# Growth of $Si_{1-x}Ge_x$ by rapid thermal chemical vapor deposition and application to heterojunction bipolar transistors

J. C. Sturm, P. V. Schwartz, E. J. Prinz, and H. Manoharan
*Department of Electrical Engineering, Princeton University, Princeton, New Jersey 08544*

(Received 26 November 1990; accepted 26 February 1991)

Rapid thermal chemical vapor deposition has been applied towards the growth of Si and $Si_{1-x}Ge_x$ structures on a 100 Å scale. In this paper the relative merits of gas switching versus temperature switching for the growth of such structures are discussed. Active temperature control in the 600–700 °C range using infrared transmission for temperature measurement is demonstrated. The growth technique is applied to 45 Å period superlattices with individual layer temperature control, and to heterojunction bipolar transistors with near-ideal electrical characteristics.

## I. INTRODUCTION

The steady progress in the scaling of integrated circuit features to reduced dimensions has been accompanied by the introduction of new techniques for the more precise control of the formation of vertical junctions. A key driver for such vertical profile control has been the bipolar transistor. In the 1970's typical basewidths were $\sim 1\ \mu m$, and the junctions were fabricated by diffusion. In the 1980's typical basewidths were on the order of 1000's of Å, and critical profiles were formed by ion implantation. Because of problems such as straggle, channeling, and diffusion during annealing, it appears unlikely that ion implantation can be used to create arbitrary junction profiles on a scale much below 1000 Å. Furthermore, the formation of heterojunctions such as $Si_{1-x}Ge_x$ layers on silicon will also be difficult by implantation.

In this paper the technique of rapid thermal chemical vapor deposition (RTCVD) is applied towards the growth of layers on a 100 Å scale with an interface abruptness of $\sim 10$ Å. Rapid temperature switching and rapid gas switching approaches are contrasted, and the use of infrared transmission for monitoring water temperature for feedback control is demonstrated. The technique is then applied to the growth of a superlattice structure and heterojunction bipolar transistors (HBTs).

## II. GROWTH APPARATUS

The work described in this paper was performed in a homemade reactor shown schematically in Fig. 1. The growth chamber consists of a cylindrical quartz tube with a diameter of 175 mm. On one end the tube diameter is reduced to $\sim 25$ mm and the gas inlet connection is made with a compression O-ring fitting. The other end is connected via a flange and O rings to a stainless steel assembly, through which the exhaust gas is pumped by a mechanical rotary vane pump with a pumping speed of 28 cfm. A butterfly valve is used between the chamber and the pump for control of the reactor pressure. The wafers are also loaded into the reactor from this end, but a load lock and transfer mechanism are used so that the growth chamber is not vented to atmospheric pressure during the sample loading. The wafer is suspended in the growth chamber by several quartz pins without a susceptor. It is heated from one side by a bank of twelve 6 kW tungsten–halogen lamps, with the entire tube and lamps enclosed in a water-cooled gold-plated reflector. Although 72 kW are available at full power, over 30 kW is rarely used.

The process gas flows are adjusted by mass flow controllers, and five-ported valves are used so that the process gas flows can be rapidly switched on or off without having to wait for the flow controllers to stabilize. The process gas flows are first established into a vent line which bypasses the reactor. Rapid gas switching can then be accomplished by switching this flow from the vent line to a line which flows to the reactor chamber (or vice versa to turn off a gas). The hydrogen is purified by conventional diffusion through palladium, and the other process gases are filtered only for particulates.

Typical epitaxial growth conditions are a pressure of 6 Torr with a 3 lpm hydrogen carrier flow and 26 sccm of dichlorosilane. Diborone and phosphine (each $\sim 50$ ppm in hydrogen) are used for *in situ* doping, and 0.8% germane in hydrogen is used in addition to dichlorosilane for silicon–germanium alloy growth.

## III. TEMPERATURE SWITCHING VERSUS GAS SWITCHING

Initial epitaxial growth experiments in a lamp-heated, susceptor-free, chemical vapor deposition (CVD) reactor were called limited reaction processing (LRP).[1] The central fea-

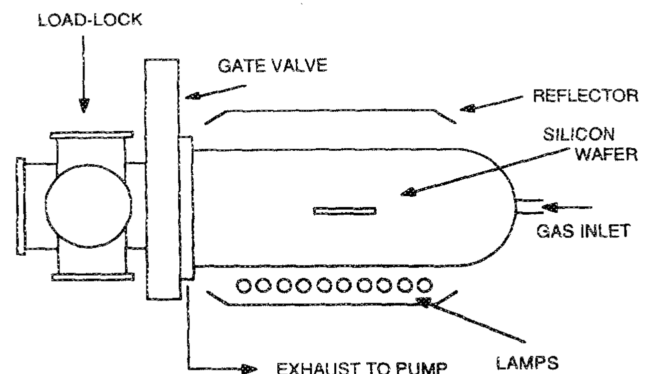


FIG. 1. Schematic diagram of the RTCVD reactor used in this work.

