Progress In Electromagnetics Research, Vol. 174, 75–88, 2022

Deep Insight into Channel Engineering of Sub-3 nm-Node P-Type Nanosheet Transistors with a Quantum Transport Model

Afshan Khaliq[1], Shuo Zhang[1, *], Jun Z. Huang[2], Kai Kang[3], and Wen-Yan Yin[1, *]

Abstract—Based on a self-consistent Schrödinger-Poisson solver and top-of-the-barrier model, a quantum transport simulator of p-type gate-all-around nanosheet FET is developed. The effects of material (Si/Ge), stress, crystallographic orientation, and cross-sectional size are deeply explored by numerical simulations for the device performance at the sub-3 nm technology node. A strain-dependent 6-band $k \cdot p$ Hamiltonian is incorporated into the model for a more accurate calculation of E-k dispersion in the strain-perturbed valence band structure, where the curvature, energy shift, and splitting of subbands are investigated in detail for hole transport properties. Further, the effect of channel engineering is comprehensively analyzed, by evaluating density-of-states effective mass, average injection velocity, mobility, current density distributions, and the current-voltage characteristics. An effective performance improvement from 2 GPa compressive stress is obtained in [100]/(001) and [110]/(001) channels, with a 7% enhancement of ON-current in Ge nanosheet FETs. While a wider channel cross-section improves the drive current by increasing the effective channel width, a smaller cross-sectional width yields an average increase up to 29% in the ON-state injection velocity due to stronger quantum confinement.

1. INTRODUCTION

As transistor sizes shrink to the nanoscale, alternative structures and devices are investigated for further extending Moore's Law, among which, nanosheet (NS) gate-all-around (GAA) architecture [1] has attracted immense attention recently because of its flexibility in device sizing and drivability of current compared to FinFETs [2, 3]. Although the GAA structure surrounding the nanosheet channel helps in achieving good electrostatic control, reducing leakage currents, and suppressing short-channel effects [4], a key issue faced by the development of nanosheet transistor is its relatively worse performance in p-type device w.r.t n-type counterpart, due to overall lower hole mobility in the IV and III–V classes of materials [5]. To achieve more balanced performance between n- and p-type devices and improve the entire nanosheet technology, p-type nanosheet field-effect transistors (pNS-FETs) need more in-depth research. However, accurate modeling of hole transport is more complicated than that of its electron counterpart, since the coupling of the first three valence bands in bulk materials, defined as a light hole (LH), heavy hole (HH), and "spin-split off (SO)", introduces a strong non-parabolicity and anisotropy in the band structure. The situation is even more complicated when the device sizes are drastically scaled down to several nanometers, and the band structure is remarkably influenced by the crystal orientation, stress, and quantum confinement [6]. In the quantum transport regime, the modulation in the band structure dispersion creates variations in the effective mass (m_{eff}), LH-HH band splitting, and density-of-states (DOS), which further modify hole transport properties [7]. It has been reported that

Received 12 April 2022, Accepted 25 May 2022, Scheduled 9 June 2022
* Corresponding author: Shuo Zhang (zhangshuo94@zju.edu.cn), Wen-Yan Yin (wyyin@zju.edu.cn).
[1] Innovative Institute of Electromagnetic Information and Electronic Integration (EIEI), College of Information and Electronic Engineering, Zhejiang University, Hangzhou 310027, China. [2] MaxLinear Inc., Carlsbad, CA 92008, USA. [3] School of Electronic Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China.

mobility is closely related to the injection velocity near the top-of-the-barrier [8, 9], and the average ballistic velocity is a critical predictor of drive current and overall device performance in highly-scaled transistors at sub-3 nm node.

Based on the above concerns, the channel engineering of pNS-FETs including that of crystal orientation, stress, and cross-sectional dimensions, needs extensive and profound consideration in a quantum transport scope. Strain/stress engineering takes a key position among all the technological innovations since it is cost-effective and its benefit on the device performance is comparatively large [6, 8, 10]. While tensile uniaxial strain/stress improves the device performance in n-type transistors [11], compressive strain/stress provides more significant hole mobility enhancement in p-type devices, because of its squeezing effect on the band structure yielding a decrease in overall m_{eff} and an increase in group velocity [12, 13]. This enhancement is much less in the biaxially strained case. Besides, as a continuous design parameter in a NS channel, the cross-sectional dimension will affect both the effective width and quantum confinement energy [14]. Therefore, it needs further optimization along with the engineering of channel crystalline orientation configuration. However, recent studies on GAA transistors are mostly limited to n-type devices [8, 15], and some of them use semi-classical approaches for the study [16]. Therefore, in this work, the pNS-FET is studied in a quantum transport model. The $k \cdot p$ Hamiltonian based on the Poisson-Schrödinger self-consistent solver and top-of-barrier ballistic transport model handle quantum mechanical effects rigorously, at the same time guarantee a good balance between computational load and accuracy. The band structure, density-of-states effective mass (m_{DOS}), injection velocity, mobility, and current densities, as well as current-voltage (I-V) characteristics, are evaluated to analyze the impact of compressive stress, cross-sectional size, and channel/wafer crystal orientation, in the Si and Ge pNS-FETs. Such an in-depth device exploration will provide helpful insight into the pNS-FET design and optimization strategies at extremely scaled technology nodes.

2. DEVICE STRUCTURE AND SIMULATION MODEL

2.1. Device Structure

The three-dimensional (3-D) schematic structure of the pNS-FET is illustrated in Fig. 1(a). The simulated region in Fig. 1(b) is confined along the channel width (y-axis) and the channel thickness (z-axis) directions, whereas the hole transport direction is along the channel length (x-axis) direction. According to the International Roadmap for Devices and Systems (IRDS) 2020, the channel length (L_g) is set at 7 nm, and the thickness of the NS-FET (T_{sh}) is fixed at 3 nm for the consideration of the sub-3 nm node. The NS is surrounded by the dielectric layer with an effective oxide thickness $EOT = 1$ nm, through which the gate controls the Si or Ge channel with near-intrinsic carrier concentration. In order to assess crystal orientation effects, [100]/(001), [110]/(001) and [111]/($\bar{1}$10) channel/wafer orientations are considered. As a key parameter in the NS-FET structure, the width (W_{sh}) is varied at 5-, 7-, and 10 nm to evaluate its impacts on the device performance. The cross-sectional area of the rectangular channel is $W_{sh} \times T_{sh}$. For each crystallographic orientation and W_{sh}, the simulation is repeated with uniaxial compressive stress varied from 0 to 2 GPa along the channel, as shown in Fig. 1(c). Following the standard convention, the uniaxial stress is defined as the stress which is non-zero only along one of the principal axes.

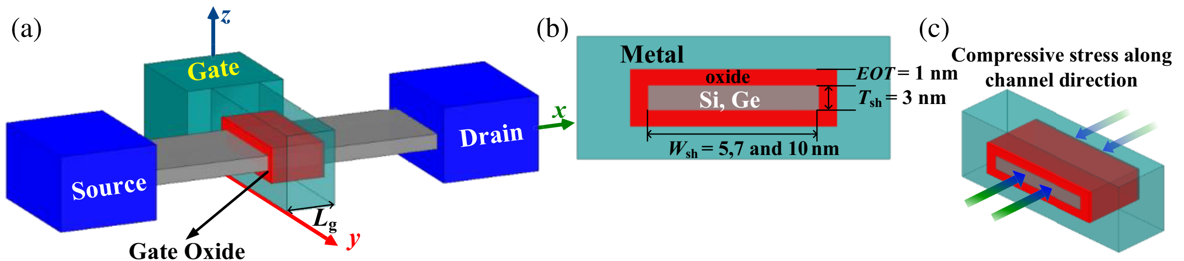

Figure 1. (a) 3-D geometry of the single pNS-FET structure, (b) 2-D cross-section cut at the middle of the Si/Ge channel, and (c) illustration of applied uniaxial compressive stress along the channel direction.

