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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 July 2007 (26.07.2007)

PCT

(10) International Publication Number
WO 2007/084465 A1

(51) International Patent Classification:
H01L 21/338 (2006.01) *H01L 29/778* (2006.01)
H01L 29/423 (2006.01) *H01L 21/335* (2006.01)

(21) International Application Number: PCT/US2007/001045

(22) International Filing Date: 16 January 2007 (16.01.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11/333,726 17 January 2006 (17.01.2006) US
11/493,069 26 July 2006 (26.07.2006) US

(71) Applicant *(for all designated States except US)*: **CREE, INC.** [US/US]; 4600 Silicon Drive, Durham, North Carolina 27703 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **SMITH, Richard, Peter** [US/US]; 242 Sweet Bay Place, Carrboro, North Carolina 27510 (US). **SHEPPARD, Scott, Thomas** [US/US]; 101 Autumn Lane, Chapel Hill, North Carolina 27516 (US). **ALLEN, Scott** [US/US]; 2740 Weaver Hill Drive, Apex, North Carolina 27502 (US).

(74) Agent: **MYERS BIGEL SIBLEY & SAJOVEC, P.A.**; P.O. Box 37428, Raleigh, North Carolina 27627 (US).

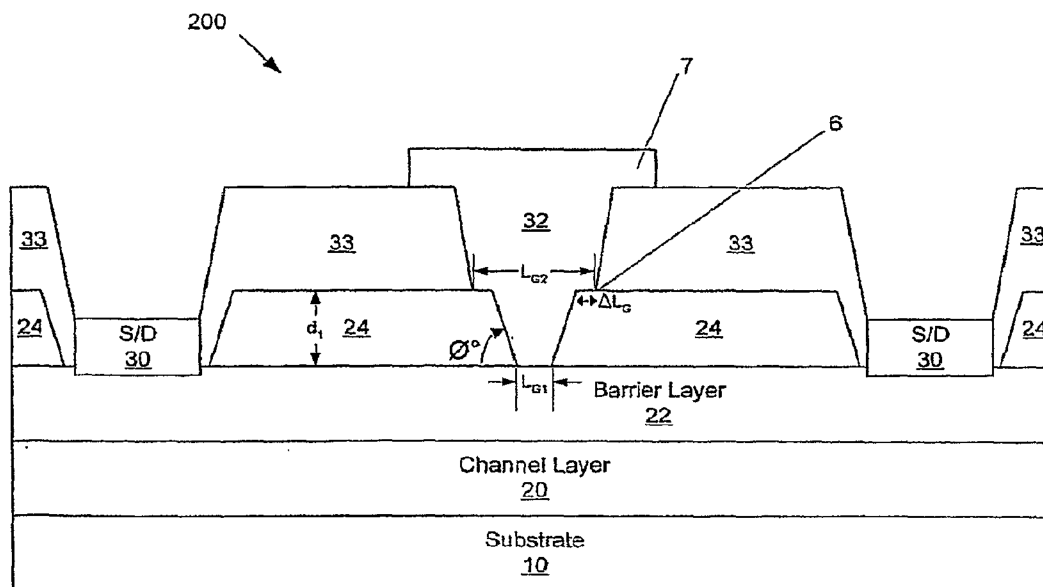
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR,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

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

[Continued on next page]

(54) Title: METHODS OF FABRICATING TRANSISTORS INCLUDING SUPPORTED GATE ELECTRODES AND RELATED DEVICES



(57) Abstract: Transistors are fabricated by forming a protective layer having an opening extending therethrough on a substrate, and forming a gate electrode in the opening. A first portion of the gate electrode laterally extends on surface portions of the protective layer outside the opening, and a second portion of the gate electrode is spaced apart from the protective layer and laterally extends beyond the first portion. Related devices and fabrication methods are also discussed.

WO 2007/084465 A1

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

METHODS OF FABRICATING TRANSISTORS INCLUDING SUPPORTED GATE ELECTRODES AND RELATED DEVICES

FIELD OF THE INVENTION

[0001] The present invention relates to semiconductor devices, and, more particularly, to methods of fabricating transistors and related devices.

BACKGROUND OF THE INVENTION

[0002] Materials such as silicon (Si) and gallium arsenide (GaAs) have found wide application in semiconductor devices for lower power and, in the case of Si, lower frequency applications. However, these more familiar semiconductor materials may not be well suited for higher power and/or high frequency applications, for example, due to their relatively small bandgaps (*e.g.*, 1.12 eV for Si and 1.42 for GaAs at room temperature) and/or relatively small breakdown voltages.

[0003] In light of the difficulties presented by Si and GaAs, interest in high power, high temperature and/or high frequency applications and devices has turned to wide bandgap semiconductor materials such as silicon carbide (2.996 eV for alpha SiC at room temperature) and the Group III nitrides (*e.g.*, 3.36 eV for GaN at room temperature). These materials, typically, may have higher electric field breakdown strengths and higher electron saturation velocities as compared to gallium arsenide and/or silicon.

[0004] A device of particular interest for high power and/or high frequency applications is the High Electron Mobility Transistor (HEMT), which is also known as a modulation doped field effect transistor (MODFET). In a HEMT device, a two-dimensional electron gas (2DEG) may be formed at the heterojunction of two semiconductor materials with different bandgap energies. The smaller bandgap material may have a higher electron affinity than the wider bandgap material. The 2DEG is an accumulation layer in the undoped ("unintentionally doped") smaller bandgap material, and can contain a relatively high sheet electron concentration, for example, in excess of 10^{13} carriers/cm^2. Additionally, electrons that originate in the wider bandgap semiconductor may transfer to the 2DEG, allowing a relatively high electron mobility due to reduced ionized impurity scattering. This combination of

relatively high carrier concentration and relatively high carrier mobility can give the HEMT a relatively large transconductance, and may provide a performance advantage over metal-semiconductor field effect transistors (MESFETs) for high-frequency applications.

[0005] High electron mobility transistors fabricated in the gallium nitride/aluminum gallium nitride (GaN/AlGaN) material system can generate large amounts of RF power due to a combination of material characteristics, such as relatively high breakdown fields, relatively wide bandgaps, relatively large conduction band offset, and/or relatively high saturated electron drift velocity. A major portion of the electrons in the 2DEG may be attributed to polarization in the AlGaN.

[0006] HEMTs in the GaN/AlGaN system have already been demonstrated. For example, U.S. Patents 5,192,987 and 5,296,395 describe AlGaN/GaN HEMT structures and methods of manufacture. In addition, U.S. Patent No. 6,316,793, to Sheppard et al., which is commonly assigned and incorporated by reference herein, describes a HEMT device having a semi-insulating silicon carbide substrate, an aluminum nitride buffer layer on the substrate, an insulating gallium nitride layer on the buffer layer, an aluminum gallium nitride barrier layer on the gallium nitride layer, and a passivation layer on the aluminum gallium nitride active structure. Moreover, U.S. Patent Application Publication No. US2005/0170574 to Sheppard et al., which is also commonly assigned and incorporated by reference herein, describes a HEMT device including a protective layer and/or a low damage recess fabrication technique which may reduce damage to the semiconductor in the gate region of the transistor that may occur during an anneal of ohmic contacts of the device.

[0007] One step in the fabrication of HEMT devices is the formation of the gate electrode. Conventional methods of gate electrode formation may include depositing a dielectric, etching through the dielectric using a mask and/or other sacrificial layer, and depositing a T-shaped gate electrode (referred to as a "T-gate") into the etched portion of the dielectric. However, in such conventional methods, gaps may be formed between the edges of the dielectric and the gate electrode, due to isotropy of the dielectric etch. This may be detrimental to device operation, as the unpassivated semiconductor surface exposed by the gap may cause current collapse and/or drift in the device. In addition, although a passivation layer may be formed in the gap after formation of the gate, the passivation properties of such a post-gate

passivation layer may be inferior to that of the initial pre-gate dielectric. This may be due to the fact that the pre-gate dielectric can be formed at relatively high temperatures, which may not be feasible once the gate metallization has been deposited.

[0008] Accordingly, other methods of gate electrode formation have been developed to prevent formation of such a gap. For example, the mask and/or other sacrificial layer may be removed from the dielectric prior to formation of the gate electrode in the etched portion of the dielectric. As such, the gate electrode may completely fill the etched portion of the dielectric, and "wings" or sidelobes of the gate electrode may be formed directly on the dielectric surface. Such a gate electrode is referred to as a dielectrically-supported T-gate, or a "gamma" gate. However, because portions of the dielectric may be sandwiched between the sidelobes and the active channel, gate-drain capacitance (c_{gd}) and/or gate-source capacitance (c_{gs}) may be increased, which may detrimentally affect device performance.

SUMMARY OF THE INVENTION

[0009] According to some embodiments of the present invention, methods for fabricating a transistor may include forming a protective layer on a substrate. The protective layer may have an opening extending therethrough. A gate electrode may be formed in the opening, for example, directly on opposing sidewalls of the protective layer. A first portion of the gate electrode may laterally extend on surface portions of the protective layer outside the opening, and a second portion of the gate electrode may be spaced apart from the protective layer and may laterally extend beyond the first portion.

[0010] In some embodiments, before the gate electrode is formed, a second layer may be formed on the protective layer. The second layer may have a second opening extending therethrough that is wider than the first opening. The gate electrode may be formed in the second opening, such that the second portion of the gate electrode may laterally extend on surface portions of the second layer outside the second opening.

[0011] In other embodiments, the second layer may be a material having a lower dielectric index than the protective layer.

[0012] In some embodiments, the second layer may be formed on the protective layer and may include a recess extending through the second layer. The

recess may expose a portion of the protective layer. The recess in the second layer may be widened to define the second opening exposing the first opening and exposing surface portions of the protective layer on opposite sides of the first opening.

[0013] In other embodiments, the protective layer and the second layer may be formed of different materials. The protective layer may be patterned using the second layer as a mask to form the first opening extending through the protective layer prior to widening the recess in the second layer. The recess in the second layer may be widened by symmetrically expanding the recess in the second layer. As such, the second opening and the first opening may be self-aligned.

[0014] In some embodiments, the second layer may be a photoresist layer. The recess in the second layer may be widened by ashing the second layer using an oxygen plasma and/or by hard baking the substrate. In other embodiments, the second layer may be a polymide layer.

[0015] In other embodiments, the second layer may be a sacrificial layer. As such, the second layer may be removed after forming the gate electrode. A passivation layer may be formed on the protective layer and on the gate electrode after removing the second layer.

[0016] In still other embodiments, a channel layer may be formed on the substrate, and a barrier layer may be formed on the channel layer. The protective layer may be formed on the barrier layer. The gate electrode may be formed to extend through the opening in the protective layer to contact the barrier layer. A junction between the channel layer and the barrier layer may define a heterojunction.

[0017] In some embodiments, the channel layer may have a lower bandgap than the barrier layer.

[0018] In other embodiments, first and second ohmic contact regions may be formed on the barrier layer adjacent to and spaced apart from the protective layer. As such, the protective layer may be between the first and second ohmic contact regions. The first and second ohmic contact regions may be formed by patterning the protective layer to expose portions of the barrier layer, forming ohmic metal regions on the exposed portions of the barrier layer adjacent to and spaced apart from the patterned protective layer, and annealing the ohmic metal regions.

[0019] In some embodiments, the barrier layer may be formed of Group-III nitride layer, and the channel layer may be formed of an undoped Group-III nitride layer. For example, the barrier layer may be formed of aluminum gallium nitride

(AlGaN), and the channel layer may be formed of gallium nitride (GaN). The protective layer may be a dielectric material, such as silicon nitride (SiN), aluminum nitride (AlN), and/or silicon dioxide (SiO_2).

[0020] According to further embodiments of the present invention, a method of fabricating a transistor may include forming a first layer on a substrate. The first layer may be a dielectric material. A second layer may be formed on the first layer. The second layer may have a hole therein exposing a portion of the first layer. A first opening extending through the first layer may be formed using the second layer as a mask. After forming the first opening, the hole in the second layer may be widened to form a second opening. The second opening may expose the first opening, as well as surface portions of the first layer on opposite sides of the first opening. A gate electrode may be formed in the first and second openings. For example, the gate electrode may be formed directly on opposing sidewalls of the first opening. First portions of the gate electrode may extend on surface portions of the first layer outside the first opening, and second portions of the gate electrode may extend on surface portions of the second layer outside the second opening.

[0021] In some embodiments, the first portions of the gate electrode may laterally extend on surface portions of the first layer outside the opening. The second portions of the gate electrode may laterally extend on the surface portions of the second layer outside the second opening beyond the first portions of the gate electrode.

[0022] In other embodiments, the first layer and the second layer may be formed of different materials. For example, the first layer may be formed of a dielectric material, such as silicon nitride, aluminum nitride, and/or silicon dioxide. The second layer may be formed of a material having a lower dielectric index than the first layer.

[0023] In some embodiments, the hole in the second layer may be widened by symmetrically expanding the hole in the second layer, such that the second opening and the first opening may be self-aligned.

[0024] In other embodiments, the second layer may be a photoresist layer. The hole in the second layer may be widened by ashing the second layer using an oxygen plasma and/or hard baking the substrate. In still other embodiments, the second layer may be a polyimide layer.