 © 1991 American Vacuum Society 

ture of LRP is the use of the wafer temperature (as opposed to the gas flows in conventional CVD) to start and stop the growth reaction [Fig. 2(a)]. The flows of reactive gases were first established when the wafer was cold, and hence no growth was occurring. To start the growth, the wafer temperature was then rapidly switched to the growth temperature (900–1000 °C in initial experiments) at a ramp rate of about 300 °C/s. To stop the growth, the lamps were switched off and the wafer temperature rapidly dropped, effectively turning off the thermally activated growth reaction. By changing gas flows while the wafer was cold, multiple layers, such as $i$–$p^+$–$i$ doping structures or epitaxial silicon–oxide–polysilicon structures for metal–oxide-semiconductor (MOS) capacitors could be grown.[2,3] The fundamental advantage of the LRP approach is that the thermal exposure of the wafer is held to an absolute minimum, and hence any excess thermal diffusion that would occur if the wafer were exposed to the process temperature for some time before or after the actual growth is avoided.

These initial results yielded an interface abruptness that was nearly indistinguishable by conventional secondary ion mass spectroscopy (SIMS) measurements from that in samples grown by molecular-beam epitaxy (MBE) at much lower temperatures (~600 °C). However, by varying the energy of the SIMS sputtering beam, it is possible to separate the true interface abruptness from SIMS ion mixing artifacts. Such measurements on thin $i/p/i$ layers (each layer ≈500 Å) grown at 900 and 1000 °C by LRP showed the doping interface abruptnesses to be 45 and 160–215 Å/decade, respectively.[4] Clearly, this is not sufficient for well resolved 100 Å layers. To demonstrate that these transition widths are consistent with conventional process modeling (SUPREM-III[5]), the growth of alternating 100 Å intrinsic silicon and boron-doped ($1\times10^{18}$ cm$^{-3}$) silicon layers was simulated at 1000 °C. A typical LRP growth rate of 0.25 μm/min at 1000 °C was used, and the only time at the process temperature was that for the actual growth, as in LRP (2.4 s/layer). The resulting profiles (Fig. 3) clearly show that because of excess thermal diffusion, growth temperatures of 1000 °C cannot be used for devices on the 100 Å scale, even with rapid temperature switching.

At lower temperatures the growth rate of silicon by CVD falls off rapidly, and obeys an Arrhenius relationship with a typical activation energy of 2.0 eV. Data for our system are shown in Fig. 4. 2 eV is a smaller activation energy than those of typical substitutional diffusers in silicon (~3–5 eV). Therefore, although longer growth times are required at lower temperatures, thermal diffusion effects will be reduced. The improvement is evident in Fig. 3, which shows the simulated profile for growth of the same $p/i/p/i$ structure at 800 °C (using a growth rate of 200 Å/min). Based on this work, we can assume that a growth temperature of about 800 °C is an upper limit to the growth of well-resolved structures on a 100 Å scale.

At temperatures under 800 °C, typical growth rates are 100 Å/min or less, requiring 100 s or more for 100-Å thick layers. With such long growth times, several extra seconds at the growth temperature do not significantly contribute to any thermal diffusion. This removes the motivation to switch the growth reaction on and off by changing the wafer temperature. Instead, in our experiments the wafer is first brought to the growth temperature in pure hydrogen, and then the growth is switched on and off using fast switching of

