Optimization of external poly base sheet resistance in 0.13 µm quasi self-aligned SiGe:C HBTs


IMEC, Kapeldreef 75, B-3001 Leuven, Belgium

¹ Also with KULeuven-ESAT, Kasteelpark Arenberg 10, B-3001 Leuven, Belgium.

ABSTRACT

This paper investigates the optimization of the external polysilicon base sheet resistance of quasi self-aligned (QSA) SiGe:C HBTs from a 0.13 µm BiCMOS process. Taking advantage of optimized implant conditions to improve the doping of the external base poly, and using an optimized non-selective epitaxy process with improved growth rate ratio of 1.7 between the polycrystalline silicon and monocrystalline silicon of the base, the maximum oscillation frequency $f_{\text{max}}$ reaches 300 GHz.

Keywords : SiGe:C HBT ; Maximum oscillation frequency ; Non-selective epitaxy

1. Introduction

SiGe:C heterojunction bipolar transistors (HBTs) have demonstrated cut-off frequencies, $f_t$ and $f_{\text{max}}$, higher than 200 GHz with quasi self-aligned (QSA) architecture [1-3] and fully self-aligned (FSA) architecture [4-9]. These performances and the compatibility of SiGe:C HBTs in the standard CMOS platform qualify SiGe:C BiCMOS technology as an effective approach for demanding high speed applications like 60 GHz ultra wide band communication and 77 GHz car radar [10]. QSA HBTs using non-selective epitaxial growth (NSEG) for the base layer deposition take advantage of low topography, low process complexity and ease of integration. As shown in Fig. 1, in the QSA HBT, the base consists of the intrinsic base, the external base and the link base between the intrinsic and the external part. NSEG forms the intrinsic base and the external base in one step and provides low process complexity compared to selective epitaxial growth (SEG) of the base.

The base resistance $R_b$ directly influences $f_{\text{max}}$ of HBTs as described in Eq. (1) [11]. Figs. 2 and 3 collect the data from the measurements in our experiments before the optimization of the external poly base, which we are going to demonstrate in this paper. The symbols in Figs. 2 and 3 represent the measurement results, and the arrowed lines serve as eye-assist lines for the optimization. Fig. 2 shows the trend that when $f_t$ and $C_c$ keep unchanged, $f_{\text{max}}$ increases as $R_b$ drops. In QSA HBTs, $R_b$ is the sum of the intrinsic base, the link base, and the external base resistance and therefore depends on the external base sheet resistance ($R_{\text{Sh,ext}}$), as shown in Fig. 3. In order to keep the high $f_t$ unchanged, we focus on the external part of the base resistance to improve $f_{\text{max}}$ of the device.

The base layer of QSA HBTs is grown monocrystalline in the active regions, and polycrystalline in the field regions. Device requirements determine the composition of the base layer in the active regions. Consequently, the composition of the polycrystalline external base regions is frozen. To avoid the complex process of SEG of the base, we focus here on the optimization of implant conditions and the thickness of the external base of our non-selective process. We will first describe our QSA HBT technology, and then report the optimization of the implant and growth conditions of the external base layer. The optimized process results in a thin intrinsic base for high peak $f_t$ and a thick highly doped external poly base for low $R_b$, and hence high $f_{\text{max}}$.

$$f_{\text{max}} = \sqrt{\frac{f_t}{8\pi R_b C_c}}$$  \hspace{1cm} (1)

2. Optimization of the external base sheet resistance

2.1. Fabrication of QSA SiGe:C HBTs

The QSA SiGe:C NPN HBTs in this study are integrated in a 0.13 µm SiGe:C BiCMOS technology [1,12-14]. The fabrication process starts with the subcollector implantation and an $\pi$-type epitaxial layer grown on top of the p-substrate, followed by the formation of deep and shallow trenches. After fabrication of the CMOS devices, the protection layer is opened and the processing of the bipolar transistor starts. The epitaxial base is non-selectively grown as a stack of thin undoped Si buffer layer, the actual SiGe:C base layer, and a thin undoped Si cap layer. This layer grows as a monocrystalline stack in the active region, and as a polycrystalline stack on top of the shallow trench isolation (STI) field oxide. A thin base oxide and nitride layer are deposited and the emitter...
window opened by etching the nitride. An arsenic in-situ doped poly emitter is deposited after an HF-dip, followed by masked etch of the polyemitter and the nitride. The external base is implanted with the photoresist still on top of the emitter. The implantation is self-aligned to the patterned poly emitter (quasi self-aligned to the emitter window). The implant dose is 4 \times 10^{15}/\text{cm}^2. After removal of the photore sist on top of the emitter and formation of nitride spacers, the poly base is patterned. The final junction anneal is a spike RTA at 1085 °C. Nickel silicidation completes the front-end of line process.

