(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 August 2010 (19.08.2010)

PCT

(10) International Publication Number
WO 2010/092555 A2

(51) International Patent Classification: Not classified

(21) International Application Number:
PCT/IB2010/050668

(22) International Filing Date:
15 February 2010 (15.02.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/152,704 15 February 2009 (15.02.2009) US

(72) Inventor; and
(71) Applicant : ASA, Gil [IL/IL]; 104 Trumpeldor St., 30900 Zichron Yaakov (IL).

(74) Agent: DR. MARK FRIEDMAN LTD.; Moshe Aviv Tower, 54th Floor, 7 Jabotinsky St., 52520 Ramat Gan (IL).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

WO 2010/092555 A2

(54) Title: FOUR-TRANSISTOR AND FIVE-TRANSISTOR BJT-CMOS ASYMMETRIC SRAM CELLS

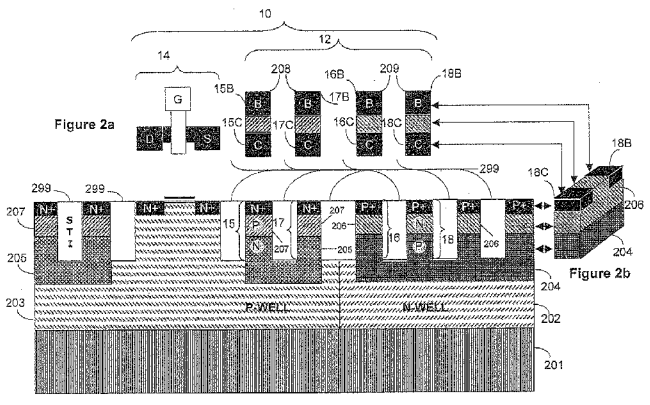


Figure 2a

Figure 2b

Figure 2

Layers of Figure 2

201 – low doping p-type substrate (part of standard CMOS)
202 – N-well (part of standard CMOS)
203 – P-well (part of standard CMOS)
204 – new p-type layer connected to V+ and forms the common emitter of the PNP
205 – new n-type layer connected to ground (common emitter for NPN)
206 – new n-type layer that forms the base (B) of the PNP
207 – new p-type layer that forms the base (B) of the NPN
208 – high doping p-type for PMOS drain/source and PNP collector (C)
209 – high doping n-type for NMOS drain/source and NPN collector (C)

(57) Abstract: A memory cell comprises asymmetric retention elements formed of bipolar junction transistors integrated with a CMOS transistor. The BJT transistors of the retention element may be vertically stacked. In one embodiment, the N region of two adjacent NPN BJT transistors may be connected to ground and may form a common emitter of the NPN BJT transistors while the P region of two adjacent PNP BJT transistors may be connected to high voltage and may form a common emitter of the PNP BJT transistors. For further compactness in one embodiment a base of one transistor doubles as a collector of another transistor. The retention element may have only a single bit line and a single write line, with no negative bit line. In some embodiments, a single inverter and only three transistors may form the retention element. Memory space may be cut approximately in half.

FOUR-TRANSISTOR AND FIVE-TRANSISTOR BJT-CMOS ASYMMETRIC SRAM CELLS

Priority Information

This patent application claims the benefit of and is a continuation-in-part of U. S. Provisional Patent Application No. 61/152,704 filed February 15, 2009.

FIELD AND BACKGROUND OF THE INVENTION

The present invention relates to memory sells for integrated circuits, and, more particularly to memory cells whose retention elements integrate vertically stacked bipolar transistor transistors ("BJT") with CMOS transistors.

The scale down tendency of the integrated circuit industry is causing an increasing technological barrier. One of the solutions is three-dimensional integrated circuits. These integrated circuits increase the manufacturing complexity significantly.

FIG. 1A shows an ordinary prior art six-transistor memory cell. As can be seen, it has two complementary bit lines, a bit line and a negative bit line. These two complementary bit lines are also tied to differential sense amplifier that has the ability to speed up the loaded bit lines.

Memory blocks hold 50-60% of dies area. Consequently, compacting the area of memory is equivalent to scaling down the size of the chip. The most serious effort to reduce the SRAM area (Static Random Access Memory) has been focused on the effort to reduce the number of transistors per memory cell. Such proposed solutions include the one-transistor SRAM (US Patent No. 6765830) and the Thyristor based SRAM (US Patent No. 6944051). All such proposed solutions, however, are not really SRAM, but rather are DRAM (Dynamic RAM) with SRAM interface. These memory cells require ongoing refresh cycles and significant amount of area for control and emulation of the SRAM interface. Consequently, these memory cells do not have the same applicability as an SRAM. Although these memory cells may achieve approximately 80% of area saving per cell, only less then 40% area saving will hold for full block of such memory with all the peripheral overhead.

It is noted that some of the SRAM cells designed in the past have demonstrated bipolar transistor cells or integration of such BJT (Bipolar Junction Transistors) with MOS transistors or other passive devices (US Patent No. 4926378 and US Patent No. 4845674).

There is a compelling need to have an apparatus that will significantly improve the compactness of integrated circuits, especially for SRAM. There is a further compelling need to have an apparatus that will also reduce the power requirements for such integrated circuits.

SUMMARY OF THE PRESENT INVENTION

One aspect of the present invention is directed to volatile memory, comprising a collection of memory cells, each memory cell comprising a retention element and a CMOS transistor, non-transistor semiconductor components including wires and substrate, the retention elements being asymmetric and comprising bipolar junction transistors, wherein each bipolar junction transistor is vertically stacked, the retention element operatively engaged to the CMOS transistor.

A further aspect of the present invention is directed to a memory cell for volatile memory, comprising a first pair of NPN bipolar junction transistors, wherein each NPN bipolar junction transistor in the first pair is vertically stacked; a second pair of PNP bipolar junction transistors, wherein each PNP bipolar junction transistor in the second pair is vertically stacked; and a CMOS transistor, the first and second pair of bipolar junction transistors forming a retention element operatively engaged to the CMOS transistor.

A still further aspect of the present invention is directed to a memory cell of a volatile memory, the memory cell comprising a retention element comprising four or fewer bipolar junction transistors, each bipolar junction transistor being vertically stacked, the retention element operatively enagged to an adjacent CMOS transistor.

