

Description

Title of Invention: SEMICONDUCTOR DEVICE, AND PRODUCTION METHOD THEREFOR

Technical Field

[0001] The present invention relates to a semiconductor device and a production method therefor.

Background Art

[0002] A typical semiconductor device includes a column-shaped contact provided to electrically connect a gate electrode on a substrate to a wiring layer disposed above the gate electrode (see Non-Patent Literature 1, for example). A semiconductor device provided with a plurality of contacts may include a memory gate structure and a select gate structure. In the memory gate structure, for example, a lower gate insulating film, a charge storage layer, an upper gate insulating film, and a memory gate electrode are stacked in this order. The select gate structure is provided on a sidewall of the memory gate structure through a sidewall spacer. The memory gate structure and the select gate structure are provided on an active region (substrate surface). The contacts are provided on the memory gate structure and the select gate structure.

[0003] For example, in such a semiconductor device, voltage is applied from various wiring layers to sites such as the memory gate electrode and a select gate electrode of the select gate structure through the contacts. This voltage application causes a quantum tunneling effect due to a voltage difference between the surface of the substrate and this memory gate electrode G100. Charge is injected into a charge storage layer EC by the quantum tunneling effect.

[0004] The select gate structure, which is provided on the sidewall of the memory gate structure through the sidewall spacer, allows control of the select gate electrode separately from the memory gate electrode through voltage application from a contact installation portion to the select gate electrode.

[0005] For example, as illustrated in FIG. 13, in such a semiconductor device 100, a contact installation portion 102 integrated with the select gate electrode (not illustrated) is provided on an element isolation layer 101 adjacent to the active region (not illustrated). In the semiconductor device 100, the charge storage layer EC, an upper gate insulating film 23b, and the memory gate electrode G100 of the memory gate structure extend over the element isolation layer 101. The contact installation portion 102 is disposed on sidewalls of the charge storage layer EC, the upper gate insulating film 23b, and the memory gate electrode G100 through a

sidewall spacer 105. Sites such as the memory gate electrode G100 and the contact installation portion 102 are covered by an interlayer insulating layer 120. An upper wiring layer 112 is provided on another interlayer insulating layer 121 above the interlayer insulating layer 120.

[0006] The contact installation portion 102 includes a contact C100 vertically provided on a flat contact installation surface 102c, and is electrically connected to the upper wiring layer 112 through the contact C100. The contact installation portion 102 applies, to the select gate electrode in the active region, voltage applied through the upper wiring layer 112.

[0007] The contact C100 of the semiconductor device 100 electrically connects the contact installation portion 102 to the upper wiring layer 112. In addition, a contact C101 electrically connects an impurity diffusion region (not illustrated) formed in, for example, an active region (not illustrated) to a wiring layer 113 above.

[0008] Typically, in the semiconductor device 100, another interlayer insulating layer 123 is disposed above the interlayer insulating layer 121 including the wiring layers 112 and 113. The interlayer insulating layer 123 includes another wiring layer 114. In the semiconductor device 100, a contact C102 electrically connects the wiring layers 113 and 114. For example, voltage applied

to the topmost wiring layer 114 is then applied to an impurity diffusion layer on the surface of the substrate through the contact C102, the wiring layer 113, and the contact C101 in this order.

Citation List

Non-Patent Literature

[0009]

Non-Patent Literature 1: "Semiconductor production: Renesas Electronics", [online], Searched through the Internet on October 8, 2014 (URL: http://japan.renesas.com/company_info/fab/line/line12.html)

Summary of Invention

Technical Problem

[0010] The select gate electrode (not illustrated) adjacent to the memory gate electrode G100 through the sidewall spacer 105, and the contact installation portion 102 integrated with the select gate electrode are produced in the following manner. First, the memory gate structure covered by the sidewall spacer 105 is formed on the active region. This process involves formation of the charge storage layer EC, the upper gate insulating film 23b, and the memory gate electrode G100, which are covered by the sidewall spacer 105, on the element isolation layer 101.

[0011] Subsequently, a layer-shaped conductive layer is formed on the entire surfaces of the active region and the element isolation layer 101. Then, a resist is formed in a region of the element isolation layer 101 where the contact installation portion 102 is to be formed. Thereafter, the conductive layer is etched back to form, on the active region, a sidewall-shaped select gate electrode along the sidewall spacer 105. At this time, the contact installation portion 102, which is connected to the select gate electrode, is formed on the element isolation layer 101 while the conductive layer remains intact in the region where the resist is formed.

[0012] The contact installation portion 102 formed in this manner includes a base 102a provided with the flat contact installation surface 102c on which the contact C100 can be vertically provided. In addition, the contact installation portion 102 includes a mounting portion 102b extending over a top portion of the memory gate electrode G100 from the base 102a. Accordingly, in the semiconductor device 100, the interlayer insulating layer 120, on which the memory gate electrode G100 and the contact installation portion 102 are disposed, inevitably has the larger thickness as a result of the formation of the mounting portion 102b protruding upward from the top portion of the memory gate electrode G100.

[0013] With this configuration, in the conventional semiconductor device 100, the contact C101, through which

a substrate surface in a memory well is connected to the upper wiring layer 113, has the larger height in accordance with the increase in the thickness of the interlayer insulating layer 120. Accordingly, the contact C101 has a larger aspect ratio (contact height/contact diameter), which leads to an increased contact resistance value.

[0014] When the thickness of the interlayer insulating layer 120 is reduced to achieve a reduced aspect ratio and to avoid the increase in the contact resistance value of the contact C101, the distance between a top portion of the contact installation portion 102 and each of the upper wiring layers 112 and 113 is reduced. As a result, contact failure potentially occurs between the contact installation portion 102 and the wiring layer 113 above, to which different voltages are applied.

[0015] The present invention has been made under such circumstances and aims at providing a semiconductor device that avoids increase in a contact resistance value and also avoids contact failure with a wiring layer, and a production method for the semiconductor device.

Solution to Problem

[0016] In order to solve the above-described problems, a semiconductor device of the present invention includes: a memory gate structure including a lower gate insulating film, a charge storage layer, an upper gate insulating

film, and a memory gate electrode stacked in this order; a contact installation structure including at least the charge storage layer, the upper gate insulating film, and a separated memory gate electrode made of a layer same as a layer of the memory gate electrode, and electrically insulated from the memory gate structure; a sidewall-shaped select gate electrode disposed on each of a sidewall of the memory gate structure through a sidewall spacer and a sidewall of the contact installation structure through the sidewall spacer, the sidewall-shaped select gate electrode being disposed continuously between the memory gate structure and the contact installation structure; and a contact vertically provided and extending over the sidewall spacer and the sidewall-shaped select gate electrode from a top portion of the contact installation structure.

[0017] The semiconductor device of the present invention is the above-described semiconductor device and the sidewall-shaped select gate electrode includes a first select gate electrode and a second select gate electrode. The first select gate electrode has a sidewall shape and is disposed along the sidewall spacer of one of the sidewalls of the memory gate electrode, and the second select gate electrode has a sidewall shape and is disposed along the sidewall spacer on the other sidewall of the memory gate electrode, and the first select gate electrode is electrically insulated from the second

select gate electrode. The memory gate structure has a linear shape. The contact installation structures include a first contact installation structure disposed near one end of the memory gate structure in a longitudinal direction of the memory gate structure, and a second contact installation structure disposed near the other end of the memory gate structure in the longitudinal direction. The contacts include a first contact vertically provided and extending over the sidewall spacer and the first select gate electrode from a top portion of the first contact installation structure, and a second contact vertically provided and extending over the sidewall spacer and the second select gate electrode from a top portion of the second contact installation structure.

[0018] Another method for producing a semiconductor device of the present invention includes: a contact installation structure formation step of stacking layers of a lower gate insulating film, a charge storage layer, an upper gate insulating film, and a memory gate electrode in this order on a substrate, and then performing patterning to form a memory gate structure and a contact installation structure, the memory gate structure including the lower gate insulating film, the charge storage layer, the upper gate insulating film, and the memory gate electrode stacked in this order, the contact installation structure including at least the

charge storage layer, the upper gate insulating film, and a separated memory gate electrode which are made of layers same as the layers of the memory gate electrode and stacked in this order, the contact installation structure being electrically insulated from the memory gate structure; a sidewall spacer formation step of

forming a sidewall spacer along each sidewall of the memory gate structure and the contact installation structure; a select gate electrode formation step of forming a conductive layer covering the memory gate structure and the contact installation structure, the sidewalls of the memory gate structure and the contact installation structure being covered by the sidewall spacers, and then etching back the conductive layer to form a sidewall-shaped select gate electrode disposed continuously between the sidewall of the memory gate structure and the sidewall of the contact installation structure through the sidewall spacers; and a contact formation step of forming at least one contact vertically provided and extending over the select gate electrode from a top portion of the contact installation structure.

Advantageous Effects of Invention

[0019] According to the present invention, a contact is provided and extends over a select gate electrode from a top portion of a contact installation structure having a configuration same as that of a memory gate structure. This configuration includes no mounting portion on a top portion of the memory gate structure unlike a conventional case. Accordingly, the distance to an upper wiring layer is reduced to achieve a reduced aspect ratio, and hence increase in a contact resistance value is avoided. In addition, the contact installation structure

can be further spaced from the upper wiring layer to avoid contact failure with the wiring layer.

Brief Description of Drawings

[0020] FIG. 1 is a schematic view illustrating a planar layout of a semiconductor device produced by a production method of the present invention.

FIG. 2 is a cross-sectional view illustrating a cross section taken along line A-A' in FIG. 1.

FIG. 3 is a cross-sectional view illustrating a cross section taken along line B-B' in FIG. 1.

FIG. 4A is a cross-sectional view illustrating a cross section taken along line C-C' in FIG. 1; and FIG. 4B is a cross-sectional view illustrating a cross section taken along line D-D' in FIG. 1.

FIG. 5A is a schematic view illustrating semiconductor-device production step (1); FIG. 5B is a schematic view illustrating semiconductor-device production step (2); and FIG. 5C is a schematic view illustrating semiconductor-device production step (3).

FIG. 6A is a schematic view illustrating semiconductor-device production step (4); FIG. 6B is a schematic view illustrating semiconductor-device production step (5); and FIG. 6C is a schematic view illustrating semiconductor-device production step (6).

FIG. 7 is a cross-sectional view illustrating a cross section taken along line D-D' in FIG. 1 in semiconductor-device production step (4).

FIG. 8A is a schematic view illustrating semiconductor-device production step (7); FIG. 8B is a schematic view illustrating semiconductor-device production step (8); and FIG. 8C is a schematic view illustrating semiconductor-device production step (9).

FIG. 9A is a schematic view illustrating semiconductor-device production step (10); and FIG. 9B is a schematic view illustrating semiconductor-device production step (11).

FIG. 10 is a schematic view illustrating a state in which select gate electrodes are placed over the planar layout in FIG. 1 and indicating where select gate electrode cutout portions are to be formed.

FIG. 11 is a cross-sectional view illustrating a cross section taken along line D-D' illustrated in FIG. 10.

FIG. 12A is a cross-sectional view illustrating a cross section taken along line A-A' in FIG. 1 in semiconductor-device production step (12); and FIG. 12B is a cross-sectional view illustrating a cross section taken along line in B-B' FIG. 1 in semiconductor-device production step (12).

