

(19)日本国特許庁(JP)

(12)公開特許公報(A)

(11)公開番号

特開2022-163107

(P2022-163107A)

(43)公開日 令和4年10月25日(2022.10.25)

(51)国際特許分類	F I
H 0 1 L 27/11578(2017.01)	H 0 1 L 27/11578
H 0 1 L 21/336(2006.01)	H 0 1 L 29/78 3 7 1
H 0 1 L 27/11551(2017.01)	H 0 1 L 27/11551
H 0 1 L 27/11526(2017.01)	H 0 1 L 27/11526
H 0 1 L 27/11573(2017.01)	H 0 1 L 27/11573
審査請求 有 請求項の数 44 O L (全43頁) 最終頁に続く	

(21)出願番号	特願2022-121900(P2022-121900)	(71)出願人	518109985
(22)出願日	令和4年7月29日(2022.7.29)		サンライズ メモリー コーポレイション
(62)分割の表示	特願2021-12587(P2021-12587)の分割		アメリカ合衆国カリフォルニア州 9 4 5
原出願日	平成28年11月4日(2016.11.4)		3 8 ・フレモント・レイクビュー プールバード 4 6 8 3 1
(31)優先権主張番号	62/260,137	(74)代理人	110001379
(32)優先日	平成27年11月25日(2015.11.25)		特許業務法人 大島特許事務所
(33)優先権主張国・地域又は機関	米国(US)	(72)発明者	ハラリ、エリ
(31)優先権主張番号	62/363,189		アメリカ合衆国カリフォルニア州 9 5 0
(32)優先日	平成28年7月15日(2016.7.15)		7 0 ・サラトガ・ヒル アベニュー 2 0
(33)優先権主張国・地域又は機関	米国(US)		2 3 8
(31)優先権主張番号	15/220,375		
(32)優先日	平成28年7月26日(2016.7.26)		
	最終頁に続く		

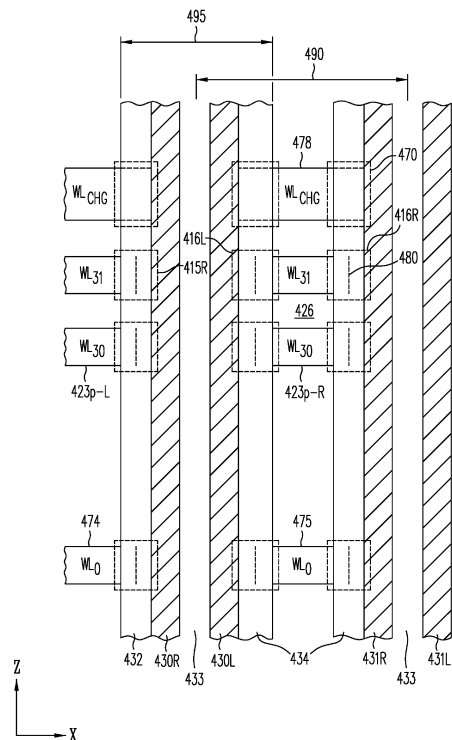
(54)【発明の名称】 3次元垂直NORフラッシュ薄膜トランジスタストリング

(57)【要約】 (修正有)

【課題】小さい読み出しレイテンシを有し、読み出し妨害及びプログラム妨害状態から受ける影響を低減し、1ビット当たりのコスト低くするメモリ構造を提供する。

【解決手段】メモリ構造は、半導体基板の上に形成された半導体材料の複数のアクティブ列430L、430R、431L、431Rを含む。各アクティブ列は、半導体基板の平坦な表面に直交する第1の方向に沿って延在し、かつ、第1の高濃度ドーブ領域、第2の高濃度ドーブ領域及び第1の高濃度ドーブ領域及び第2の高濃度ドーブ領域の両方に隣接する1つ以上の低濃度ドーブ領域を含む。アクティブ列は、2次元アレイに配置され、各アクティブ列における1つ以上の表面上に設けられた電荷トラップ層432、434と、各々が第3の方向に沿って長手方向に延びる複数の導体を含む。

【選択図】図4B



【特許請求の範囲】

【請求項 1】

メモリ構造であって、

平坦な表面を有し、メモリ回路をサポートするための回路が形成された半導体基板と、前記半導体基板の上に形成された半導体材料の複数のアクティブ列であって、各アクティブ列は、前記半導体基板の平面に直交する第 1 の方向に沿って延在し、かつ第 1 の高濃度ドーパ領域、第 2 の高濃度ドーパ領域、及び前記第 1 の高濃度ドーパ領域及び前記第 2 の高濃度ドーパ領域の両方に隣接する 1 つ以上の低濃度ドーパ領域を含み、前記アクティブ列は、第 2 の方向に沿って延在する複数行のアクティブ列と、第 3 の方向に沿って延在する前記複数行のアクティブ列とを有する 2 次元アレイに配置され、前記第 2 の方向及び前記第 3 の方向は、前記半導体基板の前記平面に平行である、該複数のアクティブ列と、各前記アクティブ列における 1 つ以上の表面上に設けられた電荷トラップ材料と、

10

複数の導体のスタックであって、前記複数の導体のスタックの各々は、第 3 の方向に沿って長手方向に延びる 2 列の前記アクティブ列の間に設けられ、前記電荷トラップ材料及び前記複数の導体は、全体で、第 1 の方向に沿って設けられた複数の可変閾値薄膜トランジスタを形成する、該複数の導体のスタックとを備え、

前記複数の可変閾値薄膜トランジスタの各々は、(i)ワード線を提供し、各可変閾値薄膜トランジスタのゲート端子として機能する、前記導体のスタックにおける前記導体のなかの関連する 1 つの導体と、(i i)各可変閾値薄膜トランジスタのチャネルとして機能する、前記アクティブ列の 1 つの前記低濃度ドーパ領域の部分と、(i i i)前記低濃度ドーパ領域の前記部分と前記導体との間の電荷トラップ材料と、(i v)それぞれビット線及びソース線を提供し、各可変閾値薄膜トランジスタのドレイン端子及びソース端子として機能する、前記アクティブ列の前記第 1 の高濃度ドーパ領域及び第 2 の高濃度ドーパ領域とを含むことを特徴とするメモリ構造。

20

【請求項 2】

請求項 1 に記載のメモリ構造であって、

読み出し動作の開始時に、前記ソース線は所定の電圧にプリチャージされ、

前記ソース線は、読み出し動作中に前記ソース端子と前記ゲート端子との間の少なくとも所定の電圧差を維持するのに十分な容量を提供することを特徴とするメモリ構造。

30

【請求項 3】

請求項 1 に記載のメモリ構造であって、

前記アクティブ列は、絶縁誘電体材料またはエアギャップによって互いに分離されることを特徴とするメモリ構造。

【請求項 4】

請求項 1 に記載のメモリ構造であって、

前記導体のスタックの各々のなかの前記複数の導体は、絶縁誘電体材料またはエアギャップによって互いに絶縁されていることを特徴とするメモリ構造。

【請求項 5】

請求項 1 に記載のメモリ構造であって、

各アクティブ列に関連する前記可変閾値薄膜トランジスタは、1 つ以上の N O R 薄膜トランジスタストリングに並列に構成されることを特徴とするメモリ構造。

40

【請求項 6】

請求項 2 に記載のメモリ構造であって、

前記アクティブ列に関連する全ての前記可変閾値薄膜トランジスタが非導通のとき、前記ソース線がフロート状態になり、前記容量は、前記アクティブ列の前記可変閾値薄膜トランジスタに仮想電圧源を提供することを特徴とするメモリ構造。

【請求項 7】

請求項 6 に記載のメモリ構造であって、

メモリ回路をサポートするための前記回路は、前記可変閾値薄膜トランジスタの 1 つに選択的に接続されて、読み出し動作中に前記可変閾値薄膜トランジスタが導通することに

50

起因する電圧降下を検出するための電圧降下検出器を含むことを特徴とするメモリ構造。

【請求項 8】

請求項 6 に記載のメモリ構造であって、

前記容量は、1 つ以上の導体による寄与を含むことを特徴とするメモリ構造。

【請求項 9】

請求項 6 に記載のメモリ構造であって、

前記アクティブ列の 1 つ以上の可変閾値薄膜トランジスタが、前記容量を所定の電圧に充電する専用のプリチャージトランジスタとして機能すること特徴とするメモリ構造。

【請求項 10】

請求項 6 に記載のメモリ構造であって、

プログラミング動作中または読み出し動作中に、前記仮想電圧源は、導電状態の可変閾値薄膜トランジスタに電流を供給することを特徴とするメモリ構造。

10

【請求項 11】

請求項 1 に記載のメモリ構造であって、

グローバルビットまたはソース線を更に備え、各グローバルビットまたはソース線は、前記第 2 の方向に沿って延在する導体を含み、前記グローバルビットまたはソース線は、各アクティブ列のソース線またはビット線を前記半導体基板内の前記回路に選択的に相互接続することを特徴とするメモリ構造。

【請求項 12】

請求項 11 に記載のメモリ構造であって、

前記グローバルビットまたはソース線は、前記半導体基板の前記平坦な表面と前記アクティブ列のアレイとの間、または前記アクティブ列のアレイの上に設けられることを特徴とするメモリ構造。

20

【請求項 13】

請求項 12 に記載のメモリ構造であって、

前記第 2 の方向に沿って延びる複数行のアクティブ列をなす前記アクティブ列には、前記グローバルビットまたはソース線のうちの第 1 のグローバルビットまたはソース線及び第 2 のグローバルビットまたはソース線が用いられ、前記第 1 のグローバルビットまたはソース線は、前記複数行のアクティブ列内の 1 つおきのアクティブ列の前記ビット線に接続し、前記第 2 のグローバルビットまたはソース線は、前記第 1 のグローバルビットまたは

30

【請求項 14】

請求項 12 に記載のメモリ構造であって、

前記複数行のアクティブ列における隣接するアクティブ列は、共通ワード線としての前記導体の 1 つを共有し、前記共通ワード線は、前記隣接するアクティブ列のうちの一方のアクティブ列の第 1 の NOR ストリングの薄膜トランジスタ、及び前記隣接するアクティブ列のうちの他方のアクティブ列の第 2 の NOR ストリングの薄膜トランジスタに対して用いられることを特徴とするメモリ構造。

【請求項 15】

請求項 1 に記載のメモリ構造であって、

前記導体は、前記半導体基板の前記平坦な表面にそれぞれ平行な複数の平面に配置され、

各導体は、アクティブ列の隣接する行の間に位置する延長部分を有し、

前記延長部分は、前記第 2 の方向に沿って延在し、同一平面上の隣接する前記導体の前記延長部分間の距離は前記電荷トラップ材料の 2 倍未満であることを特徴とするメモリ構造。

40

【請求項 16】

請求項 1 に記載のメモリ構造であって、

薄膜トランジスタの、前記導体、前記電荷トラップ材料及びチャンネルとして機能する前

50

記低濃度ドーブ領域は、前記可変閾値薄膜トランジスタのチャンネルとして機能する前記低濃度ドーブ領域から前記電荷トラップ材料への電子のトンネリングがより促進され、消去中に前記ワード線の導体から前記電荷トラップ材料への電子のトンネリングが妨げられるような所定の曲率で湾曲していることを特徴とするメモリ構造。

【請求項 17】

請求項 1 に記載のメモリ構造であって、

前記電荷トラップ材料は、ブロッキング層、ストレージ層及びトンネル誘電体層を含むことを特徴とするメモリ構造。

【請求項 18】

メモリ構造を作製するための方法であって、

半導体基板の中及び上にメモリ回路をサポートするための回路を形成するステップであって、前記半導体基板は平坦な表面を有する。該回路を形成するステップと、

前記半導体基板の前記平坦な表面上に第 1 の絶縁層を設けるステップと、

前記第 1 の絶縁層に、前記半導体基板の中及び上の前記回路に電氣的に接続するためのコンタクトを設けるステップであって、前記コンタクトは、前記第 1 の絶縁層を貫通して第 1 の方向に延び、前記第 1 の方向は前記半導体基板の平面と直交している、該コンタクトを設けるステップと、

複数の導体スタックを形成するステップであって、各導体は、長手方向に、前記半導体基板の前記平面に平行な第 2 の方向に延び、前記導体は、介在する絶縁層によって互いに絶縁されており、前記複数の導体スタックは複数のトレンチによって互いに分離されており、前記複数のトレンチは、前記第 2 の方向に沿った行と、前記半導体基板の前記平面に平行な第 3 の方向に沿った行とを有するアレイ状に、前記第 1 の方向に沿って前記第 1 の絶縁層に達するように画定される、該複数の導体スタックを形成するステップと、

前記複数のトレンチの側壁上に電荷トラップ層を堆積するステップと、

前記複数のトレンチの底部及び前記第 1 の絶縁層で前記電荷トラップ層を異方性エッチングするステップと、

前記電荷トラップ層の表面上に、前記半導体基板に向けて前記第 1 の方向に沿って延在する第 1 の導電型の低濃度にドーブされたポリシリコンを堆積し、

前記複数のトレンチを高速エッチング誘電体層で充填するステップと、

前記第 1 の絶縁層に達するシャフトを形成するために、前記高速エッチング誘電体層をフォトリソグラフィパターンニング及び異方性エッチングするステップと、

前記低濃度にドーブされたポリシリコンの部分に接触する、前記シャフトの側壁上に、前記第 1 の導電型と反対の第 2 の導電型の高濃度にドーブされたポリシリコンの層を設けるステップと、

前記導体スタックの頂部から前記高濃度にドーブされたポリシリコンを除去し、前記導体スタックの上に第 2 の絶縁層を設けるステップとを含むことを特徴とする方法。

【請求項 19】

請求項 18 に記載の方法であって、

前記低濃度にドーブされたポリシリコンはアモルファスシリコンとして堆積され、続いて多結晶シリコンにアニールされることを特徴とする方法。

【請求項 20】

請求項 18 に記載の方法であって、

前記導体は、低抵抗材料の群から選択された材料を含み、前記低抵抗材料の群は、タングステンまたは他の屈折性金属、N+ドーブされたポリシリコン、P+ドーブされたポリシリコン、ニッケルシリサイド、コバルトシリサイド、及びタングステンまたは他のシリサイド及び高濃度にドーブされたポリシリコンのシリサイド及びシリサイド、及びそれらの組み合わせからなる群から選択されることを特徴とする方法。

【請求項 21】

請求項 18 に記載の方法であって、

隣接する両シャフトの高濃度にドーブされたポリシリコンは、共用ソース線及び共有ド

10

20

30

40

50

ラインまたはビット線を形成し、前記低濃度にドーブされたポリシリコンは共有チャンネルを形成し、前記導体のスタックは、NORストリング内の薄膜トランジスタの個々のワード線を形成することを特徴とする方法。

【請求項 2 2】

請求項 1 8 に記載の方法であって、

前記導体のスタックの各々における前記平坦な表面から更に離れた連続する導体の各々は、前記導体のスタックにおいて前記第 3 の方向に沿って延びる距離がその直前の導体よりもより小さいことを特徴とする方法。

【請求項 2 3】

請求項 2 2 に記載の方法であって、

前記導体のスタックの各々のなか全ての導体が形成された後、前記介在する絶縁層を貫通するバイアをエッチングして、各導体を前記半導体基板の前記回路に接続するステップを更に含むことを特徴とする方法。

【請求項 2 4】

請求項 1 8 に記載の方法であって、

前記導体は、前記第 3 の方向に沿って長手方向に延びる導体ストリップを形成するようにパターンングされ、各導体ストリップの所定の部分が所定の凹曲面の曲率半径を有することを特徴とする方法。

【請求項 2 5】

請求項 1 8 に記載の方法であって、

前記複数の導体スタックを形成するステップは、前記介在する絶縁層として複数の絶縁層を設けるステップであって、前記介在する絶縁層同士は、介在する犠牲層によって互いに分離される、該複数の絶縁層を設けるステップと、

前記複数の絶縁層の間に形成された前記犠牲層を除去して、前記絶縁層の間にキャビティを形成するステップと、

前記キャビティに高導電性の導体を充填するステップとを含むことを特徴とする方法。

【請求項 2 6】

請求項 2 5 に記載の方法であって、

前記電荷トラップ層が前記複数のトレンチの側壁上に堆積された後に、前記キャビティ及び前記導体が形成され、前記低濃度にドーブされたポリシリコンが前記電荷トラップ層の表面上に堆積されることを特徴とする方法。

【請求項 2 7】

請求項 1 8 に記載の方法であって、

前記電荷トラップ層は、ブロッキング層、ストレージ層及びトンネル誘電体層を含むことを特徴とする方法。

【発明の詳細な説明】

【技術分野】

【0001】

[関連出願との相互参照]

本出願は、(i) 2 0 1 5 年 1 1 月 2 5 日に 出 願 さ れ た 同 時 係 属 中 の 「 3 次 元 垂 直 N O R フラッシュ薄膜トランジスタストリング」と題する米国仮特許出願第 6 2 / 2 6 0 1 3 7 号 (同 時 係 属 米 国 仮 特 許 出 願 I) 、 (i i) 2 0 1 6 年 7 月 2 6 日 に 出 願 さ れ た 同 時 係 属 中 の 「 垂 直 制 御 ゲ ー ト を 備 え た 積 層 水 平 ア ク テ ィ ブ ス ト リ ッ プ に 構 成 さ れ た マ ル チ ゲ ー ト N O R フラッシュ薄膜トランジスタストリング」と題する米国特許出願第 1 5 / 2 2 0 3 7 5 号 (同 時 係 属 米 国 特 許 出 願 I) 、 (i i i) 2 0 1 6 年 7 月 1 5 日 に 出 願 さ れ た 同 時 係 属 中 の 「 容 量 結 合 型 非 揮 発 性 薄 膜 ト ラ ン ジ ス タ ス ト リ ン グ 」 と 題 する 米 国 仮 特 許 出 願 第 6 2 / 3 6 3 1 8 9 号 (同 時 係 属 米 国 仮 特 許 出 願 I I) 、 及 び

(i v) 2 0 1 6 年 8 月 2 6 日 に 出 願 さ れ た 同 時 係 属 中 の 「 3 次 元 容 量 結 合 型 非 揮 発 性 薄 膜 ト ラ ン ジ ス タ ス ト リ ン グ 」 と 題 する 米 国 特 許 出 願 第 1 5 / 2 4 8 4 2 0 号 (同 時 係 属 米

10

20

30

40

50

国特許出願 I I) に関連し、それら優先権の利益を主張する。同時係属米国仮特許出願 I I 、同時係属米国仮特許出願 I I 、同時係属米国特許出願 I I 、及び同時係属米国特許出願 I I の開示内容は、参照によってその全内容が本明細書に組み込まれるものとする。

【 0 0 0 2 】

本発明は、高密度メモリ構造に関する。本発明は特に、水平ワード線を有する垂直ストリップに形成された薄膜のような、相互接続された薄膜記憶素子によって形成された高密度メモリ構造に関する。

【 背景技術 】

【 0 0 0 3 】

本開示では、いくつかのメモリ回路構造について記載する。これらの構造は、従来の製造プロセスを用いてプレーナ半導体基板（例えば、シリコンウェハ）上に製造することができる。本明細書の明確化のために、「垂直」という用語は、半導体基板の表面に垂直な方向を指し、「水平」という用語は、半導体基板の表面に平行な任意の方向を指すものとする。