A yet still further aspect of the present invention is directed to a retention element of a volatile memory, the retention element comprising first and second NPN bipolar junction transistors; and first and second PNP bipolar junction transistors, wherein each of the four transistors is vertically stacked, wherein a base of a first NPN transistor doubles as a collector of a first PNP transistor and wherein a base of a second NPN transistor doubles as a collector of a second PNP transistor.

A still further aspect of the present invention is directed to a retention element for a volatile memory, comprising a first transistor being a bipolar junction transistor;

a second transistor being a CMOS transistor; a third transistor being either a bipolar junction transistor or a CMOS transistor, wherein the retention element is capable of functioning as a retention element without the retention element containing a fourth transistor.

A still further aspect of the present invention is directed to a RAM memory for a chip, comprising a collection of memory cells, each memory cell comprising a retention element and a CMOS transistor, the retention elements being asymmetric and comprising bipolar junction transistors, the retention element operatively engaged to the CMOS transistor, wherein a memory proportion of the chip is equal to or less than 30 per cent of an area of the chip. wherein each NPN bipolar junction transistor in the first pair shares an N layer that sits on a P well of an NMOS transistor and wherein each PNP bipolar junction transistor in the second pair shares a P layer that sits on an N well of a PMOS transistor.

These and other features, aspects and advantages of the present invention will become better understood with reference to the following drawings, descriptions and claims.

BRIEF DESCRIPTION OF THE DRAWINGS

Various embodiments are herein described, by way of example only, with reference to the accompanying drawings, wherein:

FIG. 1A is a schematic of a six transistor memory cell of the prior art using CMOS technology;

FIG. 1B is a schematic of a compact memory cell using four bipolar junction transistors and having one write line and one bit line;

FIG. 2 shows a vertical cross-sectional view of a memory cell having a bipolar junction transistor retention element integrated into a CMOS environment, in accordance with one embodiment of the present invention;

FIG. 2A shows a top plan view of the transistors of the memory cell of FIG. 2;

FIG. 2B shows an isometric view of a transistor of the memory cell of FIG. 2;

FIG. 3 shows a vertical cross-sectional view of a memory cell having a bipolar junction transistor retention element that is especially compact, in accordance with one embodiment of the present invention and also shows two schematic parts of the retention element;

FIG. 3A is a schematic of the memory cell of FIG. 3 and showing how to connect the two schematic parts of FIG. 3;

FIG. 4 shows a vertical cross-sectional view of a memory cell whose retention element may be formed of only three transistors integrated into a CMOS environment, in accordance with one embodiment of the present invention;

FIG. 4A is a schematic of the memory cell of FIG. 4; and

FIG. 5 is a schematic of a memory cell that is a "mirror image" of the memory cell of FIG. 4-FIG. 4A, in accordance with one embodiment of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

The following detailed description is of the best currently contemplated modes of carrying out the invention. The description is not to be taken in a limiting sense, but is made merely for the purpose of illustrating the general principles of the invention, since the scope of the invention is best defined by the appended claims.

The present invention generally provides a memory cell for RAM memory and for other volatile memory such as a flip-flop and a register file. The cell may comprise a retention element and either an adjacent NMOS or adjacent PMOS transistor operatively engaged to the retention element. The retention element may be assymetric and may be comprised of four bipolar junction transistors, each of which may be vertically stacked. Other than the retention elements the chip may use CMOS transistors. The NPN bipolar junction transistors may share an N layer and the PNP bipolar junction transistor may share a P layer. For further compactness, in one embodiment a base of one BJT transistor may double as a collector of another BJT transistor. The retention element may have only a single bit line and a single write line, with no negative bit line. In some embodiments, a single inverter and only three transistors, including one or two BJT transistors, may form the retention element.

In contrast to the prior art, in which memory cells may either use CMOS transistors or bipolar junction transistors (BJT), the memory cell of the present invention may involve integrating a BJT retention element into a chip that may otherwise be filled with CMOS transistors. In further contrast to the prior art, in which SRAM cells may have 6 transistors, the memory cells of the present invention may have only 4 or 5 transistors. In further contrast to the priort art, in which BJT technology is power hungry, the BJT retention element of the present invention may draw a very reasonable amount of current and power (e.g. 10% of the widely used 6 transistor cell in terms of power). This may be for three reasons. First, because in contrast to the prior art, in which BJT retention elements may typically have a high current gain on the order of magnitude of 100, the retention element of

the present invention may suffice with a current gain on the order of magnitude of one or two. Second, because in contrast to prior art BJTs that are relatively large, BJTs of the present invention may be much smaller since the BJTs of the present invention must meet the critical dimensions of CMOS environment in which they are embedded, which is 100 times as small as BJT critical dimensions. Third, because in contrast to the prior art, in which the voltage on the base-emitter junction may be approximately 1.0 volts since BJTs are typically not used for memory but for amplifiers which may require higher voltage, the base-emitter junction of the BJT retention element of the present invenion may have a relatively small forward voltage of only approximately 0.7 volts or less, and preferably 0.6 volts. In still further contrast to the prior art, in which transistors of retention elements in a CMOS environment may be laid out horizontally, the transistors of the retention element of the present invention may be vertically stacked, which may make it more compact. In further contrast to the prior art, two NPN or PNP transistors may share a common emitter. In still further contrast to the prior art, the SRAM cells of the present invention may be able to save 50% of the area relative to the six transistor cell SRAM block (see, e.g. FIG. 1A) with the same operating speed. This may reduce the cost of making the chip in two ways: first because cost is directly proportional to the area of the chip and second because reducing the area of the chip may increase the yield (i.e. the proportion of chips that can be considered "good chips") which independently reduces the cost of manufacturing the chip. In still further contrast to prior art, in which additional materials involved adding additional items of critical dimension, the added materials in the chip pursant to the present invention may include nothing of critical dimensions. In further contrast to the prior art, in which retention elements of memory are symmetrical, requiring matching transistors, something that is harder and harder to achieve due to miniaturization, the retention elements of the present invention may be asymmetric. In yet still further contrast to the prior art, in which the base of one transistor in the retention element is independent of the collector of the second transistor in the retention element, in at least one embodiment of the retention element of the present invention (see, e.g. FIG. 3), the base of one transistor may also double as the collector of a second transistor. This may further add to the compactness of the cell. In still further contrast ot the prior art, in which two inverters are used in a retention element, the retention element of the present invention, in at least two embodiments (see, e.g. FIG. 4A and FIG. 5), may have a single inverter.