FIG. 13 is a cross-sectional view illustrating a cross section of a conventional semiconductor device including a contact installation portion.

Description of Embodiments

[0021] Hereinafter, embodiments of the present invention will be described in the following order.

1. Configuration of semiconductor device of the present invention
 - 1-1. Planar layout of semiconductor device
 - 1-2. Cross sections of sites of semiconductor device
 - 1-3. Principles of operation to inject charge into charge storage layer in selected memory cell
 - 1-4. Principles of operation not to inject charge into charge storage layer in nonselected memory cell in which high charge storage gate voltage is applied to memory gate electrode
2. Production method for semiconductor device
3. Operations and effects
4. Production methods according to other embodiments without third photomask processing step
5. Other embodiments

[0022] (1) Configuration of semiconductor device of the present invention

(1-1) Planar layout of semiconductor device

FIG. 1 is a schematic view illustrating a planar layout of a semiconductor device 1 of the present

invention. FIG. 1 mainly illustrates a planar layout of memory gate structures 4a and 4b, first select gate structures 5a and 5b, second select gate structures 6a and 6b, contact installation structures 10a, 11a, 10b, and 11b, and select gate electrode cutout portions 13, 14, 15, and 16, which are all disposed in a memory circuit region ER1. FIG. 1 also mainly illustrates a planar layout of logic gate structures 7a and 7b disposed in a peripheral circuit region ER2. FIG. 1 omits illustrations of, for example, a sidewall spacer disposed on each sidewall of the memory gate structures 4a and 4b and the contact installation structures 10a, 11a, 10b, and 11b, any sidewall structure formed in the first select gate structures 5a and 5b and the second select gate structures 6a and 6b, and any element isolation layer disposed in a memory well W1 and logic wells W2 and W3, which will be described later.

[0023] In the present invention, is the contact installation structures 10a, 11a, 10b, and 11b include characteristic configuration. The following description will be first made on the overall configuration of the semiconductor device 1 in which the contact installation structures 10a, 11a, 10b, and 11b are disposed. Specific configurations of the contact installation structures 10a, 11a, 10b, and 11b will be described in detail later in "(1-2) Cross sections of sites of semiconductor device".

[0024] The semiconductor device 1 includes the memory circuit region ER1 and the peripheral circuit region ER2 on a semiconductor substrate (not illustrated). For example, the P-type memory well W1 is disposed in the memory circuit region ER1. The P-type logic well W2 and the N-type logic well W3 are disposed in the peripheral circuit region ER2.

[0025] The memory circuit region ER1 includes a memory cell region ER11 between gate-contact cutout regions ER12 and ER13. The memory cell region ER11 includes a plurality of memory cells 3a, 3b, 3c, 3d, 3e, and 3f disposed in a matrix of rows and columns. The memory cells 3a, 3b, 3c, 3d, 3e, and 3f all have identical configurations, and thus the following description will be made mainly on the memory cells 3a and 3b disposed along line A-A'.

[0026] The memory cell 3a has a configuration in which the memory gate structure 4a is disposed between the first select gate structure 5a and the second select gate structure 6a through a sidewall spacer (not illustrated). In the present embodiment, the memory gate structure 4a, which extends across the memory cells 3a, 3c, and 3e of the first column, and the memory gate structure 4b, which extends across the remaining memory cells 3b, 3d, and 3f of the second column, are formed linearly and disposed next to each other. A contact C4a (C4b) is vertically provided on the memory gate structure 4a (4b) and

connected to a memory gate line (not illustrated) to allow application of a memory gate voltage from the memory gate line to a memory gate electrode G1a (G1b) through the contact C4a (C4b).

[0027] In the memory cell region ER11, the first select gate structure 5a (5b) and the second select gate structure 6a (6b) are formed linearly and disposed next to the memory gate structure 4a (4b). The first select gate structure 5a (5b) includes a first select gate electrode G2a (G2b), and the second select gate structure 6a (6b) includes a second select gate electrode G3a (G3b). The first select gate electrode G2a (G2b) and the second select gate electrode G3a (G3b) respectively have sidewall shapes along sidewall spacers on sidewalls of the memory gate electrode G1a (G1b). The first select gate electrode G2a (G2b) and the second select gate electrode G3a (G3b) are disposed on the perimeter surrounding the memory gate electrode G1a (G1b), and electrically insulated from each other at the select gate electrode cutout portions 13 and 14 (15 and 16). The first select gate electrode G2a (G2b) and the second select gate electrode G3a (G3b) are not provided at the select gate electrode cutout portions 13 and 14 (15 and 16).

[0028] Two source regions D1 and D3 are symmetrically disposed at an interval therebetween on a surface of the memory well W1 in the memory cell region ER11, and a

plurality of drain regions D2 are disposed between the source regions D1 and D3. In the memory cell region ER11, the memory cells 3a, 3c, and 3e of the first column are disposed between the source region D1 and the drain regions D2. The memory cells 3b, 3d, and 3f of the second column are disposed between the drain regions D2 and the source region D3. The memory cells 3a, 3c, and 3e and the memory cells 3b, 3d, and 3f are symmetrically disposed with respect to the drain regions D2. Each of the memory cells 3a, 3c, and 3e between the source region D1 and the drain regions D2 has a configuration in which the memory gate structure 4a is disposed between the first select gate structure 5a and the second select gate structure 6a. Each of the memory cells 3b, 3d, and 3f between the drain regions D2 and the source region D3 has a configuration in which the memory gate structure 4b is disposed between the second select gate structure 6b and the first select gate structure 5b.

[0029] The source region D1 on the surface of the memory well W1 is disposed along the first select gate structure 5a and extends to regions adjacent to the first select gate structure 5a in accordance with positions of the memory cells 3a, 3c, and 3e of the first column. The source region D1 is shared by the memory cells 3a, 3c, and 3e arranged in line. A contact C1 is connected to a source line (not illustrated) and vertically provided on

the source region D1 to allow application of a source voltage from the source line through the contact C1.

[0030] The drain regions D2 on the surface of the memory well W1 between the second select gate structures 6a and 6b are disposed in regions adjacent to the second select gate structures 6a and 6b in accordance with positions of the memory cells 3a and 3b (3c, 3d, 3e, and 3f) adjacent to the drain regions D2, respectively. The drain region D2 is shared by the adjacent memory cells 3a and 3b (3c and 3d; or 3e and 3f). A contact C2 is connected to a bit line (not illustrated) and vertically provided on the drain region D2 to allow application of a bit voltage through the contact C2 from the bit line. In FIG. 1, the bit line (not illustrated) is shared by the memory cells 3a and 3b (3c and 3d; or 3e and 3f) arranged in a row direction to allow application of an identical bit voltage to the memory cells 3a and 3b (3c and 3d; or 3e and 3f) on a row-by-row basis.

[0031] The source region D3 on the surface of the memory well W1 is symmetrical to the source region D1. Similar to the source region D1, the source region D3 extends to regions adjacent to the first select gate structure 5b. The source region D3 is shared by the memory cells 3b, 3d, and 3f of the second column. A contact C3 is vertically provided on the source region D3 and connected to the same source line for the source region D1. Accordingly, an identical source voltage is applied through the

contacts C1 and C3 to the memory cells 3a, 3b, 3c, 3d, 3e, and 3f disposed in the memory cell region ER11.

[0032] The two memory gate electrodes G1a and G1b, which are disposed next to each other in the memory cell region ER11, extend linearly and next to each other in the gate-contact cutout region ER12 adjacent to the memory cell region ER11 and the gate-contact cutout region ER13 adjacent to the memory cell region ER11. One end of each of the memory gate electrodes G1a and G1b is disposed in the gate-contact cutout region ER12 and the other end of each of the memory gate structures 4a and 4b is disposed in the gate-contact cutout region ER13.

[0033] In the present embodiment, the first select gate electrode G2a, the memory gate electrode G1a, and the second select gate electrode G3a included in the memory cells 3a, 3c, and 3e of the first column are symmetrical to the second select gate electrode G3b, the memory gate electrode G1b, and the first select gate electrode G2b included in the memory cells 3b, 3d, and 3f of the second column. Thus, the following description of the gate-contact cutout regions ER12 and ER13 will be made with the first select gate electrode G2a, the memory gate electrode G1a, and the second select gate electrode G3a included in the memory cells 3a, 3c, and 3e of the first column.

[0034] The contact installation structure 10a separated and insulated from the memory gate electrode G1a is

disposed in the gate-contact cutout region ER12. In the present embodiment, the contact installation structure 10a has a strip shape and disposed on a straight line along a longitudinal direction of the memory gate electrode G1a. The first select gate electrode G2a extending from the memory cell region ER11 has a four-sided shape and is disposed in the gate-contact cutout region ER12. The contact installation structure 10a is disposed in a central region surrounded by the first select gate electrode G2a through a sidewall spacer, and thus the contact installation structure 10a is adjacent to the first select gate electrode G2a through the sidewall spacer.

[0035] In the gate-contact cutout region ER12, a contact C5a is vertically provided in a region extending from above the contact installation structure 10a, over the sidewall spacer and the first select gate electrode G2a, to a surface (substrate surface) of substrate. This configuration allows application of a first select gate voltage to the first select gate electrode G2a from a first select gate line (not illustrated) through the contact C5a.

[0036] In addition, the gate-contact cutout region ER12 includes a select gate electrode cutout portion 13 between a part of the first select gate electrode G2a having the four-sided shape and an end of the linear second select gate electrode G3a extending from the

memory cell region ER11. Through the select gate electrode cutout portion 13, the part of the first select gate electrode G2a having the four-sided shape faces the end of the second select gate electrode G3a at a predetermined distance, and the select gate electrode cutout portion 13 electrically separates the first select gate electrode G2a from the second select gate electrode G3a. With this configuration, the select gate electrode cutout portion 13 blocks voltage application from the first select gate electrode G2a to the second select gate electrode G3a when the first select gate voltage is applied to the first select gate electrode G2a through the contact C5a in the gate-contact cutout region ER12.

[0037] The contact installation structure 11a separated and insulated from the memory gate electrode G1a is disposed in the gate-contact cutout region ER13. In the present embodiment, similar to the contact installation structure 10a described above, the contact installation structure 11a has a strip shape and is disposed on a straight line along the longitudinal direction of the memory gate electrode G1a.

[0038] The second select gate electrode G3a extending from the memory cell region ER11 has a four-sided shape in the gate-contact cutout region ER13. The contact installation structure 11a is disposed in a central region surrounded by the second select gate electrode G3a through a sidewall spacer, and accordingly, the contact

installation structure 11a is adjacent to the second select gate electrode G3a through the sidewall spacer.

[0039] In the gate-contact cutout region ER13, a contact C6a is vertically provided in a region extending from the contact installation structure 11a, over the sidewall spacer and the second select gate electrode G3a, to the substrate surface. This configuration allows application of a second select gate voltage to the second select gate electrode G3a from a second select gate line (not illustrated) through the contact C6a.