2.2. Simulation Approach

As shown in Fig. 2, transport properties and device performance are evaluated by an in-house developed simulator, based on the Poisson-Schrödinger self-consistent solver [17] and the ballistic top-of-the-barrier transport model. The band structure of NS is rigorously calculated using a strain-dependent effective Hamiltonian, where Bir-Pikus strain Hamiltonian is calculated and incorporated into the 6-band $k \cdot p$ Hamiltonian, thereby the total Hamiltonian is given as

$$H = H_0 + H(\varepsilon) + H_{so} - qV(y, z), \tag{1}$$

where H_0 is the unstrained 6-band $k \cdot p$ Hamiltonian and H_{so} is the spin split-off Hamiltonian. Bir and Pikus modeled strain as a perturbation [18] and constructed a strain matrix $H(\varepsilon)$ as

$$H(\varepsilon) = \begin{bmatrix} l\varepsilon_{xx} + m(\varepsilon_{yy} + \varepsilon_{zz}) & n\varepsilon_{xy} & n\varepsilon_{xz} \\ n\varepsilon_{xy} & l\varepsilon_{yy} + m(\varepsilon_{yy} + \varepsilon_{zz}) & n\varepsilon_{yz} \\ n\varepsilon_{xz} & n\varepsilon_{yz} & l\varepsilon_{zz} + m(\varepsilon_{yy} + \varepsilon_{zz}) \end{bmatrix}, \tag{2}$$

where l, m, and n signify the deformation potential constants for the valance band and (ε_{xx}, ε_{yy}, ε_{zz}, $2\varepsilon_{yz}$, $2\varepsilon_{xz}$, $2\varepsilon_{xy}$) is the strain vector in Voigt's notation [19].

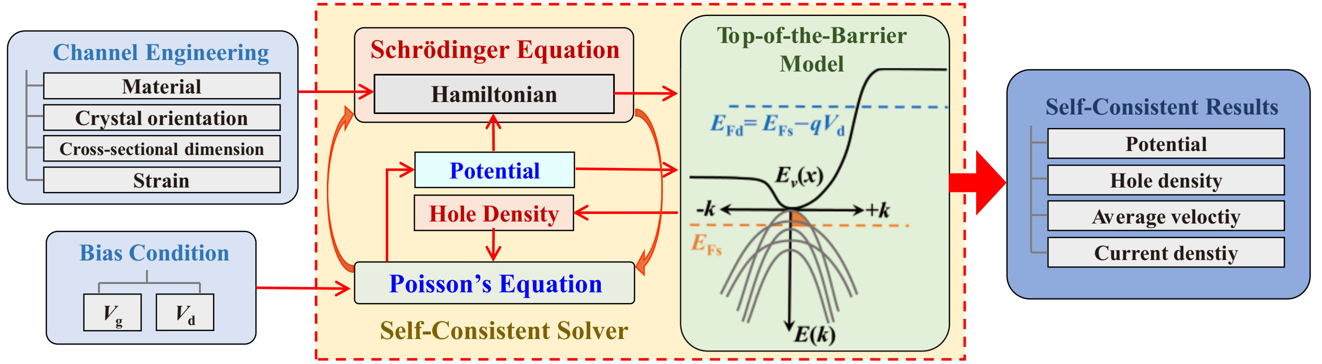

Figure 2. Flowchart of the simulation framework.

Schrödinger equation is solved by the above Hamiltonian, coupling with Poisson's equation to get a self-consistent solution. Hole density is fed into Poisson's equation [20] based on the ballistic injection model, to solve for the electrostatic potential $V(y, z)$ in the NS cross-section, which is, in turn, fed back into Schrödinger equation as potential energy, through Hamiltonian on-site diagonal elements. This keeps iterating until self-consistency is reached. Based on the self-consistent solution, the E-k dispersion of valence band structure is further calculated, and m_{DOS}, average velocity, hole concentration as well as current density are finally obtained from the top-of-the-barrier model [21]. To study the drain current versus gate voltage (I_d-V_g) characteristics of the device, OFF-state current (I_{OFF}) and supply voltage (V_{DD}) is set at 0.01 μA/μm and 0.7 V, respectively, according to IRDS 2020 for high-performance applications at sub-3 nm node [22]. I_d-V_g curves are shifted by aligning I_{OFF}, and this can be done by work function adjustment of the metal gate. The ON-state current (I_{ON}) is extracted at bias condition $V_g = V_d = V_{DD}$, where V_d is the drain voltage.

3. NUMERICAL RESULTS AND DISCUSSION

3.1. Valance Band Structure

The band structure of the pNS-FET is calculated after the convergence of Schrödinger-Poisson iterations at different V_g. As holes in pNS-FETs are free to move only in the x-direction and are confined in both y- and z-directions, the NS has a 1-D E-k dispersion that can be estimated by quantizing both y- and z-components of k. The valance band structure of strained Si and Ge channels with different cross-sections is calculated and compared for three different channel/wafer orientations and stress strengths.

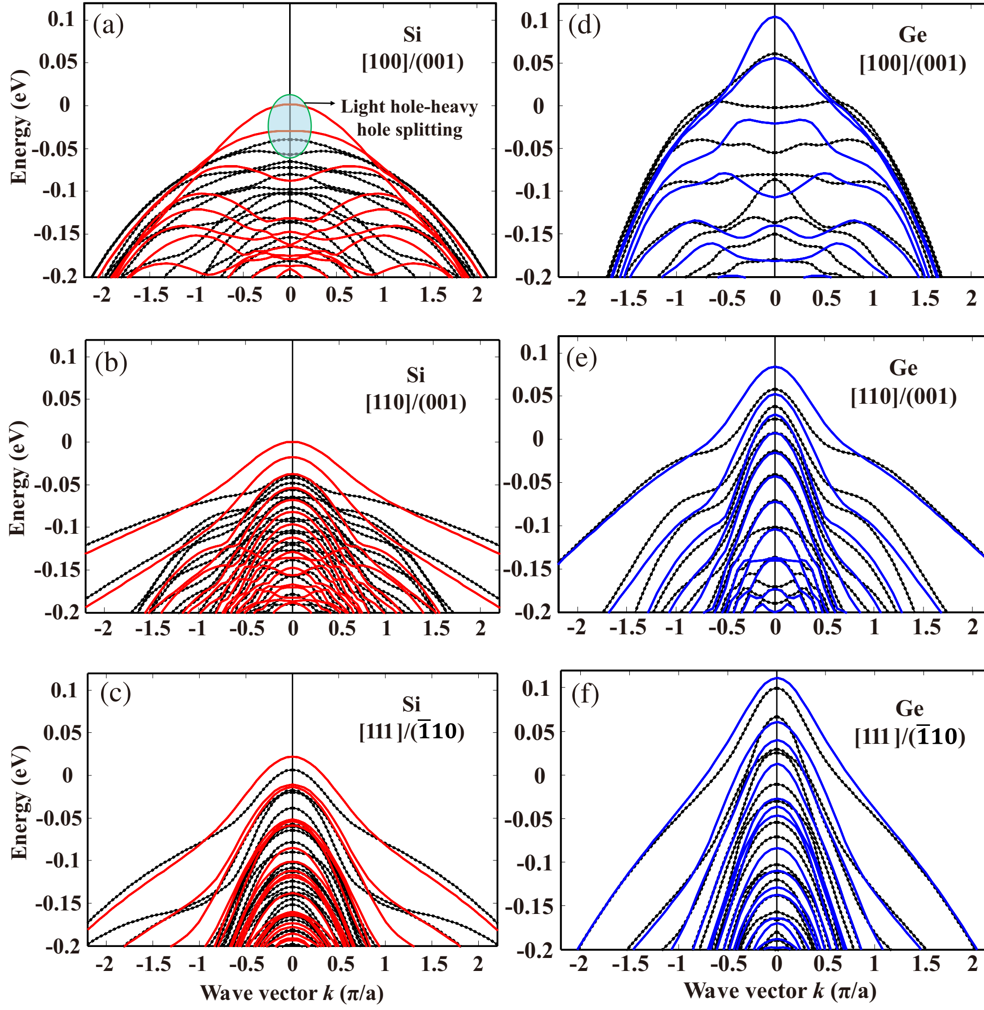

Figure 3. E-k dispersion of the valance band for $10\,\text{nm} \times 3\,\text{nm}$ cross-section of the pNS-FET at $V_g = 0.7\,\text{V}$, for (a)–(c) Si (red), and (d)–(f) Ge (blue), with [100]/(001), [110]/(001), and [111]/($\bar{1}$10) configurations. The unstrained case is shown by dotted black lines, and the 2 GPa compressive stress case is displayed by solid lines.