The term "vertically stacked" means the three regions of the bipolar junction transistor are stacked vertically. The term "area" in the context of a chip or in the context of

the memory portion of a chip refers to two-dimensional area of a chip and the two dimensions refer to length and width on a surface and do not include the depth of the chip.

When a phrase refers to "a third transistor being either a bipolar junction transistor or a CMOS transistor" the word "either" in that phrase intends to exclude a possibility of the third transistor sometimes being a BJT transistor and sometimes being a CMOS transistor.

The principles and operation of the memory cell and retention element of the present invention may be better understood with reference to the drawings and the accompanying description.

FIG. 1B shows a cell that is more compact that the prior art cell shown in FIG. 1A and has only one WL (write line) and one bit line. It uses Bipolar Junction Transistors (BJT). BJT transistors 01 and 04 form an inverter. BJT transistors 02 and 03 form a second inverter. The cell 1 is formed by two cross-coupled bipolar inverters.

FIG. 2 is a cross-sectional view of a memory cell 10 which may be formed of an asymmetric retention element 12 and a CMOS transistor 14. A collection of memory cells 10 may form a RAM for a chip or may form other volatile memory. Retention element 12 may be comprised of four bipolar junction transistors 15, 16, 17, 18 integrated into and operatively engaged to a CMOS environment such as NMOS transistor 14. Besides the CMOS transistor located outside the retention element 12, the volatile memory may include non-transistor semiconductor components such as wires and substrate.

As seen from FIG. 2, retention element 12 may include a first pair of NPN bipolar junction transistors 15, 17, wherein each NPN bipolar junction transistor in the first pair may be vertically stacked in terms of its three regions. In addition, retention element 12 may include a second pair of PNP bipolar junction transistors 16, 18, wherein each PNP bipolar junction transistor in the second pair may likewise be vertically stacked. Each NPN transistor 15, 17 and each PNP transistor 16, 18 may occupy an area smaller than an area occupied by a single smallest NMOS in the volatile memory.

The schematic memory cell structure shown in FIG. 1B may be used to create the retention element 12 shown in FIG. 2. The layers of the memory cell 10 shown in FIG. 2 have been labeled. For example, low doping p-type substrate 201 is part of a standard CMOS. Other standard parts of CMOS environment may include N-well 202 and P-well 203. FIG. 2A shows a high doping p-type region/layer 208 for PMOS drain/source, PNP collectors 16C, 18C and NPN base contacts 15B, 17B. The high-doping n-type region/layer 209 serves for NMOS drain/source, NPN collectors 15C, 17C and PNP base contacts 16B, 18B, all shown in FIG. 2. In FIG. 2, the real base lies beneath the collector and above the

emitter and the term "base contact" means that 15B, 17B, 16B, 18B denote the contact to the real base layer 206 for PNP and to real base layer 207 for NPN.

A RAM or other volatile memory of a chip may be formed of a collection of memory cells where each memory cell may be comprised of a retention element in accordance with one embodiment of the present invention, together with adjacent CMOS transistors, such as a NMOS or PMOS transistor.

As seen from FIG. 2, retention element 12 may be located adjacent a CMOS transistor which may be an NMOS transistor 14 such as is shown in FIG. 2 and which may alternatively be a PMOS transistor in other versions. Accordingly, as further seen from FIG. 2, each NPN bipolar junction transistor in the first pair may share a common emitter, such as an N layer 205 that sits on a P well 203 of an NMOS transistor of the CMOS environment. The common emitter may be at the lowest of the three regions of the NPN bipolar junction transistor. Likewise, each PNP bipolar junction transistor in the second pair may share a common emitter, such as a P layer 204 that sits on an N well 202 of a PMOS transistor. Again, the common emitter may be at the lowest of the region of the PNP bipolar junction transistor.

Accordingly, in this embodiment, four new layers 204, 205, 206, 207 may be implanted into the CMOS environment, two new vertical layers 204 and 206 for the PNP transistors and two new vertical layers 205 and 207 for the NPN transistors. New p-type layer 204 may be connected to high voltage V+ and may form the common emitter of the PNP transistors 16, 18. New n-type layer 206 may form the base (B) of the PNP transistors 16, 18. New n-type layer 205 may be connected to ground and form the common emitter for NPN transistors 15, 17. New p-type layer 207 may form the base (B) of the NPN transistors 15, 17. The term "new" when used in calculating "new" layers means something above and beyond the implants required in the standard CMOS manufacturing process. It is assumed that any CMOS retention element, or even a BJT retention element laid out horizontally, would occupy one top surface layer of the memory area of the chip. Accordingly, the top layers are not deemed new.

As seen in FIG. 2, the entire power supply of the retention element 12 may be conducted in buried layers 204, 205 and may not require additional metal resources (which also may take up some area). The BJT transistors may be completely encapsulated by surrounding shallow trench insulator ("STI") 299. The added layers may be simply engraved into the existing silicon N wells or P wells of the existing CMOS environment sitting on the substrate. The added implants/layers need no extraordinary critical dimensions or special

skills over those needed in the standard CMOS process. Items of critical dimension may not need to be added in order to make the retention elements of the present invention. Although adding even four additional layers may increase the cost of making the chip by approximately 8%, a cost savings that may be of a higher order of magnitude may be achieved from the compactness and increased yield of the retention element. Furthermore, in some embodiments (see, e.g. FIG. 3) the number of additional layers can be cut in half.

FIG. 2B is a top perspective view further depicting the positions of the regions of one of the BJT transistors in retention element 12 of FIG. 2.

FIG. 3 shows an embodiment of a memory cell 2 that may be integrated into the CMOS technology of volatile memory and that may be even more compact that the memory cell 10 of FIG. 2 since memory cell 2 may use a retention element 31 that may be even more compact than retention element 12 of FIG. 2. As seen from the elliptical circles drawn in FIG. 3, a base of one vertically stacked bipolar junction transistor may double as a collector of an adjacent vertically stacked bipolar junction transistor. For example, a base (region 302) of an NPN transistor (101) in the first pair of NPN bipolar junction transistors may double as a collector of a PNP transistor (102) in the second pair of PNP bipolar junction transistors. Likewise, a base of a PNP transistor (103) in the second pair of PNP bipolar junction transistors may double as a collector of an NPN transistor (104) in the first pair of NPN bipolar junction transistors. Accordingly, cell 2 may require only two additional implants (i.e. regions or layers) 301, 302 in addition to the implants in the standard CMOS manufacturing process. The implants may require no extraordinary critical dimensions or special skills over those needed in the standard CMOS manufacturing process. The retention element of FIG. 3, as in FIG. 2 may be asymmetrical and may have no negative bit line and one bit line.

