Description

FLOATING GATE MEMORY CELL WITH CHARGE LEAKAGE PREVENTION

TECHNICAL FIELD

The invention relates to floating gate memory cells with improved reliability and a method for improving the manufacture of floating gate memory cells. More specifically, the invention relates to salicide floating gate memory cells with reduced charge leakage.

BACKGROUND ART

As MOS transistors are increasingly miniaturized, it has become necessary to reduce the sheet resistance of silicon and polysilicon electrodes, especially at the junction between a silicon/polysilicon region and a metal lead via. One method of doing this is to fuse metal with a surface layer of silicon or polysilicon. The fusion of metal and silicon or metal and polysilicon is termed silicide. Of particular interest are processes which allow formation of self-aligned silicide, termed salicide, on the gate, drain and source regions of a transistor simultaneously. It is possible to form salicide on some parts of an IC and not on others. However the construction of salicide requires additional oxide etching steps on all parts of the IC regardless of whether certain parts of the IC do not require silicide. These additional oxide etching steps can degrade the reliability of a floating gate memory cell, even if the memory cell does not receive silicide, as explained below. Before discussing how various salicide processes degrade a floating gate memory cell's reliability, it may be helpful to first explain the basic salicide process as it applies to single gate MOS transistors, and to explain some of the issues regarding the construction of salicide on conventional MOS transistors.

The basic salicide process can be described with reference to the conventional silicide MOS transistor shown in Fig. 1. Transistor 11 has a lightly doped drain, LDD, structure since such drain architecture is often necessary at sub-micron technologies. Although transistor 11 is shown to be an n-channel type, it would be understood by those in the art that such a silicide structure could readily be applied to a p-channel transistor.

Control gate 27 is formed atop a gate oxide layer 20 over a p-type silicon substrate 13. Lightly doped n- source region 15b and lightly doped n- drain region 17b are self-aligned on either side of control gate 27. Oxide sidewall spacers 23 and 25 are then formed on both sides of control gate 27.

Oxide sidewall spacers 23 and 25 serve two functions. First, they serve to self-align the formation of heavily doped n+ source region 15a and heavily doped n+ drain region 17a at a predetermined distance from control gate 27. This predetermined distance is defined by the desired length of the lightly doped n- regions 15b and 17b. The lengths of lightly doped n- regions 15b and 17b are selected such that short-channel effects are mitigated and transistor action of device 11 is enhanced while simultaneously raising the operating voltage of device 11. Secondly, oxide sidewall spacers 23 and 25 serve to define the formation of self-aligned silicide 29 on heavily doped n+ regions 15a and 17a and on control gate 27. Oxide sidewall spacers 23 and 25 also prevent silicide 29b on heavily doped n+ regions 15a and 17a from making contact with control gate 27 or with silicide 29a on control gate 27.

Once source region 15 and drain region 17 are formed, a metal film used for the formation of silicide is deposited over the entire surface of transistor 11. The metal may be a refractory metal such as titanium or a group-VIII metal. After deposition of the selected metal film, the wafer in which transistor 11 is located is

heated. The metal film reacts to the heat by annealing with any exposed silicon and polysilicon to form silicide, but the metal film will not react with exposed oxide. Therefore, a layer of silicide 29b is formed over silicon regions 15a and 17a, and another layer of silicide 29a is formed over polysilicon control gate 27. However, silicide is not formed over oxide sidewall spacers 23 and 25 or over the lightly doped n- source 15b and drain 17b regions, which are protected by oxide sidewall spacers 23 and 25. Any unreacted metal is then selectively removed by using an etchant that will not attack silicide 29, silicon substrate 13 or oxide sidewall spacers 23 and 25.

Use of titanium metal to form titanium silicide, $TiSi_2$, has been noted to have some advantages over other refractory metals since $TiSi_2$ exhibits low resistivity and can reliably form silicide on both single-crystal silicon and polycrystalline silicon through a thermal reaction. Titanium silicide, however, has some drawbacks.

One unfavorable consequence of using titanium metal in the formation of silicide 29 is that titanium metal may inhibit sidewall spacers 23 and 25 from properly isolating silicide 29a on control gate 27 from silicide 29b on source region 15a and drain region 17a. Under certain conditions, it has been found that silicon from source 15, drain 17 and control gate 27 of an MOS transistor will diffuse into a titanium metal film covering sidewall spacers 23 and 25. With reference to Fig. 2, when the wafer is heated, the silicon diffused into the titanium metal film covering sidewall spacers 23 and 25 will form a lateral silicide layer 28 over sidewall spacers 23 and 25. Lateral silicide 28 may grow to electrically short control gate 27 with source region 15a or drain region 17a. This problem is termed bridging.

It has been found that if the titanium metal film is annealed in a nitrogen, N, ambient, the titanium

metal film will absorb a large amount of the nitrogen. This retards the diffusion of silicon into the titanium metal film and thereby mitigates the bridging problem during the annealing process step. As long as sidewall spacers 23 and 25 are of sufficient size, use of a nitrogen atmosphere will retard the diffusion of silicon enough to prevent lateral silicide 28 from bridging across sidewall spacers 23 and 25.

This poses a problem for the micro-miniaturization of conventional MOS switch transistors such as transistor 11. As transistors are further reduced in size, the lightly doped n- regions 15b and 17b of the source and drain, respectively, need to be reduced in length for proper scaling and optimum performance. However, the lengths of the lightly doped n- regions 15b and 17b are defined by the size of sidewall spacers 23 and 25, respectively, and the minimum size of sidewall spacers 23 and 25 is limited by the need to prevent bridging.

With reference to Fig. 3, the minimum size of sidewall spacer 23 and 25 needed to prevent bridging can be much larger than the desired reduced length of the lightly doped region 15b and 17b needed for proper scaling. This can result in transistor 11 having oversized lightly doped -n regions 15b and 17b and less than optimum performance.

U.S. Pat. 5,208,472 to Su et al. teaches a method of addressing this problem. With reference to Fig. 4, Su et al. teach forming oxide sidewall spacers 23 and 25 of transistor 11 in two process steps. During a first process step, a first part, 23a and 25b, of the oxide spacers is formed to a size determined by the optimal size of lightly doped regions 15b and 17b, respectively. During a second process step, a second part, 23b and 25b, of the oxide spacers is formed over the first parts 23a and 25a, respectively, to extend the final size of the combined oxide spacers 23 and 25 to an appropriate size required for preventing bridging.

