Integration and electrical evaluation of epitaxially grown Si and SiGe channels for vertical NAND Memory applications

[1]E. Capogreco, R. Degraeve, J. G. Lisoni, [1]K. V. Luong, A. Arreghini, M. Toledano-Luque, A. Hikavyy, [2]T.Numata, [1]K. De Meyer, G. Van den bosch and J. Van Houdt

Imec, Kapeldreef 75 Leuven Belgium, [1]also with ESAT, KULeuven, Leuven Belgium, [2]assignee at imec from Toshiba Corporation.

Abstract — **Epitaxially grown Si and $Si_{0.6}Ge_{0.4}$ are integrated as replacement of poly-Si channel in vertical cylindrical transistors for vertical NAND memory application, in order to investigate the impact of the grain boundaries on current conduction. Epi-Si outperforms both poly-Si and Epi-SiGe channels, resulting in the best conduction, with large improvement on both sub-threshold swing and transconductance (g_m). The experimentally observed g_m bimodal distribution for epi-Si is corroborated and explained through a resistive network model: lower g_m conduction occurs when current needs to cross a high resistance boundary, whereas higher g_m is obtained when this boundary is not present.**

I. Introduction

Vertical NAND memory technology is currently in production as replacement of planar floating gate beyond the 15 nm node [1,2]. In the past few years, several approaches have been developed to vertically stack NAND cells [3], however in the industrially most relevant schemes the channel is always formed after the gate stack through filling a vertical cylindrical memory hole [4]-[6]. Therefore, poly-Si is the typical channel material. Unfortunately, the conduction in poly-Si is a major concern, since the drive current for the read operation (I_D) is low and unstable, and presents a broad distribution [7,8]. Current is hampered by grain boundaries and Si-SiO_2 interface defects leading to a percolating conduction through the defective material.

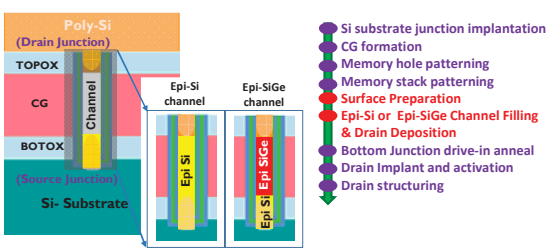

Fig. 1: Schematic view of our vertical NAND single test vehicle. Added/modified steps w.r.t. poly-Si-based process flow are highlighted in red.

A possible solution to boost I_D was already explored in [9], and consists in engineering the poly-Si grain size and in curing the interfacial defects through thermal annealing. Nevertheless, single crystalline channel material should provide the ideal performance, as grain boundaries are absent.

In this study we integrate and electrically evaluate epitaxially grown Si-based materials to replace poly-Si channel in vertical cylindrical transistors for vertical NAND memories. Beside pure epi-Si, also epi-SiGe, with a Ge concentration of 40%, is assessed. Then, current conduction is electrically analysed and results are interpreted trough a resistive network model, to clarify the impact of material defects on I_D.

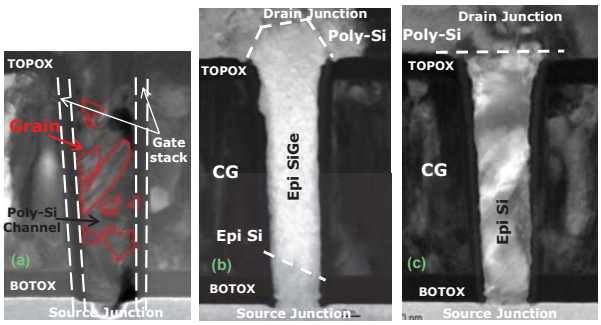

Fig. 2: Dark Field STEM on (a) poly-Si channel, (b) epi-$Si_{0.6}Ge_{0.4}$ channel, (c) epi-Si channel.

