

**METHOD OF FABRICATING A SILICON TUNNELING FIELD EFFECT
TRANSISTOR (TFET) WITH HIGH DRIVE CURRENT**

TECHNICAL FIELD

[0001] The present disclosure relates generally to devices and methods of fabrication of semiconductor devices, and more particularly to the fabrication of tunneling field-effect transistors (TFETs) having high drive current.

BACKGROUND

[0002] As conventional CMOS transistors are scaled down, power dissipation due to off-state leakage current increases. In an effort to reduce this power consumption, tunneling field effect transistors (TFET) have been proposed because of the extremely low off-state leakage current capability. Additional technical background information regarding TFETs may be found in "Tunneling Field-Effect Transistors (TFETs) With Subthreshold Swing (SS) Less Than 60 mV/dec," Choi, W.Y., et al., IEEE Electron Device Letters, Vol. 28, No. 8, August 2007; "Double-Gate Strained-Ge Heterostructure Tunneling FET (TFET) With Record High Drive Currents and <60mV/dec Subthreshold Slope," Krishnamohan, Tejas, et al., IEEE 2008; "Impact of SOI, Si_{1-x}Ge_xOI and GeOI Substrates On CMOS Compatible Tunnel FET Performance," Mayer, F., et al., IEEE 2008; and "Green Transistor - A V_{DD} Scaling Path For Future Low Power ICs," Hu, Chenming, et al., IEEE 2008, which are incorporated herein by reference.

[0003] Referring to FIGURE 1A, there is shown a conventional TFET. Instead of both source and drain regions having the same conductivity type (e.g., P-type or N-type) as done in conventional MOSFETs, the TFET is constructed to include a P+ type drain/source

and N+ type source/drain in a substrate. As will be appreciated, a conventional TFET may act as either a p-channel or n-channel depending on the gate voltage value.

[0004] Referring to FIGURES 1B and 1C, in the OFF-state, the gate bias voltage (V_{gs}) is zero with insufficient band bending to allow tunneling. In this state, the leakage current is extremely low and is due to the drift of minority carriers (similar to that of a reverse biased P-I-N diode). In the ON-state (usually with the P+ source grounded and the N+ drain at a fixed positive bias), the gate induces sufficient band bending resulting in a reduction of the tunneling barrier width to less than 5 nanometers (nm). This leads to substantial band-to-band tunneling of electrons from the valence band of the P+ source region to the conduction band of the N+ drain region.

[0005] Advantages of conventional TFETs include a sub-60 mV subthreshold slope, low off-state leakage current, control of short channel effects, and an intrinsic channel that minimizes any dopant fluctuation issues. In contrast, the main disadvantage of conventional TFETs is the low drive current because the current originates from band-to-band tunneling. In other words, the drive current (drain current I_d) capability of conventional TFETs is much

lower than conventional MOSFETs. FIGURE 2 illustrates a general comparison of the I_d - V_g curves between a conventional TFET and conventional MOSFET.

[0006] To increase the drive current, it has been proposed to utilize silicon-germanium (SiGe) in the source region near the tunneling region. FIGURES 3A, 3B and 3C illustrate the typical tunneling width for silicon and SiGe in the source region. As shown in FIGURE 3B, the increase in the valence bandgap (E_v) resulting from a SiGe source region decreases/narrows the tunneling barrier width. However, because conventional SiGe processes usually result in a gap or space between the SiGe source region and the gate stack due to an offset spacer, the tunneling barrier width is still determined by the valence bandgap of silicon (in the channel). Inclusion of the spacer gap causes an increase (shown as L_H) in the tunnel barrier width - effectively negating most of the benefits gained by utilizing SiGe in the source region.

[0007] Accordingly, there is a need for an improved TFET structure (and method of fabricating) having increased drive current capabilities.

SUMMARY

[0008] In accordance with one embodiment, there is provided a method of forming a tunneling field effect (TFET) semiconductor device. The method includes providing a substrate comprising a first substrate material and providing a gate structure on the substrate, the gate structure comprising a gate and a gate dielectric. A drain region of a first type is formed adjacent the gate structure and a source region of a second type is formed adjacent the gate structure, wherein at least a portion of the source region is positioned underneath the gate dielectric such that the gate dielectric overlaps the source region.

