

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

H01L 29/792 (2006.01)

H01L 21/336 (2006.01)



[12] 发明专利说明书

专利号 ZL 03824003.3

[45] 授权公告日 2008 年 9 月 17 日

[11] 授权公告号 CN 100420036C

[22] 申请日 2003.9.23 [21] 申请号 03824003.3

[30] 优先权

[32] 2002.10.9 [33] US [31] 10/267,153

[86] 国际申请 PCT/US2003/030588 2003.9.23

[87] 国际公布 WO2004/034426 英 2004.4.22

[85] 进入国家阶段日期 2005.4.8

[73] 专利权人 飞思卡尔半导体公司

地址 美国得克萨斯

[72] 发明人 高利尚卡尔·钦达劳雷

保罗·A·英格尔索尔

克莱格·T·斯维夫特

亚历山大·B·豪伊夫勒

[56] 参考文献

US2002/0024078A1 2002.2.28

JP58-115863A 1983.7.9

CN1286495A 2001.3.7

EP0513923A2 1992.11.19

JP2001-189391A 2001.7.10

审查员 章放

[74] 专利代理机构 中国国际贸易促进委员会专利
商标事务所

代理人 康健峰

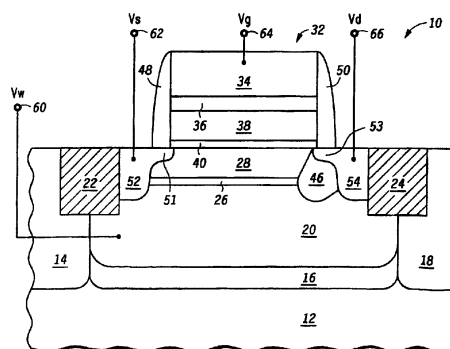
权利要求书 3 页 说明书 15 页 附图 4 页

[54] 发明名称

非易失性存储器件及其制造方法

[57] 摘要

在根据本发明的一种实施方案中，一种半导体器件(10)具有均匀布植于半导体衬底(20)中的具有第一导电型的高掺杂层(26)，其中沟道区(28)位于衬底(20)的上表面和高掺杂层(26)之间。在可选实施方案中，一种半导体器件(70)具有反掺杂沟道(%)和在沟道下方的防击穿区(74)。栅叠层(32)在衬底(20)上形成。具有第二导电型的源(52)和漏(54, 53)布植于衬底中。最后得到的非易失性存储单元提供低的自然阈电压，以使读周期过程中的阈电压漂移达到最小。另外，具有第二导电型并以某个角度在漏侧布植的晕轮区(46)可以用来协助允许较高编程速度的热载流子注入。



1. 一种半导体器件，包括：

半导体衬底；

在半导体衬底中的在半导体衬底的表面下方第一距离处形成的具有第一导电型的第一高掺杂层；

在半导体衬底上形成的第一绝缘层；

在第一绝缘层上形成的电荷存储层；

在电荷存储层上形成的第二绝缘层；

在半导体衬底的第一预先确定区中形成的具有第二导电型的源；

在半导体衬底的第二预先确定区中形成的具有第二导电型的漏，其中，所述第一高掺杂层不延伸到所述源和所述漏的深度之下；

在第一绝缘层下方在所述源和所述漏之间的沟道区；

仅在第一绝缘层的漏侧并贯穿所述漏而且在第一绝缘层的下方离第一绝缘层的边缘第二距离处形成的具有第一导电型的第二高掺杂层，其中所述第二高掺杂层增加第二距离内的掺杂物梯度；

其中所述半导体器件是非易失性存储单元并且在访问该非易失性存储单元的过程中，在沟道区内的沟道区边缘处形成耗尽区来掩蔽在第二距离内增加的掺杂物梯度。

2. 根据权利要求 1 的半导体器件，其中直接在第一绝缘层下方并在第一高掺杂层上方的衬底具有比第一高掺杂层的掺杂物浓度低的掺杂物浓度。

3. 一种用于形成半导体器件的方法，包括步骤：

提供半导体衬底；

在半导体衬底中的在半导体衬底的表面下方第一距离处形成具有第一导电型的第一高掺杂层；

在半导体衬底上形成第一绝缘层；

在第一绝缘层上形成电荷存储层；

在电荷存储层上形成第二绝缘层；

在半导体衬底的第一预先确定区中形成具有第二导电型的源;

在半导体衬底的第二预先确定区中形成具有第二导电型的漏, 其中, 所述第一高掺杂层不延伸到所述源和所述漏的深度之下; 以及

仅在第一绝缘层的漏侧并贯穿所述漏而且在第一绝缘层的下方离第一绝缘层的边缘第二距离处形成具有第一导电型的第二高掺杂层, 其中所述第二高掺杂层增加第二距离内的掺杂物梯度;

其中在第一绝缘层下方在所述源和所述漏之间形成沟道区, 并且其中所述半导体器件是非易失性存储单元并且在访问该非易失性存储单元的过程中, 在沟道区内的沟道区边缘处形成耗尽区来掩蔽在第二距离内增加的掺杂物梯度。

4. 一种半导体器件, 包括:

半导体衬底;

在半导体衬底中的在半导体衬底的表面下方第一距离处形成的第一P型高掺杂层, 其中所述第一P型高掺杂层使用铟作为掺杂物来形成;

在半导体衬底上形成的第一绝缘层;

在第一绝缘层上形成的电荷存储层;

在电荷存储层上形成的第二绝缘层;

在半导体衬底的第一预先确定区中形成的N型源;

在半导体衬底的第二预先确定区中形成的N型漏, 其中, 所述第一P型高掺杂层不延伸到所述N型源和所述N型漏的深度之下;

在第一绝缘层下方在所述N型源和所述N型漏之间的沟道区;

仅在第一绝缘层的漏侧并贯穿所述N型漏而且在第一绝缘层的下方离第一绝缘层的边缘第二距离处形成的第二P型高掺杂层, 其中所述第二P型高掺杂层使用铟作为掺杂物来形成。

5. 一种半导体器件, 包括:

半导体衬底;

在半导体衬底中的在半导体衬底的表面下方第一距离处形成的具有第一导电型的高掺杂层;

在半导体衬底上形成的氧化物-氮化物-氧化物结构;

在氧化物-氮化物-氧化物结构上形成的栅电极;

在半导体衬底的第一预先确定区中形成的具有第二导电型的源;

在半导体衬底的第二预先确定区中形成的具有第二导电型的漏, 其中, 所述高掺杂层不延伸到所述源和所述漏的深度之下;

在氧化物-氮化物-氧化物结构下方在所述源和所述漏之间的沟道区; 以及

仅在氧化物-氮化物-氧化物结构的漏侧并贯穿所述漏而且在氧化物-氮化物-氧化物结构下方离氧化物-氮化物-氧化物结构的边缘第二距离而形成的具有第一导电型的倾角晕轮区; 其中所述倾角晕轮区增加第二距离内的掺杂物梯度;

其中所述半导体器件是非易失性存储单元并且在访问该非易失性存储单元的过程中, 在沟道区内的沟道区边缘处形成耗尽区来掩蔽在第二距离内增加的掺杂物梯度。

非易失性存储器件及其制造方法

技术领域

本发明一般地涉及半导体器件，更特别地涉及供存储单元中使用的半导体器件。

背景技术

在基于 SONOS（硅-氧化物-氮化物-氧化物-硅）的非易失性存储（NVM）单元中，向氮化物的热载流子注入（HCI）可以用来编程具有高阈值电压（ V_t ）状态和低 V_t 状态的存储单元。高效的 HCI 编程要求高的沟道区掺杂和陡峭的漏结；但是，由于具有高的沟道区掺杂，读干扰被加重。也就是，低 V_t 状态下存储单元的重复读取不断地增加存储单元的 V_t 。 V_t 可能增加到存储单元的状态可以从低 V_t 状态改变到高 V_t 状态的一点，从而导致存储单元的可靠性失败。因此，需要一种在重复读取过程中具有增强可靠性的存储单元。

