

NANO-SENSE AMPLIFIER FOR MEMORY

FIELD OF THE INVENTION

The invention generally relates to memories, and more particularly to a sense amplifier of a series of memory cells.

BACKGROUND OF THE INVENTION

5 A conventional sense amplifier generally addresses about 256 to 1,024 memory cells via a line, a so-called bitline. The conventional sense amplifier is more specifically a differential amplifier operating with a bitline and a complementary bitline which is used as a reference line.

10 A conventional technique for increasing the performances of dynamic DRAM memory consists in reducing the number of cells addressed by a sense amplifier (reference is also made to a reduction in the length of the bitline). However, a larger number of sense amplifiers has to be provided in order to
15 address the whole of the cells making up the memory, which is expressed by a loss of global efficiency insofar that management of the memory consumes useful surface area to the detriment of the actual memory.

In order to react to this loss of efficiency, the article
20 "A 500 MHz Random Cycle, 1.5 ns Latency, SOI Embedded DRAM Macro Featuring a Three-Transistor Micro Sense Amplifier", J Barth et al., ISSCC (2007), Pages: 486-487 shows an architectural hierarchy based on sense amplifiers, so-called micro-sense amplifiers, addressing via a local bitline, few

cells (32 typically) but consisting of very few (typically three) transistors.

It will be noted that this article proposes an architecture in which the different memory cells are made on a silicon-on-insulator substrate (SOI). The different micro-sense amplifiers are, as for them, formed either on a bulk substrate or a SOI substrate.

Each micro-sense amplifier has two input terminals connected to two main bitlines RBL (Read Bitline) and WBL (Write Bitline) which will control the read/write operations of the cells addressed in parallel via the local bitline LBL connected to the output terminal of the micro-sense amplifier.

Resorting to two main bitlines proves to be a problem in that the architecture proposed by this article cannot be transposed for producing standalone memories and thus remains limited to the production of embedded memories.

Further, each micro-sense amplifier can only address a reduced number of memory cells (16 to 32 typically) so that a relatively large number of micro-sense amplifiers (from 64k to 128k for a 2 Mbit memory) has to be resorted to. In spite of the relatively reduced size of the micro-amplifier (only 3 transistors), the drawback mentioned earlier of significant surface consumption for the sense amplification function, to the detriment of the memory function, therefore partly remains.

SHORT DESCRIPTION OF THE INVENTION

The object of the invention is to propose a technique which allows the drawbacks of the architectural hierarchy proposed in the aforementioned article to be overridden.

For this purpose, the invention proposes, according to a first aspect, a sense amplifier for a series of cells of a rewritable memory, including:

- a writing stage comprising a CMOS inverter, the input of which is directly or indirectly connected to an input terminal

of the sense amplifier, and the output of which is connected to an output terminal of the sense amplifier intended to be connected to a local bitline addressing the cells of said series;

- 5 - a reading stage comprising a sense transistor, the gate of which is connected to the output of the inverter and the drain of which is connected to the input of the inverter.

According to a first embodiment, the input of the writing stage is directly connected to the input terminal of the inverter, said input terminal being intended to be connected to a main bitline which will address a plurality of sense amplifiers in parallel.

According to a second embodiment, the reading stage comprises an additional transistor, complementary to the sense transistor, the additional transistor and the sense transistor forming a CMOS inverter, the input of which is connected to the output of the writing stage and the output of which is connected to the input of the inverter of the writing stage.

In this second embodiment, the input of the writing stage may be indirectly connected to the input terminal of the inverter via a decoding stage comprising a transistor, the drain of which is connected to the input terminal of the sense amplifier and the source of which is connected to the input of the writing stage.

25 One or more transistors of the sense amplifier are multigate transistors. The biasing of one of the gates of the multigate transistors can be modified during operations for writing, reading and retaining data in the cells of said series.

30 The sense amplifier according to the first aspect of the invention is preferentially made on a semiconductor-on-insulator substrate comprising a thin layer of semiconducting material separated from a base substrate by an insulating layer, and each of the transistors has a back control gate

formed in the base substrate below the channel and capable of being biased for modulating the threshold voltage of the transistor.

According to a second aspect, the invention relates to a method for controlling a sense amplifier according to the first aspect of the invention made on a SeOI substrate, in which the bias of the back control gates is modified during operations for writing, reading and retaining data in the cells of said series.

During a reading operation, the threshold voltage of the transistors of the writing stage may be increased and the threshold voltage of the transistor(s) of the reading stage may be decreased by controlling the bias of their back control gates.

During a writing operation, the threshold voltage of the transistors of the writing stage may be decreased and the threshold voltage of the transistor(s) of the reading stage may be increased by controlling the bias of their back control gates.

During a retention operation, the local bitline may be discharged via one of the transistors of the writing stage, for which the bias of the back control gate is controlled so as to lower its threshold voltage.

According to still another aspect, the invention relates to a matrix array of cells connected to local bitlines and to word lines, characterized in that it includes sense amplifiers according to the first aspect of the invention.

In the matrix array of cells, the sense amplifiers can be arranged in one or more sense amplifier banks and the matrix array can further comprise a column decoder placed next to each bank and configured to drive said bank directly. Further logic circuits can be placed next to the sense amplifier banks or the column decoder.

According to another aspect, the invention relates to a memory including a matrix array of cells.

The memory cells and the sense amplifiers of the memory are preferentially made on a semiconductor-on-insulator substrate.

5 According to still another aspect, the invention relates to a main sense amplifier intended to be connected via a main bitline to a plurality of sense amplifiers according to the first embodiment of the first aspect of the invention, characterized in that it comprises a stage for amplifying the
10 signal delivered by a cell during a reading operation, and a switchable high impedance inverter stage for sending back the amplified signal onto the main bitline following the reading operation.

 According to a further aspect, the invention relates to
15 sense amplifier including:

 - a writing stage comprising a CMOS inverter, the input of which is directly or indirectly connected to an input terminal of the sense amplifier, and the output of which is connected to an output terminal of the sense amplifier intended to be
20 connected to a local line,

 - a reading stage comprising a sense transistor, the gate of which is connected to the output of the inverter and the drain of which is connected to the input of the inverter;
and wherein one or more transistors of the writing stage and
25 of the reading stage are independent double gate transistors.

