

Partial Replacement Silicide Gate

[0001] This application claims priority to provisional application serial number 60/575,764, filed May 28, 2004, entitled "Partial Replacement Silicide Gate" which application is incorporated herein by reference.

TECHNICAL FIELD

[0002] This invention relates generally to semiconductor devices, and more specifically to the formation of CMOS transistors.

BACKGROUND

[0003] Modern day metal-oxide-semiconductor field-effect transistors (MOSFET) typically utilize gates made of polysilicon. One disadvantage in utilizing polysilicon gates for MOSFETs is the depletion effect that affects polysilicon gates. At inversion, a polysilicon gate generally experiences depletion of carriers in the area of the polysilicon adjacent the gate dielectric. The depletion effect reduces the effective gate capacitance of the MOSFET incorporating a polysilicon gate. Ideally, it is desirable that the gate capacitance is high. The higher the gate capacitance, the more charge is accumulated on both sides of the gate capacitor, and therefore more charge is accumulated in the channel. As more charge is accumulated in the channel, the drain-source current becomes higher when the transistor is biased.

[0004] Figure 1 illustrates, in simplified form, a cross-sectional view of a MOSFET that has a polysilicon gate electrode 8. Due to the polysilicon material of gate electrode 8, the depletion effect causes charged carriers to accumulate near interface 4 between gate electrode 8 and the gate dielectric 2. Accordingly, the effective gate capacitance decreases. This is due to an

increase in the effective distance between the charges accumulated on both sides of the gate dielectric 2. The effective distance that separates the charge on both sides of gate dielectric 2 becomes larger due to the depletion of charge, in the polysilicon gate 8, in the vicinity of interface 4. Accordingly, polysilicon gate electrode 8 causes the effective gate capacitance to decrease.

[0005] Silicide metal gates have been proposed as a promising replacement for polysilicon gates for MOS transistors. Figure 2 illustrates a cross-sectional view of a MOSFET with a silicided polysilicon gate. The gate has a layer of silicide 10 with an underlying layer of polysilicon 8. Polysilicon 8 typically exists because of under silicidation. While silicide 10 may help decrease the resistance of the gate, charge is still depleted in the vicinity of interface 4 between polysilicon 8 and gate oxide 2, thereby causing a smaller effective gate capacitance.

[0006] Fully silicide gates were developed to solve the above-mentioned problem. Like pure metals, fully silicided gates eliminate the gate-depletion effect. Fully silicided gates are formed by a two-step process. First, silicide is formed in the source/drain region, and the hard-mask prevents silicide formation on gate. After a liner oxide or nitride deposition, the gate is subject to a CMP process to expose the gate polysilicon. A second silicide is formed on gate resulting in silicide metal gate.

[0007] The conventional formation process of metal silicide gates has some disadvantages. Incomplete gate silicidation and variations in the gate silicide phase are serious challenges. Non-uniform silicidation will impact on device work function, threshold voltage, equivalent oxide thickness, gate leakage, and gate resistance. Figure 3 illustrates an example of the silicidation variation. MOS transistors 11 and 13 are formed in separated regions having different device density. Due to CMP variations, the thickness of the gate electrodes 14 and 16 are different.

During the silicide, the thick difference H_v causes either over silicidation or under silicidation. A small device may be silicided fully, but a large device based on the same condition may not. The gate difference may also come from nMOS and pMOS. For example, using the same process, an nMOS gate has a thickness of about 210Å, and a pMOS has a thickness of about 530Å. The thickness difference comes from the different poly doping type. This poly thickness difference is too high for the silicidation process applied to nMOS and pMOS at the same time. Another silicidation variation occurs at wafer level. After poly gate etch back, there may be about 100Å poly height difference between devices at the wafer center and wafer edge.

[0008] Besides the silicidation variations, the traditional two-step fully silicided gate process is complicated, increasing cost and decreasing yield. Therefore, new method for forming fully silicided gates is needed.