附图说明

本发明作为例子来说明并且不被附图所限制，附图中相似参考数字指示相似元素，并且其中：

图 1 说明根据本发明一种实施方案的其中形成有阱布植和沟道布植的半导体衬底的横截面视图；

图 2 说明根据本发明一种实施方案的在半导体衬底上形成有栅叠层的图 1 半导体衬底的横截面视图；

图 3 说明根据本发明一种实施方案的在形成晕轮布植之后的图 2 栅叠层的横截面视图；

图 4 说明根据本发明一种实施方案的在半导体衬底内形成源和漏区以及扩展区并且沿着栅叠层的侧壁形成侧壁隔离物之后的图 3 半导体器

件;

图 5 说明根据本发明一种可选实施方案的其中形成有阱布植的半导体衬底的横截面视图;

图 6 说明根据本发明一种实施方案的在半导体衬底和沟道布植上形成有第一氧化物层、氮化物层, 以及第二氧化物层的图 5 半导体衬底的横截面视图;

图 7 说明根据本发明一种实施方案的在形成栅叠层之后的图 6 半导体衬底的横截面视图; 以及

图 8 说明根据本发明一种实施方案的在半导体衬底内形成源和漏区以及扩展区并且沿着栅叠层的侧壁形成侧壁隔离物之后的图 7 半导体器件。

本领域技术人员应当认识到, 图中的元素为了简单和清楚而说明并不一定按比例画出。例如, 图中一些元素的尺寸可能相对于其他元素被放大以帮助增进本发明实施方案的理解。

具体实施方式

在本发明的一种实施方案中, 可以用作 NVM 存储单元的半导体器件形成有防击穿 (APT) 区和任选漏侧的高掺杂区 (晕轮)。如果存在, 晕轮区导致半导体器件的沟道区和漏区之间增加的掺杂物梯度。APT 区使沟道区可以具有相对低的掺杂物浓度或者关于 APT 区相反地掺杂, 这通过降低自然 V_t 使读干扰 (即读周期过程中的阈电压漂移) 达到最小。因此, 使用晕轮区和 APT 区使半导体器件的高效热载流子注入编程可以保持同时减小读干扰。

图 1 说明包括具有隔离沟 22 和 24、包围 N 型阱 14 和 18、在隔离沟 22 和 24 之间的隔离 N 型阱 16, 以及掩膜层 30 的半导体衬底 12 的半导体器件 10。应当注意, 隔离沟 22 和 24、包围 N 型阱 14 和 18、隔离 N 型阱 16, 以及掩膜层 30 的形成在本领域是众所周知的, 因此在这里仅简单地描述。隔离沟 22 和 24 在衬底 12 中形成, 然后包围 N 型阱 14 和 18 被形成。隔离沟 22 和 24 可以包括任何类型的绝缘材料, 例如

氧化物、氮化物等，或它们的任意组合。在形成包围 N 型阱 14 和 18 之后，形成图案的掩膜层 30 用来定义隔离沟 22 和 24 之间的开口。应当注意，形成图案的掩膜层 30 可以是任何类型的掩膜层，例如光刻胶层、硬掩膜等。然后，隔离 N 型阱 16 在衬底 12 中形成。在形成隔离 N 型阱 16 之后，被隔离 P 型阱 20 在隔离 N 型阱 16 中形成，使得 P 型阱 20 与衬底 12 隔离。

在形成被隔离 P 型阱 20 之后，防击穿 (APT) 区 26 和沟道区 28 在隔离沟 22 和 24 之间形成。(应当注意，APT 区 26 和沟道区 28 可以按任意顺序形成。) 沟道区 28 和 APT 区 26 被形成，使得沟道区 28 位于衬底 12 的上表面和 APT 区 26 之间，并且 APT 区 26 位于沟道区 28 和被隔离 P 型阱 20 之间。(应当注意，APT 区 26 也可以称作高掺杂区 26。) 用于形成 APT 区 26 的掺杂物被选择，使得它不会显著地扩散到沟道区 28 中。箭头 31 说明掺杂物均匀地施加到衬底 12 上。APT 区 26 和沟道区 28 两者的布植方向基本上垂直于衬底 12。也就是，方向离垂直不大于大约 10 度。还应当注意，APT 区 26 的掺杂物浓度大于被隔离 P 型阱 20 的掺杂物浓度。

在一种实施方案中，APT 区 26 和沟道区 28 被形成，使得沟道区 28 的掺杂物浓度小于 APT 区 26 的掺杂物浓度。在一种实施方案中，APT 区 26 和沟道区 28 使用 P 型掺杂物，例如硼或铟来形成。在该实施方案中，沟道区 28 的掺杂物浓度可以比 APT 区 26 的掺杂物浓度低 10~50 倍。因此，APT 区 26 可以用大约 30~50 千电子伏特 (keV) 的能量和大约 $1 \times 10^{12}/\text{cm}^2 \sim 1 \times 10^{14}/\text{cm}^2$ 的剂量来布植，并且沟道区 28 可以用大约 5~30keV 的能量和大约 $1 \times 10^{11}/\text{cm}^2 \sim 1 \times 10^{13}/\text{cm}^2$ 的剂量来布植。应当注意，在一种实施方案中，不同的 P 型掺杂物可以用于沟道区 28 和 APT 区 26，例如硼用于沟道区 28 并且铟用于 APT 区 26。可选地，相同的 P 型掺杂物可以用于两个区。

在所说明的实施方案中，半导体衬底 12 是容积衬底。在该实施方案中，衬底 12 是包含半导体的衬底，并且可以包括硅、砷化镓、硅锗等，或者它们的任意组合。可选地，衬底 12 可以是绝缘体上外延硅

(SOI)衬底(没有显示),具有底半导体层、在底半导体层上方的埋入绝缘层,以及顶半导体层。在该实施方案中应当注意,包围N型阱14和18以及隔离N型阱16是不必需的。也就是,被隔离P型阱20相当于SOI衬底的顶半导体层。在该实施方案中,埋入绝缘层可以是氧化硅层,并且顶和底半导体层可以由硅、锗、砷化镓等制成。

图2说明在去除掩膜层30以及在隔离阱22和24之间在沟道区28上形成SONOS栅叠层32之后的半导体器件10,其中SONOS栅叠层32包括在沟道区28上形成的第一氧化物40、在第一氧化物40上形成的氮化物38、在氮化物38上形成的第二氧化物36,以及在第二氧化物36上形成的栅34。(应当注意,第一氧化物40、氮化物38,以及第二氧化物36可以称作氧化物-氮化物-氧化物结构。)掩膜层30可以使用常规工艺来去除。在形成栅叠层32时,第一氧化物层分别使用化学汽相沉积(CVD)或热氧化工艺在半导体衬底12上沉积或生长。可选地,第一氧化物层可以通过物理汽相沉积(PVD)、原子层沉积(ALD)、热氧化等,或上述方法的组合来形成。然后,氮化物层沉积在第一氧化物层上。氮化物层可以通过CVD、PVD、ALD等,或它们的组合来形成。第二氧化物层分别使用化学汽相沉积(CVD)或热氧化工艺在氮化物层上沉积。可选地,第二氧化物层可以通过物理汽相沉积(PVD)、原子层沉积(ALD)、热氧化等,或上述方法的组合来形成。栅层通过CVD、PVD、ALD等,或它们的组合沉积在所形成的第二氧化物层上。然后使用常规的掩膜和刻蚀工艺,第一氧化物层、氮化物层、第二氧化物层,以及栅层可以被形成图案并刻蚀,以形成最后得到的栅叠层32。(应当注意,在可选实施方案中,叠层的每个层可以单独形成图案并刻蚀,以形成最后得到的栅叠层32。)在一种实施方案中,最后得到的栅叠层32(同样地,在栅叠层32下方的沟道区28的一部分)具有大约0.35微米~0.06微米的长度。

栅叠层32的栅34可以是任何导电材料,例如多晶硅或包含金属的材料,并且可以称作控制栅。第一氧化物40和第二氧化物36可以是任何电介质,例如绝缘材料或绝缘材料的叠层,例如氧化硅、氮氧化合

