Collector Transport in SiGe HBTs Operating at Cryogenic Temperatures

Hanbin Ying, *Student Member, IEEE*, Jason Dark, Anup P. Omprakash, *Student Member, IEEE*, Brian R. Wier, *Student Member, IEEE*, Luwei Ge, Uppili Raghunathan, *Student Member, IEEE*, Nelson E. Lourenco, *Member, IEEE*, Zachary E. Fleetwood, *Student Member, IEEE*, Martin Mourigal, Dragomir Davidovic, and John D. Cressler, *Fellow, IEEE*

***Abstract*— This paper provides insight into the transport mechanisms of the collector current in SiGe HBTs operating at cryogenic temperatures and compares three technology generations of devices. Based on the experimental data, a method to differentiate direct tunneling from quasi-ballistic transport is proposed. Measurements indicate that direct tunneling becomes more significant at cryogenic temperatures. The effects of technology scaling on the direct tunneling were investigated using TCAD. Direct tunneling was found to be sensitive to the base width and the Ge profile. It is predicted that without an increase in the Ge content, direct tunneling may dominate over quasi-ballistic transport at the limits of technology scaling.**

***Index Terms*— Cryogenic, HBT, scaling, SiGe, transport.**

I. Introduction

SILICON–germanium heterojunction bipolar transistors (SiGe HBTs) have recently gained attention due to their potential use in quantum computing readout circuits [1]. Quantum computing readout, together with other physics experiments such as charge sensing, detecting magnetic switching, or electron counting, creates a need for amplifiers that directly interface with quantum mechanical devices operating at deep cryogenic temperatures (<4 K and even down to tens of mK) [2]–[4]. In order to connect the interior of the cryogenic refrigerators to room temperature instrumentation, experimental setups historically use lengthy cables that present significant loading on the already tiny output signals [5]. It was found that by adding a cryogenic preamplifier at mK temperatures, the system signal-to-noise ratio can be increased by as much as 100 times [1]. Overall, these interface applications demand low noise, low-to-medium speed (up to several hundred MHz), and most importantly very low power. The cooling power of cryogenic refrigerators used for research purposes is typically on the order of a few watts at 4.2 K and $< 1\ \mu$W at 100 mK (as per vendor quotes from Lake Shore Cryotronics, Advanced Research Systems, Janis Research Company, Quantum Design, and [6]), while that for commercial dilution refrigerators (DRs) can be 0.5–1 mW at 100 mK [7], [8], increasing rapidly as the base temperature rises. In all cases, a significant portion of this power dissipation (the exact percentage is usually unknown) is used to cool mechanical fixtures, such as sample plates, which leaves much less cooling capacity for the sample and interface circuits under study. Since the circuits are placed close to the samples (e.g., qubits or quantum mechanical structures) within the same cooling system, the limitations on cooling demand minimization of heat dissipation from the circuits in order to maintain the cold ambient temperature for the samples. For an optimistic scenario (100 mK using a commercial DR), interface circuits should at most operate in the range of 10–100 μW, if not less. For SiGe HBTs to successfully serve these emerging applications, it is important to understand the cryogenic transport mechanisms of SiGe HBTs at these extremely low-power levels.

Previous studies have shown an evidence for quasi-ballistic (nonequilibrium) transport and trap-assisted tunneling (TAT) in early generations of SiGe HBTs [9], [10]. Recently, at cryogenic temperatures as low as 70 mK, SiGe HBTs exhibited nonclassical characteristics in their measured currents, where a portion of the collector current was explained using a direct tunneling mechanism from emitter to collector [11], [12]. The same mechanism was discovered independently in a different SiGe technology, suggesting that the mechanism is prevalent in advanced SiGe HBTs [13]. In light of this newly discovered mechanism, this paper proposes a simple method to differentiate direct tunneling mechanisms from quasi-ballistic transport. This provides new insight into how the technology scaling of SiGe HBTs affects the collector

Manuscript received March 9, 2018; revised May 23, 2018 and June 20, 2018; accepted July 4, 2018. Date of publication July 18, 2018; date of current version August 21, 2018. The work of H. Ying and J. D. Cressler was supported by the Laboratory Directed Research and Development (LDRD) Program at Sandia National Laboratories (operated by NTES of Sandia, a wholly owned subsidiary of Honeywell International Inc., for DOE's NNSA) under Contract DE-NA-0003525. The work of J. Dark and D. Davidovic was supported by DOE under Contract DE-FG02-06ER46281. The review of this paper was arranged by Editor K. Kalna. *(Corresponding author: Hanbin Ying.)*

H. Ying, A. P. Omprakash, B. R. Wier, U. Raghunathan, Z. E. Fleetwood, and J. D. Cressler are with the School of Electrical and Computer Engineering, Georgia Tech, Atlanta, GA 30332-0250 USA (e-mail: ying.hanbin@gatech.edu; cressler@ece.gatech.edu).

J. Dark, L. Ge, M. Mourigal, and D. Davidovic are with the School of Physics, Georgia Tech, Atlanta, GA 30332 USA.

N. E. Lourenco is with the Advanced Concepts Laboratory, Georgia Tech Research Institute, Atlanta, GA 30332-0866 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TED.2018.2854288

current in Gummel characteristics at cryogenic temperatures. The transport mechanisms responsible for the base current, though clearly important, will, for brevity, be presented in a separate paper.