U.S. Pat. 5,508,212 to Wang et al. teaches a different approach to address the same problem. With reference to Fig. 5, Wang et al. teach implanting nitrogen, N, at a large angle into a titanium metal film, Ti, covering transistor 11. The implanted nitrogen N forms a layer titanium nitride, TiN, reaching into the titanium metal film Ti before the heating step to anneal the titanium metal film Ti and silicon into silicide. Wang et al. explain that due to the high angle of nitrogen implantation, the nitrogen N will reach deeper into the titanium film Ti in areas over oxide spacers 23 and 25 than in areas over source 15, drain 17 and gate 27. This results in a titanium nitride layer TiN that covers the side of oxide spacers 23 and 25, but does not reach the source 15, drain 17 and gate 27 regions. As a result, even less silicon diffuses into the titanium film Ti in the region over oxide spacers 23, 25 and thereby prevents lateral silicide 28 from bridging across the source 15a and drain 17d to control gate 27 while achieving a smaller minimum size for oxide spacers 23, 25. Although this approach reduces the minimum size of oxide spacers 23 and 25 required for preventing bridging, the achieved reduced minimum size may not necessarily be equal to that required for optimal performance. A tradeoff between performance and the prevention of bridging may still be needed.

U.S. Pat. 5,322,809 to Moslehi teaches a different problem afflicting the use of salicide in the micro-miniaturization of conventional MOS transistors. Moslehi explains that as a typical MOS transistor is miniaturized, the source and drain regions of the transistor need to be made more shallow to maintain proper size scaling and performance. For example, MOS transistors having channel lengths less than 0.8 μm require source and drain regions having a depth smaller than 0.25 μm, as explained in Silicon Processing for the VLSI Era, vol. 2 by S. Wolf, page 154. As further explained by Wolf on page 160, special care and added

process steps, such as the use of barrier layers, should be used in the formation of silicide on shallow source/drain junctions which are smaller than 0.2 μm to prevent excess Si in the source/drain regions from being consumed by the silicide.

This is also explained by Moslehi who expounds upon some of the difficulties of forming silicide on shallow source/drain junctions. Moslehi states that because the source and drain regions are shallow, the formation of silicide over the source and drain may consume so much silicon in the source and drain regions as to be detrimental to the transistor. Moslehi also assserts that one cannot simply reduce the depth of the formed silicide with the reduction in the depth of the source and drain regions because the control gate still requires a large amount of silicide formation to reduce its ohmic resistance. Rather than forming the silicide on the control gate in separate process steps from the silicide on the source and drain regions, Moslehi teaches a method of slowing down the formation of silicide over the source and drain regions without affecting the rate of silicide formation over the control gate. It is thereby still possible to form silicide on the source, drain and control gate simultaneously in a salicide process.

To accomplish this, Moslehi proposes that a thin silicide boundary, preferably a thin nitride layer, be laid over the source and drain regions after the formation of the sidewall spacers. The silicide boundary is not laid over the control gate. The silicide boundary is thin enough to slow the formation of silicide over the source and drain regions, but not thick enough to completely inhibit it.

In order to prevent the formation of the silicide boundary over the control gate, an oxide mask is placed over the control gate prior to formation of the sidewall spacers and therefore prior to the laying of the silicide boundary over the source and drain regions.

After the silicide boundary has been formed on the source and drain regions, the oxide mask covering the gate is removed and the entire device is then covered with a refractory metal. Moslehi recommends making the sidewall spacers of the same material as the silicide boundary.

The problems discussed above regarding the construction of a micro-miniature salicide MOS transistor generally do not afflict the construction of a salicide floating gate memory cell which is not optimized for microminiaturization.

With reference to Fig. 6, a floating gate memory cell 31 typically has a control gate 37 stacked on top of a floating gate 35 with interpoly oxide 30 in between, and an additional gate oxide 31 under floating gate 35. Control gate 37 and interpoly oxide 30 typically have similar thicknesses as the control gate 27 and oxide layer 20 of the typical MOS transistor 11 of Figs. 1-5. As a result, the dual gate structure of a floating gate memory cell 31 is much larger than gate structure 27 of the conventional MOS transistors 11 discussed above. This results in floating gate memory cell 31 having oxide sidewall spacers 39 and 38 which are taller and wider than the sidewall spacers 23 and 25 of conventional single gate MOS transistors. Therefore, any lateral silicide growth 33 will generally not extend far enough to make contact and cause bridging between the silicide 36b on the source 32 or silicide 36b on the drain 34 and the silicide 36a on the control gate 37.

Furthermore, since a floating gate memory cell 31 needs to withstand voltages of 2½ to 4 times the main power supply, Vcc, during various phases of operation, they cannot use shallow source and drain junctions nor can they be miniaturized to the extent of conventional MOS transistors, which are designed to withstand at most a Vcc of typically 3V to 5V. As a result, floating gate memory cell 31 does not require a lightly doped drain structure. Therefore, floating gate memory cell 31 does not have the contradicting requirements of large sidewall

spacers to prevent bridging and small sidewall spacers for sizing of a lightly doped region, as discussed by Su et al. and Wang et al. Additionally, their need to withstand 2½ to 4 times Vcc at their electrodes, also forces their source 32 and drain 34 regions to be necessarily much deeper than in a conventional MOS transistor. Therefore, floating gate memory cell 31 does not have shallow source and drain regions and is not subject to the problems afflicting the use of salicide in a conventional micro-miniaturized transistor having shallow source and drain regions, as discussed by Moslehi.

On the other hand, if floating gate memory cells and microminiature transistors are to be constructed on the same device and use common process steps such that the floating gate memory array is not constructed in process steps separate from the microminiature transistors, then the sidewall spacer size typically has to be a compromise for both, or optimized only for the floating gate cells or for the microminiaturized transistors. If microminiaturized transistors and floating gate memory cells are indeed constructed using common process steps and the spacer size is indeed optimized for microminiature transistors, then the spacers on the floating gate cells would be smaller than those shown in Fig. 6, but would still generally be larger than those on the microminiature transistors. Additionally, the floating gate memory cells would still not be subject to the problems afflicting the use of silicide in shallow source and drain regions since they would still need to withstand voltage stresses of 2½ to 4 times Vcc.