II. Experimental Setup

Fig. 1 shows the schematic view of the single-layer test vehicle used in our investigation. The poly-Si-based process flow, already described in [10], is for this study modified to replace the channel with epi-Si and epi-$Si_{0.6}Ge_{0.4}$. Devices are fabricated by gate-first and channel-last approach in 300 mm wafers by patterning cylindrical holes with a depth of 260 nm . ~15 nm thick gate dielectric is deposited at the hole sidewall, leading to an inner channel diameter of ~50 nm. Gate stack is capped with a thin amorphous silicon sacrificial layer, used to protect the oxides during the opening of the bottom of the hole, required to provide the connection to the substrate [10]. The epi-channel growth is selective to Si and therefore proper preparation of the surface is required prior to growth. For this purpose, the Si substrate is cleaned as described in [11] in order to remove its native oxide, and the residual amorphous Si sacrificial layer has to be completely removed to allow an epitaxial growth exclusively from the bottom of the hole. This is done by using HCl vapor at 900 °C. The epi-channels are formed by using a reduced pressure chemical vapor deposition system. Epi-Si is grown from the substrate up to the top

978-1-4673-6933-6/15/$31.00 ©2015 IEEE
Authorized licensed use limited to: University of Texas at Austin. Downloaded on May 09,2024 at 21:15:43 UTC from IEEE Xplore. Restrictions apply.

Micron Ex. 1034, p. 1
Micron v. YMTC
IPR2024-00911

surface at a growth temperature of 810 °C; then, poly-Si is deposited in situ on top of it, to form the drain.

For the case of epi-$Si_{0.6}Ge_{0.4}$, the same reactor is used but an extra process step is introduced, consisting in 50 nm of epi-Si grown at the bottom of the memory holes with the conditions reported above; this allows to keep unmodified the source junction for more direct electrical assessment. Then, epi-growth is switched to $Si_{0.6}Ge_{0.4}$, at a temperature of 700 °C. Finally, the drain is formed by in-situ poly-Si deposition, as for the epi-Si case.

Fig. 2 shows Dark Field STEM images of poly-Si, epi-$Si_{0.6}Ge_{0.4}$ and epi-Si channels. It can be seen that the samples have been grown epitaxially (Fig. 2(b) and (c)), and present twins and stacking faults, which are more evident for epi-Si channel (Fig. 2(c)). Poly-Si channel (Fig. 2(a)) is characterized by many grains (highlighted in red).

III. Electrical characterization

Fig. 3 compares the shape of typical I_D-V_G characteristics for the different channel materials, showing the large improvement of epi-Si over both poly-Si and also epi-$Si_{0.6}Ge_{0.4}$ devices. In Fig. 4, the statistical distribution of I_D at fixed overdrive, subthreshold swing (STS) and threshold voltage (V_{th}), measured for a large population of devices, are compared for all channels. When comparing poly-Si to epi-Si channel, it is clear that the latter outperforms by far the first one. Single crystal conduction and absence of grain boundaries result in high I_D (Fig. 4(a)). Moreover, a clear improvement in STS (Fig. 4(b)) and V_{th} closer to 0 V (Fig. 4(c)) are also observed. The STS distribution in epi-Si devices appears to be bimodal: a tightly distributed mode is centered around 70 mV/dec, similar to in planar epi-Si devices [8], and a more widely distributed mode is observed, with STS-values that stretch up to ~300mV/dec.

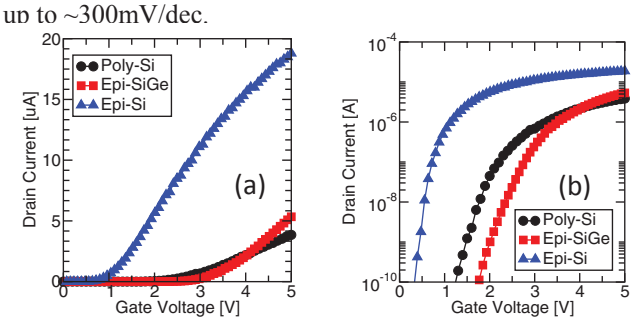

Fig. 3: Typical I_D-V_G in (a) linear scale and (b) logarithmic scale. The electrical characterization is performed by sweeping the gate up to 5 V while keeping 1 V at the drain.