This paper also investigates the effects of generational scaling on the collector current of SiGe HBTs operating at cryogenic temperatures. Conventional scaling of SiGe HBTs leads to large improvements in peak unity-gain cutoff frequency (f_T) and maximum oscillation frequency (f_{MAX}), but such improvements are targeted for high-injection operation near room temperature [14], [15]. For low-power quantum interface applications at cryogenic temperatures, it is crucial to know how temperature scaling affects the low-to-medium injection region due to the presence of nonclassical transport mechanisms. Therefore, three generations of SiGe HBTs were measured and simulated in TCAD at cryogenic temperatures in order to examine the sensitivity of transport mechanisms to process parameters and provide insight for future profile designs.

II. Device Technology and Measurement Setup

The devices investigated in this paper are from GLOBALFOUNDRIES BiCMOS 5 AM (0.5 μm), 8 HP (130 nm), and 9 HP (90 nm) technologies and are referred to as first [16], third [17], [18], and fourth [19] generation devices. All devices presented are high-performance (HP) device variants from each generation; meaning, their collector profiles were optimized for maximum speed. The HP devices for the first, third, and fourth generations have peak f_T/f_{MAX} values of 50/65, 200/270, and 300/350 GHz, respectively, at 300 K. For simplicity, only n-p-n SiGe HBTs will be discussed. C-B-E-B-C layout structures, with an emitter length in the range of 1–10 μm were measured for each generation, and the presented results were verified to be both reproducible and consistent across multiple samples.

The measurement was performed in a Quantum Design Physical Property Measurement System (PPMS) DynaCool system. The system was configured to allow quick sweeps of temperature with a maximum cooling power of 5 mW at 1.5 K. Test samples were die-attached on custom gold packages using indium solder and connected to the packages by gold wirebonds. The connections from the PPMS system were adapted to triaxial cabling to minimize the residual noise and fed into the Agilent 4156C Semiconductor Parameter Analyzer for device characterization.

III. Existing Transport Theories

Three different transport mechanisms have been shown to exist in the collector current of SiGe HBTs operating at cryogenic temperatures: quasi-ballistic transport [9], TAT [10], and direct tunneling [12], [13]. A qualitative illustration of the three transport mechanisms with respect to bias is shown in the conduction band diagram in Fig. 1. When the emitter–base voltage (V_{BE}) is small, no conduction is possible, because carriers do not have sufficient energy to surmount the base potential barrier, and the base width is too large for any tunneling process. As V_{BE} is increased,

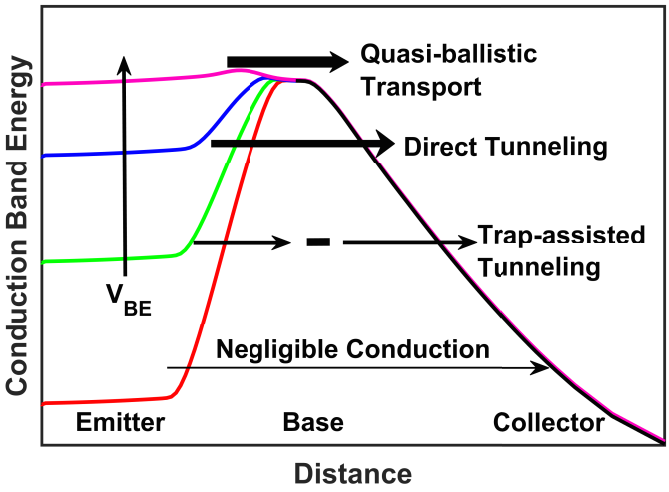

Fig. 1. Qualitative illustration of various transport mechanisms from conduction band of SiGe HBTs under various V_{BE} values.

however, the conduction band energy in the emitter is raised relative to the base. This effectively reduces both the base width and the barrier height for electrons in the emitter. If the base width is small and trap levels exist in the base, electrons can tunnel from the emitter into the trap states in the base and subsequently tunnel into the collector to generate transport current. With increasing V_{BE}, direct tunneling may become possible if the base width is small enough, such that electrons can tunnel directly (instead of via traps) from emitter to collector. At higher V_{BE}, electrons are brought close to the top of the base conduction band, and direct conduction occurs.

Direct conduction occurs via either drift-diffusion (in a thick base) or quasi-ballistic transport (in a thin base). For an extremely thin base, ballistic transport may occur [20]–[22]. Despite formal proof, quasi-ballistic transport was inferred to exist in SiGe HBTs and accounts for the low-temperature nonideal behavior in the collector current [9]. In quasi-ballistic transport, the collector current can be modeled phenomenologically as drift-diffusion but with an electron temperature higher than the ambient temperature. At cryogenic temperatures, the lack of phonons results in less energy exchange between the electron ensemble and the lattice, which yields reduced cooling power for the electrons. This can explain the higher electron temperature. It is generally observed that the drift-diffusion $1/T$ scaling of collector current initially holds until the temperature is "low," at which point the current is better modeled with an effective electron temperature rather than an ambient temperature [23]. For this reason, in this paper, quasi-ballistic transport and drift-diffusion are assumed to be of the same form (both are here referred to as direct conduction), and the deviation of $1/T$ slope with cooling is assumed to come from quasi-ballistic transport.