物、金属氧化物、氮化物等，或它们的任意组合。氮化物 38 可以是氮化硅、氮氧化合物、或者已知具有电荷陷阱使得电荷可以存储于其中的任何其他材料。因此，第一氧化物 40 和第二氧化物 36 也可以分别称作第一和第二绝缘层或分别地底和顶电介质，并且氮化物 38 可以称作电荷存储层、存储元件或电介质。

虽然栅叠层 32 说明为 SONOS 叠层，在可选实施方案中，栅叠层 32 可以是任何类型的 NVM 栅叠层。例如，栅叠层 32 可以用浮动栅叠层（没有显示）来代替，其具有在隔离沟 22 和 24 之间在沟道区 28 上形成的隧道电介质、在隧道电介质上形成的浮动栅、在浮动栅上形成的控制电介质，以及在控制电介质上形成的控制栅。在形成浮动栅叠层时，隧道电介质层通过 CVD、PVD、ALD、热氧化等，或它们的组合在半导体衬底 12 上面形成。隧道电介质层可以是任何绝缘材料，例如氧化物（例如二氧化硅）、氮化物、氮氧化合物、金属氧化物等。然后隧道电介质层使用常规工艺来形成图案并刻蚀，以在沟道区 28 上面形成浮动栅叠层的隧道电介质（其中隧道电介质位于与图 2 中说明的栅叠层 32 的氧化物 40 相似的位置）。

然后，浮动栅层通过 CVD、PVD、ALD 等，或它们的组合在半导体衬底 12 和隧道电介质上形成。在一种实施方案中，浮动栅层可以是任何导电材料，例如多晶硅、金属等。在另一种实施方案中，浮动栅层可以是许多纳米晶体（即分立存储单元），例如在纳米晶体 NVM 器件中。然后，浮动栅层使用常规工艺来形成图案并刻蚀，以在隧道电介质上面形成浮动栅叠层的浮动栅。

然后，控制电介质层通过 CVD、PVD、ALD、热氧化等，或它们的组合在半导体衬底 12 和浮动栅上形成。然后，控制电介质层使用常规工艺来形成图案并刻蚀，以在浮动栅上面形成浮动栅叠层的控制电介质。应当注意，控制电介质是任选的，并不是在所有浮动栅器件中都形成。如果存在，控制电介质层可以是任何绝缘材料，例如氧化物（例如二氧化硅）、氮化物、金属氧化物、高介电常数材料（即具有大于大约 4 并小于大约 15 的介电常数的材料）等，或它们的组合。然后，控制栅

层通过 CVD、PVD、ALD 等，或它们的组合在半导体衬底 12 和控制电介质上形成。控制栅层可以是任何导电材料，例如多晶硅或包含金属的材料。使用常规的掩膜和刻蚀工艺，控制栅层被形成图案并刻蚀，以在控制电介质上面形成浮动栅叠层的控制栅。（应当注意，在可选实施方案中，不是分别形成图案并刻蚀浮动叠层的每个层，层的组合或所有层可以使用相同图案和刻蚀工艺来形成图案并刻蚀，以便减小形成最后得到的浮动栅叠层所需的工艺步骤。）

现在参考图 3，形成图案的掩膜层 42 使用常规的掩膜工艺来形成。应当注意，掩膜层 42 可以是任何类型的掩膜层，例如光刻胶或硬掩膜。形成图案的掩膜层 42（也称作布植掩膜）掩蔽半导体器件 10 的源侧（位于栅叠层 32 的第一侧，源区随后将在那里形成），同时暴露半导体器件 10 的漏侧（位于栅叠层 32 的与第一侧相对的第二侧，漏区随后将在那里形成）。如图 3 中所说明的，倾角布植 44 用来形成晕轮区 46，其在栅叠层 32 下方延伸从栅叠层 32 的第一边开始测量的距离 47。在一种实施方案中，距离 47 至多大约 500 埃。倾角布植 44 具有相应的布植角 θ ，其中 θ 是从垂直向开始计算的。在一种实施方案中， θ 大约为 20~60 度，更优选地大约 30~40 度。因此，布植 44 的角度足以增加在栅叠层 32 下方区域 45 处晕轮区 46 中的掺杂物浓度，使得它大于沟道区 28 的掺杂物浓度。在一种实施方案中，晕轮区 46 使用 P 型掺杂物，例如硼或铟，以大约 10~50keV 的能量并具有大约 $1 \times 10^{12}/\text{cm}^2 \sim 1 \times 10^{14}/\text{cm}^2$ 的剂量来布植。（应当注意，可选地晕轮区 46 可以称作倾角晕轮 46 或者高或重掺杂物区 46。并且，晕轮区 46 的掺杂物浓度一般大于被隔离 P 型阱 20 的掺杂物浓度。）

图 4 说明在去除掩膜层 42 以及形成侧壁隔离物 48 和 50、源和漏延伸区 51 和 53，以及源和漏区 52 和 54 之后的半导体器件 10。掩膜层 42 可以使用常规工艺步骤来去除。在去除掩膜层 42 之后，源延伸区 51 和漏延伸区 53 使用常规的掩膜和布植工艺来形成。应当注意，延伸区 51 和 53 延伸到沟道区 28 中，并且每个都在栅叠层 32 的一部分的下方。在一种实施方案中，N 型掺杂物，例如砷、磷或锑以大约 30~70keV 的

能量并具有大约 $1 \times 10^{14}/\text{cm}^2 \sim 1 \times 10^{15}/\text{cm}^2$ 的剂量来布植，以形成延伸区 51 和 53。漏延伸区 53 被形成，使得它不会延伸超过晕轮区 46。应当注意，在形成漏延伸区 53 之后，导致从沟道区 28 到漏延伸区 53 的递增的掺杂物梯度。虽然没有晕轮区 46 时从沟道区 28 到漏延伸区 53 存在递增的掺杂物梯度，晕轮区 46 的存在进一步增加该掺杂物梯度。并且，晕轮区 46 的存在允许沟道区 28 中相对低的掺杂物浓度。

在形成延伸区 51 和 53 之后，隔离物 48 和 50 使用常规工艺步骤沿着栅叠层 32 的侧壁形成。这些隔离物，例如可以包括任何绝缘材料，例如氧化物或氮化物。可选地，隔离物 48 和 50 可以不存在。如果隔离物 48 和 50 不存在，那么源和漏区 52 和 54 可以不形成，使得延伸区 51 和 53 分别用作源和漏区。但是，随着隔离物 48 和 50 的存在，源和漏区可以使用另外的布植步骤来形成。在一种实施方案中，N 型掺杂物，例如砷、磷或锑以大约 $10 \sim 30\text{keV}$ 的能量并具有大约 $1 \times 10^{15}/\text{cm}^2 \sim 5 \times 10^{16}/\text{cm}^2$ 的剂量来布植，以形成源区 52 和漏区 54。应当注意，漏和源区 52 和 54 不在隔离沟 22 和 24 下方延伸。还应当注意，APT 26 的深度被选择，使得它不在源和漏区 52 和 54 的深度下方延伸。虽然没有显示，更多的常规工艺可以用来完成半导体器件 10。例如，触点可以在源区 52、栅 34、漏区 54，以及被隔离 P 型阱 20 上形成。并且，其他半导体器件级可以在半导体器件 10 的下方或上方形成。

如图 4 中所示， V_w 60 对应于施加到被隔离 P 型阱 20 的电压， V_s 62 对应于施加到源区 52 的电压， V_g 64 对应于施加到栅 34 的电压，并且 V_d 66 对应于施加到漏区 54 的电压。在所说明的实施方案中，半导体器件 10 可以用作 NVM 存储器（没有显示）内的 NVM 存储单元。如这里所使用的，高 V_t 状态对应于存储单元的编程状态，而低 V_t 状态对应于存储单元的擦除状态。（但是应当注意，在可选实施方案中编程和擦除状态可以反过来。）

半导体器件 10 通过从氮化物 38 中移走电子，导致半导体器件 10 具有低 V_t （例如低于大约 2 伏特）来擦除。许多已知的方法可以用来使半导体器件 10 置于低 V_t 状态，例如 Fowler-Nordheim 隧道、热空穴