10

【 0 0 0 4 】

「 3 次元垂直 N A N D ストリング」など多種の高密度不揮発性メモリ構造が従来技術で知られている。これらの高密度メモリ構造の多くは、堆積された薄膜（例えば、ポリシリコン薄膜）から形成され、「メモリストリング」のレイとして構成された薄膜蓄積トランジスタを用いて組織化される。メモリストリングの 1 タイプは、N A N D メモリストリング又は単に「N A N D ストリング」と呼ばれる。N A N D ストリングは、多数の直列接続された薄膜メモリトランジスタ（「T F T」）からなる。直列接続された T F T のいずれかの内容を読み出し又はプログラミングするには、ストリング内の全ての直列接続された T F T のアクティブ化が必要である。薄膜 N A N D トランジスタは、単結晶シリコン中に形成された N A N D トランジスタよりも導電率が低いので、長い N A N D ストリングを通して導通する必要がある低い読み出し電流では、読み出しアクセスが遅くなる（すなわち、長い待ち時間をもたらす）。

20

【 0 0 0 5 】

別のタイプの高密度メモリ構造は、N O R メモリストリングまたは「N O R ストリング」と呼ばれる。N O R ストリングは、その各々が共有ソース領域及び共有ドレイン領域にそれぞれ接続された多数のメモリトランジスタを含む。したがって、N O R ストリング内の読み出し電流は、N A N D ストリングを通る読み出し電流よりはるかに小さい抵抗で導通するように、N O R ストリングが並列に接続されている。N O R ストリング内のメモリトランジスタを読み出しまたはプログラミングするためには、そのメモリトランジスタのみをアクティブ化（すなわち「オン」または導通）する必要があり、N O R ストリング内の他の全てのメモリトランジスタは休止状態（すなわちオフまたは非導通）に置かれ得る。その結果、N O R ストリングにより、読み出されるべきアクティブ化されたメモリトランジスタのより迅速な感知が可能になる。従来の N O R トランジスタは、チャンネルホットエレクトロン注入技術によってプログラミングされ、その場合、電子がソース領域とドレイン領域との間の電圧差によってチャンネル領域内で加速され、適切な電圧が制御ゲートに印加されると、制御ゲートとチャンネル領域との間の電荷トラップ層に注入される。チャンネルホットエレクトロン注入プログラミングは、チャンネル領域を流れる比較的大きな電流（電子の流れ）を必要とするため、並列にプログラミングすることができるトランジスタの数が制限される。ホットエレクトロン注入によってプログラミングされたトランジスタとは異なり、ファウラー・ノルドハイム・トンネリングまたは直接トンネリングによってプログラミングされたトランジスタでは、制御ゲートと、ソース領域及びドレイン領域との間に印加される高電界によってチャンネル領域から電荷トラップ層に電子が注入される。ファウラー・ノルドハイム・トンネリングと直接トンネリングは、チャンネルホットエレクトロン注入よりも効率的で、大規模並列プログラミングが可能であるが、このようなトンネリングは、プログラム妨害条件の影響をより受けやすい。

30

40

【 0 0 0 6 】

50

3次元NORメモリアレイは、2011年3月11日に出願され2014年1月14日に公報発行された「3D NORアレイのメモリアーキテクチャ」と題する、H. T. Lueによる米国特許第8630114号明細書（特許文献1）に開示されている。

【0007】

2015年9月21日に出願され2016年3月24日に公開された「3D NORアレイのメモリアーキテクチャ」と題する、Haibing Pengによる米国特許出願公開第2016/0086970号明細書（特許文献2）は、基本NORメモリグループのアレイからなる不揮発性NORフラッシュメモリデバイスを開示しており、このデバイスでは、個々のメモリセルが、半導体基板に平行な水平方向に沿って積層され、ソース電極とドレイン電極が、導電チャネルの一方の側または両側に配置される全ての電界効果トランジスタによって共有される。

10

【0008】

3次元垂直メモリ構造は、例えば、2013年1月30日に出願され2014年11月4日に公報発行された「コンパクトな3次元垂直NAND及びその製造方法」と題する、Alsmeyerらによる米国特許第8878278号明細書（特許文献3）に開示される。特許文献3（Alsmeyer）は、様々なタイプの高密度NANDメモリ構造、例えば、「テラビット・セル・アレイ・トランジスタ」（TCAT）NANDアレイ（図1A）、「パイプ形状のビット・コスト・スケーラブルな（P-BiCS）フラッシュ・メモリ」（図1B）及び「垂直NAND」メモリストリング構造などを開示している。同様に、2002年12月31日に出願され2006年2月28日に公報発行された「直列接続されたトランジスタストリングを組み込んだプログラブルメモリアレイ構造の製造方法」と題する、Walkerらによる米国特許第7005350号明細書（特許文献4）（Walker I）もまた、多数の3次元高密度NANDメモリ構造を開示している。

20

【0009】

2005年8月3日に出願され2009年11月3日に公報発行された「デュアルゲートデバイス及び方法」と題する、Walkerによる米国特許第7612411号明細書（特許文献5）（Walker II）は、「デュアルゲート」メモリ構造を開示し、この構造では共有アクティブ領域が、共有アクティブ領域の反対側に形成された2つのNANDストリング内の独立制御される記憶素子を提供する。

30

【0010】

3次元NORメモリアレイは、米国特許第8630114号明細書（特許文献1）に開示されている。

【0011】

垂直ポリシリコンゲートによって制御される水平NANDストリングを含む3次元メモリ構造は、Kimら、2009（非特許文献1）に開示されている。垂直ポリシリコンゲートを有する水平NANDストリングも含む、別の3次元メモリ構造は、Lueら、2010（非特許文献2）に開示されている。

【0012】

図1Aは、従来技術による3次元垂直NANDストリング101及び102を示す。図1Bは、従来技術による3次元垂直NANDストリングの基本回路表現140を示す。具体的には、図1Aの垂直NANDストリング101及び102及びそれらの垂直NANDストリング150の回路表現は、それぞれが（基板の表面に沿って直列に接続された32個以上のトランジスタを接続する各トランジスタではなく）基板に対して垂直になるように90度回転されている本質的に従来技術の水平NANDストリングである。垂直NANDストリング101及び102は、基板から高層ビルのように積み上がっているストリング構成で直列に接続された薄膜トランジスタ（TFET）であり、各TFETは、記憶素子と、ワード線導体の隣接スタック内のワード線導体の1つによって提供される制御ゲートとを有する。図1Bに示すように、垂直NANDストリングの最も簡単な実施形態では、TFET15及び16は、それぞれ別々のワード線WL0及びWL31によって制御されるNANDストリング150の最初及び最後のメモリトランジスタである。信号BLSによってア

40

50

クティブ化されるビット線選択トランジスタ11と、信号SSによってアクティブ化される接地選択トランジスタ12は、読み出し、プログラム、プログラム禁止及び消去の各動作の間に、垂直NANDストリング150内のアドレス指定されたTF Tを、端子14において対応するグローバルビット線GBLに接続し、端子13においてグローバルソース線(GSL)に接地させる。任意の1つのTF T(例えば、TF T17)の内容の読み出しまたはプログラミングのためには、垂直NANDストリング150内の32個全てのTF Tのアクティブ化し、各TF Tを読み出し禁止及びプログラム禁止の状態にすることが必要である。このような条件により、垂直NANDストリングに設けることができるTF Tの数が64個以下または128個以下のTF Tに制限される。更に、垂直NANDストリングが形成されるポリシリコン薄膜は、単結晶シリコン基板に形成された従来のNANDストリングよりチャネル移動度が非常に低く、したがってより高い抵抗率を有し、このため従来のNANDストリングの読み出し電流と比較して読み出し電流が低くなる。

10

【0013】

「垂直構造半導体メモリ装置及びその製造方法」と題する、Hwangによる米国特許出願公開第2011/0298013号明細書(特許文献6)では、そのFIG.4Dにおいて、ラップアラウンドスタックワード線(図1Cの150として本明細書に再掲されている)によってアドレス指定される3次元垂直NANDストリングのブロックを示す。

【0014】

1996年7月23日に出願され1998年6月16日に公報発行された「非対称電荷トラップを利用したメモリセル」と題する、Eitanによる米国特許第5768192号明細書(特許文献7)は本発明の一実施形態で採用されるタイプのNROM型メモリトランジスタ動作を開示している。

20

【0015】

2010年10月11日に出願され2011年9月27日に公報発行された「非対称電荷トラップを利用したメモリセル」と題する、Zvi Or-Bachらによる米国特許第8026521号明細書(特許文献8)は、第1の層及び第2の層が水平に配向されたトランジスタを含む層転写された単結晶シリコンの第1の層及び第2の層を開示している。この構造では、水平方向に配向されたトランジスタの第2の層は、水平に配向されたトランジスタの第1の層を覆い、水平方向に配向されたトランジスタの各グループは側部ゲートを有する。

30

【0016】

従来の不揮発性メモリトランジスタ構造を有するが、保持時間が短いトランジスタは、「準揮発性(quasi-volatile)」と呼ぶことができる。これに関連する点として、従来の不揮発性メモリはデータ保持時間が数十年を超える。単結晶シリコン基板上の平面準揮発性メモリトランジスタは、H.C. Wannら、1995(非特許文献3)に開示されている。また、準揮発性メモリを有する準揮発性3-D NORアレイが、上述の米国特許第8630114号明細書(特許文献6)に開示されている。

【先行技術文献】

【特許文献】

【0017】

40

【特許文献1】米国特許第8630114号明細書

【特許文献2】米国特許出願公開第2016/0086970号明細書

【特許文献3】米国特許第8878278号明細書

【特許文献4】米国特許第7005350号明細書

【特許文献5】米国特許第7612411号明細書

【特許文献6】米国特許出願公開第2011/0298013号明細書

【特許文献7】米国特許第5768192号明細書

【特許文献8】米国特許第8026521号明細書

【非特許文献】

【0018】

50

【非特許文献1】W. Kim et al., "Multi-layered Vertical gate NAND Flash Overcoming Stacking Limit for Terabit Density Storage" ("Kim"), (2009) Symposium on VLSI Tech. Dig. of Technical Papers, pp 188-189

【非特許文献2】H.T. Lue et al., "A Highly Scalable 8-Layer 3D Vertical-gate (VG) TFT NAND Flash Using Junction-Free Buried Channel BE-SONOS Device," 2010 Symposium on VLSI: Tech. Dig. Of Technical Papers, pp.131-133

【非特許文献3】H.C. Wann and C.Hu, "High-Endurance Ultra-Thin Tunnel Oxide in Monos Device Structure for Dynamic Memory Application", IEEE Electron Device letters, Vol. 16, No. 11, November 1995, pp 491-493

10

【発明の概要】

【課題を解決するための手段】

【0019】

本発明の一実施形態によれば、3次元垂直NORフラッシュメモリストリング（「マルチゲート垂直NORストリング」または単に「垂直NORストリング」）と呼ばれる高密度メモリ構造が提供される。垂直NORストリングは、並列に接続された複数の薄膜トランジスタ（「TFT」）を含み、共通ソース領域及び共通ドレイン領域がそれぞれ略垂直方向に延びている。更に、垂直NORストリングは、垂直NORストリングにおけるそれぞれのTFTを制御する複数の水平制御ゲートを含む。垂直NORストリングのTFTが並列に接続されると、垂直NORストリングにおける読み出し電流は、同じ数のTFTのNANDストリングによる読み出し電流よりもはるかに小さい抵抗で導通する。垂直NORストリングにおけるTFTのいずれか1つを読み出しまたはプログラムするために、そのTFTのみがアクティブにされる必要があり、垂直NORストリングにおける他の全てのTFTは非導通のままであり得る。結果として、垂直NORストリングにより、より速い感知が可能となり、プログラム妨害または読み出し妨害状態を最小限にしながら、より多くの（例えば、数百またはそれ以上）TFTを含めることができる。

20

【0020】

一実施形態では、垂直NORストリングの共有ドレイン領域はグローバルビット線（「電圧 V_{bl} 」）に接続され、垂直NORストリングの共有ソース領域はグローバルソース線（「電圧 V_{ss} 」）に接続される。あるいは、第2の実施形態では、共有ドレイン領域のみが供給電圧にバイアスされたグローバルビット線に接続され、共有ソース領域は共有ソース領域内の電荷量によって決定される電圧にプリチャージされる。プリチャージを行うため、共有ソース領域の寄生容量Cをプリチャージする1つ以上の専用TFTを設けることができる。

30

【0021】

本発明の一実施形態によれば、マルチゲートNORフラッシュ薄膜トランジスタアレイ（「マルチゲートNORストリングアレイ」）は、シリコン基板の表面に垂直に延びる垂直NORストリングのアレイとして編成される。各マルチゲートNORストリングアレイは、行をなすように配列された多数の垂直アクティブ列を含み、各列は第1の水平方向に沿って延在し、各アクティブ列は、第1の導電型の高濃度ドーパされた2つの垂直なポリシリコン領域を有し、これらは第2の導電型にドーパされていないかまたは低濃度でドーパされた1つ以上の垂直ポリシリコン領域によって分離される。高濃度にドーパされた領域はそれぞれ共通のソースまたはドレイン領域を形成し、それぞれが第1の水平方向に直交する1つ以上の水平導体のスタックとともに、低濃度にドーパされた領域はそれぞれ複数のチャンネル領域を形成する。電荷トラップ材料は、少なくともアクティブ列内のTFTのチャンネル領域を覆う記憶素子を形成する。各スタック内の水平導電線は、互いに電氣的に絶縁されており、アクティブ列の記憶素子及びチャンネル領域の上に制御ゲートを形成する。このようにして、マルチゲートNORストリングアレイは、ストレージTFTの3次元アレイを形成する。

40

【0022】

50

一実施形態では、サポート回路が半導体基板内に形成され、サポート回路及び半導体基板の上に形成された複数のマルチゲートNORストリングアレイをサポートする。サポート回路は、アドレスエンコーダ、アドレスデコーダ、センスアンプ、入出力ドライバ、シフトレジスタ、ラッチ、比較セル、電源線、バイアス及び基準電圧発生器、インバータ、NAND、NOR、排他的論理和やその他の論理ゲート、他のメモリ素子、シーケンサ、ステートマシンなどを含む。マルチゲートNORストリングアレイは、複数のマルチゲートNORストリングアレイを有する複数のブロックの回路として編成することができる。

【0023】

本発明の実施形態によれば、垂直NORストリング内のTFETの閾値電圧の変動は、同一または別のマルチゲート垂直NORストリングアレイ内に1つ以上の電氣的にプログラム可能な基準垂直NORストリングを提供することによって補償することができる。垂直NORストリングに固有のバックグラウンドリーク電流は、読み出し中のTFETの結果を、プログラム可能な参照垂直NORストリングで同時に読み出されるTFETの結果と比較することによって、読み出し動作中に実質的に中和することができる。いくつかの実施形態では、垂直NORストリングの各TFETは、各制御ゲートとその対応するチャンネル領域との間の容量性結合を増幅し、それによってチャンネル領域から電荷トラップ材料（すなわち記憶素子）の消去中に制御ゲートから電荷トラップ材料への電荷注入を減少させることができる。この有利な容量結合は、垂直NORストリングの各TFETに1ビット以上を記憶するのに特に有用である。別の実施形態では、各TFETの電荷トラップ材料は、格納されたデータのリフレッシュを必要とするより低い保持時間ではあるが、高い書き込み/消去サイクル耐久性を有するように改変された構造を有し得る。しかし、垂直NORストリングアレイに要求されるリフレッシュは、従来のダイナミックランダムアクセスメモリ（DRAM）よりもはるかに少ないことになるので、本発明のマルチゲートNORストリングアレイは、いくつかのDRAMアプリケーションにおいても動作することができる。このような垂直NORストリングを使用すると、従来のDRAMと比較して実質的により低いビット当たりコスト対ビット性能指数、及び従来のNANDストリングアレイと比較して大幅に小さい読み出しレイテンシが可能となる。

【0024】

別の実施形態では、垂直NORストリングは、NROM/ミラービットTFETストリングとしてプログラムされ、消去され、読み出され得る。

【0025】

従来技術の垂直NANDストリングではなく垂直NORストリングとしてTFETを構成することで、(i)ダイナミックランダムアクセスメモリ（DRAM）アレイのレイテンシに近くものにできる小さい読み出しレイテンシ、(ii)長いNANDフラッシュストリングに関連する読み出し妨害及びプログラム妨害状態から受ける影響の低減及び(iii)NANDフラッシュストリングと比較して、低い1ビット当たりのコストという効果が得られる。

【0026】

本発明は、添付の図面と併せて、以下の詳細な説明を参照することにより、よりよく理解されよう。

【図面の簡単な説明】

【0027】

【図1A】図1Aは、従来技術における3次元垂直NANDストリング101及び102を示す図である。

【図1B】図1Bは、従来技術における3次元垂直NANDストリングの基本回路表現140を示す図である。

【図1C】図1Cは、ラップアラウンドスタックワード線（本図の150）によってアドレス指定される3次元垂直NANDストリングのブロックの3次元表現を示す図である。

【図2】図2は、概念化されたメモリ構造100を示す図であって、本発明の一実施形態によるメモリセルの3次元構成を示しており、これらのメモリセルは、垂直NORストリ

10

20

30

40

50

ング形態で設けられ、各垂直NORストリングは、多数の水平ワード線のうちの1つによって制御されるメモリセルを有する。

【図3A】図3Aは、アクティブ列に形成された垂直NORストリング300のZ-Y平面内の基本回路表現を示す図であり、垂直NORストリング300は不揮発性ストレージTF Tの3次元アレイ(配列)を表し、本発明の一実施形態によれば、ローカルソース線(L S L)355及びローカルビット線(L B L)354を共有する各TF Tは、それぞれグローバルビット線(G B L)314及びグローバルソース線(G S L)313によってアクセスされる。

【図3B】図3Bは、アクティブ列に形成された垂直NORストリング305のZ-Y平面内の基本回路表現を示す図であり、垂直NORストリング305は、不揮発性ストレージTF Tの3次元アレイを表し、本発明の一実施形態によれば、寄生キャパシタCを有する共有ローカルソース線355上の電圧(「 V_{SS} 」)を設定するための専用プリチャージTF T370を含む。

【図3C】図3Cは、1つまたは複数のプログラミングされた閾値電圧を有し、寄生キャパシタ360に接続された動的な不揮発性ストレージトランジスタ317の基本回路表現を示す図であり、キャパシタ360は、ソース端子(ソース線)355に仮想電圧 V_{SS} を一時的に保持するようにプリチャージされ、制御ゲート323pの電圧が閾値電圧を超える電圧まで上昇すると、電圧 V_{SS} の放電によってトランジスタ317の閾値電圧を動的に検出できる。

【図4A】図4Aは、本発明の一実施形態のZ-Y平面内の断面図であり、各列は、図3Aまたは図3Bのいずれかに示す基本回路表現を有する垂直NORストリングを形成し得る。

【図4B】図4Bは、本発明の一実施形態のZ-X平面内の断面図であり、アクティブ列430R、430L、431R及び431L、電荷トラップ層432及び434、及びワード線423p-L及び423p-Rを示す。

【図4C】図4Cは、本発明の一実施形態の垂直NORストリング対491及び492のZ-X平面における基本的な回路表現を示す図である。

【図5】図5は、本発明の一実施形態のZ-Y平面内の断面図であり、アクティブ列531の垂直NORストリング、グローバルビット線514-1(G B L 1)、グローバルソース線507(G S L 1)、及び共通ボディバイアスソース506(V_{bb})の間の接続を示す。

