Observation of Electron Velocity Overshoot in Sub-100-nm-channel MOSFET's in Silicon

S. Y. CHOU, STUDENT MEMBER, IEEE, DIMITRI A. ANTONIADIS, SENIOR MEMBER, IEEE, AND HENRY I. SMITH, SENIOR MEMBER, IEEE

Abstract—**n-channel MOSFET's with channel lengths from 75 nm to 5 μm were fabricated in Si using combined X-ray and optical lithographies, and were characterized at 300, 77, and 4.2 K. Average channel electron velocities v_e were extracted according to the equation $v_e = g_{ml}/C_{ox}$, where g_{ml} is the intrinsic transconductance and C_{ox} is the capacitance of the gate oxide. We found that at 4.2 K the average electron velocity of a 75-nm-channel MOSFET is 1.7×10^7 cm/s, which is 1.8 times higher than the inversion layer saturation velocity reported in the literature, and 1.3 times higher than the saturation velocity in bulk Si at 4.2 K. As channel length increases, the average electron velocity drops sharply below the saturation velocity in bulk Si. These experimental results strongly suggest velocity overshoot in a 75-nm-channel MOSFET.**

I. Introduction

OVERSHOOT of electron velocity in Si following rapid changes of electric field was first predicted by Ruch in 1972 using the Monte Carlo calculation of electron scattering [1]. Later, the same effect was also shown by using the relaxation approximation [2] and the electron temperature method [3]. Velocity overshoot is due to the nonequivalence of electron momentum and energy relaxation times, and is expected to occur in a time scale which is shorter than the energy relaxation time. Since the electron energy relaxation time in Si is rather short, velocity overshoot in MOSFET's is expected to become significant only when the distance between source and drain is shorter than 100 nm [1]. To search for velocity overshoot, we fabricated n-channel MOSFET's with channel lengths from 75 nm to 5 μm. The devices were characterised at 300, 77, and 4.2 K. At 4.2 K, we found that a device with a channel length of 75 nm in saturation had an average electron velocity of 1.7×10^7 cm/s, which is 1.8 times higher than the value reported by Fang and Fowler [4] for inversion layers in Si, and 1.3 times higher than the saturation velocity in bulk Si [5] at the same temperature. Moreover, we found that the average electron velocity dropped sharply as channel length increased, strongly suggesting velocity overshoot in the 75-nm-channel device.

II. Experimental

N-channel MOSFET's were fabricated in p-type 20 Ω·cm (100) Si, using X-ray and optical direct-step-on-wafer lithographies [6]. To increase the surface acceptor concentration to 5×10^{17} cm^{-3}, boron was implanted at 30 keV and a dose of 2×10^{13} cm^{-2}, followed by oxidation at 900°C in dry oxygen ambient for 40 min. Instead of using a self-aligned gate structure, PMMA resist lines of widths from 100 nm to 5 μm, patterned by X-ray lithography, were used to mask the channels during source–drain implantation of As, with energy of 30 keV and a dose of 7×10^{15} cm^{-2}. To prevent Al spiking through the shallow junctions at contact areas, deeper junctions were formed under the source and drain contacts by implanting P at 160 keV and a dose of 3×10^{15} cm^{-2}, using photoresist as the implantation mask. After the implants, the photoresist, the PMMA, and the original SiO_2 were removed, and the wafers were oxidized at 900°C in dry oxygen ambient for 25 min. Afterwards, Al gates were patterned and 300 nm of CVD oxide was deposited over the samples. Finally, contact cuts and Al contacts were made to the Al gates and the source–drain regions. The calculated junction depth was about 0.1 μm [7]. The resulting total source-drain series resistance was 4.4 Ω·mm. The thickness of gate oxide measured from capacitors fabricated on the same chip was 11 nm.

Channel lengths were determined by both what we call the "closed-channel" method and by an electrical measurement method. The "closed-channel" method relies on accurate SEM measurement of the width of PMMA lines used as the implantation mask. Because of our X-ray mask fabrication technique [6], [8], the linewidth varied systemetically around each nominal width value across the X-ray mask area by about ±10 percent. Due to lateral dopant spreading during implantation and diffusion during oxidation, devices that came from the narrower PMMA lines (100–160 nm) had their channels closed and behaved like resistors. As the width of PMMA lines widened further, devices started showing transistor behavior. We found that resistor-transistor transistions occurred at PMMA linewidths equal to 160 nm ± 10 nm. This gives a lateral spread of 80 nm for each junction, which is consistent with the TEM results by Sheng and Marcus [9].