SUMMARY OF THE INVENTION

[0009] The preferred embodiment of the present invention presents a method of forming a fully silicided gate of a CMOS transistor.

[0010] In accordance with one aspect of the present invention, a method of forming a silicide metal gates having uniform silicide thickness is presented. A gate dielectric is formed over a substrate. A silicon-containing layer is formed over the gate dielectric. A dielectric layer is formed over the silicon-containing layer. A top layer is formed over the dielectric layer. The gate dielectric, the silicon-containing layer, the dielectric layer, and the top layer are patterned into a gate stack. A spacer is formed along an edge of the gate stack. The top layer and the dielectric layer are removed. A metal layer is deposited on the silicon-containing layer and silicided. After the formation of the gate, the top surface of the resulting gate electrode is lower than the top edge of the spacer.

[0011] By using the preferred embodiment of the present invention, the thickness of the gate silicide is substantially uniform whether the transistor is an nMOS or a pMOS transistor, in a dense region or a less dense region, at the center of a wafer or at an edge of the wafer. The gate silicidation process can thus be better controlled.

BRIEF DESCRIPTION OF THE DRAWINGS

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

[0013] Figure 1 illustrates a conventional polysilicon gate transistor;

[0014] Figure 2 illustrates a conventional transistor with a partially silicided gate;

[0015] Figure 3 illustrates an example of the gate thickness variation; and

[0016] Figures 4 through 16 illustrate intermediate stages of forming a preferred embodiment of the present invention having a fully silicided-gate structure.

DETAILED DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS

[0017] 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.

[0018] The intermediate stages of a preferred embodiment of the present invention are illustrated in Figures 4 through 16. In Figure 4, substrate 32 is preferably a semiconductor. More preferably, substrate 32 is formed of silicon, either a bulk silicon wafer or a layer of silicon formed on an insulating layer such as the well-known silicon-on-insulator (SOI) structure. Substrate 32 may also be formed of other semiconductor or insulator materials such as strained silicon, strained SOI, silicon-germanium, strained silicon germanium, silicon-germanium on insulator, strained silicon-germanium on insulator, Ge, strained-Ge, germanium-on-insulator (GeOI), strained GeOI, strained semiconductor, compound semiconductor, and multi-layers semiconductor.

[0019] A gate dielectric layer 33 is formed on the substrate 32. In one embodiment, gate dielectric 33 comprises materials having dielectric values (k value) greater than about 9, such as oxide, oxy-nitride, oxygen containing dielectric, nitrogen containing dielectric, and combinations and multi-layers thereof. In other embodiments, it comprises materials having k values greater than about 9, such as HfO_2 , HfSiO_x , HfSiON , ZrO_2 , Al_2O_3 , TiO_2 , Ta_2O_5 , La_2O_3 , CeO_2 , $\text{Bi}_4\text{Si}_2\text{O}_{12}$, WO_3 , Y_2O_3 , LaAlO_3 , $\text{Ba}_{1-x}\text{Sr}_x\text{TiO}_3$ (x less than 1), PbTiO_3 , BaTiO_3 , SrTiO_3 , PbZrO_3 ,

PST, PZN, PZT, PMN, metal-oxide, metal-silicate, metal-nitride, transition metal-oxide, transition metal-silicate, transition metal-nitride, and combinations and multi-layers thereof.

[0020] A layer 34 is deposited on the dielectric 33. In the preferred embodiment, layer 34 is a polysilicon layer. In alternative embodiments, layer 34 contains at least one layer of silicon-containing material and other underlying layer or layers. The silicon-containing layer resides at the top of the layer 34 and is preferably a polysilicon, but also can be amorphous silicon. The layers underlying the silicon-containing material can be a material comprising oxide, oxynitride, silicon nitride or their combinations. For illustration purposes, layer 34 will be referred as polysilicon 34 although it may be a composite layer. Polysilicon 34 has a preferred thickness of about 50Å to about 1500Å, more preferably about 100Å to about 400Å.