【図6A】図6Aは、本発明の一実施形態のXY平面内の断面図であり、図4Cに関連して説明したように、垂直NORストリング451aのTF T685(T_L)及び垂直NORストリング対491の垂直NORストリング451bのTF T684(T_R)を示し、図6Aにおいては、グローバルビット線614-1は、ローカルビット線L B L-1の一つおきにアクセスし、トランジスタチャネル領域656Lの所定の湾曲675により、プログラミング中の各制御ゲートと対応するチャネルとの間の容量結合が増幅されている。

【図6B】図6Bは、本発明の一実施形態のXY平面内の断面図であり、図4Cに関連して説明したように、垂直NORストリング対491の垂直NORストリング451bのTF T684(T_R)とアクティブ領域を共有する垂直NORストリング451aのTF T685(T_L)を示し、図6Bにおいては、グローバルビット線614-1はローカルビット線654(L B L-1)の一つおき(奇数番目)のビットにアクセスし、グローバルビット線614-2はローカルビット線657-2(L B L-2)の一つおき(偶数番目)のビットをアドレス指定し、ローカルソース線L S L-1及びL S L-2は、仮想電源電圧 V_{SS} を供給するためにプリチャージされる。

【図6C】図6Cは、本発明の一実施形態のX-Y平面内の断面図であり、ワード線群を各々が含む専用ワード線スタック623pと、(ピラーすなわち柱状部である)ローカル垂直ピラービット線654(Z方向に沿って延びる)及びローカル垂直ピラーソース線655(Z方向に沿って延びる)とを示しており、各ワード線スタックのワード線群の各ワード線は垂直NORストリングのTF Tを包み込む(「ラップアラウンド」する)ように

10

20

30

40

50

延在し、ローカル垂直ピラービット線 6 5 4 及びローカル垂直ピラーソース線 6 5 5 は、それぞれグローバル水平ビット線 6 1 4 及びグローバル水平ソース線 6 1 5 によってアクセスされ、また図 6 C においては、隣接するワード線スタック 6 2 3 p は、エアギャップ 6 1 0 または別の誘電体分離によって互いに絶縁されている。

【図 6 D】図 6 D は、本発明の一実施形態の X - Y 平面内の断面図であり、垂直 NOR ストリングが千鳥状に近接して配置された形態を示しており、図 6 C に示すものと同様に、ワード線スタック 6 2 3 p が共有され、プリチャージされた寄生キャパシタ 6 6 0 の各々が、プリチャージされた仮想供給電圧 V_{SS} を供給する。

【図 7 A】図 7 A は、本発明の一実施形態による、マルチゲート NOR ストリングアレイの製造プロセスで形成される中間構造の断面図である。

10

【図 7 B】図 7 B は、本発明の一実施形態による、マルチゲート NOR ストリングアレイの製造プロセスで形成される中間構造の断面図である。

【図 7 C】図 7 C は、本発明の一実施形態による、マルチゲート NOR ストリングアレイの製造プロセスで形成される中間構造の断面図である。

【図 7 D】図 7 D は、本発明の一実施形態による、マルチゲート NOR ストリングアレイの製造プロセスで形成される中間構造の断面図である。

【図 8 A】図 8 A は、本発明の一実施形態の読み出し動作の概略図であり、ここでは、垂直 NOR ストリングのローカルソース線 (LSL) はハードワイヤードされており、図 8 A において、「 $W L_s$ 」は選択されたワード線上の電圧を表し、垂直 NOR ストリング内の全ての非選択ワード線 (「 $W L_{NS}$ 」) は、読み出し動作中に 0 V に設定される。

20

【図 8 B】図 8 B は、ローカルソース線がプリチャージ仮想電圧 V_{SS} でフロート状態になる実施形態の読み出し動作の概略図であり、図 8 B においては、「 $W L_{CHG}$ 」は、プリチャージされたトランジスタ (例えば、図 3 C のプリチャージされたトランジスタ 3 1 7 または 3 7 0) 上のゲート電圧を表す。

【発明を実施するための形態】

【0028】

図 2 は、垂直 NOR ストリングに設けられたメモリセル (または記憶素子) の 3 次元構成を示す、概念化されたメモリ構造 1 0 0 を示す。概念化されたメモリ構造 1 0 0 では、本発明の一実施形態によれば、各垂直 NOR ストリングは、対応する水平ワード線によってそれぞれ制御されるメモリセルを含む。基板層 (図 2 の 1 0 1) は、例えば、当業者によく知られている、集積回路を製造するために使用される従来のシリコンウェハであり得る。この詳細な説明では、(図 2 に示すような) デカルト座標系は、議論を容易化する目的のためだけに採用される。この座標系の下では、基板層 (図 2 の 1 0 1) の表面は X - Y 平面に平行な平面とみなされる。したがって、本明細書で使用する「水平」という語は、X - Y 平面に平行な任意の方向を指し、「垂直」は Z 方向を指す。

30

【0029】

図 2 において、Z 方向の各垂直列は、垂直 NOR ストリング (例えば、垂直 NOR ストリング 1 2 1) における記憶素子または T F T を表す。縦方向の NOR ストリングは、X 方向に並んだ行として規則的に配列されている。(当然ながら、Y 方向に並んだ行の配列としても同様な配置が見られる。) 垂直 NOR ストリングの記憶素子は、垂直ローカルソース線及び垂直ローカルビット線 (図示せず) を共有する。水平ワード線のスタック (例えば、 $W L_{123}$) は、Y 方向に沿って延び、各ワード線は、Y 方向に沿ってワード線に隣接して位置する垂直 NOR ストリングの対応する T F T の制御ゲートとして働く。グローバルソース線 (例えば、 $G S L_{122}$) 及びグローバルビット線 (例えば、 $G B L_{124}$) は、一般に、概念化されたメモリ構造 1 0 0 の底部の下または上部の上のいずれかにおいて X 方向に沿って延びるように設けられる。あるいは、信号線 $G S L_{122}$ 及び $G B L_{124}$ の両方が、概念化されたメモリ構造 1 0 0 の下か、その上部の上にルーティングされてもよく、これらの信号線の各々は、アクセストランジスタ (図示せず) によって個々の垂直 NOR ストリングのローカルソース線及びローカルビット線に選択的に接続することができる。従来技術の垂直 NAND ストリングとは異なり、本発明の垂直 NOR スト

40

50

リングでは、その記憶素子のいずれか1つに書き込みまたはそこから読み出しするときに、垂直NORストリングにおける他の記憶素子のアクティブ化は行われない。図2に示されているように、単に説明のための例示として、概念化されたメモリブロック（メモリ構造）100は、垂直NORストリングの4×5配列からなるマルチゲート垂直NORストリングアレイであり、各NORストリングは、通常、32以上の記憶素子とアクセス選択トランジスタを有する。概念的な構造として、メモリブロック（メモリ構造）100は、本発明のメモリ構造のある目立った特徴を抽象化したものに過ぎない。図2には、各垂直NORストリングが複数の記憶素子を有する垂直NORストリングの4×5構成として示されているが、本発明のメモリ構造は、X方向及びY方向のいずれかに沿って各行に任意の数の垂直NORストリングを有することができ、各垂直NORストリングには任意の数の記憶素子が存在し得る。例えば、X方向とY方向の両方に沿って列になして配列された数千の垂直NORストリングが存在することができ、各NORストリングは、例えば、2、4、8、16、32、64、128またはそれ以上の個数の記憶素子を有する。

10

【0030】

図2の各垂直方向のNORストリング（例えば、垂直NORストリング121）の記憶素子の数は、垂直NORストリングに制御ゲートを提供するワード線（例えば、WL123）の数に対応する。ワード線は、それぞれがY方向に沿って延びる細長い金属ストリップとして形成される。ワード線は、互いに重ね合わされて積層され、その間の誘電体絶縁層によって互いに電氣的に絶縁される。各スタック内のワード線は、任意の数であってもよいが、好ましくは2の整数乗（すなわち2のn乗（nは整数））である。ワード線の数に対する2のべき乗の選択は、従来のメモリ設計における慣例的な手法に従う。バイナリアドレスをデコードすることによって、アドレス指定可能な各メモリユニットにアクセスすることが慣例的な手法である。この慣例的な手法を採用するか否かは好みの問題であり、従う必要はない。例えば、本発明の範囲内で、概念化されたメモリ構造100は、X及びY方向の各行に沿って、必ずしも2のn乗（nは任意の整数）ではない数であるM個の垂直NORストリングを有することができる。以下に説明されるいくつかの実施形態では、2つの垂直NORストリングが垂直ローカルソース線及び垂直ローカルビット線を共有することができるが、2つの垂直NORストリングそれぞれの記憶要素は2つの別個のワード線スタックによって制御される。これは、垂直方向のNORストリングの記憶密度を事実上2倍にする。

20

30

【0031】

図2の概念化されたメモリ構造100は、単にメモリセルの構成を示すために提供されているので、X、Y、Zのいずれの方向にも特定のスケールで描かれてはいない。

【0032】

図3Aは、アクティブな列に形成された垂直NORストリング300のZ-Y平面における基本的な回路表現を示し、垂直NORストリング300は、不揮発性ストレージTF Tの3次元アレイを表し、本発明の一実施形態によれば、各TF Tは、ローカルソース線355及びローカルビット線354を共有する。この詳細な説明では、用語「アクティブ領域」、「アクティブ列」または「アクティブストリップ」は、アクティブデバイス（例えば、トランジスタまたはダイオード）がその上に存在し得る1つ以上の半導体材料の領域、列、またはストリップを指す。図3Aに示すように、垂直NORストリング300はZ方向に沿って延び、垂直ローカルソース線355と垂直ローカルドレインすなわちビット線354との間に並列に接続されたTF T 316及び317を含む。ビット線354とソース線355は離間しており、その間の領域（すなわち、ボディ領域356）は、垂直方向のNORストリングでTF Tのチャンネル領域を提供する。記憶素子は、（ボディ領域356のなかの）チャンネル領域と各水平ワード線323pとの交差部に形成され、pはワード線スタック内のワード線のインデックスであり、この例では、pは0から31までの任意の値をとることができる。ワード線はY方向に沿って延びている。ローカルビット線354は、ビット線アクセス選択トランジスタ311を介して水平グローバルビット線（GBL）314に接続され、水平グローバルビット線（GBL）314はX方向に沿っ

40

50

て延び、ローカルビット線 354 をアクセスビット線供給電圧 V_{b1} に接続する。ローカルソース線 355 は、水平グローバルソース線 (GSL) 313 を介してソース電源電圧 V_{ss} に接続されている。ローカルソース線 355 と GSL 313 との間を接続するために、任意選択でソース選択トランジスタ (図 3 A には図示せず) を設けることができる。任意選択のソース選択トランジスタは、当業者には知られているように、基板 (例えば、図 2 の半導体基板 101) に、または基板の上でかつメモリ構造 100 の下において実現可能なソースデコード回路によって制御することができる。アクティブ列のボディ領域 356 は、端子 331 で基板バイアス電圧 V_{bb} に接続され得る。基板バイアス電圧 V_{bb} は、例えば、消去動作中に使用することができる。 V_{bb} 供給電圧は、マルチゲート垂直 NOR ストリングアレイの全体に印加することができ、または垂直 NOR ストリングの 1 つ以上の行に選択的に印加するようにデコードすることができる。 V_{bb} 電源電圧をボディ領域 356 に接続する線は、好ましくはワード線の方向に沿って延びる。

10

【0033】

図 3 b は、アクティブ列に形成された垂直 NOR ストリング 305 の Z - Y 平面内の基本回路表現を示しており、垂直 NOR ストリング 305 は、不揮発性ストレージ TFT の 3 次元アレイを表し、これらの TFT は、共有ローカルソース線 355 上の電圧 (「 V_{ss} 」) を瞬時に設定するための、(選択に応じて) 専用のプリチャージ TFT 370 を含み、共有ローカルソース線 355 は、本発明の一実施形態によれば寄生容量 (キャパシタ 360 によって表される) を有する。図 3 A の垂直 NOR ストリング 300 とは異なり、垂直 NOR ストリング 305 は GSL 313 を実装しておらず、一時的に V_{ss} ボルトの電圧を保持する寄生キャパシタ 360 をプリチャージするプリチャージトランジスタ 370 で GSL を置き換えている。このプリチャージ方式の下では、グローバルソース線 (例えば、図 3 A のグローバルソース線 313) 及びそのデコード回路が不要となり、それにより、製造プロセス及び回路レイアウトの両方を単純化し、各垂直 NOR ストリング用に非常にタイトなフットプリントを提供することができる。図 3 c は、その通常の記憶機能に加えて、専用のプリチャージトランジスタ 370 のプリチャージ機能を実行するために使用することができる不揮発性ストレージ TFT 317 の構造を強調して示す。TFT 317 のための動的読み出し動作については、TFT 317 の記憶素子 334 にプログラムされるいくつかの閾値電圧のうちの一つの正しいものを感知することと関連して後で説明する。

20

30

【0034】

図 4 A は、並列アクティブ列 431 及び 432 を示す Z - Y 平面内の断面図であり、その各々は、本発明の一実施形態によれば、図 3 A または図 3 B のいずれかに示される基本回路表現を有する垂直 NOR ストリングを形成することができる。図 4 A に示すように、アクティブな列 431 及び 432 はそれぞれ、低濃度 P ドープまたは非ドープのチャンネル領域 456 によって分離された垂直 N + ドープ局所ソース領域 455 及び垂直 N + ドープローカルドレインすなわちビット線領域 454 を含む。P ドープされたチャンネル領域 456、N + ドープされたローカルソース領域 455、及び N + ドープされたローカルドレインまたはビット線領域 454 は、ボディバイアス電圧 V_{bb} 、ソース電源電圧 V_{ss} 、及びビット線電圧 V_{b1} にそれぞれバイアスされる。本発明のいくつかの実施形態では、アクティブストリップが十分に薄い (例えば 10 nm 以下) 場合など、ボディバイアス電圧 V_{bb} の使用は任意選択である。十分に薄いアクティブストリップの場合、電圧 V_{bb} が垂直 NOR ストリングに沿って TFT のチャンネル領域に固体供給電圧を供給しないように、アクティブ領域は制御ゲート上の適切な電圧の下で容易に完全空乏化される。アクティブ領域である 431 及び 432 を電氣的に絶縁する分離領域 436 は、誘電体絶縁材またはエアギャップのいずれかであり得る。WL0 - WL31 (及び任意選択で WLCHG) とそれぞれ符合が付されたワード線 423 p の垂直スタックは、アクティブな列 431 及び 432 に形成された垂直 NOR ストリングにおける TFT に制御ゲートを提供する。ワード線スタック 423 p は、典型的には、酸化シリコン (例えば、 SiO_2) またはエアギャップで形成された誘電体層 426 によって互いに電氣的に絶縁された Y 方向に沿って

40

50

延びる細長い金属導体（例えば、タングステン、シリサイドまたはシリサイド）として形成される。ワード線 4 2 3 p と P ドープされたチャネル領域 4 5 6 との間に電荷トラップ材料（図示せず）を設けることによって、不揮発性記憶素子を各ワード線 4 2 3 p と各 P ドープチャネル領域 4 5 6 との交差部に形成することができる。例えば、図 4 A は、破線のボックス 4 1 6 によって、不揮発性記憶素子（またはストレージトランジスタ） $T_0 \sim T_{31}$ が形成され得る場所を示す。破線のボックス 4 7 0 は、専用のプリチャージトランジスタが形成される場所を示し、全てのトランジスタ $T_0 \sim T_{31}$ がオフ状態にあるときに、共通ローカルビット線領域 4 5 4 から共通ローカルソース線領域 4 5 5 に電荷を転送することを可能にする。

【0035】

図 4 B は、本発明の一実施形態による、アクティブな列 4 3 0 R、4 3 0 L、4 3 1 R 及び 4 3 1 L、電荷トラップ層 4 3 2 及び 4 3 4、ならびにワード線スタック 4 2 3 p - L 及び 4 2 3 p - R を示す、Z - X 平面内の断面図である。図 4 A と同様に、図 4 B の垂直ワード線スタック 4 2 3 p - L 及び 4 2 3 p - R の各々は、細い導体のスタックを示し、ここで、p はスタック内のワード線（例えば、ワード線 $W L_0 \sim W L_{31}$ ）にラベル付けするインデックスである。図 4 B に示すように、各ワード線は、ワード線の反対側の（領域 4 9 0 内の）隣接するアクティブ列 4 3 0 L 及び 4 3 1 R 上に形成された垂直 NOR ストリングにおける不揮発性 TFT の制御ゲートとして働く。例えば、図 4 B において、ワード線スタック 4 2 3 p - R 内のワード線 $W L_{31}$ は、アクティブ列 4 3 0 L 上のトランジスタ 4 1 6 L とアクティブ列 4 3 1 R 上のトランジスタ 4 1 6 R の両方の制御ゲートとして機能する。隣接するワード線スタック（例えば、ワード線スタック 4 2 3 p - L 及び 4 2 3 p - R）は、後述するように、連続するワード線層をエッチングすることによって形成されるトレンチの幅である距離 4 9 5 だけ分離される。ワード線層を介してエッチングされたトレンチの内側に、アクティブ列 4 3 0 R 及び 4 3 0 L、ならびにそれぞれの電荷トラップ層 4 3 2 及び 4 3 4 が続いて形成される。電荷トラップ層 4 3 4 は、ワード線スタック 4 2 3 p - R と垂直アクティブ列 4 3 1 R、4 3 0 L との間に設けられている。以下で詳述するように、トランジスタ 4 1 6 R のプログラミング中、電荷トラップ層 4 3 4 に注入された電荷は、破線のボックス 4 8 0 内の電荷トラップ層 4 3 4 の部分にトラップされる。トラップされた電荷は、TFT 4 1 6 R の閾値電圧を変化させ、アクティブな列 4 3 1 R 上のローカルソース領域 4 5 5 とローカルドレイン領域 4 5 4 との間に流れる読み出し電流を測定することによって検出することができる（これらの領域は、例えば図 4 A のアクティブな列の直交断面に示される）。いくつかの実施形態では、プリチャージワード線 4 7 8（すなわち、 $W L_{CHG}$ ）が、ローカルソース線 4 5 5 の寄生容量 C を充電するために使用されるプリチャージ TFT 4 7 0 の制御ゲートとして設けられる（図 3 B のキャパシタ 3 6 0 及び図 4 A のローカルソース線 4 5 5 参照）を接地またはソース電源電圧 V_{SS} に接続する。便宜上、電荷トラップ層 4 3 4 は、プリチャージトランジスタ 4 7 0 にも記憶素子を提供するが、それ自体はメモリトランジスタとして使用されない。また、アクティブな列 4 3 1 R に形成されたメモリトランジスタ $T_0 \sim T_{31}$ のいずれかを用いてプリチャージを行ってもよい。これらのメモリトランジスタのうちの一つ以上は、それらの記憶機能に加えて、プリチャージトランジスタの機能を実行することができる。プリチャージを実行するために、ワード線または制御ゲート上の電圧は、プログラム可能な最高閾値電圧よりも数ボルト高く一時的に上昇し、ローカルビット線 4 5 4 に印加される電圧 V_{SS} をローカルソース線 4 5 5 に転送することができる（図 4 A）。メモリトランジスタ $T_0 \sim T_{31}$ がプリチャージ機能を実行することにより、個別の専用プリチャージ TFT 4 7 0 の必要性がなくなる。しかし、プリチャージ機能を実行しているときに、このようなメモリ TFT の閾値電圧を過度に妨害しないように注意する必要がある。

