MOSFET Device With Localized Stressor

TECHNICAL FIELD

[0001] The present invention relates generally to semiconductor devices, and more particularly, to metal-oxide-semiconductor field-effect transistors and methods of manufacture.

BACKGROUND

[0002] Size reduction of metal-oxide-semiconductor field-effect transistors (MOSFET), including reduction of the gate length and gate oxide thickness, has enabled the continued improvement in speed, performance, density, and cost per unit function of integrated circuits over the past few decades. To further enhance transistor performance, MOSFET devices have been fabricated using strained channel regions located in portions of a semiconductor substrate. Strained channel regions allow enhanced carrier mobility to be realized, thereby resulting in increased performance when used for n-channel (NMOSFET) or for p-channel (PMOSFET) devices. Generally, it is desirable to induce a tensile strain in the n-channel of an NMOSFET transistor in the source-to-drain direction to increase electron mobility and to induce a compressive strain in the p-channel of a PMOSFET transistor in the source-to-drain direction to increase hole mobility. There are several existing approaches of introducing strain in the transistor channel region.

[0003] In one approach, semiconductor alloy layers, such as silicon-germanium or silicon-germanium-carbon, are formed below an overlying thin semiconductor layer, wherein the semiconductor alloy layer has a different lattice structure than the overlying semiconductor layer. The difference in the lattice structure imparts strain in the overlying semiconductor layer to increase carrier mobility.

[0004] This approach, however, can be difficult to process in addition to presenting junction leakage concerns as a result of the blanket semiconductor alloy layer. The level of germanium in the epitaxially grown semiconductor alloy layer can be difficult to control. In addition, the presence of a blanket semiconductor alloy layer allows an unwanted interface between the source/drain regions to exist, possibly introducing junction leakage

[0005] In another approach, strain in the channel is introduced by creating a recess in the substrate in the source/drain regions. A layer of SiGe is epitaxially grown within the recessed regions, thereby introducing strain in the channel. The amount of stress may be increased by increasing the Ge concentration during the growth process. To increase the Ge concentration in the recessed area, however, creates process challenges. For example, increasing the Ge concentration during the epitaxial growth results in a higher density of dislocations and defects in the SiGe layer. Degraded selectivity and deposition process windows are also of concern.

[0006] Therefore, there is a need for an efficient and cost-effective method to induce strain in the channel region such that the performance characteristics of transistors are enhanced.

SUMMARY OF THE INVENTION

[0007] These and other problems are generally reduced, solved or circumvented, and technical advantages are generally achieved, by embodiments of the present invention, which provides a strained semiconductor device to improve the operating characteristics of the semiconductor device and a method of manufacture.

[0008] In an embodiment of the present invention, a metal-oxide-semiconductor field-effect transistor (MOSFET) having localized stressors formed in the source/drain regions is provided. The localized stressors comprise a stress-inducing layer formed in recesses in the source/drain regions. The stress-inducing layer comprises a first semiconductor material and a second semiconductor material. A process is performed on the stress-inducing layer such that the concentration of the second semiconductor material is increased.

[0009] In another embodiment of the present invention, the stress-inducing layer is formed over a non-recessed region of the source/drain regions.

[0010] In another embodiment of the present invention, a first method of forming a MOSFET having localized stressors formed in the source/drain regions is provided. The first method comprises forming a stress-inducing layer in the source/drain regions (either recessed or non-recessed), wherein the stress-inducing layer comprises a first semiconductor material and a second semiconductor material. A treatment may then be performed. The treatment may be any process, such as an oxidation or a nitridation process, that causes an ambient to react with the first semiconductor material and forces the second semiconductor material further down into the stress-inducing layer.

[0011] In yet another embodiment of the present invention, a second method of forming a MOSFET having localized stressors formed in the source/drain regions is provided. The second method comprises forming a stress-inducing layer in the source/drain regions (either recessed or non-recessed), wherein the stress-inducing layer comprises a first semiconductor material and a second semiconductor material. A reaction layer is formed over the stress-inducing layer and a treatment is performed to cause the reaction layer to react with the first semiconductor material. As a result, the second semiconductor material is forced further into the stress-inducing layer.