Figs. 3(a)–(c) show the E-k dispersion of the valance band for the $10\,\text{nm} \times 3\,\text{nm}$ cross-section with [100]/(001), [110]/(001) and [111]/($\bar{1}$10) channel/wafer orientations for stress (2 GPa) and unstrained Si, while Figs. 3(d)–(f) show the corresponding results for Ge as a comparison. The non-parabolic and anisotropic behaviors of different subbands are all captured by the model and shown in the E-k dispersion. Stress causes the mixing of the valence bands such that the calculated hole subbands exhibit a combination of HH, LH, and SO characteristics [21]. The compressive stress induces an upward shift and splitting into the subbands with the gradient of the E-k map changed. For the ultra-scaled cross-sectional dimensions in NS, the degeneracy between LH and HH bands is lifted by the strong quantum confinement, which is called LH-HH splitting [23], as shown in Fig. 3(a). Stress strongly affects all the factors that control the transport properties of holes, including the m_{eff}, confinement energy, and LH-HH splitting. In this section, the effect of stress on these parameters is examined. In all the considered cases, the energy of the first subband is raised by the applied stress, shown in Fig. 3. At the same time, the compressive stress compresses almost all the band structures towards the $\mathbf{\Gamma}$ point and thus increases the curvatures, especially for [100]/(001) and [110]/(001) orientation configurations. This, in general causes the m_{eff} of those subbands to be lighter and the group velocities higher. Specifically, the LH-HH splitting, obtained from the energy spacing of the first two subbands, is also increased. All these effects

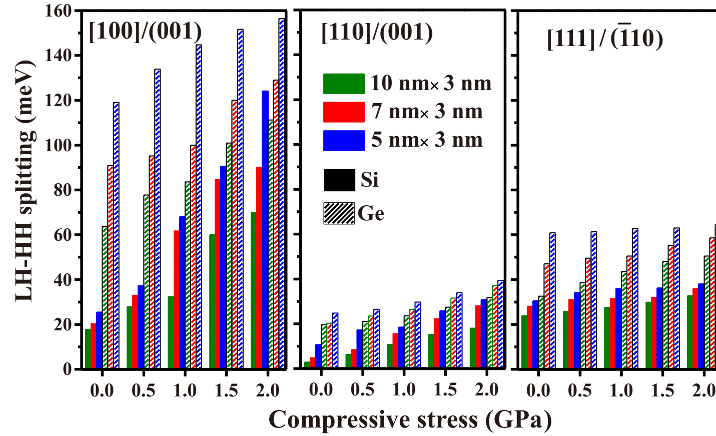

Figure 4. The LH-HH splitting versus compressive stress (varying from 0 to 2 GPa), for three different channel/wafer orientations [100]/(001), [110]/(001) and [111]/($\bar{1}$10) with cross-sectional areas 10 nm × 3 nm (green), 7 nm × 3 nm (red), and 5 nm × 3 nm (blue). The solid-filled bars are for Si, and pattern-filled ones are for Ge.

are caused by the increase in confinement energy, due to the application of compressive stress. The extracted values of LH-HH splitting as a function of the compressive stress for different channel cross-sections and channel/wafer orientations are summarized in Fig. 4. Ge exhibits a larger splitting over Si, and it can be also seen that stress, crystal orientation, and cross-sectional area, all have a significant impact on the LH-HH splitting. In all cases, the LH-HH splitting is enlarged by reducing the cross-sectional area because the shrinking cross-section increases the confinement energy. A noticeable increase in LH-HH splitting due to stress can also be seen from the band structures in Fig. 3. The difference of this splitting induced by stress is huge when the orientation configuration is varied. [100]/(001) shows the strongest splitting together with its increase resulting from stress, in comparison with the other two orientation configurations, while [110]/(001) yields the smallest one. For example, with [100]/(001) orientations, a strong LH-HH splitting up to 156 meV can be achieved by applying a 2 GPa compressive stress in the Ge NS with its cross-sectional area reduced to 5 nm × 3 nm. On the contrary, for [110]/(001) orientation, a splitting of only 40 meV/32.7 meV for Ge/Si, respectively, is obtained even with the same stress strength and cross-section. The LH-HH splitting less than thermal energy ($k_BT = 26$ meV at room temperature, where k_B is the Boltzmann constant and T is the temperature) is considered as a weak splitting, and its impact on the transport properties in the NS is trivial. The LH-HH splitting, combined with the variations of LH and HH band curvatures, significantly determines the overall hole m_{eff} and thus the transport property of the channel. When the LH subband is shifted to the higher energy w.r.t the HH band, the increase of this splitting is beneficial to the transport performance by reducing the overall m_{eff}. However, its effect becomes negative if the HH band is in a higher energy window relative to the LH band. Therefore, the band structure of [100]/(001) configuration earns the most considerable benefits from the effects of stress, including the increases of both curvature (lower m_{eff}) and LH-HH splitting. Further, the LH-HH splitting can substantially influence other performance parameters, such as the velocity of the carriers (discussed in the next section).

3.2. Ballistic Velocity Calculation along the Transport Direction

The injection velocity, defined as the ballistic carrier velocity at the top-of-the-barrier near the source, is one of the main driving forces for improving the device performance [7]. Therefore, the hole injection velocity in a NS channel is evaluated to identify the effects of stress, crystal orientation, and cross-sectional area. The average ballistic velocity ($v_{avg\text{-}x}$) is calculated for the pNS-FET using the ballistic top-of-the-barrier model [21], and it is contributed by the carriers from both source and drain, the densities of which are in equilibrium with Fermi-Dirac distribution. Under ballistic conditions, positive k_x states are filled by the holes with positive velocities statistically determined by the source Fermi level

E_{Fs}, whereas negative k_x states are occupied by the holes with negative velocities, injected from the drain (referring to the drain Fermi level E_{Fd}). For a high drain voltage ($V_d \gg k_BT/q$, where q is the elementary charge), E_{Fd} becomes much greater than E_{Fs}, and thus the holes in the drain reservoir are faced with a very high potential barrier, leading to a suppressed thermionic emission from the drain to the top of the barrier, where the Fermi-Dirac probability determined by E_{Fd} is negligible. In such a scenario, the transport in the channel is dominated by the holes with positive velocities from the source injection [24]. The ballistic hole velocity at each k_x-state can be calculated from the E-k dispersion yielded by the $k \cdot p$ Hamiltonian in (1), and it is expressed as

$$v_x(k_x) = -\frac{1}{\hbar}\frac{\mathrm{d}E(k_x)}{\mathrm{d}k_x}, \tag{3}$$

where $E(k_x)$ is the energy at state k_x for different subbands. To compute $v_{avg\text{-}x}$ of the pNS-FET along channel direction, all the k_x-states with positive $v_x(k_x)$ are taken into calculation. The hole occupancy in the valence band is expressed as $[1-f(E(k_x), E_{Fs})]$ for a specific k_x-state. Finally, $v_{avg\text{-}x}$ is calculated as a weighted average of $v_x(k_x)$, given by [25].