【0036】

アクティブな列 4 3 0 R 及び 4 3 0 L は、図 4 B では、エアギャップまたは誘電体絶縁 4 3 3 によって分離された 2 つの別個のアクティブな列として示されているが、隣接する垂直 N + ローカルソース線は、単一の共有垂直ローカルソース線によって実現されてもよ

10

20

30

40

50

い。同様に、垂直N+ローカルドレインすなわちビット線も、単一の共有垂直ローカルビット線によって実現されてもよい。このような構成は、「垂直NORストリング対」を提供する。その構成では、アクティブな列430L及び430Rは、1つのアクティブな列内の2つの枝(ブランチ)(したがって、「ペア」と見なすことができる。垂直NORストリング対は、アクティブな列430R及び430Lと、反対側のワード線スタック423p-L及び423p-Rとの間に介在する電荷トラップ層432及び434を介して二重密度記憶を提供する。実際、アクティブな列430R及び430Lは、エアギャップまたは誘電体絶縁材433を除去することによって1つのアクティブストリングに統合され得るが、単一のアクティブな列の2つの対向する面に実装されるNOR TFFストリングの対(ペア)を形成する。このような構成は、アクティブな列の反対面に形成されたTFFが別個のワード線スタックによって制御され、別個の電荷トラップ層434及び432から形成されるので、同様の倍密度記憶を達成する。別個の薄いアクティブな列430R及び430Lを別々に維持する(すなわち、それらを1つのアクティブな列にマージする代わりに別々に維持する)ことは有利であり、その理由は、各アクティブな列のTFFが併合された列よりも薄く、したがってそれらのTFFは、適切な制御ゲート電圧条件の下で、より容易に完全に空乏化することができ、アクティブな列(図4A)の垂直ソース領域455と垂直ドレイン領域454との間のソース-ドレイン副閾値リーク電流を実質的に低減するからである。非常に長い垂直NORストリング(例えば、128TFF以上)であっても、超薄型(したがって抵抗性が高い)アクティブ列を有することが可能であるが、その理由は、垂直NORストリングのTFFは並列接続されているため、そして、ストリング中のTFFが直列に接続されているため、ストリング中のTFFのいずれかを感知するために全てスイッチオンする必要があるNAND型TFFストリングの高抵抗とは対照的に、一度に多くのTFFのうちの一つだけがスイッチオンされるためである。例えば、32-TFF垂直NORストリングでは、トランジスタ T_{30} (図4A)を読み出すことが可能となるために、チャンネル領域456のチャンネル長はわずか20nmでよく、この長さは、NANDストリングの対応するチャンネル長(32倍すなわち640nmの長さを有する)と比較して短い。

10

20

【0037】

図4Cは、本発明の一実施形態による垂直NORストリング対491及び492のZ-X平面における基本的な回路表現を示す。図4Cに示すように、垂直NORストリングs451b及び452aは、図4Bのアクティブストリップ430L及び431Rの垂直NORストリング用に示された方法で、共通ワード線スタック423p-Rを共有する。共通接続されたそれぞれのローカルビット線に対して、垂直NORストリング対491及び492には、それぞれアクセス選択トランジスタ411を介してグローバルビット線414-1(GBL1)に、アクセス選択トランジスタ414を介してグローバルビット線414-2(GBL2)が用いられる。それぞれの共通に接続されたローカルソース線のためには、垂直NORストリング対491及び492はそれぞれグローバルソース線413-1(GSL1)及びグローバルソース線413-2(GSL2)が用いられる(ソース線選択用のアクセス選択トランジスタも同様に提供されるが、図4Cには示されていない)。図4Cに示すように、垂直NORストリング対491は、ローカルソース線455、ローカルビット線454、及び任意選択のボディ接続456を共有する垂直NORストリング451a及び451bを含む。したがって、垂直NORストリング対491は、図4Bのアクティブ列430R及び430L上に形成された垂直NORストリングを表す。ワード線スタック423p-L及び423p-R(この例では、31p0)は、垂直NORストリング451a及び垂直NORストリング451bの制御ゲートをそれぞれ提供する。スタック内のゲートを制御するワード線は、基板に形成されたデコード回路によってデコードされ、アドレス指定されたTFF(すなわちアクティブ化されたワード線)及びアドレスされていないTFF(すなわち、ストリング中の他の全ての非アクティブ化ワード線)に適切な電圧が印加されることが確実となる。図4Cは、図4Bのアクティブ列430L及び431R上のストレージトランジスタ416L及び416Rには、同じワー

30

40

50

ド線スタック 4 2 3 p - R がどのように用いられるかを示す。したがって、垂直 NOR ストリング対 4 9 1 の垂直 NOR ストリング 4 5 1 b 及び垂直ストリング対 4 9 2 の垂直 NOR ストリング 4 5 2 a は、図 4 B のアクティブ列 4 3 0 L 及び 4 3 1 R 上に形成された隣接する垂直 NOR ストリングに対応する。垂直 NOR ストリング 4 5 1 a のストレージトランジスタ（例えば、ストレージトランジスタ 4 1 5 R）には、ワード線スタック 4 2 3 p - L が用いられる。

【 0 0 3 8 】

別の実施形態では、図 4 C のハードワイヤードグローバルソース線 4 1 3 - 1、4 1 3 - 2 が排除され、共有 N + ローカルソース線 4 5 5（垂直 NOR ストリングの両方 4 5 1 a 及び 4 5 1 b に共通）とその多数の関連するワード線 4 2 3 p - L 及び 4 2 3 p - R との間の寄生キャパシタ C によって置き換えられる。3 2 個の T F T の垂直スタックにおいて、3 2 本のワード線の各々は寄生キャパシタに寄与して総寄生キャパシタ C を提供し、プリチャージ T F T 4 7 0 によって供給される電圧を一時的に保持して仮想ソース電圧 V_{ss} 読み出しまたはプログラミング動作の比較的短い持続時間の間に行われる。この実施形態では、キャパシタ C に一時的に保持される仮想ソース電圧は、アクセストランジスタ 4 1 1 及びプリチャージトランジスタ 4 7 0 を介してグローバルビット線 G B L 1 からローカルソース線 4 5 5 に供給される。代替的に、それらの記憶機能に加えて、ワード線電圧をプログラムされた最高電圧よりも瞬間的に高くすることによって垂直 NOR ストリング内のメモリ T F T のうちの 1 つ以上がローカルソース線 4 5 5 をプリチャージするために使用される場合には、専用のプリチャージトランジスタ 4 7 0 を省略することができる。しかし、ストレージ T F T をこの目的のために使用する場合、ストレージ T F T の過剰プログラミングを避けるために注意を払わなければならない。仮想 V_{ss} 電圧を使用することは、ハードワイヤードグローバルソース線（例えば、G L S 1、G L S 2）及びそれらに付随する復号回路及びアクセストランジスタを排除し、それによってプロセスフロー及び設計上の課題を実質的に単純化し、大幅にコンパクトな垂直 NOR ストリングを実現するという重要な利点を提供する。

【 0 0 3 9 】

図 5 は、本発明の一実施形態による、アクティブ列 5 3 1 のグローバル NOR 型ビット線 5 1 4 - 1（G B L 1）、グローバルソース線 5 0 7（G S L 1）及び共通ボディバイアスソース 5 0 6（ V_{bb} ）への垂直 NOR ストリングの接続を示す Z Y 平面内の断面図である。図 5 に示すように、ビット線アクセス選択トランジスタ 5 1 1 は、G B L 1 をローカルビット線 5 5 4 に接続し、埋め込みコンタクト 5 5 6 は、アクティブストリップ上の P ボディ領域を、基板内のボディバイアスソース 5 0 6（ V_{bb} ）に任意選択で接続する。ビット線アクセス選択トランジスタ 5 1 1 は、図 5 のアクティブ列 5 3 1 の上に形成されている。しかし、代替的に、ビット線アクセス選択トランジスタ 5 1 1 は、アクティブ列 5 3 1 の底部または基板 5 0 5（図 5 には示されていない）に形成されてもよい。図 5 において、ビット線アクセス選択トランジスタ 5 1 1 は、例えば、アクセス選択ワード線 5 8 5 とともに、N + / P - / N + ドープされたポリシリコンスタックの独立した島状部に形成することができる。ワード線 5 8 5 を選択するために十分に大きな電圧が印加されると、P チャネルが反転され、それによってローカルビット線 5 5 4 が G B L 1 に接続される。ワード線 5 8 5 は、垂直 NOR ストリングの T F T への制御ゲートとして働くワード線 5 2 3 p と同じ方向（すなわち、Y 方向）に沿って延びる。ワード線 5 8 5 は、ワード線 5 2 3 p とは別個に形成することができる。一実施形態では、G B L 1 は X 方向に沿って（すなわち、ワード線の方向に垂直に）水平に延び、ビット線アクセス選択トランジスタ 5 1 1 はローカルビット線 5 5 4 へのアクセスを提供し、ローカルビット線 5 5 4 は、G B L 1 によって提供される多くの垂直 NOR ストリングのただ 1 つのローカルビット線である。読み出し及びプログラム動作効率を高めるために、マルチゲート NOR ストリングアレイにおいて、数千のグローバルビット線を用いて、ワード線 5 8 5 によってアクセスされる数千の垂直 NOR ストリングのローカルビット線に並列にアクセスすることができる。図 5 において、ローカルソース線 5 5 5 はコンタクト 5 5 7 を介してグローバ

ルソース線 5 1 3 - 1 (G S L 1) に接続され、例えば基板 5 0 5 の回路をデコードすることによってデコードされ得る。あるいは、既に説明したように、グローバルソース線は、ローカルソース線 5 5 5 に仮想ソース電圧 V_{SS} を供給し、T F T 5 7 0 を介してローカルソース線 5 5 5 の寄生キャパシタ 5 6 0 (すなわち、寄生キャパシタ C) を一時的にプリチャージすることによって除去することができる。

【 0 0 4 0 】

基板 5 0 5 に形成されたサポート回路は、アドレスエンコーダ、アドレスデコーダ、センスアンプ、入出力ドライバ、シフトレジスタ、ラッチ、比較セル、電源線、バイアス及び基準電圧発生器、インバータ、N A N D、N O R、排他的論理和やその他の論理ゲート、他のメモリ素子、シーケンサ、ステートマシンなどを含む。マルチゲート N O R ストリングアレイは、複数のマルチゲート N O R ストリングアレイを有する複数のブロックの回路として編成することができる。

10

【 0 0 4 1 】

図 6 A は、図 4 C に関連して上述したように、垂直 N O R ストリング対 4 9 1 の垂直 N O R ストリング 4 5 1 a の T F T 6 8 5 (T_L) 及び垂直 N O R ストリング 4 5 1 b の T F T 6 8 4 (T_R) のを示す X - Y 平面内の断面図である。

図 6 に示すように、T F T 6 8 4 及び 6 8 5 は、N + ローカルソース領域 6 5 5、及び N + ローカルドレインまたはビット線領域 6 5 4 を共有し、両方の領域は、Z 方向に沿って細長いピラーをなすように延びる。(N + ローカルソース領域 6 5 5 は図 4 A のローカルソース線 4 5 5 に対応し、N + ローカルドレイン領域 6 5 4 は図 4 A のローカルビット線 4 5 4 に対応する。) この実施形態では、P ドープされたチャンネル領域 6 5 6 L 及び 6 5 6 R は、ローカルソースピラー 6 5 5 とローカルドレインピラー 6 5 4 との間に一对のアクティブストリングを形成し、それらは分離領域 6 4 0 によって互いに分離され、Z 方向に沿って延びる。ワード線 6 2 3 p - L (W L 3 1 - 0)、6 2 3 p - R (W L 3 1 - 1) とチャンネル領域 6 5 6 L、6 5 6 R の外側との間には電荷トラップ層 6 3 4 が形成される。電荷トラップ層 6 3 4 は、例えば、トンネル誘電体 (例えば、二酸化シリコン) の薄膜、続いて窒化シリコンのような電荷トラップ材料の薄層または非導電性誘電材料に埋め込まれた導電性ナノドット、または分離されたフローティングゲートからなるトランジスタのゲート誘電体材料であり得、O N O (O x i d e - N i t r i d e - O x i d e) などのブロック誘電体の層または酸化アルミニウムまたは酸化ハフニウムなどの高誘電率膜、またはそのような誘電体の組み合わせがキャップされる (上層をなす)。ソース - ドレイン導電部は、ワード線 6 2 3 p - L 及び 6 2 3 p - R によってそれぞれ制御され、電荷トラップ層 6 3 4 の外側に制御ゲートを形成する。T F T 6 8 4 (T_R) をプログラムまたは読み出しするとき、ワード線 6 2 3 p - L に適切な禁止電圧を維持することによって T F T 6 8 5 (T_L) をオフにする。同様に、T F T 6 8 5 (T_L) をプログラムまたは読み出しするとき、ワード線 6 2 3 p - R に適切な禁止電圧を維持することによって T F T 6 8 4 (T_R) をオフにする。

20

30

【 0 0 4 2 】

図 6 A に示す実施形態では、ワード線 6 2 3 p - L 及び 6 2 3 p - R の輪郭形状は、消去時の逆トンネリング効率を低下させ、プログラミング中に T F T 6 8 4 及び 6 8 5 へのトンネル効率を高めるような輪郭形状に形成されている。具体的には、当業者に知られているように、チャンネル領域 6 5 6 R の湾曲 6 7 5 によって、プログラミング中にアクティブチャンネルのポリシリコンとトンネル誘電体との間の界面における電界を増幅し、消去中にワード線とブロック誘電体との間の界面における電界を減少させる。この機能は、マルチレベルセル (M L C) 構成で T F T トランジスタ 1 つあたり 2 ビット以上を格納する場合に特に役立つ。この技術を使用すると、各 T F T に 2、3 または 4 ビットまたはそれ以上を格納することができる。実際、T F T 6 8 4 及び 6 8 5 は、記憶状態の連続体を有するアナログストレージ T F T として使用することができる。プログラミングシーケンス (後述する) の後、電子は、破線 6 8 0 によって概略的に示されるように、電荷トラップ層 6 3 4 にトラップされる。図 6 A において、グローバルビット線 6 1 4 - 1 及び 6 1

40

50

4 - 2 は、ワード線 6 2 3 p - R 及び 6 2 3 p - L に垂直に延び、図 4 C のビット線 4 1 4 - 1 及び 4 1 4 - 4 に対応する垂直 NOR ストリングの上または下のいずれかに設けられる。図 2 に関連して上述したように、ワード線は、X 方向に沿ってメモリブロック（メモリ構造）1 0 0 の全長にわたって延び、グローバルビット線は Y 方向に沿ってメモリブロック（メモリ構造）1 0 0 の幅にわたって延びる。重要な点として、図 6 A において、ワード線 6 2 3 p - R は、ワード線 6 2 3 p - R の反対側の 2 つの垂直 NOR ストリングの T F T 6 8 4 及び 6 8 3 によって共有される。したがって、グローバルビット線 6 1 4 - 1（G B L 1）はローカルドレインまたはビット線領域 6 5 7 - 1（「奇数アドレス」）に接触し、T F T 6 8 4 及び 6 8 3 は独立して読み出したりはプログラムされることを可能にするために、グローバルビット線 6 1 4 - 2（G B L 2）は、ローカルドレインまたはビット線領域 6 5 7 - 2（「偶数アドレス」）に接触する。この効果を達成するために、グローバルビット線 6 1 4 - 1 及び 6 1 4 - 2 に沿ったコンタクトは千鳥状に互い違いに配置され、各グローバルビット線は、X 方向行に沿って 1 つおきの垂直 NOR ストリング対と接触する。

10

【0043】

同様に、マルチゲート NOR ストリングアレイの底部または上部に位置するグローバルソース線（図 6 A には示されていない）は、グローバルビット線に平行に延び、偶数または奇数アドレスで垂直方向 NOR ストリング対のローカルソース線に接触し得る。代替的に、寄生キャパシタ C（すなわち、キャパシタ 6 6 0）の一時的な仮想電源電圧 V_{ss} へのプリチャージが使用される場合、グローバルソース線を設ける必要はなく、デコードスキーム及びプロセスの複雑さが単純化される。

20

【0044】

図 6 A は、垂直 NOR ストリング対が積み重ねられたワード線で提供される可能ないくつかの実施形態のうちの一つのみを示す。例えば、チャンネル領域 6 5 6 R の湾曲 6 7 5 を更に大きな湾曲にすることができる。逆に、図 6 B の実施形態に示すように、そのような湾曲は完全に除去（すなわち真っ直ぐに）され得る。図 6 B の実施形態では、図 6 A の分離領域 6 4 0 の間隔は、チャンネル領域 6 5 6 L 及び 6 5 6 R を単一の領域 6 5 6（L + R）に併合することによって低減または完全になくなり、デュアルチャンネル構成（例えば、T F T 6 8 5（ T_L ）及び 6 8 4（ T_R ）が、同じアクティブストリップの対向する面に存在する。）を犠牲にすることなくより大きな面積効率を達成することができる。図 6 A、図 6 B の実施形態では、ワード線を共有する垂直 NOR ストリングは、互いに対して千鳥状に互い違いに配置され（図示せず）るが、それらは、各垂直 NOR ストリングの有効フットプリントを低減させるために、互いに近接して配置することができる。図 6 A 及び図 6 B は、グローバルビット線 6 1 4 - 1 と N + ドープローカルドレインビット線ピラー 6 5 4（L B L - 1）との間のコンタクトを介した直接接続を示しているが、このような接続はビット線アクセス選択トランジスタ（例えば図 5 のビット線アクセス選択トランジスタ 5 1 1、図 6 A 及び図 6 B には示されていない）を使用して達成することもできる。

30

【0045】

図 6 A 及び図 6 B の実施形態では、N + ドープローカルドレイン領域 6 5 4 とそれに隣接するローカル N + ドープソース領域 6 5 8（図 4 A の分離領域 4 3 6 に対応）との間の誘電体分離を、例えばワード線 6 2 3 p - R 及び 6 2 3 p - L が 2 つの背中合わせの電荷トラップ層の厚さよりも小さく、堆積処理中に電荷トラップ層がそれらに併合されるように画定することによって確立することもできる。この結果形成される堆積された電荷トラップ層の併合状態によって所望の誘電体分離が形成される。代替的には、隣接するストリングの N + ピラー 6 5 4 から一方のストリングの N + ピラー 6 5 8 を隔離するギャップ 6 7 6（エアギャップまたは誘電体充填）を作り出す（すなわち、図 4 A に示すギャップ（分離領域 4 3 6）を作り出す）べく、N + ポリシリコンの高アスペクト比のエッチングを用いることによって隣接するアクティブストリング間の分離を達成することができる。

40

【0046】

従来技術の垂直 NAND ストリングと本発明の垂直 NOR ストリングとの間の対照的な

50

点を述べると、いずれのタイプのデバイスも、制御ゲートと同様のワード線スタックを有する薄膜トランジスタを使用するが、そのトランジスタの向きが異なっていることであり、従来技術のNANDストリングでは、垂直アクティブストリップの各々が、直列に接続された32個、48個またはそれ以上の個数のTFETを有し得る。これとは異なり、本発明の垂直NORストリングを形成する各アクティブ列は、並列に接続された1組または2組の複数の個数(32個、48個またはそれ以上の個数)のTFETを有し得る。従来技術のNANDストリングでは、いくつかの実施形態におけるワード線は、一般的には、アクティブストリップを包み込む(「ラップアラウンド」する)ように延在する。本発明の垂直NORストリングのいくつかの実施形態では、図4C、図6A及び図6Bに示すように、各アクティブストリップに対して個別に指定された左右のワード線が使用され、それによって各グローバルビット線に対する倍増(すなわちペア)した記憶密度が達成される。本発明の垂直NORストリングは、プログラム妨害または読み出し妨害の問題が生じず、従来技術のNANDストリングの遅いレイテンシの問題もない。したがって、垂直方向のNORストリングにおいて、垂直NANDストリングにおけるより多数個のTFETを備えることができる。しかし、垂直NORストリングは、長い垂直ソース及びドレイン拡散部(例えば、図4Aに示すローカルソース領域455及びローカルドレイン領域454)間の低閾値または他のリークの影響をより受け易い。