[0009] In accordance with another embodiment, there is provided a tunneling field-effect transistor (TFET) having a substrate, a gate structure on the substrate, the gate structure comprising a gate and a gate dielectric, a drain region of a first type adjacent the gate structure, and a source region of a second type adjacent the gate structure, wherein at least a portion of the source region is positioned underneath the gate dielectric such that the gate dielectric overlaps the source region.

[0010] In yet another embodiment there is provided a method of fabricating a tunneling field-effect transistor (TFET) device. The

method includes providing a substrate having a first substrate material and forming a gate structure thereon, the gate structure comprising a gate and a gate dielectric. A drain region of a first type is formed adjacent the gate structure and an intermediate source region is formed adjacent the gate structure. At least a portion of the source region is positioned underneath the gate dielectric such that the gate dielectric overlaps the source region. The method includes selectively removing substantially all of the intermediate source region and forming a cavity positioned beneath the gate dielectric and disposed between the substrate and the gate dielectric and exposing a portion of the substrate. A first source region is formed on the exposed portion of the substrate and a second source region is formed above the first source region, with the second region of a second type different from the first type, wherein at least a one of a portion of the first source region and a portion of the second source region is disposed within the cavity.

[0011] Other technical features may be readily apparent to one skilled in the art from the following figures, descriptions, and claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] For a more complete understanding of the present disclosure, and the advantages thereof, reference is now made to the following descriptions taken in conjunction with the accompanying drawings, wherein like numbers designate like objects, and in which:

[0013] FIGURES 1A, 1B and 1C illustrate a conventional tunneling field effect transistor (TFET) and representations of the OFF and ON states of the TFET, respectively;

[0014] FIGURE 2 illustrates a general comparison of the I_d - V_g curves between a conventional TFET and conventional MOSFET;

[0015] FIGURES 3A, 3B and 3C are diagrams illustrating the difference in tunneling widths for silicon and SiGe in the source region of a TFET;

[0016] FIGURE 4 is a cross-sectional view of TFET in accordance with the present disclosure; and

[0017] FIGURES 5A-5J are cross-sectional views illustrating various steps of a method or process in accordance with the present disclosure.

DETAILED DESCRIPTION

[0018] Reducing the tunneling barrier width and increases drive current (i.e., drive current I_d). Increasing band-to-band tunneling (BTBT) also increases drive current. It has been determined that various structures and device modifications can be implemented to increase drive current: (1) material having a lower E_g (bandgap voltage) in the tunneling region, (2) an abrupt junction at the tunneling region (for generating a higher E-field with more band bending), (3) high doping concentration in the source region, and (4) positioning the source region closer to the channel.

[0019] The present disclosure provides a TFET device (and methods of fabrication) with increased drive current and other performance enhancements. Though these will be described hereinafter, some of the advantages of the present disclosure are directed to the use of silicon channel instead of SiGe (increased reliability), the use of SiGe in the source region (not in drain region), controlling SiGe profile to achieve higher drive current, in-situ doped P+ SiGe source providing an abrupt junction and high source doping concentration, and SiGe source with high Ge content to achieve high drive current. As will be appreciated, not all

structures or methods described herein are required to form the TFET of the present disclosure in order to obtain increased drive current (or any other advantages described herein).

[0020] Referring to FIGURE 4, there is depicted a cross-sectional view of a TFET device structure 100 formed in or on a substrate 110 in accordance with the present disclosure. The TFET 100 includes the substrate 110, a gate stack 120, a source region 130 and a drain region 140. The gate stack 120 includes a gate 122 and a gate dielectric layer 124. The gate stack 120 may further include sidewall spacers (not shown) and the gate 122 may be formed of polysilicon. As will be appreciated, the structures and regions shown in the FIGURES are not drawn to scale or actual form, and are for illustrative purposes.

[0021] In one embodiment, the substrate 110 is silicon, and in other embodiments, it may include, for example, silicon-germanium, silicon-on-insulator (SOI), or other suitable semiconductor substrate materials, now known or later developed. The substrate 110 may include silicon (e.g., n-type, p-type, or no type) provided in a single well or twin-well process, and may further include an epitaxial layer.

[0022] In the embodiment shown, the source region 130 is formed

of silicon-germanium (SiGe) with p-type dopants (or impurities) and the drain region 140 is formed of silicon (but could be SiGe) with n-type dopants. Further, the source region 130 includes a first source region 134 of a first type and a second source region 132 of a second type. In one embodiment, the first source region 134 is undoped SiGe and the second source region 132 is P+ doped SiGe.