[0040] In addition, the gate-contact cutout region ER13 includes the select gate electrode cutout portion 14 between a part of the second select gate electrode G3a having the four-sided shape and an end of the linear first select gate electrode G2a extending from the memory cell region ER11. Accordingly, the part of the second select gate electrode G3a having the four-sided shape is separated and electrically insulated from the end of the first select gate electrode G2a by the select gate electrode cutout portion 14 in the gate-contact cutout region ER13. With this configuration, when the second select gate voltage is applied to the second select gate electrode G3a through the contact C6a in the gate-contact cutout region ER13, the select gate electrode cutout portion 14 blocks voltage application from the second select gate electrode G3a to the first select gate electrode G2a.

[0041] Accordingly, in the memory circuit region ER1, the contact installation structure 10a and the first select gate electrode G2a, which are connected to the contact C5a, are electrically insulated by the select gate electrode cutout portions 13 and 14 from the contact installation structure 11a and the second select gate electrode G3a, which are connected to the contact C6a. This electrical insulation allows separate control of the first select gate electrode G2a and the second select gate electrode G3a.

[0042] In the gate-contact cutout regions ER12 and ER13, the second select gate electrode G3b, the memory gate electrode G1b, and the first select gate electrode G2b on the second column respectively have configurations same as those of the first select gate electrode G2a, the memory gate electrode G1a, and the second select gate electrode G3a on the first column described above. Similar to the first column, the second select gate electrode G3b, the memory gate electrode G1b, and the first select gate electrode G2b include the contact installation structures 10b and 11b and the select gate electrode cutout portions 15 and 16.

[0043] In the memory circuit region ER1, the second select gate electrode G3b of the second column is disposed next to the second select gate electrode G3a of the first column. That is, the positions of the first

select gate electrode G2b and the second select gate electrode G3b are reversed in the left-right direction. [0044] With this configuration, the contact installation structure 11b connected to a contact C6b, through which voltage is applied to the second select gate electrode G3b of the second column, is disposed in the gate-contact cutout region ER12. The contact installation structure 10b connected to a contact C5b, through which voltage is applied to the first select gate electrode G2b of the second column, is disposed in the gate-contact cutout region ER13.

[0045] As for the second select gate electrode G3b, the memory gate electrode G1b, and the first select gate electrode G2b, the contact installation structure 10b and the first select gate electrode G2b connected to the contact C5b are separated and electrically insulated by the select gate electrode cutout portions 15 and 16 from the contact installation structure 11b and the second select gate electrode G3b connected to the contact C6b. This electrical insulation allows separate control of the first select gate electrode G2b and the second select gate electrode G3b.

[0046] The following describes the peripheral circuit region ER2 adjacent to the above-configured memory circuit region ER1. In the present embodiment, the peripheral circuit region ER2 is disposed at a position adjacent to the memory cell region ER11 of the memory

circuit region ER1, but the present invention is not limited this configuration. The peripheral circuit region ER2 may be disposed at various positions such as a position adjacent to the gate-contact cutout region ER12, a position adjacent to the gate-contact cutout region ER13, or a position adjacent to a region between the memory cell region ER11 and the gate-contact cutout region ER12.

[0047] A plurality of peripheral circuits 18 and 19 are disposed in the peripheral circuit region ER2. The peripheral circuit 18 has, for example, an N-type metal-oxide-semiconductor (MOS) transistor structure in the P-type logic well W2. In such a case, the logic well W2 includes a logic gate structure 7a to allow application of a logic gate voltage to the logic gate structure 7a through a contact C8.

[0048] In the logic well W2, impurity diffusion regions D4 and D5 are disposed in regions adjacent to the logic gate structure 7a such that the logic gate structure 7a is disposed between the impurity diffusion regions D4 and D5. A contact C9 is vertically provided on the impurity diffusion region D4. A contact C10 is vertically provided on the impurity diffusion region D5.

[0049] The peripheral circuit 19 has, for example, a P-type MOS transistor structure in the N-type logic well W3. In such a case, the logic well W3 includes a logic gate

structure 7b to allow application of a logic gate voltage to the logic gate structure 7b through a contact C12.

[0050] In the logic well W3, impurity diffusion regions D6 and D7 are disposed in regions adjacent to the logic gate structure 7b such that the logic gate structure 7b is disposed between the impurity diffusion regions D6 and D7. A contact C13 is vertically provided on the impurity diffusion region D6. A contact C14 is vertically provided on the impurity diffusion region D7.

[0051] (1-2) Cross sections of sites of semiconductor device

FIG. 2 is a cross-sectional view illustrating cross sections of the memory cells 3a and 3b in the memory cell region ER11 and the peripheral circuits 18 and 19 in the peripheral circuit region ER2, taken along line A-A' in FIG. 1. The semiconductor device 1 includes a semiconductor substrate S. The memory well W1 is disposed on the semiconductor substrate S in the memory circuit region ER1, and the logic wells W2 and W3 are disposed on the semiconductor substrate S in the peripheral circuit region ER2.

[0052] In the present embodiment, in the memory well W1, the two memory cells 3a and 3b are disposed in a region including the line A-A'. The drain regions D2, on each of which the contact C2 is vertically provided, are disposed on the substrate surface between the memory cells 3a and 3b. Since the memory cells 3a and 3b are

symmetrically disposed and have identical configurations, the memory cell 3a will be described below.

[0053] The memory cell 3a includes, the memory gate structure 4a, the first select gate structure 5a, and the second select gate structure 6a, which are disposed over the memory well W1. For example, the memory gate structure 4a is included in an N-type transistor structure. The first select gate structure 5a is included in an N-type MOS transistor structure. The second select gate structure 6a is included in an N-type MOS transistor structure.

[0054] The source region D1 and each drain region D2 are spaced at a predetermined distance on the surface of the memory well W1. A source voltage from a source line is applied to the source region D1 through the contact C1 (FIG. 1), and a bit voltage from a bit line is applied to the drain region D2 through the contact C2. In the present embodiment, impurity concentration is $1.0\text{E}21/\text{cm}^3$ or more in each of the source region D1 and the drain region D2. In the memory well W1, the impurity concentration of a surface region (for example, a region extending to the depth of 50 nm from the surface) in which a channel layer is formed is set to $1.0\text{E}19/\text{cm}^3$ or less, preferably is $3.0\text{E}18/\text{cm}^3$ or less, by impurity implantation performed in a production process.

[0055] The memory gate structure 4a includes a charge storage layer EC containing, for example, silicon nitride

(Si_3N_4), silicon oxynitride (SiON), or alumina (Al_2O_3). The charge storage layer EC is disposed over the memory well W1 between the source region D1 and the drain region D2 through a lower gate insulating film 23a made of an insulating member such as SiO_2 . The memory gate structure 4a further includes the memory gate electrode G1a on the charge storage layer EC through an upper gate insulating film 23b made of an insulating member such as SiO_2 . The lower gate insulating film 23a and the upper gate insulating film 23b of the memory gate structure 4a insulate the charge storage layer EC from the memory well W1 and the memory gate electrode G1a.

[0056] A sidewall spacer 27a made of an insulating member is disposed along a first sidewall of the memory gate structure 4a. The memory gate structure 4a is adjacent to the first select gate structure 5a through the sidewall spacer 27a. This sidewall spacer 27a, which is disposed between the memory gate structure 4a and the first select gate structure 5a, has the thickness enough to insulate the memory gate structure 4a from the first select gate structure 5a.

[0057] The first select gate structure 5a includes a gate insulating film 25a made of an insulating member on the memory well W1 between the sidewall spacer 27a and the source region D1. The gate insulating film 25a has the thickness of less than or equal to 9 nm, preferably less than or equal to 3 nm. The first select gate electrode

G2a, to which the first select gate line is connected, is disposed on the gate insulating film 25a.

[0058] The sidewall spacer 27a made of an insulating member is disposed along a second sidewall of the memory gate structure 4a. The memory gate structure 4a is adjacent to the second select gate structure 6a through the sidewall spacer 27a. This sidewall spacer 27a, which is disposed between the memory gate structure 4a and the second select gate structure 6a, has the thickness same as that of the sidewall spacer 27a between the memory gate structure 4a and the first select gate structure 5a, which is enough to insulate the memory gate structure 4a from the second select gate structure 6a.

[0059] The second select gate structure 6a includes a gate insulating film 25b made of an insulating member on the memory well W1 between the sidewall spacer 27a and the drain region D2. The gate insulating film 25b has the thickness of less than or equal to 9 nm, preferably less than or equal to 3 nm. The second select gate electrode G3a, to which the second select gate line is connected, is disposed on the gate insulating film 25b.

[0060] The first select gate electrode G2a and the second select gate electrode G3a, which are disposed along the sidewalls of the memory gate electrode G1a through the sidewall spacers 27a, are formed by etching back a conductive layer in a production step, which will be described later. Each of the first select gate electrode

G2a and the second select gate electrode G3a has a sidewall shape in which the height of the first select gate electrode G2a or the second select gate electrode G3a decreases with the distance from the memory gate electrode G1a such that the top portion of the first select gate electrode G2a or the second select gate electrode G3a is sloped (angled) toward the memory well W1.

[0061] A sidewall structure SW made of an insulating member is disposed on each of the sidewall of the first select gate structure 5a and the sidewall of the second select gate structure 6a. An extension region D1a is disposed in a region of a surface of the memory well W1 below one of the sidewall structures SW, and an extension region D2a is disposed in a region of the surface of the memory well W1 below the other sidewall structure SW.

[0062] In the present embodiment, when impurity concentration is $1 \times 10^{19}/\text{cm}^3$ or less in a region extending to the depth of 50 nm from the surface of the memory well W1 between the first select gate electrode G2a and the second select gate electrode G3a, the gate insulating films 25a and 25b are each formed to have thicknesses of less than or equal to 9 nm through a subsequent production step. When impurity concentration is $3 \times 10^{18}/\text{cm}^3$ or less in the region extending to the depth of 50 nm from the surface of the memory well W1 between the first select gate electrode G2a and the second select gate

electrode G3a, the gate insulating films 25a and 25b are each formed to have thicknesses of less than or equal to 3 nm through a subsequent production step.

[0063] The memory cell 3b has a configuration similar to that of the memory cell 3a. The memory cell 3b includes the first select gate structure 5b and the second select gate structure 6b on the memory well W1 between the source region D3 and the drain region D2. The memory gate structure 4b is disposed between the first select gate structure 5b and the second select gate structure 6b through the sidewall spacers 27a. In the memory cell 3b, the sidewall structure SW is disposed on each of the sidewall of the first select gate structure 5b and the sidewall of the second select gate structure 6b.

Extension regions D3a and D2b are disposed in regions of the surface of the memory well W1 below the respective sidewall structures SW.

[0064] The memory well W1 disposed in the memory circuit region ER1 is electrically insulated from the logic well W2 disposed in the peripheral circuit region ER2 by an element isolation layer 20. The logic well W2 disposed in the peripheral circuit region ER2 is electrically insulated from the logic well W3 by another element isolation layer 20. In the present embodiment, the peripheral circuit 18 having an N-type MOS transistor structure is disposed in the logic well W2, whereas the

peripheral circuit 19 having a P-type MOS transistor structure is disposed in the logic well W3.

[0065] The logic gate structure 7a in the logic well W2 includes a logic gate electrode G5 disposed through a gate insulating film 29a between a pair of the impurity diffusion regions D4 and D5 disposed on the substrate surface. The sidewall structure SW is disposed on a sidewall of the logic gate structure 7a. Extension regions D4a and D5a are disposed on in regions of the substrate surface below the respective sidewall structures SW.