10

【0047】

本発明の垂直NORストリングの2つの追加の実施形態を図6C及び図6Dに示す。これらの実施形態では、各ワード線スタック内の全てのワード線が垂直アクティブストリップを包み込む(ラップアラウンドする)。

20

【0048】

図6Cでは、垂直方向のNORストリングが、金属ワード線のスタック及びワード線間の誘電体分離層をエッチングすることによって形成される空隙の内部に形成される。製造プロセスフローは、垂直NORストリングにおけるトランジスタが、垂直NANDストリングにおいて直列ではなく、互いに並列に提供されるという点を除いて、従来技術の垂直NANDストリングと同様である。垂直NORストリングにおけるトランジスタは、非ドーブまたは低濃度ドーブされたチャネル領域656を備え、垂直NORストリングに沿って全てのTFETのための共有ローカルソース線655(LSL)及び共有ローカルビット線654(ドレイン)(LBL)を提供する空隙の深さ全体に延びるN+ドーブ垂直ピラーによって形成される。電荷蓄積素子となる電荷トラップ領域634は、チャネル656とワード線スタック623pとの間に配置され、垂直アクティブストリップに沿って2個、4個、8個、32個、64個以上のTFETのスタック(例えば、デバイス685(T₁₀))を形成する。図6Cの実施形態では、ワード線スタック623pはY方向に延び、個々の水平ストリップ(WL31-0)、(WL31-1)は、誘電体分離またはエアギャップ610によって互いに分離されている。グローバルビット線614(GBL)及びグローバルソース線615(GSL)は、ワード線に垂直なX方向に沿って行方向に水平に延びている。グローバルビット線614の各々は、メモリアレイの下またはその上のいずれかに配置され得るアクセス選択トランジスタ(図5の511、ここでは図示せず)を介し、垂直ストリップの行に沿ってローカルビット線ピラー654(LBL)にアクセスする。同様に、各グローバルソース線615は、その行に沿ってローカルソース線ピラーにアクセスする。図6A及び図6Bに示された構造は、図6Cの実施形態では単一の垂直NORストリングによって占められた概ね同じ領域に一对の垂直NORストリングを適合させることができ、図6Cに示す各垂直NORストリングにおける各TFETは、2つの平行な導電チャネル(すなわち、チャネル領域656a及び656b)を有し、したがって、より多くの電荷を蓄積し、読み出し電流を増加すなわち倍増させることができる。

30

40

【0049】

図6Dは、本発明の一実施形態によるラップアラウンドワード線を有するよりコンパクトな垂直NORストリングを示す。図6Dに示すように、垂直NORストリングは千鳥状に近接して配置されているので、ワード線スタック623p(WL31-0)を、より多

50

くの垂直NORストリングによって共有させることができる。この千鳥状配置は、ローカルソース線ピラー655(LSL)の寄生容量C(すなわち、キャパシタ660)の使用によって可能になる。以下で説明するように、読み出し及びプログラム動作中に仮想電圧 V_{SS} を一時的に保持するためにキャパシタ660をプリチャージすることによって、ハードワイヤードグローバルソース線(例えば、図6cのGSL615)が必要なくなる。図6C及び図6Dの垂直NORストリングは、従来技術の垂直NANDストリング(例えば、図1CのNANDストリング)と比較して、それ自体の面積効率は有効に高めないこともあるが、このような垂直NORストリングは、垂直NANDストリングよりも非常に長い。例えば、本発明の垂直NORストリングは、各スタック内の長さ128個分~512個分またはそれ以上のTFTのストリングを十分にサポートすることができるが、その

10

【0050】

[製造プロセス]

図7A、図7B、図7C及び図7Dは、本発明の一実施形態による、マルチゲートNORストリングアレイの製造プロセスで形成される中間構造の断面図である。

【0051】

図7Aは、本発明の一実施形態による、低抵抗層723pが基板701上に形成された後の半導体構造700のZ-Y平面内の断面図を示す。この例では、pは0~31の整数であり、32本のワード線のそれぞれを表す。図7Aに示すように、半導体構造700は低抵抗率層723-0~723-31を含む。半導体基板701は、例えば、垂直NORストリングを形成する前に、メモリ構造700のための支持回路が形成されているPドーパブルクシリコンウェハを表す。このようなサポート回路は、アナログ及びデジタル論理回路の両方を含むことができる。そのようなサポート回路のいくつかの例は、シフトレジスタ、ラッチ、センスアンプ、比較セル、電源ライン、バイアス及び基準電圧発生器、インバータ、NAND、NOR、排他的ORやその他の論理ゲート、入出力ドライバ、ビット線及びワード線デコーダを含むアドレスデコーダ、他のメモリ素子、シーケンサ及びステートマシンを含み得る。これらのサポート回路を提供するために、当業者に知られているように、従来のNウェル、Pウェル、トリプルウェル(図示せず)、 N^+ 拡散領域(例えば領域707-0)及び P^+ 拡散領域(例えば領域706)、低電圧トランジスタ及び高電圧トランジスタ、キャパシタ(コンデンサ)、抵抗器、ダイオード及び相互接続部が設けられる。

20

30

【0052】

サポート回路が半導体基板701の中及び上に形成された後、絶縁層708が設けられ、この絶縁層として例えば厚い二酸化シリコンの堆積または成長され得る。いくつかの実施形態では、グローバルソース線713-0を含む1つ以上の金属相互接続層が形成されてもよく、これは所定の方向に沿って延びる水平の細いストリップ(帯状体)として提供され得る。グローバルソース線713-0は、エッチングされた開口714を介して基板701内の回路707に接続されている。以下、本明細書では説明を容易にするために、グローバルソース線はX方向に沿って延びていると仮定する。金属相互接続線は、1層または複数層の堆積された金属層上にフォトリソグラフィパターンニング及びエッチングを適用することによって形成できる。(あるいは、これらの金属相互接続線は、従来の銅またはタングステンのダマシ配線プロセスなどの従来のダマシ配線プロセスを使用して形成できる。)次いで、厚い誘電体層709を堆積し、続いて従来の化学機械研磨(CMP)を用いて平坦化する。

40

【0053】

導体層723-0~723-31が連続して形成され、各導体層は、その下の層及びその上の層と介在する絶縁層726によって絶縁されている。図7Aでは、32本の導体層が示されているが、このような層は任意の数だけ設けることができる。実際には、設けることができる導体層の数は、プロセス技術に依存し、例えば複数の導体層及びそれらの間

50

の誘電体分離層 726 を切断することを可能にする、十分に制御された異方性エッチングプロセスが利用可能か否かで決まってくる。例えば、導体層 723 p は、まず厚さ 1 ~ 2 nm の窒化チタン層 (TiN) を堆積させた後、厚さ 10 ~ 50 nm のタングステン (W) または類似の高融点金属、またはシリサイド (とりわけニッケル、コバルトまたはタングステンのシリサイド、またはサリサイド等) の層を堆積させ、その後、酸化アルミニウム (Al₂O₃) などのエッチング停止材料の薄層を堆積させることによって形成することができる。各導体層は、堆積後にブロック 700 でエッチングされるか、または従来のダマシン配線プロセスによるブロックとして堆積される。図 7 A に示す実施形態では、各連続する導体層 723 p は、Y 方向において、直前の金属層の端部に接してない距離 (凹部 727 の凹み) だけ延在して (すなわち、そこから凹んで) いて、全ての導体層は、プロセスの後の段階で構造 700 の上部から接触可能となるように形成される。しかし、図 7 A の階段状導体スタックを形成するのに必要なマスクング及びエッチングステップの数を減らすために、個々の導体面を個別にマスクしエッチングして露出した凹部 727 の表面を形成することを必要としない、当業者に公知の他のプロセス技術を用いることによって、複数の導体層に対して同時に凹部 727 の表面を形成することが可能である。導体層が堆積され、エッチングされた後、対応する誘電体分離層 726 が堆積される。誘電体分離層 726 は、例えば、15 nm ~ 50 nm の厚さの二酸化シリコンであり得る。従来の CMP は、次の導体層を堆積するために各誘電体層の表面を準備する。ブロック 700 のスタック内の導体層の数は、垂直 NOR ストリング内の少なくともメモリ TFT の数と、それに加えられた追加の導体層に対応し、当該追加の導体層は、プリチャージ TFT (例えば、図 5 のプリチャージ TFT 575) などの非メモリ TFT の制御ゲートとして、またはビット線アクセス選択 TFT (例えば、図 5 の 585 ビット線アクセス選択 TFT 511) の制御ゲートとして使用され得る。

10

20

30

40

50

【0054】

誘電体分離層 710 及びハードマスク層 715 が次に堆積される。ハードマスク 715 は、導体層 723 p をエッチングして、まだ形成されていないワード線の長いストリップを形成するようにパターン化される。ワード線 623 p - R、623 p - L に対するマスクングパターンの一例が図 6 に示されており、当該パターンは、分離部 (ギャップ) 676 の位置で互い向かって接近するワード線内の延長部及び所望の湾曲 675 を生成するための各ワード線内の凹部などのフィーチャを含む。導体層 723 p の底部の誘電体層 709 に達するまで、連続した導体層 723 p 及びそれらに入り込む誘電体分離層 726 を介して異方性エッチングすることによって深いトレンチが形成される。多数の導体層がエッチングされるので、フォトレジストマスクそれ自体は、多数の連続するエッチングを通じて所望のワード線パターンを保持するのに十分にロバスト (頑強) ではない可能性がある。ロバストなマスクを提供するためには、当業者に知られているように、ハードマスク層 715 (例えば、カーボン) が好ましい。エッチングは、誘電体材料 709、またはグローバルソース線上のランディングパッド 713、または基板 701 で終了してもよい。ランディングパッド 713 をエッチングから保護するためのエッチング停止バリア膜 (例えば、酸化アルミニウム) を設けることが好都合であり得る。

【0055】

図 7 B は、本発明の実施形態による半導体構造 700 の Z - X 平面の断面図であって、図示のように、連続的な導体層 723 p 及び対応する誘電体層 726 をエッチングして、誘電体層 709 に達するトレンチ (例えば、深いトレンチ 795) が形成される。図 7 B では、図示のように、導体層 723 p が異方性エッチングされて、深いトレンチ 795 によって互いに分離された導体スタック 723 p - R 及び 723 p - L を形成している。この異方性エッチングは高アスペクト比のエッチングである。最良の結果を達成するために、当業者に知られているように、異なる層の材料がエッチングされるので、エッチング化学物質を導体材料エッチングと誘電体エッチングとの間で交互に変えることが必要になり得る。スタックの底部で結果的に形成されるワード線の導体幅及びトレンチ間隔が、スタックの最上部または最上部のワード線の対応する幅及び間隔とほぼ同じとなるように、いず

れの層のアンダーカットも回避されるべきなので、多段階エッチングの異方性が重要である。当然のことながら、スタック内の導体層の数が多いほど、多数の連続したエッチングを通じて厳しいパターン許容差を維持することがより困難になる。例えば64または128以上の導体層を介してエッチングすることに伴う困難を軽減するために、エッチングは、例えばそれぞれ32層のセクションごとに行うことができる。別個にエッチングされた部分は、例えば、非特許文献1に教示されているように、一体に縫合することができる。

【0056】

導体材料（例えば、タングステンまたは他の加工しにくい材料）の複数の導体層723pを通して行うエッチングは、介在する絶縁層726のエッチングよりもずっと困難で時間がかかる。このため、導体層723pの複数のエッチングの必要性を排除する代替プロセスを採用することができる。このプロセスは、第1に、当業者に周知であり、図7Bの導体層723pを、容易にエッチング可能な材料の犠牲層に置換することからなる。例えば、絶縁層726は二酸化シリコンであってもよく、犠牲層（図7bでは723pとして示される空間を占める）は、窒化シリコンまたは別の高速エッチング誘電材料であってもよい。次に、ONON（酸化物-窒化物-酸化物-窒化物）交互誘電体層を介して、深いトレンチを異方性エッチングして、デュアル誘電体（二重誘電体）の高いスタックを形成する。製造プロセスの後の段階（後述する）において、これらのスタックは、ポリシリコンのアクティブ垂直ストリップによって支持され、犠牲層を、好ましくは選択的な化学的または等方性エッチングによってエッチング除去可能とする。このようにして形成されたキャビティは、導体材料のコンフォーマルな堆積によって充填され、その結果、介在する絶縁層726によって分離された導体層723pが得られる。

【0057】

図7Bの構造が形成された後、電荷トラップ層734及びポリシリコン層730が、エッチングされた導体ワード線スタックの垂直側壁にコンフォーマルに連続して堆積される。得られた構造のZ-X平面における断面を図7Cに示す。図7Cに示されているように、電荷トラップ層734は、例えば、高誘電率の誘電体膜（例えば、酸化アルミニウム、酸化ハフニウム、または二酸化ケイ素及び窒化ケイ素のいくつかの組み合わせなど）からなる厚さが5~15nmのブロック誘電体732aを最初に堆積することによって形成される。その後、電荷トラップ材料732bが4~10nmの厚さに堆積される。電荷トラップ材料732bは、例えば、窒化シリコン、シリコンリッチな酸窒化物、誘電体膜に埋め込まれた導電性ナノドット、または同じ垂直アクティブストリップを共有する隣接するTFETから絶縁された薄い導電性フローティングゲートであってもよい。次いで、電荷トラップ732bは、2~10nmの範囲の厚さの堆積されたコンフォーマルな薄いトンネル誘電体膜（例えば、二酸化シリコン層、または酸化シリコン-窒化シリコン-酸化シリコン（「ONO」）三重層）によってキャップされてもよい。電荷トラップ層734から形成される記憶素子は、SONOS、TANOS、ナノドットストレージ、分離フローティングゲート、または当業者に知られている任意の適切な電荷トラッピングサンドイッチ構造のいずれかであり得る。電荷トラップ層734の結合厚さは、典型的には15~25nmである。

【0058】

電荷トラップ層734の堆積後、マスキング工程を用いて、トレンチ795の底部の電荷トラップ層734及び誘電体層709を通した異方性エッチングにより、トレンチ795の底部にコンタクト開口が形成され、コンタクト開口は、ソース電源電圧 V_{SS} のための下部グローバルソース線ランディングパッド713（図7B参照）を、またはグローバルビット線電圧 V_{BL} （図示せず）の領域を、またはバックバイアス供給電圧 V_{BB} に接触するためのP+領域706（図7C参照）を底部としている。一部の実施形態では、このエッチングステップの前に、トレンチ795の底部における電荷トラップ材料734のコンタクト開口エッチングの間にトンネル誘電体層732cの垂直表面を保護するために、ポリシリコンの超薄膜（例えば、2~5nmの厚さ）の堆積を行う。一実施形態では、各グローバルソース線は、垂直NORストリング対の行にお

10

20

30

40

50

いて交互の位置にのみ接続される。例えば、図 5 において、奇数番目のアドレスワード線については、電気コンタクト（例えば、コンタクト開口 5 5 7）がエッチングされて、N + ドープされたローカルソース線（例えば、図 5 のローカルソース線 5 5 5）をグローバルソース線 5 1 3 - 1 に接続する。同様に、偶数番目のアドレスワード線については、電気コンタクトがエッチングされて、垂直 NOR ストリング対の列における N + ドープされたローカルソース線をグローバルソース線 5 1 3 - 2（図 5 には図示せず）に接続する。寄生キャパシタ C（すなわち、図 5 のキャパシタ 5 6 0）を介して仮想 V_{SS} を使用する実施形態では、トレンチ 7 9 5 の底部の電荷トラップ層 7 3 4 をエッチングするステップをスキップすることができる。

【0059】

10

その後、ポリシリコン薄膜 7 3 0 が 5 ~ 10 nm の厚さに堆積される。図 7 C では、ポリシリコン薄膜 7 3 0 が、トレンチ 7 9 5 の対向する側壁に、それぞれ 7 3 0 R 及び 7 3 0 L と表示されている。ポリシリコン薄膜 7 3 0 はドープされていないか、または好ましくは、ホウ素を $1 \times 10^{16} / \text{cm}^3 \sim 1 \times 10^{17} / \text{cm}^3$ の範囲のドーピング濃度にドープされ、内部に形成される TFT がより大きい固有の閾値電圧を有し得るようにされる。トレンチ 7 9 5 は、その対向する側壁に電荷トラップ層 7 3 4 及びポリシリコン薄膜 7 3 0 を収容するのに十分な幅である。ポリシリコン 7 3 0 の堆積に続いて、上述したスタック内の犠牲層がエッチング除去され、こうして形成されたキャビティが、コンフォーマルに堆積された導体層 7 2 3 p で充填される（図 7 C）。

【0060】

20

図 7 B に示すように、トレンチ 7 9 5 は Y 方向に沿って延びる。独立したワード線スタック 7 2 3 p - L 及び 7 2 3 p - R の形成後、一実施例では、半導体構造 7 0 0 は、16000 個またはそれ以上の個数の並列ワード線スタックを有することができ、それぞれが、各スタックの長手方向に沿って形成された 8000 個またはそれ以上の個数のアクティブ列、すなわち 16000 個の TFT（スタックの両側に 8000 個の TFT）の制御ゲートとして機能する。各スタック内に 64 本のワード線があると、このようなマルチゲート垂直 NOR ストリングアレイのそれぞれに最終的に 160 億個の TFT が形成される。各 TFT が 2 つのデータビットを記憶する場合、そのようなマルチゲート垂直 NOR ストリングアレイは 32 ギガビットのデータを記憶する。このようなマルチゲート垂直 NOR ストリングアレイ（それに加えて、スペアアレイ）は、1 つの半導体基板上に約 32 個形成

30

【0061】

図 7 D は、一実施形態における図 7 C の構造の上側表面の X - Y 平面における断面図である。ワード線 7 2 3 p - L と 7 2 3 p - R との間には、垂直に堆積された P ドープされたポリシリコン構造（すなわちアクティブな列）の 2 つの側壁 7 3 0 L 及び 7 3 0 R が存在する。側壁 7 3 0 L と 7 3 0 R との間の深い空隙 7 4 0 は、高速エッチング絶縁誘電体材料（例えば、二酸化シリコン、または液体ガラス、または炭素ドープ酸化シリコン）で充填され得る。次いで、上側表面は、従来の CMP を使用して平坦化され得る。次に、フォトリソグラフィ工程により、開口 7 7 6 及び 7 7 7 が露出され、続いて高アスペクト

比の選択エッチングが行われ、露出領域 7 7 6 及び 7 7 7 において高速エッチング誘電体材料がトレンチ 7 9 5 の底に達するまで掘り込まれる。エッチング中の過度のパターン劣化を回避するために、このエッチングステップにおいてハードマスクが必要とされることがある。掘り込まれた空隙は、インサイチュ（in situ）N + ドープされた（その場で N + ドープされた）ポリシリコンで充填される。N + ドープは、露出した空隙内の非常に薄く低濃度ドープされたアクティブポリシリコンピラー 7 3 0 L 及び 7 3 0 R に拡散して N + ドープされた領域を形成する。あるいは、空隙をインサイチュ N + ドープポリシリコンで充填する前に、空隙内の低濃度ドープポリシリコンを、短時間の等方性プラズマエッチングまたは選択的ウェットエッチングによってエッチング除去することができる。次に、CMP または上側表面エッチングにより、上側表面から N + ポリシリコンが除去され