[0021] Figure 5 illustrates a gate etch stop layer (GESL) 36 formed on polysilicon 34. In the preferred embodiment, GESL 36 is an oxide. In other embodiments, it is a dielectric comprising polysilicon, SiON, SiN and combinations thereof. The thickness is preferably between about 10Å to about 100Å, more preferably about 30Å to about 60Å. The GESL 36 is used as an etching stop layer in subsequent steps to prevent the underlying polysilicon 34 from being etched. GESL 36 can be formed using methods such as local oxidation or other known methods.

[0022] GESL 36 and polysilicon 34 are then pre-doped, which is symbolized as arrows 37 in Figure 6. The dopant depends on the type of MOS transistors to be formed. For nMOS, the dopant can be antimony, arsenic, or phosphorous. For PMOS, the dopant can be boron, indium, or aluminum. The doping concentration is preferably about $1\text{E}15\text{cm}^{-2}$ to about $9\text{E}17\text{cm}^{-2}$. Doping polysilicon 34 ensures that both nMOS and pMOS have correct work functions when fully silicided.

[0023] Figure 7 illustrates a top layer 38 deposited on the GESL layer 36. In the preferred embodiment, top layer 38 is a polysilicon and will be referred to as top layer 38. In other embodiments, it can be a dielectric layer. However, it will be a different material from the underlying GESL 36 so that a selective etching can be performed to the top layer 38 and stop at GESL 36. The top layer 38 has a preferred thickness of about 100Å to about 1200Å, more preferably about 500Å to about 1000Å.

[0024] The stacked structure formed in previous steps, including the top layer 38, GESL 36, polysilicon-containing layer 34 and gate dielectric 33 is then patterned and etched as shown in Figure 8 into a gate stack. An implantation can be optionally performed to form lightly doped source/drain (LDD) regions 40.

[0025] Figure 9 illustrates a pair of spacers 44 formed along the sidewalls of the gate stack comprising regions 33, 34, 36 and 38. Spacers 44 serve as self-aligning masks for subsequent source/drain formation steps, as described below. The spacers may be formed by well-known methods such as blanket depositing a dielectric layer over the entire region including substrate 32 and top layer 38, then anisotropically etching to remove the dielectric from the horizontal surfaces and leaving spacers 44. Spacers 44 can be oxide, oxy-nitride, nitride or combinations and multi-layers. They preferably have a width of less than about 600Å. It is noted that the materials of the spacers 44 is different from that of top poly 38 so that selective etching can be conducted.

[0026] In the preferred embodiment, source and drain regions 45 are formed by implanting the desired dopant into substrate 32. In other embodiments, the source and drain regions 45 can be formed by forming recesses in substrate 32, and epitaxially growing semiconductor materials in recesses. The methods of forming source and drain regions 45 are well known in the art.

[0027] Figure 10 illustrates a photo resist 50 on the formed structure. Photo resist 50 is masked and developed to expose the top layer 38, as shown in Figure 11. The top layer 38 is removed through selective etching. Since gate spacers 44 have a material different from that of top poly 38, lateral etching of the spacers is prevented. GESL 36 protects the underlying layer 34 from being etched. Photo resist 50 is then completely removed. GESL 36 is selectively etched as shown in Figure 12, exposing the surface of the polysilicon-containing layer 34. In other embodiments, CESL 36 may be etched away before removing photo resist 50.