2.2. Optimization of the external poly base implant

The external base is implanted through the base oxide into the thin poly base layer. Therefore, the optimization of the external base implantation was investigated in four aspects, which are implant dose, implant energy, thickness of the base oxide and the Si cap layer. Increasing the dose above 4 \times 10^{15}/\text{cm}^2 with single implant does not reduce the sheet resistance by more than 5%, even when the dose is doubled or tripled. The implant energy is critical, as too much doping will be lost in the base oxide at too low energy, and too much will be implanted below the base poly at too high energy. As the base layer is thin, the process window is narrow.

The thickness of the base oxide and the Si cap has impact on the position of the peak of the implantation. If the peak of the implantation is nicely in the middle of the poly base layer, a maximum amount of dopants is incorporated in the layer. However, if the peak is shifted towards one of the edges of the layer, a lot of dopants are lost either in the base oxide or in the field oxide underneath the poly base. As shown in Fig. 4, with single implantation at a constant energy of 10 keV, the sheet resistance is lower with a thinner base oxide or a thicker Si cap. The Si cap however accommodates the dopants diffused from the emitter after RTA, and its thickness is optimized in order to position the Si/SiGe heterojunction with respect to the base/emitter pn-junction. Therefore, the Si cap cannot be increased beyond this optimum thickness to reduce the external base sheet resistance, because this would seriously impact the current gain and \( f_t \).

The external base sheet resistance can however be reduced by using a higher implant dose on the condition that the implant is more homogeneously spread over the thickness of the layer to lower the amorphization and hence to improve the activation of the dopants. Therefore we investigate multiple implantations, each using a dose of 4 \times 10^{15}/\text{cm}^2 but at different energies. The impact of multiple implantations is shown in Fig. 4. Compared to a single implantation, double implantation using 10keV and 15keV as implant energy, results in a higher active dopant concentration. Adding a third implant dose at 7 keV does not significantly further improve the sheet resistance. Therefore, a double implant at 10 keV and 15 keV is the option we choose.

2.3. Optimization of the epitaxy process

Fig. 5 shows that, including Ge or not in the poly base does not influence \( R_{sh,ext} \). Only the thickness of the total stack has an impact on \( R_{sh,ext} \). For instance, using an intrinsic base thickness of 47 nm including 12 nm Si cap, a minimal poly base thickness of 60 nm is required in order to guarantee a \( R_{sh,ext} \) lower than 300 \( \Omega/\text{sq} \), enabling \( f_{\text{max}} \) to reach 300 GHz. For the QSA architecture, which uses non-selective epitaxy for the base deposition, a high growth rate ratio between polycrystalline silicon and monocrystalline silicon is needed to achieve a thicker external poly base while keeping the intrinsic base thin. Epitaxy was performed using a standard horizontal cold wall ASM Epsilon™ 2000 reactor, a chemical vapor deposition system designed for production applications. Silane and germane were used as precursors in hydrogen ambient. The growth of the intrinsic monocrystalline SiGe:C base stack requires 6 epitaxial deposition steps for a total thickness of 47 nm in the present architecture when a Si cap thickness of 12 nm is used. Because near the half (\( \sim 47\% \)) of the stack thickness consists of Si deposition, a cost-effective approach is to focus the optimization effort on the growth conditions at 650 °C for the buffer layer and for the cap layer. We started the development work on blank wafers, using Si (100) substrates and template wafers with 500 nm oxide. The growth rates of the monocrystalline layer and of the polycrystalline layer were studied as function of the silane partial pressure for different values of the total pressure. In the total pressure \( p_t \) range 20-120 Torr, the mono-Si growth rate as well as the poly-Si growth rate increase linearly with the silane partial pressure \( p_{\text{SiH}_4} \) and the increasing rate decreases for larger \( p_t \). As summarized in Fig. 6, the results show that the growth rate ratio between mono-Si and poly-Si deposition does not depend on \( p_{\text{SiH}_4} \) but on the hydrogen partial pressure \( p_{\text{H}_2} \), the value of which is practically identical to \( p_t \) due to the weak ratio \( p_{\text{SiH}_4}/p_{\text{H}_2} \) [15]. The high quality of the as-grown monocrystalline layers was confirmed by optical and electron microscope inspections. Using a total pressure of 120 Torr, the growth rate ratio for the Si buffer layer and the Si cap layer was increased up to 1.7 which enabled a total poly base thickness of at least 60 nm. The optimized process was finally applied to patterned wafers.