40

50

、領域 754 (N+) 及び 755 (N+) に高濃度 N+ ポリシリコンのパイロン (塔状部) が残る。これらの N+ パイロンは、結果として形成される垂直 NOR ストリングにおける TFT のための共有垂直ローカルソース線及び共有垂直ローカルビット線を形成する。

【0062】

次に、誘電体分離層が堆積され、フォトリソグラフィ・マスキング及びエッチングステップを使用してパターン化される。エッチングステップにより、垂直ローカルビット線を水平グローバルビット線に接続するコンタクトを開く (例えば、図 6 に示すように、奇数番目アドレスではコンタクト (ビット線 657-1 に接続) からストリングスまで、偶数番目アドレスでは 657-2 からストリングスまで)。低抵抗の金属層 (例えば、タンゲステン) が堆積される。次に、堆積された金属をフォトリソグラフィ及びエッチング工程を用いてパターン化して、グローバルビット線 (例えば、図 6 に示すような、奇数番目アドレスにストリングス用のグローバルビット線 614-1 (GBL1)、ストリングス用グローバルビット線 614-2 (GBL2)) を形成する。あるいは、グローバルビット線は、従来の銅ダマシ配線プロセスを使用して形成することができる。全てのグローバルビット線ならびにワード線スタックの全ての金属層 723p (図 7A) は、エッチングされたピアホールによって、当業者には周知のように、基板内のワード線及びビット線デコード及び感知回路に接続される。スイッチ及び感知回路、デコーダ及び基準電圧源は、個々にまたはビット線及びワード線のいくつかで共有する形でグローバルビット線及びグローバルワード線に設けることができる。

【0063】

いくつかの実施形態では、ビット線アクセス選択トランジスタ (図 5 の 511) 及び関連する制御ゲートワード線 (例えば、図 5 のワード線 585) が、当業者に知られるように独立した垂直 N+P-N+ トランジスタとして形成され、これは、奇数番目及び偶数番目のグローバルビット線 (例えば、図 6A のビット線 614-1 及び 614-2) を、交互の奇数及び偶数アドレス (例えば、図 6A のローカルビット線 657-1 及び 657-2) において垂直 NOR ストリングに選択的に接続する。

【0064】

[読み出し動作]

垂直 NOR ストリングの TFT は並列に接続されているので、本発明の全ての実施形態において、アクティブ列 (垂直 NOR ストリング対が形成されたアクティブな列を含む) 内の全ての TFT は、好ましくは、共有ローカルソース線と共有ローカルビット線 (例えば、図 4C に示すローカルビット線 454 及びローカルソース線 455) との間の読み出し動作中のリーク電流を抑制するように、エンハンスメントモードにあるべき、すなわち、各 TFT が正のゲート-ソース間閾値電圧を持つべきである。エンハンスメントモード TFT は、約 1V のネイティブ TFT 閾値電圧を目標とする、典型的には $1 \times 10^{16} \sim 1 \times 10^{17} / \text{cm}^3$ の濃度のホウ素でチャネル領域 (例えば、図 7C の P チャネル領域 756) をドーピングすることによって達成される。このような TFT では、アクティブ列の垂直 NOR ストリング対の非選択ワード線は全て 0V に保持される。代替的には、読み出し動作は、共有ローカル N+ドレイン線 (例えば、図 4C のローカルソース線 455) 上の電圧を約 1.5V に上昇させ、共有ローカル N+ドレイン線 (例えば、ローカルビット線 454) を約 2V に保ち、選択されていない全てのローカルワード線を 0V に保持する。このような構成は、ワード線をソースに対して -1.5V に設定することと等価であり、それによって、例えば TFT がわずかに過剰消去された場合に生じる、わずかに空乏モードの閾値電圧の TFT によるリーク電流を抑制する。

【0065】

垂直 NOR ストリングの TFT を消去した後、過剰消去された (すなわち、空乏モードの閾値電圧を有する) 垂直 NOR ストリングの TFT をエンハンスメントモードの閾値電圧に戻すために、ソフトプログラミング動作が必要とされ得る。図 5 では、P チャネルがボディバイアスソース 506 のバックバイアス電圧 (V_{bb}) に接続される任意選択のコンタクト 556 が示されている (図 4C にボディ接続領域 456 としても示されている)

。共有電圧源と共有N + ドレイン/ローカルビット線との間のサブスレッショルドリーク電流を低減するために、各アクティブ列内のTFTの閾値電圧を変調するために V_{bb} に負電圧を使用することができる。いくつかの実施形態では、制御ゲートが0Vに保持されるトンネル消去TFTに対して、消去動作中に正の V_{bb} 電圧を使用することができる。

【0066】

垂直NORストリング対のTFTに記憶されたデータを読み出すために、マルチゲートNORストリングアレイ内の全てのワード線を0Vに保持することによって、垂直NORストリング対の両方の垂直NORストリング上の全てのTFTが最初に「オフ」状態に置かれる。アドレス指定された垂直NORストリングは、デコード回路を使用して、共通のワード線に沿っていくつかの垂直NORストリングの中の感知回路を共有することができる。あるいは、各垂直NORストリングは、グローバルビット線（例えば、図4CのGBL1）を介して専用の感知回路に直接接続されてもよい。後者の場合、同じワード線平面を共有する1つ以上の垂直NORストリングを並列に感知することができる。アドレス指定された垂直NORストリングのそれぞれのローカルソース線は $V_{ss} \sim 0V$ で設定されるが、この設定は、図8Aに概略的に示されるようにハードワイヤードグローバルソース線（例えば、図4CのGSL1）を介して行われるか、または図8Bに概略的に示されるように、プリチャージ中に $V_{b1} \sim 0V$ を瞬間的にフローティングローカルソース線455または355の寄生容量C（例えば、キャパシタ460またはキャパシタ360）に送り込むプリチャージトランジスタ（例えば、図4Cのプリチャージトランジスタ470または図3Cのトランジスタ317）を介して仮想 $V_{ss} \sim 0V$ として行われるかのいずれかである。

【0067】

プリチャージトランジスタ470をオフにした直後に、ローカルビット線（例えば、図4Cのローカルビット線454）は、ビット線アクセス選択トランジスタ（例えば、図4Cのビット線アクセス選択トランジスタ411または図5のアクセス選択トランジスタ511）を介して $V_{b1} \sim 2V$ に設定される。 $V_{b1} \sim 2V$ はまた、アドレス指定された垂直NORストリングに対するセンスアンプにおける電圧でもある。このとき、アドレス指定されたワード線は、0Vから典型的には約6Vまでの小さなインクリメント電圧ステップで上昇するが、奇数番目アドレスTFT及び偶数番目アドレスTFTの両方の非選択ワード線は、垂直NORストリング対の0Vのままである。図8Aのハードワイヤード V_{ss} の実施形態では、アドレス指定されたTFTは一実施例では2.5Vの閾値電圧にプログラムされているので、ローカルビット線LBLの電圧 V_{b1} は選択されたTFTを介して WLS が2.5Vを超えるとすぐにローカルソース線（ V_{ss} ）に接続され、選択されたグローバルビット線に対応するセンスアンプで検出される電圧降下（図8Aに破線矢印で示す）をもたらす。図8Bの仮想 V_{ss} の実施形態では、プリチャージトランジスタワード線 $WLCHG$ が瞬間的にオンにされ、読み出しシーケンスの開始時にフローティングローカルソース線 LSS を0Vにプリチャージする。次に、選択されたワード線 WLS は、その電圧インクリメントステップを経て、プログラムされた2.5Vを超えるとすぐに、選択されたTFTはそのローカルビット線上の電圧を $V_{b1} \sim 2V$ から瞬間的に低下させる。この電圧ディップ（降下）（図8Bに破線の矢印で示す）は、選択されたローカルビット線に接続されたグローバルビット線のセンスアンプによって検出される。当業者に知られているように、選択されたTFTのプログラムされた閾値電圧を正確に読み出すための他の代替方式が存在する。仮想電圧 V_{ss} を一時的に保持するための寄生容量Cに依存する実施形態では、垂直スタックが高ければ高いほど容量Cは大きくなり、したがってホールド時間は長くなり、選択されたセンスアンプに供給される読み出し信号は大きくなる点は指摘しておくべきであろう。Cを更に増加させるために、一実施形態では、容量を増加させることを第1の目的とする1つ以上のダミー導体を垂直ストリングに追加することが可能である。

【0068】

MLC実装（すなわち、各TFTが2ビット以上を記憶する「マルチレベルセル」実装

10

20

30

40

50

）の場合、アドレス指定されたTFTは、いくつかの電圧（例えば、1V（消去状態）、2.5V、4Vまたは5.5V）のうちの1つにプログラムされている可能性がある。アドレス指定されたワード線WL₅は、センスアンプでTFTの導通が検出されるまで、電圧インクリメントステップで電圧が上昇する。あるいは、単一のワード線電圧を印加することができ（例えば、最大約6ボルト）、ローカルビット線LBL（V_{b1}）の放電率は、記憶されたマルチビットの電圧状態を表すいくつかのプログラム可能な基準電圧からの放電率と比較することができる。このアプローチは、連続した状態に対して拡張することができ、効果的にアナログストレージを提供します。プログラマブル基準電圧は、マルチゲート垂直NORストリングアレイ内に配置された専用の基準垂直NORストリングに記憶され、読み出し、プログラム中の特性及びバックグラウンドリークが厳密に追跡される。垂直NORストリング対では、2つの垂直NORストリングのうち一方のTFTのみを各読み出しサイクルで読み出すことができ、そのとき他方の垂直NORストリング上のTFTは「オフ」状態（すなわち、全てのワード線が0V）に置かれる。読み出しサイクル中、垂直NORストリングのTFTのうち1つだけに読み出し電圧が印加されるので、読み出し妨害条件は本質的に存在しない。

10

【0069】

本発明の一実施形態の一例では、64個のTFTと1個以上のプリチャージTFTが垂直NORストリング対の各垂直NORストリングに設けられ得る。

各ワード線は、そのローカル垂直N+ソース線ピラーとの交差部において、キャパシタを形成する（例えば、図6Aのキャパシタ660を参照）。このようなキャパシタの典型的な容量は、例えば、 1×10^{18} F（ファラド）である。垂直NORストリングの両方の垂直NORストリングに全てのキャパシタを含めると、全体的な分布容量Cは約 1×10^{16} ファラドになり、この容量は、読み出しサイクル（典型的にプリチャージ動作の直後に1マイクロ秒未満で完了する）中にローカルソース線がプリチャージされたソース電圧（V_{ss}）を保存するのに十分である。ビット線アクセス選択トランジスタ411及びプリチャージTFT470による充電時間は数ナノ秒程度であり、この充電時間は読み出しレイテンシに顕著には加わらない。直列接続された多数のTFTが導通することが必要とされるNANDストリングの読み出し動作とは異なり、垂直NORストリングにおけるTFTからの読み出しは、読み出し動作が垂直NORストリングにおけるTFTのうち1つのみの導通しか伴わない。

20

30

【0070】

本発明の垂直NORストリングの読み出しレイテンシに寄与する主な要因が2つ存在し、それは（a）グローバルビット線（例えば、図6AのGBL614-1）の抵抗R_{b1}及び容量C_{b1}に関連するRC時間遅延、及び（b）アドレス指定されたTFTが導通を開始したときのローカルビット線（例えば、LBL-1）上の電圧降下V_{b1}に対するセンスアンプの応答時間である。グローバルビット線に関連するRC時間遅延は、例えば、16000の垂直NORストリングの場合、数十ナノ秒程度である。従来技術の垂直NANDストリング（例えば、図1BのNANDストリング）のTFTを読み出すための読み出しレイテンシは、32子以上の直列接続されたTFT及びグローバルビット線の容量C_{b1}を放電する選択トランジスタを通る電流によって決定される。対照的に、本発明の垂直NORストリングでは、読み出し電流放電C_{b1}は、ビット線アクセス選択トランジスタ411と直列の1つのアドレス指定されたトランジスタ（例えば、図4Aのトランジスタ416L）を介して供給され、これにより、ローカルビット線電圧（V_{b1}）の放電がより迅速になる。この結果、はるかに短いレイテンシが達成される。

40

【0071】

図4Cでは、一度に1つのTFT（例えば、垂直NORストリング451b内のTFT416L）が読み出されると、垂直NORストリング対491の垂直NORストリング451a及び451bのいずれかの他の全てのTFTが「オフ」状態に保持され、それらのワード線は0Vに保持される。垂直NORストリング対492の垂直NORストリング452a内のTFT416Rは、ワード線W31をTFT416Lと共有しているが、垂直

50

NORストリング452aにはグローバルビット線414-2が用いられるので、TF T 416LはTF T 416Lと同時に読み出され、ストリング451bには、グローバルビット線414-1が用いられる(図6A及び図6Bは、グローバルビット線614-1及び614-2が隣接する垂直NORストリング対がどのように提供されるかを示す)。

【0072】

一実施形態では、ワード線スタックは、32の平面に設けられた32本以上のワード線を含む。1つのマルチゲート垂直NORストリングアレイでは、各平面は16000個のTF Tを制御する8000本のワード線を含み、各ビット線は専用のセンスアンプに接続されていれば16000本のグローバルビット線を通して並列に読み出すことができる。あるいは、複数のグローバルビット線がデコード回路を介してセンスアンプを共有する場合、16000個のTF Tはいくつかの連続した読み出しサイクルにわたって読み出される。大量の放電TF Tを並列に読み出すと、チップの接地電源(V_{SS})に電圧バウンスが発生し、読み出しエラーが発生する可能性がある。しかし、ローカルソース線にプリチャージされた寄生キャパシタCを使用する(すなわち、垂直NORストリングのための仮想ソース電圧(V_{SS}))を提供する)実施形態は、そのような接地電圧バウンスが排除される点で特に有利である。

これは、垂直NORストリングの仮想電源電圧が独立しており、チップの接地電源に接続されていないためである。

【0073】

[プログラム(書き込み)及びプログラム禁止動作]

アドレス指定されたTF Tのプログラミングは、選択されたワード線(例えば、ワード線423p-R)とアクティブチャンネル領域(例えば、図4Aのボディ領域456のアクティブチャンネル領域)との間に高いプログラミング電圧が印加されたときの、TF Tのチャンネル領域(例えば、図4bで430Lとして示すチャンネル領域)から電荷トラップ層(例えば、電荷トラップ層434)への電子のトンネリング(直接トンネリングまたはファウラー・ノルドハイム・トンネリング)によって達成され得る。トンネリングは非常に効率的であり、TF Tをプログラムするのに必要な電流は殆どないため、低電力消費で数万個のTF Tの並列プログラミングを達成することができる。トンネリングによるプログラミングは、例えば、20V、100マイクロ秒のパルスが必要とすることがある。好ましくは、プログラミングは、約14Vで始まり、約20Vの高さになる一連のより短い持続時間の段階的な電圧パルスによって実施される。段階的な電圧パルスを用いることで、TF Tの電氣的ストレスが低減し、意図したプログラムされた閾値電圧のオーバーシュートが回避される。

【0074】

各プログラミング高電圧パルスの後、アドレス指定されたトランジスタは、それがその目標閾値電圧に達したかどうかを調べるために読み出される。目標閾値電圧に達していない場合、選択されたワード線に印加された次のプログラミングパルスは、典型的には数百mVだけインクリメントされる。このプログラムベリファイシーケンスは、0Vをアクティブ列(例えば、図4Bの列430L)のローカルビット線(例えば、図4Aのローカルビット線454)に印加させた状態で、1つのアドレスされたワード線(すなわち、制御ゲート)に繰り返し印加される。これらの高ワード線プログラミング電圧では、TF T 416Lのチャンネル領域は反転され、かつ0Vに保持されて、電子がTF T 416Lの電荷蓄積層にトンネリングする。読み出し感知が、アドレス指定されたTF Tがその目標閾値電圧に到達したことを示すとき、アドレス指定されたTF Tはさらなるプログラミングが禁止されなければならないが、同じワード線を共有する他のTF Tは、それらのより高い目標閾値電圧にプログラミングを続けることができる。例えば、垂直NORストリング451b内のTF T 416Lをプログラムする場合、垂直NORストリング451b及び451a内の他の全てのTF Tのプログラミングは、その全てのワード線を0Vに保持することによって禁止されなければならない。

【0075】

10

20

30

40

50

目標閾値電圧に達した後、TFT416Lに対してさらなるプログラミングを禁止するために、半選択電圧（すなわち約10V）がローカルビット線454に印加される。10Vがチャンネル領域に印加され、20Vが制御ゲート上に印加された状態では、正味10Vのみが電荷トラップ層に印加されるので、ファウラー・ノルドハイム・トンネリング電流は重要ではなく、最大20Vまでの階段状のパルス電圧の残りのシーケンス中には、TFT416Lに対して有意な更なるプログラミングは行われぬ。ワード線WL31上のプログラミング電圧パルスを上昇させ続けながらローカルビット線454を10Vに上昇させることにより、同じ選択されたワード線を共有する垂直NORストリング上の全てのTFTは、それらのより高い目標閾値電圧に正しくプログラムされる。数十万個のTFTをマルチレベルセルストレージにおけるそれらの様々な目標閾値電圧状態と並列に正確にプログラムするためには、「プログラム・読み出し・プログラム禁止」のシーケンスが不可欠である。このような個々のTFTの過剰プログラミングのプログラム禁止は、次のより高い目標閾値電圧状態の閾値電圧によるオーバーステアまたはマージを引き起こす可能性がある。TFT416R及びTFT416Lは同じワード線を共有するが、それらは異なる垂直NORストリング対452及び451に属する。それぞれのビット線電圧がGBL1及びGBL2を介して供給され、独立して制御されるので、同じプログラミングパルス電圧シーケンスでTFT416L及びTFT416Rの両方をプログラムすることが可能である。例えば、TFT416Lは、いつでもプログラミングすることができ、またTFT416Lは、いつでもプログラミングすることを禁止することができる。垂直NORストリング対491の垂直NORストリングs451a及び451bは、別々のワード線423p-L及び423p-Rによってそれぞれ制御されるので、これらのプログラミング及びプログラミング禁止電圧条件を満たすようにでき、各ローカルビット線の電圧は、他の全ての垂直NORストリング対とは独立して設定することができる。プログラミング中に、アドレス指定されたワード線スタック内またはアドレス指定されていないワード線スタック内の選択されていないワード線は、0V、半選択電圧10V、またはフロート状態にすることができる。グローバルソース線（例えば、図4CのGSL1）がソースアクセス選択トランジスタ（図4Cには図示せず）を介してアクセスされる実施形態では、プログラミング中にアクセス選択トランジスタがオフになり、プログラミング及びプログラミング禁止の間、ローカルソース線455の電圧がローカルビット線454の電圧に追従する。図4Cのキャパシタ460によって表されるその寄生容量Cによってローカルソース線上の電圧が供給される実施形態についても同様である。図4Cの実施形態では、グローバルソース線は存在するがソースアクセス選択トランジスタは存在せず、アドレス指定されたストリングのグローバルソース線413-1に印加される電圧は、好ましくは、プログラミング及びプログラミング禁止の間、アドレスされたグローバルビット線414-1の電圧を追跡すべきである。

