-2})\) at 50 keV, and annealed at 900 °C for 30 min to allow for diffusion and activation of the phosphorous ions. On top of the polysilicon, 40 nm of tetra-ethyl-ortho-silicate was deposited and the gate layout was defined by electron-beam lithography and anisotropic dry etching. Contact areas were created by implanting the source and drain regions with arsenic ions \((5 \times 10^{15} \text{ cm}^{-2})\) at 50 keV. Finally, the arsenic ions were activated by heating the sample to 900 °C. Note that diffusion of arsenic ions under the gate reduces the effective channel length.

We performed measurements at temperatures from 1.8 K to room temperature. For the low-temperature experiments, the device was bonded and mounted in a 4 He cryostat. The conductance was measured by applying an ac current at 70 Hz. After amplification and filtering, the ac voltage across the sample was detected at the same frequency with a digital spectrum analyzer and was kept well below \(k_B T/e\). At room temperature the transistor characteristics were measured using a probe station equipped with a HP 4145 curve tracer.

III. ROOM-TEMPERATURE RESULTS

MOSFETs were investigated with widths \(W\) and lengths \(L\) ranging from 0.1 to 8.0 \(\mu\)m. At room temperature, the desired long-channel behavior was found for devices with lengths of 0.8 \(\mu\)m or more. All channels shorter than 0.2 \(\mu\)m suffered from strong short-channel effects. For the intermediate lengths, these effects were also observed except for the narrow devices. Figures 2(a) and 2(b) display the drain current \(I_D\) versus gate voltage \(V_G\) at room temperature for two devices with \(L=0.5 \mu\)m, differing only in width. The 2 \(\mu\)m wide MOSFET clearly shows short-channel effects: the subthreshold swing \(S = \ln 10 \times dV_G/d(\ln I_D)\) rises rapidly with drain voltage and, at high biasing, the gate can no longer block the current (i.e., punch-through). In contrast, the 0.2 \(\mu\)m wide MOSFET possesses good subthreshold characteristics: both the subthreshold swing and the threshold voltage are only slightly affected by the drain voltage. Figure 3(a) shows the subthreshold swing versus channel width at \(V_D = 0.1 \text{ V}\) for \(L=0.3 \mu\)m and \(L=2.0 \mu\)m. A linear dependence on channel width is found, corresponding with a constant mobility of 830 \(\text{cm}^2/\text{V s}\).
channel has a much higher swing which points to a degradation of the subthreshold characteristics. Only for the narrowest devices (W<1.0 μm) does the swing approach that of the long channel. Clearly, for the narrow devices, the side gates are effective in depleting the channel, thereby suppressing the short-channel effects and improving the performance. These observations are in line with work done on quite different types of devices by Hisamoto et al. 2 Quite recently, Leobandung et al. reported the suppression of short-channel effects in wire-channel wrap-around gate structures. 2 Although their measurements were done on devices with a doping concentration more than two orders higher than ours and a geometry more difficult to fabricate, we obtain similar results for the suppression. Furthermore, our experimental findings are in agreement with theoretical work on three-dimensional gating in MOSFETs by Yan et al. 11 and numerical calculations by Miyano et al. 12 Although calculated for different dimensions, the results obtained by Miyano et al. reproduce the channel-width dependence of the subthreshold swing displayed in Fig. 3(a). Finally, Fig. 3(b) displays the maximum transconductance at V_D=0.1 V for the 2.0 μm long MOSFET. We observe a linear increase with channel width over the entire range. From this we conclude that the mobility is constant and amounts to 830 cm^2/Vs, approximately 30% smaller than the typical mobility in the bulk.