IV. Direct Tunneling Versus Quasi-Ballistic Transport

The collector current density (J_C) versus V_{BE} of SiGe HBTs from three technology generations across temperature is shown in Fig. 2. At high temperatures, the current is linear on a

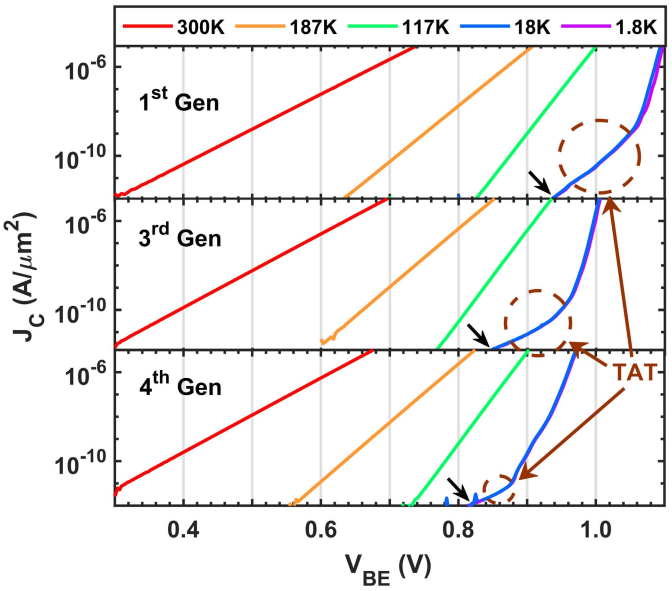

Fig. 2. Measured collector current density versus V_{BE} of all three generations from 300 to 1.8 K. Below 18 K, the curves overlap. The TAT region is circled, with its onset marked with arrows.

log scale (exponential) as drift-diffusion transport dictates. At both 1.8 and 18 K, a nonideal current at low injection is observed for all generations. In the first generation device, the nonideal current below 10 nA/μm^2 was shown to be driven by TAT [10]. Since a similar slope and location of the nonideal current are observed in the third and fourth generations, it is plausible that the same TAT mechanism is present in the more advanced generations. In the present investigation, the region with a smaller slope, circled in Fig. 2, is assumed to be due to a TAT mechanism. Section V will discuss the effects of scaling more in depth. For now, the main question to address is under which conditions the other mechanisms, namely direct tunneling and quasi-ballistic transport, exist, and more importantly, how to differentiate between them in the measurement.

To distinguish between the two, it is recognized that quasi-ballistic transport should be only weakly dependent on the base width. In quasi-ballistic transport, the majority of carriers travel across the base without scattering. Therefore, a small change in base width does not change the already small scattering rate for carriers. In other words, the collector current component from quasi-ballistic transport should be invariant to changes in the base width. On the other hand, direct tunneling current is proportional to the transmission probability P for a rectangular base barrier, according to [24]

$$P \propto e^{-2W\sqrt{2m(U-E)/\hbar^2}} \tag{1}$$

where m is the effective mass of the electron, U is the potential energy, E is the energy of the electron ($E < U$ inside the potential barrier), and W is the barrier width. Although the base barrier is not rectangular, the exponential dependence still applies, and a change in barrier width (base width) is expected to result in an exponential change in the tunneling current [12], [13]. Therefore, the difference between direct

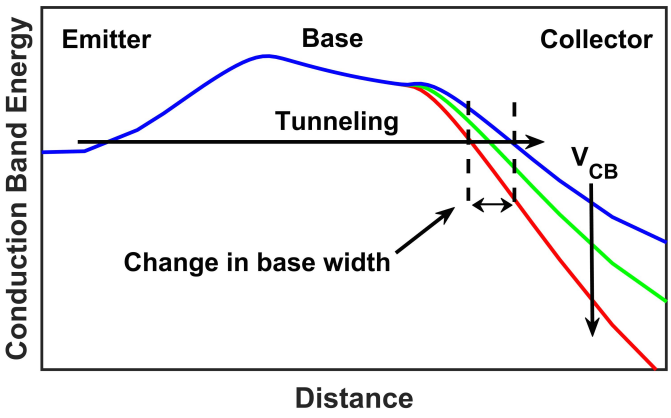

Fig. 3. Qualitative illustration of conduction band diagram of SiGe HBTs showing the variation of base barrier width under multiple V_{CB} values.

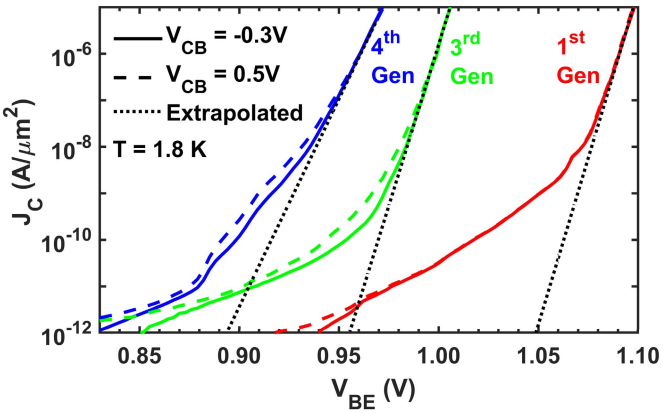

Fig. 4. Collector current from the Gummel characteristics at 1.8 K for $V_{CB} = -0.3$ and 0.5 V. Extrapolations of quasi-ballistic current (dotted lines) estimate the turn-ON voltage for this mechanism.

tunneling versus the quasi-ballistic transport can be revealed by measuring the collector current for different base widths.