Another concern involving the use of silicide in a salicide process is that often times it is necessary to have silicide on some regions of an IC, but not on others. As explained above, silicide tends to reduce the sheet resistance of silicon and polysilicon electrodes, but some circuits require electrodes with high resist-

ances. Such devices include, for example, ESD and latch-up protection circuitry, resistors and I/O circuitry. The conventional method of selectively growing silicide on some circuits of an IC and not on others is to first deposit an oxide layer over the entire IC after all transistors have been constructed, but before starting any salicide processes. A photoresist pattern covering those circuits which should not receive silicide is placed over the IC, and the oxide layer is etched off of all exposed regions which are not covered by the photoresist pattern. The photoresist pattern is then removed, leaving an oxide layer boundary over only those circuits which should not receive silicide. Removal of the oxide layer, however, can impact the integrity of the oxide spacers and introduce structural abnormalities into a floating gate memory cell, which have been identified by the inventors as leading to a reduction in a memory cell's expected life span.

It is an object of the present invention to provide a process for a salicide floating gate memory cell with improved data retention.

It is another object of this invention to provide a salicide floating gate memory cell structure with a reduced failure rate.

It is another object of this invention to provide a floating gate memory cell structure well suited for further miniaturization.

It is yet another object of the present invention to provide a process for selectively growing salicide in some areas of a memory IC and not in others while not degrading the reliability of the memory cells or salicide transistors.

DISCLOSURE OF THE INVENTION

The above objects have been met in a floating gate memory cell which corrects the cause of a previously unknown charge leakage problem identified by the inventors of the present invention.

When designating certain areas of a floating gate memory IC to receive silicide and other areas not to receive silicide, a layer of oxide is first deposited over all active areas of the IC. The oxide layer is then etched off of all areas which are to receive silicide. If the memory cell is not designated to receive silicide, then it remains covered by the oxide layer during the salicide process steps. After the salicide process steps, the oxide layer is removed. Thus, all devices require that they be covered by an oxide mask and all require that the oxide mask be later removed, either before or after the salicide process steps. Removal of the oxide layer can thin out areas of an oxide re-growth, or re-oxide growth, surrounding the floating gate of a memory cell. The re-oxide growth has previously been made to protect the polysilicon from damage during subsequent ion implantations and other process steps. It has been found, however, that the re-oxide growth can develop regions thin enough for charge to leak out of the floating gate of a memory cell even if it is still thick enough to protect the polysilicon from the subsequent process steps.

If the memory cell is designated to receive silicide and therefore has prior art oxide sidewall spacers covering the re-oxide growth, then it has been found that when the layer of oxide is etched off the floating gate memory cell in preparation for the salicide process steps, the memory cell's prior art oxide sidewall spacers can develop areas which are partially etched down to the re-oxide growth, which encases the floating gate. Thus, the re-oxide growth surrounding the floating gate can experience thinning regardless of whether it is covered by prior art oxide sidewall spacers, or not. Although the reduction in the prior art oxide sidewall spacers of the floating gate memory cells does not generally lead to the type of silicide related failures discussed above in relation to a single gate MOS switch transistor, the inventors of the present invention have

found that this reduction in the thickness of the re-oxide growth can lead to a previously unrecognized source of charge leakage from the floating gate.

A control gate and floating gate stack typically have a re-oxide growth covering the sides of the gates. The re-oxide growth encapsulates the floating gate and provides a protective coating to all polysilicon gates during subsequent ion implantations and other process steps. However it has been found that if the re-oxide growth is thinned down below a certain point of about 100 Å, it can provide a path for slow charge leakage out of the floating gate even though it is still thick enough to protect the polysilicon gates from damage during subsequent process steps. Even if the prior art oxide sidewall spacers remain tall enough to prevent bridging between the control gate and the source and drain electrodes of a memory cell, a reduction in the width of the prior art oxide sidewall spacers resulting from the removal of the oxide mask layer prior to, or after the, salicide process steps can cause the above described thinning of the re-oxide growth. This can create a conduction path out of the floating gate if there are any anomalous protrusions from the poly 1 floating gate or any abnormally formed poly layers. The reduction in the width of the prior art sidewall spacers can also cause isolation topography issues. Although the re-oxide growth is not completely etched away and thus still encapsulates the floating gate, the re-oxide growth may develop thinned regions through which charge may tunnel out of the floating gate.

If the non-volatile memory cell is of the electrically erasable type, then it will typically have a thin oxide tunnel region under the floating gate between the source and drain through which charge is moved into and out of the floating gate. The disturbances on the re-oxide growth at the side walls of the floating gate effectively form additional, uncontrollable charge tunnel regions. Consequently, the memory cell can experience

charge loss. This will result in a floating gate memory cell of reduced performance and lower reliability.

Therefore, a non-volatile, floating gate memory cell is susceptible to this previously unrecognized charge leakage problem regardless of whether is receives silicide or not. As long as other transistors, or itself, are subjected to a prior art silicide process, a memory cell's re-oxide growth may experience severe thinning around its floating gate.

A memory array is typically given a projected cycling and speed rating based on results of initial speed tests on its memory cells under stressed conditions. Considering that the thinned regions of the re-oxide growth surrounding the floating gate may not affect a memory cell's initial performance, since the thinned regions provide for a gradual and not catastrophic charge leakage path, a memory cell's initial performance may indicate a high cycling endurance inspite of having charge leakage paths in its re-oxide growth. Nonetheless, the charge leakage will lead to a gradual loss of data and to a premature demise of the floating gate memory array. Thus, a memory cell's initial performance test may no longer be a reliable indicator of a memory's future performance.

Thus, the invention proposes a method and structure for a floating gate memory cell in an IC with selective salicide construction which maintains the integrity of the re-oxide growth surrounding the floating gate regardless of whether the memory cell receives silicide or not.

The thinning of the re-oxide growth occurs around the dual gate structure of a memory cell. But charge leakage out of the memory cell occurs only at thinned out regions of the re-oxide growth on the side walls of the floating gate. This is because a dual gate memory cell stores charge only in its floating gate. It is not necessary for a voltage potential to be applied to the memory cell in order to induce charge leakage.