[0066] The logic well W3, which has a conduction type different from that of the logic well W2, has a configuration similar to that of the logic well W2. The logic gate structure 7b in the logic well W3 includes a logic gate electrode G6 disposed through a gate insulating film 29b between a pair of the impurity diffusion regions D6 and D7 in the substrate surface. The sidewall structure SW is disposed on a sidewall of the logic gate structure 7b. Extension regions D6a and D7a are disposed in regions of the substrate surface below the respective sidewall structures SW.

[0067] An interlayer insulating layer 21 covers sites such as the first select gate structures 5a and 5b, the memory gate structures 4a and 4b, the second select gate structures 6a and 6b, the contact C2, and the logic gate structures 7a and 7b of the semiconductor device 1 so as

to insulate these sites from each other. Silicide SC covers surfaces of other various sites such as the source regions D1 and D3 and the drain region D2.

[0068] FIG. 3 is a cross-sectional view taken along line B-B' in FIG. 1 and illustrates a cross section including the select gate electrode cutout portions 13 and 15 in the gate-contact cutout region ER12 of the memory circuit region ER1. As illustrated in FIG. 3, the select gate electrode cutout portions 13 and 15 are disposed on the element isolation layer 20 in the memory well W1.

[0069] For example, in a region in which the select gate electrode cutout portion 15 is disposed, the sidewall-shaped second select gate electrode G3b is disposed on a first sidewall of the memory gate structure 4b through the sidewall spacer 27a. However, the first select gate electrode G2b and the second select gate electrode G3b are not disposed on a second sidewall of the memory gate structure 4b, but only an insulation wall 27b including a sidewall spacer and a sidewall structure is disposed on the second sidewall.

[0070] In the present embodiment, in a region in which the select gate electrode cutout portion 13 closer to the memory gate structure 4a is disposed, the sidewall-shaped first select gate electrode G2a is disposed on a first sidewall of the memory gate structure 4a through the sidewall spacer 27a. However, the first select gate electrode G2a and the second select gate electrode G3a

are not disposed on the second sidewall of the memory gate structure 4a, and only an insulation wall 27b including a sidewall spacer and a sidewall structure is disposed on the first sidewall. In each of the regions in which the select gate electrode cutout portions 13 and 15 are disposed, a recess 30 is formed on a surface of the element isolation layer 20 by removing a part of a substrate surface in a production process.

[0071] The following describes the contact installation structures 10a, 11a, 10b, and 11b, which include characteristic configurations of the present invention. Since the contact installation structures 10a, 11a, 10b, and 11b all have identical configurations, the following description will be made on the contact installation structure 10a. FIG. 4A is a cross-sectional view illustrating a cross section of the contact installation structure 10a in the gate-contact cutout region ER12 of the memory circuit region ER1, which is taken along line C-C' in FIG. 1. FIG. 4B is a cross-sectional view illustrating a cross section of the contact installation structure 10a, which is taken along line D-D' orthogonal to line C-C' in FIG. 1.

[0072] As illustrated in FIGS. 4A and 4B, the contact installation structure 10a is disposed on the substrate surface of the element isolation layer 20 in the memory well W1. In the contact installation structure 10a, the charge storage layer EC, the upper gate insulating film

23b, and a memory gate electrode (separated memory gate electrode) G8a are stacked in this order. The memory gate electrode (separated memory gate electrode) G8a is made of a layer same as that of the memory gate electrode G1a. Although the contact installation structure 10a includes the same components such as the charge storage layer EC, the upper gate insulating film 23b, and the memory gate electrode G8a as those of the memory gate structure 4a, no quantum tunneling effect, which is due to a large voltage difference, occurs below the memory gate electrode G8a, and hence no charge is injected into the charge storage layer EC.

[0073] In the present embodiment, the charge storage layer EC, the upper gate insulating film 23b, and the memory gate electrode G8a included in the contact installation structure 10a are made of layers same as those of the charge storage layer EC, the upper gate insulating film 23b, and the memory gate electrode G1a included in the memory gate structure 4a, respectively, and thus have thicknesses identical to those of the memory gate structure 4a.

[0074] As illustrated in FIG. 4A, the sidewall-shaped first select gate electrode G2a is disposed along the sidewall spacer 27c on the sidewall of the contact installation structure 10a. The contact C5a is vertically provided in a region extending from a part of a flat top portion of the memory gate electrode G8a, over

the sidewall spacer 27c and the first select gate electrode G2a, to the substrate surface. A part of the contact C5a is disposed in a vertical manner on the flat top portion of the memory gate electrode G8a, and a part of the contact C5a is disposed in a vertical manner on the flat surface of the element isolation layer 20.

Hence, the contact C5a is installed stably.

[0075] The contact C5a is disposed over the first select gate electrode G2a and between the memory gate electrode G8a of the contact installation structure 10a and the element isolation layer 20. This configuration allows the contact C5a to be constantly in contact with the surface of the first select gate electrode G2a, for example, when the contact C5a is formed by a photolithography process and has a positioning error relative to the first select gate electrode G2a during the photolithography process. Hence, the contact installation structure 10a is electrically connected to the first select gate electrode G2a and has a stable electric resistance without being affected by the photolithography process.

[0076] Unlike a conventional case, the contact installation structure 10a includes no mounting portion on a top portion of a memory gate electrode, but includes the charge storage layer EC, the upper gate insulating film 23b, and the memory gate electrode G8a, which are the same as those of the memory gate structure 4a. With

this configuration, the contact installation structure 10a has the height substantially identical to that of the memory gate structure 4a. The contact C5a reliably connects the sidewall-shaped first select gate electrode G2a, which is disposed along the sidewall spacer 27a on the sidewall of the memory gate structure 4a, to an upper wiring layer (not illustrated).

[0077] Accordingly, the distance from the substrate surface to the upper wiring layer in the contact installation structure 10a is determined based on the height of the memory gate structure 4a. Since the contact installation structure 10a includes no mounting portion on the top portion of the memory gate electrode unlike a conventional case, the thickness of the interlayer insulating layer 21 is reduced to avoid increase in the aspect ratio of the contact extending from the substrate surface to the upper wiring layer.

[0078] As illustrated in FIG. 4B, the first select gate electrode G2a is disposed without a gap in a region GP1 through which the sidewall spacer 27a, which is disposed along a sidewall at an end of the memory gate electrode G1a, faces the sidewall spacer 27c disposed along a sidewall at an end of the contact installation structure 10a. Accordingly, the first select gate electrode G2a is disposed continuously between the contact installation structure 10a and the memory gate electrode G1a.

[0079] With this configuration, in a case where the first select gate voltage is applied to the contact C5a, which extends from the contact installation structure 10a and over the sidewall spacer 27c and the first select gate electrode G2a, the first select gate voltage is applied to the sidewall-shaped first select gate electrode G2a, which is adjacent to the memory gate electrode G1a through the sidewall spacer 27a.

[0080] In the present embodiment, the first select gate electrode G2a is formed in the region GP1, through which the sidewall spacer 27a on the sidewall of the memory gate electrode G1a faces the sidewall spacer 27c on the sidewall of the contact installation structure 10a, by etching back the conductive layer in the production process. As a result, the thickness of the first select gate electrode G2a becomes the smallest at its substantially center portion, which is most distant from the sidewall spacers 27a and 27c facing each other, between the sidewall spacers 27a and 27c.

[0081] In other words, in the region GP1, the first select gate electrode G2a has a v-shaped concave top surface, and the top surface is angled toward the substrate surface as it approaches the center portion between the sidewall spacers 27a and 27c. The silicide SC is disposed on, for example, the surfaces of the memory gate electrode G1a, the contact installation structure 10a, and the first select gate electrode G2a.

[0082] In the semiconductor device 1, the memory gate electrodes G1a and G1b, the contact installation structures 10a, 11a, 10b, and 11b, the sidewall spacers 27a and 27c, the first select gate electrode G2a and G2b, and the second select gate electrodes G3a and G3b are disposed so that the following relation holds: $D_p < (2 \times D_{sp}) + (2 \times D_{sw})$. As illustrated in FIGS. 1 and 4B, D_p represents a clearance (distance) between the sidewall of the memory gate electrode G1a and the sidewall of the contact installation structure 10a in the region GP1, through which the sidewall spacer 27a on the sidewall of the memory gate electrode G1a faces the sidewall spacer 27c on the sidewall of the contact installation structure 10a. As illustrated in FIGS. 1 and 4A, D_{sw} represents the thickness of the first select gate electrode G2a from the sidewall spacer 27c disposed on the sidewall of the memory gate electrode G1a to the sidewall structure SW, and D_{sp} represents the thickness of the sidewall spacer 27c between the memory gate electrode G1a of the contact installation structure 10a and the first select gate electrode G2a.

[0083] When the above relation holds for the semiconductor device 1, the first select gate electrode G2a (G2b) and the second select gate electrode G3a (G3b) are each disposed without a gap in the region GP1 between the sidewall spacer 27a on the sidewall of the memory gate electrode G1a (G1b) and the sidewall spacer 27c on

the sidewall of the contact installation structures 10a and 11a (10b and 11b) facing the sidewall spacer 27a.

[0084] In the present embodiment, the memory gate electrode G1a and the contact installation structure 10a are disposed on one straight line, but the present invention is not limited to this arrangement. Other various arrangements may be employed as long as the first select gate electrode G1a is formed without a gap in the region GP1 between the sidewall spacer 27a on the sidewall of the memory gate electrode G1a and the sidewall spacer 27c on the sidewall of the contact installation structure 10a. The sidewall spacers 27a and 27c face each other.

[0085] For example, although the memory gate electrode G1a faces the contact installation structure 10a, a central line of the memory gate electrode G1a may not be aligned with a central line of the contact installation structure 10a, or the memory gate electrode G1a and the contact installation structure 10a may not be disposed on one straight line.

[0086] The memory gate electrode G1a and the contact installation structure 10a have the identical width, but the present invention is not limited to this configuration. The width of the contact installation structure 10a may be less than or greater than that of the memory gate electrode G1a. The contact installation structure 10a has a rod shape in a planar layout, but the

present invention is not limited to this shape. The contact installation structure 10a may take various outline shapes such as an L shape or a J shape.

[0087] (1-3) Principles of operation to inject charge into charge storage layer in selected memory cell

The following gives simple description of a case in which data is programmed to the memory cell 3a by injecting charge into, for example, the charge storage layer EC of the memory cell 3a in the semiconductor device 1 of the present invention. In this case, as illustrated in FIG. 2, in the memory cell (may be referred to as a selected memory cell) 3a in which charge is injected into the charge storage layer EC, the charge storage gate voltage of 12 V is applied from the memory gate line (not illustrated) to the memory gate electrode G1a of the memory gate structure 4a through the contact C4a (FIG. 1) so as to form the channel layer (not illustrated) along a region of the surface of the memory well W1 facing the memory gate electrode G1a.

[0088] Then, the gate OFF voltage of 0 V is applied from the first select gate line (not illustrated) to the first select gate electrode G2a of the first select gate structure 5a through the contact C5a (FIG. 1), and the source OFF voltage of 0 V is applied to the source region D1. Accordingly, no channel layer is formed in a region of the surface of the memory well W1 facing the first select gate electrode G2a. The first select gate

structure 5a blocks the electrical connection between the source region D1 and the channel layer below the memory gate structure 4a, and prevents voltage application from the source region D1 to the channel layer below the memory gate structure 4a.