[0012] It should be appreciated by those skilled in the art that the conception and specific embodiment disclosed may be readily utilized as a basis for modifying or designing other structures or processes for carrying out the same purposes of the present invention. It should also be realized by those skilled in the art that such equivalent constructions do not depart from the spirit and scope of the invention as set forth in the appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0013] The object and other advantages of this invention are best described in the preferred embodiment with reference to the attached drawings that include:

[0014] Figures 1-6 illustrate various process steps of fabricating a MOSFET device having localized stressors positioned over the source/drain regions in accordance with a first method embodiment of the present invention; and

[0015] Figures 7-8 illustrate various process steps of fabricating a MOSFET device having localized stressors positioned over the source/drain regions in accordance with a second method embodiment of the present invention.

DETAILED DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS

[0016] The making and using of the presently preferred embodiments are discussed in detail below. It should be appreciated, however, that the present invention provides many applicable inventive concepts that can be embodied in a wide variety of specific contexts. The specific embodiments discussed are merely illustrative of specific ways to make and use the invention, and do not limit the scope of the invention.

[0017] Figures 1-6 illustrate a method embodiment for fabricating a semiconductor device having a strained channel region in accordance with an embodiment of the present invention. Embodiments of the present invention illustrated herein may be used in a variety of circuits. Referring first to Figure 1, a wafer 100 having a gate insulator layer 110 and a gate electrode 112 formed on a substrate 114 is shown in accordance with an embodiment of the present invention. The substrate 114 may comprise bulk silicon, doped or undoped, or an active layer of a semiconductor-on-insulator (SOI) substrate. Generally, an SOI comprises a layer of a semiconductor material, such as silicon, formed on an insulator layer. The insulator layer may be, for example, a buried oxide (BOX) layer or a silicon oxide layer. The insulator layer is provided on a substrate, typically a silicon or glass substrate. Other substrates, such as a multi-layered or gradient substrate may also be used.

[0018] The gate insulator layer 110 and the gate electrode 112 may be formed and patterned as is known in the art on the substrate 114. The gate insulator layer 110 is preferably a high-K dielectric material, such as silicon oxide, silicon oxynitride, silicon nitride, an oxide, a nitrogen-containing oxide, a combination thereof, or the like. Preferably, the gate insulator layer 110 has a relative permittivity value greater than about 4. Other examples of such materials include

aluminum oxide, lanthanum oxide, hafnium oxide, zirconium oxide, hafnium oxynitride, or combinations thereof.

[0019] In the preferred embodiment in which the gate insulator layer 110 comprises an oxide layer, the gate insulator layer 110 may be formed by any oxidation process, such as wet or dry thermal oxidation in an ambient comprising an oxide, H_2O, NO, or a combination thereof, or by chemical vapor deposition (CVD) techniques using is tetra-ethyl-ortho-silicate (TEOS) and oxygen as a precursor. In the preferred embodiment, the gate insulator layer 110 is about 8 Å to about 50 Å in thickness, but more preferably about 16 Å in thickness.

[0020] The gate electrode 112 preferably comprises a conductive material, such as a metal (*e.g.*, tantalum, titanium, molybdenum, tungsten, platinum, aluminum, hafnium, ruthenium), a metal silicide (*e.g.*, titanium silicide, cobalt silicide, nickel silicide, tantalum silicide), a metal nitride (*e.g.*, titanium nitride, tantalum nitride), doped poly-crystalline silicon, other conductive materials, or a combination thereof. In one example, amorphous silicon is deposited and recrystallized to create poly-crystalline silicon (poly-silicon). In the preferred embodiment in which the gate electrode is poly-silicon, the gate electrode 112 may be formed by depositing doped or undoped poly-silicon by low-pressure chemical vapor deposition (LPCVD) to a thickness in the range of about 400 Å to about 2500 Å, but more preferably about 1500 Å.

[0021] The gate insulator layer 110 and the gate electrode 112 may be patterned by photolithography techniques as is known in the art. Generally, photolithography involves depositing a photoresist material, which is then masked, exposed, and developed. After the photoresist mask is patterned, an etching process may be performed to remove unwanted portions of the gate dielectric material and the gate electrode material to form the gate insulator layer 110 and the gate electrode 112 as illustrated in Figure 1. In the preferred embodiment in

which the gate electrode material is poly-crystalline silicon and the gate dielectric material is an oxide, the etching process may be a wet or dry, anisotropic or isotropic, etch process, but preferably is an anisotropic dry etch process.