[0023] The source region 130 extends to a point that is substantially underneath the gate dielectric layer 122. In one embodiment, the region 130 extends to the edge of the dielectric layer 122 (aligned). In another embodiment, the region 130 extends beyond the edge, and to a point beneath the layer 122 (overlapped).

In the overlapped configuration, the gate dielectric layer 122 overlaps or is overlapping the source region 130. Referring to FIGURE 4, although the second source region 134 is shown as being in the overlapped configuration, while the first source region 132 is shown as being aligned with the edge of the gate dielectric layer 122, this is for illustration only. Other configurations may be implemented, for example, in which a portion of the first source region 132 is also disposed beneath the gate dielectric layer 122.

The amount of overlap suitable for providing benefits as described herein vary depending on device parameters, such as doping

concentration, oxide thickness, gate length. In one embodiment, the source region 130 overlaps with the dielectric layer 122. In another embodiment, the distance of overlap ranges from about 1 nm to about 4 nm. In yet another embodiment, the overlap is about 2 nm.

[0024] The positioning of the source region 130 at least partially underneath the gate dielectric layer 122 reduces the tunneling barrier width, thereby leading to a higher drive current. Also shown in FIGURE 4 is a dotted circle representing, generally, the tunneling region of the TFET 100.

[0025] As will be understood, the TFET 100 may be formed using conventional processes. Moreover, more than one TFET may be fabricated on the substrate.

[0026] Now referring to FIGURES 5A-5J, there are shown cross-sectional views of a process in accordance with this disclosure for fabricating or forming the TFET 100. With specific reference to FIGURE 2A, there is provided the substrate 110. The gate stack 120 is formed in accordance with well-known semiconductor fabrication techniques. In the embodiment shown, the gate dielectric layer 122 may be silicon oxide and the gate 124 may be silicon nitride; however, other material(s) may be utilized for these structures.

As noted previously, sidewall spacers (not shown) may optionally be constructed.

[0027] Referring to FIGURE 5B, a next step in the process includes doping certain regions of the substrate (these will eventually be the source/drain regions 130, 140 and are shown by dotted lines). N⁺ dopants are selectively introduced into the substrate 110 to form an intermediate source region 502 and the drain region 140. In one embodiment, the N⁺ dopants are introduced using implantation; however, alternative doping processes known to those skilled in the art may be utilized. It will be understood additional thermal processing may be performed to diffuse the N⁺ dopant material within the substrate 110, as desired. In one embodiment, implantation is performed in a substantially vertical direction with respect to the substrate 110 (i.e., perpendicular to the substrate). In another embodiment, implantation is performed at an angle (i.e., tilted), as shown in FIGURE 5B, causing N⁺ dopants to disperse in a manner that forms at least a portion of the intermediate N⁺ source region underneath the gate dielectric layer 122. Any suitable angle and power for implantation may be utilized to generate the desired distance of source/channel overlap. In various embodiments, the angle is about 5 degrees or

greater, 10 degrees or greater, or may be 20 degrees or greater. In another embodiment, a combination of vertical and tilted implantations may be used.

[0028] Referring to FIGURES 5C and 5D, a mask layer 510 is formed over the drain region 140, as shown. The mask layer 510 may be a silicon oxide layer. The mask layer 510 is typically formed by depositing oxide over the substrate 110 followed by mask definition and etch (using conventional processes). After the mask layer 510 is formed, the substrate surface is selectively etched to remove the intermediate source drain region 502 (or at least a substantial portion of it). This results in a cavity underneath the gate dielectric layer 122, as shown in FIGURE 5D. In other words, the gate dielectric layer 122 overhangs an empty volume positioned between the substrate 110 and the gate dielectric layer 122.

[0029] Referring to FIGURES 5E, a layer of material is formed on the exposed portion of the substrate 110 (which will eventually form the source region 130). As will be appreciated, this layer of material is the first source region 134 shown in FIGURE 4. In one embodiment, the first source region or layer 134 is an epitaxial layer of SiGe formed in-situ. The main purpose of the undoped SiGe

layer 132 is to ensure the tunneling region is formed of SiGe. As will be appreciated, the higher the Ge concentration, the lower the tunneling barrier width. In other embodiments, the layer 134 may be formed in a different manner, and may further include different types of material, such as Ge, GeSn or SiGeSn.