Epi-$Si_{0.6}Ge_{0.4}$ performs worse than epi-Si; it shows a better I_D than poly-Si, but the highest STS and V_{th}. This is a clear indication of high trap density at the SiGe/SiO_2 interface: higher interface trap density will delay the formation of the inversion layer, resulting in high STS and V_{th}.

Given the scaled nature of our devices, featuring a channel diameter of ~50 nm, individual defect charging events can be distinguished when the IV-curve is recorded at a slow V_G ramp rate and with high voltage resolution (1 mV/step). The transconductance (g_m) is then extracted from these I_D-V_G curves in between charging events [8]. As a result, the intrinsic g_m value is obtained, and its statistical distribution is shown in Fig. 5 for all types of channels.

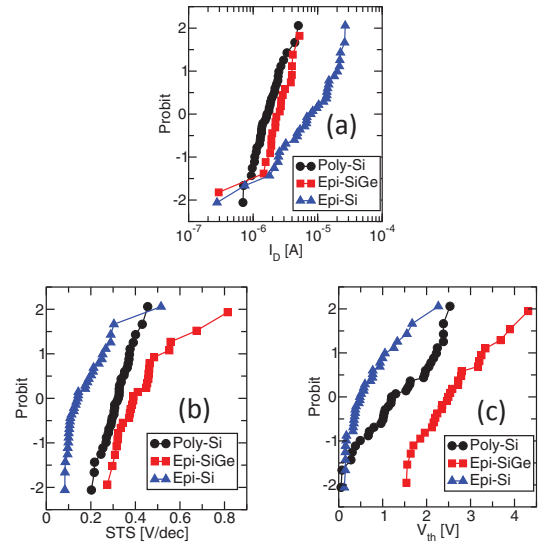

Fig. 4: Statistical distribution of (a) I_D at fixed overdrive of V_{th} + 2 V, (b) STS and (c) V_{th} recorded at a target current of 30 nA of poly-Si, epi-$Si_{0.6}Ge_{0.4}$ and epi-Si channel devices.

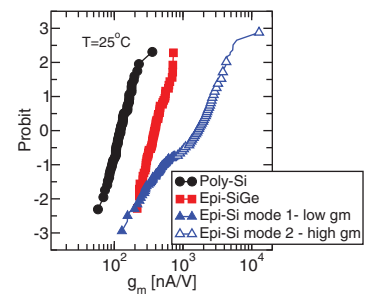

Fig. 5: Statistical distribution of g_m for poly-Si, epi- $Si_{0.6}Ge_{0.4}$ and epi-Si devices. Epi-Si g_m is bimodal corresponding to either grain boundary-controlled g_m (mode 1) or single-grain-controlled g_m (mode 2).

The highest transconductance is observed in epi-Si, but also epi-$Si_{0.6}Ge_{0.4}$ has three times higher g_m over poly-Si. Comparing this result with the I_D-distribution in Fig. 4(a), it is concluded that the low drain current in epi-$Si_{0.6}Ge_{0.4}$ is largely caused by defect charging while the intrinsic conduction properties are markedly improved with respect to poly-Si. Note that – as in Fig. 5 – the epi-Si transconductance distribution is bimodal. In the next section we will demonstrate that the low-transconductance mode in epi-Si is due to crystal stacking faults, while the high-transconductance mode corresponds to the single-grain behavior with a carrier mobility of ~100-120 $cm^2/(V \cdot s)$, in

Authorized licensed use limited to: University of Texas at Austin. Downloaded on May 09,2024 at 21:15:43 UTC from IEEE Xplore. Restrictions apply.

agreement with values extracted from single-crystal Si FinFET [7].