IV. MODELING

To illustrate the effects of the side gates, we calculated the potential and carrier distributions in a 0.2 μm long channel, using the finite-element simulator PADDY, developed at Philips Research Laboratories (Eindhoven). For the geometry of our device and simply incorporating abrupt junctions, the three-dimensional Poisson equation was solved for a range of gate voltages, while keeping the source and drain grounded. By comparing the calculated distribution of the electrons with the concentration of impurities, one can judge whether the channel is conducting under vanishing bias conditions. Here, we make a comparison between two devices with W=0.5 μm and W=0.1 μm. Figure 4 shows the electron concentrations versus position along a vertical line through the center of the channel (see inset Fig. 4(a)). For the wide MOSFET, the electron concentration at the location near the gate oxide tracks the gate voltage. However, at larger distances from the gate, the electron concentration is virtually constant and independent of gate voltage as a result of screening. Due to the overlap of the source and drain depletion layers, the concentration of electrons at the buried oxide always exceeds the channel impurity concentration. Despite the action of the gate, the region near the buried oxide remains inverted for all gate voltages and consequently the channel cannot be pinched off. An entirely different situation is met for the narrow channel. Here, the electron concentration varies almost uniformly with gate voltage. Because the side gates are located closer to each other, they can deplete the full channel, prevent inversion, and turn off the channel. Hence, the simulations confirm the observed suppression of short-channel effects by the side gates.

V. VOLUME INVERSION

The uniformity found for the calculated carrier distributions in the narrow channel justifies a discussion of the consequences of volume inversion in our devices. Volume inversion—the creation of carrier inversion in the bulk of the SOI wire, rather than only at the Si/SiO_2 interfaces—may lead to better device characteristics because it reduces surface scattering and increases mobility. In most submicron devices, short-channel effects are suppressed by a high doping concentration in the channel, which reduces the chances to achieve volume inversion. On the other hand, in our device the long-channel behavior is restored by the presence of the side gates while a low doping concentration can be maintained. The device is therefore excellent for studying the effect of volume inversion, especially on the mobility. However, in Fig. 3(b) the maximum transconductance is observed to be linear over the entire range of W and thus suggests a constant mobility. Of course, one expects that the carrier mobility for an entirely inverted channel is enhanced only if the bulk mobility largely exceeds that at the interface. This is clearly not the case in our device, as we noted earlier. The subthreshold swings observed in our experiments are in good agreement with theoretical calculations on SOI MOSFETs. Wouters et al. 14 showed that the theoretical lower limit for bulk MOSFETs of 60 mV/decade still holds for gate-all-
The gate voltage is around threshold. A quantum dot can be formed in the inversion layer into disjunct segments when the potential in the oxide, a nonuniform potential landscape exists, breaking up the dot into segments. A plausible scenario for the formation of the quantum dot is the following: due to fixed charges in the channel and in the oxide, a nonuniform potential landscape exists, breaking up the inversion layer into disjunct segments when the gate voltage is around threshold.

VI. LOW-TEMPERATURE EXPERIMENTS

Strikingly different characteristics are observed at low temperatures. Figure 5(a) displays the conductance versus gate voltage at 1.8 K for a device with $L = 0.2 \mu m$ and $W = 0.5 \mu m$. A series of equidistant peaks in the conductance is observed. This regularity and the observation that the conductance oscillations are hardly affected by magnetic fields up to 8 T, makes it plausible to identify this phenomenon as Coulomb-blockade oscillations. This phenomenon is generally found in quantum-dot systems where a small island (the quantum dot) is weakly coupled to the environment through tunnel barriers. Every conductance peak then corresponds with the addition of a single electron to the quantum dot. The occurrence of Coulomb oscillations in our MOSFETs thus marks the presence of a quantum dot in the rectangular silicon channel. A plausible scenario for the formation of the dot is the following: due to fixed charges in the channel and in the oxide, a nonuniform potential landscape exists, breaking up the inversion layer into disjunct segments when the gate voltage is around threshold.

A quantum dot can be formed in case one segment limits the conductance and the tunnel barriers connecting the segment to the source and drain are sufficiently high. Neglecting the quantization of the internal levels of the dot, the peak spacing then simply equals $e/C_G$, where $C_G$ is the capacitance from dot to gate. For the present MOSFET [Fig. 5(a)], the peak spacing is 31 mV on average, which corresponds in that limit to $C_G = 5.2 \text{ aF}$.