[0025] In some embodiments, the first layer may be patterned by selectively etching the portion of the first layer exposed by the hole using a low-damage etching technique to form the first opening.

[0026] In other embodiments, the hole in the second layer may be widened by selectively etching the second layer using a wet etch to form the second opening.

[0027] In some embodiments, a channel layer may be formed on the substrate, and a barrier layer may be formed on the channel layer. The first layer may be formed on the barrier layer. The gate electrode may be formed to extend through the first opening in the first layer to contact the barrier layer. A junction between the channel layer and the barrier layer may define a heterojunction.

[0028] According to other embodiments of the present invention, a transistor may include a substrate, a protective layer on the substrate, and a gate electrode. The protective layer may have an opening extending therethrough, and the gate electrode may be in the opening. For example, the gate electrode may be directly on opposing sidewalls of the opening in the protective layer. The gate electrode may include a first portion laterally extending on surface portions of the protective layer outside the opening, and a second portion spaced apart from the protective layer and laterally extending beyond the first portion.

[0029] In some embodiments, the transistor may include a second layer on the protective layer. For example, the second layer may be a sacrificial layer. The second layer may have a lower dielectric index than the protective layer. The second layer may have a second opening extending therethrough that is wider than the first opening. The gate electrode may be in the second opening, and the second portion of the gate electrode may laterally extend on portions of the second layer outside the second opening. The first and second openings may define a stair-step profile.

[0030] In other embodiments, the transistor may include a passivation layer on the protective layer and on the gate electrode. The protective layer and the passivation layer may be formed of a same material.

[0031] In still other embodiments, the transistor may include a channel layer on the substrate, and a barrier layer on the channel layer. The protective layer may be on the barrier layer. The gate electrode may extend through the opening in the protective layer to contact the barrier layer. The channel layer and the barrier layer may be configured to provide a High Electron Mobility Transistor (HEMT). For

example, the channel layer and the barrier layer have different lattice constants, and as such, may provide a pseudomorphic HEMT.

[0032] In some embodiments, the transistor may include first and second ohmic contact regions on the barrier layer. The first and second ohmic contact regions may be on opposite sides of the gate electrode and spaced apart from the protective layer.

[0033] The protective layer may have a thickness of at least about a thickness of the ohmic contact regions.

[0034] In other embodiments, the barrier layer may be a Group-III nitride layer, and the channel layer may be an undoped Group-III nitride layer. For example, the barrier layer may be aluminum gallium nitride (AlGaN), and the channel layer may be gallium nitride (GaN). The protective layer may be a dielectric material. For example, the protective layer may be silicon nitride, aluminum nitride, and/or silicon dioxide. In other embodiments, the substrate may be silicon carbide.

[0035] According to some embodiments of the present invention, a method of fabricating a transistor may include forming a protective layer having a first opening extending therethrough. A dielectric layer may be formed on the protective layer. The dielectric layer may have a second opening extending therethrough that is wider than the first opening. A gate electrode may be formed in the first and second openings such that a first portion of the gate electrode laterally extends on surface portions of the protective layer outside the first opening, and a second portion of the gate electrode laterally extends beyond the first portion on portions of the dielectric layer outside the second opening.

[0036] In some embodiments, the dielectric layer may have a lower dielectric index than the protective layer. For example, the dielectric layer may be an oxide layer. More particularly, the dielectric layer may be a high-temperature oxide (HTO) layer, and the protective layer may be a high-purity nitride (HPN) layer. For instance, the protective layer may be stoichiometric silicon nitride, and the dielectric layer may be silicon dioxide.

[0037] In other embodiments, the dielectric layer may be a high-quality oxide layer. The high-quality oxide layer may have a dielectric index of less than about 1.5.

[0038] In some embodiments, the dielectric layer may have a thickness greater than that of the protective layer. For example, the dielectric layer may have a thickness of about 500 Angstroms (Å) to about 3000 Å, and the protective layer may

have a thickness of about 200 Å to about 2000 Å. The thickness and/or the dielectric index of the dielectric layer may be sufficient to reduce the gate-to-drain and/or gate-to-source capacitance of the device.

[0039] In other embodiments, forming the dielectric layer may include depositing an oxide layer on the protective layer using a chemical vapor deposition (CVD) process. The oxide layer may be deposited at a temperature of greater than about 500 °C.

[0040] In some embodiments, forming the dielectric layer having the second opening extending therethrough may include forming the dielectric layer on the protective layer including a recess extending therethrough. The recess may expose a portion of the protective layer. The recess in the dielectric layer may be widened to define the second opening. The second opening may expose the first opening, and may expose surface portions of the protective layer on opposite sides of the first opening.

[0041] In other embodiments, the protective layer and the dielectric layer may be formed of different materials. Accordingly, prior to widening the recess in the dielectric layer, the protective layer may be patterned using the dielectric layer as a mask to form the first opening extending through the protective layer.

[0042] In some embodiments, the recess in the dielectric layer may be widened by symmetrically expanding the recess in the dielectric layer so that the second opening and the first opening are self-aligned. For example, the recess may be symmetrically expanded by etching the dielectric layer using wet oxide etchant that is selective with respect to the protective layer, such as a buffered oxide etchant (BOE).

[0043] In other embodiments, the method of fabricating the transistor may include forming a channel layer, and forming a barrier layer on the channel layer. The protective layer may be formed on the barrier layer. The gate electrode may be formed extending through the first opening in the protective layer to contact the barrier layer. A junction between the channel layer and the barrier layer may define a heterojunction.

[0044] In some embodiments, first and second ohmic contact regions may be formed on the barrier layer adjacent to and spaced apart from the protective layer so that the protective layer is between the first and second ohmic contact regions. More particularly, the dielectric layer and the protective layer may be patterned to expose first and second portions of the barrier layer. After patterning the dielectric layer and

the protective layer, the first and second ohmic contact regions may be respectively formed on the first and second portions of the barrier layer adjacent to and spaced apart from the patterned protective layer. First and second ohmic metal regions may be annealed to form the first and second ohmic contacts.

[0045] In other embodiments, before forming the dielectric layer, the protective layer may be patterned to expose the first and second portions of the barrier layer. The dielectric layer may be formed on the protective layer and on the first and second portions of the barrier layer. The dielectric layer may be patterned to define first and second recesses therein respectively exposing the first and second portions of the barrier layer. In addition, the first and second recesses in the dielectric layer may be symmetrically expanded using a wet oxide etchant, such as a buffered oxide etchant (BOE).

[0046] In some embodiments, the dielectric layer may be patterned to expose first and second portions of the protective layer. The protective layer may be patterned using the dielectric layer as a mask to expose the first and second portions of the barrier layer.

[0047] In other embodiments, the gate electrode may be formed in the first opening directly on opposing sidewalls of the protective layer.

[0048] In some embodiments, a dielectric liner may be formed in the first opening on opposing sidewalls of the protective layer. For example, the dielectric liner may be a silicon dioxide liner or a nitride liner. After forming the dielectric liner, the gate electrode may be formed in the first opening directly on the dielectric liner.

[0049] In some embodiments, the barrier layer may be formed of Group-III nitride layer, and the channel layer may be formed of an undoped Group-III nitride layer. For example, the barrier layer may be formed of aluminum gallium nitride (AlGaN), and the channel layer may be formed of gallium nitride (GaN). The protective layer may be a dielectric material, such as silicon nitride (SiN), aluminum nitride (AlN), and/or silicon dioxide (SiO_2).

[0050] According to further embodiments of the present invention, a transistor may include a protective layer having a first opening extending therethrough, and a dielectric layer on the protective layer. The dielectric layer may have a second opening extending therethrough that is wider than the first opening. The transistor may also include a gate electrode in the first and second openings. The gate electrode

may include a first portion laterally extending on surface portions of the protective layer outside the first opening, and a second portion laterally extending beyond the first portion on portions of the dielectric layer outside the second opening. For example, the first and second openings may define a stair-step profile.

[0051] In some embodiments, the dielectric layer and the protective layer may be formed of different materials. The dielectric layer may have a lower dielectric index than the protective layer. For example, the dielectric layer may be an oxide layer. More particularly, the dielectric layer may be a high-temperature oxide (HTO) layer, and the protective layer may be a high-purity nitride (HPN) layer. For instance, the protective layer may be stoichiometric silicon nitride, and the dielectric layer may be silicon dioxide.

[0052] In other embodiments, the dielectric layer may be a high-quality oxide layer. The high-quality oxide layer may have a dielectric index of less than about 1.5.

[0053] In still other embodiments, the dielectric layer may have a thickness greater than that of the protective layer. For example, the dielectric layer may have a thickness of about 500 Angstroms (Å) to about 3000 Å, and the protective layer may have a thickness of about 200 Å to about 2000 Å. The thickness and/or the dielectric index of the dielectric layer may be sufficient to reduce the gate-to-drain and/or gate-to-source capacitance of the device.

[0054] In some embodiments, the transistor may further include a channel layer and a barrier layer on the channel layer. The protective layer may be on the barrier layer, and the gate electrode may extend through the first opening in the protective layer to contact the barrier layer. The channel layer and the barrier layer may be configured to provide a High Electron Mobility Transistor (HEMT). For example, the channel layer and the barrier layer have different lattice constants, and as such, may provide a pseudomorphic HEMT.

[0055] In other embodiments, the transistor may include first and second ohmic contact regions on the barrier layer at opposite sides of the gate electrode and spaced apart from the protective layer. For example, the dielectric layer may extend between the first and second ohmic contact regions and the protective layer.

[0056] In some embodiments, the gate electrode may be directly on opposing sidewalls of the first opening in the protective layer.

[0057] In other embodiments, a dielectric liner may extend in the first opening between a sidewall of the protective layer and the gate electrode. The gate electrode may be directly on the dielectric liner.

[0058] In some embodiments, the barrier layer may be a Group-III nitride layer, and the channel layer may be an undoped Group-III nitride layer. For example, the barrier layer may be aluminum gallium nitride (AlGaN), and the channel layer may be gallium nitride (GaN). The protective layer may also be a dielectric material. For example, the protective layer may be silicon nitride, aluminum nitride, and/or silicon dioxide. In other embodiments, the substrate may be silicon carbide.

BRIEF DESCRIPTION OF THE DRAWINGS

[0059] **Figures 1** to **9** are cross-sectional views illustrating intermediate fabrication steps in operations for fabricating transistor devices according to some embodiments of the present invention.

[0060] **Figure 10** is a cross-sectional view illustrating a transistor device according to some embodiments of the present invention.

[0061] **Figures 11** to **18** are cross-sectional views illustrating intermediate fabrication steps in operations for fabricating transistor devices according to further embodiments of the present invention.

DETAILED DESCRIPTION OF EMBODIMENTS OF THE INVENTION

[0062] The present invention now will be described more fully hereinafter with reference to the accompanying drawings, in which embodiments of the invention are shown. However, this invention should not be construed as limited to the embodiments set forth herein. Rather, these embodiments are provided so that this disclosure will be thorough and complete, and will fully convey the scope of the invention to those skilled in the art. In the drawings, the thickness of layers and regions are exaggerated for clarity. Like numbers refer to like elements throughout.

[0063] It will be understood that when an element such as a layer, region or substrate is referred to as being "on" or extending "onto" another element, it can be directly on or extend directly onto the other element or intervening elements may also be present. In contrast, when an element is referred to as being "directly on" or extending "directly onto" another element, there are no intervening elements present. It will also be understood that when an element is referred to as being "connected" or

"coupled" to another element, it can be directly connected or coupled to the other element or intervening elements may be present. In contrast, when an element is referred to as being "directly connected" or "directly coupled" to another element, there are no intervening elements present.

[0064] It will also be understood that, although the terms first, second, etc. may be used herein to describe various elements, these elements should not be limited by these terms. These terms are only used to distinguish one element from another. For example, a first element could be termed a second element, and, similarly, a second element could be termed a first element, without departing from the scope of the present invention.

[0065] Furthermore, relative terms, such as "lower" or "bottom" and "upper" or "top," may be used herein to describe one element's relationship to another element as illustrated in the Figures. It will be understood that relative terms are intended to encompass different orientations of the device in addition to the orientation depicted in the Figures. For example, if the device in one of the figures is turned over, elements described as being on the "lower" side of other elements would then be oriented on "upper" sides of the other elements. The exemplary term "lower", can therefore, encompasses both an orientation of "lower" and "upper," depending of the particular orientation of the figure. Similarly, if the device in one of the figures is turned over, elements described as "below" or "beneath" other elements would then be oriented "above" the other elements. The exemplary terms "below" or "beneath" can, therefore, encompass both an orientation of above and below.

[0066] The terminology used in the description of the invention herein is for the purpose of describing particular embodiments only and is not intended to be limiting of the invention. As used in the description of the invention and the appended claims, the singular forms "a", "an " and "the" are intended to include the plural forms as well, unless the context clearly indicates otherwise. It will also be understood that the term "and/or" as used herein refers to and encompasses any and all possible combinations of one or more of the associated listed items. It will be further understood that the terms "comprises" and/or "comprising," when used in this specification, specify the presence of stated features, integers, steps, operations, elements, and/or components, but do not preclude the presence or addition of one or more other features, integers, steps, operations, elements, components, and/or groups thereof.