The structure of cell 2 may also be viewed as two back-to-back thyristors . As seen in FIG. 3, the PNPN (thyristor) may be formed vertically, which may make it possible to obtain the memory cell 2 of FIG. 3A in high density. Conceptually, the structure shown in FIG. 3 may also be thought of as four transistors folded together.

In general, the BJT transistors of the retention elements of the present invention may not be hungry for a large amount of power or current. In the bipolar junction transistors of the retention elements of the present invention, the current gain on the transistors may be low, for example 1 or 2 (e.g. less than 3). It may be that the BJT transistors may be used as switches rather than as amplifiers or drivers. In addition, the transistors are very small, much smaller than usual BJT transistors since they are in a CMOS environment and the critical

dimensions of CMOS transistors are approximately one hundred times smaller than the critical dimensions of BJT technology (the same being true of the dimensions of the transistors shown in FIGS. 4, 5 relative to usual BJT transistor dimensions). The BJT transistors in the retention elements of the present invention may be only slightly wider than the width of the gate of the NMOS into which the retention element is integrated. In addition, a relatively small forward bias or forward voltage may be maintained on the base-emitter junction. The cell 2 may work on low-level power supply to preserve low power and to prevent latch up. For example, forward voltage may be maintained at approximately 0.7 volts or less, and preferably approximately 0.6 volts. As can be seen for example in FIG. 4A, the voltage may be controlled by the power supply. The retention cell 2 shown in FIG. 3 using the vertical PNP/NPN structure may result in better immunity to soft errors since most of the radiation-generated carriers must cross two back-to-back PN junctions before they could penetrate the transistor base, which is on top (and radiation tends to hit below). In some embodiments, the additional implants 301, 302 may be placed under the CMOS transistors to avoid bulk noise and soft error upset.

FIG. 4 and FIG. 4A show an embodiment in which a retention element for a RAM memory of a chip may save area of the chip by virtue of the fact that the retention element may be comprise only three transistors. The embodiments of FIG. 4, FIG. 4A and FIG. 5 together teach a retention element for a volatile memory, comprising a first transistor being a bipolar junction transistor; a second transistor being a CMOS transistor; and a third transistor being either a bipolar junction transistor or a CMOS transistor, wherein the retention element may be capable of functioning as a retention element without the retention element containing a fourth transistor.

In the retention element of the memory cell 3 shown in FIGS. 4-4A, the first transistor 402 may be an NPN transistor, the second transistor 403 may be a PMOS transistor and the third transistor 401 may be either a PMOS transistor or a PNP transistor. The "0" data bit may be preserved using a loop. In the retention element of the memory cell 3A shown in FIG. 5, in contrast, the first transistor 502 may be a PNP transistor, the second transistor 503 may be an NMOS transistor and the third transistor 501 may be either a NMOS transistor or an NPN transistor. In the cell 3A shown in FIG. 5, "1" data bit may be preserved using a loop. In both FIGS. 4-4A and FIG. 5, the retention element contains only one inverter comprised of the first transistor and the third transistor.

In the embodiment of FIG. 4 and FIG. 4A a resistor (in place of transistor 401) may connect the high power supply to the intermediate node 405. As can be appreciated from

FIG. 4, the retention element of FIG. 4A (and of FIG. 5) may be integrated into a CMOS environment of a memory chip. FIG. 4 shows a cross sectional view of the cell 3 of FIG. 4A. Since cell 3 may utilize only one NPN transistor it may require only one additional implant. Moreover, even this single implant can in certain embodiments be made part of the CMOS environment by applying the shallow P well on the deep N well to form the vertical NPN transistor (402). Note that transistor 401 is only a pull up element and can readily take the form of a lateral PNP transistor since gate 405 can become a dummy gate (the same is true for 406). Gates 407 and 406 may form part of CMOS transistors. Note also that the N well of 401 and 403 may be connected to core voltage since the other N well (under 402) may be grounded. Therefore these N wells 401, 403 should be separated with care.

The cell 3 of FIG. 4A may have only 4 transistors. Writing '1' may cause node 405 to rise to core voltage which is higher then Vs. As a result, transistor 403 may enter a state of deep cut off, and NPN transistor 402 (with no base current) may also enter a state of cut off. In this case, the voltage in node 405 may remain core voltage because it may have a high impedance path to the core supply (401). This high impedance can take the form of large resistor or PNP/PMOS transistor with appropriate base/gate bias. The pass transistor 404 can be NMOS or PMOS. Writing '0' may force zero voltage at node 405 which may turn on PMOS 403 and then NPN transistor 402 may conduct and maintain node 405 at logical '0'. In this case, a small DC current may run through the large impedance of transistor 401.

FIG. 5 shows a schematic of a three-transistor BJT retention element called memory cell 3A, which may use equivalent designs to the structure of the memory cell 3 of FIG. 4 and FIG. 4A but with the opposite pull up and pull down implementation. In memory cell 3A of FIG. 5, the high impedance 401 may be replaced by the high impedance 501 that pulls down. Again, transistor 504 can be PMOS instead of NMOS. When the system writes "1" into compact memory cell 3A shown in FIG. 5, the input node 505 may obtain V_{core} level that may turn ON NMOS transistor 503 because $V_{core} > Vs$. This state may be retained until the system writes "0", at which point pulling down node 505 will cut off NMOS transistor 503 thereby also cutting off NPN transistor 502, and the memory cell may hold "0". The input node 505 (and node 405 in FIG. 4A) may have a high resistance path towards the power supply thus giving these cells good noise immunity.

The process in FIG. 2 and FIG. 4 may be used for any application utilizing CMOS integration with low gain BJT and not only for memory applications.