注入、直接隧穿等。

半导体器件 10 通过将电子存储于氮化物 38 中，导致半导体器件 10 具有高 V_t （例如高于大约 4 伏特）来编程。因此，半导体器件 10 可以通过施加漏电压（ V_d ）和源电压（ V_s ），其中 V_d 比 V_s 高大约 3~5 伏特来编程。例如，在一种实施方案中，可以使用 1 伏特的 V_s 和 4 伏特的 V_d 。在该实施方案中，施加大约 5~10 伏特的栅电压（ V_g ）和大约 0~-3 伏特的阱电压（ V_w ）。在施加有上述电压的半导体器件 10 的编程过程中，热载流子在漏耗尽区中产生，其中一些通过氧化物 40 注入到氮化物 38 中。这导致增加半导体器件 10 的 V_t 。应当注意，由晕轮区 46 和漏延伸区 53 产生的掺杂物梯度增强该热载流子注入，从而保持半导体器件 10 的高效热载流子编程。该效率被保持，即使沟道区 28 具有相对低的掺杂物浓度（大约 $1 \times 10^{16}/\text{cm}^3 \sim 1 \times 10^{17}/\text{cm}^3$ ）。此外，沟道区 28 的相对低的掺杂物浓度减小半导体器件 10 的自然 V_t 从而改善读干扰，这将在下面描述。

半导体器件 10 的自然 V_t 指的是在将任何电荷置入氮化物 38 之前的阈值电压。对于较高的自然 V_t ，读干扰被退化。（应当注意，如这里所使用的，读干扰描述当低 V_t 存储单元被不断读取时阈值电压（ V_t ）的逐渐增加，即读周期过程中的阈值电压漂移。）因此，当自然 V_t 增加时，离存储单元失败的时间减少。也就是，当自然 V_t 增加时，对存储单元的较少次数的读取由于从低 V_t 到高 V_t 的漂移而导致引失败。因此，通过降低自然 V_t ，低 V_t 状态的读干扰被改善（即阈值电压漂移减小）。例如，参考回图 4，半导体器件 10 的读取可以通过施加比 V_s 高大约 0.5~1.5 伏特的 V_d 来执行。例如，在一种实施方案中， V_s 可以是 0 伏特并且 V_d 可以是 1 伏特。在该实施方案中，足以在沟道区 28 中产生大约 10~30 微安培的电流的 V_g 和 V_w 被施加。例如，在一种实施方案中，可以使用 2 伏特的 V_g 和 0 伏特的 V_w 。（应当注意，在该实例中提供的电压根据源电压（ V_s ）来给出。也就是，在该实例中，如果 V_s 增加 1 伏特，那么 V_d 、 V_g 和 V_w 也增加 1 伏特。）在读取或访问已擦除半导体器件 10 的过程中（即半导体器件 10 处于低 V_t 状态），反型层在

沟道区 28 中形成, 并且耗尽区 (没有显示) 在漏区 54 和漏延伸区 53 的周围形成。该耗尽区基本上掩蔽在晕轮区 46 中形成的掺杂物梯度, 从而防止晕轮区 46 的较高掺杂物增加半导体器件 10 的 V_t 。这样, V_t 保持在低 V_t 状态, 从而通过减小 V_t 漂移而改善读干扰。

因为如上所述栅叠层 32 的长度大约 0.35~0.06 微米, 短沟道泄漏可能在半导体器件 10 的编程过程中发生。但是, 高掺杂 APT 区 26 也用来减小该短沟道泄漏, 从而减小功率损耗和提高编程效率。

图 5~8 说明本发明的可选实施方案, 其中不是使用相同导电型的掺杂物来形成沟道区 28 和 APT 区 26, 而是使用不同导电型的掺杂物的两个布植步骤可以用来形成沟道区 86 和 APT 区 74。也就是, 在该可选实施方案中, 沟道区 28 和 APT 区 26 可以分别用沟道区 86 和 APT 区 74 来代替, 它们以与上述沟道区 28 和 APT 区 26 类似的方式工作, 以允许半导体器件的高效热载流子注入编程并同时减小读干扰。并且, 如将在下面描述的, 在该可选实施方案中, 晕轮区 46 可以不存在。(应当注意, 在下面图 5~8 的描述中, 与在图 1~4 的描述中使用的参考数字相同的参考数字指示相似或类似的元素。)

图 5 说明包括具有隔离沟 22 和 24、包围 N 型阱 14 和 18、在隔离沟 22 和 24 之间的隔离 N 型阱 16, 以及形成图案的掩膜层 30 的半导体衬底 12 的半导体器件 10。应当注意, 隔离沟 22 和 24、包围 N 型阱 14 和 18、隔离 N 型阱 16, 以及掩膜层 30 的形成与上面参考图 1 所描述的相同, 因此将不在这里参考图 5 再次描述。在形成隔离沟 22 和 24、包围 N 型阱 14 和 18、形成图案的掩膜层 30、隔离 N 型阱 16, 以及被隔离 P 型阱 20 之后 (其中在上面参考图 1 提供的相同的描述、材料和可选方案在这里参考图 5 而应用), APT 区 74 在被隔离 P 型阱 20 中在隔离沟 22 和 24 之间形成。(应当注意, APT 区 74 也可以称作高掺杂区 74)。

箭头 72 说明掺杂物均匀地施加到衬底 12 上。APT 区 74 的布植的方向基本上垂直于衬底 12。也就是, 方向离垂直不大于大约 10 度。在一种实施方案中, APT 区 74 使用 P 型掺杂物, 例如硼或铟来形成。例

如, APT 区 74 可以用大约 $30 \sim 50\text{keV}$ 和大约 $1 \times 10^{12}/\text{cm}^2 \sim 1 \times 10^{14}/\text{cm}^2$ 的剂量来布植。还应当注意, APT 区 74 和被隔离 P 型阱 20 的掺杂物具有相同的导电型, 并且 APT 区 74 的掺杂物浓度大于被隔离 P 型阱 20 的掺杂物浓度。例如, 在一种实施方案中, APT 区 74 的掺杂物浓度比被隔离 P 型阱 20 的掺杂物浓度高大约 $2 \sim 100$ 倍。例如, APT 区 74 掺杂物浓度可以是大约 $5 \times 10^{17}\text{cm}^{-3} \sim 5 \times 10^{18}\text{cm}^{-3}$, 并且被隔离 P 型阱 20 的掺杂物浓度可以是大约 $5 \times 10^{16}\text{cm}^{-3} \sim 5 \times 10^{17}\text{cm}^{-3}$ 。

图 6 说明在去除形成图案的掩膜层 30 和形成第一氧化物层 80、氮化物层 82, 以及第二氧化物层 84 之后的半导体器件 70。应当注意, 掩膜层可以如上面参考图 2 所描述而去除。在所说明的实施方案中, 第一氧化物层 80 分别使用化学汽相沉积 (CVD) 或热氧化工艺在半导体衬底 12 上沉积或生长。可选地, 第一氧化物层可以通过物理汽相沉积 (PVD)、原子层沉积 (ALD)、热氧化等, 或上述方法的组合来形成。然后, 氮化物层 82 沉积在第一氧化物层 80 上。氮化物层 82 可以通过 CVD、PVD、ALD 等, 或它们的组合来形成。然后, 第二氧化物层 84 分别使用化学汽相沉积 (CVD) 或热氧化工艺在氮化物层 82 上沉积。可选地, 第二氧化物层 84 可以通过物理汽相沉积 (PVD)、原子层沉积 (ALD)、热氧化等, 或上述方法的组合来形成。

在形成第二氧化物层 84 之后, 形成图案的掩膜层 76 用来定义在隔离沟 22 和 24 之间的开口。应当注意, 形成图案的掩膜层 76 可以是任何类型的掩膜层, 例如光刻胶层、硬掩膜等。在形成图案的掩膜层 76 形成之后, 沟道区 86 在被隔离 P 型阱 20 中形成。在一种实施方案中, 沟道区 86 使用 N 型掺杂物, 例如砷、磷或锑来形成。该 N 型掺杂物可以用大约 $5 \sim 70\text{keV}$ 的能量和大约 $1 \times 10^{11}/\text{cm}^2 \sim 5 \times 10^{13}/\text{cm}^2$ 的剂量来布植。在所说明的实施方案中, N 型掺杂物抵消 APT 区 74 的已有 P 型掺杂物的一部分, 以形成沟道区 86。结果, 沟道区 86 具有第一导电型 (例如在该实施方案中为 N 型) 并位于衬底 12 的上表面和 APT 区 74 之间, 并且 APT 区 74 具有第二导电型 (例如在该实施方案中为 P 型) 并位于沟道区 86 和被隔离阱 20 之间。应当注意, 为了使 N 型掺杂物适