[0067] Embodiments of the invention are described herein with reference to cross-section illustrations that are schematic illustrations of idealized embodiments (and intermediate structures) of the invention. As such, variations from the shapes of the illustrations as a result, for example, of manufacturing techniques and/or tolerances, are to be expected. Thus, embodiments of the invention should not be construed as limited to the particular shapes of regions illustrated herein but are to include deviations in shapes that result, for example, from manufacturing. For example, an implanted region illustrated as a rectangle will, typically, have rounded or curved features and/or a gradient of implant concentration at its edges rather than a binary change from implanted to non-implanted region. Likewise, a buried region formed by implantation may result in some implantation in the region between the buried region and the surface through which the implantation takes place. Thus, the regions illustrated in the figures are schematic in nature and their shapes are not intended to illustrate the actual shape of a region of a device and are not intended to limit the scope of the invention.

[0068] Unless otherwise defined, all terms used in disclosing embodiments of the invention, including technical and scientific terms, have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs, and are not necessarily limited to the specific definitions known at the time of the present invention being described. Accordingly, these terms can include equivalent terms that are created after such time. It will be further understood that terms, such as those defined in commonly used dictionaries, should be interpreted as having a meaning that is consistent with their meaning in the present specification and in the context of the relevant art and will not be interpreted in an idealized or overly formal sense unless expressly so defined herein. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety.

[0069] Embodiments of the present invention may be particularly well suited for use in nitride-based HEMTs, such as Group III-nitride based devices. As used herein, the term "Group III nitride" refers to those semiconducting compounds formed between nitrogen and the elements in Group III of the periodic table, usually aluminum (Al), gallium (Ga), and/or indium (In). The term also refers to ternary and quaternary compounds such as AlGaN and AlInGaN. As is well understood by those in this art, the Group III elements can combine with nitrogen to form binary (*e.g.*,

GaN), ternary (*e.g.*, AlGaN, AlInN), and quaternary (*e.g.*, AlInGaN) compounds. These compounds may all have empirical formulas in which one mole of nitrogen is combined with a total of one mole of the Group III elements. Accordingly, formulas such as $Al_xGa_{1-x}N$, where $0 \leq x \leq 1$, may be used to describe these compounds.

[0070] Suitable structures for GaN-based HEMTs that may utilize embodiments of the present invention are described, for example, in commonly assigned U.S. Patent 6,316,793 and U.S. Patent Publication No. 2002/0066908A1 filed July 12, 2001 and published June 6, 2002, for "ALUMINUM GALLIUM NITRIDE/GALLIUM NITRIDE HIGH ELECTRON MOBILITY TRANSISTORS HAVING A GATE CONTACT ON A GALLIUM NITRIDE BASED CAP SEGMENT AND METHODS OF FABRICATING SAME", United States Patent Publication No. 2002/0167023A1 to Smorchkova *et al.*, published November 14, 2002, entitled "GROUP-III NITRIDE BASED HIGH ELECTRON MOBILITY TRANSISTOR (HEMT) WITH BARRIER/SPACER LAYER", and United States Patent Application Serial No. 10/617,843 filed July 11, 2003 for "NITRIDE-BASED TRANSISTORS AND METHODS OF FABRICATION THEREOF USING NON-ETCHED CONTACT RECESSES," the disclosures of which are hereby incorporated herein by reference in their entirety.

[0071] Some embodiments of the present invention may arise from a realization that, for devices operating at relatively high frequencies (for example, above the X-band frequency range), and particularly for millimeter wave radio frequency (RF) devices, a tradeoff may exist between minimizing peak electric fields and/or surface-related trapping and maintaining relatively low capacitance values. Accordingly, some embodiments of the present invention provide a transistor device that combines some of the positive aspects of a T-gate (such as relatively low gate resistance and/or relatively low c_{gd} and c_{gs}) with some of the positive attributes of a gamma gate (such as pre-gate passivation and/or peak electric field reduction).

[0072] **Figures 1** to **9** are cross-sectional views illustrating exemplary intermediate fabrication steps in methods for fabricating transistor devices according to some embodiments of the present invention. Referring now to **Figure 1**, a substrate **10** is provided on which a transistor device may be formed. A channel layer **20** is formed on the substrate **10**, and a barrier layer **22** is formed on the channel layer **20**. The substrate **10** may be a semi-insulating silicon carbide (SiC) substrate that

may be, for example, the 4H polytype of silicon carbide. Other silicon carbide candidate polytypes may include the 3C, 6H, and 15R polytypes. The term "semi-insulating" is used descriptively herein, rather than in an absolute sense. In some embodiments of the present invention, the silicon carbide bulk crystal may have a resistivity equal to or higher than about 1×10^5 Ω-cm at room temperature.

[0073] Silicon carbide has a much closer crystal lattice match to Group III nitrides (which may be employed in the channel layer **20** and/or the barrier layer **22**) than does sapphire (Al_2O_3), which may be a common substrate material for Group III nitride devices. The closer lattice match may result in Group III nitride films of higher quality than those generally available on sapphire. Silicon carbide also has a relatively high thermal conductivity, and as such, the total output power of Group III nitride devices formed on silicon carbide may not be as limited by thermal dissipation of the substrate as similar devices formed on sapphire and/or silicon. Also, semi-insulating silicon carbide substrates may provide for device isolation and reduced parasitic capacitance. Exemplary SiC substrates that may be used in some embodiments of the present invention are manufactured by, for example, Cree, Inc., of Durham, N.C., the assignee of the present invention, and methods for producing such substrates are described, for example, in U. S. Patent Nos. Re. 34,861; 4,946,547; 5,200,022; and 6,218,680, the disclosures of which are incorporated by reference herein in their entireties. Similarly, techniques for epitaxial growth of Group III nitrides have been described in, for example, U. S. Patent Nos. 5,210,051; 5,393,993; 5,523,589; and 5,292,501, the disclosures of which are also incorporated by reference herein in their entireties.

[0074] It is to be understood that, although silicon carbide may be employed as a substrate, embodiments of the present invention may utilize any suitable substrate for the substrate **10**, such as sapphire (Al_2O_3), aluminum nitride (AlN), aluminum gallium nitride (AlGaN), gallium nitride (GaN), silicon (Si), gallium arsenide (GaAs), LGO, zinc oxide (ZnO), LAO, indium phosphide (InP), and the like.

[0075] Optional buffer, nucleation and/or transition layers (not shown) may also be provided on the substrate **10**. For example, an AlN buffer layer may be provided to provide an appropriate crystal structure transition between a silicon carbide substrate and the remainder of the device. Additionally, strain balancing transition layer(s) may also be provided as described, for example, in commonly assigned United States Patent Publication 2003/0102482A1, filed July 19, 2002 and

published June 5, 2003, and entitled "STRAIN BALANCED NITRIDE HETEROJUNCTION TRANSISTORS AND METHODS OF FABRICATING STRAIN BALANCED NITRIDE HETEROJUNCTION TRANSISTORS, and United States Provisional Patent Application Serial No. 60/337,687, filed December 3, 2001 and entitled "STRAIN BALANCED NITRIDE HETEROJUNCTION TRANSISTOR," the disclosures of which are incorporated by reference herein as if set forth fully herein.

[0076] Still referring to **Figure 1**, a channel layer **20** is provided on the substrate **10**. The channel layer **20** may be deposited on the substrate **10** using buffer layers, transition layers, and/or nucleation layers as described above. The channel layer **20** may be under compressive strain. Furthermore, the channel layer **20** and/or buffer, nucleation, and/or transition layers may be deposited by metal-organic chemical vapor deposition (MOCVD) or by other techniques known to those of skill in the art, such as molecular beam epitaxy (MBE) and/or hydride vapor phase epitaxy (HVPE). In some embodiments of the present invention, the channel layer **20** may be a Group III-nitride layer, such as GaN. The channel layer **20** may also include other Group III-nitride layers, such as indium gallium nitride (InGaN), aluminum indium gallium nitride (AlInGaN), or the like. The channel layer **20** may be undoped (*i.e.*, "unintentionally doped"), and may be grown to a thickness of greater than about 20 Å. The channel layer **20** may also be a multi-layer structure, such as a superlattice or combinations of GaN, AlGaN, or the like.

[0077] As further shown in **Figure 1**, a barrier layer **22** is provided on the channel layer **20**. For example, the barrier layer **22** may be deposited on the channel layer **20**. The barrier layer **22** may be a Group III-nitride layer, such as $Al_xGa_{1-x}N$ (where $0 < x < 1$). The barrier layer **22** may also include other Group III-nitride layers, such as AlInGaN, AlN, and/or combinations of layers thereof. The barrier layer **22** may, for example, be from about 0.1 nm to about 100 nm thick, but may not be so thick as to cause cracking or substantial defect formation therein. In certain embodiments of the present invention, the barrier layer **22** may be a highly-doped n-type layer. For example, the barrier layer **22** may be doped to a concentration of less than about 10^{19} cm^{-3}.

[0078] Some embodiments of the present invention may be applicable in high electron mobility transistors (HEMTs). More particularly, the channel layer **20** and the barrier layer **22** may be formed of materials having different bandgaps, such that

an interface between the channel layer and the barrier layer defines a heterojunction. For example, the channel layer **20** may have a bandgap that is less than the bandgap of the barrier layer **22**. As such, the energy of the conduction band edge of the channel layer **20** may be less than the energy of the conduction band edge of the barrier layer **22** at the junction between the channel **20** and barrier **22** layers, and the channel layer **20** may have a greater electron affinity than the barrier layer **22**. For example, where both the channel layer **20** and the barrier layer **22** are formed of Group III-nitride layers, the channel layer **20** may be a GaN layer, and the barrier layer **22** may be a AlGaN layer.

[0079] Examples of layers according to certain embodiments of the present invention are described in United States Patent Publication No. 2002/0167023A1, to Smorchkova *et al.*, entitled "GROUP-III NITRIDE BASED HIGH ELECTRON MOBILITY TRANSISTOR (HEMT) WITH BARRIER/SPACER LAYER" the disclosure of which is incorporated herein by reference as if set forth fully herein. In particular embodiments of the present invention, the barrier layer **22** may have a thickness, Al composition, and/or doping sufficient to induce a significant carrier concentration at the interface between the channel layer **20** and the barrier layer **22** through polarization effects when the barrier layer **22** is buried under ohmic contact metal. Also, the barrier layer **22** may be thick enough to reduce or minimize scattering of electrons in the channel due to ionized impurities deposited at the interface between the barrier layer **22** and a subsequently formed protective layer.

[0080] In addition, in other embodiments of the present invention, the channel layer **20** and the barrier layer **22** may have different lattice constants. For example, the barrier layer **22** may be a relatively thin layer having a smaller lattice constant than the channel layer **20**, such that the barrier layer **22** "stretches" at the interface between the two. Accordingly, a pseudomorphic HEMT (pHEMT) device may be provided.

[0081] **Figure 2** illustrates formation of a protective layer **24** on the barrier layer **22**. The protective layer **24** may be dielectric material, such as silicon nitride (Si_xN_y), aluminum nitride (AlN), silicon dioxide (SiO_2), and/or other suitable protective material. Other materials may also be utilized for the protective layer **24**. For example, the protective layer **24** may also include magnesium oxide, scandium oxide, aluminum oxide and/or aluminum oxynitride. Furthermore, the protective

layer **24** may be a single layer or may include multiple layers of uniform and/or non-uniform composition.

[0082] The protective layer **24** may be blanket formed on the barrier layer **22**. For example, the protective layer **24** may be a silicon nitride (SiN) layer formed by high quality sputtering and/or plasma-enhanced chemical vapor deposition (PECVD). The protective layer **24** may have a thickness of about 30 nm, however, other thickness layers may also be utilized. For example, the protective layer may be sufficiently thick so as to protect the underlying layer during a subsequent anneal of ohmic contacts. Layers as thin as two or three monolayers may be sufficient for such purposes. However, in general, the protective layer **24** may have a thickness of from about 10 nm to about 500 nm. Also, a high quality SiN protective layer may be grown in-situ with the MOCVD growth of the group III nitride layers.

[0083] In particular embodiments of the present invention, the protective layer **24** may be SiN. The SiN protective layer may be formed by PVD and/or CVD and may be non-stoichiometric in compressive or tensile strain. For example, the SiN protective layer may have a stress of between about -100 MPa and about 100 MPa. In certain embodiments of the present invention, the SiN protective layer may have an index of refraction at a 633 nm wavelength of from about 1.6 to about 2.2. In particular embodiments, the index of refraction of the SiN protective layer may be 1.98 ± 0.05.

[0084] **Figure 3** illustrates formation of ohmic contact regions **30** on the barrier layer **22**. Referring now to **Figure 3**, the protective layer **24** is patterned to expose portions of the barrier layer **22**, and first and second ohmic contact regions **30** are formed on the barrier layer **22**. For example, windows may be etched into the protective layer **24** to expose the underlying barrier layer **22**. The windows may be etched utilizing a patterned mask and a low-damage etch with respect to the barrier layer **22**. Examples of low damage etch techniques may include etching techniques other than reactive ion etching, such as inductively coupled plasma or electron cyclotron resonance (ECR) or downstream plasma etching with no DC component to the plasma. For example, for a SiO_2 barrier layer, a wet etch with buffered hydrofluoric acid may be used. A selective etch of SiN and/or SiO_2 to an etch stop layer, such as ITO, SCO, MgO or the like, followed by a low damage removal of the etch stop layer may also be performed. For a SiN barrier layer, SiO_2 may be used as an etch stop layer. In such embodiments, the protective layer **24** may include the SiN,

AlN, and/or SiO_2 layers, as well as the etch stop layer. Thus, in certain embodiments of the present invention, the protective layer **24** may include multiple layers.