$$v_{avg\text{-}x} = \frac{\sum_{v_x(k_x)>0}[v_x(k_x)(1-f(E(k_x), E_{Fs}))]}{\sum_{v_x(k_x)>0}[1-f(E(k_x), E_{Fs})]}, \tag{4}$$

where the weight is the Fermi-Dirac probability, given as

$$[1-f(E(k_x), E_{Fs})] = \left[\exp\left(\frac{E_{Fs}-E(k_x)}{k_BT}\right)+1\right]^{-1}. \tag{5}$$

Self-consistent simulations are carried out, and $v_{avg\text{-}x}$ is calculated at ON-state with the obtained self-consistent potential. Fig. 5 shows the ON-state $v_{avg\text{-}x}$ as a function of stress, for [100]/(001), [110]/(001) and [111]/($\bar{1}$10) channel/wafer orientations, where Ge has an overall higher $v_{avg\text{-}x}$ w.r.t Si, due to its lighter m_{eff}. As seen from (4), $v_{avg\text{-}x}$ is mainly determined by two factors, $v_x(k_x)$ and the hole occupancy. The former can be estimated from the slope of the subband, whereas the latter is proportional to the energy level of the state. For both materials and all the three orientation configurations, $v_{avg\text{-}x}$ is improved by decreasing cross-sectional dimensions. This is because the larger confinement energy caused by the reduction in cross-sectional dimensions increases the energy spacing of those subbands, leading to fewer HH subbands populated in the high energy window, and thus

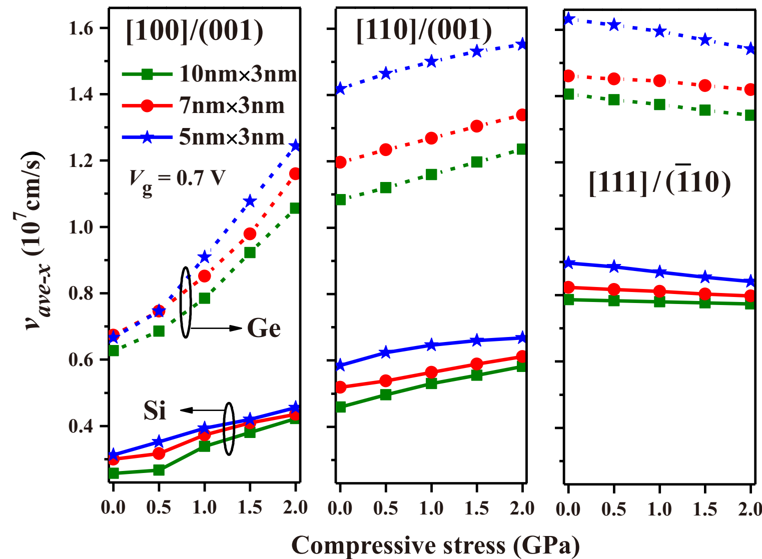

Figure 5. The average hole ballistic velocity ($v_{avg\text{-}x}$) at $V_g = 0.7\,\mathrm{V}$, as a function of the applied compressive stress (from 0 to 2 GPa), for different channel/wafer orientation configurations and cross-sectional sizes in both Si (solid line) and Ge (dashed line) nanosheet channels.

a larger weight of higher v_x in the expression of $v_{avg\text{-}x}$ is obtained. In Si NS, the cross-sectional size reduction from 10 nm × 3 nm to 5 nm × 3 nm yields an average increase in $v_{avg\text{-}x}$ of 12%, 16% and 17% for [100]/(001), [110]/(001) and [111]/($\bar{1}$10) orientation configurations, respectively, whereas in Ge channel, the corresponding percentage increases of $v_{avg\text{-}x}$ are 15%, 29%, and 17% on average. [100]/(001) channel shows the smallest $v_{avg\text{-}x}$ in comparison with the other two, however, it obtains the most significant enhancement of $v_{avg\text{-}x}$ from stress. In contrast, despite the overall highest $v_{avg\text{-}x}$, [111]/($\bar{1}$10) configuration shows degradation of $v_{avg\text{-}x}$ from 0.89×10^7 to 0.84×10^7 cm/s in the Si channel, and from 1.632×10^7 to 1.54×10^7 cm/s in the Ge channel, as the stress strength is increased from 0 to 2 GPa. The difference in the stress effect for different orientations can be directly interpreted from their band structures in Fig. 3. Most of the subbands yield higher $v_x(k_x)$ after the application of stress in [100]/(001) NS, with LH-HH splitting further increasing the contribution of light-m_{eff} subbands, while for [111]/($\bar{1}$10) NS, the slopes and curvatures of subbands in the higher energy range are kind of decreased by stress (bands near the valence band top are stretched by stress), leading to a decreased $v_x(k_x)$ and increased contribution of heavy-m_{eff} bands. Owing to the effect of stress, the highest $v_{avg\text{-}x}$ of 1.552×10^7 cm/s can be obtained by applying a 2 GPa compressive stress in the Ge channel with a 5 nm × 3 nm cross-sectional area and [110]/(001) crystal orientation configuration.

At a technology node below 3 nm, the extremely-scaled device size leads the hole transport to the ballistic limit, where the ballistic hole mobility ($\mu_{\text{ballistic}}$) is purely determined by the channel length and injection velocity, so it can be defined as

$$\mu_{\text{ballistic}} = \frac{q v_{ave\text{-}x} L_g}{2 k_B T}. \tag{6}$$

More information about the mobility calculation can be found in [26]. In a ballistic regime, strain engineering is a vital performance knob for pNS-FETs in terms of mobility enhancement. With 2 GPa compressive uniaxial stress, Ge nanosheet with [110]/(001) crystal orientations and 5 nm × 3 nm cross-sectional area show a peak $\mu_{\text{ballistic}}$ of 208 cm^2/(V·s), while a $\mu_{\text{ballistic}}$ of 90 cm^2/(V·s) is achieved for its Si counterpart. Note that all these values are obtained with $L_g = 7$ nm at sub-3 nm node. The percentage increase/decrease in $\mu_{\text{ballistic}}$ is calculated for both Ge and Si nanosheets with 2 GPa stress relative to the unstrained case and is listed in Table 1 for each channel/wafer orientation configuration and cross-sectional size. It is shown that stress has the most significant enhancement on $\mu_{\text{ballistic}}$ in [100]/(001) channel (up to 86.7% increase w.r.t. unstrained case in Ge nanosheet), while its effect on $\mu_{\text{ballistic}}$ of [111]/($\bar{1}$10) channel is negative, where a maximum decrease of 6.2% in $\mu_{\text{ballistic}}$ is observed in a 5 nm × 3 nm Si nanosheet cross-section. This observation is consistent with the previous analysis of v_{ave-x} behavior. Moreover, for [110]/(001) configuration, stress shows more significant effects in Si nanosheet w.r.t Ge, and in a wider cross-section, whereas, for [100]/(001), more substantial benefit from stress is obtained in Ge nanosheet and in a smaller cross-section.

Table 1. Percentage increase/decrease of ballistic hole mobility with stress.