Charge leakage out of the floating gate is uncontrollable and is driven by an inherent built-in potential in the floating gate resulting from the charge stored within it. Thus to prevent charge leakage, one needs to protect the integrity of the re-oxide growth only at the side walls of the floating gate.

This is accomplished by forming an oxide-protective cover on the sides of the floating gate and over its re-oxide growth. The oxide-protective cover has the property of being resistant to the etchant which is used to remove the oxide mask layer prior to salicide process steps. When the oxide mask layer is laid over the floating gate in preparation for designating some areas of an IC for salicide growth, the oxide-protective cover forms a barrier between the oxide mask layer and the re-oxide growth. When the oxide mask layer is etched off the memory cell, either before or after the salicide processes depending on whether the memory cell is to receive silicide, the oxide protective cover over the re-oxide growth will not be affected by the etchant and thereby maintain the integrity of the re-oxide growth underneath it. Preferably the protective cover is a nitride structure made thick enough to prevent all silicide from penetrating it and reaching the re-oxide growth on the side wall of the floating gate.

If it is desired that silicide also be formed on the memory cell, then the oxide-protective nitride cover may be structured into sidewall spacers for the floating gate memory cell instead of using prior art sidewall spacers made of oxide. In this case, a hydrogen fluoride etchant is used to remove the oxide mask off the memory cell. The hydrogen fluoride etchant is highly selective of nitride and will remove the oxide mask without much attacking the nitride sidewall spacers. A titanium layer, or some other oppropriate metal film, is laid over the memory cell followed by a heat annealing step.

In this manner, it is possible to still grow an oxide mask layer over the entire IC prior to formation of the salicide and then to etch away the oxide mask layer in selected areas without causing thin tunnel regions in the re-oxide growth at the sides of the memory's floating gate. The process of the present invention also prevents a degradation in the width of the nitride sidewall spacers, thereby also protecting the integrity of the lightly doped regions of an LDD MOS switch transistor.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a prior art salicide MOS transistor.

Fig. 2 is a prior art salicide MOS transistor showing bridging of silicide from the control gate to the source and drain regions.

Fig. 3 is a prior art salicide MOS transistor having large oxide spacers to prevent silicide bridging.

Fig. 4 is a prior art salicide MOS transistor with oxide spacers formed in two steps.

Fig. 5 is a prior art salicide MOS transistor with a titanium nitride boundary layer.

Fig. 6 is a prior art salicide floating memory device.

Figs. 7-19 are process steps to form a floating gate memory device in accord with the present invention.

Fig. 20 is an example of floating gate memory cells in accord with the present invention having a non-uniform gate structure.

BEST MODE OF CARRYING OUT THE INVENTION

With reference to Fig. 7, three devices will be constructed in a common substrate 48, which is part of a wafer, not shown. The three devices will be n-type devices, but this is for illustrative purposes only and it would be understood by a person versed in the art that the following process and structure can readily be extended to p-type and CMOS architecture. One device will be a salicide floating gate memory cell and the other two

devices will be n-type enhancement mode, MOS switch transistors with one of the two being a salicide n-type MOS switch transistors and the other not receiving silicide. Although the following process will be demonstrated in the construction of three devices, it would be understood that many more devices could be constructed following the same process steps in a common wafer.

In the following best mode illustration, both of the n-type MOS switch transistors have an LDD structure, but if it were desired that some MOS switch transistors not have an LDD structure, then certain process steps may be omitted and additional masking steps may be necessary. However, these basic process variations are considered to be within the scope of the present invention and within the ability of a person versed in the art. In the following best mode illustration, the floating gate memory cell, which would typically be part of a large array of memory cells, is constructed concurrently with the two n-type, enhancement mode MOS switch transistors. However, if desired, it would be possible to construct the memory cell array separately by taking appropriate masking steps to isolate a memory cell array region of an IC from all periphery regions of the IC. The periphery circuitry of the memory IC would then be constructed after completion of the memory array, or vise-versa.

Reference characters 40 and 50 will identify the two n-type MOS transistors. Transistor 40 will be a salicide transistor and transistor 50 will not receive any silicide. Reference character 60 will identify the floating gate memory cell, which in the preferred embodiment receives silicide.

In the preferred mode, construction of floating gate memory cell 60 is begun before initiating construction of transistors 40 and 50. After substrate 48 has been properly cleaned and the active areas have been defined and isolated, a mask 70 is laid over areas where

transistors 40 and 50 are to be constructed. A protective buffer oxide layer, not shown, may optionally be grown on substrate 48 prior to laying mask 70. An oxide layer 72 is grown on the surface of substrate 48 in a region where non-volatile gate memory cell 60 is to be constructed. If non-volatile memory cell 60 is to be an EPROM cell, then oxide layer 72 would have a preferred thickness of 120 Å to 250 Å, but if non-volatile memory cell 60 is to be an EEPROM cell, then oxide layer 72 would have a preferred thickness of 50 Å to 100 Å. A first polysilicon layer 74 is laid over oxide layer 72. First polysilicon layer 74 will later be patterned to form the floating gate of memory cell 60 and oxide layer 72 will be patterned into the memory cell's gate oxide under the floating gate.

With reference to Fig. 8, mask 70 is removed and a thick oxide layer 76 is grown on top of the areas where devices 40-60 are to be constructed, including over first polysilicon layer 74. Thick oxide layer 76 is made thick enough to prevent charge tunneling, and has a preferred thickness of 120 Å to 250 Å. A second polysilicon layer 78 is then deposited over thick oxide layer 76. A second mask and a polysilicon gate pattern, not shown, configure the gates of devices 40-60 during an etchant step, resulting in the structure shown in Fig. 9.

With reference to Fig. 9, the etchant step configures thick oxide layer 76 and second polysilicon layer 78 to form thick gate oxides 49 and 59, as well as control gates 43 and 53 of transistors 40 and 50, respectively. Control gate 43 has a preferred channel length of 0.2 μm to 1.5 μm. Gate 53 has a similarly preferred channel length of 0.2 μm to 1.5 μm. The same etchant step also forms interpoly oxide 68, control gate 63, gate oxide 61 and floating gate 69 of memory cell 60. Memory cell 60 has a preferred channel length of 0.25 μm to 1.0 μm and control gate 63 forms a primary capacitive connection to floating gate 69.