[0085] Still referring to **Figure 3**, using a subsequent photolithography step and evaporation, ohmic metal is formed on the exposed portions of the barrier layer **22**. The ohmic metal is patterned so as to be smaller than the window in the protective layer **24**, and the ohmic metal is annealed to provide the first and second ohmic contact regions **30**. As such, the edges of the ohmic contact regions **30** may be spaced apart from the adjacent protective layer **24**.

[0086] The anneal may be a relatively high temperature anneal. For example, the anneal may be performed at a temperature of greater than about 900 °C. By using such an ohmic contact anneal, the resistance of the ohmic contact regions **30** may be reduced from a relatively high resistance to, for example, less than about 1 Ω-mm. Thus, as used herein, the term "ohmic contact" may refer to a non-rectifying contact that has a contact resistance of less than about 1 Ω-mm. The presence of the protective layer **24** during the high temperature process steps may reduce and/or inhibit damage to the barrier layer **22** that may be caused by such steps. Thus, for example, the sheet resistance of the gate region **21** (*i.e.*, the length of the channel between the ohmic contact regions **30**) after the high temperature ohmic contact anneal may be substantially the same as the sheet resistance of the gate region **21** as-grown (*i.e.*, before the contact anneal).

[0087] The ohmic contact regions **30** may be spaced apart from the protective layer **24** by a distance that is sufficient to allow for misalignment tolerances in the formation and/or patterning of the ohmic contact metal. For example, the edges of the ohmic contact regions **30** may be spaced apart from the protective layer **24** by a distance of about 0.1 micrometer (μm) to about 0.2 μm. It may be desirable to prevent contact between the ohmic metal and the protective layer **24** to reduce the likelihood of the metal diffusing into the protective layer **24** during subsequent heating steps, which may result in a short between a gate contact and the ohmic contact region(s) **30**. However, the gaps between the ohmic contact regions **30** and the protective layer **24** need not be so large as to defeat the protective purpose of the protective layer **24** (and/or substantially degrade the performance of the device), but may not be so small as to risk contact between the ohmic material and the protective layer **24**. Thus, in certain embodiments of the present invention, the gaps may be in the range of about 0.1 micrometer (μm) to about 0.5 μm.

[0088] **Figure 4** illustrates the formation of a second layer **31** on the ohmic contact regions **30** and the protective layer **24**. The second layer **31** may be a dielectric layer having a lower dielectric index than the protective layer **24**. For instance, the second layer **31** may be a high-quality oxide layer, such as a high-temperature oxide (HTO) layer. The second layer **31** may also be a sacrificial layer, which may be removed in a subsequent step. For example, the second layer **31** may be a photoresist layer. Also, the second layer **31** may be a polymide layer. As shown in **Figure 4**, the second layer **31** includes a hole or recess **3** that extends through the second layer **31** and exposes a portion of the protective layer **24**. For example, the second layer **31** may be photolithographically patterned to form the recess **3**.

[0089] **Figure 5** illustrates the formation of an opening or "window" in the protective layer **24**, where the gate electrode may be formed in a subsequent step. As shown in **Figure 5**, an opening **4** is formed extending through the protective layer **24** to expose a portion of the barrier layer **22**. More particularly, the exposed portion of the protective layer **24** is patterned using the second layer **31** as a mask to form the opening **4**. The opening **4** may be formed by selectively etching the portion of the protective layer **24** exposed by the recess **3** in the second layer **31** using a low damage etch technique, as described above, to facilitate the formation of a low-leakage Schottky gate contact on the exposed surface of the barrier layer **22**. Note that the opening **4** in the protective layer **24** may be wider than the recess **3** in the second layer **31**, due to isotropy (*i.e.*, undercutting) of the etch. In particular embodiments, the ohmic contacts **30** may provide source and drain contacts, and the opening **4** may be offset between the source and drain contacts such that the opening **4**, and subsequently the gate contact, may be closer to the source contact than the drain contact.

[0090] Referring now to **Figure 6**, the recess **3** in the second layer **31** is widened to define a second opening **5** that is wider than the first opening **4**. As such, the second opening **5** exposes the first opening **4** as well as surface portions of the protective layer **24** on opposite sides of the first opening **4**. The recess **3** in the second layer **31** may be symmetrically expanded so that the second opening **5** and the first opening **4** are self-aligned. For example, where the second layer **31** is formed of a different material than the protective layer **24**, the recess **3** in the second layer **31** may be widened by selectively etching the second layer **31** using a wet etching technique. More particularly, where the second layer **31** is an oxide layer, the recess **3** may be

widened by selectively etching the second layer **31** using a buffered oxide etchant (BOE). In addition or in an alternative, where the second layer **31** is a photoresist layer, the recess may be widened by ashing the second layer **31** using an oxygen plasma and/or hard baking the device, for example, at a temperature above 100 degrees Celsius (C) but below 150 degrees C.

[0091] Accordingly, as shown in **Figures 5** and **6**, the protective layer **24** may be selectively etched using the second layer **31** as a mask to form the first opening **4**, and the recess **3** in the second layer **31** may be selectively widened around the first opening **4** to form the second opening **5**. As such, the first opening **4** and the second opening **5** may be self-aligned. In some embodiments, the second layer **31** may be any material which may be selectively etched relative to the protective layer **24**.

[0092] **Figure 7** illustrates the formation of a gate electrode **32**. As shown in **Figure 7**, the gate electrode **32** is formed in the second opening **5** and the first opening **4**, and extends through the protective layer **24** to contact the exposed portion of the barrier layer **22**. The gate electrode **32** may be formed in the first opening **4** directly on opposing sidewalls of the protective layer **24**, such that a gap may not be formed between the two. Suitable gate materials may depend on the composition of the barrier layer **24**. However, in certain embodiments, conventional materials capable of making a Schottky contact to a nitride based semiconductor material may be used, such as Ni, Pt, $NiSi_x$, Cu, Pd, Cr, TaN, W and/or WSiN.

[0093] Still referring to **Figure 7**, first portions **6** (also referred to as "wings" or "sidelobes") of the gate electrode **32** laterally extend on surface portions of the protective layer **24** outside the first opening **4**, and second portions 7 of the gate electrode **32** that are spaced vertically apart from the protective layer **24** laterally extend on surface portions of the second layer **31** outside the second opening **5**. As used herein, the term "laterally" refers to a direction that is substantially parallel with respect to a surface of the substrate. The second portions **7** of the gate electrode **32** may laterally extend beyond the first portions **6**. Because the first opening **4** and the second opening **5** may be self-aligned by the processes described thus far, the length by which the first portions **6** of the gate electrode **32** extend onto the protective layer **24** may be controlled. As such, according to some embodiments of the present invention, gate-to-drain capacitance (c_{gd}) and/or gate-to-source capacitance (c_{gs}) of the transistor device that may be caused by the extension of the gate electrode **32** onto the

protective layer **24** may also be controlled. Accordingly, a modified gamma gate **32** having self-aligned sidelobes **6** may be formed.

[0094] **Figure 8** illustrates removal of the sacrificial second layer **31**. As shown in **Figure 8**, the second layer **31** is removed after the gate electrode **32** is formed. The second layer **31** may be formed using conventional techniques, depending on the composition of the second layer **31**.

[0095] **Figure 9** illustrates the formation of a passivation layer **34**. The passivation layer **34** may provide environmental protection for the transistor device. For example, the passivation layer **34** may be a thicker nitride or relatively low-k encapsulation layer, and may be formed to improve coverage of the protective layer **24**. The passivation layer **34** may be blanket deposited on the structure of **Figure 8** after removal of the second layer **31**. As such, the second portions **7** of the gate electrode **32** laterally extend on portions of the passivation layer **34**. In particular embodiments, the passivation layer **34** may be deposited so as to substantially fill the gaps between the protective layer **24** and the ohmic contacts **30**. In certain embodiments of the present invention, the passivation layer **34** may be silicon nitride, aluminum nitride, silicon dioxide, and/or an oxynitride. Furthermore, the passivation layer **34** may be a single or multiple layers of uniform and/or non-uniform composition.

[0096] **Figure 10** is a cross-sectional view illustrating a high electron mobility transistor (HEMT) device **100** according to further embodiments of the present invention. As shown in Figure **10**, the gate electrode **32** of the HEMT device **100** extends through the protective layer **24** to contact the barrier layer **22**. The gate electrode **32** includes first portions **6** laterally extending on surface portions of the protective layer **24** outside the opening therein, and second portions **7** spaced vertically apart from the protective layer **24** by the passivation layer **34** and laterally extending on portions of the passivation layer **34** beyond the first portions **6**. In other words, the first and second portions **6** and **7** of the gate electrode **32** define a stair-step profile. The second portions **7** of the gate electrode **32** may be spaced apart from the protective layer **24** by about 1000 Angstroms (Å) to about 6000 Å. The gate electrode **32** also extends directly on opposing sidewalls of the opening in the protective layer **24**, such that a gap may not be present between the gate electrode **32** and the protective layer **24**. The first portions **6** of the gate electrode **32** may also be formed directly on the protective layer **24**.

[0097] In addition, because the first opening (in the protective layer **24**) and the second opening (in the second layer **31** of **Figure 6**) may be self-aligned by the processes described above, a width of the second opening (illustrated as L_{G2}) may be self-aligned and symmetric around a width of the first opening, which defines the gate length L_{G1}. As such, the gate length L_{G1} may be tuned for an operational frequency of interest based on formation of the opening in the protective layer **24**. For example, the gate length L_{G1} may be about 0.05 micrometer (μm) to about 2.0 μm. In addition, a length ΔL_G by which the first portion **6** of the gate electrode **32** extends onto the protective layer **24** may also be controlled, by widening the opening in the sacrificial second layer, as described above. The thickness d_1 of the protective layer **24** and the length ΔL_G of first portion of the gate electrode **32** may be optimized to reliably support the peak gate-to-drain voltage V_{GD} and to minimize and/or reduce the gate-to-drain capacitance c_{gd} and the gate-to-source capacitance c_{gs}. For example, the length ΔL_G may be about 0.1 micrometer (μm) to about 0.6 μm, and the thickness d_1 may be about 300 Angstroms (Å) to about 5000 Å. In addition, the angle Ø of the sidewalls of the protective layer **24** defining the first opening may also be optimized, for example, between about 45 degrees to about 90 degrees.

[0098] As described above, in some embodiments, the protective layer **24** may be SiN, the barrier layer **22** may be AlGaN, and the channel layer **20** may be GaN. The transistor **100** also includes a nucleation layer **15** between the substrate **10** and the channel layer **20**, as also discussed above.

[0099] Also, **Figure 10** illustrates that the ohmic contact regions may be source/drain contact regions **30**. The protective layer **24** may be formed to a thickness that is at least about as thick as the ohmic contact regions. For example, the protective layer **24** may be formed to a thickness in the range of about 500 Angstroms (Å) to about 5000 Å. In addition, the passivation layer **34** may substantially fill the gap between the protective layer **24** and the source/drain regions **30**. In certain embodiments of the present invention, the passivation layer **34** may be formed of the same material as the protective layer **24**. As such, the passivation layer **34** may be silicon nitride, aluminum nitride, silicon dioxide and/or an oxynitride. Furthermore, the passivation layer **34** may be a single or multiple layers of uniform and/or non-uniform composition.

[0100] In other embodiments of the present invention, the second layer **31** may not be a sacrificial layer. As such, second layer **31** may not be removed, and the

passivation layer **34** may not be formed. Thus, **Figure 7** illustrates a high electron mobility transistor (HEMT) device according to other embodiments of the present invention. For example, as noted above, the second layer **31** may be a dielectric layer. More particularly, the second layer **31** may be an oxide layer, such as a HTO layer and/or other high-quality oxide layer, and the protective layer **24** may be a nitride layer, such as a high-purity nitride (HPN) layer, and/or another protective layer that may be selectively etched with respect to the oxide layer. Such an oxide layer may provide a lower impurity concentration and/or other improved electrical characteristics in comparison to the passivation layer **34**. Accordingly, further embodiments of the present invention will be described with reference to **Figures 11** to **18**.

[0101] **Figures 11** to **18** are cross-sectional views illustrating examples of intermediate fabrication steps in methods for fabricating transistor devices according to further embodiments of the present invention. Referring now to **Figure 11**, a barrier layer **22** and a channel layer **20** may be formed a substrate **10** on which a transistor device may be formed, as described above with reference to **Figure 1**. For example, the substrate **10** may be a semi-insulating silicon carbide (SiC) substrate. Optional buffer, nucleation and/or transition layers (not shown) may also be provided on the substrate **10**, for example, to provide an appropriate crystal structure transition between a silicon carbide substrate and the remainder of the device. The channel layer **20** may be single or multi-layer Group III-nitride layer, such as a superlattice. The barrier layer **22** may be a Group III-nitride layer, such as $Al_xGa_{1-x}N$ (where $0 < x < 1$). The barrier layer **22** may also include other Group III-nitride layers, such as AlInGaN, AlN, and/or combinations of layers thereof.