[0089] In the second select gate structure 6a, the second select gate voltage of 1.5 V is applied from the second select gate line (not illustrated) to the second select gate electrode G3a through the contact C6a (FIG. 1), and the charge storage bit voltage of 0 V is applied to the drain region D2. Accordingly, a channel layer is formed in the memory well W1 facing the second select gate electrode G2a of the second select gate structure 6a, and the memory well W1 becomes a conductive state. The memory well W1 in the conductive state electrically connects the channel layer below the memory gate structure 4a to the drain region D2, and sets the charge storage bit voltage of 0 V to the channel layer below the memory gate structure 4a. In this case, the substrate voltage of 0 V, which is same as the charge storage bit voltage, is applied to the memory well W1.

[0090] Accordingly, since the memory gate electrode G1a has a voltage of 12 V and the channel layer below the memory gate structure 4a has a voltage of 0 V, a large voltage difference of 12 V occurs between the memory gate electrode G1a and the channel layer. This voltage difference causes the quantum tunneling effect, by which

charge is injected into the charge storage layer EC to program data.

[0091] (1-4) Principles of operation not to inject charge into charge storage layer in nonselected memory cell in which high charge storage gate voltage is applied to memory gate electrode

In the semiconductor device 1 produced by a production method of the present invention, for example, when no charge is to be injected into the charge storage layer EC of the memory cell 3a, a high charge storage gate voltage same as that for data programming is applied to the memory gate electrode G1a. With this voltage application, the first select gate structure 5a blocks the electrical connection between the source region D1 and the channel layer below the memory gate structure 4a, and the second select gate structure 6a blocks the electrical connection between the drain region D2 and the channel layer below the memory gate structure 4a. Accordingly, charge is prevented from being injected into the charge storage layer EC of the memory gate structure 4a.

[0092] In this case, the charge storage gate voltage of 12 V is applied to the memory gate electrode G1a of the memory gate structure 4a in the memory cell (also referred to as a nonselected memory cell) 3a, in which no charge is to be injected into the charge storage layer EC. The charge storage gate voltage is then applied

(transmitted) to the memory well W1, thereby forming a channel layer along a region of the surface of the memory well W1 facing the memory gate electrode G1a.

[0093] The gate OFF voltage of 0 V is applied from the first select gate line (not illustrated) to the first select gate electrode G2a of the first select gate structure 5a through the contact C5a (FIG. 1). The source OFF voltage of 0 V is applied to the source region D1. Accordingly, the memory well W1 facing the first select gate electrode G2a of the memory cell 3a is in a non-conductive state, and thus the first select gate structure 5a blocks the electrical connection between the source region D1 and the channel layer below the memory gate structure 4a.

[0094] The second select gate voltage of 1.5 V is applied from the second select gate line (not illustrated) to the second select gate electrode G3a of the second select gate structure 6a through the contact C6a (FIG. 1). The OFF voltage of 1.5 V is applied to the drain region D2. Accordingly, the memory well W1 facing the second select gate electrode G3a is in a non-conductive state, and thus the second select gate structure 6a blocks the electrical connection between the drain region D2 and the channel layer below the memory gate structure 4a.

[0095] As described above, a region of the memory well W1 below the first select gate structure 5a and a region of the memory well W1 below the second select gate structure

6a become non-conductive states on both sides of the memory gate structure 4a of the memory cell 3a. The non-conductive states block the electrical connection between the drain region D2 and the channel layer facing the memory gate electrode G1a and formed on the surface of the memory well W1, and the electrical connection between the source region D1 and the channel layer. Accordingly, a depletion layer is formed around the channel layer.

[0096] In the memory gate structure 4a, a capacitor (hereinafter referred to as a gate-insulator capacitor) C2 is achieved by a three-layer configuration of the upper gate insulating film 23b, the charge storage layer EC, and the lower gate insulating film 23a. A capacitor (hereinafter referred to as a depletion-layer capacitor) C1 of the depletion layer is formed in the memory well W1 and surrounds the channel layer. The capacitors C1 and C2 are connected to each other in series. For example, when the gate-insulator capacitor C2 has a capacitance three times as large as that of the depletion-layer capacitor C1, the channel layer has a channel potential V_{ch} calculated to be 9 V by a formula below.

[0097] [Formula 1]

$$V_{ch} = (\text{Memory gate voltage } MV - \text{Substrate voltage } CV) \times \frac{\text{Capacitance of gate-insulator capacitor } C2}{\text{Capacitance of depletion-layer capacitor } C1 + \text{Capacitance of gate-insulator capacitor } C2}$$

[0098] Thus, in the memory gate structure 4a, the channel potential V_{ch} of the channel layer surrounded by the

depletion layer in the memory well W1 is 9 V when the charge storage gate voltage of 12 V is applied to the memory gate electrode G1a. Accordingly, the memory gate electrode G1a and the channel layer have a small voltage difference of 3 V therebetween. As a result, the quantum tunneling effect does not occur, and thus charge is prevented from being injected into the charge storage layer EC.

[0099] In the memory cell 3a, an impurity diffusion region having a high impurity concentration is not formed in a region of the memory well W1 between the memory gate structure 4a and the first select gate structure 5a and in a region of the memory well W1 between the memory gate structure 4a and the second select gate structure 6a. Thus, the depletion layer is reliably formed around the channel layer formed near the surface of the memory well W1. The depletion layer prevents application of the channel potential V_{ch} from the channel layer to the gate insulating films 25a and 25b of the first select gate structure 5a and the second select gate structure 6a.

[0100] Accordingly, in the memory cell 3a, the depletion layer prevents the channel potential V_{ch} of the channel layer from being applied to the gate insulating films 25a and 25b when the gate insulating films 25a and 25b of the first select gate structure 5a and the second select gate structure 6a have thicknesses reduced in accordance with a low bit voltage applied to the drain region D2 and a

low source voltage applied to the source region D1. Thus, the insulation breakdown of the gate insulating films 25a and 25b due to the channel potential V_{ch} is avoided.

[0101] (2) Production method for semiconductor device

The following production steps for the semiconductor device 1 having the above-described configuration allows production of the contact installation structures 10a, 11a, 10b, and 11b, and the first select gate electrodes G2a and G2b and the second select gate electrodes G3a and G3b, which are controlled separately from each other, through a small number of photomask steps. Each of FIGS. 5A, 5B, and 5C illustrates a cross section taken along line A-A' in FIG. 1. In this step, first, as illustrated in FIG. 5A, after the semiconductor substrate S is prepared, the element isolation layers 20 composed of an insulating member are formed at predetermined positions such as boundaries of the memory circuit region ER1 and the peripheral circuit region ER2 by, for example, a shallow trench isolation (STI) method.

[0102] Subsequently, to perform impurity implantation, a sacrificial oxide film 30a is formed on a surface of the semiconductor substrate S by, for example, a thermal oxidation method. Then, the P-type logic well W2 and the N-type logic well W3 are formed by implanting P-type impurities and N-type impurities into the peripheral circuit region ER2 by, for example, an ion implantation technique.

[0103] Subsequently, a resist is patterned by a photolithography technique and an etching technique using a first photomask (not illustrated) dedicated to fabrication of the memory circuit region ER1. As illustrated in FIG. 5B, in which any corresponding part to that in FIG. 5A is denoted by an identical reference numeral, the memory circuit region ER1 is exposed to form a resist Rm1 covering the peripheral circuit region ER2.

[0104] Subsequently, the memory well W1 is formed by implanting P-type impurities only into the memory circuit region ER1 by using the patterned resist Rm1. In addition, a channel formation layer (not illustrated) is formed on the substrate surface facing the memory gate electrodes G1a and G1b and the sidewall spacer 27a (FIG. 2) to be formed later, by implanting N-type impurities into a surface of the memory circuit region ER1. Then, the resist Rm1 is used to remove the sacrificial oxide film 30a from the memory circuit region ER1 by using, for example, hydrofluoric acid (a first photomask processing step).

[0105] In the first photomask processing step, in a case in which a P-type substrate is used as the semiconductor substrate S, the step of forming the memory well W1 by implanting P-type impurities into the semiconductor substrate S may be omitted.

[0106] Subsequently, after the resist Rm1 is removed, as illustrated in FIG. 5C, in which any corresponding part

to that in FIG. 5B is denoted by an identical reference numeral, an ONO film, in which layers of the lower gate insulating film 23a, the charge storage layer EC, and the upper gate insulating film 23b are stacked in this order, is formed on the entire surfaces of the memory circuit region ER1 and the peripheral circuit region ER2. Then, a layer-shaped memory gate electrode conductive layer 35, which is to be formed into the memory gate electrodes G1a and G1b, is formed on the upper gate insulating film 23b. Subsequently, a protection insulating film 30b made of an insulating member is formed on the memory gate electrode conductive layer 35 by, for example, the thermal oxidation method and a chemical vapor deposition (CVD) method.

[0107] Subsequently, a resist is patterned by the photolithography technique and the etching technique using a second photomask (not illustrated) dedicated to fabrication of the memory circuit region ER1. As illustrated in FIG. 6A, in which any corresponding part to that in FIG. 5C is denoted by an identical reference numeral, a resist Rm2 is formed only where the memory gate structures 4a and 4b are to be formed and where the contact installation structures 10a, 11a, 10b, and 11b are to be formed. Then, the memory gate electrode conductive layer 35 is patterned by using the resist Rm2 to form the memory gate electrodes G1a and G1b, and the memory gate electrodes G8a, G9a, G8b, and G9b, which are

separated from the memory gate electrodes G1a and G1b and smaller in size than the memory gate electrodes G1a and G1b (second photomask processing step).

[0108] In the present embodiment, the patterning of the memory gate electrode conductive layer 35 with the use of the resist Rm2 is performed such that the memory gate electrode G1a (G1b) and the memory gate electrodes G8a and G9a (G8b and G9b), which are separated from the memory gate electrode G1a (G1b) and having the small size, are disposed on one straight line.

[0109] As illustrated in FIG. 7, an inter-electrode region (spacing) GP2 is disposed between the sidewall of the memory gate electrode G1a (G1b) and each of the sidewalls of the memory gate electrodes G8a and G9a (G8b and G9b) having the small sizes, which are formed by using the resist Rm2. The inter-electrode regions GP2 face each other at a predetermined distance.

[0110] Subsequently, after the resist Rm2 is removed, the upper gate insulating film 23b and the charge storage layer EC that are exposed in regions other than the memory gate electrodes G1a and G1b and the memory gate electrodes G8a, G9a, G8b, and G9b are removed (ON film is removed) in this order as illustrated in FIG. 6B, in which any corresponding part to that in FIG. 6A is denoted by an identical reference numeral. Accordingly, the upper gate insulating film 23b and the charge storage layer EC remain at positions corresponding to the memory

gate electrodes G1a and G1b and the memory gate electrodes G8a, G9a, G8b, and G9b, which has been formed by patterning.

[0111] As a result, the lower gate insulating film 23a, the charge storage layer EC, the upper gate insulating film 23b, and the memory gate electrode G1a (G1b) are stacked in this order to form the memory gate structure 4a (4b) in the memory circuit region ER1. In the gate-contact cutout regions ER12 and ER13, the charge storage layer EC, the upper gate insulating film 23b, and the memory gate electrode G1a (G1b), which are same as those of the memory gate structure 4a (4b), are stacked in this order to form the contact installation structures 10a and 11a (10b and 11b) on the element isolation layer 20 (contact structure formation step).