After formation of the gate structures of devices 40-60, a re-oxidation step forms an oxide re-growth, or re-oxide layer, 73 over the surface of substrate 48 and over the gates of devices 40, 50 and 60, as shown in Fig. 10. Traditionally, re-oxide layer 73 is grown to protect the surfaces of silicon substrate 48 and polysilicon layers 43, 53, 63 and 69 from damage during subsequent process steps. But the inventors have found that if the thickness of re-oxide 73 at the sidewalls of the floating gate 69 fall below a certain point, it can affect the integrity of floating gate 69 even if re-oxide 73 remains thick enough to still protect the surface of floating gate 69 from damage during subsequent process steps, as explained below. A mask layer 71 is then laid over memory cell 60. At this point, lightly doped -n regions 45b/47b and 55b/57b are self-aligned on either side of control gates 43 and 53, respectively. Mask 71 protects memory cell 60 during formation of the lightly doped -n regions.

A first arsenic ion implantation "A" forms lightly doped n- regions 45b and 47b in transistor 40 and lightly doped regions 55b and 57b in transistor 50. Lightly doped regions 45b/47b and 55b/57b have a preferred ion concentration of $10^{15}cm^{-2}$ to $10^{19}cm^{-2}$ and a preferred depth of 0.15 μm to 0.3 μm. Mask 71 is then removed off floating gate memory cell 60.

With reference to Fig. 11, a new masking layer 75 is placed over transistors 40 and 50, and a second arsenic ion implantation "B" forms heavily doped n+ source 65 and drain 67 regions in floating gate memory cell 60 to a preferred ion concentration of 10^{20} cm^{-2} to 10^{21} cm^{-2} and a preferred depth of 0.3 μm to 0.6 μm.

Masking layer 75 is then removed, resulting in the structure of Fig. 12. Fig. 12 shows transistors 40 and 50 with partially constructed source regions 45b and 55b, respectively, and partially constructed drain regions 47b and 57b, respectively. Re-oxide layer 73 still covers devices 40-60. Construction of sidewall

spacers in preparation of salicide process steps is now initiated.

With reference to Fig. 13, a nitride layer 77 is laid over MOS transistors 40 and 50 and over floating gate memory cell 60 by mean of an LPCVD reactant or CVD reactant and etchant step. Use of nitride in the construction of the sidewall spacer for floating gate memory 60 prevents charge leakage off of floating gate 69 due to a previously unrecognized problem identified by the inventors of the present invention, as explained below. Nitride layer 77 is used to form an oxide-protecting cover, or coating, over re-oxide layers 73 of devices 40-60. If memory cell 60 was designated not to receive silicide, nitride layer 77 would still be used to form the oxide-protecting cover over re-oxide layer 73. The formed nitride cover would protect re-oxide layer 73 from thinning due to the removal of a later oxide masking layer shown below. Thus, a nitride, oxide-protecting cover is formed over the re-oxide layer 73 at the sides of floating gate 69 regardless of whether the memory cell is subjected to subsequent salicide process steps or not.

With reference to Fig. 14, nitride layer 77 is etched down to the substrate to form sidewall spacers 44/46 for transistor 40, sidewall spacers 54/56 for transistor 50 and sidewall spacers 64/66 for memory cell 60. Sidewall spacers serve as an oxide-protecting cover over re-oxide layer 73 at the sides of floating gate 69. Formation of the nitride sidewall spacers results in the removal of the re-oxide off the tops of control gates 43, 53 and 63, but this causes no structural damage to the control gates since later passivation layers provide protection for the top of the control gates. Removal of re-oxide layer 73 from the tops of control gates 43, 53 and 63 also does not affect the reliability of devices 40-60 since the control gates do not store charge and thus are not subject to the charge leaking problem affecting floating gate 69.

With reference to Fig. 14, a resist mask 78 is placed over memory cell 60 followed by a third heavily doped +n arsenic ion implantation "C" applied to all device 40-60. In transistor 40, sidewall spacers 44 and 46 define the length of lightly doped -n regions 45b and 47b, respectively. Ion implant "C" forms heavily doped +n source region 45a and heavily doped +n drain region 47a to finish the lightly doped drain, LDD, construction of transistor 40. Similarly, in transistor 50, sidewall spacers 54 and 56 define the length of lightly doped -n source region 55b and lightly doped -n drain region 57b, while allowing formation of self-aligned, heavily doped n+ source region 55a and heavily doped +n drain region 57a. Heavily doped regions 45a, 47a, 55a and 57a have a preferred ion concentration of 10^{20} cm^{-2} to 10^{21} cm^{-2} and a preferred depth of 0.2 μm to 0.4 μm.

In Fig. 15, an oxide layer 79 is deposited over devices 40-60. Oxide layer 79 will be removed off of devices which need a silicide layer and remain on devices which do not need a silicide layer. In the present illustrative case, devices 40 and 60 will undergo salicide processes while transistor 50 will receive no silicide. Therefore, a fifth mask 81 is laid over transistor 50 and not over devices 40 and 60. Oxide layer 79 is then etched away off of devices 40 and 60 using hydrogen fluoride, HF, an etchant which attacks oxide and is highly selective of nitride. If it were desired that memory cell not receive silicide, then mask layer 81 would also be placed over memory cell 60.

If the sidewall spacers were made of oxide, as is typical in the prior art, it would be difficult to control the etching of oxide layer 79 without also affecting the oxide sidewall spacers. Figs. 16 and 17 illustrate some of the consequences of the removal of oxide layer 79 off of devices 40 and 60 if the sidewall spacers of the memory cell 60 and transistor 40 were constructed of oxide. Fig. 16 shows prior art transistors 40′ and 50′, and Fig. 17 shows three examples of possible

deformations of memory cell 60' which the inventors have identified as sources of charge leakage off of floating gate 69'. For the sake of brevity and clarity, all elements in Figs. 16 and 17 similar to those of Fig. 15 are identified with similar reference characters plus the addition of a prime symbol.

In Figs. 16a and 16b, devices 40' and 50' are representative of devices 40 and 50 of Fig. 15, respectively. Since device 50' is not intended to receive salicide, protective oxide layer 79' is shown to remain on top of device 50' in Fig. 16a while titanium metal layer 83' is laid over all devices, including device 40' and 60'. Titanium metal layer 83' will form a self-aligned silicide layer on all exposed silicon and polysilicon areas following a heat annealing step, as explained above. Since all silicon surfaces of device 50' are protected under oxide mask 79', no silicide will form on its source 55', drain 57' or control gate 53'. Device 40', however, is intended to receive silicide, and Fig. 16a shows that protective oxide mask 79' is etched off prior to the laying of titanium metal layer 83'.