[0102] The channel layer **20** and the barrier layer **22** may be formed of materials having different bandgaps, such that an interface between the channel layer and the barrier layer defines a heterojunction. More particularly, the channel layer **20** may have a bandgap that is less than the bandgap of the barrier layer **22**. For example, where both the channel layer **20** and the barrier layer **22** are formed of Group III-nitride layers, the channel layer **20** may be a GaN layer, and the barrier layer **22** may be a AlGaN layer. In addition, the channel layer **20** and the barrier layer **22** may have different lattice constants. For example, the barrier layer **22** may be a relatively thin layer having a smaller lattice constant than the channel layer **20**, such

that the barrier layer **22** "stretches" at the interface between the two. Accordingly, a pseudomorphic HEMT (pHEMT) device may be provided.

[0103] Still referring to **Figure 11**, a protective layer **24** is formed on the barrier layer **22**. The protective layer **24** may be dielectric material, such as silicon nitride (Si_xN_y), aluminum nitride (AlN), silicon dioxide (SiO_2), and/or other suitable protective material. The protective layer **24** may be a single layer or may include multiple layers of uniform and/or non-uniform composition. Other materials may also be utilized for the protective layer **24**.

[0104] In some embodiments of the present invention, the protective layer **24** may be a high purity nitride (HPN) layer. As described herein, a HPN layer may refer to a nitride layer having significantly lower overall impurity levels as compared to some other nitride layers. More particularly, the impurity levels of a HPN layer may generally be one to three orders of magnitude smaller than the impurity levels in nitride layers grown by PECVD and/or by sputtering. For example, HPN may refer to nearly stoichiometric silicon nitride (*i.e.*, Si_3N_4) having relatively low impurity concentrations in the Si_3N_4 layer itself as well as at the interface between the Si_3N_4 layer and the underlying layer. This may be achieved by depositing the HPN layer in the same reactor as the underlying layer and using the conditions described further herein.

[0105] For example, the protective layer **24** may include a high purity SiN layer as described in U.S. Patent Application Serial No. 11/286,805 filed on November 23, 2005 entitled "GROUP III NITRIDE SEMICONDUCTOR DEVICES WITH SILICON NITRIDE LAYERS AND METHODS OF MANUFACTURING SUCH DEVICES," the disclosure of which is incorporated herein by reference as if fully set forth herein. The SiN layer may be formed, for example, by Low Pressure Chemical Vapor Deposition (LPCVD) and/or Metal-Organic Chemical Vapor Deposition (MOCVD). In particular, pursuant to certain embodiments of the present invention, a SiN protective layer **24** may be grown at relatively high temperatures (*e.g.,* above about 700 °C) *in situ* with the MOCVD growth of the Group III nitride layers. For example, where the channel layer **20** is GaN, the protective layer **24** may be grown in the GaN epitaxial reactor without cooling the substrate **10** from the growth temperatures of the Group III nitride layers. In particular embodiments, the SiN layers may be grown at a temperature in the range of about 900 °C to about 1000

°C. Such high-temperature growth may also facilitate reducing the impurity levels in the SiN layer and at the interface between a Group III nitride layer and the SiN layer. Additionally, high growth rates may be employed which may facilitate reducing the levels of background reactor impurities incorporated into the SiN layer.

[0106] Forming the SiN layer *in situ* may also reduce the levels of impurities that are incorporated into the top surface of the uppermost Group III nitride layer and/or into the SiN layer itself. In particular, when the device is removed from the reactor and the SiN layer is formed via a post-MOCVD growth process such as, for example, sputtering or PECVD, a number of different mechanisms may introduce impurities. For example, as discussed in detail in U.S. Patent No. 6,498,111, if hydrogen is present in an MOCVD reactor during the growth of a Group III nitride layer, the hydrogen may tend to incorporate into the Group III nitride layer during cooling of the reactor following growth. Likewise, exposure of the device to the atmosphere upon removal from the reactor may allow for incorporation of oxygen atoms. In addition, various other impurities may be introduced, particularly adjacent the outer surfaces of the device, as a result of handling of the device and/or chemical cleaning of the device. Impurities may also be added if post-growth processing such as wet etching, electrode deposition, annealing steps, etc. are performed prior to deposition of the SiN protective/passivation layer. These impurities may change the surface states at the interface between the Group III nitride layer and the SiN layer in ways that may be undesirable and/or difficult to control/reproduce. For example, the presence of impurities can increase trapping at the interface between the SiN layer and the underlying Group III nitride layer, thereby increasing the sheet resistance of the channel.

[0107] In certain embodiments of the present invention, the silicon nitride layer may be formed using a high purity silicon gas (*e.g.*, 99.99% pure) source. For example, high purity silane (SiH_4) may be used as a source gas in the growth of the SiN layer(s). As is known to persons of skill in the art, silane is often used as a source for silicon dopants in the growth of n-doped Group III nitride layers. Typically, diluted silane gas is used in such applications as it is less expensive and easier to use than pure silane, which may be highly combustible. The use of such pure silane may facilitate reducing the level of impurities at, for example, the interface between the Group III nitride layer and the SiN layer and/or within the SiN layer, which may, in

certain circumstances, improve the performance and/or reproducibility of the device. In particular, a higher-quality (*i.e.*, more pure) SiN layer may help reduce or minimize trapping within the body of the insulative layer, thereby providing a higher breakdown critical field. Pure disilane (Si_2H_6) or other pure silicon gas sources may also be used instead of high purity silane (SiH_4). When such a pure silane gas source is included with the reactor, it may still be desirable to include a diluted silane source as well, so that diluted silane gas may be used as the dopant gas source during the growth of, for example, an n-doped or co-doped Group III nitride layer.

[0108] In some embodiments, the high purity SiN layer **24** may be stoichiometric. The term "stoichiometry" may refer to the molar proportions of elements in stoichiometric compounds. In such stoichiometric compounds, the molar proportions are whole numbers. For example, the stoichiometry of silicon and nitrogen in Si_3N_4 is about 3:4. The stoichiometry of a SiN layer may also affect the index of refraction of the layer. In certain embodiments of the present invention, a SiN protective layer **24** may have an index of refraction at a 633 nm wavelength of from about 1.6 to about 2.2. Stoichiometric SiN may also be characterized by a relatively low etch rate in a buffered oxide etch (BOE). For example, the etch rate of stoichiometric SiN in BOE may be less than about 10 Angstroms per minute. In contrast, PECVD silicon nitride may have a BOE etch rate on the order of about 100 Angstroms per minute, and sputtered silicon nitride may have a BOE etch rate on the order of about 1000 Angstroms or more per minute. The stoichiometry of a SiN layer may be adjusted, for example, by adjusting the relative flow rates of SiH_4 and NH_3 source gases in a CVD process. Moreover, when formed at relatively high temperatures, CVD-grown SiN tends to be stoichiometric.

[0109] In some embodiments of the present invention, the protective layer **24** may include a HPN layer that is sufficiently thick so as to reduce and/or inhibit damage to the barrier layer **22** during subsequent high temperature process steps. For instance, the HPN layer may have a thickness of from about 200 Angstroms (Å) to about 2000 Å. In addition, the protective layer **24** may include multiple layers of HPN and silicon dioxide (SiO_2) and/or plasma enhanced chemical vapor deposition (PECVD) silicon nitride (SiN). It will be understood that the terms "Si_xN_y," "SiN" and "silicon nitride" are used herein interchangeably to refer to both stoichiometric and non-stoichiometric silicon nitride. However, embodiments of the present

invention are not limited by these examples, for example, any HPN can be used as the high quality dielectric protective layer **24**. More generally, the material of the protective layer **24** should have appropriate interface characteristics, should be capable of withstanding relatively high temperatures, and should be capable of being removed without significantly damaging the underlying barrier layer **22.**

[0110] **Figure 12** illustrates formation of a dielectric layer, such as an oxide layer **33**, on the protective layer **24**. For example, the oxide layer **33** may be deposited on the protective layer **24** using a chemical vapor deposition (CVD) process at a temperature of greater than about 500 °C. In addition and/or alternatively, the oxide layer **33** may be formed *in situ*. The oxide layer **33** may be formed to a thickness that is greater than the thickness of the protective layer **24**. For example, the oxide layer **33** may have a thickness of about 300 Angstroms (Å) to about 5000 Å.

[0111] The oxide layer **33** may also have a lower dielectric index than the protective layer **24**. For instance, the oxide layer **33** may be a high-quality oxide layer, such as a high-temperature oxide (HTO) layer. The high-quality oxide layer **33** may have a dielectric constant or dielectric index of less than about 1.5. In addition, the high-quality oxide layer **33** may have a relatively low impurity concentration, for example, as may be measured by a secondary ion mass spectroscopy (SIMS) analysis. As such, a high-quality oxide layer according to some embodiments of the present invention may have improved electrical characteristics as compared to a layer formed by a PECVD process, such as the post-gate passivation layer **34** of **Figure 10.**

[0112] For example, in some embodiments, the HTO layer may be a silicon dioxide (SiO_2) layer deposited by a LPCVD process at a temperature of about 900 °C. The LPCVD nature of the deposition may allow the oxide layer **33** to be formed while reducing and/or minimizing negative effects on the underlying semiconductor layer(s). In addition, the relatively high temperature of deposition may allow for the formation of a relatively dense oxide layer **33**, with measured dielectric strength of about 8-9 MV/cm (*i.e.*, about 80-90% of the theoretical value for SiO_2).

[0113] In certain embodiments of the present invention, the oxide layer **33** may be stoichiometric SiO_2. The SiO_2 layer may have an index of refraction at a 633 nm wavelength of from about 1.36 to about 1.56. In particular embodiments, the index of refraction of an SiO_2 protective layer may be 1.46 ± 0.03 as measured by ellipsometry. In addition, the SiO_2 layer may provide reduced levels of charges, traps,

and other electrically active non-idealities (impurities) therein, which may, in certain circumstances, improve the performance and/or reproducibility of the device. In particular, a high-quality (*i.e.*, more pure) SiO_2 layer may help reduce or minimize trapping within the body of the insulative layer, thereby providing a higher breakdown critical field. More generally, the oxide layer **33** may be selected to provide a thickness, impurity concentration, dielectric index, and/or other characteristics that may be sufficient to reduce gate-to-drain and/or gate-to-source capacitance in a HEMT device.

[0114] Referring now to **Figure 13**, the oxide layer **33** and the protective layer **24** are patterned to expose portions of the barrier layer **22**. For example, utilizing a patterned mask (not shown), the oxide layer **33** may be patterned to define recesses **9** exposing first and second portions of the protective layer **24**, and the protective layer **24** may be patterned using the oxide layer **33** as a mask to define windows **8** exposing portions of the barrier layer **22**. More particularly, a dry etching process may be used to define the recesses **9** in the oxide layer **33**, and a low-damage wet etching process may be used to selectively etch the protective layer **24** to define the windows **8**. A selective etch of the protective layer **24** to an etch stop layer, such as ITO, SCO, MgO or the like, followed by a low damage removal of the etch stop layer may also be performed. In such embodiments, the protective layer **24** may include SiN, AlN, and/or SiO_2 layers, as well as the etch stop layer. Alternatively, with both layers **33** and **24** deposited, a low-damage dry etch may be used to etch through both the oxide layer **33** and the protective layer **24** in a single step. Examples of such low damage etch techniques may include etching techniques other than reactive ion etching, such as inductively coupled plasma or electron cyclotron resonance (ECR) or downstream plasma etching with no DC component to the plasma.

[0115] Also, in some embodiments, the protective layer **24** may be patterned to define the windows **8** exposing portions of the barrier layer **22** before forming the oxide layer **33**. As such, the oxide layer may be formed on the protective layer **24** and on the first and second portions of the barrier layer **22**. The oxide layer may then be patterned to define recesses **9** exposing portions of the barrier layer **22**. The recesses **9** may be narrower than the windows **8** etched into the protective layer **24**. As such, portions of the oxide layer **33** may remain along the sidewalls of the windows **8** in the protective layer **24**. Accordingly, when ohmic contact regions are formed on the

barrier layer 22 in a subsequent step, the edges of the ohmic contact regions may be spaced apart from the adjacent protective layer 24 due to the presence of the oxide layer 33 therebetween.