Configuration / Cross-section	**[100]/(001)**		**[110]/(001)**		**[111]/($\bar{1}$10)**	
	Si	**Ge**	**Si**	**Ge**	**Si**	**Ge**
10 nm × 3 nm	29.7%	68%	20.9%	14%	−1.6%	−4.5%
7 nm × 3 nm	34%	72%	15%	11.9%	−3%	−2.8%
5 nm × 3 nm	44%	86.7%	12.3%	9.4%	−6.2%	−5.5%

3.3. Density-of-States Effective Mass

DOS depends on the E-k dispersion and determines the number of available energy states in a system regulated by Fermi-Dirac distribution. The size quantization effect modifies the E-k relationship leading to the formation of subbands. An analytical expression of 1-dimensional (1D) DOS in valence band

under the m_{eff} approximation for a simple parabolic band, can be used here to qualitatively evaluate and interpret the DOS and hole sheet density of NS from the self-consistent solution. The 1-D DOS in the valence band of a nanosheet is given as

$$\mathrm{DOS} = \frac{\sqrt{2m_{\mathrm{DOS}}}}{\pi\hbar}\sqrt{\frac{1}{(E_v - E)}}, \tag{7}$$

where m_{DOS} is the density-of-states effective mass, E_v is the valence band maximum, and $\hbar$ is the reduced Planck constant. Based on the ballistic top-of-the-barrier model [20], the hole density per unit length (N_S) of a pNS-FET can be expressed as

$$N_s = N_s^{v+}(E_{Fs}) + N_s^{v-}(E_{Fd}), \tag{8}$$

where N_S consists of both $N_s^{v+}(E_{Fs})$, which accounts for the holes injected from the source with positive velocities and a Fermi-Dirac distribution based on E_{Fs}, and $N_s^{v-}(E_{Fd})$ as those injected from the drain with negative velocities. Since $E_{Fd} = E_{Fs} - qV_d$, at a high V_d ($\gg k_BT/q$), $N_s^{v-}(E_{Fd})$ is negligible compared to $N_s^{v+}(E_{Fs})$ and thus,

$$N_s(V_d \gg k_BT/q) = N_s(V_d = 0)/2 = N_s^{v+}(E_{Fs}). \tag{9}$$

Regarding the symmetry of the valence band, N_S at a high V_d can be further expressed as

$$N_s = \frac{1}{2}\int_{-\infty}^{E_v} \mathrm{DOS} \times (1 - f(E, E_{Fs}))dE. \tag{10}$$

Therefore, with the input of DOS and Fermi-Dirac distribution,

$$N_s = \frac{\sqrt{m_{\mathrm{DOS}}/2}}{\pi\hbar}\int_{-\infty}^{E_v} \frac{(E_v - E)^{-1/2}}{1 + e^{\frac{E_{Fs}-E}{k_BT}}}dE \tag{11}$$

By numerical calculation of the integral at $E_{Fs} = 0$, this expression yields the relation between N_S and m_{DOS} for NS with different materials, stress strengths, and cross-sections (resulting in different band structures). Further with the corresponding N_S obtained at each V_g from the self-consistent device simulation, m_{DOS} can be extracted to reflect the overall transport property that different curvatures and energy splitting of band structures provide.

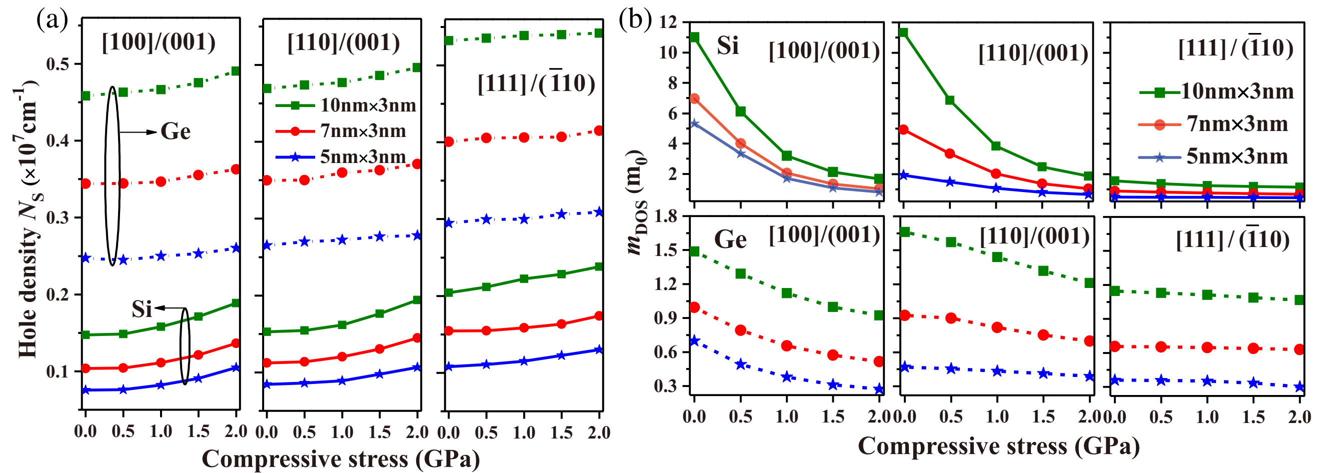

Figure 6. (a) N_S and (b) the corresponding m_{DOS}, versus the strength of compressive stress (from 0 to 2 GPa) for both Si (solid line) and Ge (dashed line) pNS-FETs. Comparison is shown for [100]/(001), [110](001) and [111]/($\bar{1}$10) channel/wafer orientation configurations with different channel cross-sectional sizes at ON-state $V_g = 0.7$ V.

Figure 6(a) shows the N_S calculated from the self-consistent device simulation at ON-state ($V_g = 0.7\,\mathrm{V}$) for different crystal orientation configurations and cross-sectional sizes in Si and Ge pNS-FETs. The increase in N_S induced by the increase of both stress and cross-sectional area can be seen clearly. Ge channel contains a much larger amount of holes at ON-state w.r.t its Si counterpart, and $[111]/(\bar{1}10)$ channel yields a higher N_S in comparison with the other two orientation configurations. Fig. 6(b) compares the extractedm_{DOS} among different cases corresponding to the N_S in Fig. 6(a). Si shows an overall higher m_{DOS} w.r.t Ge, and a larger cross-section increases m_{DOS} by involving more subbands contributing to the transport. Most importantly, it is shown that stress decreases m_{DOS} by increasing the energy separation of subbands in the valence band structure, as already shown in Fig. 3. The effective decrease of m_{DOS}, due to the application of stress, is more prominent in Si w.r.t Ge, because of a more significant upward energy shift of Si band structure also observed in Fig. 3. Among all the orientation configurations, m_{DOS} of $[111]/(\bar{1}10)$ undergoes a minuscule effect of stress, resulting from its relatively insignificant band structure variation.

3.4. Hole Current Density in the Channel Cross-Section

Based on (4) and the cross-sectional potential profile obtained from self-consistent simulation, $v_{avg\text{-}x}$ at each grid point (y, z) in the channel is calculated. The contour of $v_{avg\text{-}x}(y, z)$ across the channel cross-section is shown in Fig. 7, for [100]/(001), [110]/(001) and $[111]/(\bar{1}10)$ Ge nanosheet channel at $V_g = 0.7\,\mathrm{V}$, where the effect of 2 GPa compressive stress is evaluated by comparing with an unstrained case. Among different orientation configurations, $[111]/(\bar{1}10)$ yields an overall higher $v_{avg\text{-}x}$ (up to $1.8 \times 10^7\,\mathrm{cm/s}$) across the whole cross-section, while [100]/(001) shows the lowest $v_{avg\text{-}x}$ (less than $1.44 \times 10^7\,\mathrm{cm/s}$), along with the largest variation ($\sim 0.15 \times 10^7\,\mathrm{cm/s}$) in the central region of the cross-section. For all the cases, the peak values of $v_{avg\text{-}x}(y, z)$ are located near the interfaces between channel and oxide, reaching the maximum at the corners. Notably, a 2 GPa stress increases the gradient of $v_{avg\text{-}x}$ profile in $[111]/(\bar{1}10)$ channel, while decreases that in the other two orientation configurations. In addition, an obvious enhancement of $v_{avg\text{-}x}$ over the whole cross-section is observed in [100]/(001) and [110]/(001) channels. It is worth mentioning that the channel cross-sectional dimension scaling also causes some marginal improvement in $v_{avg\text{-}x}(y, z)$.