[0030] Referring to FIGURE 5F, after the first source region layer 134 is formed, another layer of material is formed above the exposed portion of the layer 134 (which together will eventually form source region 130). In one embodiment, the layer 132 is formed on the surface of the layer 134. As will be appreciated, this second layer of material is the second source region 132 shown in FIGURE 4. In one embodiment, the second source region or layer 132 is an epitaxial layer of SiGe formed in-situ with P⁺ species. It is desirable to have a high germanium (Ge) concentration or content within the second source region 132. In one embodiment, the Ge concentration at this formation process is about 30% or higher (e.g., Si_{0.7}Ge_{0.3}). Other concentrations of Ge may be used, such as about 20% or higher, or about 15% or higher. However, in the formation of SiGe using a conventional selective epitaxial growth (SEG) process, a concentration higher than about 30% may result in undesirable dislocations.

[0031] As will be appreciated, the steps shown in FIGURES 5E and 5F relating to the formation of the SiGe layers 132, 134 (or re-formation of the source region 130) may include additional processing, such as etching to provide the proper profile and thickness to the layers 132, 134. In various embodiments, the thickness of layer 134 is about 5 nm or less, ranges from about 1 nm to about 4 nm, and in yet another embodiment, is about 2 nm.

[0032] It will be understood that the formation of the source region 130 into the first source region 134 (undoped SiGe) different from the second source region 132 (P+ SiGe) provides an abrupt junction in the tunneling region. The dopant gradient should be relatively steep.

[0033] Referring now to FIGURES 5G and 5H, an amorphous germanium layer 520 is formed over at least the source region 130 on the P+ SiGe layer 132. After formation, an anneal process is performed that drives additional germanium into the P+ SiGe second source region 132 to increase the Ge concentration (e.g., greater than 30% when $\text{Si}_{0.7}\text{Ge}_{0.3}$ is used). In one embodiment, the anneal process is performed using laser irradiation. Laser annealing induces surface melting of both amorphous Ge film and the SiGe layer 132 underneath (a silicon substrate will not melt as its

melting point is higher than Ge and SiGe). This process causes inter-diffusion of liquid Ge and silicon and results in a melt having a higher Ge concentration. The melt then re-crystallizes epitaxially with the SiGe layer as a template. The laser anneal process can be controlled to selectively melt the layer 132, but not the layer 134 (or not much) or Si substrate. This may be achieved because the melting point of Ge is lower than that of the layer 132 and the substrate. As a result, the second source region 132 is a P+ SiGe region with a Ge concentration greater than 30%.

[0034] Now turning to FIGURES 5I and 5J, the unreacted Ge layer 520 is removed, and the silicon oxide mask layer 510 is removed. The resulting device is the TFET structure 100 shown in FIGURE 4. It will be understood that the gate stack 120 may be a dummy gate stack which can be removed and a replaced with another gate stack.

[0035] The order of steps or processing can be changed or varied from that described above. Though not shown, additional devices and structures are typically formed on/in the substrate along with the TFET. It will be understood that well known processes have not been described in detail and have been omitted for brevity.

Although specific steps, insulating materials, conductive materials and apparatuses for depositing and etching these materials may have been described, the present disclosure may not be limited to these specifics, and others may be substituted as is well understood by those skilled in the art.

[0036] It may be advantageous to set forth definitions of certain words and phrases used throughout this patent document. The terms "include" and "comprise," as well as derivatives thereof, mean inclusion without limitation. The term "or" is inclusive, meaning and/or. The phrases "associated with" and "associated therewith," as well as derivatives thereof, mean to include, be included within, interconnect with, contain, be contained within, connect to or with, couple to or with, be communicable with, cooperate with, interleave, juxtapose, be proximate to, be bound to or with, have, have a property of, or the like.

[0037] While this disclosure has described certain embodiments and generally associated methods, alterations and permutations of these embodiments and methods will be apparent to those skilled in the art. Accordingly, the above description of example embodiments does not define or constrain this disclosure. Other changes, substitutions, and alterations are also possible without departing

from the spirit and scope of this disclosure, as defined by the following claims.

WHAT IS CLAIMED IS:

1. A method of forming a tunneling field effect (TFET) semiconductor device, the method comprising:

providing a substrate comprising a first substrate material;
providing a gate structure on the substrate, the gate structure comprising a gate and a gate dielectric;

forming a drain region of a first type adjacent the gate structure; and

forming a source region of a second type adjacent the gate structure, wherein at least a portion of the source region is positioned underneath the gate dielectric such that the gate dielectric overlaps the source region.