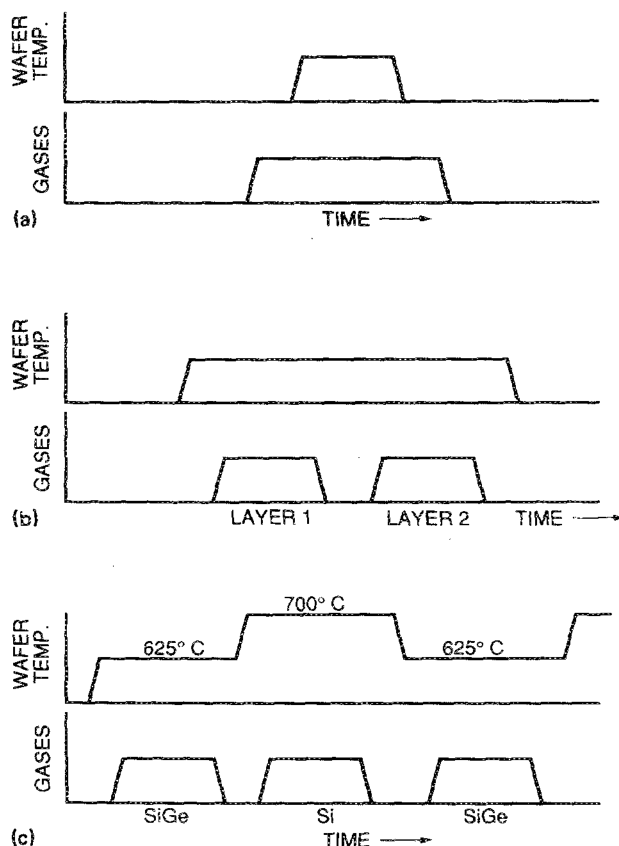


FIG. 2. Schematic diagram of temperature and gas flows vs time for (a) limited reaction processing, (b) RTCVD with a conventional CVD approach, and (c) RTCVD with individual layer temperature control.

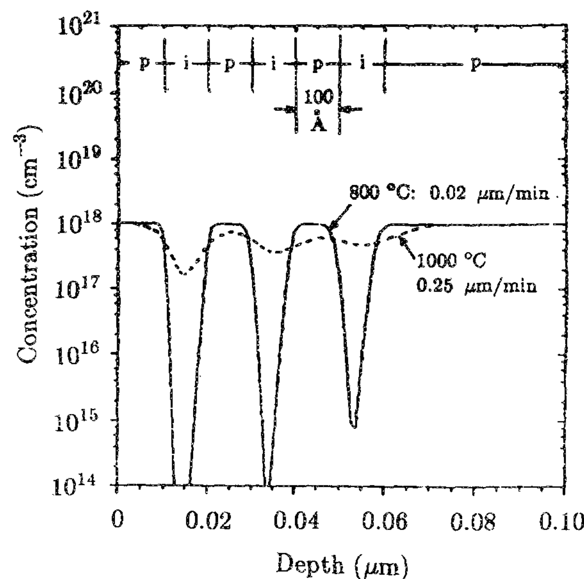


FIG. 3. Simulated dopant profiles for a $p/i/p/i$ structure grown at 1000 and 800 °C with no excess thermal exposure.

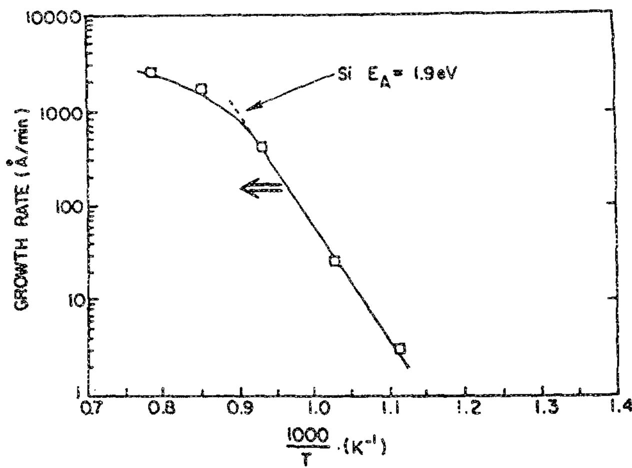


FIG. 4. Silicon growth rate (from 0.9% dichlorosilane in hydrogen at 6 Torr) vs inverse temperature in our reactor.

the process gases [Fig. 2(b)]. This approach removes any risk of low quality growth below the desired temperature, as might occur with LRP during the temperature ramp up or ramp down. *P–n* junctions can be formed by simply switching the doping gases, and a low-temperature growth interruption is not required. Given our chamber volume of ~10 l, a pressure of 6 Torr, and a carrier gas flow of 3 slpm, the gas residence time is only 2 s. Coupled with a growth rate of 100 Å/min, this residence time gives control on a scale of better than 10 Å.

The ability to rapidly switch the sample temperature is still useful for optimizing the growth temperature of individual layers in a multilayer structure. For example, to achieve a silicon growth rate of 30 Å/min, a temperature of 700 °C is required for our growth conditions. To avoid islanding and to achieve metastable strained layers beyond the critical thickness during SiGe growth on Si, we have found that a temperature under 650 °C (typically 600–625 °C) is required. Reasonable (~100 Å/min) growth rates are still possible at 600–625 °C because of a fortuitous catalytic reaction of germane on silicon CVD growth.[6,7] In conventional CVD, one would have to grow both the Si and SiGe layers at a single nonoptimum temperature. However, with RTCVD, one can grow one layer, turn off the reactive process gas flows (leaving the hydrogen carrier on), rapidly ramp directly to the new growth temperature, and then start the growth of the new layer [Fig. 2(c)].