[0116] **Figure 14** illustrates formation of ohmic contacts **30** on the barrier layer **24**. Referring now to **Figure 14**, metal may be deposited on the portions of the barrier layer **22** exposed by the windows **8** and the recesses **9**, for example, by evaporation, to provide ohmic metal regions on the barrier layer **22**. Suitable metals may include refractory metals, such as titanium (Ti), tungsten (W), titanium tungsten (TiW), silicon (Si), titanium tungsten nitride (TiWN), tungsten silicide (WSi), rhenium (Re), Niobium (Nb), nickel (Ni), gold (Au), aluminum (Al), tantalum (Ta), molybdenum (Mo), nickel silicide (NiSi), titanium silicide (TiSi), titanium nitride (TiN), tungsten silicon nitride (WSiN), platinum (Pt) and the like. Any unwanted metal may be removed using, for example, a solvent. The ohmic metal regions may be patterned to be smaller than the window in the protective layer **24**, and may be annealed to provide the first and second ohmic contact regions **30**. As such, the ohmic contact regions **30** may be spaced apart from the protective layer **24** by a distance that is sufficient to allow for misalignment tolerances in the formation and/or patterning of the ohmic contact metal. For example, the edges of the ohmic contact regions **30** may be spaced apart from the protective layer **24** by a distance of about 0.1 micrometer (μm) to about 0.5 μm.

[0117] Still referring to **Figure 14**, the anneal may be a relatively high temperature anneal. For example, the anneal may be performed at a temperature of greater than about 900 °C. By using such an ohmic contact anneal, the resistance of the ohmic contact regions **30** may be reduced from a relatively high resistance to, for example, less than about 1 Ω-mm. Thus, as used herein, the term "ohmic contact" may refer to a non-rectifying contact that has a contact resistance of less than about 1 Ω-mm. The presence of the protective layer **24** during the high temperature process steps may reduce and/or inhibit damage to the barrier layer **22** that may be caused by such steps. Thus, for example, the sheet resistance of the gate region **21** (*i.e.*, the length of the channel between the ohmic contact regions **30**) after the high temperature ohmic contact anneal may be substantially the same as the sheet resistance of the gate region **21** as-grown (*i.e.* before the contact anneal).

[0118] It will be appreciated that, in some embodiments, it may not be necessary to anneal the deposited metal in order to form the ohmic contacts **30**. For

example, where the ohmic contacts **30** are formed on implanted source/drain regions in the barrier layer **22**, the metal may be ohmic as-deposited. Since a contact anneal may not be required, it may be acceptable for the metal of the ohmic contacts **30** to contact the protective layer **24** . Thus, in some embodiments of the invention, a lithography step that may otherwise be required to ensure that the ohmic contacts **30** are spaced away from the protective layer **24** may be avoided as a result of the implantation of impurities into the barrier layer **22**. In addition, since the ohmic contacts **30** may be formed on the implanted regions, the ohmic contacts may have a lower resistivity than ohmic contacts formed on non-implanted regions. Thus, the on-resistance of devices formed according to some embodiments of the invention may be lowered. Moreover, some embodiments of the present invention may employ semiconductor fabrication methods as discussed in currently commonly assigned United States Patent Application Serial No. 11/302,062 (Attorney Docket No. 5308-619) entitled SEMICONDUCTOR DEVICES INCLUDING IMPLANTED REGIONS AND PROTECTIVE LAYERS AND METHODS OF FORMING THE SAME, and United States Patent Application Serial No. 11/434,853 (Attorney Docket No. 5308-635) entitled SEMICONDUCTOR DEVICES INCLUDING SELF ALIGNED REFRACTORY CONTACTS AND METHODS OF FABRICATING THE SAME, the disclosures of which are incorporated by reference herein in their entireties.

[0119] **Figure 15** illustrates the formation of a hole or recess **3** in the oxide layer **33**. More particularly, a mask (not shown) is formed on the ohmic contacts **30** and the oxide layer **33**, and the oxide layer **33** is patterned to form the recess **3**. The oxide layer **33** may be selectively patterned, for example, by one of the etch techniques described above. The recess **3** extends through the oxide layer **33** and exposes a portion of the protective layer **24**.

[0120] **Figure 16** illustrates the formation of an opening or "window" in the protective layer **24** where the gate electrode may be formed in a subsequent step. As shown in **Figure 16**, an opening **4** is formed extending through the protective layer **24** to expose a portion of the barrier layer **22**. More particularly, the exposed portion of the protective layer **24** is patterned using the oxide layer **33** as a mask to form the opening **4**. The opening **4** may be formed by selectively etching the portion of the protective layer **24** exposed by the recess **3** in the oxide layer **33** using a low damage etch technique, as described above, to thereby facilitate the formation of a low-

leakage Schottky gate contact on the exposed surface of the barrier layer **22**. Note that the opening **4** in the protective layer **24** may be wider than the recess **3** in the oxide layer **33**, due to isotropy (*i.e.*, undercutting) of the etch. In particular embodiments, the ohmic contacts **30** may provide source and drain contacts, and the opening **4** may be offset between the source and drain contacts such that the opening **4**, and subsequently the gate contact, may be closer to the source contact than the drain contact.

[0121] It is to be understood that, in some embodiments, a low-damage dry etch may be used to etch through both the oxide layer **33** and the protective layer **24** in a single step to define the recess **3** in the oxide layer **33** and the opening **4** in the protective layer **24** to expose the barrier layer **22**. In addition, in some embodiments, the oxide layer **33** and the protective layer **24** may be patterned to define the recesses **3** and **9** and the openings **4** and **8** using a single mask. More generally, although illustrated and described with reference to particular intermediate fabrication steps, some steps may be combined, divided, and/or eliminated in particular embodiments of the present invention.

[0122] Referring now to **Figure 17**, the recess **3** in the oxide layer **33** is widened to define a second opening **5** that is wider than the first opening **4**. As such, the second opening **5** exposes the first opening **4** as well as surface portions of the protective layer **24** on opposite sides of the first opening **4**. The recess **3** in the oxide layer **33** may be symmetrically expanded so that the second opening **5** and the first opening **4** are self-aligned. For example, the recess **3** in the oxide layer **33** may be widened by selectively etching the oxide layer **33** using a buffered oxide etchant (BOE). In addition, in some embodiments, the recesses **9** in the oxide layer **33** used in the formation of the ohmic contacts **30** may also be symmetrically expanded by the buffered oxide etchant (BOE).

[0123] Accordingly, as shown in **Figures 16** and **17**, the protective layer **24** may be selectively etched using the oxide layer **33** as a mask to form the first opening **4**, and the recess **3** in the oxide layer **33** may be selectively widened around the first opening **4** to form the second opening **5**. As such, the first opening **4** and the second opening **5** may be self-aligned. In some embodiments, the oxide layer **33** may be formed of another high-quality dielectric material which may be selectively etched relative to the protective layer **24**.

[0124] **Figure 18** illustrates the formation of a gate electrode **32** to provide a high electron mobility transistor (HEMT) device **200** according to further embodiments of the present invention. As shown in **Figure 18**, the ohmic contact regions may be source/drain contact regions **30**. Also, the gate electrode **32** is formed in the second opening **5** and the first opening **4**, and extends through the protective layer **24** to contact the exposed portion of the barrier layer **22**. The gate electrode **32** may be formed in the first opening **4** directly on opposing sidewalls of the protective layer **24**, such that a gap may not be formed between the two. In addition and/or alternatively, a dielectric liner (not shown), such as a high-quality silicon dioxide liner or a nitride liner, may be formed in the first opening on opposing sidewalls of the protective layer **24**. The gate electrode **32** may be formed directly on the dielectric liner to provide a metal-insulator-semiconductor (MIS) HEMT device.

[0125] Suitable gate materials may depend on the composition of the barrier layer **24**. However, in certain embodiments, conventional materials capable of making a Schottky contact to a nitride based semiconductor material may be used, such as Ni, Pt, $NiSi_x$, Cu, Pd, Cr, TaN, W and/or WSiN. In addition, in some embodiments, the gate material may also be deposited on the source/drain contact regions **30** during the formation of the gate electrode **32**. More particularly, the gate material may be deposited in the first opening **4** in the protective layer **24**, and in the second opening **5** and the recesses **9** in the oxide layer **33**. The gate material may then be patterned to define the gate electrode **32** such that portions thereof remain on the source/drain contact regions **30** in the recesses **9**. Forming the gate material (or other conductive layer) on the source/drain contact regions **30** may facilitate contact to the relatively small source/drain contact regions **30**, for example, for probe access and/or wire bonding.

[0101] Still referring to **Figure 18**, first portions **6** (also referred to as "wings" or "sidelobes") of the gate electrode **32** laterally extend on surface portions of the protective layer **24** outside the first opening **4**, and second portions **7** of the gate electrode **32** that are spaced vertically apart from the protective layer **24** laterally extend on portions of the oxide layer **33** outside the second opening **5**. As used herein, the term "laterally" refers to a direction that is substantially parallel with respect to a surface of the substrate. The second portions **7** of the gate electrode **32** may laterally extend beyond the first portions **6**. In other words, the first and second

portions **6** and **7** of the gate electrode **32** define a stair-step profile. Accordingly, a modified gamma gate **32** having self-aligned sidelobes **6** may be formed.

[0102] Also, as shown in **Figure 18**, because the first opening **4** and the second opening **5** may be self-aligned by widening the opening in the oxide layer **33** as described above, the length ΔL_G by which the first portions **6** of the gate electrode **32** extend onto the protective layer **24** may be controlled. In addition, the second portions **7** of the gate electrode **32** may be spaced apart from the protective layer **24** based on the thickness of the oxide layer **33**. As such, according to some embodiments of the present invention, gate-to-drain capacitance (c_{gd}) and/or gate-to-source capacitance (c_{gs}) of the transistor device that may be caused by the extension of the gate electrode **32** onto the protective layer **24** may also be controlled. The relatively low dielectric index of the high-quality oxide layer **33**, may further contribute to the reduction in gate-to-drain and/or gate-to-source capacitance.

[0103] The thickness d_1 of the protective layer **24** and the length ΔL_G of first portion of the gate electrode **32** may also be optimized to reliably support the peak gate-to-drain voltage V_{GD} and/or to minimize and/or reduce the gate-to-drain capacitance c_{gd} and/or the gate-to-source capacitance c_{gs}. For example, the length ΔL_G may be about 0.1 micrometer (μm) to about 0.6 μm, and the thickness d_1 may be about 200 Angstroms (Å) to about 2000 Å. In addition, the thickness d_2 of the oxide layer **33** may be about 500 Angstroms (Å) to about 3000 Å. The angle Ø of the sidewalls of the protective layer **24** defining the first opening may also be optimized, for example, between about 45 degrees to about 90 degrees. Also, a width of the second opening (illustrated as L_{G2}) may be self-aligned and symmetric around a width of the first opening, which defines the gate length L_{G1}. As such, the gate length L_{G1} may be tuned for an operational frequency of interest based on formation of the opening in the protective layer **24**. For example, the gate length L_{G1} may be about 0.05 micrometer (μm) to about 2.0 μm.

[0104] Thus, some embodiments of the present invention may combine many of the positive qualities of a T-gate with those of a dielectrically-supported gate electrode, or gamma gate. Accordingly, a modified gamma gate structure with self-aligned wings/sidelobes is provided having relatively low gate resistance, relatively low gate-to-drain and/or gate-to-source capacitance, and relatively high gain. In addition, gaps between the gate electrode and the protective layer may be reduced and/or avoided, and the peak electric fields may be reduced.

[0105] While embodiments of the present invention have been described herein with reference to particular HEMT structures, the present invention should not be construed as limited to such structures, and may be applied to formation of gate electrodes in many different transistor structures, such as pseudomorphic HEMTs (pHEMTs) (including GaAs/AlGaAs pHEMTs) and/or GaN MESFETs. More generally, embodiments of the present invention may be applied in any electronic device where a gate electrode may be formed and/or where preservation of the as-grown surface properties may be imperative. For example, silicon nitride and/or other removable encapsulant can be used in the fabrication sequence of AlGaN metal-semiconductor-metal (MSM) detectors, RF limiter diode structures, and/or power switching Schottky diodes.

[0106] Also, additional layers may be included in transistor devices while still benefiting from the teachings of the present invention. Such additional layers may include GaN cap layers, as for example, described in Yu et al., "Schottky barrier engineering in III–V nitrides via the piezoelectric effect," Applied Physics Letters, Vol. 73, No. 13, 1998, or in U.S. Patent Publication No. 2002/0066908A1 published June 6, 2002, for "ALUMINUM GALLIUM NITRIDE/GALLIUM NITRIDE HIGH ELECTRON MOBILITY TRANSISTORS HAVING A GATE CONTACT ON A GALLIUM NITRIDE BASED CAP SEGMENT AND METHODS OF FABRICATING SAME," the disclosures of which are incorporated by reference as if set forth fully herein. In some embodiments, insulating layers such as SiN_x, or relatively high quality AlN may be deposited for making a MISHEMT and/or passivating the surface. The additional layers may also include a compositionally graded transition layer or layers. In addition, the barrier layer **22** described above may also include multiple layers. Thus, embodiments of the present invention should not be construed as limiting the barrier layer to a single layer but may include, for example, barrier layers having combinations of GaN, AlGaN and/or AlN layers. For example, a GaN, AlN structure may be utilized to reduce or prevent alloy scattering.

[0107] In the drawings and specification, there have been disclosed typical embodiments of the invention, and, although specific terms have been employed, they have been used in a generic and descriptive sense only and not for purposes of limitation.

THAT WHICH IS CLAIMED:

1. A method of fabricating a transistor, the method comprising:

forming a protective layer, the protective layer having an opening extending therethrough; and

forming a gate electrode in the opening so that a first portion of the gate electrode laterally extends on surface portions of the protective layer outside the opening and so that a second portion of the gate electrode spaced apart from the protective layer laterally extends beyond the first portion.