The current density profile $j(y, z)$ in the channel cross-section can be calculated by the product of hole density $N(y, z)$ and average carrier velocity $v_{avg\text{-}x}(y, z)$, at each grid point (y, z), and it is given as

$$j\left(y, z\right) = q \times N\left(y, z\right) \times v_{avg\text{-}x}\left(y, z\right). \tag{12}$$

With the equation above, $j(y, z)$ is calculated at each bias point, and the total current is obtained by integrating it over the whole cross-sectional area, for a pNS-FET with different channel materials, orientation configurations, cross-sectional sizes, and stress strengths. Fig. 8 shows the distributions of ON-state $j(y, z)$ in the channel cross-section of a Ge pNS-FET with 2 GPa compressive stress, for both a wide (10 nm × 3 nm) and narrow (5 nm × 3 nm) nanosheet channel. As seen clearly, the pattern of ON-state current density in the channel cross-section is mainly determined by the wave functions of those states in subbands contributing most to the transport. At a V_{DD} of only 0.7 V and in such a strongly confined structure where the subband spacing is large, the main contribution of mobile holes is from the top few subbands. In addition, smaller cross-section results in a higher current density, and a strong differentiation of current distribution is seen among different channel orientation configurations. [100]/(001) accounts for an overall smallest current density, the peaks of which are mainly located near the two side surfaces interfaced with the oxide, whereas $[111]/(\bar{1}10)$ yields the largest current density yet with the most non-uniform and complex distribution. However, [110]/(001) channel shows a comparably large current density with the evenest distribution.

By integrating $j(y, z)$ over the entire channel cross-section, the total current in a sub-3 nm-node pNS-FET is calculated and plotted as a function of V_g in Fig. 9 for different materials, orientation configurations, stress strengths, and cross-sectional areas. It is shown that Ge NS outperforms its Si counterpart due to its lighter m_{eff}, and a 2 GPa compressive stress leads to an improvement of I_{ON} with differentiation in crystal orientation configuration. While the magnitude of I_{ON} is shown to be the largest for $[111]/(\bar{1}10)$ channel, due to its largest N_s and $v_{avg\text{-}x}$, the enhancement of I_{ON} by stress is minuscule, and this results from the stress-induced decrease of $v_{avg\text{-}x}$ as shown in Fig. 5, which offsets the corresponding increase of N_S in Fig. 6(a). On the other hand, [100]/(001) channel obtains

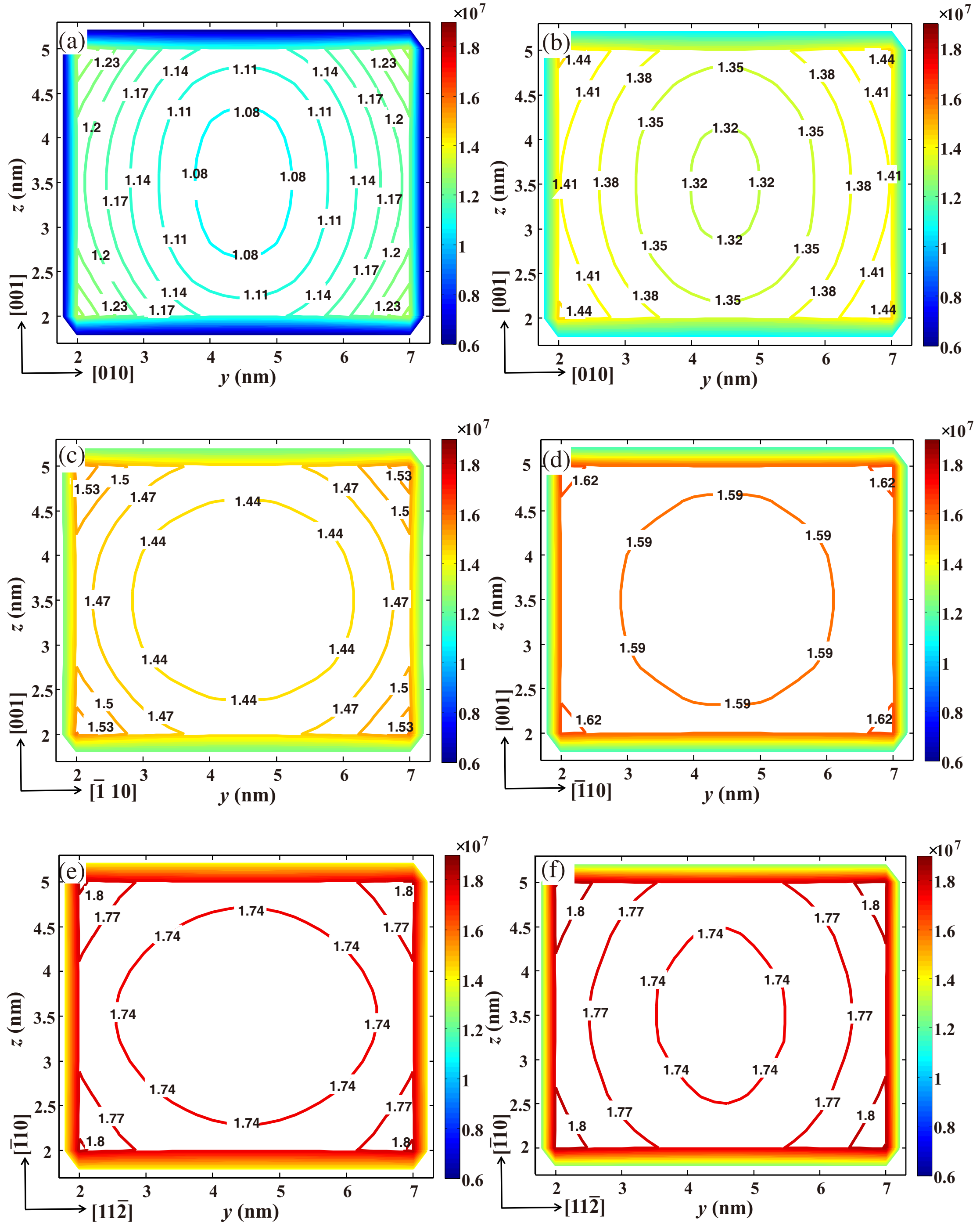

Figure 7. The distribution of average hole ballistic velocity $v_{avg\text{-}x}(yz)$ in a $5\,\mathrm{nm} \times 3\,\mathrm{nm}$ Ge channel cross-section, with channel/wafer orientation configuration (a), (b) [100]/(001), (c), (d) [110]/(001) and (e), (f) [111]/($\bar{1}$10). The left column is the unstrained case, while the right column is with a compressive stress of 2 GPa. The quantization surfaces for each orientation are also displayed.

the largest enhancement of I_{ON} from the application of 2 GPa compressive stress because of substantial improvement in both N_s and $v_{avg\text{-}x}$. Moreover, considering the total current flowing through the channel, Fig. 9(b) shows the benefit of larger effective channel width in drive current for a $10\,\mathrm{nm} \times 3\,\mathrm{nm}$ NS channel, which is yet degraded by its weaker quantum confinement and electrostatic gate control.