With reference to Fig. 16a, after removal of oxide layer 79, a titanium metal film 83' is laid over all devices. If oxide sidewall spacers 44' and 46' are reduced in size during the removal of oxide layer 79, two problems may result. The first is the formation of lateral silicide bridging as explained above. But assuming that oxide sidewall spacers 44' and 46 are not reduced in height enough for lateral silicide bridging between gate 43' and source 45a' and drain 47a' regions, the inventors have identified an additional source of performance degradation not related to the height of oxide sidewall spacers 44' and 46', but rather due to a reduction in the width of oxide sidewall spacers 44' and 46'.

The width reduction of oxide sidewall spacers 44' and 46' causes a degradation in the lightly doped drain structure of transistor 40'. If the width of oxide

sidewall spacers 44' and 46' is reduced, they will pull away by an amount "L" from their respective boundary abutting the lightly doped -n drain region 47b' and the heavily doped +n drain region 47a' or the lightly doped -n source region 45b' and the heavily doped +n source region 45a'.

After a heat annealing step, the titanium metal film 40' reacts with the exposed silicon of source 45', drain 47' and control gate 43'. As explained above, some of the silicon from control gate 43', source 45' and drain 47 diffuse over oxide sidewall spacers 44' and 46' during the heat annealing treatment resulting in the formation of partial vertical silicide over oxide spacers 44' and 46'. After the formation of silicide, all excess titanium metal is removed resulting in the structure of Fig. 16b.

With reference to Fig. 16b, prior art oxide sidewall spacers 44' and 46' are shown to be tall enough to prevent vertical silicide 95' from causing any bridging problems, and silicide 93b' on the source 45a' and drain 47a' regions is shown not to be deep enough to cause any problems, as discussed above in reference to the identified prior art. However, the width of the prior art sidewall spacers 44' and 46' has been reduced by an amount L. Silicide 93b' is formed on the exposed surface of the lightly doped n- drain 47b' and lightly doped n- source 45b' regions. This causes partial shunting of the lightly doped n- regions 45b' and 47b' with their respective heavily doped n+ regions 45a' and 47a'. Consequently, the resistance, and therefore the effective length, of the lightly doped regions 45b' and 47b' is reduced. Since the lengths of the lightly doped n- regions 45b' and 47b' enhance the transistor action of device 40' as well as raise the operating source-to-drain, Vds, voltage of the device, a reduction in the effective length of lightly doped regions 45b' and 47b results in a transistor having a lower breakdown voltage, reduced performance and lower reliability.

The inventors have also identified a source of charge loss on prior art floating gate memory cells due to a reduction in the width of prior art oxide sidewall spacers. Since single gate MOS switch transistors such as devices 40' and 50' of Fig. 16a are not required to store charge on their control gates 43' and 53', respectively, this charge loss problem does not afflict single gate switch transistors and was therefore not previously identified in the art of MOS switch transistor architecture.

The problems of using oxide in the formation of sidewall spacers in a salicide process are fundamentally different in floating gate memory cells than in single gate, enhancement mode, switch transistors. As explained above, floating gate memory cells are generally more resistant to silicide bridging over sidewall spacers than single gate enhancement mode transistors due to their use of larger sidewall spacer, and floating gate memory cells are more resistant to the depth of silicide growth over source and drain region than single gate switch transistor due to the use of deeper source and drain regions to resist high voltages, which are not applied to signal gate enhancement mode transistors. Also, since floating gate transistors generally do not require a lightly doped drain structure, they do not suffer from the shunting of the lightly doped n- regions of a source or drain to its respective heavily doped n+ region. Thus, a floating gate memory cell will generally not experience the structural degradation associated with a single gate enhancement mode transistor in a salicide process.

Similarly, the problem afflicting floating gate memory cells in a salicide process identified by the inventors is inherently absent in the performance and structure of single gate enhancement mode transistors. As is known in the art, floating gate memory cells store information in the form of trapped charge in their floating gate. The inventors have identified a source of charge leakage off of a memory cell's floating gate,

which can lead to premature cell failure and an overall degradation in a memory array's performance. The source of charge leakage is due to various memory cell deformations which can lead to a thinning of the re-oxide layer surrounding the floating gate. Single gate enhancement mode transistors do not require any such trapping of charge for proper operation. In fact, single gate enhancement mode transistors are designed to avoid charge trapping in their control gate since such trapping would alter their behavior characteristic and introduce structural defects into their gate oxide which would reduce their life expectancy.

Figs. 17A-17C illustrate three examples of possible memory cell deformations which can lead to a thinning of re-oxide layer 73' surrounding floating gate 69' and to charge leakage due to the use of prior art oxide sidewall spacers 64' and 66'. The problem afflicting floating gate memory cells does not result from the use of silicide directly. Rather, the problem is a result of the process steps required to form silicide. Therefore, Figs. 17A-17C illustrate the source of charge leakage identified by the inventors without making any reference to the location of silicide itself.

With reference to Fig. 17A, 60a' depicts a prior art floating gate memory cell at the process step last depicted in Fig. 15 where protective oxide layer 79' has been laid over all cells in a chip and certain areas of the chip are being designated to receive silicide. Prior art floating gate memory cell 60b' depicts the state of the memory cell after the removal of protective oxide layer 79' in preparation for the laying of a titanium metal film and the formation of silicide.

A memory array comprises hundreds of thousands, or millions, of memory cells. It is difficult to create a memory array consisting of only perfectly formed memory cells. A percentage of the memory cells within the memory array will have structural abnormalities. Fig. 17a depicts memory cell 60a' with an abnormally shaped

floating gate 69'. Re-oxide 73' is grown on exposed silicon areas. Since re-oxide 73' is formed before to the formation of the prior art oxide sidewall spacer 64' and 66', re-oxide 73' is shown to surround control gate 63', floating gate 69' and an area under prior art oxide sidewall spacers 64' and 66'.