To vary the base width without changing the other characteristics of devices, collector–base voltage V_{CB} is used as a tuning parameter. As shown in Fig. 3, increasing V_{CB} shifts the collector conduction band energy (E_C) down, effectively reducing the tunneling barrier width. Similarly, decreasing V_{CB} increases the barrier width. Experimentally, the Gummel characteristics were measured at $V_{CB} = -0.3$ V and $V_{CB} = 0.5$ V. As shown in Fig. 4, apart from the increase of current at low injection due to band-to-band tunneling in the collector–base junction, the collector current of the first generation device is invariant to changes in V_{CB}. However, a portion of the collector current in the third and fourth generation devices changes with V_{CB}.

To quantify this change in current, the normalized collector current at multiple V_{CB} values is plotted versus collector current density (J_C) in Fig. 5. Clearly, in the third and fourth generation devices, collector current density from 10^{-11} to 10^{-7} A/μm^2 is very sensitive to V_{CB}, and this sensitivity disappears toward higher J_C. In the first generation, however, the sensitivity is much smaller throughout the J_C range. The strong sensitivity can be explained by the strong dependence

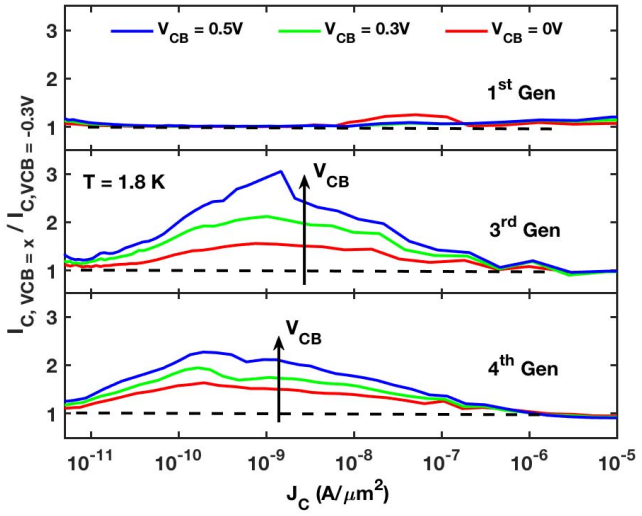

Fig. 5. Ratio of collector current at various V_{CB} values to the current at $V_{CB} = -0.3$ V at 1.8 K in each generation.

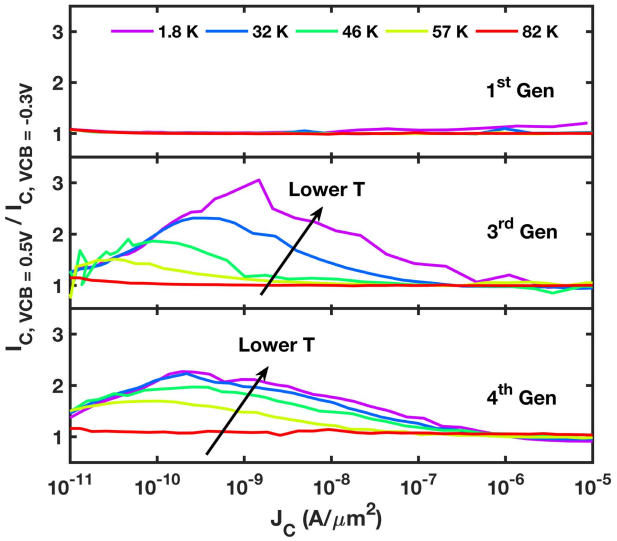

Fig. 6. Ratio of collector current under $V_{CB} = 0.5$ V and $V_{CB} = -0.3$ V from 82 to 1.8 K in each generation.

of direct tunneling current on the base width (i.e., V_{CB}), while the weak sensitivity indicates quasi-ballistic transport dominates, which is independent of the base width. In other words, quasi-ballistic transport is present in all generations, but direct tunneling is present only in the third and fourth generation devices. This makes intuitive sense.

There are two interesting observations to be made. First, it may seem surprising that the ratio of current is larger in the third than the fourth generation device, although the third generation device should have less tunneling due to its larger base width. This can be understood since the ratio of the current is proportional to the ratio of the tunneling probability P. This ratio, from (1), is proportional to $e^{\Delta W}$, where ΔW is the change in base width. In other words, the ratio is proportional to the absolute change of the base width, not its percentage change. Compared with the fourth generation, the collector and base doping are lower in the third generation, causing a larger change in depletion width for the same change in V_{CB}. This causes the effective base width to change more significantly, resulting in a larger change in tunneling current for the third generation, as can be seen in Fig. 5. The tunneling current, however, is still larger in the fourth generation if we compare them at a fixed V_{BE} (barrier height) value because of its smaller base width. The second observation is that direct tunneling can potentially degrade the device transconductance g_m, as shown by the smaller slope of current in Fig. 4. Therefore, the presence of direct tunneling requires more careful considerations for cryogenic circuit designs.