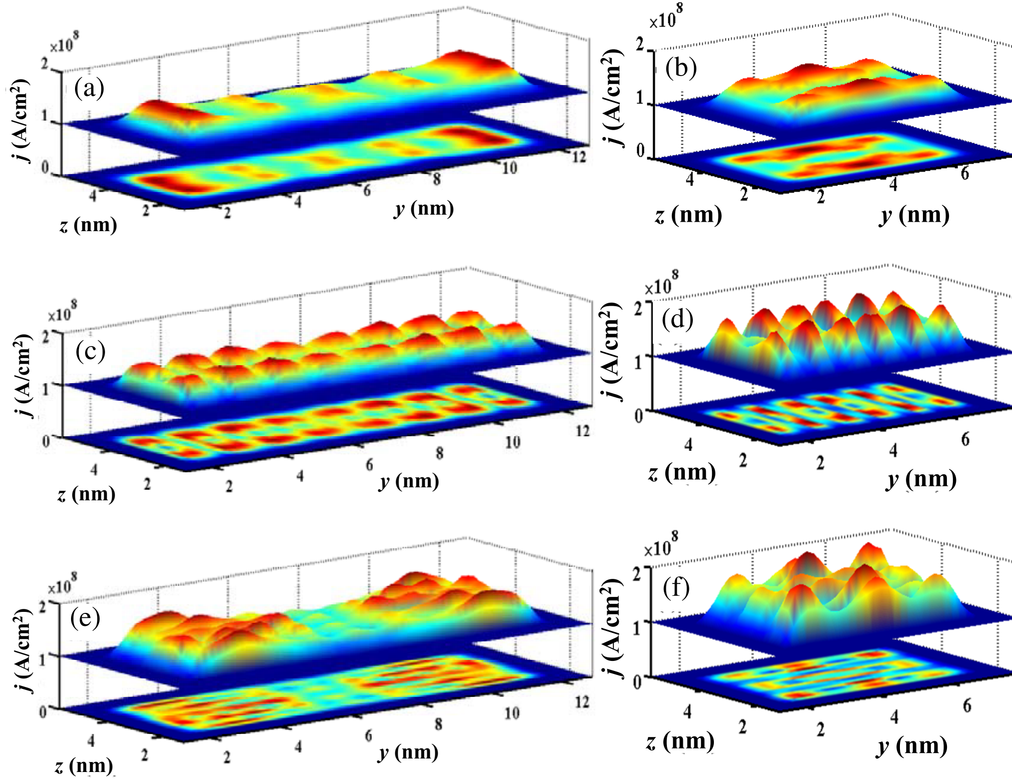

Figure 8. The distribution of ON-state ($V_g = V_d = 0.7\,\mathrm{V}$) current density $j(y, z)$ in the channel cross section of a Ge pNS-FET with (a), (b) [100]/(001), (c), (d) [110]/(001) and (e), (f) [111]/($\bar{1}$10) channel/wafer orientation configuration, respectively, where a compressive stress of 2 GPa is applied. (a), (c) and (e) on the left are for 10 nm × 3 nm cross section, while (b), (d) and (f) on the right are for 5 nm × 3 nm.

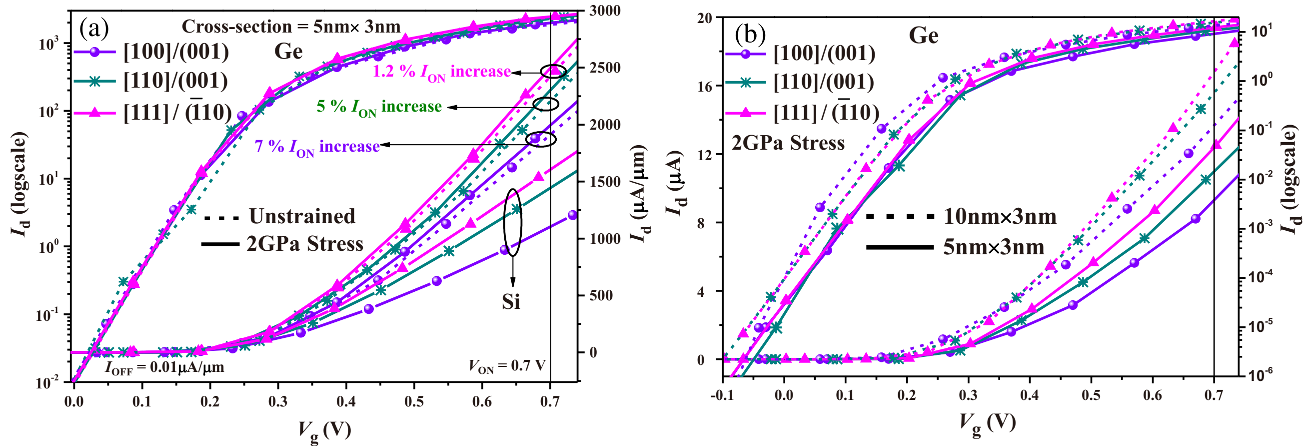

Figure 9. (a) I_d-V_g characteristics for a Ge pNS-FET with 5 nm × 3 nm channel cross-section, considering [100]/(001), [110]/(001) and [111]/($\bar{1}$10) orientation configurations. The comparison between 2 GPa compressive stress (solid line) and unstrained (dashed line) cases is shown with the percentage increase of I_{ON} calculated. The corresponding results of stressed Si pNS-FETs are also plotted for comparison, and I_d is normalized by the channel width. (b) I_d-V_g characteristics for a Ge pNS-FET with 2 GPa compressive stress, compared between 5 nm × 3 nm (solid line) and 10 nm × 3 nm (dashed line) channel cross-sections for all considered orientation configurations. Current flowing through the channel is without normalization by width.

4. DISCUSSION

The nanoscale device simulation using the non-equilibrium Green's function (NEGF) simulator has a lot of computational complexity, where the self-energy and Green's function calculations constitute the most expensive parts of the whole simulation. For a case study, the simulation results of an I_d-V_g curve with 10 bias points for Si NS-FET with $10\,\text{nm} \times 3\,\text{nm}$ channel cross-section size considering (100)/[001] orientation configuration, which takes 83660.1 s (23 hours) by NEGF NS-FET simulator and 43210.8 s (12 hours) by the top-of-the-barrier model using supercomputer (Intel (R)-Xeon (R) CPU E5-2630 @2.40 GHz, two processor RAM = 64 GB). Also, a significant reduction of 93.1% in memory consumption is achieved by using the top-of-the-barrier model. Thus using the top-of-the-barrier model can greatly reduce the computational time and cost-efficiency. However, in the NEGF method, the on-state current consists of both thermionic and tunneling current, while in the ballistic top-of-the-barrier model direct source-to-drain tunneling (SDT) cannot be handled which limits its accuracy. Another fact that cannot be dealt with in the top-of-the-barrier model is the effect of scattering [27] such as the channel/insulator interface roughness called surface roughness, which affects the performance of NS-FETs. The roughness introduces electrostatic potential variations inside the channel, which behave as a scattering potential for carriers. In this study, the ballistic transport is taken into account therefore the scattering effects are not considered and lie outside the scope of this paper and will be published elsewhere.

5. CONCLUSION

Exploiting our in-house developed quantum transport simulator, this work has comprehensively investigated the engineering of stress, crystallographic orientation, and cross-sectional size for both Si and Ge NS channels in a p-type NS-FET at a sub-3 nm technology node. Device simulations were carried out by the self-consistent Schrödinger-Poisson solver based on a stress-dependent 6-band $k \cdot p$ Hamiltonian, combined with a ballistic top-of-the-barrier model. The E-k dispersion of valence band structure was rigorously calculated, where the band curvature, LH-HH splitting, and subband energy shift were analyzed for interpreting the effects of channel engineering on hole transport properties. In a quasi-ballistic regime, m_{DOS}, average injection velocity, and hole density distribution were also evaluated and further projected on the simulation of I-V characteristics of the device. By compressing the band structure, shifting up LH subbands, and increasing the LH-HH splitting, stress is shown to effectively decrease m_{DOS} and enhance hole transport in both [100]/(001) and [110]/(001) channels. In particular, [100]/(001) Ge NS channel obtains the most significant decrease of m_{DOS} and increase of $v_{avg\text{-}x}$ at a compressive stress of 2 GPa, resulting in an enhancement of mobility and ON-current up to 86.7% and 7%, respectively, for a $5\,\text{nm} \times 3\,\text{nm}$ cross-section. Although [111]/($\bar{1}$10) channel shows an overall better performance over the other two orientation configurations, it is insensitive to the tuning of compressive stress. A wider cross-section of NS indeed improves current drivability due to larger effective channel width; however, its lower current density induced by weaker quantum confinement mitigates this gain.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (Grant Nos. 62071418 and 61931007), and the Science Challenge Project (Grant No. TZ2018002).