当地抵消 APT 区 74 的一部分, 沟道区 86 中的 N 型掺杂物浓度应当高于 APT 区 74 中的 P 型掺杂物浓度。

在形成沟道区 86 之后, 沟道区 86 的净掺杂浓度在一种实施方案中大约为 $0 \sim 5 \times 10^{18} \text{cm}^{-3}$ 。这里所使用的净掺杂浓度指的是一种导电型的掺杂物与另一种导电型的掺杂物之间的绝对差。例如, 为沟道区 86 提供的净掺杂浓度指的是 APT 区 74 的 P 型掺杂物和沟道区 86 的 N 型掺杂物之差的绝对值。在本发明的一种实施方案中, 沟道区 86 中 P 型掺杂物的浓度减去沟道区 86 中 N 型掺杂物的浓度小于或等于被隔离 P 型阱 20 中的净掺杂浓度。应当注意, 沟道区 86 中 P 型掺杂物的浓度减去沟道区 86 中 N 型掺杂物的浓度可以提供负数, 并具有大于被隔离 P 型阱 20 中的净掺杂浓度的绝对值。在本发明的另一种实施方案中, 沟道区 86 中 P 型掺杂物的浓度减去沟道区 86 中 N 型掺杂物的浓度可以提供负数, 并具有小于被隔离 P 型阱 20 中的净掺杂浓度的绝对值。在可选实施方案中, 可以在 APT 区下方的区域中具有非均匀阱掺杂, 使得 APT 掺杂浓度小于阱浓度的最大值。

应当注意, 如图 6 中所示, 沟道区 86 在形成第一氧化物层 80、氮化物层 82, 以及第二氧化物层 84 之后形成。但是, 在可选实施方案中, 沟道区 86 可以在形成这些层之前形成。也就是, 在形成参考图 5 描述的 APT 区 74 之后, 随后的布植步骤可以用于使用相同的形成图案的掩膜层 30 来形成沟道区 86。因此, 在该实施方案中, 形成图案的掩膜层 76 将不需要。

图 7 说明在形成栅叠层 32 之后的半导体器件 70。在氮化物层 82 上形成第二氧化物层 84 之后, 形成图案的掩膜层 76 被移走 (例如, 使用常规工艺)。然后, 栅层通过 CVD、PVD、ALD 等, 或它们的组合沉积在所形成的第二氧化物层 84 上。然后使用常规的掩膜和刻蚀工艺, 第一氧化物 80、氮化物层 82、第二氧化物层 84, 以及栅层可以被形成图案并刻蚀, 以形成最后得到的栅叠层 32。也就是, 第一氧化物层 80 的刻蚀导致第一氧化物 40, 氮化物层 82 的刻蚀导致氮化物 38, 第二氧化物层 84 的刻蚀导致第二氧化物 36, 并且栅层的刻蚀导致栅 34。(应

当注意，在可选实施方案中，叠层的每个层可以单独形成图案并刻蚀，以形成最后得到的栅叠层 32。例如，氧化物层 80 和 84 以及氮化物层 82 可以在形成沟道区 86 之前被形成图案并刻蚀。) 在一种实施方案中，最后得到的栅叠层 32 (同样地，栅叠层 32 下方的沟道区 86 的一部分) 具有大约 0.35 微米~0.06 微米的长度。(应当注意，上面参考第一氧化物 40、氮化物 36、第二氧化物 36，以及栅 34 而提供的描述，包括材料和可选方案也适用与图 7 的栅叠层 32。)

虽然栅叠层 32 以图 7 中的 SONOS 叠层说明，在可选实施方案中，栅叠层 32 可以是任何类型的 NVM 栅叠层，如上面参考图 3 所描述的。因此，上面为栅叠层 32 提供的所有描述也适用于该实施方案。也就是，上面参考图 3 的栅叠层 32 描述的所有形成方法、材料，以及可选方案在这里再次应用于栅叠层 32。例如，栅叠层 32 可以用上面所描述的浮动栅叠层 (没有显示) 来代替。但是应当注意，如果栅叠层 32 用浮动栅叠层代替的话，浮动栅可能太厚而不允许用于形成沟道区 86 的布植的适当穿透。因此，在使用浮动栅叠层的实施方案中，沟道区 86 可以在形成 APT 区 74 之后和形成浮动栅叠层的任何部分之前形成。

在一种实施方案中，在形成栅叠层 32 之后，晕轮区例如晕轮区 46 可以被隔离 P 型阱 20 中形成，如上面参考图 3 所描述的。也就是，在形成栅叠层 32 之后，形成图案的掩膜层 42 可以用来形成晕轮区 46，如上面参考图 3 所描述的。在该实施方案中，晕轮区 46 (在图 7 和 8 中没有显示) 将与沟道区 86 和 APT 区 74 (而不是沟道区 28 和 APT 区 26) 相邻。但是，参考图 3 为晕轮区 46 和倾角布植 44 而描述的相同的形成方法、材料，以及可选方案可以应用于具有沟道区 86 和 APT 区 74 代替沟道区 28 和 APT 区 26 的当前实施方案。应当注意，在图 5~8 的当前实施方案中，由于用来形成沟道区 86 和 APT 区 74 的反掺杂方法，晕轮区 46 可以是不需要的。

图 8 说明在去除掩膜层 76、形成栅叠层 32、形成晕轮区 46，以及形成侧壁隔离物 48 和 50、源和漏延伸区 51 和 53，以及源和漏区 52 和 54 之后的半导体器件 70。应当注意，上面为晕轮区 46、侧壁隔离物 48

和 50、源和漏延伸区 51 和 53，以及源和漏区 52 和 54 提供的相同描述在这里参考图 8 而应用。也就是，参考图 4 而描述的相同的形成方法、材料，以及可选方案应用于图 8。还应当注意，在图 8 中晕轮区 46 被显示，因此图 8 的半导体器件 70 类似于图 4 的半导体器件 10，除了图 4 的沟道区 28 和 APT 区 26 用沟道区 86 和 APT 区 74 来代替之外，使得晕轮区 46 与沟道区 86 和 APT 区 74 相邻。但是应当注意，在可选实施方案中，晕轮区 46 可以不存在。在该可选实施方案中，沟道区 86 和 APT 区 74 将与漏延伸区 53 和漏区 54 相邻。

如图 8 中所示（类似于图 4）， V_w 60 对应于施加到被隔离 P 型阱 20 的电压， V_s 62 对应于施加到源区 52 的电压， V_g 64 对应于施加到栅 34 的电压，并且 V_d 66 对应于施加到漏区 54 的电压。在所说明的实施方案中，半导体器件 70 可以用作 NVM 存储器（没有显示）中的 NVM 存储单元。如这里所使用的，高 V_t 状态对应于存储单元的编程状态，而低 V_t 状态对应于存储单元的擦除状态。（但是应当注意，在可选实施方案中，编程和擦除状态可以反过来。）

半导体器件 70 的编程和擦除操作与上面参考图 4 的半导体器件 10 所描述的相同。例如，在使用上面参考半导体器件 10 的编程而描述的电压的半导体器件 70 的编程过程中，热载流子在漏耗尽区中产生，其中一些通过氧化物 40 注入到氮化物 38 中。这导致增加半导体器件 70 的 V_t 。应当注意，如果晕轮区 46 存在，那么由晕轮区 46 和漏延伸区 53 形成的掺杂物梯度增强该热载流子注入，从而保持半导体器件 70 的高效热载流子编程。该效率被保持，即使沟道区 86 相对于 APT 区 74 被反掺杂。此外，沟道区 86 的反掺杂减小半导体器件 70 的 V_t ，从而改善读干扰，如将在下面描述的。