The following technical details illustrate one non-limiting possible realization of the retention elements depicted in FIGS. 1-5. The following process parameters may be applied in accordance with process simulations (for example in the following order):

- - Grounded N-well (same doping and depth of equivalent ordinary CMOS process).
- First compensated P-type implant layer: doping of 2-8e17/cm^3, depth of 200nm.
- Second compensated N-type implant layer: doping 2-6e17/cm^3, depth of 160nm.
- Standard gate formation.
- Standard drain and source formation (simultaneously forming the base and emitter of the bipolar devices).

Operation parameters of a single compact cell (2) at 65nm:

- Cell estimated lay out area: 0.25u^2
- Cell internal node capacitance: 0.5fF
- Cell operating power: 0.6V
- Read out operating power: 0.9V

While the invention has been described with respect to a limited number of embodiments, it will be appreciated that many variations, modifications and other applications of the invention may be made. Therefore, the claimed invention as recited in the claims that follow is not limited to the embodiments described herein.

WHAT IS CLAIMED IS:

1. Volatile memory, comprising:
a collection of memory cells, each memory cell comprising a retention element and a CMOS transistor,
non-transistor semiconductor components including wires and substrate,
the retention elements being asymmetric and comprising bipolar junction transistors, wherein each bipolar junction transistor is vertically stacked, the retention element operatively engaged to the CMOS transistor.

2. The volatile memory of claim 1, wherein a plurality of bipolar junction transistors share a common emitter at a base of the retention element.

3. The volatile memory of claim 1, wherein a pair of NPN transistors comprises a first vertically stacked NPN transistor and a second vertically stacked NPN transistor, each NPN transistor occupying an area smaller than an area occupied by a single smallest NMOS in the volatile memory; and
a pair of PNP transistors comprising a first vertically stacked PNP transistor and a second vertically stacked PNP transistor, each PNP transistor occupying an area smaller than an area occupied by the single smallest NMOS.

4. The volatile memory of claim 1, wherein a voltage on a base-emitter junction of each bipolar junction transistor is maintained at 0.7 volts or less.

5. The volatile memory of claim 4, wherein the voltage is maintained at 0.6 volts or less.

6. The volatile memory of claim 1, wherein the transistors in a pair of NPN bipolar junction transistors share an N layer and wherein the transistors in a pair of PNP bipolar junction transistors share a P layer.

7. The volatile memory of claim 1, wherein a current gain on each of the bipolar junction transistors is in the order of magnitude of 1.

8. A memory cell for volatile memory, comprising:
a first pair of NPN bipolar junction transistors, wherein each NPN bipolar junction transistor in the first pair is vertically stacked;
a second pair of PNP bipolar junction transistors, wherein each PNP bipolar junction transistor in the second pair is vertically stacked; and
a CMOS transistor, the first and second pair of bipolar junction transistors forming a retention element operatively engaged to the CMOS transistor.

9. The memory cell of claim 8, wherein each NPN bipolar junction transistor in the first pair shares an N layer that sits on a P well of an NMOS transistor and wherein each PNP bipolar junction transistor in the second pair shares a P layer that sits on an N well of a PMOS transistor.

10. The memory cell of claim 8, wherein a base of an NPN transistor in the first pair of NPN bipolar junction transistors doubles as a collector of a PNP transistor in the second pair of PNP bipolar junction transistors.

11. The memory cell of claim 8, wherein a base of a PNP transistor in the second pair of PNP bipolar junction transistors doubles as a collector of an NPN transistor in the first pair of NPN bipolar junction transistors.

12. The memory cell of claim 8, wherein a base of one vertically stacked bipolar junction transistor doubles as a collector of an adjacent vertically stacked bipolar junction transistor.

13. The memory cell of claim 8, wherein the retention element has no negative bit line and one bit line.

14. The memory cell of claim 8, wherein each NPN bipolar junction transistor of the first pair occupies an area smaller than an area occupied by a single smallest NMOS in the volatile memory and wherein each PNP bipolar junction transistor of the second pair occupies an area smaller than an area occupied by the single smallest NMOS.

15. A memory cell of a volatile memory, the memory cell comprising:

a retention element comprising four or fewer bipolar junction transistors, each bipolar junction transistor being vertically stacked, the retention element operatively enagged to an adjacent CMOS transistor.

16. The memory cell of claim 15, wherein a current gain on each of the bipolar junction transistors is less than three.

17. The memory cell of claim 15, wherein a voltage on a base-emitter junction of each bipolar junction transistor is maintained at 0.7 volts or less.

18. The memory cell of claim 15, wherein the N region of two adjacent NPN bipolar junction transistors is connected to ground and forms a common emitter of the NPN bipolar junction transistors and wherein the P region of two adjacent PNP bipolar junction transistors is connected to high voltage and forms a common emitter of the PNP bipolar junction transistors.

19. The memory cell of claim 15, wherein an N region of a bipolar junction transistor sits on a P well of an NMOS transistor and wherein a P region of a bipolar junction transistor sits on an N well of a PMOS transistor.

20. A retention element of a volatile memory, the retention element comprising:

first and second NPN bipolar junction transistors; and

first and second PNP bipolar junction transistors, wherein each of the four transistors is vertically stacked, wherein a base of a first NPN transistor doubles as a collector of a first PNP transistor and wherein a base of a second NPN transistor doubles as a collector of a second PNP transistor.

21. The retention element of claim 20, wherein there are no additional transistors.

22. A retention element for a volatile memory, comprising:

a first transistor being a bipolar junction transistor;

a second transistor being a CMOS transistor;
a third transistor being either a bipolar junction transistor or a CMOS transistor, wherein the retention element is capable of functioning as a retention element without the retention element containing a fourth transistor.

23. The retention element of claim 22, wherein the first transistor is a NPN transistor, the second transistor is a PMOS transistor and the third transistor is either a PMOS transistor or a PNP transistor.

24. The retention element of claim 22, wherein a "0" data bit is preserved using a loop.

25. The retention element of claim 22, wherein the first transistor is a PNP transistor, the second transistor is an NMOS transistor and the third transistor is either a NMOS transistor or an NPN transistor.

26. The retention element of claim 22, wherein a "1" data bit is preserved using a loop.

27. The retention element of claim 22, wherein the retention element contains only one inverter.

28. The retention element of claim 27, wherein the inverter is comprised of the first transistor and the third transistor.