 Yet according to another aspect, the invention relates to a matrix array of cells comprising sense amplifiers and column decoders, wherein the sense amplifiers are arranged in a non-staggered fashion in one or more sense amplifier banks and
30 wherein a column decoder is placed next to each bank and configured to drive said bank directly.

SHORT DESCRIPTION OF THE DRAWINGS

Other aspects, objects and advantages of the present invention will become better apparent upon reading the

following detailed description of preferred embodiments thereof, given as a non-limiting example, and made with reference to the appended drawings wherein:

- Fig. 1 illustrates the architectural hierarchy applied in a first embodiment of the invention;
- Fig. 2 illustrates a nano-sense amplifier according to the first embodiment of the invention;
- Figs. 3a and 3b illustrate exemplary transistors having back control gates;
- Fig. 4a illustrates a possible topology of a nano-sense amplifier according to Fig. 2;
- Fig. 4b illustrates a possible organization of nano-sense amplifiers according to Fig. 2 along columns of a memory matrix;
- Fig. 5 illustrates a main sense amplifier intended for transferring data towards/from a plurality of nano-sense amplifiers according to Fig. 2;
- Fig. 6 illustrates a nano-sense amplifier according to a second embodiment of the invention.
- Fig. 7 illustrates the state-of-the-art array architecture;
- Fig. 8a illustrates an innovative array architecture using two banks of nano-sense amplifiers according to the invention;
- Fig. 8b illustrates a detailed organization of the busses for decoding each of the nano-sense amplifiers of the banks in the architecture of Fig. 8a;
- Fig. 9a illustrates an extended innovative array architecture using two banks of nano-sense amplifiers according to the invention;
- Fig. 9b illustrates a detailed organization of the busses for decoding each of the nano-sense amplifiers of the banks in the architecture of Fig. 9a.

DETAILED DESCRIPTION OF THE INVENTION

The invention according to a first aspect relates to a sense amplifier (called a nano-sense amplifier) of a series of cells of a memory.

5 An architectural hierarchy applied in a first embodiment of the invention is illustrated in Fig. 1 and according to which a nano-sense amplifier nSA covers a series of memory cells via a local bitline LBL, typically between 128 and 512 memory cells.

10 The nano-sense amplifier nSA transfers data towards/from a main sense amplifier MSA via a main bitline MBL. The main bitline MBL covers a series of nano-sense amplifiers nSA, typically between 16 and 32 nano-sense amplifiers nSA.

15 The main sense amplifier MSA is moreover connected to a line, a so-called dataline DL, on which read/written data circulate in the memory cells.

It will be understood from the description which follows of the first embodiment of the invention that a nano-sense amplifier nSA provides the reading and writing functions, while the refreshing function is achieved by the main sense amplifier MSA which forms the interface between the nano-sense
20 amplifier nSA and the peripheral circuits.

It will be noted from now on that the nano-sense amplifier nSA has a single connection (the main bitline MBL) with the main sense amplifier MSA, with which it is possible to meet
25 the constraints in terms of metallization width both in the field of embedded memories and in that of standalone memories.

As illustrated in Fig. 2, a nano-sense amplifier nSA according to the first embodiment of the invention only includes three transistors T1, T2, T3.

30 The nano-sense amplifier nSA more specifically comprises an input terminal E connected to the main bitline MBL and an output terminal S connected to the local bitline LBL via the input terminal E.

The nano-sense amplifier nSA comprises a writing stage comprising a CMOS inverter, the input of which is directly connected to the input terminal E and the output of which is connected to the output terminal S.

5 The CMOS inverter comprises, between two terminals for applying power supply potentials, VDD and GND respectively, a transistor with a first type of channel in series with a transistor with a second type of channel.

10 In the illustrated embodiment, the inverter comprises a P channel transistor T1, the source of which is connected to the potential VDD (high state) and a N channel transistor T2, the source of which is connected to the potential GND (low state).

The gates of the transistors of the inverter are connected together and connected to the main bitline MBL. The middle
15 point of the serial combination of the transistors T1 and T2 (the output of the inverter) is as for it connected to the local bitline LBL.

The nano-sense amplifier nSA further comprises a reading stage formed by a transistor T3, here an N channel transistor,
20 the gate of which is connected to the output of the inverter and to the local bitline LBL, the drain of which is connected to the input of the inverter and to the main bitline MBL, and the source of which is connected to a terminal for applying a power supply potential PGND.

25 Within the scope of a preferred embodiment of the invention, the nano-sense amplifier nSA is made on a semiconductor-on-insulator substrate comprising a thin layer of semiconducting material separated from a base substrate by an insulating layer. Each of the transistors has a back
30 control gate formed in the base substrate below the channel and capable of being biased in order to modulate the threshold voltage of the transistor. The bias respectively applied to the back control gate of transistors T1, T2 and T3 is illustrated on the figure by references Vbg1, Vbg2 and Vbg3.

It will be noted that the modulation of the threshold voltage is particularly advantageous in that it allows the operation of a transistor to be forced so as to ensure, depending on the circumstances, that it is actually blocked or
 5 conducting. With this, the operation of the nano-sense amplifier nSA may be made more secure, and therefore the operation margin thereof may be increased (which allows association of a larger number of memory cells with the nano-sense amplifier nSA).

10 In Fig. 3a, a P channel transistor is illustrated, having a back control gate BG_P arranged in the base substrate under the BOX insulating layer so as to be positioned facing the front control gate G. The back control gate BG_P is formed here by a P type doped region isolated from the base substrate by an N
 15 conductivity well C_N .

In Fig. 3b, an N channel transistor is illustrated, having a back control gate BG_N arranged in the base substrate under the insulating layer so as to be positioned facing the front control gate G. The back control gate BG_N is formed here by an
 20 N type doped region insulated from the base substrate by a P conductivity well C_P .

In Figs. 3a and 3b, the transistors are fully depleted SeOI transistors defined by a channel/body region that is thin enough to allow the field induced by the top control gate
 25 (under usual voltage conditions) to reach the buried oxide layer BOX.