IV. Statistical analysis of g_m

Resistive network model

In previous work [12], a resistive network model was developed to simulate the statistical behavior of the g_m in scaled poly-Si channel devices. In this section, this model will be applied to provide a thorough understanding of the origin of the shape of the g_m distributions shown in Fig. 5. In particular, the bimodal behavior observed for epi-Si is interpreted.

In summary, the current flow between source and drain close to the cylindrical channel surface is modeled using a 2-D resistive network. The 3-D cylindrical geometry is forced through cyclic boundary conditions. A Monte-Carlo algorithm generates (in 2-D) Si grains from randomly positioned nucleation centers. The network resistor value inside a crystal grain ($R_{c\text{-}Si}$) is increased by a factor γ for resistors at the grain boundaries (R_{bound}), eq. 1:

$$R_{bound} = \gamma . R_{c-Si} \quad (1)$$

The value of $R_{c\text{-}Si}$ depends on the gate voltage using a EKV-type of relationship [12].

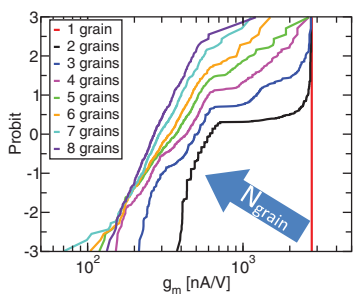

Fig. 6: Simulated g_m distributions considering an exact number of grains (N_{grains}) with no extra statistical distribution. The single grain g_m value is taken from FinFET data [7].

Applying this model, we have already shown that the variance of the g_m measured on a large number of identical devices increases when the average number of grains in the channel is decreased [12]. For an epi-grown device with one single grain, a distribution with very small variance would be expected, since no grain boundary-induced variability exists. As measured data present a bimodal distribution, in the next section we explored the possibility to explain this result by introducing more grains in the simulation.

Simulations of Epi-Si and epi-$Si_{0.6}Ge_{0.4}$

In Fig. 6, the simulated g_m distributions for 1 to 8 grains are shown. No Poisson distribution of the number of grains was introduced. The parameter γ is constant and no variations from any source other than the geometric distribution of grains is assumed. Consequently, for a 1-grain configuration, a single value corresponding to the maximum possible g_m is found. As the number of grains increases, bimodal and multimodal distributions are generated. For 8 grains, the multiple modes can no longer be distinguished and an apparent single-mode distribution with wide spread is obtained.

The origin of the multiple modes is traced back to the number and the spatial configuration of the grains in the device. As an example, in Fig. 7, devices with exactly 2 grains give rise to a bimodal distribution. The extremes correspond to a completely blocking grain boundary (mode-1, lowest g_m) and a perfectly source-to-drain aligned grain boundary (mode-2, highest g_m~maximum possible value).

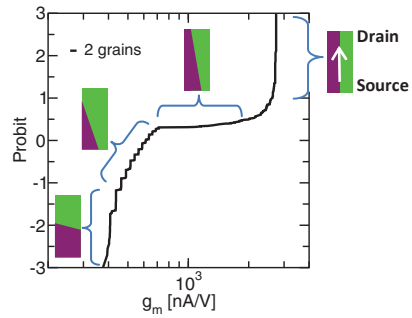

Fig. 7: Simulated g_m distribution using exactly 2 grains. The modal behavior is related to the possible grain configurations (schematically illustrated with the green/purple rectangles).