As an illustration of the role of disorder in the formation of the quantum dot in our MOSFET channels, we discuss the low-temperature results obtained on a series of devices from the same batch with lengths and widths ranging from 0.1 to 2 $\mu m$. We never found Coulomb oscillations in devices longer or wider than 0.5 $\mu m$. For smaller samples, an irregular multiple-peak structure was observed, except for a few devices that did show clear Coulomb oscillations above. We could not identify any relation between the period of the oscillations and the size of our MOSFET. This is illustrated in Figs. 5(b) and 5(c) which show the conductance versus gate voltage for two different MOSFETs from the same batch. Although the two MOSFETs are identical in size ($L = 0.2 \mu m$, $W = 0.1 \mu m$), the average peak-to-peak distance differs considerably. Furthermore, the conductance of the first MOSFET [Fig. 5(b)] is almost two orders lower. These observations clearly indicate that the quantum dot is not primarily the result of the specific layout of the device but is produced by the random potential landscape as it exists in the channel. In this context, it is relevant to note that the observed electrical properties appeared to be very reproducible. The curves plotted in Figs. 5(c) and 5(d) are obtained from the same device but taken on different occasions separated by a thermal cycle to room temperature. In fact, the conductance characteristics were studied over several months and remained essentially the same, even after thermal cycling. This is in contrast with previous experimental findings on Coulomb oscillations in disordered silicon and GaAs/AlGaAs wires, where the periods were generally found to change after thermal cycling. Note that in this earlier work Coulomb-blockade oscillations were observed at mK temperatures, implying a much smaller Coulomb energy, and conceivably a larger impact of small changes on the impurity configuration. Our experiments suggest that the potential landscape in our MOSFETs is remarkably robust. Further evidence for the rigidity of the potential landscape in submicron MOSFETs comes from noise measurements done by Mueller et al. that clearly point to current percolation paths as a result of fixed charges even at room temperature.

VII. DISCUSSION

In Sec. VI, we concluded that the conductance oscillations at low temperatures are the result of a disordered potential. This provides a plausible explanation for finding single electron tunneling only in the smallest devices. In order to have clear conductance oscillations, it is essential that the electrical properties are governed by one single island. A small device is then advantageous, since a narrow channel limits the number of parallel percolation paths, while a short length prevents the existence of multiple quantum dots in series. However, single electron tunneling at elevated temperatures requires extremely short devices which will gener-
ally suffer from short-channel effects. Our results show that the presence of the side gates strongly suppresses these effects and restores the long-channel behavior.

The results of the simulations in Fig. 4 indicate that the carrier inversion is not limited to the interface but extends into the entire volume of the channel. It is quite conceivable that this situation persists down to low temperatures. The quantum dot can then no longer be considered to be a disk-like island but rather is a three-dimensional box. At this point, it remains uncertain whether this will substantially affect the electrical properties at low temperature.

VIII. CONCLUSIONS

To summarize, measurements are presented on submicron SOI MOSFETs having a gate on three sides of the channel. At room temperature, a strong suppression of short-channel effects was achieved for the narrowest channels. This is in agreement with calculations of the three-dimensional carrier distribution. According to the simulations, the side gates are able to maintain a sufficient barrier between source and drain for the narrowest devices. At low temperature, the formation of a quantum dot in the disordered potential landscape was demonstrated, making these submicron devices act as a single electron transistor. Our findings prove that the quantum dot is the result of the disordered potential landscape rather than the gate geometry.

ACKNOWLEDGMENTS

The authors would like to thank F. Wollenberg and P. Jurrius (Utrecht) for technical assistance and A. Heringa (Philips Research, Eindhoven) for device simulations. This work was supported by the Dutch foundation “Stichting voor Fundamenteel Onderzoek der Materie” (FOM) and the foundation for Dutch Scientific Research (NWO).