2. The method of Claim 1, further comprising:

before forming the gate electrode, forming a second layer on the protective layer, the second layer having a second opening extending therethrough that is wider than the first opening,

wherein forming the gate electrode further comprises forming the gate electrode in the second opening so that the second portion of the gate electrode laterally extends on surface portions of the second layer outside the second opening.

3. The method of Claim 2, wherein the second layer comprises a material having a lower dielectric index than the protective layer.

4. The method of Claim 2, wherein forming the second layer having the second opening extending therethrough comprises:

forming the second layer on the protective layer including a recess extending therethrough and exposing a portion of the protective layer; and

widening the recess in the second layer to define the second opening exposing the first opening and exposing surface portions of the protective layer on opposite sides of the first opening.

5. The method of Claim 4, wherein the protective layer and the second layer comprise different materials, and further comprising:

patterning the protective layer using the second layer as a mask to form the first opening extending through the protective layer prior to widening the recess in the second layer.

6. The method of Claim 5, wherein widening the recess in the second layer comprises:

symmetrically expanding the recess in the second layer so that the second opening and the first opening are self-aligned.

7. The method of Claim 4, wherein the second layer comprises a photoresist layer.

8. The method of Claim 7, wherein widening the recess in the second layer comprises:

ashing the second layer using an oxygen plasma.

9. The method of Claim 7, wherein widening the recess in the second layer comprises:

hard baking the second layer.

10. The method of Claim 2, further comprising:

removing the second layer after forming the gate electrode.

11. The method of Claim 10, further comprising:

forming a passivation layer on the protective layer and on the gate electrode after removing the second layer.

12. The method of Claim 1, wherein forming the gate electrode further comprises:

forming the gate electrode in the opening directly on opposing sidewalls of the protective layer.

13. The method of Claim 1, further comprising:

forming a channel layer; and

forming a barrier layer on the channel layer,

wherein forming the protective layer comprises forming the protective layer on the barrier layer, wherein forming the gate electrode comprises forming the gate

electrode extending through the opening in the protective layer to contact the barrier layer, and wherein a junction between the channel layer and the barrier layer define a heterojunction.

14. The method of Claim 13, further comprising:

forming first and second ohmic contact regions on the barrier layer adjacent to and spaced apart from the protective layer so that the protective layer is between the first and second ohmic contact regions.

15. The method of Claim 14, wherein forming the first and second ohmic contact regions comprises:

patterning the protective layer to expose portions of the barrier layer;

forming ohmic metal regions on the exposed portions of the barrier layer adjacent to and spaced apart from the patterned protective layer; and

annealing the ohmic metal regions.

16. The method of Claim 1, wherein the protective layer comprises a dielectric material including silicon nitride, aluminum nitride, and/or silicon dioxide.

17. The method of Claim 1, further comprising:

before forming the gate electrode, forming a dielectric layer on the protective layer, the dielectric layer having a second opening extending therethrough that is wider than the first opening,

wherein forming the gate electrode further comprises forming the gate electrode in the second opening so that the second portion of the gate electrode laterally extends on surface portions of the dielectric layer outside the second opening.

18. The method of Claim 17, wherein the dielectric layer has a lower dielectric index than the protective layer.

19. The method of Claim 18, wherein the dielectric layer comprises an oxide layer.

20. The method of Claim 19, wherein the dielectric layer comprises a high-temperature oxide (HTO) layer.

21. The method of Claim 20, wherein the protective layer comprises a high-purity nitride (HPN) layer.

22. The method of Claim 21, wherein the protective layer comprises stoichiometric silicon nitride, and wherein the dielectric layer comprises silicon dioxide.

23. The method of Claim 17, wherein the dielectric layer comprises a high-quality oxide layer having a dielectric index of less than about 1.5.

24. The method of Claim 17, wherein the dielectric layer has a thickness greater than that of the protective layer.

25. The method of Claim 24, wherein the dielectric layer has a thickness of about 500 Angstroms (Å) to about 3000 Å, and wherein the protective layer has a thickness of about 200 Å to about 2000 Å.

26. The method of Claim 17, wherein forming the dielectric layer having the second opening extending therethrough comprises:

forming the dielectric layer on the protective layer including a recess extending therethrough and exposing a portion of the protective layer;

patterning the protective layer using the dielectric layer as a mask to form the first opening extending through the protective layer; and

widening the recess in the dielectric layer to define the second opening exposing the first opening and exposing surface portions of the protective layer on opposite sides of the first opening,

wherein the protective layer and the dielectric layer comprise different materials, and wherein widening the recess in the dielectric layer comprises symmetrically expanding the recess in the dielectric layer so that the second opening and the first opening are self-aligned by etching the dielectric layer using a wet oxide etchant, wherein the wet oxide etchant is selective with respect to the protective layer.

27. The method of Claim 17, further comprising:

forming a channel layer; and

forming a barrier layer on the channel layer,

wherein forming the protective layer comprises forming the protective layer on the barrier layer, wherein forming the gate electrode comprises forming the gate electrode extending through the first opening in the protective layer to contact the barrier layer, and wherein a junction between the channel layer and the barrier layer define a heterojunction, and further comprising:

patterning the dielectric layer and the protective layer to expose first and second portions of the barrier layer; and

after patterning the dielectric layer and the protective layer, respectively forming first and second ohmic contact regions on the first and second portions of the barrier layer adjacent to and spaced apart from the patterned protective layer so that the protective layer is between the first and second ohmic contact regions.

28. The method of Claim 27, wherein patterning the dielectric layer and the protective layer comprises:

before forming the dielectric layer, patterning the protective layer to expose the first and second portions of the barrier layer, wherein forming the dielectric layer comprises forming the dielectric layer on the protective layer and on the first and second portions of the barrier layer; and then

patterning the dielectric layer to define first and second recesses therein respectively exposing the first and second portions of the barrier layer.

29. The method of Claim 27, wherein patterning the dielectric layer and the protective layer comprises:

patterning the dielectric layer to expose first and second portions of the protective layer; and

patterning the protective layer using the dielectric layer as a mask to expose the first and second portions of the barrier layer.

30. The method of Claim 17, wherein forming the gate electrode comprises:

forming a dielectric liner in the first opening on opposing sidewalls of the protective layer; and

after forming the dielectric liner, forming the gate electrode in the first opening directly on the dielectric liner.

31. A method of fabricating a transistor, the method comprising:

forming a first layer, wherein the first layer comprises a dielectric material;

forming a second layer on the first layer, wherein the second layer has a hole therein exposing a portion of the first layer;

forming a first opening extending through the first layer using the second layer as a mask;

after forming the first opening, widening the hole in the second layer to form a second opening exposing the first opening and exposing surface portions of the first layer on opposite sides of the first opening; and

forming a gate electrode in the first and second openings, wherein first portions of the gate electrode extend on surface portions of the first layer outside the first opening and second portions of the gate electrode extend on surface portions of the second layer outside the second opening.

32. The method of Claim 31, wherein the first portions of the gate electrode laterally extend on surface portions of the first layer outside the opening, wherein the second portions of the gate electrode laterally extend on the surface portions of the second layer outside the second opening beyond the first portions of the gate electrode.

33. The method of Claim 31, wherein forming the gate electrode further comprises:

forming the gate electrode in the first opening directly on opposing sidewalls thereof.

34. The method of Claim 31, wherein the first layer comprises silicon nitride, aluminum nitride, and/or silicon dioxide.

35. The method of Claim 31, wherein the first layer and the second layer comprise different materials, and wherein the second layer comprises a material having a lower dielectric index than the first layer.

36. The method of Claim 31, wherein widening the hole in the second layer comprises:

symmetrically expanding the hole in the second layer so that the second opening and the first opening are self-aligned.

37. The method of Claim 31, wherein the second layer comprises a photoresist layer.

38. The method of Claim 37, wherein widening the hole in the second layer comprises:

ashing the second layer using an oxygen plasma and/or hard baking the second layer.

39. The method of Claim 31, wherein widening the hole in the second layer comprises selectively etching the second layer using a wet etch to form the second opening therein.

40. The method of Claim 31, further comprising:

forming a channel layer; and

forming a barrier layer on the channel layer,

wherein forming the first layer comprises forming the first layer on the barrier layer, wherein forming the gate electrode comprises forming the gate electrode extending through the first opening in the first layer to contact the barrier layer, and wherein a junction between the channel layer and the barrier layer define a heterojunction.

41. A transistor, comprising:

a protective layer having an opening extending therethrough; and

a gate electrode in the opening, wherein the gate electrode includes a first portion laterally extending on surface portions of the protective layer outside the

opening and a second portion spaced apart from the protective layer and laterally extending beyond the first portion.

42. The transistor of Claim 41, further comprising:

a second layer on the protective layer, the second layer having a second opening extending therethrough that is wider than the first opening,

wherein the gate electrode is in the second opening, and wherein the second portion of the gate electrode laterally extends on portions of the second layer outside the second opening.

43. The transistor of Claim 42, wherein the first and second openings define a stair-step profile.

44. The transistor of Claim 42, wherein the second layer has a lower dielectric index than the protective layer.

45. The transistor of Claim 41, further comprising:

a passivation layer on the protective layer and on the gate electrode.

46. The transistor of Claim 45, wherein the protective layer and the passivation layer comprise a same material.

47. The transistor of Claim 41, wherein the gate electrode is directly on opposing sidewalls of the opening in the protective layer.

48. The transistor of Claim 41, further comprising:

a channel layer; and

a barrier layer on the channel layer,

wherein the protective layer is on the barrier layer, wherein the gate electrode extends through the opening in the protective layer to contact the barrier layer, and wherein the channel layer and the barrier layer are configured to provide a High Electron Mobility Transistor (HEMT).

49. The transistor of Claim 48, further comprising:

first and second ohmic contact regions on the barrier layer at opposite sides of the gate electrode and spaced apart from the protective layer.

50. The transistor of Claim 49, wherein the protective layer has a thickness of at least about a thickness of the first and second ohmic contact regions.

51. The transistor of Claim 41, wherein the protective layer comprises a dielectric material including silicon nitride, aluminum nitride, and/or silicon dioxide.

52. The transistor of Claim 41, further comprising:

a dielectric layer on the protective layer, the dielectric layer having a second opening extending therethrough that is wider than the first opening,

wherein the gate electrode is in the second opening, and wherein the second portion of the gate electrode laterally extends on portions of the dielectric layer outside the second opening.

53. The transistor of Claim 52, wherein the dielectric layer and the protective layer comprise different materials, and wherein the dielectric layer has a lower dielectric index than the protective layer.

54. The transistor of Claim 53, wherein the dielectric layer comprises an oxide layer.

55. The transistor of Claim 54, wherein the dielectric layer comprises a high-temperature oxide (HTO) layer.

56. The transistor of Claim 55, wherein the protective layer comprises a high-purity nitride (HPN) layer.

57. The transistor of Claim 56, wherein the protective layer comprises stoichiometric silicon nitride, and wherein the dielectric layer comprises silicon dioxide.

58. The transistor of Claim 52, wherein the dielectric layer comprises a high-quality oxide layer having a dielectric index of less than about 1.5.

59. The transistor of Claim 52, wherein the dielectric layer has a thickness greater than that of the protective layer.

60. The transistor of Claim 59, wherein the dielectric layer has a thickness of about 500 Angstroms (Å) to about 3000 Å, and wherein the protective layer has a thickness of about 200 Å to about 2000 Å.

61. The transistor of Claim 52, further comprising:

a channel layer; and

a barrier layer on the channel layer; and

first and second ohmic contact regions on the barrier layer at opposite sides of the gate electrode and spaced apart from the protective layer,

wherein the dielectric layer extends between the first and second ohmic contact regions and the protective layer, wherein the protective layer is on the barrier layer, wherein the gate electrode extends through the first opening in the protective layer to contact the barrier layer, and wherein the channel layer and the barrier layer are configured to provide a High Electron Mobility Transistor (HEMT).

62. The transistor of Claim 52, further comprising:

a dielectric liner in the first opening between a sidewall of the protective layer and the gate electrode, wherein the gate electrode is directly on the dielectric liner.