## IV. TEMPERATURE CONTROL BY INFRARED TRANSMISSION, SUPERLATTICE GROWTH

Since the growth rate at low temperatures is an exponential function of temperature, accurate control of the absolute temperature is critical to the success of RTCVD. Exact measurement of the absolute temperature by pyrometry is virtu-

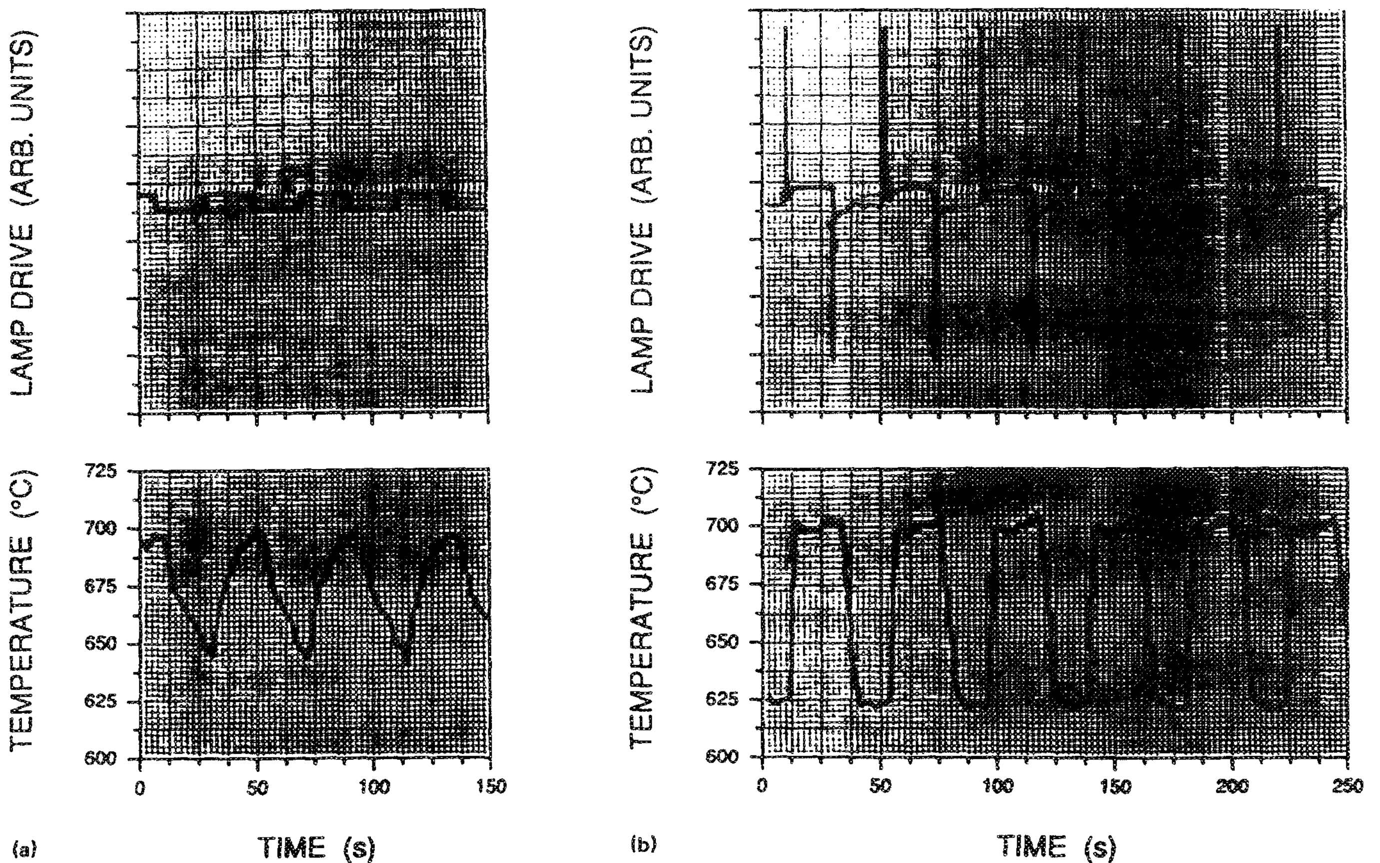


FIG. 5. Time response of the wafer temperature and lamp drive signal for (a) open-loop control, and (b) closed-loop temperature control using infrared transmission.

ally impossible because of the many factors that affect the emissivity of a silicon wafer (backside polish, temperature, field oxides, etc.).[8,9] Therefore we have developed a technique for measuring the optical absorption of the silicon substrate at 1.3 and 1.55 $\mu$m *in situ*.[10] By the use of lock-in techniques and normalization, the effects of lamp radiation interference and wafer backside roughness are completely removed from the measurements. The absorption coefficients are almost exponential functions of temperature, so that the light transmitted through the wafers can be used to easily infer the wafer temperature to within a few °C. Furthermore, the growth of thin $Si_{1-x}Ge_x$ layers on the wafer surface has little affect on transmission, so the technique can be useful for the growth of SiGe layers on silicon substrates.[11]