半导体器件 70 的自然 V_t 指的是把任何电荷置入氮化物 38 之前的阈值电压。如同半导体器件 10 的情况一样，对于半导体器件 70 的较高的自然 V_t ，读干扰被退化。因此，通过减小自然 V_t ，低 V_t 状态的读干扰可以改善（即阈值电压漂移减小）。较低的自然 V_t 减小读干扰的一种方法是通过允许低 V_t 状态的较低 V_t 。为了在半导体器件 70 的读取过程中

形成反型层，施加超过低 V_t 状态的 V_t 预先确定量的栅偏压 (V_g) (典型地称作栅过压) 是必要的。低 V_t 状态的减小的 V_t (通过沟道区 86 的反掺杂来获得)，允许在读操作过程中绝对栅偏压 (V_g) 减小，同时保持恒定的栅过压。减小的绝对栅偏压 (V_g) 将减小跨越栅叠层 32 的电场，从而导致减小的读干扰。

如果低 V_t 状态的减小的 V_t 太低 (由于沟道区 86 的反掺杂)，源漏漏电流可能在包含半导体器件 70 的存储阵列中未选择的器件中发生。未选择的器件是存储阵列中并不打算在半导体器件 70 的读操作过程中读取的那些器件。如在本领域中众所周知的，反向的阱源偏压增加低 V_t 状态的 V_t 。因此，源漏漏电流可以通过在半导体器件 70 的读操作过程中施加反向的阱源偏压到存储阵列中未选择的器件上来防止。反向的阱源偏压应当足以减小由低 V_t 状态的低 V_t 引起的源漏漏电流。例如，参考回图 8，半导体器件 70 的读取可以通过施加比 V_s 高大约 0.5~1.5 伏特的 V_d 来执行。例如，在一种实施方案中， V_s 可以是 0 伏特并且 V_d 可以是 1 伏特。在该实施方案中，足以在沟道区 28 中产生大约 10~30 微安培电流的 V_g 和 V_w 被施加。例如，在一种实施方案中，大约 1~2 伏特的 V_g 和大约 0~-3 伏特的 V_w 可以使用。应当注意，在该实例中提供的电压根据源电压 (V_s) 给出。也就是，在该实例中，如果 V_s 增加 1 伏特， V_d 、 V_g 和 V_w 也增加 1 伏特。

在读取或访问具有晕轮区 46 的已擦除半导体器件 70 (即处于低 V_t 状态的半导体器件 70) 的过程中，反型层在沟道区 86 中形成，并且耗尽区 (没有显示) 在漏区 54 和漏延伸区 53 周围形成。该耗尽区基本上掩蔽在晕轮区 46 中形成的掺杂物梯度，从而防止晕轮区 46 的较高掺杂物增加半导体 70 的 V_t 。这样， V_t 保持处于低 V_t 状态，从而通过减小 V_t 漂移而改善读干扰。并且，因为如上所述栅叠层 32 的长度大约为 0.35~0.06 微米，短沟道漏电可能在半导体器件 70 的编程过程中发生。但是，高掺杂 APT 区 74 也起到减小该短沟道漏电的作用，从而减小功率损耗并提高编程效率。

虽然本发明已关于具体导电型而描述，本领域技术人员应当认识

到，导电型可以反转。例如，源和漏以及延伸区可以是 p 型或 n 型，取决于被隔离阱的极性，以便形成 p 型或 n 型半导体器件。因此，被隔离阱 20 可能是 N 型阱而不是 P 型阱，并且源和漏区 52 和 54 以及延伸区 51 和 53 可以是 P 型。并且，在可选实施方案中，其他材料和工艺步骤可以用来形成半导体器件 10；上面所描述的那些仅作为例子提供。

在前述说明书中，本发明关于具体实施方案而描述。但是，本领域技术人员应当认识到，可以不背离在下面的权利要求书中陈述的本发明的范围而做各种修改和改变。因此，说明书和附图应当认为是说明性的而不是限制性的，并且所有这些修改打算包含在本发明的范围内。

好处、其他优点，以及问题的解决方案已在上面关于具体实施方案而描述。但是，该好处、优点、问题的解决方案，以及可能导致好处、优点或解决方案出现或变得更明显的任何元素不解释为任何或全部权利要求的苛刻、必需或关键特征或元素。如这里所使用的，术语“包括”、“包含”或任何其他变型打算覆盖非排他的结论，使得包括一系列元素的过程、方法、产品或装置不仅包括那些元素，而且可以包括没有明确列出的或这些过程、方法、产品或装置所固有的其他元素。

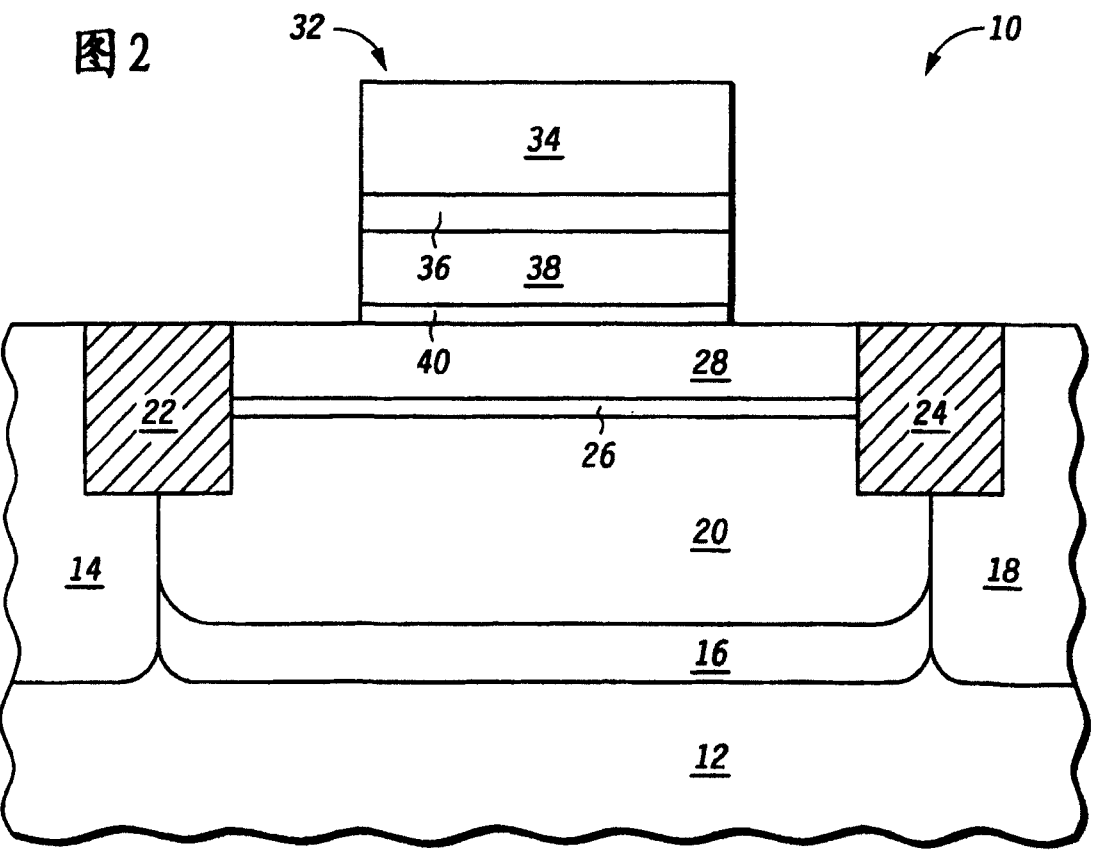
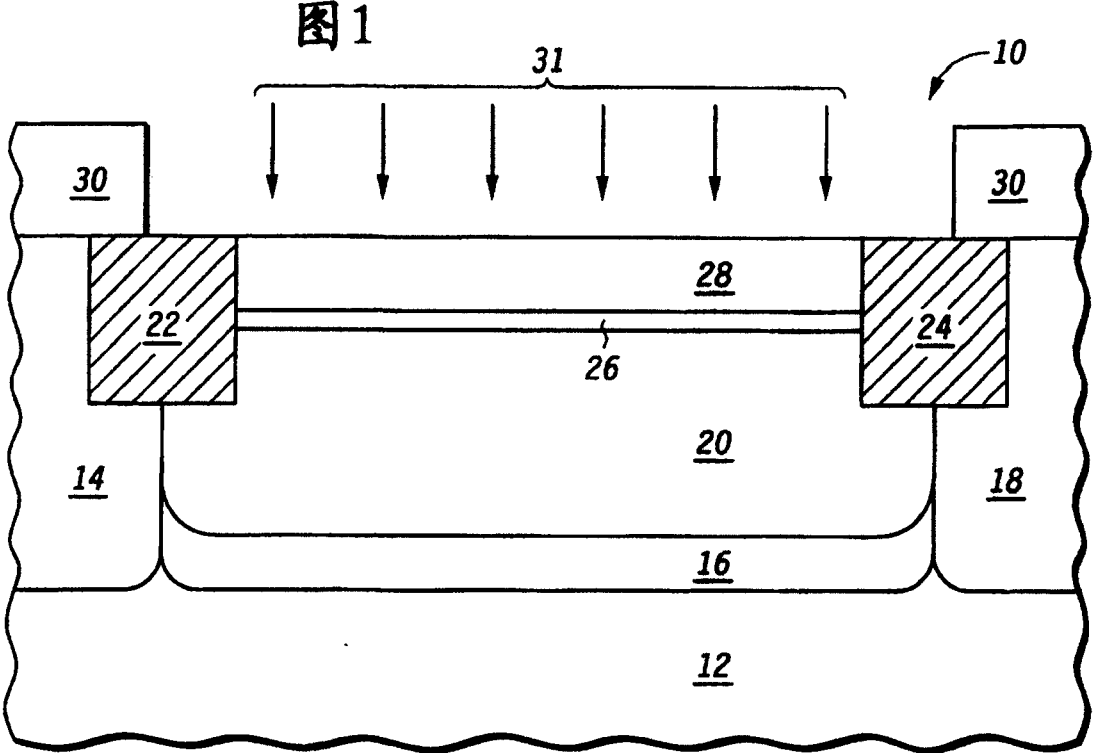