Cell 60b' depicts the condition of the prior art memory cell after the removal of protective oxide layer 79'. Since the prior art sidewall spacers 64' and 66' are likewise made of oxide, they get partially etched during removal of protective oxide layer 79'. Prior art oxide spacers 64' and 66' are shown to have much of their original height, but to have lost much of their width. Specifically, oxide spacer 66' has narrowed down to the point where it has eaten away part of the re-oxide layer 73' at the side of drain 67'. Arrow 80' identifies a charge leakage region of re-oxide layer 73'. Charge leakage region 80' is an area where re-oxide layer 73' is thinned down to a point where it is comparable or thinner than thin oxide tunneling region 61'. In effect, the thickness of re-oxide layer 73' in the area of charge leakage region 80' is reduced below 100 Å. As a result, charge which may have been moved into floating gate 69' through thin oxide tunneling region 61' can now leak or tunnel out through charge leakage region 80'. In the present embodiment, memory cell 60' is an EEPROM and therefore has a gate oxide 61' of less than 100 Å, but if memory cell 60' were an EPROM then gate oxide 61' would necessarily have a thickness greater than 120 Å. Thus, if memory cell 60' were an EPROM, then the re-oxide layer 73' in the area of charge leakage region 80' would be thinner than gate oxide 61'.

A user does not have control over charge escaping through charge leakage region 80'. When charge is stored in floating gate 69', memory cell 60b' develops a built-in potential due to the charge stored in floating gate 69'. This built-in potential drives memory cell 60b's charge loss mechanism. Charge loss through charge

leakage region 80' is gradual and not immediately noticeable. Furthermore, charge which leaks out of memory cell 60b' is typically attracted to the substrate, which is coupled to ground. This charge leakage typically does not manifest itself in any detectable circuit logic errors, but does lead to premature loss of data. Thus, this memory cell failure is independent of the presence of silicide and due only to the extreme thinning of re-oxide layer 73', which leads to the formation of charge leakage regions 80'. Therefore, floating gate memory cell 60b' may perform well during initial testing due to no silicide-related failures existing, but charge leakage region 80' will cause floating gate 69' to slowly leak charge during normal use and result in premature loss of data and a memory cell of lower endurance. Since a memory cell's endurance rating is based on initial test results and this charge leakage problem does not manifest itself during initial performance testing, a memory cell's initial performance may no longer be a reliable way of gauging its future performance.

Fig. 17B illustrates a second structural deviation which can result in the formation of a charge leakage region. With reference to Fig. 17b, memory cell 60a' shows control gate 63' having a smaller length than floating gate 69'. Again, re-oxide layer 73' surrounds both control gate 63' and floating gate 69' as well as a region under prior art oxide spacers 64' and 66'. During the removal of protective oxide layer 79', part of oxide sidewall spacers 64' and 66' are also etched away resulting in the structure of 60b'. Due to the smaller length of control gate 63', prior art oxide spacers 64' and 66' end up with a cascade structure illustrated by memory cell 60b'. It is possible for the prior art oxide sidewall spacers of some cells to be eaten more than others such that, for example, in memory cell 60b' oxide sidewall spacers 66' is eaten down to the point where it degrades re-oxide layer 73'. This forms a thinning of

the re-oxide layer 73' adjacent floating gate 69' resulting in the formation of charge leakage region 80'. Charge can then leak out of floating gate 69' through charge leakage region 80'.

With reference to Fig. 17C, a third example of a memory cell structural deviation of a memory cell which can lead to the formation of a charge leakage region is shown. Control gate 63' of memory cell 60a' is shown to have a base of length comparable to that of floating gate 69', but control gate 63' is also shown to be tapered and have a smaller length at its top than at its base. Nonetheless, re-oxide layer 73' still surrounds control gate 63' and floating gate 69', as well as regions under prior art oxide spacers 64' and 66'. However, due to the tapering of control gate 63', even more of sidewall spacers 64' and 66' is etched away during removal of protective oxide layer 79', as demonstrated by memory cell 60b'. In memory cell 60b', oxide sidewall spacer 64' is completely etched off control gate 63' and has eaten away at re-oxide layer 73' at a corner of floating gate 69' thus forming charge leakage region 80'.

It has been found that the use of nitride in the formation of sidewall spacers in floating gate cells allows a higher control of protective oxide layer 79 and thus a better control of the re-oxide layer surrounding the floating gate. By using hydrogen fluoride, HF, an etchant highly selective of nitride, it is possible to achieve a memory cell having much reduced dimensions while maintaining a high reliability in spite of lower control of memory cell structural formation at the lower dimensions. By avoiding the etching of the sidewall spacers, one can avoid over thinning of the re-oxide layer surrounding the floating gate and thereby preventing this hereunto unrecognized charge leakage problem. If nitride sidewall spacers are used, as proposed in the present invention, the problems highlighted in Figs. 16 and 17 are avoided, and following the removal of protective oxide layer 70 in Fig. 15 by using

a hydrogen fluoride etchant, one would proceed to the process step illustrated in Fig. 18.

With reference to Fig. 18, a metal film 83, preferably titanium, for reacting with exposed silicon and polysilicon and forming silicide is deposited over all devices 40-60. Transistor 50 still has oxide layer 79 covering it, and therefore does not come in contact, or react with, titanium film 83 and does not form silicide. However, oxide layer 79 has been etched off of devices 40 and 60 and titanium film 83 is in direct contact with their respective source, drain and control gate regions. Sidewall spacers 44 and 46 of transistor 40 and sidewall spacers 64 and 66 of floating gate memory cell 60 show minimal reduction in size due to the use of nitride in their construction and the nitride selective etchant used in the removal of oxide layer 79. Therefore, lightly doped -n regions 45b and 47b of transistor 40 likewise show no reduction in size. Similarly, re-oxide layer 73 at the sides of floating gate 69 experiences no thinning and does not develop any charge leakage regions. A heat annealing step of preferably 600°C to 800°C is applied for a period of 10 sec to 60 sec in a preferred nitrogen ambient. This causes titanium film 83 to react with the exposed silicon and polysilicon of devices 40 and 60, but oxide layer 79 prevent titanium film 83 from reacting with transistor 50. After annealing, any unreacted titanium 83 is removed. Oxide mask 79 is then removed by applying a hydrogen fluoride etchant to devices 40-60, and the surface of the wafer, i.e. substrate 48, is then cleaned using ammonium hydroxide resulting in the structure of Fig. 19.