To gain more insight into how direct tunneling becomes significant as temperature decreases, the ratio of collector current between $V_{CB} = 0.5$ and -0.3 V is plotted across temperature in Fig. 6. At 82 K, the ratio is mostly constant for all generations, indicating that the direct tunneling is negligible. As the temperature is lowered, a "hump" is progressively observable in the third and fourth generation devices, but not in the first generation device. Since at a fixed V_{BE} value, drift-diffusion current decreases with temperature while the tunneling current remains roughly constant with temperature, tunneling becomes the dominant mechanism at low temperatures, thereby increasing the ratio. In particular, the hump due to direct tunneling first shows up in low injection and then slowly encroaches toward high injection as the temperature is lowered. When the drift-diffusion current is replaced by quasi-ballistic current, the percentage of tunneling also stops changing, as shown in the fourth generation device. For the third and fourth generation devices, the hump at 1.8 K corresponds to a three times and two times increase, respectively, which is very significant. In comparison, the ratio is constant in the first generation, as the base width is too large for any direct tunneling. In short, direct tunneling can be observed in scaled technologies over a finite range of collector currents by simply varying V_{CB}.

V. Effects of Scaling on Transport

As SiGe HBT technology scales, there is a consequent decrease in the base width, increase in collector doping, increase in Ge mole fraction, and increase in base doping. Aside from the increase in base doping, which raises the base potential barrier, all the other modifications will lower the base potential barrier. A lower barrier enhances the direct tunneling mechanism, which means that if quasi-ballistic transport is not increased at the same rate, direct tunneling will become significant.

Quasi-ballistic transport is mostly dependent on the base barrier height. This can be understood by recognizing that most electrons do not have enough energy to overcome the base barrier at cryogenic temperatures until the barrier height is close to zero, which is when the emitter quasi-Fermi level is similar in height to the base conduction band edge. The base barrier height is directly dependent on the integrated Ge content, which is indirectly related to the peak Ge content and the width of the Ge profile. As seen from the extrapolated line in Fig. 4, the onset of the quasi-ballistic transport shifts toward

smaller V_{BE} as technology scales, because the peak Ge content increases in the more advanced generations. To decrease the turn-ON voltage of the quasi-ballistic transport, a larger Ge mole fraction is required. However, a larger Ge mole fraction requires a thinner base to maintain Ge film stability [14]. Additionally, technology scaling targets improved high-frequency operation at room temperature and typically does not exceed a 30% peak Ge content. Instead, one must also shrink the base width using decreased thermal cycles to achieve the improved performance [25]. Therefore, the onset of quasi-ballistic transport is likely fixed if room temperature scaling rules are followed. The reduced base width, however, will increase the direct tunneling current relative to quasi-ballistic current. Such increase will be visible until quasi-ballistic current eventually rises above the direct tunneling current. In other words, direct tunneling is expected to dominate the collector current up to higher V_{BE} in more scaled technology generations.

It is difficult to predict how the TAT will change with technology scaling, since TAT, which relies on the presence of traps, depends on the technology sensitive process steps, such as epitaxial growth conditions and tooling. For example, in [13], no TAT region is observed, which indicates that the specifics of the device structure design and processing conditions can play a major role in eliminating the TAT. In particular, a perimeter versus area (P/A) analysis can give more insight into the physical location of the traps in play. However, a P/A analysis is not meaningful in the present case because the standard device sizes in these highly scaled technologies allow for only a small range of P/A ratio (their emitter stripe width is fixed at minimum geometry and cannot be altered). That said, we can still estimate the energy of the trap levels based on the onset of the TAT from the Gummel characteristics. As shown in Fig. 1, quasi-ballistic transport occurs when the emitter quasi-Fermi level is close to the base conduction band edge. Since the onset of TAT occurs earlier than the onset of the quasi-ballistic transport, the trap states must be located below the base conduction band. From Fig. 2, the onset of TAT is 0.87, 0.82, and 0.81 V for the first, third, and fourth generation devices, respectively. From Fig. 4, the extrapolated onset of the quasi-ballistic transport is about 1.05, 0.96, and 0.90 V. Therefore, TAT occurs about 100–200 mV before the onset of the quasi-ballistic transport, indicating that the traps are likely located within 100–200 meV from the base conduction band edge. The good news here is that, as shown in Fig. 4, in the fourth generation devices, the TAT leakage current extends only to about 10 pA/μm^2, much lower than in the first generation devices, where it extends to 10 nA/μm^2. This may be the result of both improved process technology and film tooling (yielding reduced trap density) and, equally importantly, an earlier onset of direct tunneling (as discussed earlier). From the data, the onset of direct tunneling current appears to shift to smaller V_{BE} from the first to fourth generations, thereby masking the TAT. It is expected that, with scaling, such trends will continue and the direct tunneling will become more significant, causing the TAT current to only appear at extremely low injection.

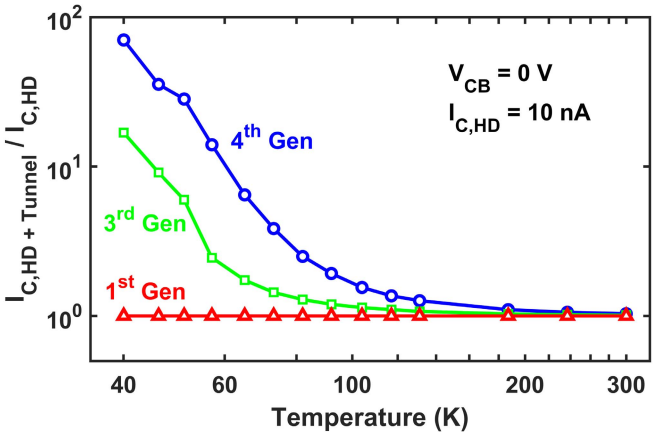

Fig. 7. Ratio of collector current from the HD model with and without the tunneling model. The collector current is extracted from forward Gummel simulation at V_{BE} corresponding to I_C =10 nA in the HD model.