29. A RAM memory for a chip, comprising:
a collection of memory cells, each memory cell comprising a retention element and a CMOS transistor, the retention elements being asymmetric and comprising bipolar junction transistors, the retention element operatively engaged to the CMOS transistor, wherein a memory proportion of the chip is equal to or less than 30 per cent of an area of the chip. wherein each NPN bipolar junction transistor in the first pair shares an N layer that sits on a P well of an NMOS transistor and wherein each PNP bipolar junction transistor in the second pair shares a P layer that sits on an N well of a PMOS transistor.

30. The RAM memory of claim 29, wherein the memory proportion is equal to or less than one-quarter of the area of the chip.

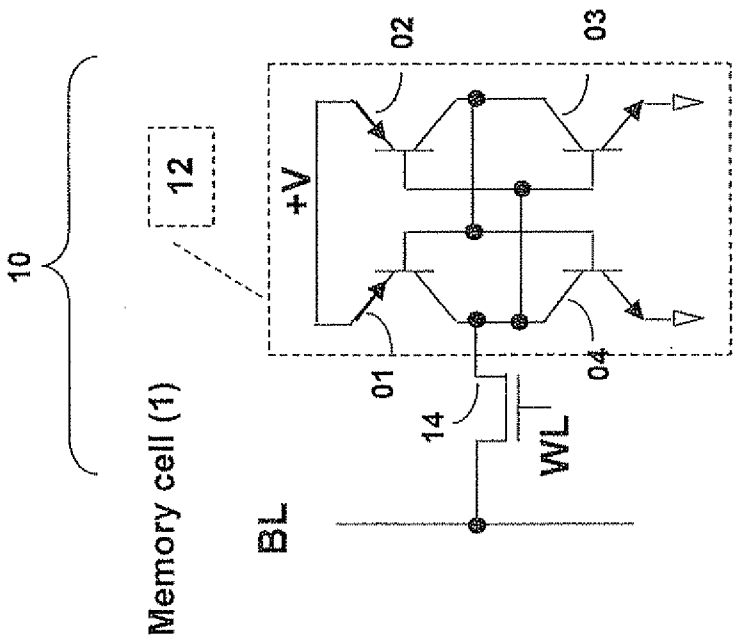


Figure 1b

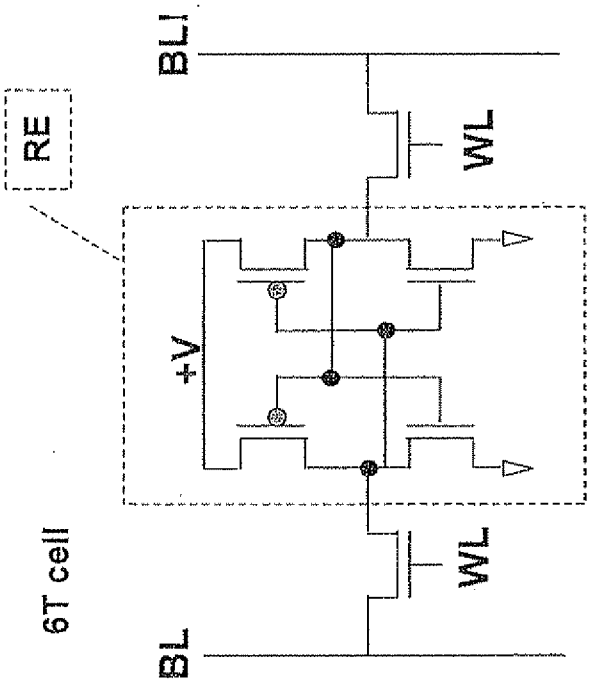


Figure 1a (prior art)