2. The method in accordance with Claim 1 wherein the substrate and drain region are formed of silicon and the source region comprises silicon germanium (SiGe).

3. The method in accordance with Claim 1 wherein forming the source region further comprises:

forming a first source region having a layer of undoped silicon germanium (SiGe); and

forming above the first source region a second source region having a layer of doped SiGe.

4. The method in accordance with Claim 3 wherein forming the second source region comprises epitaxially growing the undoped SiGe layer and forming the first source region comprises epitaxially growing the doped SiGe layer.

5. The method in accordance with Claim 4 wherein the germanium concentration within the doped SiGe layer is greater than about 20%.

6. The method in accordance with Claim 1 wherein forming the source region further comprises:

doping the source region with dopants of a first type;

selectively removing substantially all of the doped source region and forming a cavity positioned beneath the gate dielectric and between the substrate and the gate dielectric; and

epitaxially re-forming the source region using at least one layer of silicon germanium (SiGe).

7. The method in accordance with Claim 6 wherein epitaxially re-forming the source region further comprises:

forming a first source region having a layer of undoped silicon germanium (SiGe); and

forming above the first source region a second source region having a layer of doped SiGe.

8. The method in accordance with Claim 7 wherein forming the source region further comprises:

forming amorphous germanium on a surface of the SiGe layer;
and

annealing the amorphous germanium to drive germanium into the SiGe layer thereby increasing germanium content of the SiGe layer.

9. The method in accordance with Claim 6 wherein doping the source region further comprises:

implanting the dopants at an angle such that at least some of the implanted dopants are positioned beneath the gate dielectric.

10. The method in accordance with Claim 1 wherein forming the source region further comprises:

forming a layer of silicon germanium (SiGe);
forming amorphous germanium on a surface of the SiGe layer;
and

annealing the amorphous germanium to drive germanium into the SiGe layer thereby increasing germanium content of the SiGe layer.

11. A tunneling field-effect transistor (TFET) comprising:
a substrate;
a gate structure on the substrate, the gate structure comprising a gate and a gate dielectric;
a drain region of a first type adjacent the gate structure;
and
a source region of a second type adjacent the gate structure, wherein at least a portion of the source region is positioned underneath the gate dielectric such that the gate dielectric overlaps the source region.

12. The TFET in accordance with Claim 11 wherein the substrate and drain region are formed of silicon and the source region comprises silicon germanium (SiGe).

13. The TFET in accordance with Claim 11 wherein the source region further comprises:

a first source region having a layer of undoped silicon germanium (SiGe); and

a second source region having a layer of doped SiGe and disposed above the undoped SiGe layer.

14. The TFET in accordance with Claim 13 wherein the germanium concentration within the doped SiGe layer is greater than about 30%.

15. The TFET in accordance with Claim 13 wherein the undoped SiGe layer comprises undoped epitaxial SiGe and the doped SiGe layer comprises epitaxial SiGe having P-type dopants.

16. A method of fabricating a tunneling field-effect transistor (TFET) device, the method comprising:

providing a substrate comprising a first substrate material;
forming a gate structure on the substrate, the gate structure comprising a gate and a gate dielectric;

forming a drain region of a first type adjacent the gate structure; and

forming an intermediate source region adjacent the gate structure, wherein at least a portion of the source region is positioned underneath the gate dielectric such that the gate dielectric overlaps the source region;

selectively removing substantially all of the intermediate source region and forming a cavity positioned beneath the gate dielectric and disposed between the substrate and the gate dielectric and exposing a portion of the substrate;

forming on the exposed portion of the substrate a first source region;

forming above the first source region a second source region of a second type different from the first type; and

wherein at least a one of a portion of the first source region

and a portion of the second source region is disposed within the cavity.

17. The method in accordance with Claim 16 wherein the first substrate material is silicon, the drain region is silicon, and the source region comprises silicon germanium (SiGe).

18. The method in accordance with Claim 17 wherein forming the first source region further comprises epitaxially forming a doped SiGe layer, and wherein forming the second source region further comprises epitaxially forming a doped layer of silicon germanium (SiGe).

19. The method in accordance with Claim 17 further comprising:

forming amorphous germanium on a surface of the undoped SiGe layer; and

annealing the amorphous germanium to drive germanium into the SiGe layer thereby increasing germanium content of the SiGe layer.

20. The method in accordance with Claim 19 wherein the germanium concentration within the doped SiGe layer is less than about 30% prior to annealing and is greater than about 30% after annealing.