图 3

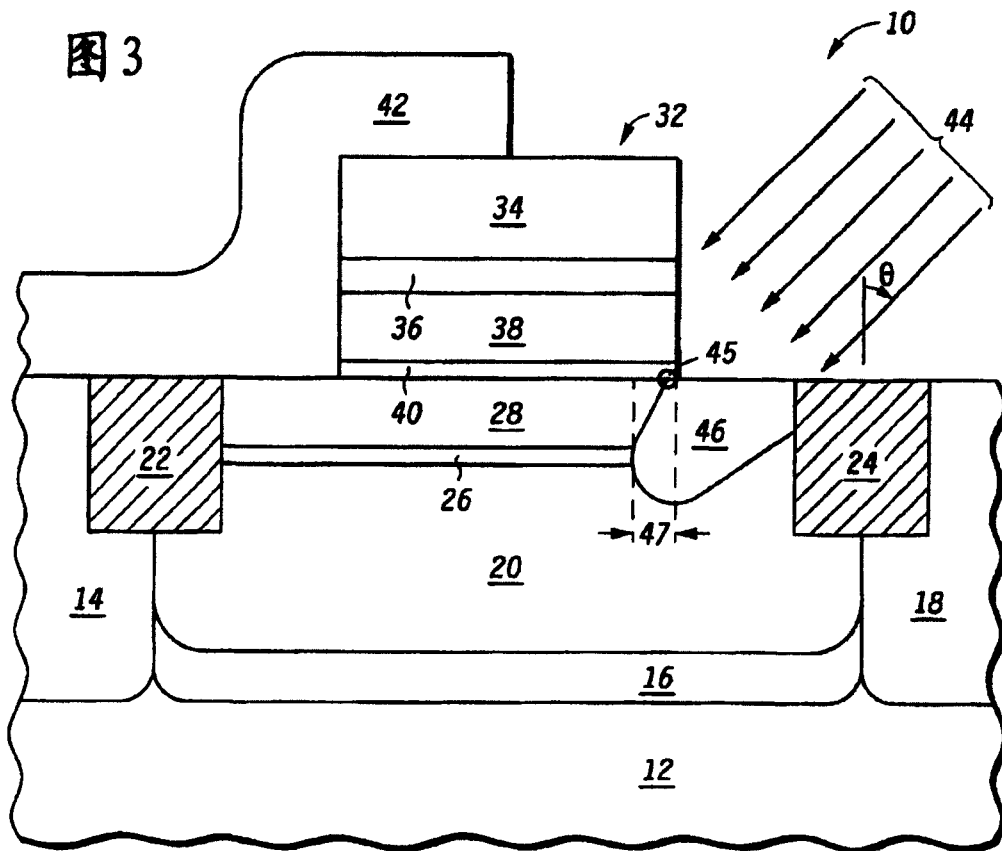


图 4

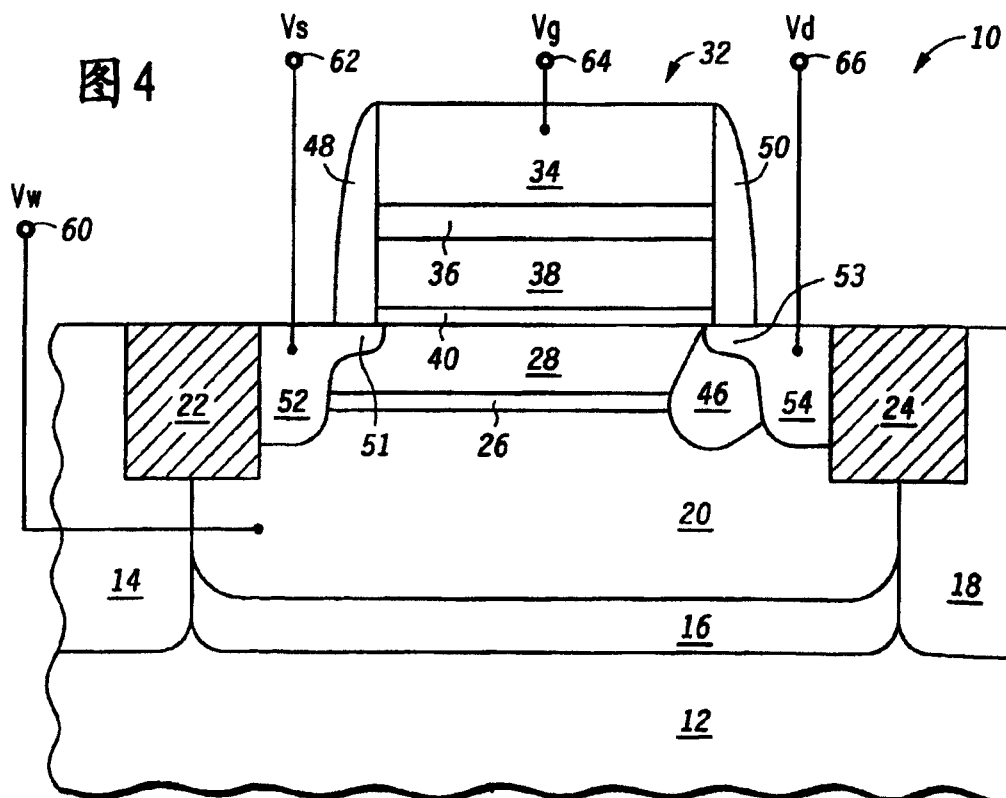


图 5

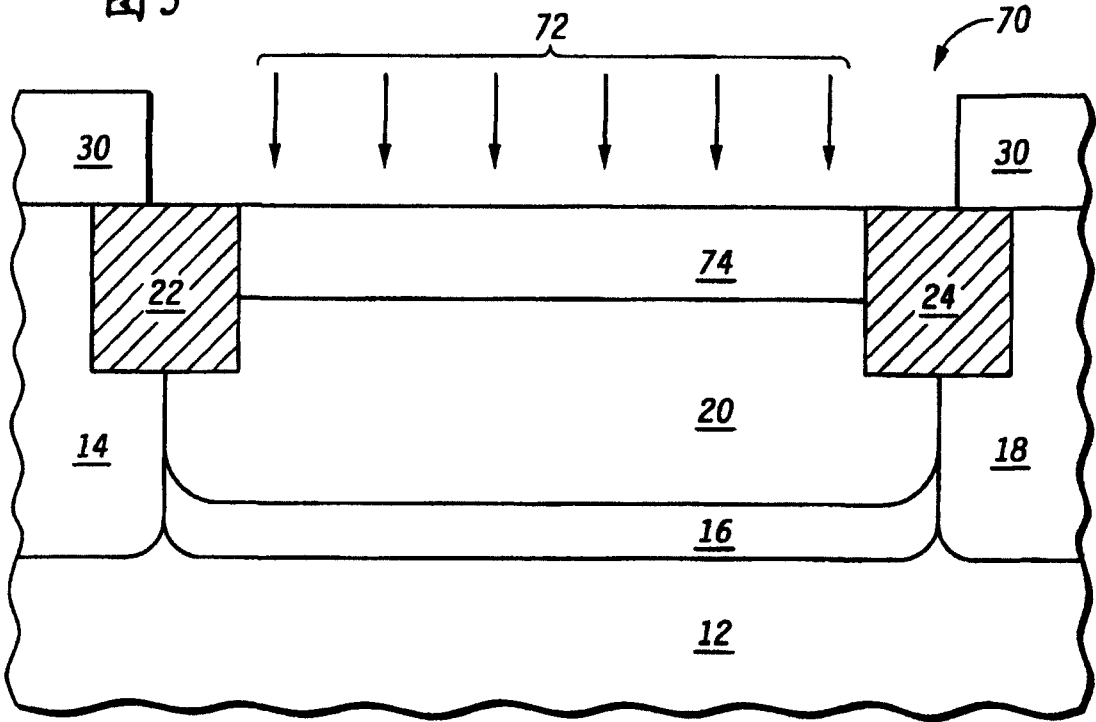