When attempting to fit the measured bimodal g_m-distribution on epi-Si (Fig. 5) with the bimodal distribution of Fig. 7, an additional lognormal spread on the maximum g_m-value has to be added, which originates from multiple other sources of variability, such as geometric device-to-device variations (channel length, width, dielectric thickness, etc...), interface trap density variations, contact and line resistance variations. After this modification, however, a good fitting of the epi-Si data is still not achieved: a simulation with exactly 2 grains does not quantitatively reproduce the measured g_m distribution. An excellent fitting can be instead obtained by assuming 66% of the devices having 1 grain and 34% having 2 grains (Fig. 8). Therefore, mode-1 at low g_m is grain boundary-limited, while mode-2 at high g_m reflects the single-crystal value without any grain boundary obstruction. Note that the γ-value of 160 is significantly lower than the values reported in poly-Si (γ=~700) [12]. This indicates that in epi-Si, the eventual grain boundaries are less 'resistive' and could be interpreted as stacking faults or twins rather than poly-Si-like grain boundaries, as also concluded from TEM inspections (Fig. 2).

Temperature dependence of g_m

In order to verify the interpretation laid out in the previous section, the temperature dependence of the g_m was measured for mode-1 and mode-2 devices. Activation energies using an Arrhenius dependence were extracted through eq. 2:

Authorized licensed use limited to: University of Texas at Austin. Downloaded on May 09,2024 at 21:15:43 UTC from IEEE Xplore. Restrictions apply.

$$E_{ACT} = \frac{\ln \frac{g_{m,T1}}{g_{m,T2}}}{\left(\frac{1}{kT_2} - \frac{1}{kT_1}\right)} \quad (2)$$

with T_1=298K and T_2=398K.

As shown in Fig. 9, the g_m for mode-1 devices increases with temperature (positive E_{ACT} = ~15 meV), a behavior typically associated with thermionic emission over grain boundaries in polycrystalline materials. The g_m for mode-2 devices on the contrary decreases with temperature (negative E_{ACT} = ~-20 meV), a behavior associated with scattering-controlled single crystal mobility. The results in Fig. 9 therefore confirm the interpretation of the bimodal behavior being caused by non-perfect epi-growth.

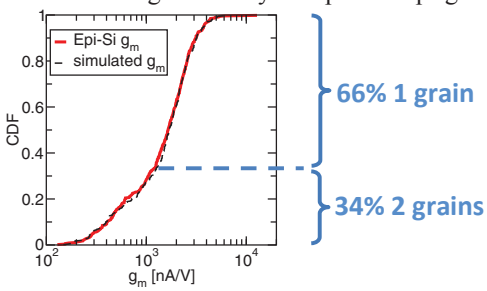

Fig. 8: Correlation between observed epi-Si g_m data and simulated g_m with distributed N_{grain} (66% single grain and 34% 2 grains configuration).

For epi-$Si_{0.6}Ge_{0.4}$ devices, the g_m-distribution is monomodal at the level of mode-1 in epi-Si (Fig. 5). Furthermore, the g_m-temperature dependence also coincides with that of mode-1 epi-Si (Fig. 9). Therefore, we conclude that the conductance in all measured epi-$Si_{0.6}Ge_{0.4}$ devices is controlled by thermionic emission over a barrier. This barrier can either come from the very defective interface with the gate stack or from the interfaces between the $Si_{0.6}Ge_{0.4}$ channel and the Si source and drain junctions (Fig. 2). Consequently, the real conductance in SiGe-crystalline material is not observed. A significant g_m increase is still expected to be achievable in SiGe-only devices.

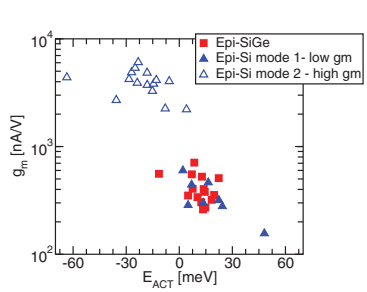

Fig. 9: Correlation between g_m and E_{ACT} for epi-$Si_{0.6}Ge_{0.4}$ and epi-Si.