【0076】

段階的に高くなる電圧プログラミングパルスの各々の後に、TFT416L及び416Rがそれらのそれぞれの目標閾値電圧に達したか否かを判定する読み出しサイクルが続く。目標閾値電圧に達している場合には、ドレイン、ソース及びボディ電圧は10Vに上昇され（あるいは、これらの電圧は10Vに近づくようにフロート状態にされる）、さらなるプログラミングが禁止されるが、ワード線WL31は、まだその目標閾値電圧に達していない同一平面上の他のアドレス指定されたTFTをプログラムし続ける。このシーケンスは、全てのアドレス指定されたTFTが正しくプログラムされるように読み出しベリファイされたときに終了する。MLCの場合、各アドレス指定されたグローバルビット線をいくつかの所定の電圧（例えば、記憶される2ビットデータの4つの異なる状態を表す0V、1.5V、3.0V、または4.5V）のうちの1つに設定し、次いで、階段状のプログラミングパルス（約20Vまで）をワード線WL31に印加することによって、複数の閾値電圧状態のうちの1つのプログラミングを加速することができる。このようにして、アドレス指定されたTFTは、有効なトンネリング電圧（すなわち、それぞれ20V、18.5V、17V、15.5V）のうちの所定の1つを受け取り、所定の閾値電圧の1

つが単一プログラミングシーケンスでT F Tにプログラムされる。その後、個々のT F Tレベルで細かい正確なプログラミングパルスを供給することができる。

【 0 0 7 7 】

[高速化全面並列プログラミング]

マルチゲート垂直N O Rストリングアレイの各ローカルソース線に固有の寄生容量Cのために、マルチゲート垂直N O Rストリングアレイの全てのローカルソース線は、高電圧パルスシーケンスを適用する前に、全ての垂直N O Rストリングに対して、(例えば、グローバルビット線G B L 1及びビット線アクセス選択トランジスタ4 1 1及びプリチャージトランジスタ4 7 0を介して)瞬間的に印加される0 V (プログラム用)または1 0 V (禁止用)を有することができる。この手順は、平面毎にワード線平面をアドレス指定することによって実行することができる。アドレス指定されたワード線平面の各々に対して、他のワード線平面上の全てのワード線を0 Vに保持しながら、そのアドレス指定されたワード線平面上の多くのまたは全てのワード線にプログラミングパルスシーケンスを印加して、アドレス指定された平面上に多数のT F Tを並列にプログラムし、次いで個々の読み出しベリファイが行うことができ、更に必要に応じて、適切にプログラムされたT F Tのローカルソース線をプログラミング禁止電圧にリセットすることができる。このアプローチは、プログラミング時間は比較的長い(すなわち、約1 0 0マイクロ秒)が、アドレス指定されたワード線平面を共有する全てのローカルソース線キャパシタをプリチャージまたは読み出しベリファイすることは1 0 0 0倍以上高速なので、大きな有利な効果をもたらす。したがって、各ワード線平面に可能な限り多くのT F Tを並列にプログラムすることが妥当である。この加速されたプログラミング機能は、シングルビットプログラミングよりもかなり遅いM L Cプログラミングで更に大きな有利な効果をもたらす。

【 0 0 7 8 】

[消去動作]

いくつかの電荷トラップ材料では、トラップされた電荷の逆トンネリングによって消去動作が実行され、この消去動作はむしろ遅くなり得るものであり、時には数十ミリ秒の2 0 Vまたはそれ以上のパルスを必要とする。したがって、消去動作は、垂直N O Rストリングアレイレベル(「ブロック消去」)で実施することができ、バックグラウンドで実行されることが多い。典型的な垂直N O Rストリングアレイは6 4のワード線平面を有し、各ワード線平面は例えば1 6 3 8 4 × 1 6 3 8 4のT F Tを制御し、合計約1 7億のT F Tを有する。したがって、1テラビットチップは、各T F Tに2ビットのデータが格納されている場合、そのような垂直N O Rアレイ配列を約3 0個含むことができる。いくつかの実施形態では、ブロック消去は、0 Vでブロック内の全てのワード線を保持しつつ、垂直N O Rストリング(例えば、図4 Cのボディ接続4 5 6及び図5のコンタクト5 5 6)において全てのT F Tによって共有されるPチャンネルに約2 0 Vを印加することによって実行することができる。消去パルスの持続時間は、ブロック内の大部分のT F Tが僅かなエンハンスメントモードの閾値電圧、すなわち0 Vと1 Vとの間で消去されるようにしなければならない。いくつかのT F Tは、オーバーシュートして空乏モード(すなわち、わずかに負の閾値電圧)に消去される。消去コマンドの一部として、過剰消去されたT F Tを、消去パルスの終了後にわずかなエンハンスメントモードの閾値電圧に戻すために、ソフトプログラミングが必要とされることがある。エンハンスメントモードにプログラムすることができない空乏モードT F Tのうちの1つを含むことがある垂直N O Rストリングは、スペアのストリングによって置き換えるべく不使用にする必要がある可能性がある。

【 0 0 7 9 】

代替的には、消去パルスをボディ(すなわち、P層)に供給するのではなく、垂直N O Rストリングアレイ内の全ての垂直N O Rストリング対上のローカルソース線及びローカルビット線(例えば、図4 Cのローカルソース線4 5 5及びローカルビット線4 5 4)の電圧が、消去パルスの持続時間の間、全てのワード線平面上の全てのワード線を0 Vに保持しつつ、約2 0 Vに上昇させる。このスキームでは、グローバルソース線及びグローバルビット線選択デコーダが、それらの接合部で2 0 Vに耐えることができる高電圧トラン

ジスタを使用する必要がある。あるいは、他の全ての平面上のワード線を 0 V に保持しながら、アドレス指定された平面上の全てのワード線に - 20 V パルスを印加することによって、アドレス指定されたワード線平面を共有する全ての T F T を一緒に消去することができる。垂直 N O R ストリング対の他の全ての電圧は 0 V に保持される。これにより、ワード線の 1 つのアドレス指定された平面によってタッチされる全ての T F T の X - Y スライスのみが消去される。

【 0 0 8 0 】

[半揮発性 N O R T F T ストリング]

垂直 N O R ストリングでの使用に適したいくつかの電荷トラップ材料（例えば、酸化物 - 窒化物 - 酸化物すなわち「 O N O 」）は、典型的には、何年ものオーダーの長期のデータ保持時間を有するが、比較的耐久性が低い（すなわち、典型的には 1 万サイクル以下のオーダーの回数の書き込み消去サイクルの後、性能が劣化する）。しかし、いくつかの実施形態では、保持時間は非常に短期間であるが、非常に向上した耐久性（例えば数分または数時間程度の保持時間、数千万回の書き込み消去サイクルが可能な耐久性）をもって電荷を蓄積する電荷トラップ材料を選択することができる。例えば、図 7 C の実施形態では、典型的には 6 ~ 8 n m の S i O₂ 層のトンネル誘電体層 7 3 2 c は、約 2 n m まで薄くすることが可能で、または別の誘電体材料（例えば、S i N）に置き換えることが可能である。はるかに薄い誘電体層により、電荷トラップ層への直接トンネリングによって電子を導入するのに、（より高い電圧を必要とするファウラー・ノルドハイム・トンネリングとは異なる）適度な電圧の使用が可能となり、この場合、電子は数分から数時間または数日間トラップされる。電荷トラップ層 7 3 2 b は、窒化シリコン、薄い誘電体膜に分散された導電性ナノドット、または分離された薄いフローティングゲートを含む他の電荷トラップ膜の組み合わせとすることができる。ブロッキング層 7 3 2 a は、二酸化ケイ素、酸化アルミニウム、酸化ハフニウム、窒化ケイ素、高誘電率誘電体、またはそれらの任意の組み合わせとすることができる。ブロッキング層 7 3 2 a は、電荷トラップ層 7 3 2 b 内の電子が制御ゲートワード線に逃げることを阻止する。トラップされた電子は、最終的に、超薄トンネル誘電体層の破損の結果として、または逆方向の直接トンネリングによって、アクティブ領域 7 3 0 R に漏れ出ることになる。電荷トラップ材料の他の組み合わせを使用することもできるが、組み合わせると、耐久性は高くなるものの、失われた電荷を補充するために定期的な書き込みまたは読み出しリフレッシュ動作を必要とする低保持の「半揮発性」ストレージ T F T になる。本発明の垂直 N O R ストリングは、比較的高速の読み出しアクセス（すなわち、低レイテンシ）を有するので、現時点では、それをダイナミックランダムアクセスメモリ（D R A M）の使用を必要とするいくつかの用途で使用することができる。本発明の垂直方向 N O R ストリングは、3 次元スタックに組み込むことができない D R A M より消費電力がはるかに小さく、D R A M が数ミリ秒ごとリフレッシュを必要とするのに対してリフレッシュサイクルは数分または数時間ごとに約 1 回だけの実行で済むので、D R A M よりもビット当たりのコストがはるかに低くなるという顕著な利点を有する。本発明の 3 次元半揮発性ストレージ T F T は、電荷トラップ材料のための適切な材料（例えば、上述のもの）を選択し、プログラム / 読み出し / プログラム禁止 / 消去条件を適切に適合させ、定期的なデータのリフレッシュを組み込むことによって達成される。

【 0 0 8 1 】

[N R O M / ミラービット N O R T F T ストリング]

本発明の別の実施形態では、垂直 N O R ストリングは、当業者に知られている 2 次元 N R O M / ミラービットトランジスタで使用されるチャンネルホットエレクトロン注入法を使用してプログラムすることができる。一例として図 4 A の実施形態を使用すると、チャンネルホットエレクトロン注入のプログラミング条件は、制御ゲート（すなわちワード線 4 2 3 p）で 8 V、ローカルソース線 4 5 5 で 0 V、ローカルドレイン線 4 5 4 で 5 V とすることができる。1 ビットを表す電荷は、ローカルビット線 4 5 4 との接合部に隣接する（ボディ領域 4 5 6 の）チャンネル領域の一端の電荷蓄積層に蓄積される。ローカルソース線

10

20

30

40

50

455及びローカルビット線454の極性を反転させることによって、第2のビットを表す電荷がプログラムされ、ローカルソース線455との接合部の隣のチャンネル領域456の反対側の端部の電荷蓄積層に記憶される。両方のビットを読み出すためには、当業者には知られるように、プログラミングの逆の順序での読み出しを必要とする。チャンネルホットエレクトロンプログラミングは、直接トンネリングまたはファウラー・ノルドハイム・トンネリングによるプログラミングよりもはるかに効率が悪いので、トンネリングで可能な超並列プログラミングに適していない。しかし、各TFTは2倍のビット密度を持ち、アーカイブメモリなどの用途においては魅力的である。NROM TFTの消去は、トラップされた電子の電荷を中和するためのバンド間トンネリング誘起ホットホール注入を利用する従来のNROM消去メカニズムを用いる、すなわちワード線に-5V、ローカルソース線455に0V、ローカルビット線454に5Vを供給することによって達成することができる。あるいは、NROM TFTは、0Vのワード線を有するボディ領域456に高い正の基板電圧 V_{bb} を印加することによって消去することができる。チャンネルホットエレクトロン注入プログラムに伴う高いプログラミング電流のため、垂直NROM TFTストリングの全ての実施形態は、図3A及び図6Cの実施形態のように、ハードワイヤードローカルソース線及びローカルビット線を使用しなければならない。

10

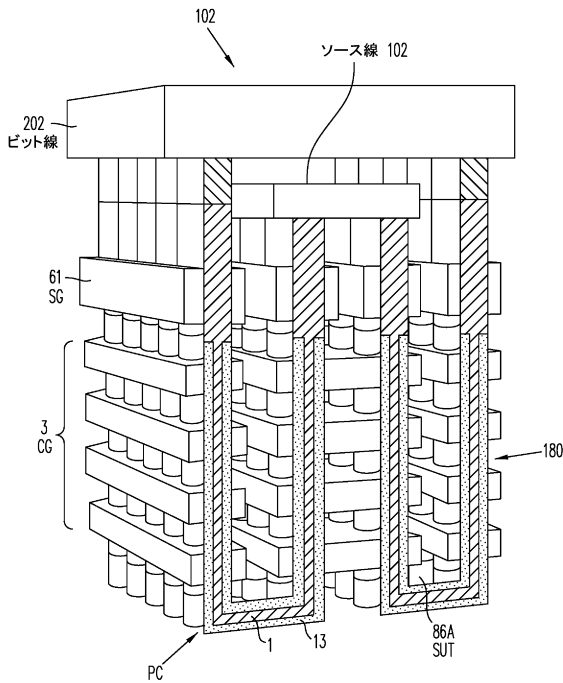
【0082】

上記の詳細な説明は、本発明の特定の実施形態を例示するために提供されるものであって、限定することを意図するものではない。本発明の範囲内で多くの変形及び変更が可能である。本発明は、特許請求の範囲の請求項に記載されている。

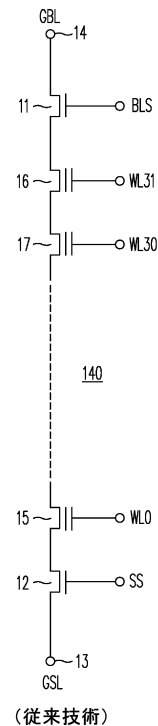
20

【図面】

【図1A】



【図1B】

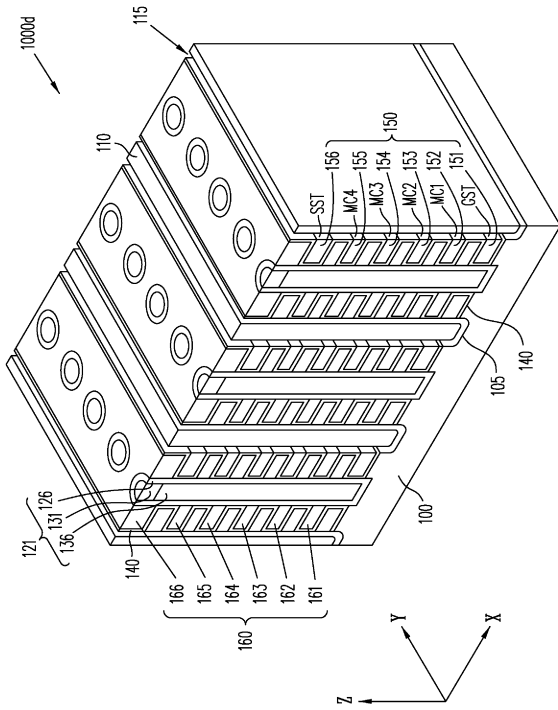


30

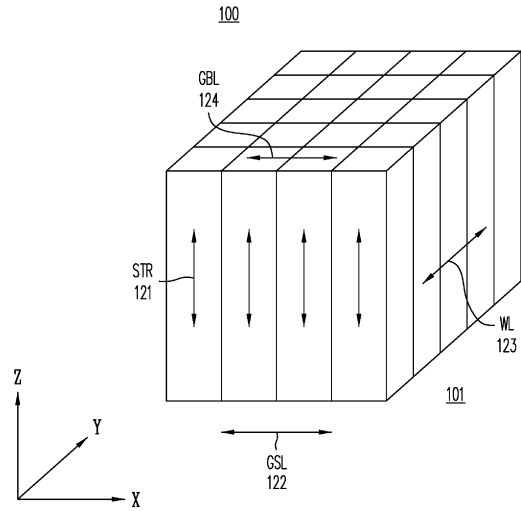
40

(従来技術)