Fig. 7 shows the \( f_t \) and \( f_{\text{max}} \) curves of the reference device before the optimization of the external base and the device with optimized \( R_b \), as function of the collector current (\( I_c \)). \( R_b \) versus \( I_c \) before and after optimization are shown in Fig. 7(b). After optimization, \( R_b \) is reduced by more than 50%, and \( f_{\text{max}} \) reaches 300 GHz while \( f_t \) stays 210 GHz.

2.4. Chemical and active concentration of dopants in the external poly base

Secondary Ion Mass Spectrometry (SIMS) is the standard technique for one dimensional depth chemical
profiling. Spreading Resistance Profiling (SRP) measurements give depth profiles of charge carrier concentration, showing the active profiles after anneal.

However, using SRP to determine the active concentration profile might lead to erroneous results for thin layers, as shown in Fig. 8. SRP results deviate from SIMS results in the tail region and are much lower than expectation. As well known, two of the main limitations of SRP are its probe contact size and probe separation, which are both in the micrometer range. For ultra-shallow layers (sub-100 nm), this leads to the need for very large correction factors to extract the underlying resistivity information from the measured raw data [16]. The thickness of the external poly base is around 60 nm. Therefore small errors on the raw data can lead to significant distortions on the final carrier profile.

Thus, Scanning Spreading Resistance Microscopy (SSRM), which uses much smaller contact size (5-10 nm) with much smaller correction factors [16-18] and hence a more accurate measurement of local active dopant concentration, was applied to characterize the boron active concentration in the thin external poly base layer during optimizing the implant conditions.

The sample measured is 63 nm poly silicon layer on top of 500 nm oxide on p-type Si-substrate. 200 nm oxide capping layer was deposited on the poly layer to avoid bevel rounding. Cleave the sample before measurement. Fig. 9 shows the 2D resistance map from SSRM measurement. The color in the map indicates the level of the resistance with reference of color bar at the right side of the picture. Fig. 9 shows that the depth profile of the active concentration in the poly base is almost homogeneous from top to bottom, despite the large differences between the grains and the grain boundaries. Fig. 9 indicates that the low concentration at the tail of the SRP measurement in Fig. 8 is an artifact, the reason of which could be that the poly base layer is too thin.

Fig. 1. SEM cross-section of the QSA SiGe:C HBT in this work

Fig. 2. $f_{max}$ versus total base resistance $R_b$ ($0.13 \times 2 \, \mu m^2$).
Fig. 3. Total base resistance $R_b$ versus $R_{sh,ext}$ ($0.13 \times 2 \, \mu m^2$).

Fig. 4. Poly base sheet resistance $R_{sh,ext}$ for single, double and triple external base implants.
**Fig. 5.** $R_{sh,ext}$ versus the poly base thickness.

**Fig. 6.** Growth rates as function of hydrogen partial pressure for mono-Si and poly-Si at different values of the silane partial pressure. The symbols represent the experimental data and the full lines and dashed lines serve as eye-guides for mono-Si and poly-Si data, respectively.
Fig. 7. $f_T$ and $f_{max}$ curve as function of $I_c$ (a) and $R_b$ versus $I_b$ (b) on devices before and after the optimization of external base.

Fig. 8. SIMS and SRP measurements of the poly base.
3. Conclusions

Optimization of the external poly base implant and growth conditions has been investigated in QSA HBTs using non-selective epitaxy growth for the base layer deposition. Optimized double implantation in the external poly base at energy $10\text{ keV}/15\text{ keV}$ allows us to induce higher active dopant concentration. A modified epitaxy process with a growth rate ratio of 1.7 between polycrystalline silicon and monocrystalline silicon has been developed in order to form a thicker external poly base layer, hence lower external base sheet resistance, while keeping the intrinsic base unchanged. This modified epitaxy process allows a further reduction of the external poly base sheet resistance. After optimization, the final $f_{\text{max}}$ reaches 300 GHz.

Acknowledgements

The authors wish to thank the IMEC pilot line for processing the wafers and IMEC materials and component analysis department for SIMS, SRP and SSRM measurements.

References