VI. TCAD Simulations

We created three 2-D models in Sentaurus TCAD corresponding to three SiGe HBT generations and used them to investigate how scaling impacts the direct tunneling current. The models were fully calibrated to the room temperature dc and ac measurement data, with Slotboom bandgap narrowing and Phillips unified mobility models [26], [27]. To simulate the tunneling current, the nonlocal tunneling model with Wentzel–Kramers–Brillouin approximations was enabled, which does not alter the room temperature calibration. The nonlocality of tunneling was incorporated as a generation/recombination rate dependent on the local quasi-Fermi level, electric field, and the potential profile along the tunneling paths [28]. The collector current was calibrated to the measurement results from 300 to 40 K using only effective tunneling mass as a free parameter. Quasi-ballistic transport was not included in the TCAD due to the lack of physical models. For simulations below 50 K, recombination in the base was found to not affect the simulated collector current (as expected for direct tunneling).

Two investigations were conducted using TCAD. The first evaluated the amount of tunneling versus conventional drift-diffusion present in the transport current as the device scales. To evaluate this, the device was simulated with either hydrodynamic (HD) and tunneling models engaged or with the HD model alone. The ratio of collector current with and without tunneling at V_{BE}, corresponding to 10 nA of current from the HD model alone, is plotted in Fig. 7. As expected, the ratio of current with or without tunneling is constant throughout the temperature range in the first generation device because the base width is too large for tunneling processes. For the third and fourth generation devices, however, the effects of direct tunneling on collector current begin to appear at as high as 180 K in the fourth generation and 130 K in the third generation. Initially, the tunneling is simply a fractional increase compared with the HD current. At lower temperatures, however, the HD current becomes smaller compared with the direct tunneling current, and the ratio between the two increases dramatically. This can be visually seen in Fig. 8,

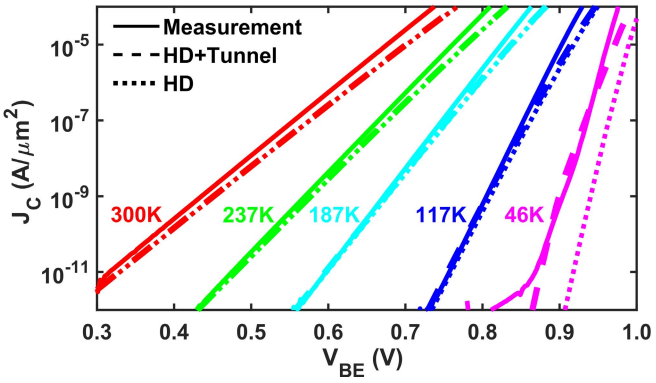

Fig. 8. Collector current of the fourth generation device from Gummel characteristics and TCAD simulations. TCAD simulations are in the HD model with or without the tunneling model.

where the collector current from the simulation with and without the tunneling model, as well as the measured data, is plotted across the temperature for the fourth generation device. At high temperatures, the tunneling does not affect the collector current, and both simulations overlap with the data. At low temperatures, however, the simulated current with only the HD model can be as much as 2–4 orders of magnitude smaller than the measured data. The HD model with direct tunneling, on the other hand, still models the data fairly well until a divergence above 1 μA/μm^2 at low temperatures, as shown in Fig. 8 for 46 K. The inaccuracies at high currents are due to the absence of the quasi-ballistic transport model, which is necessary to explain the continuous exponential increase in the measured current.

The second investigation evaluates the sensitivity of tunneling to the technology process parameters that can modify the base barrier shape. Four process parameters were chosen, namely, peak base doping, peak doping of the selectively implanted collector (SIC), peak Ge, and base width. In addition, a scenario where the base width is reduced while the base doping is increased by the same factor (to keep constant integrated base charge) was also simulated. For comparison purposes, the base width is defined as the distance between the E-B and C-B metallurgical junctions. Though clearly multiple parameters will be scaled simultaneously in the real world (e.g., vertical reduction of base/collector profiles, reduction of emitter cap layer thickness, and modification of Ge profile, and so on), it is difficult to assess the contribution of individual parameters and make a fair comparison. Instead, the sensitivity to the various individual parameters was examined through the use of TCAD simulations. The base and SIC doping profiles are assumed to be Gaussian shaped and defined by the peak and standard deviation. The base width was adjusted by varying the standard deviation of the Gaussian boron profile, which changes the E-B and C-B doping intercepts. All simulations were performed using the calibrated fourth generation TCAD model deck. The doping profile across the cutline in the intrinsic device is published in [12].