[0022] Figure 2 illustrates wafer 100 after a mask 210 has been formed over the gate insulator layer 110 and gate electrode 112 in accordance with an embodiment of the present invention. The mask 210 acts as a mask to prevent epitaxial growth of SiGe along the surface of the gate insulator layer 110 and gate electrode 112 during subsequent processing steps. Generally, epitaxial growth of SiGe involves exposing silicon to germanium under specific process conditions. The germanium reacts with exposed silicon to grow SiGe. Because the gate insulator layer 110 and the gate electrode 112 are frequently formed of silicon, *e.g.*, silicon dioxide and polysilicon, respectively, it may be desired to form a mask over the gate insulator layer 110 and the gate electrode 112 to prevent SiGe growth thereon.

[0023] The mask 210 may be formed by performing a blanket deposition of silicon nitride via a chemical vapor deposition (CVD) process, a physical vapor deposition process (PVD), an atomic layer deposition process (ALD), or the like. Once deposited, the mask 210 may be patterned via photolithography techniques. Preferably, the mask 210 is about 0 Å to about 1000 Å in thickness. It should be noted that the stressor (formed in subsequent steps) may be formed substantially adjacent to, spaced away from, or underneath the gate electrode.

[0024] Referring now to Figure 3, the wafer 100 is shown after an etching process has been performed to create recessed regions 310 in the substrate 114 on either side of the gate electrode 112 in accordance with an embodiment of the present invention. It should be noted that the portion of the substrate 114 located between the recessed regions 310 and below the gate insulator layer 110 and the gate electrode 112 defines the channel region.

[0025] It should also be noted that the process described in Figures 1-6 illustrates source/drain regions having SiGe epitaxially grown in recessed regions 310 as an example of a preferred embodiment of the present invention. The processes described in Figures 1-6 may be applied to elevated-type SiGe regions in the source/drain regions. In this alternative embodiment, SiGe is epitaxially grown on the surface of the substrate 114 without forming recesses in the substrate 114. Thereafter, the same processes described herein may be performed to push the Ge lower, thereby creating an area of higher Ge concentration.

[0026] Referring back to Figure 3, the recessed regions 310 may be formed by a plasma etch using chlorine and bromine chemistry. Preferably, the recessed regions have a depth from about 10 nm to about 200 nm. An optional anneal may be performed to facilitate silicon migration to repair any etch damage as well as to slightly smoothen the silicon surface for the subsequent epitaxy process.

[0027] Figure 4 illustrates the wafer 100 after an epitaxial growth of a stress-inducing layer 410, such as silicon germanium, in the recessed regions 310 in accordance with an embodiment of the present invention. The epitaxy process used to perform the epitaxial growth may be chemical vapor deposition, ultra-high vacuum chemical vapor deposition (UHV-CVD), or molecular beam epitaxy. The epitaxially grown materials may also extend above or below the surface of the substrate 114. A raised source/drain region is shown for illustrative purposes only. In a preferred embodiment, the stress-inducing layer 410 extends from about 0 nm to about 200 nm above the surface of the substrate 114. The stress-inducing layer 410 is preferably from about 10 nm to about 300 nm in thickness. Other materials, such as SiC, SiGeN, and SiCN, may be used.

[0028] The stress-inducing layer 410 preferably comprises a first semiconductor material and a second semiconductor material, wherein the presence of the second semiconductor material has a different lattice structure than the first semiconductor material, thereby imparting strain in the channel region. In an embodiment in which a silicon substrate is used, the first semiconductor material may be silicon and the second semiconductor material may be germanium, and therefore, the stress-inducing layer 410 may be a SiGe layer.

[0029] Figure 5 illustrates the wafer 100 after performing a treatment to increase the concentration of the second semiconductor material, *e.g.*, Ge, in the lower section of the stress-reducing layer 410, creating a depletion layer 510. In an embodiment in which the stress-inducing layer 410 comprises SiGe, the treatment may comprise exposing the stress-reducing layer 410 to a gas or liquid ambient under process conditions that causes the gas or liquid ambient to react with the silicon of the SiGe material. As a result of the reaction, the depletion layer 510 will comprise SiX, where X denotes another material, and the Ge from the depletion layer 510 is pushed further down into the stress-inducing layer 410 below the depletion layer 510, as indicated by the directional arrows in Figure 5. In a preferred embodiment the resulting stress-inducing layer 410 below the depletion layer 510 exhibits a gradient concentration of Ge such that the concentration variation from the top to the bottom of the stress-inducing layer 410 is greater than at least about 5%.

[0030] In an embodiment, the treatment may comprise a thermal (furnace/RTP), plasma, UV, implant, or the like treatment in an ambient of an oxidation-based gas, such as O_2, H_2O, H_2, N_2O, NO, O_3, or the like at a temperature from about 100° C to about 1200° C. In this embodiment, the depletion layer may comprise a film of Si_xO_y.