REFERENCES

1. Loubet, N., T. Hook, P. Montanini, C.-W. Yeung, S. Kanakasabapathy, and M. Guillorn, "Stacked nanosheet gate-all-around transistor to enable scaling beyond FinFET," *Proceedings Symposium VLSI Technology*, Vol. 5, No. 1, 14–15, 2017, doi: 10.23919/VLSIT.2017.7998183.
2. Jang, D., et al., "Device exploration of nanosheet transistors for sub-7-nm technology node," *IEEE Transactions on Electron Devices*, Vol. 64, No. 6, 2707–2713, 2017, doi: 10.1109/TED.2017.2695455.

3. Nagy, D., G. Indalecio, A. J. Garcia-Loureiro, M. A. Elmessary, K. Kalna, and N. Seoane, "FinFET versus gate-all-around nanowire FET: Performance, scaling, and variability," *IEEE Journal of Electron Devices Society*, Vol. 6, No. 1, 332–340, 2018, doi: 10.1109/JEDS.2018.2804383.
4. Zhang, S., et al., "Design considerations for Si- and Ge-stacked nanosheet pMOSFETs based on quantum transport simulations," *IEEE Transactions on Electron Devices*, Vol. 67, No. 1, 26–32, 2020, doi: 10.1109/TED.2019.2954308.
5. Sun, Y., S. E. Thompson, and T. Nishida, "Physics of strain effects in semiconductors and metal-oxide-semiconductor field-effect transistors," *Journal of Applied Physics*, Vol. 101, No. 10, 2007, doi: 10.1063/1.2730561.
6. Yao, J., et al., "Physical insights on quantum confinement and carrier mobility in Si, $Si_{0.45}Ge_{0.55}$, Ge gate-all-around NSFET for 5 nm technology node," *IEEE Journal of Electron Devices Society*, Vol. 6, 841–848, 2018, doi: 10.1109/JEDS.2018.2858225.
7. Teherani, J. T., "A comprehensive theoretical analysis of hole ballistic velocity in Si, SiGe, and Ge: Effect of uniaxial strain, crystallographic orientation, body thickness, and gate architecture," *IEEE Transactions on Electron Devices*, Vol. 64, No. 8, 3316–3323, 2017, doi: 10.1109/TED.2017.2708691.
8. Mohapatra, E., T. P. Dash, J. Jena, S. Das, and C. K. Maiti, "Strain induced variability study in gate-all-around vertically-stacked horizontal nanosheet transistors," *Physica Scripta*, Vol. 95, No. 6, 2020, doi: 10.1088/1402-4896/ab89f5.
9. Khakifirooz, A. and D. A. Antoniadis, "Transistor performance scaling: The role of virtual source velocity and its mobility dependence," *IEEE International Electron Devices Meeting*, 1–4, 2006, doi: 10.1109/IEDM.2006.346873.
10. Dash, T. P., S. Dey, S. Das, E. Mohapatra, J. Jena, and C. K. Maiti, "Strain-engineering in nanowire field-effect transistors at 3 nm technology node," *Physica E Low-Dimensional System and Nanostructures*, Vol. 118, 113964, 2020, doi: 10.1016/j.physe.2020.113964.
11. Maegawa, T., T. Yamauchi, T. Hara, H. Tsuchiya, and M. Ogawa, "Strain effects on electronic bandstructures in nanoscaled silicon: From bulk to nanowire," *IEEE Transactions on Electron Devices*, Vol. 56, No. 4, 553–559, 2009, doi: 10.1109/TED.2009.2014185.
12. Zhang, J.-H., Q.-A. Huang, H. Yu, and S.-Y. Lei, "Orientation effects in ballistic high-strained p-type Si nanowire FETs," *Sensors*, Vol. 9, No. 4, 2746–2759, 2009, doi: 10.3390/s90402746.
13. Thompson, S. E., G. Sun, K. Wu, J. Lim, and T. Nishida, "Key differences for process-induced uniaxial vs. substrate-induced biaxial stressed Si and Ge channel MOSFETs," *IEEE International Electron Devices Meeting*, 221–224, 2005, doi: 101109/IEDM.2004.1419114.
14. Sun, Y., X. Yu, R. Zhang, B. Chen, and R. Cheng, "The past and future of multi-gate field-effect transistors: Process challenges and reliability issues," *Journal of Semiconductor*, Vol. 42, No. 2, 2021, doi: 10.1088/1674-4926/42/2/0231102.
15. Stanojević, Z., V. Sverdlov, O. Baumgartner, and H. Kosina, "Subband engineering in n-type silicon nanowires using strain and confinement," *Solid State Electronics*, Vol. 70, 73–80, 2012, doi: 10.1016/j.sse.2011.11.022.
16. Zhang, L., J. He, J. Zhang, F. Liu, Y. Fu, Y. Song, and X. Zhang, "An analytic model for Ge/Si core/shell nanowire MOSFETs considering drift-diffusion and ballistic transport," *10th Int'l Symposium on Quality Electronic Design*, 2009, doi: 10.1109/ISQED.2009.4810359.
17. Xu, X. and J. Mahanty, "The influence of Si delta doping on the electronic structure of AlGaAs-GaAs-AlGaAs single quantum wells," *Journal of Physics: Condense Matter*, Vol. 6, No. 25, 4745–4762, 1994.
18. Moussavou, M., N. Cavassilas, E. Dib, and M. Bescond, "Influence of uniaxial strain in Si and Ge p-type double-gate metal-oxide-semiconductor field effect transistors," *Journal of Applied Physics*, Vol. 118, No. 11, 2015, doi: 10.1063/1.4930567.
19. Huang, J. Z., L. Zhang, P. Long, M. Povolotskyi, and G. Klimeck, "Quantum transport simulation of III-V TFETs with reduced-order $k \cdot p$ method," *Tunneling Field Effect Transistor Technology*, 151–180, 2016, doi: 10.1007/978-3-319-31653-6_6.
20. Ma, Z.-H., W. C. Chew, and L. J. Jiang, "A novel fast solver for Poisson's equation with Neumann boundary condition," *Progress In Electromagnetic Research*, Vol. 136, 195–209, 2013.

21. Neophytou, N., A. Paul, and G. Klimeck, "Bandstructure effects in silicon nanowire hole transport," *IEEE Transactions Nanotechnology*, Vol. 55, No. 6, 1286–1297, 2008, doi: 10.1109/TNANO.2008.2006272.
22. "IEEE International Roadmap for Devices and Systems — IEEE IRDS," 2020.
23. Wang, J., A. Rahman, G. Klimeck, and M. Lundstrom, "Bandstructure and orientation effects in ballistic Si and Ge nanowire FETs," *IEEE International Electron Devices Meeting*, Vol. 2005, 530–533, 2005, doi:10.1109/IEDM.2005.1609399.
24. Anantram, M. P., M. S. Lundstrom, and D. E. Nikonov, "Modeling of nanoscale devices," *Proceedings of the IEEE*, Vol. 96, No. 9, 1509–1510, 2008, doi: 10.1109/JPROC.2008.927311.
25. Datta, S., *Quantum Transport: Atom to Transistor*, Cambridge University Press, 2005.
26. Lundstrom, M. S. and D. A. Antoniadis, "Compact models and the physics of nanoscale FETs," *IEEE Transactions on Electron Devices*, Vol. 61, No. 2, 225–233, 2014, doi: 10.1109/TED.2013.2283253.
27. Wang, R., Y. Zhang, G. H. Chen, and C. Y. Yam, "Quantum mechanical modeling of electron-photon interactions in nanoscale devices," *Progress In Electromagnetic Research*, Vol. 154, 163–170, 2015.