Fig. 9 shows the normalized collector current at a fixed V_{BE} value versus the various process parameters, all simulated at 57 K. The process parameters were normalized to the values

TABLE I
SUMMARY OF PARAMETERS IN TCAD SIMULATION

Parameters	Values
Peak Base Doping	$7.8 \times 10^{19}\ cm^{-3}$
Peak SIC Doping	$9.0 \times 10^{18}\ cm^{-3}$
Peak Ge Fraction	0.27
Base Width	16 nm

given in Table I. Observe that scaling Ge affects the current density significantly, as even a slight reduction in Ge increases the base barrier width and height and thereby decreases the tunneling current exponentially. A separate simulation (not shown) reveals that at low temperatures, the Ge content at the BE junction or the Ge grading in the base alone is much less important to the magnitude of tunneling than the total integrated Ge. To understand why, we can look at a typical Ge profile, which ramps up before the EB intercept and ramps down after the CB intercept. Compared with a graded Ge profile, a box Ge profile reduces the base barrier height not only within the neutral base but also in the EB and CB depletion regions, which effectively reduces the barrier width and exponentially increases the tunneling. Therefore, if a large collector current is desired (e.g., for a large current gain β), or if the turn-ON voltage needs to be reduced, a larger Ge mole fraction (close to a box profile shape) throughout the base is preferred. There should be less concern over using a box profile in scaled technologies as the Ge film stability requirement would be more relaxed due to inherently smaller base widths. The base width reduction decreases the tunneling barrier width and allows significantly more collector current to flow. Even if the base doping is increased (higher barrier) while the base width is reduced, the current still increases significantly, because the effect of base width dominates over the effect of base doping. The effect of collector doping is also small, which means that a higher collector doping to suppress Kirk and heterojunction barrier effects for the room temperature operation will not have much impact on cryogenic operation.

VII. SUMMARY

This paper provides insight into the cryogenic collector transport in SiGe HBTs. A unified picture of transport is summarized for three technology generations, where direct tunneling can be distinguished from the quasi-ballistic transport through a simple experimental method. Among all transport mechanisms, TAT could be potentially eliminated using optimized process technology, particular device structures, and improved film tooling with lower epitrap densities. At least, the industry scaling trend so far suggests that the effects of TAT are diminishing in more advanced generations. Direct tunneling and quasi-ballistic transport are expected to become more dominant when the base width is scaled down, and the competition between the two transport mechanisms centers around the integrated Ge profile (width and Ge profile shape). Without an increase in Ge content, direct tunneling may

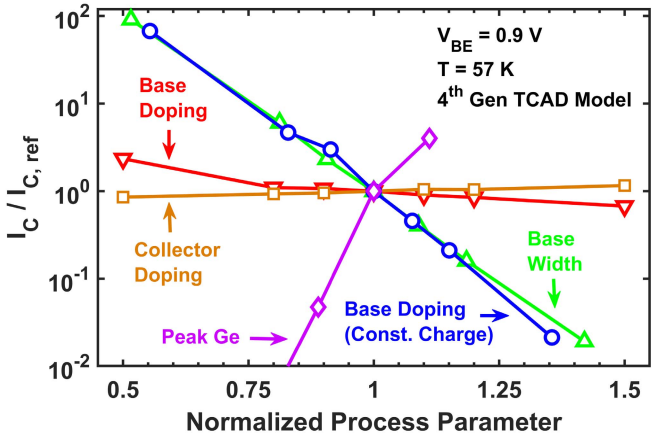

Fig. 9. Simulated collector current versus process parameters that are individually varied in TCAD. Both the current and process parameters are normalized to the initial values.

dominate over quasi-ballistic transport to a higher collector current level. Through a process parameter sensitivity analysis using TCAD, total integrated Ge content and base width are determined to be the most important factors for optimizing cryogenic collector current in SiGe HBTs.

ACKNOWLEDGMENT

The authors would like to thank T. England, M. Holmes, and S. Carr of Sandia and A. Joseph, V. Jain, N. Cahoun, T. Lamouthe, D. Harame, and the GlobalFoundries SiGe Team for their support. They would also like to thank A. Ildefonso for valuable discussions.

REFERENCES

[1] M. J. Curry *et al.*, "Cryogenic preamplification of a single-electron-transistor using a silicon-Germanium heterojunction-bipolar-transistor," *Appl. Phys. Lett.*, vol. 106, no. 20, p. 203505, 2015, doi: 10.1063/1.4921308.

[2] S. Gustavsson, I. Shorubalko, R. Leturcq, S. Schön, and K. Ensslin, "Measuring current by counting electrons in a nanowire quantum dot," *Appl. Phys. Lett.*, vol. 92, no. 15, p. 152101, 2008, doi: 10.1063/1.2892679.

[3] T. Müller *et al.*, "An *in situ* tunable radio-frequency quantum point contact," *Appl. Phys. Lett.*, vol. 97, no. 20, p. 202104, 2010, doi: 10.1063/1.3517483.

[4] L.-J. Wang *et al.*, "A graphene quantum dot with a single electron transistor as an integrated charge sensor," *Appl. Phys. Lett.*, vol. 97, no. 26, p. 262113, 2010, doi: 10.1063/1.3533021.

[5] C. K. Andersen, J. Kerckhoff, K. W. Lehnert, B. J. Chapman, and K. Mølmer, "Closing a quantum feedback loop inside a cryostat: Autonomous state preparation and long-time memory of a superconducting qubit," *Phys. Rev. A*, vol. 93, no. 1, p. 012346, 2016, doi: 10.1103/PhysRevA.93.012346.