图 6

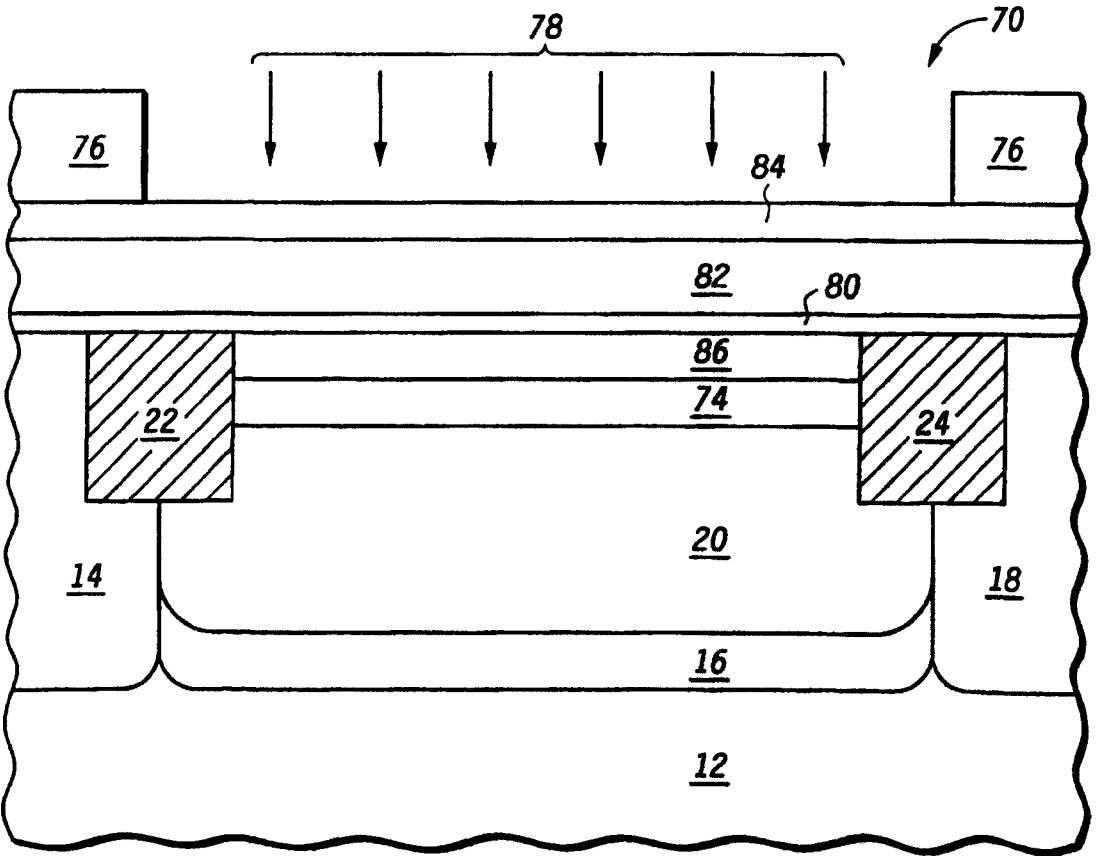


图 7

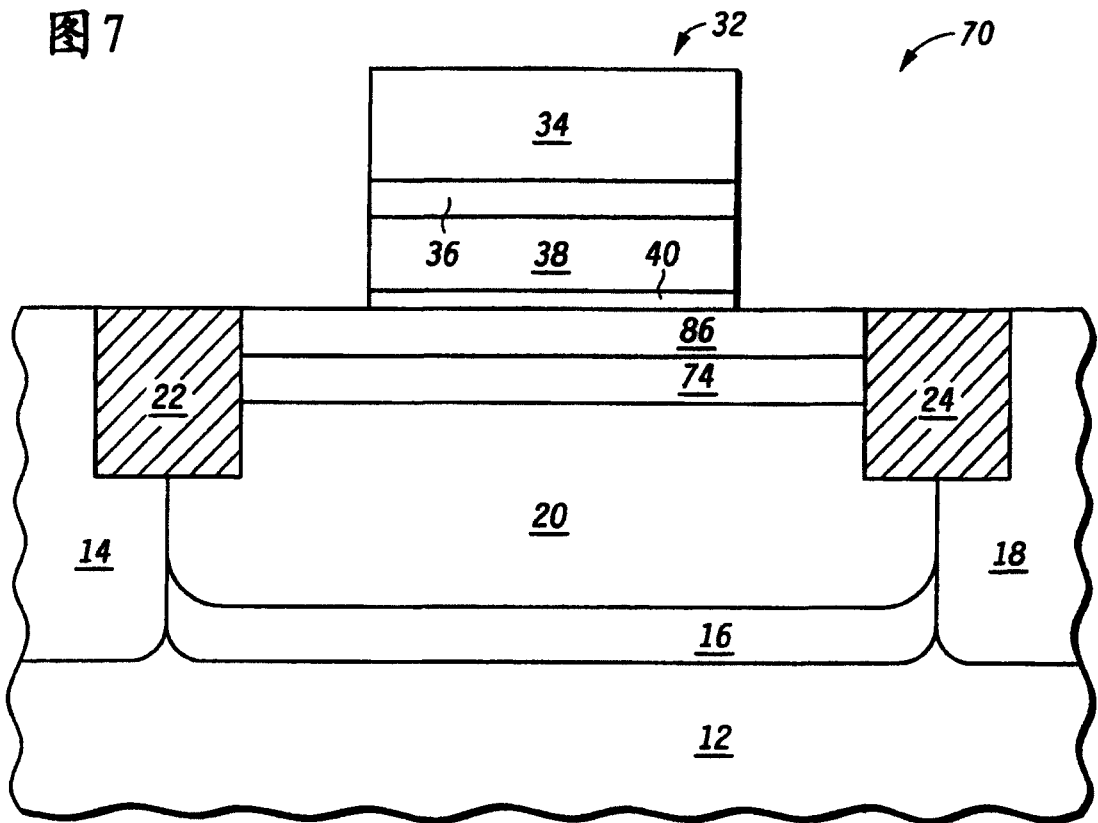


图 8