[0112] Subsequently, a protection insulating film 30c is formed on the entire surfaces of the memory circuit region ER1 and the peripheral circuit region ER2 as illustrated in FIG. 6C, in which any corresponding part to that in FIG. 6B is denoted by an identical reference numeral. The present embodiment describes a case in which the single-layer protection insulating film 30c is formed on the entire surfaces, but the present invention is not limited to a single-layer film. For example, a two-layer protection insulating film in which an oxide-based insulating film and a nitride-based insulating film

are stacked in this order may be formed on the entire surfaces.

[0113] The protection insulating film 30c thus formed is used as the sidewall spacers 27a and 27c on the sidewalls of the memory gate structure 4a (4b) and the contact installation structures 10a and 11a (10a and 11b). Thus, the protection insulating film 30c has the thickness corresponding to D_{sp} representing the thickness of the sidewall spacer 27c between the memory gate electrode G8a of the contact installation structure 10a and the first select gate electrode G2a in the above-described formula $D_p < (2 \times D_{sp}) + (2 \times D_{sw})$. The protection insulating film 30c is formed such that the above-described formula $D_p < (2 \times D_{sp}) + (2 \times D_{sw})$ is satisfied.

[0114] Subsequently, as illustrated in FIG. 8A, in which any corresponding part to that in FIG. 6C is denoted by an identical reference numeral, the protection insulating film 30c is etched back to form the sidewall spacers 27a covering the memory gate structures 4a and 4b and to form the sidewall spacers 27c covering the contact installation structures 10a, 11a, 10b, and 11b not illustrated (sidewall spacer formation step).

Subsequently, a resist is patterned by the photolithography technique and the etching technique using a third photomask (not illustrated) dedicated to fabrication of the memory circuit region ER1. This patterning forms a resist Rm3 covering the entire surface

of the peripheral circuit region ER2 but leaving the memory circuit region ER1 exposed, as illustrated in FIG. 8B, in which any corresponding part to that in FIG. 8A is denoted by an identical reference numeral.

[0115] Subsequently, impurities are implanted by using the resist Rm3 to the memory circuit region ER1, where the first select gate structures 5a and 5b (FIG. 2) and the second select gate structures 6a and 6b (FIG. 2) are to be formed. This impurity implantation forms a channel formation layer not illustrated (third photomask processing step). The channel formation layer is formed in regions of the substrate surface facing the second select gate electrodes G3a and G3b and the first select gate electrode G2a and G2b, which are to be formed later.

[0116] Subsequently, after the resist Rm3 is removed, the sacrificial oxide film 30a in the peripheral circuit region ER2 is removed by using, for example, hydrofluoric acid. Then, as illustrated in FIG. 8C, in which any corresponding part to that in FIG. 8B is denoted by an identical reference numeral, for example, the thermal oxidation method is employed to form the gate insulating films 25a and 25b at the positions where the first select gate electrode G2a and G2b (FIG. 1) and the second select gate electrodes G3a and G3b (FIG. 1) are to be formed in the memory circuit region ER1. The thermal oxidation method is also employed to form the gate insulating films 29a and 29b at the positions where the logic gate

electrodes G5 and G6 (FIG. 1) are to be formed in the peripheral circuit region ER2.

[0117] Subsequently, as illustrated in FIG. 9A, in which any corresponding part to that in FIG. 8C is denoted by an identical reference numeral, for example, a layer-shaped N-type conductive layer 37 is formed in the memory circuit region ER1 and the peripheral circuit region ER2, and a layer-shaped P-type inversion layer 38 is formed in the peripheral circuit region ER2. The conductive layer 37 is shaped into the first select gate electrode G2a and G2b, the second select gate electrodes G3a and G3b, and the logic gate electrode G5 in a subsequent processing. The inversion layer 38 is shaped into the logic gate electrode G6 in a subsequent processing.

[0118] Subsequently, a resist is patterned by the photolithography technique and the etching technique using a fourth photomask (not illustrated) dedicated to fabrication of the memory circuit region ER1. The conductive layer 37 in the memory circuit region ER1 is processed using the resist (fourth photomask processing step (select gate electrode formation photomask processing step)). As illustrated in FIG. 9B, in which any corresponding part to that in FIG. 9A is denoted by an identical reference numeral, the conductive layer 37 (FIG. 9A) exposed in the memory circuit region ER1 is etched back while the entire surface of the peripheral circuit region ER2 is covered by a resist Rm4.

Accordingly, the conductive layer 37 and the inversion layer 38, which have been covered by the resist Rm4, remain intact in the peripheral circuit region ER2. In the memory circuit region ER1, the conductive layer 37, which is being exposed in the memory circuit region ER1, is etched back. Thereby, the sidewall-shaped select gate electrodes Ga and Gb are formed along the sidewall spacers 27a on the sidewalls of the memory gate structures 4a and 4b and along the sidewall spacers 27c on the sidewalls of the contact installation structures 10a, 11a, 10b, and 11b.

[0119] FIG. 10 is a schematic view illustrating a state in which the sidewall-shaped select gate electrodes Ga and Gb, which are formed along the memory gate structures 4a and 4b and the contact installation structures 10a, 11a, 10b, and 11b, are placed over the planar layout of the memory circuit region ER1 of the completed semiconductor device 1 illustrated in FIG. 1.

[0120] As illustrated in FIG. 10, the select gate electrode Ga in a non-separated state includes an integrated region. In the integrated region, a region surrounding the memory gate electrode G1a and regions surrounding the contact installation structures 10a and 11a, which are electrically insulated from the memory gate electrode G1a, are formed integrally. The select gate electrode Ga is disposed without a gap in the region GP1, through which the sidewall spacer 27a on the

sidewall of the memory gate electrode G1a faces the sidewall spacer 27c on the sidewall of the contact installation structure 10a (11a).

[0121] In the present embodiment, the memory gate electrode G1a has a linear shape. Thus the select gate electrode Ga in the non-separated state has the integrated shape, in which a long rectangular frame surrounding the memory gate electrode G1a and short rectangular frames surrounding the contact installation structures 10a and 11a are integrated. The memory gate electrode G1a and the contact installation structures 10a and 11a extend along one direction.

[0122] In the formation of the conductive layer 37 in the memory circuit region ER1 and the formation of the select gate electrodes Ga and Gb, which are formed by etching back the conductive layer 37, the thickness of the conductive layer 37 and conditions on the etching-back of the conductive layer 37 are set to satisfy the above-described formula $D_p < (2 \times D_{sp}) + (2 \times D_{sw})$.

[0123] FIG. 11 illustrates a cross section taken along line D-D' in FIG. 10. When a production condition at each step is set such that the above-described formula holds, as illustrated in FIG. 11, the conductive layer 37 remains without a gap in the region GP1, in which the sidewall spacer 27a on the sidewall of the memory gate electrode G1a faces the sidewall spacer 27c on the sidewall of the contact installation structure 10a, after

the conductive layer 37 is etched back. As a result, the select gate electrode Ga extends between the sidewall spacer 27a on the sidewall of the memory gate electrode G1a and the sidewall spacer 27c on the sidewall of the contact installation structure 10a.

[0124] The select gate electrode Ga between the sidewall spacer 27a on the sidewall of the memory gate electrode G1a and the sidewall spacer 27c on the sidewall of the contact installation structure 10a is formed by etching back the conductive layer 37 as described above. Thus, the thickness of the first select gate electrode G2a becomes the smallest at its substantially center portion, which is most distant from the sidewall spacers 27a and 27c facing each other, between the sidewall spacers 27a and 27c. The first select gate electrode G2a has a v-shaped concave top surface, and the top surface is angled toward the substrate surface as it approaches the center portion between the sidewall spacers 27a and 27c.

[0125] As illustrated in FIG. 9B, for example, the ion implantation technique is employed to implant low concentration of N-type impurities into the memory circuit region ER1 not being covered by the resist Rm4. Accordingly, an extension region ETa is formed on the surface of the memory well W1, which has been externally exposed, and thereafter, the resist Rm4 is removed.

[0126] Subsequently, in the present embodiment, a resist is patterned by the photolithography technique and the

etching technique using a photomask (not illustrated). Then, the conductive layer 37 and the inversion layer 38 in the peripheral circuit region ER2 are patterned by using the resist to form the logic gate electrodes G5 and G6 on the gate insulating films 29a and 29b. The resist used to form the logic gate electrodes G5 and G6 is also used to simultaneously remove part of the select gate electrodes Ga and Gb in the memory circuit region ER1.

[0127] In the present embodiment, as illustrated in FIG. 12A, in which any corresponding part to that in FIG. 9A is denoted by an identical reference numeral, a resist Rr1a is disposed where the logic gate structures 7a and 7b are to be formed in the peripheral circuit region ER2. The resist Rr1a is formed in accordance with the outline shapes of the logic gate structures 7a and 7b to be formed later. Accordingly, in the peripheral circuit region ER2, the conductive layer 37 and the inversion layer 38, which have been externally exposed, are removed, whereas the conductive layer 37 and the inversion layer 38, which have been covered by the resist Rr1a, remain. As a result, in the peripheral circuit region ER2, the logic gate electrodes G5 and G6 are disposed in accordance with the outline shape of the resist Rr1a, and stacked on the gate insulating films 29a and 29b to form the logic gate structures 7a and 7b.

[0128] Almost the entire surface of the memory circuit region ER1 is covered by a resist Rr1b. The resist Rr1b,

however, includes openings only at the positions where the select gate electrode cutout portions 13, 14, 15, and 16 are to be formed. The shapes of the openings correspond to the outline shapes of the select gate electrode cutout portions 13, 14, 15, and 16.

[0129] FIG. 10 illustrates planned formation positions Pf1, Pf2, Pf3, and Pf4 at which the select gate electrode cutout portions 13, 14, 15, and 16, respectively, are to be formed by removing part of the select gate electrodes Ga and Gb. The resist Rr1b disposed in the memory circuit region ER1 includes openings only at the planned formation positions Pf1, Pf2, Pf3, and Pf4. The conductive layer of the select gate electrodes Ga and Gb exposed through the openings of the resist Rr1b is removed to form the select gate electrode cutout portions 13, 14, 15, and 16, which are the separated portions in the select gate electrodes Ga and Gb, in accordance with the outline shapes of the openings of the resist Rr1b.

[0130] FIG. 12B illustrates cross sections of the select gate electrode cutout portions 13 and 15 disposed along line B-B' in FIG. 1. As illustrated in FIG. 12B, the select gate electrodes Ga and Gb exposed through openings H1 and H3 of the resist Rr1b are removed to form the select gate electrode cutout portions 13 and 15 having the outline shapes of the openings H1 and H3 of the resist Rr1b.

[0131] In addition to the select gate electrode Gb, the sidewall spacer 27a and the gate insulating film 29b are exposed through the openings H1 and H3 of the resist Rr1b. Thus, parts of the sidewall spacer 27a and the gate insulating film 25a exposed through the openings H1 and H3 of the resist Rr1b are also removed. Accordingly, the sidewall spacer 27a is partially removed in each of regions exposed through the openings H1 and H3, and thereby a chipped portion 40 is formed near a top portion of the sidewall spacer 27a. A part of the surface of the element isolation layer 20 is removed in addition to the gate insulating film 25a, and thereby the recess 30 is formed in the element isolation layer 20.