[0028] Figure 13 illustrates a metal 54 blanket deposited covering the layer 34, spacers 44 and source and drain regions 45. Metal 54 is preferably a transitional metal, such as cobalt, titanium, nickel, tungsten, platinum, zirconium, or the like. Metal 54 can be deposited by chemical vapor deposition (CVD), sputter deposition, or physical vapor deposition (PVD), but other CVD methods, such as atomic layer chemical vapor deposition (ALCVD), metal organic chemical vapor deposition (MOCVD), and plasma enhanced chemical vapor deposition (PECVD) may also be used. The device is then heated, which causes the metal 54 to react with the underlying polysilicon-containing layer 34 and form a silicide 56, as shown in Figure 14. The process is controlled so that the silicide 56 reaches to the bottom of the polysilicon-containing layer 34 and no over silicidation occurs. Since the thickness of the polysilicon-containing layer 34 is uniform throughout a wafer, it is guaranteed that the formed silicide 56 is uniform across the wafer. The un-reacted metal is selectively removed through the use of an etchant that does not attack the silicide 56, spacers 44 and silicon substrate 32, as is well known in the art. The remaining part forms a silicide 56. Silicides 57 are also formed in source and drain regions at the same time.

[0029] Due to the removal of the top poly 38 and GESL 36, a height difference H_d , which is shown in Figure 14, is created between silicide 56 and spacers 44. The height difference H_d is preferably greater than about 200Å. It is preferred that the height of the silicide 56 formed is lower than about 700Å and more preferably lower than about 500Å. It is also preferred that the height of the silicide 56 is between about 5 percent and about 90 percent of the height of the spacers 44. More preferably, the height of the silicide 56 is between about 10 and about 35 percent of the height of the spacers 44.

[0030] A contact etching stop layer (CESL) 58 is then formed covering the silicide 56, gate spacers 44 and silicide 57 as illustrated in Figure 15. It acts as an etch stop layer to protect underlying regions from being etched during the subsequent process steps forming electrical contacts to the source/drain regions and gate electrode 56.

[0031] Also illustrated in Figure 15 is an inter-level dielectric (ILD) 60, also sometimes known as a pre-metal dielectric (PMD) or an inter-metal dielectric (IMD) layer deposited over the surface of the CESL 58. The ILD layer 60 is preferably a phospho-silicate glass (PSG) or other known dielectrics. This ILD layer 60 provides insulation between the transistor and overlying metal lines. Figure 16 illustrates the device after the formation of contact 62 in the ILD dielectric layer 60. A photo-resist material (not shown) may be formed and patterned over the ILD layer 60 in order to form contact opening 62 to gate silicide 56. Note that CESL 58 operates as an etch stop layer during the etching of ILD dielectric layer 60 and thus protects the underlying silicide 56. Next, the exposed CESL 58 in the contact openings are etched, preferably using reactive ion etching (RIE). Because CESL 58 is quite thin relative to ILD dielectric layer 60, process control and end-point detection are more closely controlled, thus limiting the

likelihood of over-etching through the underlying silicide layer 56. At the same time contact 62 is formed, contacts to source and drain regions can also be formed (not shown).

[0032] By using the preferred embodiment of the present invention, the thickness of the silicide 56 is uniform across the device. The thickness of silicide layer is also substantially equal on the gates of different MOS transistors across the chip. To guarantee the formation of substantially uniform silicides 56 across the chip, it's desired the thickness of poly layer 34 is low and uniform prior to the formation of the silicide 56. The formation and remove of layers 36 and 38 creates a thin layer 34 with uniform thickness. This in turn ensures that the silicide thickness difference between transistors of a pMOS or an nMOS, in a isolated area or a dense area, at wafer center or at wafer edge, and in a large pattern or a small pattern are all low and preferably less than about 80Å, more preferably less than about 30Å.