V. Conclusions

In this study we have integrated epi-Si and epi-$Si_{0.6}Ge_{0.4}$ as replacement of poly-Si channel in vertical NAND memories. Successful epitaxial integration scheme requires a careful preparation of the starting Si surface, achievable by properly cleaning native oxides and by removing silicon from memory hole sidewall to allow epi-growth only from the bottom opening of the memory hole. We have electrically demonstrated that epi-Si channel results in the best conduction, with large improvements on both STS and g_m over epi-SiGe and poly-Si channels, and electron mobility comparable to the one of monocrystalline Si FinFETs. We also found that g_m is bimodally distributed for epi-Si; this feature has been successfully explained by using our resistive network model: high g_m is obtained in most devices, while lower g_m conduction occurs when the current needs to cross a higher resistance boundary, which is intrinsically different from a typical grain boundaries found in poly-Si, and likely resulting from stacking faults or twins. Conduction in epi-$Si_{0.6}Ge_{0.4}$ also shows improvements over poly-Si. However the analysis of the current has shown to be controlled by defects.

We have demonstrated that the epitaxially grown channel have the potential to give the highest possible current performance in vertical NAND memories.

Acknowledgments

This work is supported by imec's Industrial Affiliation Program on Advanced Flash Memory devices.

References

[1] K. T. Park et al., "Three-dimensional 128Gb MLC vertical NAND Flash-memory with 24-WL stacked layers and 50MB/s high-speed programming", in *Proc. IEEE ISSCC*, 2014, pp. 334-335.

[2] C. Sun et al., "A workload-aware-design of 3D-NAND flash memory for enterprise SSDs", in *Proc. Int. Symp. Quality Electronic design*, 2014, pp. 554-561.

[3] J. Choi and K. Seol, "3D Approaches for Non-volatile Memory", *in Proc.*: 2011 Symposium on VLSI Technology (VLSIT), 2011, pp.178-179.

[4] H. Tanaka, et al., "Bit Cost Scalable Technology with punch and Plug Process for Ultra High Density lash Memory", in *VLSI Symp. Tech. Dig.*, 2007, pp. 14-15.

[5] R. Katsumata *et al.*, "Pipe-shaped BiCS flash memory with 16 stacked layers and multi-level-cell operation for ultra-high density storage devices", *in Proc.*: 2009 Symposium on VLSI Technology, 2009, pp.136-137.

[6] J. Jang, et al., "Vertical Cell Array using TCAT (Terabit Cell Array Transistor) technology for Ultra High Density NAND Flash Memory", in *VLSI Symp. Tech. Dig.*, 2009, pp. 192-193.

[7] E. Nowak, et al., "Intrinsic Fluctuations in Vertical NAND Flash Memories", in *VLSI Symp. Tech. Dig.*, 2012, pp. 21-22.

[8] M. Toledano-Luque, et al., "Quantitative and predictive model of reading current variability in deeply scaled vertical Poly-Si channel for 3D memories", IEDM, 2012, pp. 203-206.

[9] J. G. Lisoni, et al., "Laser Thermal Anneal of polysilicon channel to boost 3D memory performance", in *VLSI Symp. Tech. Dig.*, 2014.

[10] G. Van den bosch *et al.*, "Highly Scaled Vertical Cylindrical SONOS Cell With Bilayer Polysilicon Channel for 3-D NAND Flash Memory", IEEE Electron Device Letters, vol.32, no.11, Nov. 2011, pp.1501-1503.

[11] A. Hikavyy, et al., "Growth of high Ge content SiGe on (110) oriented Si wafers", Thin Solid Films vol. 520, 2012, pp. 3179–3184.

[12] R. Degraeve. et al., "Characterization size and defect energy distribution in vertical SONOS poly-Si channels by means of a resistive network model", IEDM, 2013, pp. 558-560.

Authorized licensed use limited to: University of Texas at Austin. Downloaded on May 09,2024 at 21:15:43 UTC from IEEE Xplore. Restrictions apply.