【 図 1 C 】



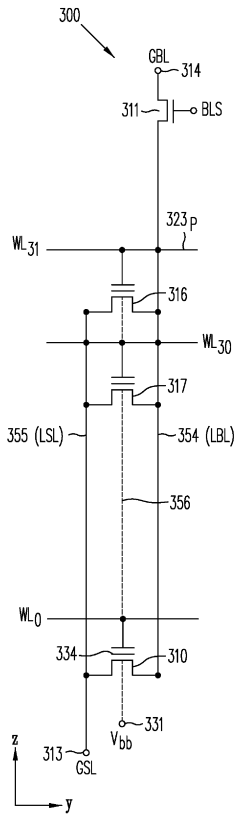
【 図 2 】



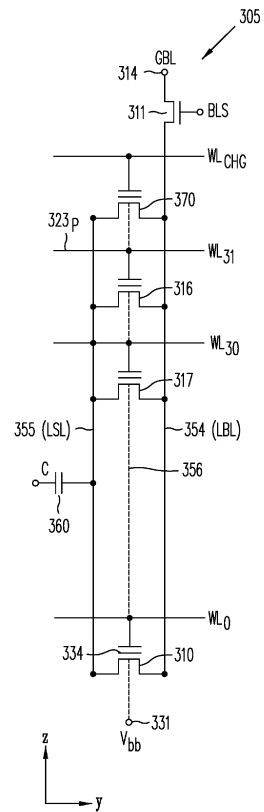
10

20

【 図 3 A 】



【 図 3 B 】

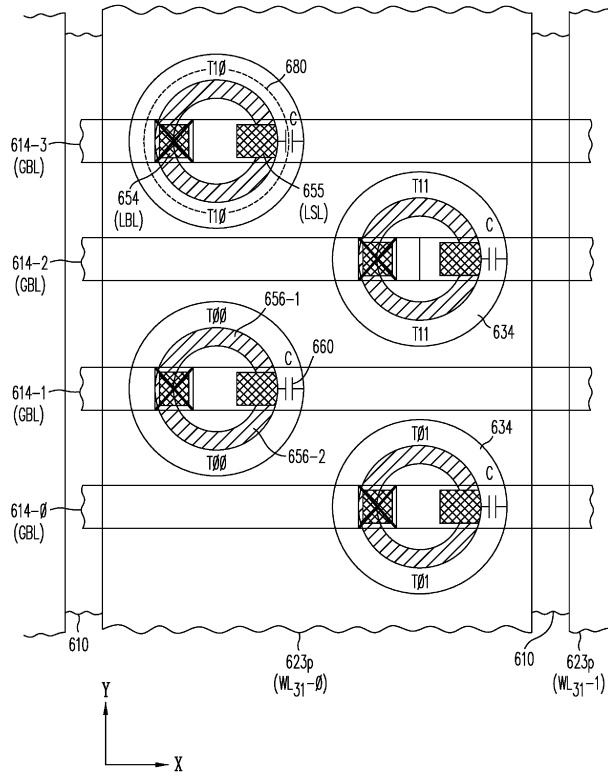


30

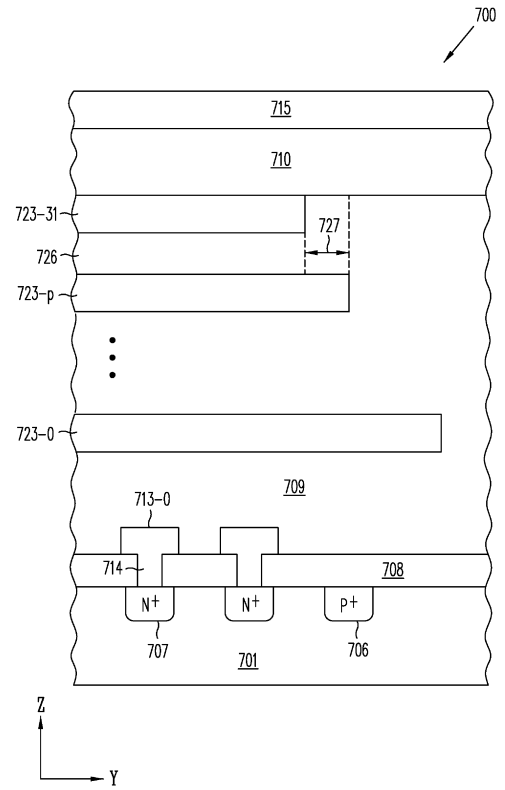
40

50

【 図 6 D 】



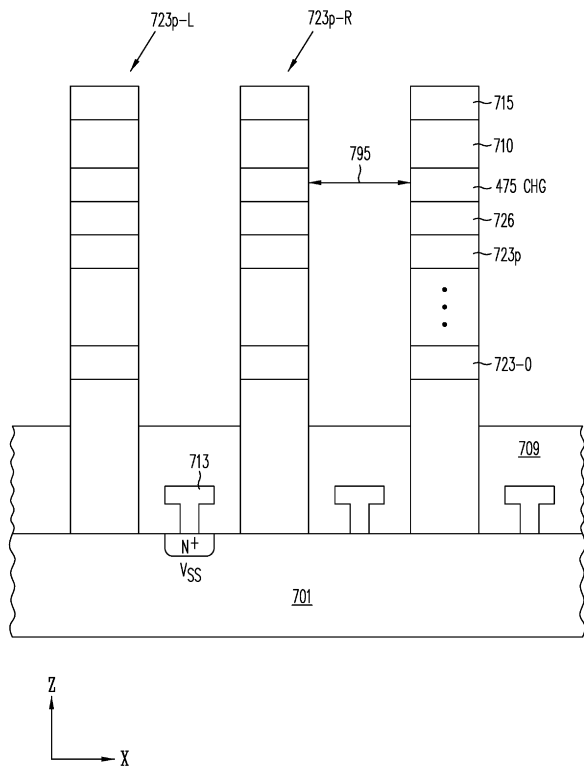
【 図 7 A 】



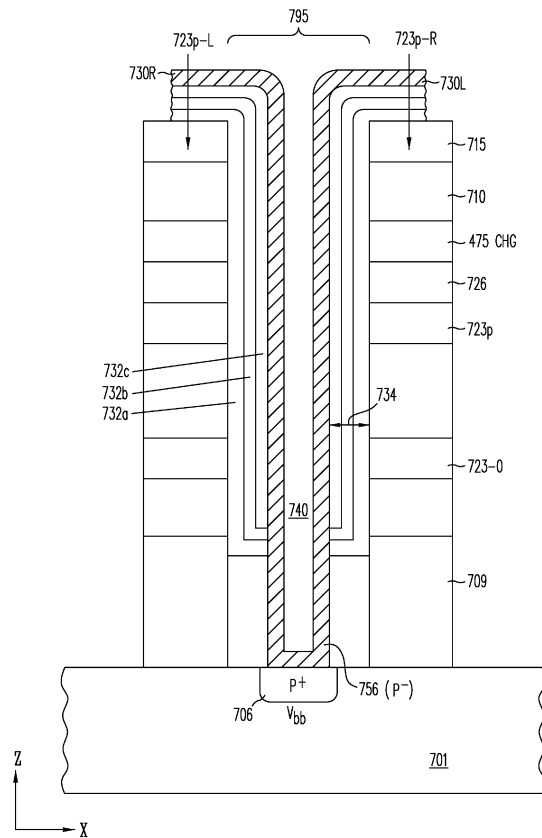
10

20

【 図 7 B 】



【 図 7 C 】

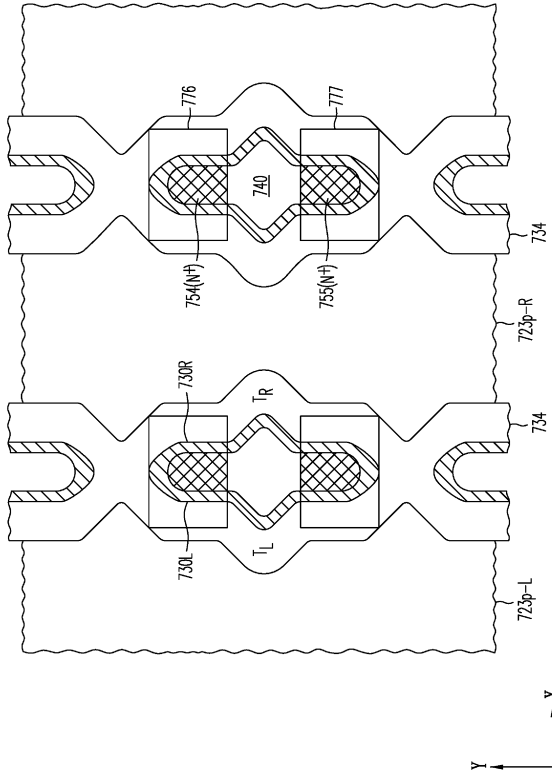


30

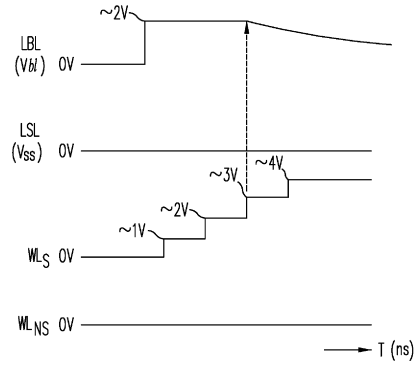
40

50

【 図 7 D 】



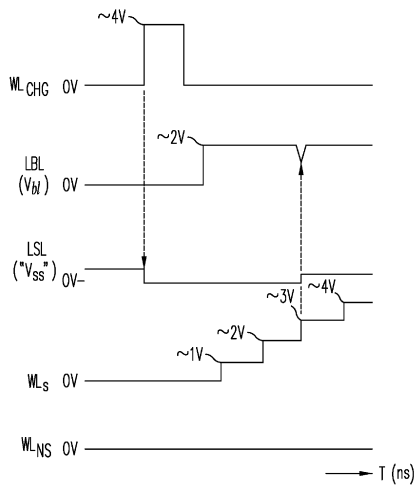
【 図 8 A 】



10

20

【 図 8 B 】



30

40

50

【手続補正書】【提出日】令和4年7月29日(2022.7.29)【手続補正1】【補正対象書類名】特許請求の範囲【補正対象項目名】全文【補正方法】変更【補正の内容】【特許請求の範囲】【請求項1】

その中またはその表面上に形成された1以上の電圧源を有する半導体基板における平坦な表面上に形成されたメモリ構造であって、

チャンネル領域、データストレージ領域、ゲート端子、第1のドレインまたはソース端子、及び第2のドレインまたはソース端子を有するストレージトランジスタであって、前記データストレージ領域の物理的状態を示す可変閾値電圧を有する、該ストレージトランジスタと、

前記ゲート端子に接続され、読み出し動作中に制御電圧を供給するワード線と、

前記第1のドレインまたはソース端子をデータ検出回路に接続するビット線と、

前記第2のドレインまたはソース端子に接続されたソース線と、

前記チャンネル領域を前記半導体基板の前記電圧源の1つに電氣的に接続して前記チャンネル領域のためにバイアス電圧を提供する導体層であって、前記ビット線および前記ソース線は、それぞれ、前記平坦な表面に実質的に垂直である第1の方向に沿って伸びる第1の導電性タイプの半導体材料の列である、メモリ構造。

【請求項2】

請求項1に記載のメモリ構造であって、

前記ソース線は、読み出し動作中に前記第2のドレインまたはソース端子と前記ゲート端子との間の少なくとも所定の電圧差を維持するのに十分な容量を提供し、

前記メモリ構造は、前記読み出し動作の前に前記容量を所定の電圧に充電するためのプリチャージトランジスタを更に備える、メモリ構造。

【請求項3】

請求項2に記載のメモリ構造であって、

前記読み出し動作中、前記制御電圧と前記所定の電圧差との和が前記可変閾値電圧を超えると、前記制御電圧によって前記ストレージトランジスタが前記容量を放電させる、メモリ構造。

【請求項4】

請求項2に記載のメモリ構造であって、

前記容量は、前記ソース線の寄生容量によって提供される、メモリ構造。

【請求項5】

請求項1に記載のメモリ構造であって、

前記ストレージトランジスタは、メモリアレイとして編成された複数のNORメモリストリングの1つにある複数のストレージトランジスタの1つであり、前記ワード線、前記導体層、前記ソース線および前記ビット線は、前記メモリアレイに関連付けられた、複数のワード線の1つ、複数のソース線の1つおよび複数のビット線の1つである、メモリ構造

【請求項6】

請求項5に記載のメモリ構造であって、

各NORメモリストリングに対して、(i)前記NORメモリストリングの前記ストレージトランジスタの前記第1のドレインまたはソース端末を提供する共通ドレイン領域と、(ii)前記NORメモリストリングの前記ストレージトランジスタの前記第2のドレインまたはソース端末を提供する共通ソース領域とを更に備える、メモリ構造。

【請求項7】

10

20

30

40

50

請求項 6 に記載のメモリ構造であって、

各々が前記平坦な表面に実質的に平行な第 2 の方向に沿って延在する、第 1 のセットの導体を更に備え、(i) 前記第 1 のセットの導体は、前記メモリアレイの前記ワード線を提供し、かつ (i i) 各 NOR メモリアレイの前記ストレージトランジスタの前記ゲート端子は、前記ワード線の対応する他の 1 つに接続された、メモリ構造。

【請求項 8】

請求項 7 に記載のメモリ構造であって、

前記第 1 のセットの導体は、前記ソースラインの各ソースラインに 1 つまたは複数のダミー導体を提供して前記ソース線の寄生容量を高める、メモリ構造。

【請求項 9】

請求項 7 に記載のメモリ構造であって、

各 NOR メモリストリングにおいて、前記ストレージトランジスタの前記データストレージ領域は、前記ストレージトランジスタの前記チャンネル領域上に提供される電荷トラップ材料の層の各一部である、メモリ構造。

【請求項 10】

請求項 9 に記載のメモリ構造であって、

前記ストレージトランジスタの前記チャンネル領域は、前記第 1 の導電性タイプとは反対である第 2 の導電性タイプの半導体材料の層の各一部である、メモリ構造。

【請求項 11】

請求項 7 に記載のメモリ構造であって、

前記 NOR メモリストリングは、前記第 2 の方向と第 3 の方向の両方に沿って配列され、前記第 3 の方向は、前記平坦な表面に実質的に平行であり、前記第 2 の方向に実質的に直交する、メモリ構造。

【請求項 12】

請求項 7 に記載のメモリ構造であって、

各 NOR メモリストリングの前記共通ソース領域は、前記 NOR メモリストリングの前記ストレージトランジスタの 1 つまたは複数の前記チャンネル領域が導通状態にされる場合を除いて、前記半導体基板に形成された前記回路に対して電氣的に浮動している、メモリ構造。

【請求項 13】

請求項 7 に記載のメモリ構造であって、

各ワード線は、前記メモリアレイの複数の前記 NOR メモリストリング間で共有された、メモリ構造。

【請求項 14】

請求項 11 に記載のメモリ構造であって、

前記第 3 の方向に沿って延在する第 2 のセットの導体を更に備え、前記第 2 のセットの導体の第 1 のグループは、前記メモリアレイの前記ビット線として機能し、その結果、各ビット線は、前記半導体基板内の前記回路の一部を、前記メモリアレイ内の 1 つの前記 NOR メモリストリングの前記ストレージトランジスタの前記共通ドレイン領域に接続する、メモリ構造。

【請求項 15】

請求項 14 に記載のメモリ構造であって、

前記第 2 のセットの導体は、前記メモリアレイと前記平坦な表面との間に形成された、メモリ構造。

【請求項 16】

請求項 14 に記載のメモリ構造であって、

前記第 2 のセットの導体は、前記メモリアレイ上に形成された、メモリ構造。

【請求項 17】

請求項 14 に記載のメモリ構造であって、

前記第 2 のセットの導体の前記第 1 のグループは、選択トランジスタを介して前記共通ド

10

20

30

40

50

レイン領域に接続された、メモリ構造。

【請求項 18】

請求項 14 に記載のメモリ構造であって、

前記第 2 のセットの導体の第 2 のグループは、導体で満たされたバイアホールを介して前記第 1 のセットの導体における対応するものに接続された、メモリ構造。

【請求項 19】

請求項 7 に記載のメモリ構造であって、

各ストレージトランジスタのデータ保持時間は 1 年未満で、プログラム/消去サイクルの耐久性は 10,000 プログラム/消去サイクルを超える、メモリ構造。

【請求項 20】

請求項 7 に記載のメモリ構造であって、

各ストレージトランジスタは、固有のエンハンスメントモードの閾値電圧を有する、メモリ構造。

【請求項 21】

請求項 7 に記載のメモリ構造であって、

前記半導体基板の前記回路は、メモリ動作のために所定の電圧を提供するための電圧源を含む、メモリ構造。

【請求項 22】

請求項 21 に記載のメモリ構造であって、

前記所定の電圧は、プログラム、プログラム禁止、読み取りおよび消去電圧のための電圧を含む、メモリ構造。

【請求項 23】

請求項 7 に記載のメモリ構造であって、

各ストレージトランジスタの前記可変閾値電圧は、ファウラー・ノルドハイム・トンネリングまたは直接トンネリングを使用して設定された、メモリ構造。

【請求項 24】

請求項 23 に記載のメモリ構造であって、

前記可変閾値電圧は、2 つ以上の充電状態の 1 つに対応するレベルに設定された、メモリ構造。

【請求項 25】

請求項 7 に記載のメモリ構造であって、

各 NOR メモリストリングにおいて、前記共通ソース領域および前記共通ドレイン領域は、それぞれ、各ストレージトランジスタの前記チャンネル領域に隣接するドーパント拡散ブロッキング層を含む、メモリ構造。

【請求項 26】

請求項 25 に記載のメモリ構造であって、

前記ドーパント拡散ブロッキング層は、3 ナノメートル未満の厚さの誘電体材料を含む、メモリ構造。

【請求項 27】

請求項 7 に記載のメモリ構造であって、

各ストレージトランジスタの前記チャンネル領域は、前記半導体基板に電氣的に接続された、メモリ構造。

【請求項 28】

請求項 27 に記載のメモリ構造であって、

各ストレージトランジスタの前記チャンネル領域は、前記第 1 の導電性タイプの半導体材料のピラーによって前記半導体基板に接続されている、メモリ構造。

【請求項 29】

請求項 27 に記載のメモリ構造であって、

前記半導体基板は、各ストレージトランジスタの前記チャンネル領域に、読み出し動作中の副閾値リークを抑制する所定のバックバイアス電圧を提供する、メモリ構造。

	10
	20
	30
	40
	50

【請求項 30】

請求項 7 に記載のメモリ構造であって、

各ストレージトランジスタの前記チャンネル領域は、横方向のホッピング伝導と蓄積された電荷のトンネリングによって消去を実行するのに十分に短い長さを有する、メモリ構造。

【請求項 31】

請求項 7 に記載のメモリ構造であって、

前記第 1 のセットの導体は、それぞれ、 N^+ ドープポリシリコン、 P^+ ドープポリシリコン、および、二酸化ケイ素、ケイ化物、またはポリサイドに関して高い仕事関数の高融点金属のいずれか 1 つを含む、メモリ構造。

【請求項 32】

請求項 7 に記載のメモリ構造であって、

各ストレージトランジスタの電荷蓄積領域は、窒化ケイ素の 1 つまたは複数の層、またはバンドギャップで設計された酸化物 - 窒化物 - 酸化物誘電体層を含む、メモリ構造。

【請求項 33】

請求項 7 に記載のメモリ構造であって、

各 NOR メモリストリングの各ストレージトランジスタは、プログラミング、プログラミング禁止、消去、または読み出し動作のために個別にアドレス指定できる、メモリ構造。

【請求項 34】

請求項 7 に記載のメモリ構造であって、

前記ストレージトランジスタは、不揮発性または準揮発性である、メモリ構造。

【請求項 35】

請求項 7 に記載のメモリ構造であって、

読み出し動作中、各 NOR メモリストリングの前記共通ソース領域の固有コンデンサは、仮想接地電圧に充電され、前記 NOR メモリストリングの前記共通ドレイン領域の固有コンデンサは、読み出し感知電圧に充電される、メモリ構造。

【請求項 36】

請求項 35 に記載のメモリ構造であって、

選択されたストレージトランジスタのゲート電極は、前記選択されたストレージトランジスタの閾値電圧の感知を可能にするために所定の電圧に上げられ、前記 NOR メモリストリングの他のすべてのストレージトランジスタの前記ゲート電極は、非導電状態に保持される、メモリ構造。

【請求項 37】

請求項 35 に記載のメモリ構造であって、

プログラミング動作中、前記共通ソース領域および前記共通ドレイン領域の前記固有コンデンサは、それぞれ、瞬間的に仮想接地電圧にプリチャージされる、メモリ構造。

【請求項 38】

請求項 7 に記載のメモリ構造であって、

選択された NOR メモリストリングの前記ストレージトランジスタの選択された 1 つにおいて、1 つまたは複数のプログラミング電圧パルスが、前記第 1 のセットの導体の選択された 1 つ以上に印可されて、前記チャンネル領域、前記共通ソース領域および前記共通ドレイン領域から前記データストレージ領域内の電荷トラップ材料への電荷の効率的なファウラー・ノルドハイム・トンネリングまたは直接トンネリングが開始され、前記第 1 のセットの導体の他のすべてのものは、前記チャンネル領域、前記共通ソース領域および前記共通ドレイン領域から選択されなかった前記データストレージ領域内の前記電荷トラップ材料への電荷の効率的なファウラー・ノルドハイム・トンネリングまたは直接トンネリングの開始を阻害する電圧に保持される、メモリ構造。

【請求項 39】

請求項 7 に記載のメモリ構造であって、

複数の NOR メモリストリングの前記ストレージトランジスタは、1 回の動作で消去される、メモリ構造。

10

20

30

40

50

【請求項 4 0】

請求項 7 に記載のメモリ構造であって、

1 つまたは複数の前記 NOR メモリストリングの 1 つまたは複数が、基準メモリストリングとして機能する、メモリ構造。

【請求項 4 1】

請求項 4 0 に記載のメモリ構造であって、

前記半導体基板内の前記回路は、前記 NOR メモリストリングの 1 つにあるストレージトランジスタから受信した信号と、前記基準メモリストリングの対応するストレージトランジスタから受信した信号とを比較する差動センスアンプを含む、メモリ構造。

【請求項 4 2】

請求項 4 0 に記載のメモリ構造であって、

前記基準メモリストリングの 1 つまたは複数のストレージトランジスタは、基準閾値電圧を設定するようにプログラムされている、メモリ構造。

【請求項 4 3】

請求項 7 に記載のメモリ構造であって、

1 つまたは複数の前記 NOR メモリストリングは、スペアストリングとして機能し、各スペアストリングは、前記メモリアレイ内の NOR メモリストリングを置き換えるように構成できる、メモリ構造。

【請求項 4 4】

請求項 4 0 に記載のメモリ構造であって、

前記基準メモリストリングのストレージトランジスタにプログラムされた閾値電圧は、マルチビットスキームの下でプログラムされた状態に対応する、メモリ構造。

10

20

30

40

50

フロントページの続き

(51)国際特許分類 F I
G 1 1 C 16/26 (2006.01) G 1 1 C 16/26 1 3 0

(33)優先権主張国・地域又は機関
米国(US)

(31)優先権主張番号 15/248,420

(32)優先日 平成28年8月26日(2016.8.26)

(33)優先権主張国・地域又は機関
米国(US)