The electrical measurement method was based on that proposed by Suciu and Johnston [10], and it gave us both electrical channel length and series resistance for each device. In the calculation, we made a reasonable assumption that the low-transverse-field mobility was the same for all channel lengths. Channel lengths determined from the "closed-channel" and the electrical-measurement methods agree with each other well.

Manuscript received August 7, 1985. This work was supported by the Joint Services Electronics Program.

S. Y. Chou is with the Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139.

D. A. Antoniadis and H. I. Smith are with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139.

0741-3106/85/1200-0665$01.00

III. Results

The devices had well-behaved characteristics down to an effective channel length of 75 nm. The I–V characteristics of such a device at 300 and 4.2 K are shown in Fig. 1.

Fig. 2. shows measured saturation transconductances of devices with different channel lengths at 300, 77 and 4.2 K. All transconductances were measured at drain–source voltage V_{DS} = 1.2 V, and were rather insensitive to V_{DS}. It is interesting to note that the transconductances of devices with channel length greater than 100 nm had temperature dependence consistent with that of the saturation velocities in bulk Si [5]. This is not surprising because the transconductance of devices with such short channel lengths is generally dominated by velocity saturation. However, the temperature dependence of transconductance of the 75-nm device is clearly anomalous.

To be more quantitative, we have calculated the average electron velocity v_e from the equation

$$v_e = g_{mi}/C_{ox}$$

where g_{mi} is the intrinsic transconductance of the device, which can be calculated from the measured transconductance [14]. The average electron velocities of devices with different channel lengths at three temperatures are given in Fig. 3. These velocities should be compared with the electron saturation velocity in a Si inversion layer. Unfortunately, at present there are no consistent data available on the saturation velocity in inversion layers over the temperature range of our experiment [4], [11]. Therefore, we have used instead the electron saturation velocity in the bulk Si as an upper limit of saturation velocity in the inversion layer. This is justified because surface scattering makes saturation velocity in an inversion layer always smaller than that in the bulk [12]. Fig. 3 shows that at 300 K the average electron velocities for all devices are lower than the upper limit. At 77 K the average electron velocities are still lower than the upper limit. However, at 4.2 K the average electron velocity of the 75-nm-channel device is 1.7×10^7 cm/s, which is 1.3 times higher than the saturation velocity in bulk Si [5], and 1.8 times higher than the inversion layer saturation velocity reported by Fang and Fowler at 4.2 K [4]. As channel length increases the average electron velocity drops quickly below the upper limit. These results strongly suggest velocity overshoot in the 75-nm-channel device.

IV. Discussion

In this section, we explain qualitively why in our case velocity overshoot is significant only at low temperatures. As mentioned in the introduction, velocity overshoot occurs on a time scale shorter than the energy relaxation time. If the time needed for electrons to cross the 75-nm channel is shorter than the energy relaxation time, velocity overshoot can become significant. Table I, third column shows the calculated electron transit times t_t across the 75-nm channel at different temperatures based on average velocities being equal to the bulk saturation velocities. The last column in Table I gives the

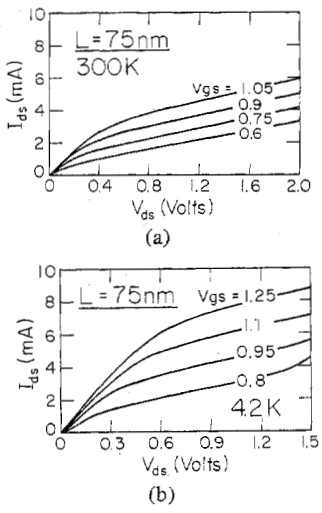

Fig. 1. I–V characteristics of a 75-nm-channel MOSFET (a) at 300 K and (b) at 4.2 K.

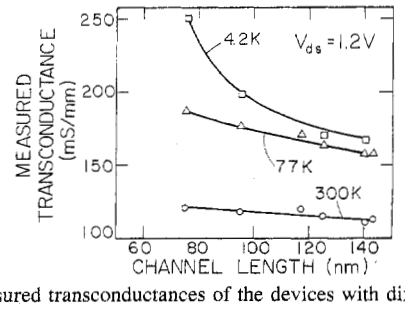

Fig. 2. Measured transconductances of the devices with different electrical channel lengths at 300, 77, and 4.2 K.