A transistor for which the channel has a conductivity of type N and a back control gate of P conductivity, has a very high threshold voltage. This threshold voltage may then be
 30 reduced by applying a positive voltage on the back control gate. A transistor for which the channel has a conductivity of type N and a back control gate of conductivity N as for it has a nominal threshold voltage which may be reduced by applying a positive voltage on the back control gate.

This variation of the threshold voltage of the transistor via the back control gate may be formulated according to $V_{th} = V_{t0} - \alpha \cdot V_{BG}$, wherein V_{th} represents the threshold voltage of the transistor, V_{BG} the voltage applied to the back control gate, 5 V_{t0} the nominal threshold voltage (which may be shifted by the work function depending on whether a back control gate of type N or P is used), and α is a coefficient related to the geometry of the transistor.

The different writing, reading and retention operations 10 which may be applied by the nano-sense amplifier of Fig. 2 are described hereafter.

Reading

During an operation for reading the datum stored in a memory cell C_i , C_j connected to the local bitline LBL, and 15 selected by means of a line, a so-called wordline (not shown), the transistor T3 of the reading stage is the sensor, while the transistors T1 and T2 of the reading stage are blocked in order to avoid any corruption of the signal from the selected cell.

20 The inverter has a voltage of $V_{DD}/2$ (potentially floating voltage) on its input E.

A high voltage is applied to the back control gate of the P channel transistor T1 (V_{bg1} is for example in the high state V_{DD}) so as to set its threshold voltage above $V_{DD}/2$ and to 25 thereby block the transistor T1. If this does not prove to be sufficient, it is also possible to reduce the power supply potential V_{DD} of the transistor T2 during this reading operation, by typically using a voltage already existing in the circuit such as $V_{DD}/2$.

30 Moreover, a low voltage is applied to the back control gate of the N channel transistor T2 (V_{bg2} is for example in the low state GND) so as to set its threshold voltage above $V_{DD}/2$ and to thereby block the transistor T1. If this does not prove to be sufficient, it is also possible to increase the power

supply potential GND of the transistor T1 during this operation, for example by a few hundred millivolts, typically by using a voltage already existing in the circuit, such as VDD/2.

5 The selected cell to be read typically provides 200-300 mV if it contains a "1" and 0 mV if it contains a "0".

A high voltage is applied to the back control gate of the N channel sense transistor T3 (Vbg3 is for example in the high state VDD), in order to lower its threshold voltage below the
10 level provided at its gate by the local bitline LBL. The power supply potential PGND of the transistor T3 is here in the low state, for example GND.

If a "1" is present on the local bitline LBL, the sense transistor T3 is conducting, and the main sense amplifier MSA
15 may then either detect the current flowing in the main bitline MBL through the transistor T3, or sense a voltage level change on the main bitline MBL (which drops since T3 is conducting) if the node was initially floating.

If a "0" is present on the local bitline LBL, the sense
20 transistor T3 is blocked and the main sense amplifier MSA may then either detect zero current on the main bitline MBL, or sense a non-modified voltage level on the main bitline MBL.

Writing

During an operation for writing a datum stored in a memory
25 cell Ci, Cj connected to the local bitline LBL, and selected by means of word line, the transistor T3 of the reading stage is blocked in order to avoid any corruption of the signal from the main sense amplifier MSA via the main bitline MBL, while the transistors T1 and T2 of the reading stage conduct as
30 efficiently as possible this signal from the main bitline MBL towards the local bitline LBL.

In the following, a distinction is made between writing a "0" (Write 0 operation), and writing a "1" (Write 1 operation).

Write 0

The inverter stage has, via the main bitline MBL, a "1" at its input.

A low state is applied to the back control gate 1 of the P channel transistor T1 (Vbg1 is for example at GND) in order to lower its threshold voltage as much as possible (preferably below $VDD/2$, for example around 100-200 mV in absolute value). The transistor T1 is blocked insofar that it receives a high state on its front control gate. The power supply potential of the transistor T1 is maintained at VDD, so that other nano-sense amplifiers in parallel (i.e. sharing the same power supply potential) may perform writing of a "1" at the same time.

A high state is applied to the back control gate of the N channel transistor T2 (Vbg2 is for example at VDD), in order to lower its threshold voltage (preferably below $VDD/2$, for example around 100-200 mV in absolute value). The transistor T2 receives a high state on its front control gate and is therefore conducting. Insofar that its threshold voltage was lowered, the conduction level of transistor T2 is increased (or further, the size of the transistor T2 may be reduced if an identical conduction level is maintained).

A low state is applied to the back control gate of the N channel transistor T3 of the reading stage (Vbg3 is for example in the low state GND), and this so that it has a high threshold voltage, preferably above $VDD/2$. As the local bitline LBL is in the low state because of the action of the transistor T2 of the writing inverter, transistor T3 is blocked and therefore does not corrupt the "1" delivered on the main bitline MBL.

The power supply potential PGND of the transistor T3 is here in the low state, for example GND. This power supply potential PGND may however be increased in order to attain a value close to $VDD/2$ (below the threshold voltage of T3) in

order to ensure that the transistor T3 is actually transparent with respect to the main bitline MBL and does not generate conflicts therein, notably during transitions between the different operations.

5 Write 1

The inverter stage has, via the main bitline MBL, a "0" at its input, and the order of the operations of transistors T1 and T2 is inverted relatively to the Write 0 operation.

Thus a high state is applied to the back control gate of the N channel transistor T2 (Vbg2 is for example at VDD) in order to lower its threshold voltage as much as possible (preferably below $VDD/2$, for example around 100-200 mV in absolute value). The transistor T2 is blocked insofar that it receives a high state on its front control gate and the power supply potential of transistor T2 is maintained at GND, so that other nano-sense amplifiers in parallel (i.e. sharing the same power supply potential) may perform writing of a "0" at the same time.

A low state is applied to the back control gate of the P channel transistor T1 (Vbg1 is for example at GND), so that it has a low threshold voltage (preferably below $VDD/2$, for example around 100-200 mV in absolute value). The transistor T1 receives a high state on its front control gate and is therefore conducting. Insofar that its threshold voltage was lowered, the conduction level of the transistor T1 is increased (or further by retaining an identical conduction level, the size of transistor T1 may be reduced).