Using this technique for probing the wafer temperature, one can then implement a feedback loop to control the lamp power. Shown in Fig. 5(a) is the typical open-loop response of our system, where the lamp drive signal is changed between its approximate value for 625 and 700 °C in 20 s intervals. The natural response time of the system is about 20 s. Shown in Fig. 5(b) are the wafer temperature and lamp drive signal which result when a simple proportional-integral-derivative (PID) feedback loop is used to control the lamp power, and the desired setpoint is alternated between 625 and 700 °C in 50 s intervals. The typical transition time is ~5 s.

The use of this temperature control technique for the optimization of growth temperature in each layer was demonstrated by the growth of a 50 period superlattice (Fig. 6). Each period consisted of a 24 Å $Si_{0.8}Ge_{0.2}$ layer grown at 625 °C (20 s growth), and a 23 Å Si layer grown at 700 °C (45 s growth). Between each two layers the temperature was directly switched from 625 to 700 °C or vice versa, with a total interruption time of ~15 s allowed for the gas purging and temperature switching. The growth time for the total superlattice was 75 min.

Shown in Fig. 7(a) is the measured x-ray diffraction data of this sample for the (400) reflection. The substrate, superlattice, and satellite peaks are clearly visible. From this data a period of 46 Å and an average superlattice composition of $Si_{0.90}Ge_{0.10}$, were extracted, in good agreement with the targeted 47 Å and $Si_{0.90}Ge_{0.10}$, respectively. One also notes that the largest satellite by far is the $n = -1$; $n = \pm 2$ and $+1$ are nonexistent. This is not indicative of asymmetry in the structure, but is indeed expected based on first principles kinematic simulations of the spectra [Fig. 7(b)] of an ideal structure with abrupt interfaces. The desired structure is also confirmed by a cross section transmission electron micrograph (TEM) shown in Fig. 8. The individual layers are well resolved, and from the micrograph one can put an upper limit of about 10 Å on the interface abruptness. The variation in the period (averaged over 10 periods) from the top to the bottom of the structure is less than 5%. From this it may be inferred that any drift in the sample temperature over the 75 min, with 100 switching events, was less than 3 °C.

## V. HETEROJUNCTION BIPOLAR TRANSISTORS

In Si/SiGe/Si *npn* HBTs, the narrow band gap base reduces the barrier seen by electrons as they travel from emitter to collector. In a device with abrupt box-like doping and germanium profiles, the collector current in a HBT compared to an otherwise identical all silicon homojunction de-

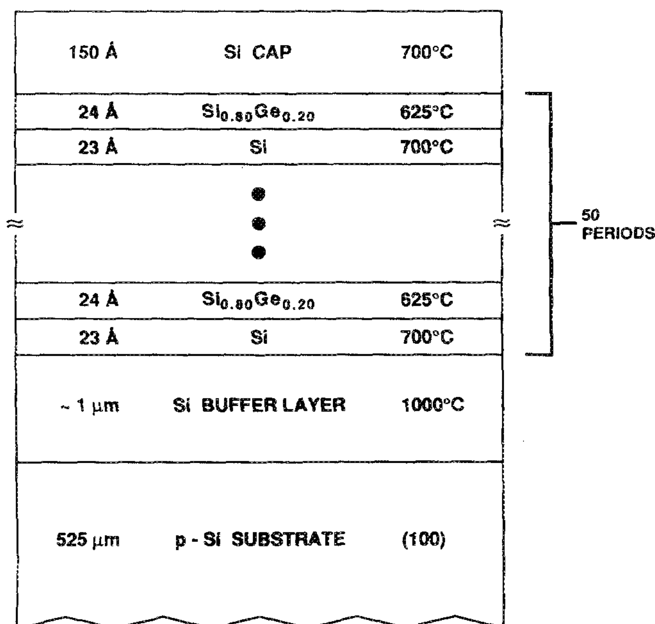


FIG. 6. Cross section of a 50 period strained layer superlattice structure grown using optimized temperatures for each layer.