[0132] In this manner, in the memory circuit region ER1, the select gate electrode Ga (Gb) is separated by removing a plurality of portions of the select gate electrode Ga (Gb) from the select gate electrode Ga (Gb). Thus, the first select gate electrode G2a (G2b) and the second select gate electrode G3a (G3b) are obtained from the integrally formed select gate electrode Ga (Gb). The first select gate electrode G2a (G2b) surrounds the contact installation structure 10a (10b) and has a sidewall shape along the sidewall spacer 27a on one sidewall of the memory gate electrode G1a (G1b). The second select gate electrode G3a (G3b) surrounds the contact installation structure 11a (11b) and has a

sidewall shape along the sidewall spacer 27a on the other sidewall of the memory gate electrode G1a (G1b).

[0133] Then, after the resists Rr1a and Rr1b are removed by, for example, ashing, a resist patterned for an N type or a P type is used to implant low concentration of N-type impurities or P-type impurities into the peripheral circuit region ER2 by, for example, the ion implantation technique. Accordingly, as illustrated in FIG. 12A (FIG. 12A illustrates the resists Rr1a and Rr1b, which are supposed to be removed in this step), the N-type extension region ETa is formed on the substrate surface in the logic well W2, which has been externally exposed. A P-type extension region ETb is formed on the substrate surface in the logic well W3, which has been externally exposed.

[0134] Subsequently, the resist is removed, followed by a step of forming the sidewall structure SW, a step of forming the source regions D1 and D3 and the drain regions D2 by implanting high concentration of N-type impurities or P-type impurities into necessary places by, for example, the ion implantation technique, a step of forming the silicide SC, and any other step. Then, the interlayer insulating layer 21 is formed to cover the memory cells 3a, 3b, 3c, 3d, 3e, and 3f, the contact installation structures 10a, 11a, 10b, and 11b, and the peripheral circuits 18 and 19.

[0135] Subsequently, a contact hole is formed in the interlayer insulating layer 21 from a top portion of the contact installation structure 10a (10b), over the first select gate electrode G2a (G2b), to the substrate surface. A contact hole is also formed in the interlayer insulating layer 21 over the second select gate electrode G3a (G3b) from a top portion of the contact installation structure 11a (11b) to the substrate surface. In addition, a contact hole is formed at any other place of the interlayer insulating layer 21, as necessary.

[0136] Subsequently, conductive material is injected into the contact holes to respectively form the column-shaped contacts C1, C2, C3, ... and so on at the contact holes. For example, in the contact installation structure 10a among the contact installation structures 10a, 11a, 11b, and 11b, the contact C5a is vertically provided and extends from a flat top portion of the contact installation structure 10a and over the first select gate electrode G2a, to the substrate surface. The contact C5a has a rectangular cross section. The semiconductor device 1 having the configuration illustrated in FIGS. 1, 2, 3, and 4 is produced by sequentially performing the above-described steps.

[0137] (3) Operations and effects

With the above-described configuration, the semiconductor device 1 has a configuration in which the charge storage layer EC, the upper gate insulating film

23b, and the memory gate electrodes G8a and G9a (G8b and G9b), which are same as those of the memory gate structure 4a (4b), are stacked in this order. The semiconductor device 1 includes the contact installation structures 10a and 11a (10b and 11b) electrically insulated from the memory gate structure 4a (4b). The semiconductor device 1 also includes the sidewall-shaped first select gate electrode G2a (G2b) and the sidewall-shaped second select gate electrode G3a (G3b), which are respectively disposed continuously between the memory gate structure 4a (4b) and the contact installation structure 10a (10b) and between the memory gate structure 4a (4b) and the contact installation structure 11a (11b). [0138] The semiconductor device 1 also includes the contact C5a (C5b) and the contact C6a (C6b). The contact C5a (C5b) is vertically provided in a region extending from the top portion of the contact installation structure 10a (10b), over the sidewall spacer 27c and the first select gate electrode G2a (G2b), to the substrate surface. The contact C6a (C6b) is vertically provided in a region extending from the top portion of the contact installation structure 11a (11b), over the sidewall spacer 27c and the second select gate electrode G3a (G3b), to the substrate surface. The contact C5a (C5b) connects the first select gate electrode G2a (G2b) to an upper wiring layer. The contact C6a (C6b) connects the second

select gate electrode G3a (G3b) to another upper wiring layer.

[0139] Thus, in the semiconductor device 1, for example, the contact C5a extends over the first select gate electrode G2a from the flat top portion of the contact installation structure 10a including the charge storage layer EC, the upper gate insulating film 23b, and the memory gate electrode G8a, which are same as those of the memory gate structure 4a. As a result, no mounting portion 102b is formed on a top portion of a memory gate structure 110 unlike a conventional case (FIG. 13).

Hence, a distance to an upper wiring layer is shortened and a smaller aspect ratio of the contact C2 or the like is achieved. Accordingly, increase in a contact resistance value is avoided. Since the semiconductor device 1 includes no mounting portion 102b on the top portion of the memory gate structure 110 unlike a conventional case, the contact installation structure 10a is further spaced from the upper wiring layer to avoid contact failure with the upper wiring layer.

[0140] In the production method for the semiconductor device 1 of the present invention, the layer-shaped memory gate electrode conductive layer 35, the layer-shaped upper gate insulating film 23b, and the layer-shaped charge storage layer EC are patterned in this order in the memory circuit region ER1, to form the memory gate structures 4a and 4b each including the

memory gate electrode G1a, the upper gate insulating film 23b, the charge storage layer EC, and the lower gate insulating film 23a. At the same time, the contact installation structures 10a, 11a, 10b, and 11b are formed by using the same layers as those of the memory gate structures 4a and 4b. The contact installation structures 10a, 11a, 10b, and 11b are electrically insulated from the memory gate structures 4a and 4b (FIGS. 6A and 7).

[0141] In the production method for the semiconductor device 1, the gate insulating films 25a, 25b, 25c, 29a, and 29b are formed in the peripheral circuit region ER2 and the memory circuit region ER1 (FIG. 8A), in which the memory gate structures 4a and 4b and the contact installation structures 10a, 11a, 10b, and 11b are formed and covered by the sidewall spacers 27a and 27c. Then, the conductive layer 37 and the inversion layer 38 are formed on the gate insulating films 25a, 25b, 25c, 29a, and 29b (FIG. 9A). Thereafter, the conductive layer 37 in the memory circuit region ER1 is etched back while the conductive layer 37 and the inversion layer 38 in the peripheral circuit region ER2 remain intact.

[0142] In this manner, in the production method for the semiconductor device 1, the sidewall-shaped select gate electrodes Ga and Gb are disposed continuously around the memory gate structures 4a and 4b and the contact

installation structures 10a, 11a, 10b, and 11b and along the sidewall spacers 27a and 27c (FIGS. 9B, 10, and 11).

[0143] In addition, in the production method for the semiconductor device 1, the conductive layer 37 and the inversion layer 38 in the peripheral circuit region ER2 are patterned by using the resist Rr1a, which is patterned using a photomask. Thereby, the logic gate electrodes G5 and G6 are formed on the gate insulating films 29a and 29b. The resists Rr1a and Rr1b, which are used to form the logic gate electrodes G5 and G6, are also used to separate the select gate electrodes Ga and Gb by removing portions of the select gate electrodes Ga and Gb in the memory circuit region ER1.

[0144] In this manner, the production method for the semiconductor device 1 forms the first select gate electrode G2a (G2b) and the second select gate electrode G3a (G3b). The first select gate electrode G2a (G2b) surrounds the contact installation structure 10a (10b). The second select gate electrode G3a (G3b) surrounds the contact installation structure 11a (11b) and is electrically insulated from the first select gate electrode G2a (G2b) (FIGS. 12A, 12B, and 13).

[0145] Accordingly, in the production method for the semiconductor device 1, the logic gate electrodes G5 and G6 in the peripheral circuit region ER2 are formed in a photomask step. Simultaneously, the select gate electrodes Ga and Gb in the memory circuit region ER1 are

separated to form the first select gate electrode G2a and G2b and the second select gate electrodes G3a and G3b that face the memory gate electrodes G1a and G1b and are electrically insulated from the memory gate electrodes G1a and G1b.

[0146] In the production method for the semiconductor device 1, the interlayer insulating layer 21 is formed to cover the memory cells 3a, 3b, 3c, 3d, 3e, and 3f, the contact installation structures 10a, 11a, 10b, and 11b, and the like. Then, contact holes are drilled to extend from the top portions of the contact installation structures 10a, 11a, 10b, and 11b, and over the first select gate electrode G2a and G2b and the second select gate electrodes G3a and G3b, respectively. Thereafter, the contact holes are filled with conductive material.

[0147] In this manner, in the present invention, the contacts C5a, C5b, C6a, and C6b are formed to extend from the top portions of the contact installation structures 10a, 11a, 10b, and 11b, and over the first select gate structures 5a and 5b and the second select gate structures 6a and 6b, respectively. The first select gate electrode G2a and G2b and the second select gate electrodes G3a and G3b are connected to wiring layers above the memory gate structures 4a and 4b through the contacts C5a, C5b, C6a, and C6b.

[0148] (4) Production methods according to other embodiments without third photomask processing step

In the above-described embodiment, a dedicated photomask process, in which resist is patterned by using photomasks dedicated to processing of the memory circuit region ER1, involves the total of four steps: the first photomask processing step, the second photomask processing step, the third photomask processing step, and the fourth photomask processing step for forming the select gate electrode (select gate electrode formation photomask processing step). However, the present invention is not limited to these four steps, and the dedicated photomask process may omit the impurity implantation of the third photomask processing step. The dedicated photomask process may only involve the three steps: the first photomask processing step, the second photomask processing step, and the select gate electrode formation photomask processing step (corresponding to the fourth photomask processing step).

[0149] In other words, it is not necessary to perform the third photomask processing step in a case where each of the completed first select gate structures 5a and 5b and second select gate structures 6a and 6b has a desired threshold voltage (V_{th}) without the impurity implantation in the third photomask processing step. In that case, the third photomask processing step is omitted.

[0150] In such a production method without the third photomask processing step, the sidewall spacer 27a covering the memory gate structures 4a and 4b (FIG. 6B)

is formed as illustrated in FIG. 8A (sidewall spacer formation step). Then, the sacrificial oxide film 30a in the peripheral circuit region ER2 is removed by using, for example, hydrofluoric acid. Thereafter, as illustrated in FIG. 8C, for example, the thermal oxidation method is employed to form the gate insulating films 25a and 25b where the first select gate electrode G2a and G2b (FIG. 1) and the second select gate electrodes G3a and G3b (FIG. 1) are to be formed in the memory circuit region ER1. The thermal oxidation method or the like is also employed to form the gate insulating films 29a and 29b where the logic gate electrodes G5 and G6 (FIG. 1) are to be formed in the peripheral circuit region ER2. Then, similar to the production method of the above-described embodiment, the semiconductor device 1 illustrated in FIG. 1 is produced through the production steps illustrated in FIGS. 9 to 12.