[0031] In another embodiment, the treatment may comprise a thermal (furnace/RTP), plasma, UV, implant, or the like treatment in an ambient of a nitridation-based gas, such as NH_3, N_2, NO, N_2O, or the like at a temperature from about 100° C to about 1200° C. In this embodiment, the depletion layer may comprise a film of Si_xN_y or $Si_xO_yN_z$. Other ambients and/or treatments that cause a reaction with the Si and little or no reaction with the Ge may also be used.

[0032] Figure 6 illustrates the wafer 100 after the mask 210 has been removed and spacers 610 have been formed in accordance with an embodiment of the present invention. The mask 210 may be removed by a wet dip in hydrofluoric acid or a dry etch process.

[0033] Spacers 610, which form spacers for performing one or more ion implants to create the source/drain regions, preferably comprise silicon nitride (Si_3N_4), or a nitrogen containing layer other than Si_3N_4, such as Si_xN_y, silicon oxynitride SiO_xN_y, silicon oxime $SiO_xN_y:H_z$, or a combination thereof. In a preferred embodiment, the spacers 610 are formed from a layer comprising Si_3N_4 that has been formed using chemical vapor deposition (CVD) techniques using silane and ammonia as precursor gases.

[0034] The spacers 610 may be patterned by performing an isotropic or anisotropic etch process, such as an isotropic etch process using a solution of phosphoric acid (H_3PO_4). Because the thickness of the layer of Si_3N_4 is greater in the regions adjacent to the gate electrode 112, the isotropic etch removes the Si_3N_4 material on top of the gate electrode 112 and the areas of substrate 114 not immediately adjacent to the gate electrode 112, leaving the spacer 610 as illustrated in Figure 6. In an embodiment, the spacers 610 are from about 1nm to about 100 nm in width.

[0035] It should be noted that a portion of the stress-inducing layer 410 may be located underneath a portion of the spacers 610, dependent upon the desired source/drain implant and electrical characteristics of the resulting source/drain regions.

[0036] Figures 7-8 illustrate a second embodiment of the present invention that utilizes a treatment of an epitaxially-grown SiGe layer in accordance with an embodiment of the present invention. The second embodiment illustrated in Figures 7-8 assume a wafer has been prepared in accordance with Figures 1-4, wherein like reference numerals refer to like elements. Accordingly, Figure 7 illustrates wafer 100 of Figure 4 after a reaction layer 710 has been formed over the stress-inducing layer 410.

[0037] The reaction layer 710 comprises a material that, under specific process conditions, reacts with the silicon in the SiGe, but reacts little or not at all with the Ge. In this manner, the reaction layer 710 can be made to react with the silicon and force the germanium lower into the SiGe layer, thereby increasing the Ge concentration. In an embodiment, the reaction layer 710 may comprise Co, Ni, Ti, or the like and may be formed by a CVD process, a PVD process, an ALD process, or the like. Other materials may be used.

[0038] Figure 8 illustrates the wafer 100 of Figure 7 after performing a treatment to cause the reaction layer 710 to react with the stress-inducing layer 410, thereby creating a reacted layer 810, in accordance with an embodiment of the present invention. The treatment may be any treatment that causes reaction layer 710 to react with the Si in the stress-inducing layer 410. As a result of the reaction, the reacted layer 810 is formed and the Ge is pushed lower into the stress-inducing layer 410, as indicated by the directional arrows in Figure 8. It should be noted that the reacted layer 810 is formed in part from a portion of the stress-inducing layer 410. Thus, the volume of the stress-inducing layer 410 decreases while the number of Ge atoms remains

relatively constant, thereby increasing the concentration of the Ge within the stress-inducing layer 410.

[0039] In an embodiment, the treatment may comprise thermal annealing in an ambient comprising, for example, an oxygen-containing gas, a nitrogen-containing gas, a combination thereof, or the like at a temperature between 100°C and 1200°C. Other treatments, such as a rapid-thermal annealing, a laser treatment, UV curing, E-beam curing, implant treatment, a plasma treatment, an ultra-violet radiation treatment, or the like, may also be used.

[0040] Thereafter, the mask 210 may be removed, spacers may be formed, and other subsequent processing may be performed as described above with reference to Figure 6.

[0041] In an embodiment, the resulting stress-inducing layer 410 is a gradient layer having a lattice constant variation between the top and the bottom of the stress-inducing layer 410 of at least 0.2%, wherein the top region exhibits a larger lattice constant than the bottom region. Furthermore, the lattice constant of the substrate in the channel region is less than lower portions of the substrate due to the stress induced by the stress-inducing layer 410.