[6] SC Group. *4K Cryocoolers*. [Online]. Available: http://www.shicryogenics.com/products/4k-cryocoolers/

[7] A. J. Matthews, M. Patton, T. Marsh, and H. van der Vliet, "A compact cryogen-free platform operating at 1 K or 50 mK," *J. Phys., Conf. Ser.*, vol. 969, no. 1, p. 012091, 2018. [Online]. Available: http://stacks.iop.org/1742-6596/969/i=1/a=012091

[8] *Cryogen Free Dilution Refrigerator System*. [Online]. Available: http://www.iceoxford.com/Cryogenic-systems/Dilution-Refrigerators-dry.htm

[9] D. M. Richey, A. J. Joseph, J. D. Cressler, and R. C. Jaeger, "Evidence for non-equilibrium base transport in Si and SiGe bipolar transistors at cryogenic temperatures," *Solid-State Electron.*, vol. 39, no. 6, pp. 785–789, 1996. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0038110195002235

[10] A. J. Joseph, J. D. Cressler, and D. M. Richey, "Operation of SiGe heterojunction bipolar transistors in the liquid-helium temperature regime," *IEEE Electron Device Lett.*, vol. 16, no. 6, pp. 268–270, Jun. 1995.

[11] H. Ying *et al.*, "Operation of SiGe HBTs down to 70 mK," *IEEE Electron Device Lett.*, vol. 38, no. 1, pp. 12–15, Jan. 2017.

[12] D. Davidović *et al.*, "Tunneling, current gain, and transconductance in silicon-germanium heterojunction bipolar transistors operating at millikelvin temperatures," *Phys. Rev. Appl.*, vol. 8, no. 2, p. 024015, Aug. 2017, doi: 10.1103/PhysRevApplied.8.024015.

[13] H. Rücker, J. Korn, and J. Schmidt, "Operation of sige HBTs at cryogenic temperatures," in *Proc. IEEE Bipolar/BiCMOS Circuits Technol. Meeting (BCTM)*, Oct. 2017, pp. 17–20.

[14] J. D. Cressler and G. Niu, *Silicon-Germanium Heterojunction Bipolar Transistors*. Boston, MA, USA: Artech House, 2002.

[15] P. S. Chakraborty *et al.*, "A 0.8 THz f_{MAX} SiGe HBT operating at 4.3 K," *IEEE Electron Device Lett.*, vol. 35, no. 2, pp. 151–153, Feb. 2014.

[16] D. C. Ahlgren *et al.*, "A SiGe HBT BiCMOS technology for mixed signal RF applications," in *Proc. Bipolar/BiCMOS Circuits Technol. Meeting*, Sep. 1997, pp. 195–197.

[17] B. A. Omer *et al.*, "A 0.13 μm BiCMOS technology featuring a 200/280 GHz ($f_T f_{max}$) SiGe HBT," in *Proc. Bipolar/BiCMOS Circuits Technol. Meeting*, Sep. 2003, pp. 203–206.

[18] B. Jagannathan *et al.*, "Self-aligned SiGe NPN transistors with 285 GHz f_{MAX} and 207 GHz f_T in a manufacturable technology," *IEEE Electron Device Lett.*, vol. 23, no. 5, pp. 258–260, May 2002.

[19] J. J. Pekarik *et al.*, "A 90nm SiGe BiCMOS technology for mm-wave and high-performance analog applications," in *Proc. IEEE Bipolar/BiCMOS Circuits Technol. Meeting (BCTM)*, Sep. 2014, pp. 92–95.

[20] S. Tanaka and M. S. Lundstrom, "A flux-based study of carrier transport in thin-base diodes and transistors," *IEEE Trans. Electron Devices*, vol. 42, no. 10, pp. 1806–1815, Oct. 1995.

[21] M. A. Stettler and M. S. Lundstrom, "A microscopic study of transport in thin base silicon bipolar transistors," *IEEE Trans. Electron Devices*, vol. 41, no. 6, pp. 1027–1033, Jun. 1994.

[22] A. A. Grinberg and S. Luryi, "Diffusion in a short base," *Solid-State Electron.*, vol. 35, no. 9, pp. 1299–1309, 1992. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0038110192901659

[23] J. C. Bardin, "Silicon-germanium heterojunction bipolar transistors for extremely low-noise applications," Ph.D. dissertation, California Inst. Technol., Pasadena, CA, USA, 2009.

[24] E. Merzbacher, *Quantum Mechanics*. Hoboken, NJ, USA: Wiley, 1998.

[25] M. Schroter *et al.*, "Physical and electrical performance limits of high-speed SiGeC HBTs—Part I: Vertical scaling," *IEEE Trans. Electron Devices*, vol. 58, no. 11, pp. 3687–3696, Nov. 2011.

[26] D. B. M. Klaassen, J. W. Slotboom, and H. C. de Graaff, "Unified apparent bandgap narrowing in *n*- and *p*-type silicon," *Solid-State Electron.*, vol. 35, no. 2, pp. 125–129, 1992. [Online]. Available: http://www.sciencedirect.com/science/article/pii/003811019290051D

[27] D. B. M. Klaassen, "A unified mobility model for device simulation—I. Model equations and concentration dependence," *Solid-State Electron.*, vol. 35, no. 7, pp. 953–959, 1992. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0038110192903257

[28] *Sentaurus Device User Guide*, Synopsys Inc., Mountain View, CA, USA, Sep. 2017.

Authors' photographs and biographies not available at the time of publication.