[0151] In the present embodiment without the third photomask processing step, the memory cells 3a, 3b, 3c, 3d, 3e, and 3f are incorporated through the production process, in which only the steps corresponding to the three photomasks are performed in addition to a typical process for producing the peripheral-circuit. In the memory cells 3a, 3b, 3c, 3d, 3e, and 3f, the first select gate electrode G2a and G2b and the second select gate electrodes G3a and G3b are formed to sandwich the memory gate electrodes G1a and G1b, and are separately

controlled. Thus, the production method without the third photomask processing step uses a reduced number of photomasks as compared to the production method of the above-described embodiment, resulting in cost reduction.

[0152] (5) Other embodiments

The present invention is not limited to the present embodiment, and various modifications are possible within the scope of the present invention. For example, the number of the memory cells 3a, 3b, 3c, 3d, 3e, and 3f, the number of the peripheral circuits 18 and 19, the number of the contact installation structures 10a, 11a, 10b, and 11b, and the number of the select gate electrode cutout portions 13, 14, 15, and 16 may be any other numbers. Each of the conduction types of the memory well W1 and the logic wells W2 and W3 may be any one of N-type and the P-type. Three or more contact installation structures 10a, 11a, ... may be provided, and three or more select gate electrode cutout portions may be provided.

[0153] In the above-described embodiment, the non-separated select gate electrodes Ga and Gb are separated by the select gate electrode cutout portions 13, 14, 15, and 16. Thereby, the first select gate electrodes G2a and G2b and the second select gate electrodes G3a and G3b, which are separately controlled, are obtained.

[0154] However, the present invention is not limited to the above configuration. The non-separated and

integrally-formed select gate electrodes Ga and Gb may be respectively used as sidewall-shaped gate electrodes surrounding the memory gate electrodes G1a and G1b, without the separations of the select gate electrodes Ga and Gb. In such a case, in FIG. 10, one of the two contact installation structures 10a and 11a, for example, the contact installation structure 10a, may be provided at the select gate electrode Ga. In such a semiconductor device, the contact C5a is vertically provided and extends from the top portion of the contact installation structure 10a and over the sidewall spacer 27a and the select gate electrode Ga. This configuration allows the select gate electrode Ga to be controlled separately from the memory gate electrode G1a through voltage application from the contact C5a to the select gate electrode Ga, thereby achieving an effect same as the above-described embodiment.

[0155] In the above-described embodiment, the first select gate electrode G2a and the second select gate electrode G3a are formed from the select gate electrode Ga by removing a part, which is the select gate electrode cutout portion, from the select gate electrode Ga. Thereby, the select gate electrode Ga is physically separated. However, the present invention is not limited to this configuration. For example, the select gate electrode cutout portion may include an inversion electrode cutout layer of a conduction type opposite to

that of the select gate electrode Ga, or an intrinsic semiconductor layer. The select gate electrode cutout portion may be used to form, at the select gate electrode, a PIN junction structure, an NIN junction structure, a PIP junction structure, an NPN junction structure, or a PNP junction structure. The formation of such a structure electrically separates the select gate electrode to obtain the first select gate electrode G2a and the second select gate electrode G3a.

[0156] In the above-described embodiment, the first select gate electrode G2a and the second select gate electrode G3a are provided to selectively apply voltage to a channel layer on the substrate surface facing the memory gate electrode G1a. However, the present invention is not limited to this configuration. The memory gate electrode G1a may include one of the first select gate electrode G2a and the second select gate electrode G3a each having a function to select the memory gate electrode G1a.

[0157] The above-described embodiment describes the semiconductor device 1 including the memory gate structure 4a, but the present invention is not limited to this configuration. The above-described embodiment is applicable to any other semiconductor device including a gate electrode and a sidewall-shaped gate electrode disposed on a sidewall of the gate electrode through a sidewall spacer.

[0158] For example, the memory gate structure 4a includes the charge storage layer EC. However, the gate structure may include no charge storage layer but include a gate electrode disposed on a substrate through a gate insulating film. The semiconductor device may include such a gate structure and a contact installation structure. The contact installation structure includes a separated gate electrode, which is made of a layer the same as that of the gate electrode, and is electrically insulated from the gate structure. In such a case, the semiconductor device has a configuration in which a sidewall-shaped gate electrode is disposed continuously between the gate structure and the contact installation structure, and a contact is vertically provided and extends from a top portion of the contact installation structure and over the sidewall spacer and the sidewall-shaped gate electrode.

[0159] In another embodiment, a charge storage layer may be provided, through a gate insulating film, between the substrate surface and a sidewall-shaped gate electrode disposed continuously between a gate structure and a contact installation structure. In such a case, a sidewall-shaped gate structure including the sidewall-shaped gate electrode has a configuration, in which a lower gate insulating film, a charge storage layer, an upper gate insulating film, and a memory gate electrode are stacked in this order. The gate structure, the

sidewall of which is provided with the sidewall-shaped gate structure through a sidewall spacer, includes a gate electrode disposed on the substrate through the gate insulating film. The contact installation structure includes a separated gate electrode made of a layer the same as that of the gate electrode.

[0160] The contact installation structures 10a and 11a, the select gate electrode cutout portions 13 and 14, and the like in the above-described embodiment may be disposed at various positions.

[0161] In the above-described embodiment, the peripheral circuits 18 and 19 may be other various peripheral circuits (direct peripheral circuits), such as a sense amplifier, a column decoder, and a row decoder, which are disposed in the same area as that of the memory cells 3a, 3b, 3c, 3d, 3e, and 3f. The peripheral circuits 18 and 19 may be other various peripheral circuits, such as a central processing unit (CPU), an application-specific integrated circuit (ASIC), and an input and output circuit, which are disposed in an area different from that of the memory cells 3a, 3b, 3c, 3d, 3e, and 3f.

Reference Signs List

[0162] 1 semiconductor device

3a, 3b, 3c, 3d, 3e, 3f memory cell

4a, 4b memory gate structure (gate structure)

5a, 5b first select gate structure

6a, 6b second select gate structure

10a, 11a, 10b, 11b contact installation structure

Ga, Gb select gate electrode (sidewall-shaped gate electrode)

G1a, G1b memory gate electrode (gate electrode)

G2a, G2b first select gate electrode (sidewall-shaped gate electrode)

G3a, G3b second select gate electrode (sidewall-shaped gate electrode)

G8a, G8b, G9a, G9b memory gate electrode
(separated memory gate electrode)

EC charge storage layer

20 element isolation layer (substrate)

23a lower gate insulating film

23b upper gate insulating film

Rr1a, Rr1b resist

W1 memory well (substrate)

W2, W3 logic well (substrate)

Claims

1. A semiconductor device comprising:

a memory gate structure including a lower gate insulating film, a charge storage layer, an upper gate insulating film, and a memory gate electrode stacked in this order;

a contact installation structure including at least the charge storage layer, the upper gate insulating film, and a separated memory gate electrode made of a layer same as a layer of the memory gate electrode, the contact installation structure being electrically insulated from the memory gate structure;

a sidewall-shaped select gate electrode disposed on each of a sidewall of the memory gate structure through a sidewall spacer and a sidewall of the contact installation structure through the sidewall spacer, the sidewall-shaped select gate electrode being disposed continuously between the memory gate structure and the contact installation structure; and

a contact vertically provided and extending over the sidewall spacer and the sidewall-shaped select gate electrode from a top portion of the contact installation structure.

2. The semiconductor device according to claim 1, wherein the sidewall-shaped select gate electrode is

disposed without a gap in a region between the sidewall spacer on a sidewall of the memory gate electrode and the sidewall spacer on a sidewall of the separated memory gate electrode facing the sidewall spacer on the sidewall of the memory gate electrode.

3. The semiconductor device according to claim 1 or 2, wherein a relation $D_p < (2 \times D_{sp}) + (2 \times D_{sw})$ holds where D_p represents a distance between the sidewall of the memory gate electrode and the sidewall of the separated memory gate electrode, D_{sw} represents a thickness of the sidewall-shaped select gate electrode from the sidewall spacer on the sidewall of the memory gate electrode, and D_{sp} represents a thickness of the sidewall spacer between the memory gate electrode and the sidewall-shaped select gate electrode.

4. The semiconductor device according to any one of claims 1 to 3, wherein

the side-wall shaped select gate electrode includes a first select gate electrode and a second select gate electrode, the first select gate electrode having a sidewall shape and disposed along the sidewall spacer of one of the sidewalls of the memory gate electrode, the second select gate electrode having a sidewall shape and disposed along the sidewall spacer on the other sidewall of the memory gate electrode, and

the first select gate electrode is electrically insulated from the second select gate electrode.

5. The semiconductor device according to claim 4, wherein

the memory gate structure has a linear shape, and the contact installation structures include a first contact installation structure disposed near one end of the memory gate structure in a longitudinal direction of the memory gate structure, and a second contact installation structure disposed near the other end of the memory gate structure in the longitudinal direction, and

the contacts include a first contact vertically provided and extending over the sidewall spacer and the first select gate electrode from a top portion of the first contact installation structure, and a second contact vertically provided and extending over the sidewall spacer and the second select gate electrode from a top portion of the second contact installation structure.

6. The semiconductor device according to claim 5, wherein each of the first contact installation structure and the second contact installation structure has a strip shape and disposed on one straight line along the longitudinal direction of the memory gate structure.

7. A method for producing a semiconductor device comprising:

a contact installation structure formation step of stacking layers of a lower gate insulating film, a charge storage layer, an upper gate insulating film, and a memory gate electrode in this order on a substrate, and then performing patterning to form a memory gate structure and a contact installation structure, the memory gate structure including the lower gate insulating film, the charge storage layer, the upper gate insulating film, and the memory gate electrode stacked in this order, the contact installation structure including at least the charge storage layer, the upper gate insulating film, and a separated memory gate electrode stacked in this order, the separated memory gate electrode being made of a layer same as the layer of the memory gate electrode, the contact installation structure being electrically insulated from the memory gate structure;

a sidewall spacer formation step of forming a sidewall spacer along each sidewall of the memory gate structure and the contact installation structure;

a select gate electrode formation step of forming a conductive layer covering the memory gate structure and the contact installation structure, the sidewalls of the memory gate structure and the contact installation structure being covered by the sidewall spacers, and then etching back the conductive layer to form a sidewall-

shaped select gate electrode disposed continuously between the sidewall of the memory gate structure and the sidewall of the contact installation structure through the sidewall spacers; and

a contact formation step of forming at least one contact vertically provided and extending over the select gate electrode from a top portion of the contact installation structure.

8. The method for producing a semiconductor device according to claim 7, wherein

in the contact installation structure formation step, the two or more contact installation structures are formed, and

in the select gate electrode formation step, a sidewall-shaped first select gate electrode and a sidewall-shaped second select gate electrode are formed as the select gate electrodes, the first select gate electrode being disposed continuously between one of the contact installation structures and the memory gate structure through the sidewall spacers, the second select gate electrode being disposed continuously between another of the contact installation structures and the memory gate structure through the sidewall spacers, and the second select gate electrode is electrically insulated from the first select gate electrode, and

in the contact formation step, one and another of the at least one contact are formed, the one contact being vertically provided and extending over the first select gate electrode from a top portion of the one of the contact installation structures, the other contact being vertically provided and extending over the second select gate electrode from a top portion of the other of the contact installation structures.