A low state is applied to the back control gate of the N channel transistor T3 of the reading stage (Vbg3 is for example in the low state GND), and this in order to increase its threshold voltage, preferably above $VDD/2$. The power supply potential PGND of the transistor T3 is here in the low state, for example GND.

As the local bitline LBL is in the high state because of the action of the transistor T1 of the writing inverter, transistor T3 is conducting, but does not corrupt the "0" delivered on the main bitline MBL since its source and its drain are both at the same value (low state GND). The power supply potential PGND may however be increased in order to attain a value close to $VDD/2$ (below the threshold voltage of T3) in order to ensure that the transistor T3 is actually transparent relatively to the main bitline MBL, and does not generate conflicts therein, notably during transitions between the different operations.

Retention

It is sought to avoid any static current, and thereby block the three transistors by ideally minimizing leakages.

In a preferential embodiment, the retention operation is applied by preparing the reading or writing operation which will follow, and this by discharging the local bitline to the low state GND which is a prior condition for writing.

The inverter has a voltage of $VDD/2$ (potentially floating voltage) on its input.

A high voltage is applied to the back control gate of the P channel transistor T1 (Vbg1 is for example in the high state VDD) so as to set its threshold voltage above $VDD/2$ and to thereby block the transistor T1. If this does not prove to be sufficient, it is also possible to reduce the power supply potential VDD of the transistor T2 during this retention operation.

Moreover, a high voltage is applied to the back control gate of the N channel transistor T2 (Vbg2 is for example in the high state VDD) so as to lower its threshold voltage (to about 100-150 mV). The transistor T2 is conducting, which allows the local bitline LBL to be discharged and to be forced to the low state GND, thereby preparing it for the next

access. As the circuit is in retention, no cell is selected and there is therefore no current.

Insofar that the local bitline LBL is in the low state GND, transistor T3 is blocked and therefore no current flows
5 through it. A low voltage may be applied to the back control gate of the transistor T3 (Vbg3 is for example in the low state GND), which increases its threshold voltage and allows minimization of the leakages.

In Fig. 4a, a possible topology of the nano-sense amplifier
10 of Fig. 2 made on two columns is illustrated. Only three metals are necessary: Metall for the interconnections of the transistors, Metal2 for the distribution of the power supply potentials PGND, GND and VDD, Metal3 for the main bitline MBL.

The width of the nano-sense amplifier corresponds to two
15 metal lines or two columns of cells in the case of a DRAM.

It will be noted that the three back control gates are active and may consequently be subject to RC delays. It is then possible to choose to regenerate them periodically, ideally at the same repeating frequency as the word line
20 drivers.

The memory cell, as for it, requires a Metall line for attaching it to the local bitline and a Metal2 line for attaching it to the word line. This leaves the Metal3 line free for letting through the main bitline MBL.

Insofar that each column of the memory matrix requires a
25 nano-sense amplifier, a possible organization may consist of arranging the nano-sense amplifiers head-to-tail as this is illustrated in Fig. 4b, a first nano-sense amplifier having on its input terminal the signal MBL_E and addressing via its local bitline a series of cells along a first (even) column, the
30 other nano-sense amplifier having on its input terminal the signal MBL_O and addressing via its local bitline a series of cells along a second (odd) column directly following adjacent to the first column in the memory matrix.

Preferential conditions for the operation of the nano-sense amplifier of Fig. 2 are threshold voltages of the transistors, close to $V_{DD}/2$ and a variation of these threshold voltages by biasing the back control gates in a range approximately
5 comprised between 100 mV and $V_{DD}/2 + 150$ mV.

These conditions involve lithography levels beginning at 55-45 nm (V_{DD} of the order of 1 V, threshold voltage of the order of 350 mV and which decreases with the scale effect; the node 45 nm having been attained in 2007-2008), and thicknesses
10 of the order of 2-20 nm for the thin layer of the SeOI substrate and of the order of 5-50 nm for the buried insulating layer of the SeOI substrate.

It will be noted that the nano-sense amplifier is mainly developed for DRAM memories. However, it operates with any
15 type of RAM memory (SRAM, PCRAM, CBRAM, ZRAM), as well as with flash memories (with, in this case, the requirement of maintaining relatively high voltages during the writing and deleting (erasing) operations; this may for example be achieved by using two transistors in series rather than a
20 single one for the transistors T1-T3).

As this was seen earlier, in the case of the DRAM, the nano-sense amplifier performs reading and writing operations but does not perform the refreshing/restoring operation; the latter being performed by the main sense amplifier MSA, a
25 possible embodiment of which is described hereafter in connection with Fig. 5. A DRAM cell actually requires being periodically refreshed in order to compensate charge losses and being restored after each read access.

The main sense amplifier MSA is connected to nano-sense
30 amplifiers via a main bitline MBL on the one hand, and to the peripheral circuits via a data line DL on the other hand.

The amplifier MSA comprises, in series between the potentials V_{DD} and GND, three transistors T6-T8. Transistor T6 is a P channel transistor, the front control gate of which is

controlled by a signal $\Phi 2$. Transistor T7 is an N channel transistor, the front control gate of which is connected to the main bitline MBL. Transistor T8 is a P channel transistor, the front control gate of which is controlled by a signal $\Phi 3$.

5 The amplifier MSA further comprises an N channel transistor T10 between the main bitline MBL and a power supply potential at $V_{DD}/2$ and the front control gate of which is controlled by a signal $\Phi 1$.

10 The amplifier MSA also comprises a P channel transistor T9 between the main bitline MBL and a power supply potential $V+$ and the front control gate of which is connected to the node A corresponding to the middle point of transistors T6 and T7.

As this will be explained in more detail subsequently, the whole of the transistors T6-T10 forms a stage for amplifying the signal delivered by a cell (on the local bitline LBL towards a nano-amplifier) during a reading operation.

15 The node A is connected to a high impedance inverter HZ1, the application of the inversion function of which is controlled by a control signal $\Phi 4$ (in Fig. 5, $\Phi 4B$ designates the complementary of $\Phi 4$).

20 The output of the high impedance inverter HZ1 is looped back onto the main bitline MBL.

As this will be detailed subsequently, the inverter HZ1 thus forms a switchable high impedance inverter stage for sending back the amplified signal on the main bitline following a reading operation.