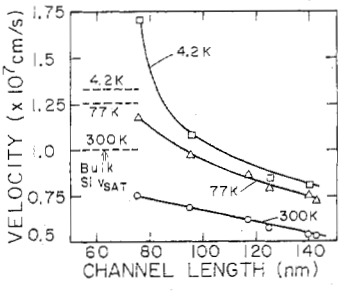

Fig. 3. Calculated average electron velocities of the devices with different electrical channel lengths at 300, 77, and 4.2 K.

TABLE I
MOMEMTUM RELAXATION AND TRANSIT TIMES FOR A 75-nm MOSFET

Temp. (K)	Sat. Velocity v_{sat} (cm/sec)	Transit Time t_t (sec)	Mobility μ_0 (cm²/V·s)	Mom. Rel. Time t_m (sec)
300	1.00×10^7	7.5×10^{-13}	330	0.4×10^{-13}
77	1.25×10^7	6.0×10^{-13}	1100	1.3×10^{-13}
4.2	1.30×10^7	5.7×10^{-13}	2200	2.7×10^{-13}

momentum relaxation time t_m calculated from $t_m = \mu_0 m^*/q$ [13], where μ_0 is the measured low-longitudinal-field electron mobility in our devices, m^* is the effective electron mass in the inversion layer, and q is the electron charge. Although the energy relaxation time cannot be obtained from the present data, it is well known that it is always longer than the momentum relaxation time [3]. We can see from Table I that at room temperature the momentum relaxation time is much shorter than the transit time. Even though the energy relaxation time is longer that the momentum relaxation time, it probably still remains shorter than the transit time. Therefore, velocity overshoot is not significant. However, at 4.2 K the momentum relaxation time is only two times smaller than the transit time, so it is quite likely that the energy relaxation times and transit times are comparable, and thus velocity overshoot becomes significant.

Acknowledgment

The authors would like to acknowledge J. Melngailis for helpful discussions, and J. Carter, P. Maciel, and other members of the technical staff at M.I.T for their assistance at various stages of device fabrication.

References

[1] J. G. Ruch, *IEEE Trans. Electron Devices*, vol. ED-19, p. 652, 1972.
[2] R. S. Huang and P. H. Ladbrooke, *J. Appl. Phys.*, vol. 48, p. 4791, 1977.
[3] R. K. Cook and J. Frey, *IEEE. Trans. Electron Devices*, vol. ED-29, p. 970, 1982.
[4] F. F. Fang and A. B. Fowler, *J. Appl. Phys.*, vol. 41, p. 1825, 1970.
[5] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. New York: Wiley, 1981, p. 46.
[6] S. Y. Chou, Henry I. Smith, and D. A. Antoniadis to be published in *J. Vac. Sci. Technol. B*, Nov./Dec. 1985.
[7] D. A. Antoniadis, S. E. Hansen, and R. W. Dutton, "SUPREM II-A program for IC process modeling and simulation," Stanford University, Stanford, CA, Tech. Rep. SEL 78-020, June 1978.
[8] S. Y. Chou, H. I. Smith, and D. A. Antoniadis, in *Proc. 29th Int. Symp. Electron. Ion Photon Beams* (Portland, OR); also *J. Vac. Sci. Technol. B*, Jan./Feb. 1986.
[9] T. T. Sheng and R. B. Marcus, *J. Elecrochem. Soc. Solid-State Sci. Tech.*, vol. 128, p. 881, 1981.
[10] P. I. Suciu and R. L. Johnston, *IEEE Trans. Electron Devices*, vol. ED-27, p. 1846, 1980.
[11] J. A. Cooper Jr. and D. F. Nelson, *IEEE Electron Device Lett.*, vol. EDL-2, p. 171, 1981.
[12] T. Ando, A. B. Fowler, and F. Stern, *Rev. Mod. Phys.*, vol. 54, p. 457, 1982.
[13] K. Seeger, *Semiconductor Physics*. New York: Spring-Verlag, 1973, p. 52.
[14] S. Y. Chou and D. A. Antoniadis, "Relationship between measured and intrisic transconductances of FET's," submitted to *IEEE Trans. Electron Devices*.