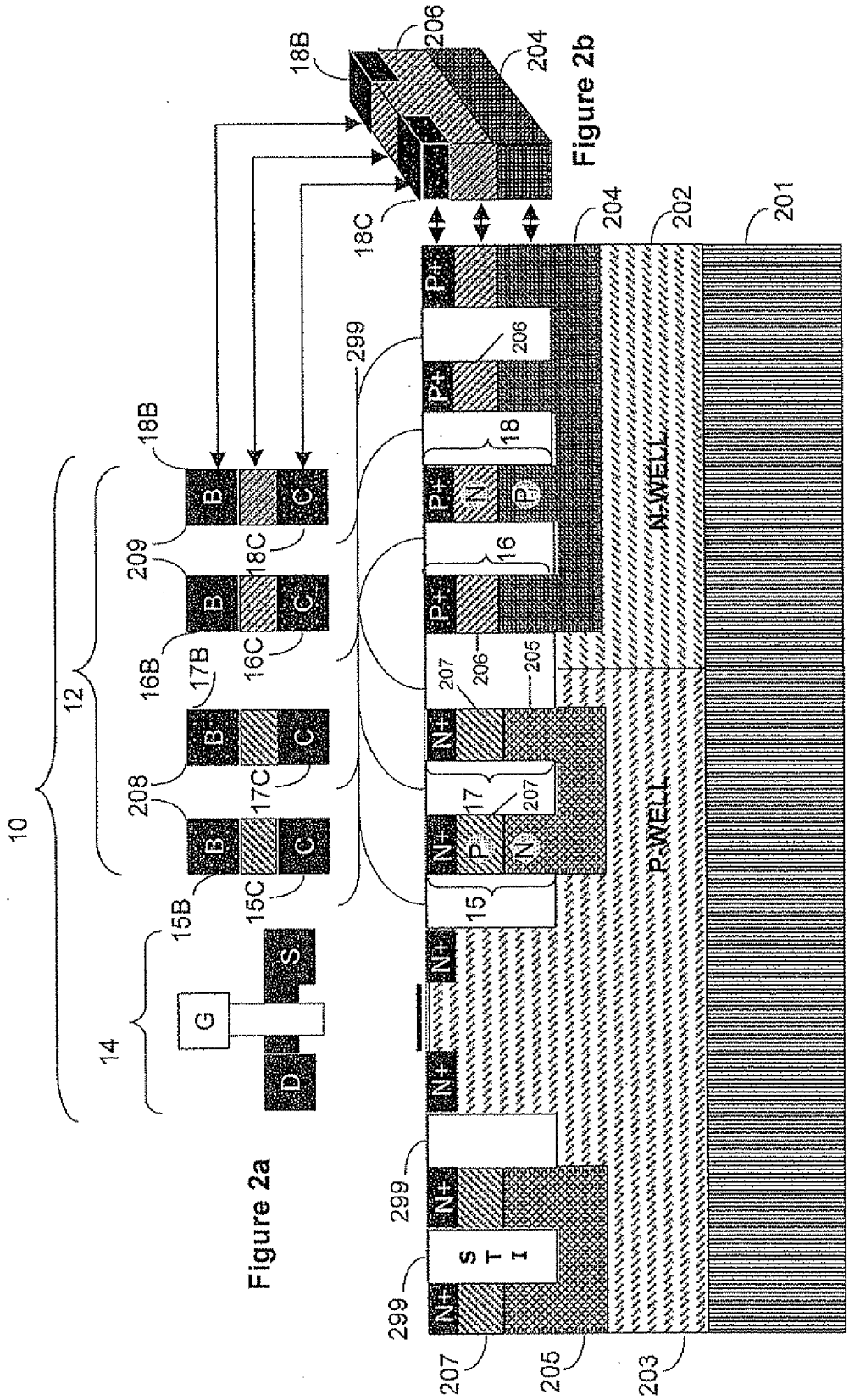


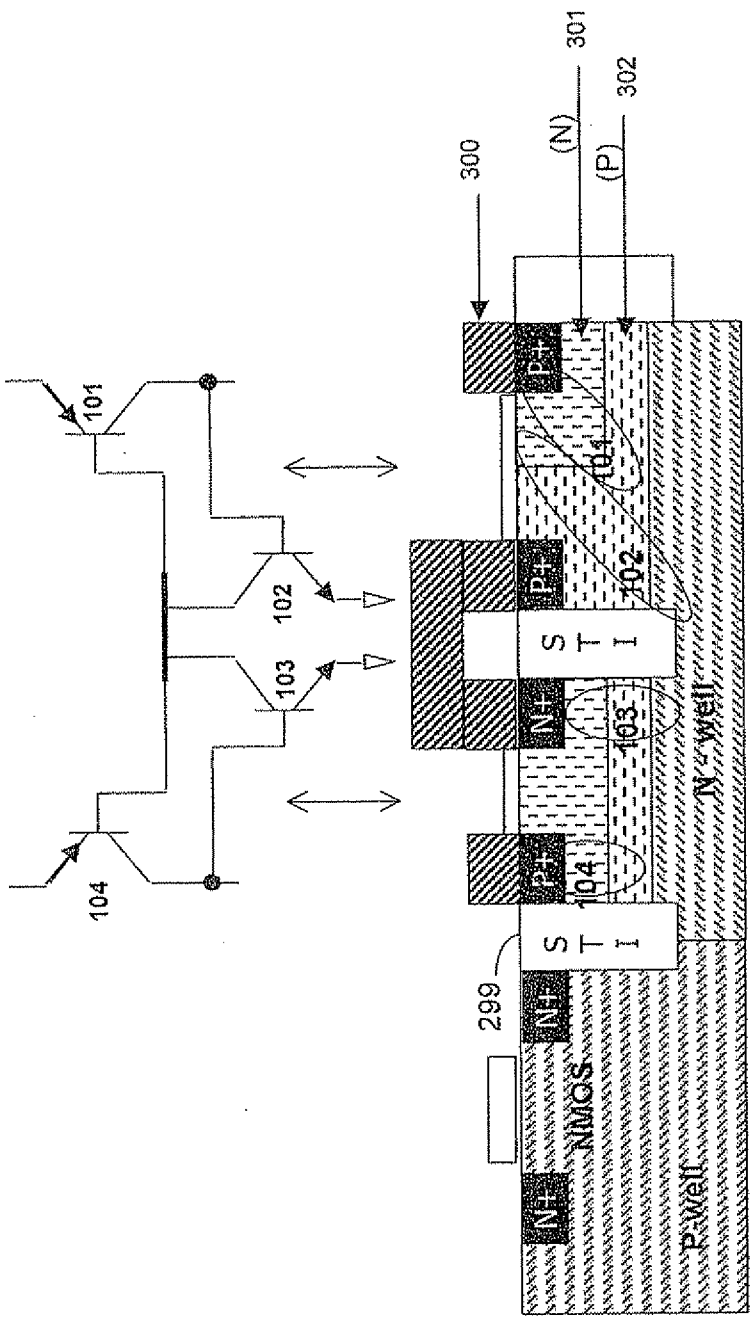
Figure 2a

Figure 2b

Figure 2

Layers of Figure 2

201 – low doping p-type substrate (part of standard CMOS)
202 – N-well (part of standard CMOS)
203 – P-well (part of standard CMOS)
204 – new p-type layer connected to V+ and forms the common emitter of the PNP
205 – new n-type layer connected to ground (common emitter for NPN)
206 – new n-type layer that forms the base (B) of the PNP
207 – new p-type layer that forms the base (B) of the NPN
208 – high doping p-type for PMOS drain/source and PNP collector (C)
209 – high doping n-type for NMOS drain/source and NPN collector (C)



Compact cell (2)