A high impedance stage HZ2, controlled by a signal $\Phi 5$ ($\Phi 5B$ designating the complementary of $\Phi 5$), enables the data line DL to be connected to the main bitline MBL.

30 Initialization of the MSA amplifier is as follows.

$\Phi 2$ is set to "1", while $\Phi 3$ is set to "0". As the transistors T6 and T8 are blocked, the set T6, T7 and T8 is floating.

The power supply potential $V+$ of the transistor T9 is moreover set to $VDD/2$.

Both stages HZ1 and HZ2 are floating, while the transistor T10 will, as for it, pre-charges the main bitline MBL to $VDD/2$, via the application of a negative pulse by the control signal $\phi 1$.

An operation for reading a "1" and for restoring this "1" implemented by the MSA amplifier is the following.

The read cell delivers a "1" on the local bitline LBL. The transistor T3 of the nano-sense amplifier nSA discharges the main bitline MBL to the low state GND. Transistor T7 is then blocked.

The power supply potential $V+$ of the transistor T9 is increased from $VDD/2$ to VDD .

The control signal $\phi 3$ of the gate of the transistor T8 is set to the high state VDD , so as to make the transistor T8 conducting. As the transistor T7 is blocked, it will cut the path towards the node A.

The control signal $\phi 2$ of the gate of the transistor T6 is set to the low state GND, so as to make the transistor T6 conducting. The potential of the node A then increases to VDD , which causes blocking of the transistor T9.

The control signal $\phi 4$ will make the HZ1 stage conducting. The latter conducts the main bitline MBL to the low state GND (by inversion of the high state of the node A), and sends back this low state to the nano-amplifier nSA, which will then rewrite a "1" (cf. previous discussion of a Write 1 writing operation by the nano-amplifier nSA).

The control signal $\phi 5$ will make the HZ2 stage conducting. The latter will then deliver the signal of the main bitline (stemming from the HZ1 stage) to the data line DL for processing by the input/output peripheral circuits (the data line DL is floating in order to accept the signal from the main bitline MBL).

An operation for reading a "0" and for restoring this "0" performed by the amplifier MS A is as follows.

The cell read delivers a "0" on the local bitline LBL. The transistor T3 of the nano-sense amplifier nSA remains blocked
5 and the main bitline MBL remains at $VDD/2$. Transistor T7 is then conducting.

The potential $V+$ of the transistor T9 is increased from $VDD/2$ to VDD .

The control signal $\Phi3$ of the gate of the transistor T8 is
10 set to the high state VDD , so as to make transistor T8 conducting.

The control signal $\Phi2$ of the gate of the transistor T6 is set to the low state GND , so as to make transistor T6 conducting.

15 The transistor T6 is a weak transistor as compared with the serial association of transistors T7 and T8, the potential of the node A drops to "0". This leads the transistor T9 to a conducting (also weak) state, this has the consequence of leading the local bitline to VDD (from $V+$). Next, the
20 transistor T7 will lead the node A closer to the low state GND .

The control signal $\Phi4$ will make the stage HZ1 conducting. The latter leads the main bitline MBL to the high state VDD (by inversion of the low state of the node A), and sends back
25 this high state to the nano-amplifier nSA, which will then rewrite a "0" (cf. previous discussion of a Write 0 writing operation by the nano-amplifier nSA).

The control signal $\Phi5$ will make the HZ2 stage conducting. The latter will then deliver the signal from the main bitline
30 MBL (stemming from the HZ1 stage) to the data line DL for processing by the input/output peripheral circuits (the data line DL is floating in order to accept the signal from the main bitline MBL).

A writing operation performed by the MSA amplifier is the following. From the point of view of the MSA amplifier, this operation is similar to a reading operation. The only difference stems from the initial conduction of the main
5 bitline MBL because of its powering from the data line via the HZ2 stage.

The following steps are the same, except with regard to the last point (transfer to the data line) which is out of context for a writing operation.

10 It will be noted that in Fig. 5, the different transistors T6-T10 were illustrated with back control gates, each being respectively connected to the corresponding front control gate. This illustration was only made for sake of clarity, and it will be understood that in practice the back control gates
15 are biased in the most suitable way for increasing the performances of the MSA amplifier.

As examples, the back control gates of transistors T7 and T9 may be biased so that these transistors have relatively high threshold voltages and are thus weaker than the other
20 transistors.

In the foregoing, an embodiment of the MSA amplifier is reported within the scope of an application to DRAM memories. It will be retained that because of problems of noise and of variability encountered with SRAM memories, the MSA amplifier
25 described earlier may advantageously be used therein for reinforcing the quality of the signal and for increasing the security of such circuits.

An MSA amplifier of the same type may be used with RAM memories (PCRAM, CBRAM, FBC-DRAM, etc.). It may also be used
30 with flash memories, while bringing attention to the requirement of supporting the relatively high voltages required by the memory cell.

A nano-sense amplifier μ SA according to a second embodiment of the invention is illustrated in Fig. 6.

This second embodiment proves to be advantageous in that, in addition to the writing and reading operations, the nano-sense amplifier nSA is also capable of ensuring refreshing/restoring operations. It is therefore not necessary
5 to resort to a main sense amplifier MSA, which is notably expressed by a gain in surface area.

Further, the metal (Metal3), which is used as a main bitline MBL within the scope of the first embodiment, may under certain circumstances prove to be too wide relatively to
10 the dimensions of the memory cell. The second embodiment allows this drawback to be overridden.

The nano-sense amplifier nSA of Fig. 6 includes a writing stage formed with a CMOS inverter formed with a P channel transistor T1, the source of which is connected to the
15 potential V1 and an N channel transistor T2, the source of which is connected to the potential V2.

The input of the inverter is connected to a node N1, which node N1 is indirectly connected to the input terminal of the nano-sense amplifier (which is intended to be connected to a
20 main bitline MBL).

The output of the inverter is connected to the output terminal of the nano-sense amplifier (which is intended to be connected to a local bitline LBL).

The nano-sense amplifier nSA of Fig. 6 further includes a
25 reading stage comprising a transistor T3, the gate of which is connected to the output of the inverter (and therefore also to the local bitline LBL) and the drain of which is connected to the input of the inverter (node N1).