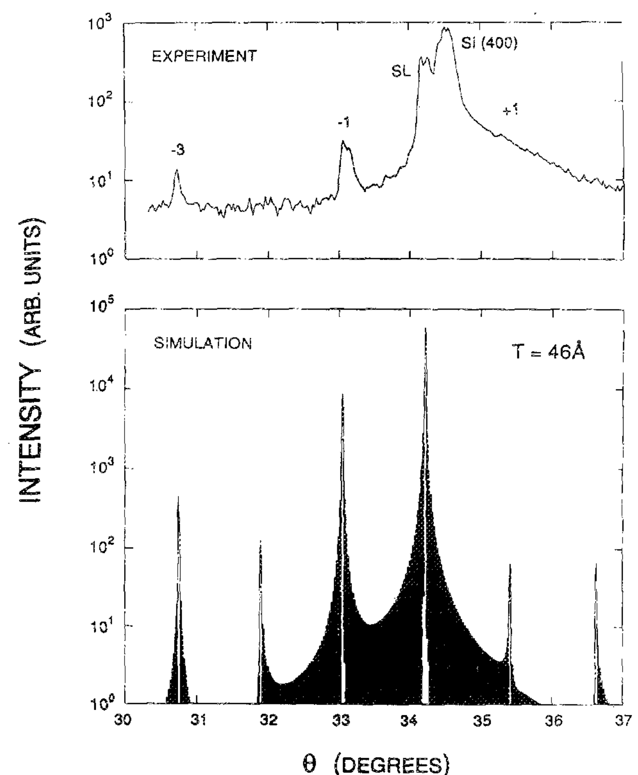


FIG. 7. Measured and simulated x-ray diffraction spectra (Cu k$\alpha$ radiation) for the superlattice structure of Fig. 6. (The "$n = -3$" peak may be from the substrate k$\beta$ reflection and not the superlattice.)

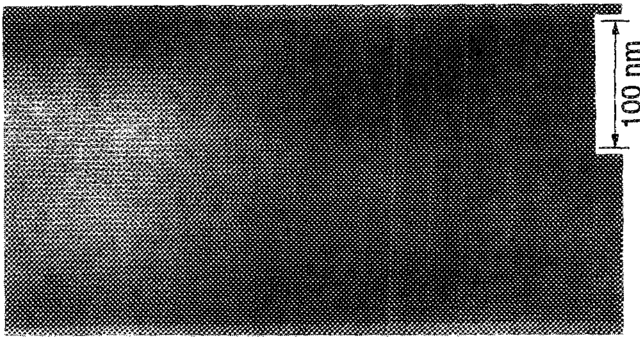


FIG. 8. Cross-section TEM of the superlattice of Fig. 6.

vice at the same bias conditions should be enhanced by a factor of $e^{\Delta E_V/kT}$, where $\Delta E_V$ is the valence band offset at the SiGe/Si interface.[12] However, because heavily doped bases (boron) and lightly doped emitters and collectors are commonly used, any small amount of boron outdiffusion will move the $p$–$n$ junctions away from the Si/SiGe interfaces and into the silicon[13] (Fig. 9). This could also be caused by autodoping, etc., during the growth of the base-emitter junction. The movement of the junction away from the interface introduces parasitic barriers into the conduction band which will drastically reduce the desired collector current enhancement (Fig. 9). By comparing the actual device performance to the simulations for various amounts of outdiffusion, one can infer the abruptness of the boron transition at the Si/SiGe interfaces.

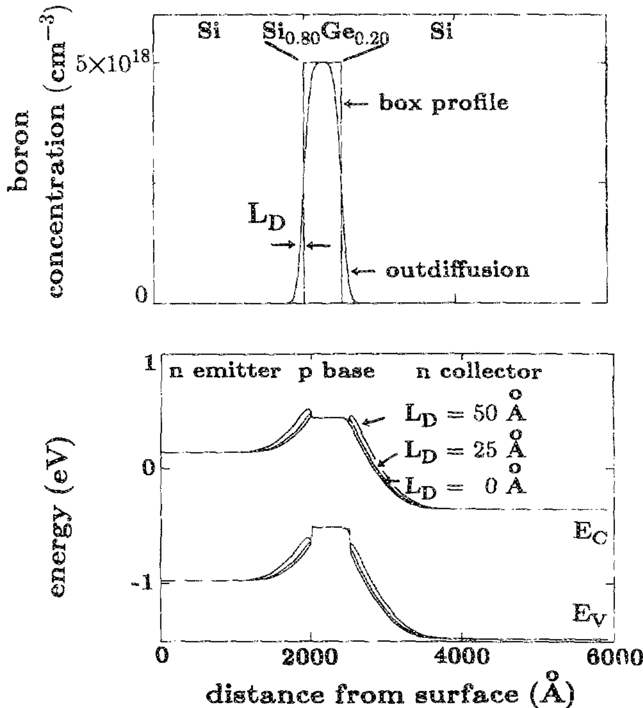


FIG. 9. Schematic view of dopant outdiffusion from a heavily doped base and its effect on band diagrams for various diffusion lengths.