With reference to Fig. 19, floating gate memory cell 60 is shown to have silicide growth 99 on its source 65, drain 67 and control gate 63. Likewise, MOS Transistor 40 is shown to have silicide 97 on its source 45a, drain 47a and control gate 43.

Fig. 19 also shows the formation of vertical silicide 103 on memory cell 60 over nitride sidewall spacers 64 and 66 in addition to the desired salicide 99. Vertical silicide 103 is caused by silicon from source region 65, drain region 67 and control gate 63 partially diffusing into titanium film 83 of Fig 18 over sidewall spacers 64 and 66. Vertical silicide 103 is too short to cause any bridging errors. More importantly, nitride sidewall spacers 64 and 66 serve as oxide-protecting covers over re-oxide layer 73 which prevent thinning of re-oxide layer 73 and thereby prevent the creation of charge leakage regions. In transistor 40, sidewall spacers 44 and 46 suffer minimal reductions in width, and thus lightly doped -n source region 45b and lightly-doped drain region 47b of transistor 40 suffered no reduction in effective length. Transistor 40 also shows some lateral silicide 101 over its nitride sidewall spacers, but sidewall spacers 44 and 46 are tall enough to prevent bridging. Similarly, nitride sidewall spacers 54 of device 50 remain relatively unaffected by the etching away of oxide mask layer 79 in Fig. 18.

In Fig. 19, memory cell 60 is shown to receive silicide 99, and its gates 63 and 69 are not misshaped. But as explained above, it is not necessary for a memory cell to be subjected to silicide process steps in order for the above described charge leakage regions to form. Fig. 20 depicts three examples of memory cells with malformed gate stack structures and which do not receive silicide and thus remain covered by oxide mask layer 79 of Fig. 15 during all salicide process steps. The structural deviations shown in Fig. 20 are similar to those depicted in Fig. 17 above.

With reference to Fig. 20A, memory cell 60 is shown to have a floating gate 69 with a protruding end. When oxide mask layer 79 is removed at the completion of all salicide process steps, as shown in Figs. 18-19, nitride layers 64 and 66 serve to protect re-oxide layer 73 and thus prevent the formation of charge leakage

regions. Similarly, Fig. 20B shows control gate 63 having a smaller length than floating gate 69, and Fig. 20C shows that control gate 63 is tapered toward the top. These structural deviations could previously have led to the formation of charge leakage regions within re-oxide layer 73 when oxide mask layer 79 is removed, as shown in Figs. 17A-17C. However, nitride layers 64 and 66 preserve the integrity of re-oxide layer 73 during removal of oxide mask layer 79 if mask layer 79 is removed with an etchant selective of nitride, such as hydrogen fluoride.

Therefore, use of an oxide-protective cover, such as nitride, over a re-oxide layer at the sides of a floating gate prior to the laying of an oxide mask in preparation of salicide process steps can greatly reduce or prevent the formation of uncontrollable charge leakage regions adjacent the floating gate. This protection is further increased by using an oxide etchant which is highly selective of nitride, such as hydrogen fluoride. As a result, the memory cell will have a higher reliability.

Claims

1. An integrated circuit memory including a floating gate memory cell with reduced charge leakage comprising:

a floating gate over a gate oxide on a substrate of first conductivity type, the length of said floating gate being determined by a first pair of walls defining the sides of said floating gate;

an interpoly oxide over said floating gate;

a control gate over said interpoly oxide, the length of said control gate being determined by a second pair of walls defining the sides of said control gate,

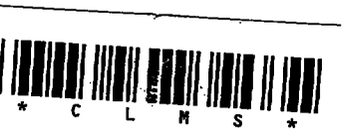
a source region of second conductivity type located proximate to a first wall of said second pair of walls of said control gate;

a drain region of second conductivity type located proximate to a second wall of side second pair of walls of said control gate, said second wall located opposite said first wall;

an oxide re-growth covering said first pair of walls of said floating gate, said oxide re-growth having a thickness sufficient for substantially preventing charge tunneling through said oxide re-growth, said oxide re-growth further being characterized by a susceptibility to a predetermined etchant; and

an insulating coating over at least one wall of said first pair of walls and covering said oxide re-growth, said insulating coating being characterized by a resistance to said predetermined etchant.

2. The floating gate memory cell of claim 1 wherein at least one wall of said first and second pair of walls has a non-uniformed profile.



3. The floating gate memory cell of claim 2 wherein said wall having a non-uniformed profile has one of an output protruding, inward receding, and inclining profile.

4. The floating gate memory cell of claim 1 wherein said control gate and floating gate have different lengths.

5. The floating gate memory cell of claim 4 wherein said control gate has a shorter length than said floating gate and said oxide re-growth covers at least part of said sides of said control gate, said insulating coating forming a cascade configuration on said sides of said control gate and said floating gate.

6. The floating gate memory cell of claim 1 wherein said insulative coating consists of nitride and said predetermined etchant is hydrogen fluoride.

7. The floating gate memory cell of claim 1 wherein said floating gate memory cell is an EPROM cell, said gate oxide has a thickness of at least 120Å, and said oxide re-growth has a thickness at least equal to said gate oxide.

8. The floating gate memory cell of claim 1 wherein said floating gate memory cell is one of an EEPROM and Flash cell, said gate oxide has a thickness of at most 100Å, and said oxide re-growth has a thickness greater than said gate oxide.

9. The floating gate memory cell of claim 1 wherein the entirety of said drain region has a substantially uniform doping level.



10. The floating gate memory cell of claim 1 further having silicide on at least one of said source region, drain region and control gate, said insulative coating further having a thickness sufficient to completely impede silicide from forming under it.

11. The floating gate memory cell of claim 18 wherein said source and drain regions are formed to a depth greater than 0.3μm and said floating gate is made to have a length greater than 0.25μm.

12. The floating gate memory cell of claim 1 wherein said floating gate memory cell is a non-silicide floating gate memory cell.

13. The integrated circuit memory of claim 1 further including an MOS transistor comprising:

a second control gate on top a second gate oxide on an area of said substrate;

a second source region proximate to said second control gate and a second drain region proximate to said second control gate, said second source and drain regions being of said second conductivity type and at least one of said source and drain regions having an LDD architecture;

a sidewall spacer defining the length of the lightly doped region of said LDD architecture, said sidewall spacer being constructed of the same material as said insulating coating.