The reading stage comprises an additional transistor T4,
30 complementary to the sense transistor, the additional transistor T4 and the sense transistor T3 forming a CMOS inverter, the input of which is connected to the output of the writing stage (and thus to the local bitline LBL) and the output of which is connected to the input of the inverter of

the writing stage (and therefore to node N1). The transistors T4 and T3 are in series between power supply potentials V4 and V3. The writing T1, T2 and reading T3, T4 inverters are thus cross coupled, which will allow the nano-sense amplifier to perform the refreshing/restoring operations alone.

To finish, the nano-sense amplifier nSA of Fig. 6 comprises a decoding stage formed by a transistor T5 (N channel transistor in the illustrated example) the gate of which is controlled by a decoding signal YDEC, the source of which is connected to a main bitline MBL and the drain of which is connected to the node N1.

Within the scope of a preferred embodiment, the nano-sense amplifier nSA of Fig. 6 is made on a semiconductor-on-insulator substrate and each of the transistors T1-T5 has a back control gate Vbg1-Vbg5 capable of being biased in order to modify the threshold voltage of the corresponding transistor.

The different writing, reading and retention operations are described hereafter which may be implemented by the nano-sense amplifier of Fig. 6.

Retention

It is sought to avoid any static current, and to prepare the reading or writing operation which will follow by discharging the local bitline LBL to the low state GND.

The power supply potential V1 is in a low state, while the back control gate of the transistor T1 is in the high state. The threshold voltage of the transistor T1 is then increased. The power supply potential V1 is in a low state (typically 0 V), while the back control gate of the transistor T2 is in the high state. The threshold voltage of the transistor T2 is then lowered.

The power supply potential V3 is in the high state (VDD), while the back control gate of the transistor T3 is in the high state. The threshold voltage of the transistor T3 is then

lowered. The power supply potential V4 is in the high state (VDD) while the back control gate of the transistor T4 is in the high state. The threshold voltage of the transistor T4 is then increased.

5 This allows the node N1 to be led to the high state, and to lead the local bitline LBL to the low state.

No decoding signal YDEC is applied to the gate of the transistor T5 (the decoder is not active during a retention operation), and the transistor T5 is blocked.

10 Reading (from the retention)

The initial conditions are the following.

The local bitline LBL is in the low state.

The power supply potential V1 is in a low state, while the back control gate of the transistor T1 is in the high state.

15 The threshold voltage of the transistor T1 is then increased, the transistor T1 then being blocked.

The power supply potential V2 is in the high state, while the back control gate of the transistor T2 is in the low state. The threshold voltage of the transistor T2 is then
20 increased, the transistor T2 then being blocked.

The power supply potential V3 is in the low state (0 V), while the back control gate of the transistor T3 is in the high state. The threshold voltage of the transistor T3 is then lowered, while the transistor T3 is blocked. The power supply
25 potential V4 is at VDD/2 or more, while the back control gate of the transistor T4 is in the low state. The threshold voltage of the transistor T4 is then reduced. The result is that the node N1 is led to VDD/2 or more.

The transistor T5 is, as for it, always blocked insofar
30 that no decoder is yet active.

Reading continues in the following way following the opening of the word line which will select the cell.

In the case when the datum to be read is a "0", the local bitline LBL remains in the low state. In this case indeed, the

local bitline LBL, as well as the cell, are in the low state (GND) so that when the transfer of charges from one to the other is allowed, nothing occurs since the balance already exists.

5 The power supply potential V1 is in the low state, while the back control gate of the transistor T1 is in the high state. The threshold voltage of the transistor T1 is then increased, the transistor T1 then being blocked. The power supply potential V2 is in the high state, while the back control gate of the transistor T2 is in the low state. The threshold voltage of the transistor T2 is then increased, the transistor T2 then being blocked.

15 The power supply potential V3 is in the low state (0 V), while the back control gate of the transistor T3 is in the high state. The threshold voltage of the transistor T3 is then lowered, while the transistor T3 is blocked since its gate connected to the local bitline LBL is at 0 V.

20 The power supply potential V4 is increased up to VDD, while the back control gate of the transistor T4 is in the low state. The threshold voltage of transistor T4 is then increased. The result is that the node N1 is led to VDD.

25 Next, V1 is increased, while the back control gate of the transistor T1 is brought to the low state. The threshold voltage of the transistor T1 is lowered; T1 being blocked insofar that its gate is connected to the node N1 which is at VDD. The transistor T2, the gate of which is also connected to the node N1 at VDD, as for it, is conducting. The local bitline LBL is then led to 0 V so that the datum may be refreshed.

30 The transistor T5 is, as for it, blocked insofar that no decoder is yet active. The transistor T5 is then made conducting when the potential of the node N1 is stable at VDD.

In the case where the datum to be read is a "1", the local bitline LBL is initially in the low state. In this case

indeed, upon opening the word line, the balance is not attained. Charges will then circulate between the cell and the local bitline LBL so as to equalize the voltages. The final voltage corresponds to a ratio between the capacitances of the local bitline LBL and of the cell (the longer the local bitline LBL, the weaker will be the signal). This voltage will be read by the nano-sense amplifier nSA.

The power supply potential V1 is in the low state, while the back control gate of the transistor T1 is in the high state. The threshold voltage of the transistor T1 is then increased, the transistor T1 then being blocked. The power supply potential V2 is in the high state, while the back control gate of the transistor T2 is in the low state. The threshold voltage of transistor T2 is then increased, the transistor T2 then being blocked.

The power supply potential V3 is in the low state (0 V), while the back control gate of the transistor T3 is in the high state. The threshold voltage of the transistor T3 is then lowered, while the transistor T3 is conducting since its gate, connected to the local bitline LBL, has a greater potential than the threshold voltage of T3.

The power supply potential V4 is increased up to VDD, while the back control gate of the transistor T4 is in the low state. The threshold voltage of the transistor T4 is then increased. The result is that the node N1 is led to 0 V.

Next, V1 is increased, while Vbgl is brought to the low state. The threshold voltage of the transistor T1 is lowered; T1 being conducting insofar that its gate is connected to the node N1 which is at 0 V. The transistor T2, the gate of which is also connected to the node N1 at 0 V, is, as for it, blocked. The local bitline LBL is then led to VDD so that the datum may be refreshed.