$Si/Si_{0.80}Ge_{0.20}/Si$ *npn* HBTs were formed by growing four sequential layers *in situ*: an $n^+$ layer ($10^{19}$ $cm^{-3}$, $\sim 1$ $\mu m$, 1000 °C), an $n$-type collector layer ($\sim 5\times10^{17}$ $cm^{-3}$, 0.3 $\mu m$, 1000 °C), a $p$-type $Si_{0.80}Ge_{0.20}$ base ($\sim 5\times10^{18}$ $cm^{-3}$, 300 Å, 625 °C), followed by the emitter ($5\times10^{17}$ $cm^{-3}$, 3000 Å, 850 °C). The relatively high silicon growth temperatures were used to provide easier control of the $n$-type doping. Contacts to the base and emitter were formed by ion implantation and annealing (800 °C, 10 min). The base-collector junction was mesa isolated and the base-emitter junction was isolated by a $p$-implant ring as described by King.[12] The SiGe layer was fully strained at the completion of the processing.

Gummel plots measured at room temperature of finished devices with an emitter size $60\times60$ $\mu m^2$ are shown in Fig. 10. The base current slope is near ideal (65–70 mV/decade) and the collector current slope is ideal (60 mV/decade). A common-emitter current gain of more than 1000 results over several decades of current. It should be noted that the base current density of our device is the same as that of an all-silicon device with a similar structure.[12] This is expected since base current is ideally dominated by hole injection from the base to the emitter, and there is virtually no change in the total valence band barrier heights from base to emitter when going from an all silicon device to a Si/SiGe/Si *npn* HBT. The Si/SiGe/Si HBTs of Ref. 12 had a base current which was about 50 times higher than the all-Si device, however. This improvement in the base currents in our devices is probably due to the low oxygen concentration in the SiGe base ($\leqslant 10^{18}$ $cm^{-3}$ in our device). The devices of Ref. 12 were grown under similar conditions and temperatures as ours, but without a load lock, which resulted in an oxygen level in the base of $\sim 10^{20}$ $cm^{-3}$. Presumably this large oxygen level caused a low lifetime which increased the recombination of electrons as they crossed the base, leading to en-

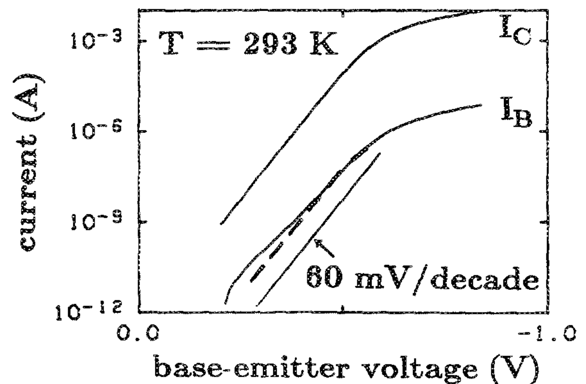


FIG. 10. Gummel plot of the $Si/Si_{0.80}Ge_{0.20}/Si$ *npn* HBT. The dotted line represents the base current of Ref. 13 for a similar all-silicon device.

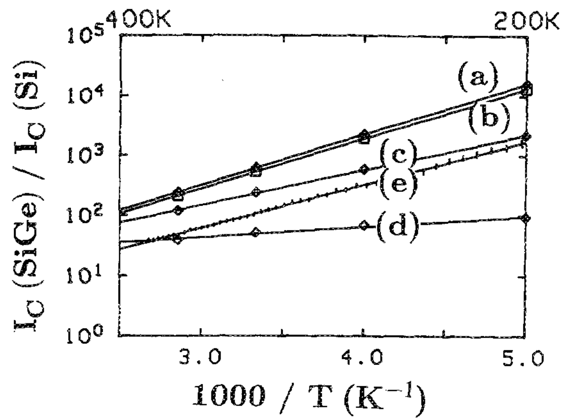


FIG. 11. Collector current enhancement vs temperature for a $Si/Si_{0.80}Ge_{0.20}/Si$ HBT. Simulation for (a) no dopant outdiffusion, (b) $L_D = 25$ Å, (c) $L_D = 50$ Å, (d) $L_D = 75$ Å, and (e) measured values.

hanced base current. Similar ideal base current results have also been reported in transistors grown by the ultrahigh vacuum (UHV)–CVD technique.[13] Our results demonstrate that ultrahigh vacuum techniques are not required to grow low-oxygen films by CVD with high lifetime, however.