[0033] 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, and 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 1. An integrated circuit comprising:
2 a plurality of first transistors;
3 a plurality of second transistors;
4 wherein each of the respective first transistors and second transistors has a gate
5 comprising a substantially fully silicided gate electrode having a top surface, a gate dielectric and
6 a spacer having a top edge;
7 wherein the gate electrode of each of the respective first and second transistors has a
8 thickness substantially smaller than the height of the respective spacer; and
9 wherein the thickness of the respective gate electrode of the first transistor and the
10 thickness of the respective gate electrode of the second transistor is substantially the same.
- 1 2. The integrated circuit of claim 1 wherein the substrate comprises a material selected from
2 the group consisting essentially of silicon-on-insulator (SOI), Silicon-Germanium-on-Insulator,
3 germanium-on-insulator (GeOI), strained semiconductor, and combinations and multi-layers
4 thereof.
- 1 3. The integrated circuit of claim 1 wherein the gate dielectric comprises a material having a
2 k value of greater than about 9.
- 1 4. The integrated circuit of claim 1 wherein the gate electrode has a height of lower than
2 about 700Å.

- 1 5. The integrated circuit of claim 1 wherein the top surface of the gate electrode is lower
2 than the top edge of the spacer by a difference of greater than about 200Å.
- 1 6. The integrated circuit of claim 1 wherein the height of the gate electrode is about 10% to
2 about 35% of the height of the respective spacer.
- 1 7. The integrated circuit of claim 1 wherein the spacer has a width of less than about 600Å.
- 1 8. The integrated circuit of claim 1 wherein the thickness of the gate electrode of any of the
2 first transistors and the thickness of the gate electrode of any of the second transistors have a
3 difference of smaller than about 80Å.
- 1 9. The integrated circuit of claim 8 wherein the first transistors are nMOS transistors and the
2 second transistors are pMOS transistors.
- 1 10. The integrated circuit of claim 8, the first transistors further comprising a dopant selected
2 from the group consisting essentially of antimony, arsenic, and phosphorous.
- 1 11. The integrated circuit of claim 8, the first transistors further comprising a dopant selected
2 from the group consisting essentially of boron, indium, and aluminum.
- 1 12. The integrated circuit of claim 8 wherein the first transistors are in a dense area and the
2 second transistors are in a less dense area.

1 13. A method of forming a metal silicide gate, the method comprising:
2 forming a gate dielectric over a substrate;
3 forming a silicon-containing layer over the gate dielectric;
4 forming a dielectric layer over the silicon-containing layer;
5 forming a top layer over the dielectric layer;
6 patterning the gate dielectric, the silicon-containing layer, the top layer, and the dielectric
7 layer into a gate stack;
8 forming a spacer, with a top edge, along a sidewall of the gate stack;
9 removing the top layer and the dielectric layer;
10 forming a metal layer on the silicon-containing layer; and
11 siliciding the silicon-containing layer with the metal layer to form a silicided gate
12 electrode.

1 14. The method of claim 13 wherein the silicon-containing layer comprises a material
2 selected from the group consisting essentially of polysilicon and amorphous silicon.

1 15. The method of claim 13 further comprising the step of:
2 doping the silicon-containing layer using a dopant selected from the group consisting
3 essentially of antimony, arsenic, and phosphorous.

1 16. The method of claim 13 further comprising the step of:
2 doping the silicon-containing layer using a dopant selected from the group consisting
3 essentially of boron, indium, and aluminum.

1 17. The method of claim 13 wherein the dielectric layer comprises an oxygen-containing
2 dielectric.

1 18. The method of claim 13 wherein the step of removing the top layer and the dielectric
2 layer comprises the steps of:
3 forming a photo resist covering the substrate; and
4 etching the top layer and the dielectric layer on the gate stack.

1 19. The method of claim 13 wherein the top surface of the gate electrode is lower than the top
2 edge of the spacer and wherein the height difference is greater than about 200Å.

1 20. The method of claim 13 wherein the height of the gate electrode is about 10% to about
2 35% of the height of the spacer.

1 21. The method of claim 13 wherein the step of siliciding the silicon-containing layer further
2 comprising the step of:
3 heating the metal layer to activate interaction between the metal layer and the silicon-
4 containing layer; and
5 removing the un-reacted metal layer.