The transistor T5 is, as for it, blocked insofar that no decoder is yet active. The transistor T5 is then made

conducting when the potential of the node N1 is stable at 0 V. The signal present on the N1 node is then transferred onto the main bitline MBL.

"Write 0" writing (from initial conditions)

5 The main bitline MBL is in the high state.

The transistor T5 is conducting, bringing the node N1 to the high state.

It will be noted that the transistor T5 may however have a substrate effect and not pass the whole of the signal on N1. This however is not very important insofar that the 4 transistors T1-T4 will amplify and restore the signal to "clean" logic levels.

V4 is brought from VDD/2 or more to the low state. The transistor T4 is then blocked.

15 V1 is brought from the low state to VDD, while Vbg1 is brought from the high state to the low state. The threshold voltage of T1 is low: T1 is blocked (N1 gate in the high state). Transistor T2 is conducting, while transistor T3 is blocked which allows the local bitline LBL to be brought to 0 V.

Next, V4 is brought from the low state to VDD. The information is now stable in the amplifier formed by the transistors T1-T4.

"Write 1" writing (from initial conditions)

25 The main bitline MBL is in the low state.

The transistor T5 is conducting, leading the node N1 to the low state.

V4 is brought from VDD/2 or more to the low state. The transistor T4 is then blocked.

30 V1 is brought from the low state to VDD, while Vbg1 is brought to the low state. The threshold voltage of T1 is low: T1 is conducting which allows the local bitline LBL to be brought to VDD.

The transistor T2 is blocked, while the transistor T3 is conducting.

Next, V4 is brought from the low state to VDD. T4 is then blocked.

5 Various physical considerations are presented hereafter as regards the nano-sense amplifier of Fig. 6.

As shown earlier, the local bitline LBL is pre-charged to GND while the main bitline LBL is pre-charged to VDD (or to the same level as V4 during a reading operation).

10 The transistors T1 and T2 are as small as possible, their performances may be increased by virtue of their back control gate, and this both when they are conducting and when they are blocked.

The transistor T3 should drive the node N1 and then the
15 main bitline MBL during a reading operation; T3 is as small as possible in order to minimize the charge of its front gate on the local bitline NBL and its performances may be increased by its back control gate for rapidly conducting N1 and MBL.

With the transistor T4, it is possible to secure the
20 refreshing operations and to override the body effect of the transistor T5. It is selected to be as small as possible.

The transistor T5 should approximately have the same size as the transistor T3, if necessary with the same choices as regards the back control gate. The transistor T5 is not
25 critical per se but requires about 8 to 128 buses in order to be able to operate as a decoder.

Eventually, the nano-sense amplifier of Fig. 6 is wider than that of Fig. 2, essentially because of the decoding bus YDEC. This bus may however be shared between two adjacent
30 matrices if these matrices store different data bits.

The nano-sense amplifier of Fig. 6 however does not require any main sense amplifier, by which the increase in size may be compensated, as far as the YDEC bus has a reasonable size.

With it, it is further possible to override the limitations due to possible congestion of the Metal3 in the case of the first embodiment, insofar that the main bitline is decoded.

It has to be noted that that the nano-sense amplifier according to the first embodiment presented on figure 2 is based on a "half latch" structure, while the nano-sense amplifier according to the second embodiment presented on figure 6 is constructed on the "full latch" structure. The proposed schematics are the preferred schematics but it is clear that these structures can be built on other schematics comprising more transistors either in the latching part or in the decoder or precharged parts.

The performance of the nano-sense amplifiers (both embodiments) is obtained through the use of a back gate on SeOI. However, it is clear that the principle can be also applied on any type of double-gate transistors having a second gate such as the multi-gate finFets for instance.

It has furthermore to be noted that the simplification of the schematics, through the reduction in number of devices and in their reduced sizes, makes the application of the nano-sense amplifier according to the invention very attractive to other (e.g. matrix array like) circuits than memories such as imagers or analogue converters for instance. In another application, the nano-sense amplifier can be used as a bus regenerator.

As indicated earlier, the nano-sense amplifier provides its own reference (threshold voltage of transistor T3) and therefore it does not need the presence of a second matrix to serve as a reference like on a conventional architecture represented on figure 7 (indeed, conventional systems need a reference Matrix for compensation of unknown offset values that are due to irregularities in the fabrication process).

In addition, as it will be discussed in further details below, since the nano-sense amplifier is very compact, it can be arranged in a periodicity meeting the minimum pitch of memory

arrays that can be attained with state of the art lithographic techniques. The additional area between adjacent memory arrays can further be used for other circuits, on-pitch or non-pitched circuits.

5 The nano-sense amplifier having a reduced width, there is no need resorting to a stacking technique, a so-called "staggering" technique, for several conventional sense amplifier circuits behind each other in order to address several adjacent columns of memory cells while taking into
10 account the pitch difference in between the sense amplifier and the cells).

As shown on the example of figure 8a, rather than making use of a single bank of 128 staggered conventional sense amplifiers (fig. 7), it is possible to make use of two banks
15 of 64 nano-sense amplifiers (here the nano-sense amplifiers are said to be effectively "on-pitch" in that they may be as large as 2 times the pitch of the cells; it may also be possible to provide perfectly "on-pitch" nano-sense amplifiers, as large as the cells).

20 It further becomes possible to locally decode these nano-sense amplifiers by placing the logic of the column decoder on the "free" edge of a bank of nano-sense amplifiers bank as it is represented on Fig.8a and Fig. 8b. For instance, if the architecture of the conventional circuit initially required
25 128 busses to decode each sense amplifier within the bank like on Fig. 7, then the new organization of the nano-sense amplifier and its decoder requires only 16 busses (8 most significant bits - MSB - and 8 least significant bits - LSB) to fulfil the same functionality.

30 This feature is particularly advantageous as the nano-sense amplifier occupies significantly less area than the conventional sense amplifier and therefore needs only few metal busses over its area to decode the bank. Most of the column decoder circuit can be integrated into the on-pitch
35 area and no longer need to be in the periphery of the memory, i.e. becomes an on-pitch circuit and thus becomes itself very small in area and has a much lower power consumption and a higher speed.