Shown in Fig. 11 is the temperature dependence of the collector current enhancement of the HBT versus inverse temperature. Also shown are simulated values for different amounts of boron outdiffusion ($L_D = \sqrt{Dt}$), assuming that the boron diffusion constant in the SiGe is the same as that in Si with the same doping level. The slope of the data matches the simulated curves for $L_D = 25$–30 Å, indicating that any broadening of the boron profile at the end of processing is characterized by an $L_D$ of at most 30 Å. The shift of the data down by a factor of 5 from the simulated values results from the fact that the effects of density of state differences between silicon and silicon–germanium were not included[14] and from uncertainty in the exact base doping. Straightforward process simulation of thermal diffusion during the high-temperature processing after growth (800 °C, 10 min) predicts that the processing contributed a $\sqrt{Dt}$ of 20 Å to the broadening. By taking the difference between this $Dt$ and the $Dt$ consistent with the final structure, one finds the interface abruptness in the as grown structure to be similar to that of an originally abrupt interface which has been broadened by $\sqrt{Dt} \approx 20$ Å. This nonideality in the as grown structure could result from thermal diffusion during the emitter growth (800 °C, 3 min) or from autodoping, slow gas switching, etc. during the interface formation. However, process simulations show that thermal diffusion during the emitter growth is responsible for virtually all of the nonabruptness in the as grown interface. On a final note, it should also be pointed out that if the Ge profile also broadened during the emitter growth and subsequent processing, the effect of the boron diffusion might be less severe, causing us to underestimate the boron transition width. Based on the diffusion data of Ref. 15 one can estimate that the germanium profile will broaden by about 10 Å for our growth and processing conditions. This effect is therefore small compared to the boron diffusion in our finished device.

## VI. CONCLUSIONS

Rapid thermal chemical vapor deposition (RTCVD) has been shown capable of growing high quality silicon and silicon–germanium layers on a 100 Å scale with an interface abruptness on the order of 10 Å. For such structures, rapid temperature switching is not required to start and stop the growth cycles, but is useful for optimizing the growth temperature of individual layers in multilayer structures. Infrared transmission may be used for accurate control of the absolute wafer temperature. HBTs with near-ideal electrical characteristics have been fabricated, demonstrating that ultrahigh vacuum technology is not required for the growth of high-quality silicon–germanium alloys.

## ACKNOWLEDGMENTS

The support of ONR (N00014-90-J-1316), NSF (ECS-86157227), and IBM (E.J. Prinz) is gratefully acknowledged. The assistance of C. Magee of Evans East for SIMS and J. McCaffrey of the Institute of Microstructural Sciences, National Research Council, Canada, for expert TEM characterization is also appreciated.

[1] J. F. Gibbons, C. M. Gronet, and K. E. Williams, Appl. Phys. Lett. **47**, 721 (1985).
[2] C. M. Gronet, J. C. Sturm, K. E. Williams, J. F. Gibbons, and S. D. Wilson, Appl. Phys. Lett. **48**, 1012 (1986).
[3] J. C. Sturm, C. M. Gronet, and J. F. Gibbons, IEEE Electron Device Lett. **EDL-7**, 577 (1986).
[4] J. E. Turner, J. Amano, C. M. Gronet, and J. F. Gibbons, Appl. Phys. Lett. **50**, 1601 (1987).
[5] C. P. Ho, S. E. Hansen, and P. M. Fahey, "SUPREM-III, a Program for Integrated Circuit Process Modelling and Simulation," Stanford Electronics Laboratories Technical Report SEL 84-001, July 1984.
[6] B. S. Meyerson, K. J. Uram, and F. K. LeGoues, Appl. Phys. Lett. **53**, 2555 (1988).
[7] P. M. Garone, J. C. Sturm, and P. V. Schwartz, Appl. Phys. Lett. **56**, 1275 (1990).
[8] T. Sato, Jpn. J. Appl. Phys. **6**, 339 (1967).
[9] D. W. Pettibone, J. R. Suarez, and A. Gat, Proc. Symp. Mater. Res. Soc. **52**, 182 (1986).
[10] J. C. Sturm, P. V. Schwartz, and P. M. Garone, Appl. Phys. Lett. **56**, 961 (1990).
[11] J. C. Sturm, P. M. Garone, and P. V. Schwartz, J. Appl. Phys. **69**, 542 (1991).
[12] C. A. King, J. L. Hoyt, and J. F. Gibbons, IEEE Trans. Electron Devices **EDL-36**, 2093 (1989).
[13] G. L. Patton, D. L. Harame, J. M. C. Stock, B. S. Meyerson, G. J. Scilla, and E. Ganin, IEEE Electron Device Lett. **EDL-10**, 534 (1989).
[14] E. J. Prinz, P. M. Garone, P. V. Schwartz, X. Xiao, and J. C. Sturm, IEEE Electron Device Lett. **EDL-12**, 42 (1991).
[15] G. F. A. van de Walle, L. J. van Ijzendoorn, A. A. van Gorkun, R. A. van der Heuvel, and A. M. L. Theunissen, Semicond. Sci. Technol. **5**, 345 (1990).