[0042] Although the present invention and its advantages have been described in detail, it should be understood that various changes, substitutions and alterations can be made herein without departing from the spirit and scope of the invention as defined by the appended claims. Moreover, the scope of the present application is not intended to be limited to the particular embodiments of the process, machine, manufacture, composition of matter, means, methods and steps described in the specification. As one of ordinary skill in the art will readily appreciate from the disclosure of the present invention, processes, machines, manufacture, compositions of matter, means, methods, or steps, presently existing or later to be developed, that perform substantially the same function or achieve substantially the same result as the corresponding

embodiments described herein may be utilized according to the present invention. Accordingly, the appended claims are intended to include within their scope such processes, machines, manufacture, compositions of matter, means, methods, or steps.

WHAT IS CLAIMED IS:

1. A semiconductor device comprising:

a gate electrode formed on a substrate; and

source/drain regions formed on either side of the gate electrode;

wherein the source/drain regions comprise a stress-inducing layer comprising a first semiconductor material and a second semiconductor material, concentration of the second semiconductor material being greater at a top surface than at a bottom surface of the stress-inducing layer.

2. The semiconductor device of claim 1, wherein the first semiconductor material is silicon and the second semiconductor material is germanium.

3. The semiconductor device of claim 1, wherein the concentration of the second semiconductor material is at least 5% greater near the top surface than near the bottom surface of the stress-inducing layer.

4. The semiconductor device of claim 1, wherein the top surface of the stress-inducing layer is raised above a surface of the substrate.

5. The semiconductor device of claim 1, wherein the stress-inducing layer is from about 10 nm to about 300 nm in thickness.

6. The semiconductor device of claim 1, wherein the stress-inducing layer is located in a recess in the substrate.

7. The semiconductor device of claim 1, wherein the stress-inducing layer is an elevated stressor such that the stress-inducing layer is formed on a non-recessed portion of the substrate.

8. The semiconductor device of claim 1, wherein the stress-inducing layer is positioned from about 0 nm to about 300 nm from the gate electrode.

9. The semiconductor device of claim 1, wherein the stress-inducing layer comprises a gradient layer having a varying lattice constant from the top surface to the bottom surface.

10. The semiconductor device of claim 9, wherein the gradient layer has a larger lattice constant near the top surface than the bottom surface.

11. The semiconductor device of claim 9, wherein a first lattice constant near the top surface is greater than 0.2% of a second lattice constant near the bottom surface of the stress-inducing layer.

12. A method of forming a semiconductor device, the method comprising:

forming a gate electrode on a substrate;

forming a stress-inducing layer on either side of the gate electrode, the stress-inducing layer comprising a first semiconductor material and a second semiconductor material; and

performing a treatment in an ambient, the treatment causing the first semiconductor material in the stress-inducing layer to react with the ambient creating a depletion layer in a top portion of the stress-inducing layer, the depletion layer being substantially free of the second semiconductor material.

13. The method of claim 12, wherein the ambient comprises an oxidation-based gas.

14. The method of claim 12, wherein the ambient comprises an oxidation-based gas or a nitridation-based gas.

15. The method of claim 12, wherein the performing the treatment is performed at a temperature from about 100° C to about 1200° C.

16. The method of claim 12, wherein the treatment comprises a thermal treatment, a plasma treatment, an ultra-violet treatment, an implant, or a combination thereof.

17. A method of forming a semiconductor device, the method comprising:

forming a gate electrode on a substrate;

forming a stress-inducing layer on either side of the gate electrode, the stress-inducing layer comprising a first semiconductor material and a second semiconductor material;

forming a reaction layer over the stress-inducing layer; and

performing a treatment, the treatment causing the reaction layer to react with the first semiconductor material in the stress-inducing layer, thereby creating a reacted layer positioned between the reaction layer and the stress-inducing layer, the reacted layer being at least partially composed of material from the stress-inducing layer, the reacted layer being substantially free of the second semiconductor material.

18. The method of claim 17, wherein the treatment comprises a thermal anneal, a rapid-thermal anneal, an ultra-violet treatment, or an E-beam curing treatment.

19. The method of claim 17, wherein the treatment is performed in an ambient comprising oxygen, nitrogen, or a combination thereof.

20. The method of claim 17, wherein the treatment is performed at a temperature from about 100°C to about 1200°C.