The integration of other circuit blocks into the on-pitch area, next to nano-sense amplifier banks can be extended for instance to the address buffers themselves as shown on Fig.9a and 9b. The address buffers can be distributed under the column decoder and the requirement for busses decreases to two groups of 3 addresses in the given example. This second row of function may be not "on pitch" as the number of circuits does not correspond to the array pitch.

Other functionalities can also be placed in a similar manner next to the nano-sense amplifiers banks, such as, for example voltage generators, pumps, analog functions, or redundancy control. There is virtually no constraint and the absence of bitline reference in the nano-sense amplifier provides flexibility to optimize all the peripheral functions of the circuit.

It has been shown that placing the column decoder next to the nano-sense amplifier bank decreases the amount of busses needed by a factor 4. Furthermore, the remaining busses are shorter as they are integrated in the on-pitch area. The immediate consequence of a local decoding approach is a gain in area as the number of bus drivers (buffers) may be reduced thanks to the proximity, and they are smaller as well as less loaded. The power consumption is also reduced for the same reasons. Speed can be increased at no extra cost because of less parasitics.

Placing other circuits can be also beneficiary for the circuit. For instance, having this opportunity to place the different voltage generators next to their load (V_{plate} , V_{BLH} , etc...) reduces or cancels most of the sources of disturbances (ohmic losses, capacitive couplings) for these functions.

It is clear that the organization shown in figures 8a-8b and 9a-9b are examples, and the person skilled in the art would contemplate that other block organizations are possible. In particular the sense amplifiers could be organized on only

one side of the array, or two blocks can be disposed on either side of the array to address the odd and even bitlines, respectively.

CLAIMS

1. A sense amplifier (nSA) of a series of cells (Ci, Cj)
5 of a memory, including:
- a writing stage comprising a CMOS inverter (T1-T2),
the input of which is directly or indirectly connected to
an input terminal of the sense amplifier, and the output
of which is connected to an output terminal of the sense
10 amplifier intended to be connected to a local bitline
(LBL) addressing the cells of said series,
- a reading stage comprising a sense transistor (T3),
the gate of which is connected to the output of the
inverter and the drain of which is connected to the input
15 of the inverter,
wherein each of the transistors has a back control gate
formed in a base substrate below the channel and capable
of being biased in order to modulate the threshold
voltage of the transistor.
- 20 2. The sense amplifier according to claim 1, wherein the
input of the writing stage is directly connected to the
input terminal of the inverter, said input terminal being
intended to be connected to a main bitline (MBL) which
25 will address a plurality of sense amplifiers in parallel.
3. The sense amplifier according to the preceding claim,
characterized in that it is made on a semiconductor-on-
insulator substrate comprising a thin layer of
30 semiconducting material separated from the base substrate
by an insulating layer.

4. A matrix array of cells according to claim 15, wherein the sense amplifiers are arranged in one or more sense amplifier banks and further comprising a column decoder placed next to each bank and configured to drive said bank directly.
5. The sense amplifier according to any of the preceding claims, wherein one or more of the transistors are multigate transistors.
6. The sense amplifier according to claim 1, wherein the reading stage comprises an additional transistor (T4) complementary to the sense transistor, the additional transistor and the sense transistor forming a CMOS inverter, the input of which is connected to the output of the reading stage and the output of which is connected to the input of the inverter of the writing stage.
7. The sense amplifier according to claim 6, wherein the input of the writing stage is indirectly connected to the input terminal of the inverter via a decoding stage comprising a transistor (T5), the drain of which is connected to the input terminal of the sense amplifier and the source of which is connected to the input of the writing stage.
8. A main sense amplifier (MSA) intended to be connected via a main bitline (MBL) to a plurality of sense amplifiers according to any of claims 1 to 7, characterized in that it comprises a step for amplifying

the signal (T6-T10) delivered by a cell during a reading operation, and a switchable high impedance inverter stage (HZ1) for sending back the amplified signal onto the main bitline following the reading operation.

5

9. A sense amplifier (nSA) including:

- a writing stage comprising a CMOS inverter (T1-T2), the input of which is directly or indirectly connected to an input terminal of the sense amplifier, and the output of which is connected to an output terminal of the sense amplifier intended to be connected to a local line,

- a reading stage comprising a sense transistor (T3), the gate of which is connected to the output of the inverter and the drain of which is connected to the input of the inverter;

and wherein one or more transistors of the writing stage and of the reading stage are independent double gate transistors,

each of the transistors has a back control gate formed in the base substrate below the channel and capable of being biased in order to modulate the threshold voltage of the transistor.

10. A matrix array of cells connected to local bitlines and to wordlines, characterized in that it includes sense amplifiers according to any of claims 1 to 7, or to claim 9.

11. A matrix array of cells according to claim 10, wherein the sense amplifiers are made on a semiconductor-on-insulator substrate.

12. A matrix array of cells according to claim 11, further comprising logic circuits placed next to the sense amplifier banks or the column decoder.

13. A matrix array of cells according to claim 12, wherein the sense amplifiers are arranged in a non-staggered fashion in one or more sense amplifier banks and wherein a column decoder is placed next to each bank and configured to drive said bank directly.

14. A memory including a matrix array of cells according to any one of claims 14 to 17.

15. A method for controlling a sense amplifier according to any of claims 1 to 7, wherein the biasing of one of the gates of the transistors is modified during operations for writing, reading and retaining data in the cells of said series.

16. A method for controlling a sense amplifier according to any of claims 1 to 7, wherein the biasing of the back control gates is modified during operations for writing, reading and retaining data in the cells of said series.

17. The method according to claim 15 or 16, wherein during a reading operation, the threshold voltage of the transistors of the writing stage is increased and the threshold voltage of the transistor(s) of the reading

stage is reduced by controlling the biasing of the back control gates thereof.

18. The method according to claim 15 or 16, wherein during
5 a writing operation, the threshold voltage of the transistors of the writing stage is reduced and the threshold voltage of the transistor(s) of the reading stage is increased by controlling the biasing of the back control gates thereof.

10

19. The method according to claim 15 or 16, wherein during
a retention operation, the local bitline is discharged
via one of the transistors of the writing stage, for
which the biasing of the back control gate is controlled
15 so as to lower its threshold voltage.