Figure 3

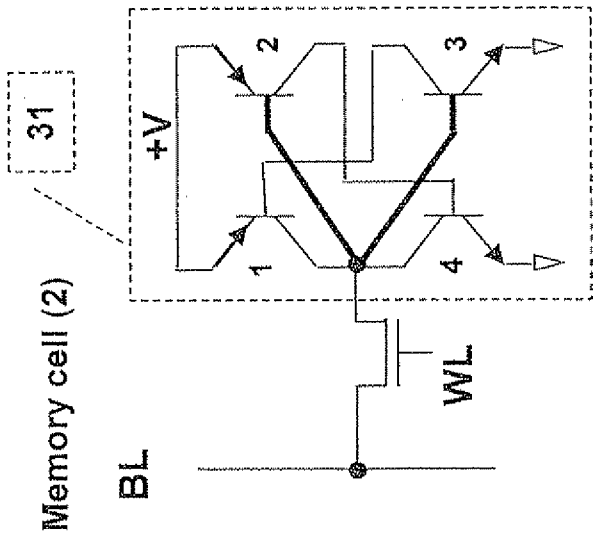


Figure 3a

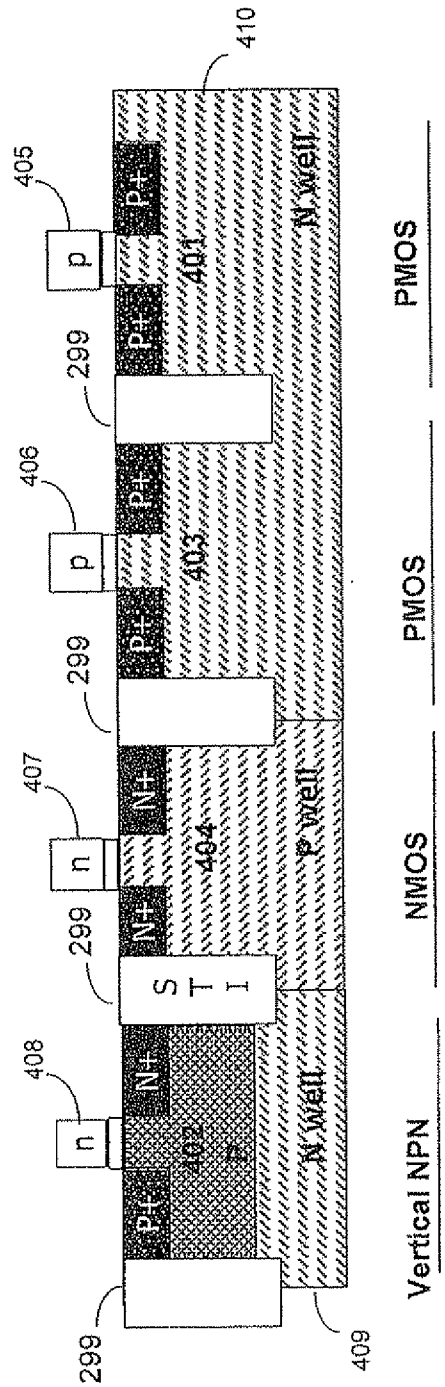

410
405
406
407
408
299
409
P+
N+
n
p
STI
401
402
403
404
N well
P well
Vertical NPN
NMOS
PMOS
PMOS

Figure 4

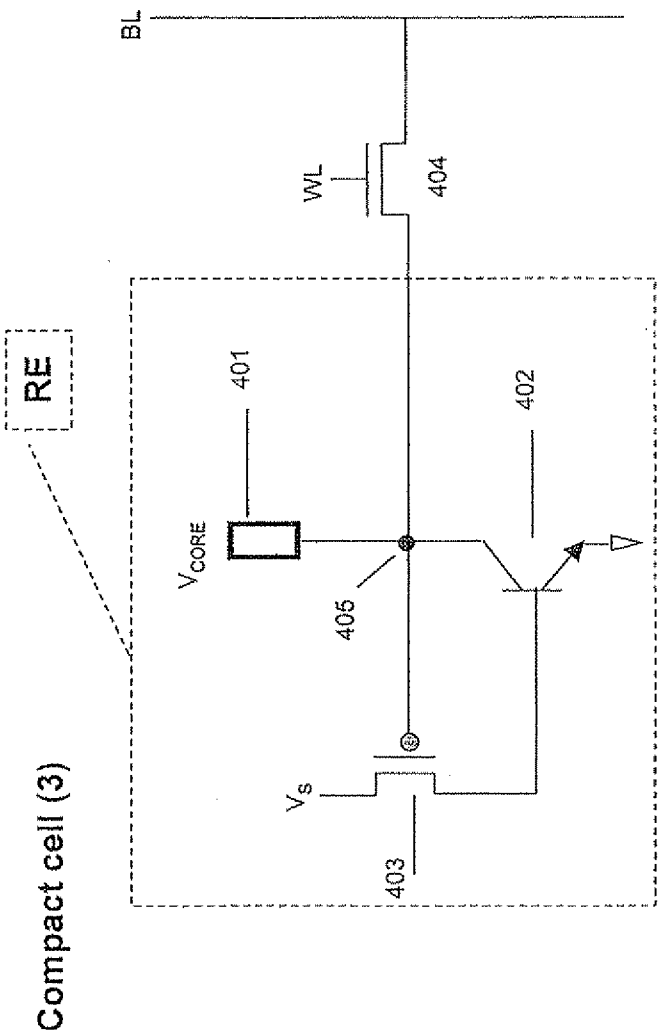


Figure 4a

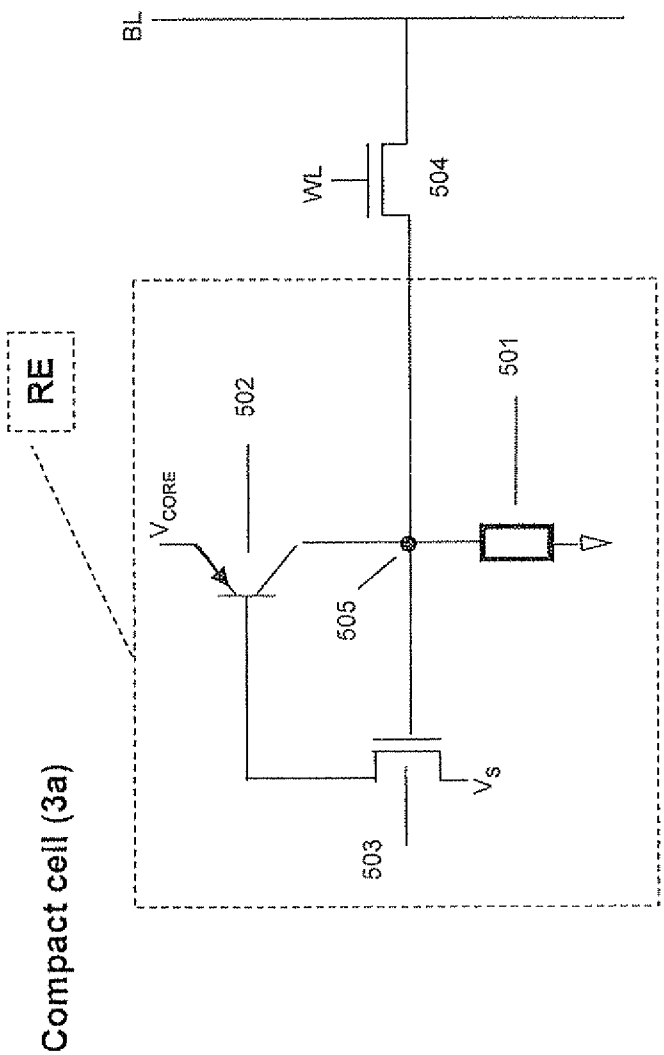


Figure 5