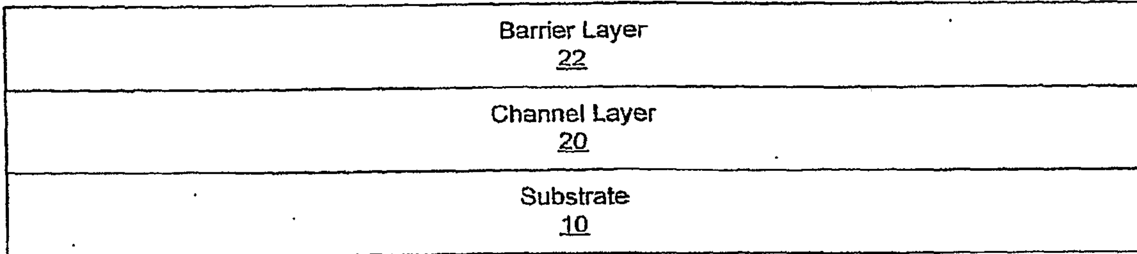


FIG. 1

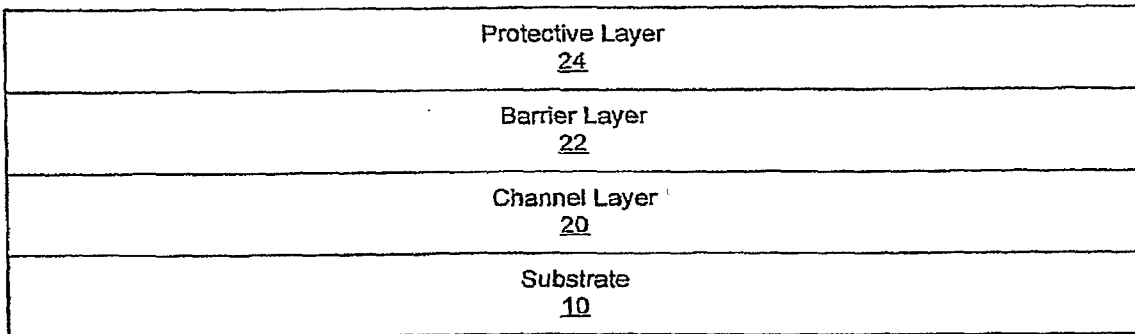


FIG. 2

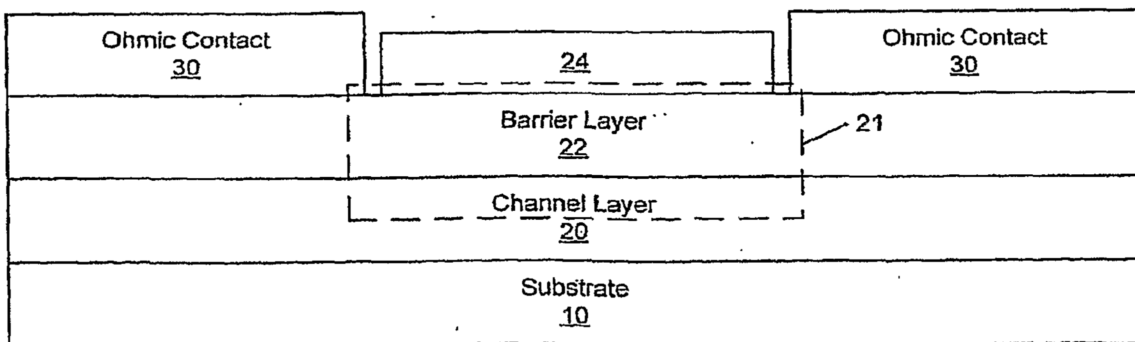


FIG. 3

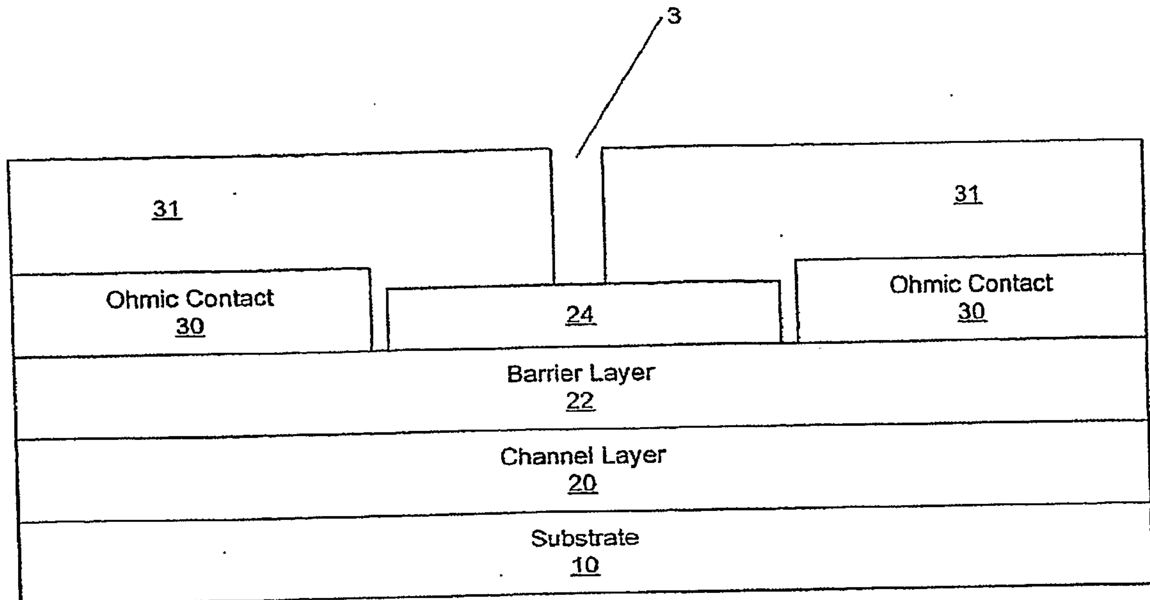


FIG. 4

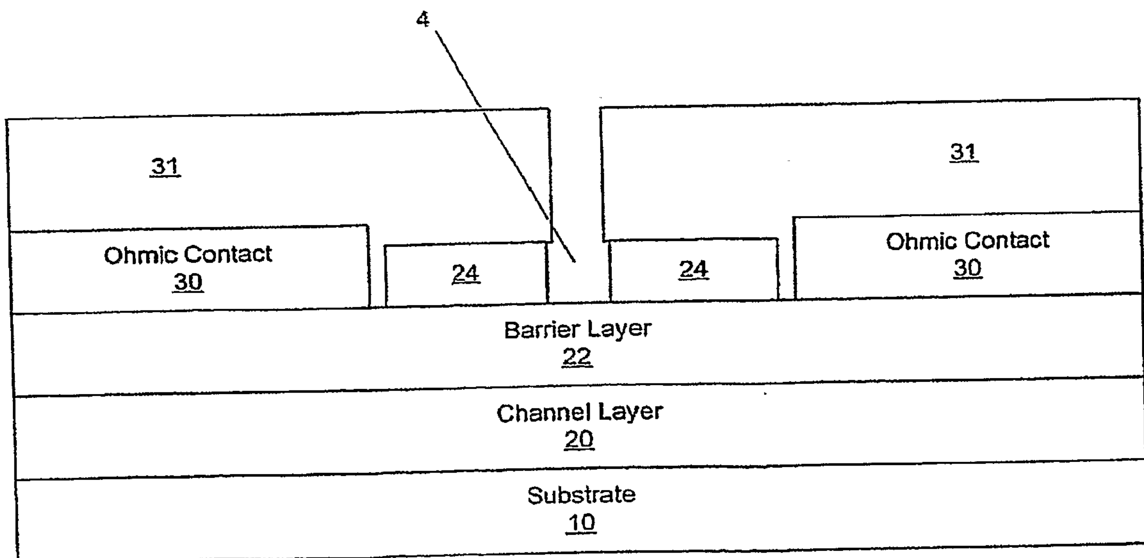


FIG. 5

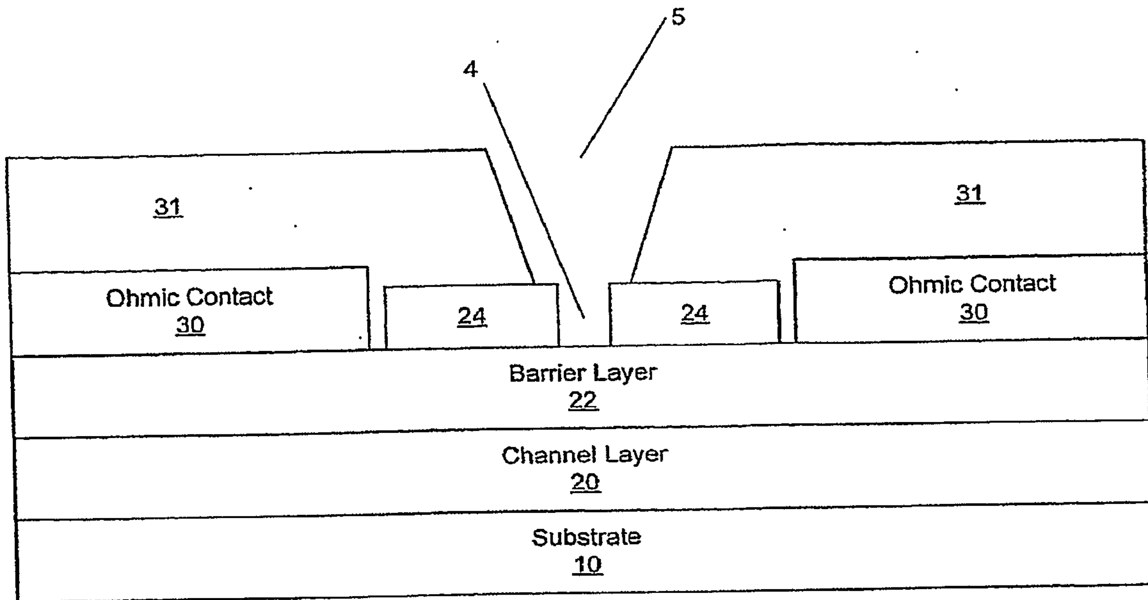


FIG. 6

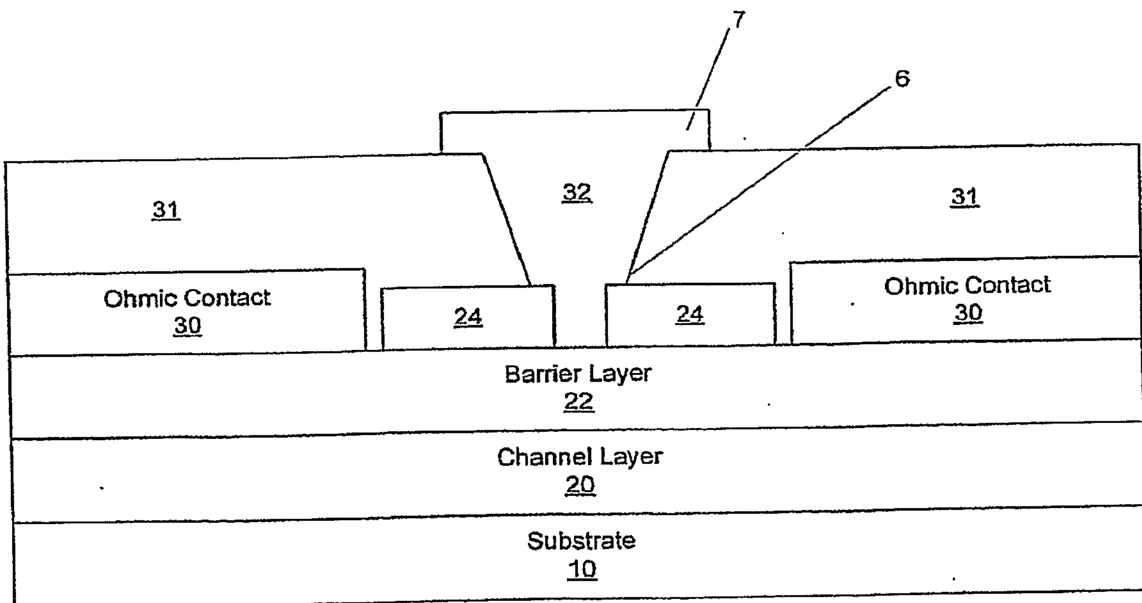


FIG. 7

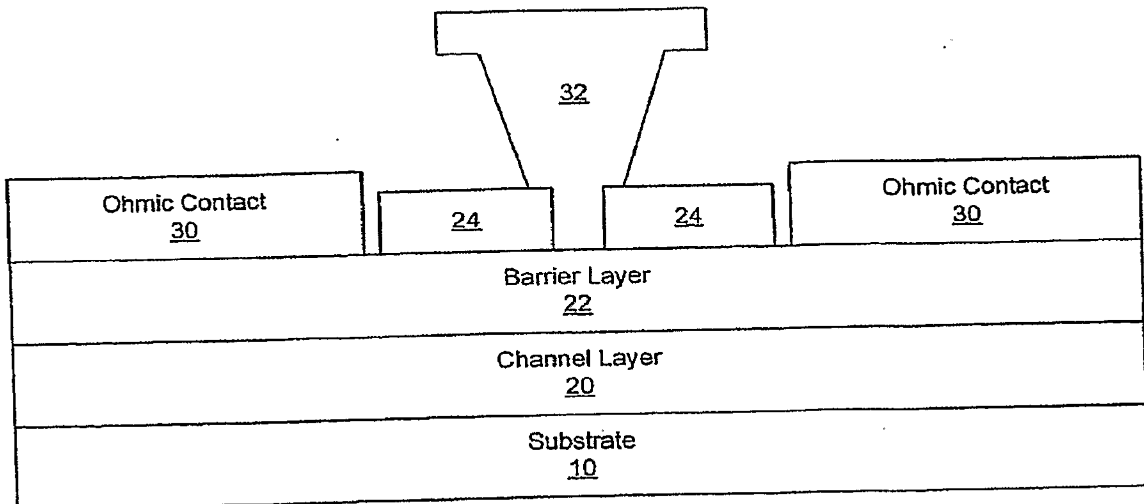


FIG. 8

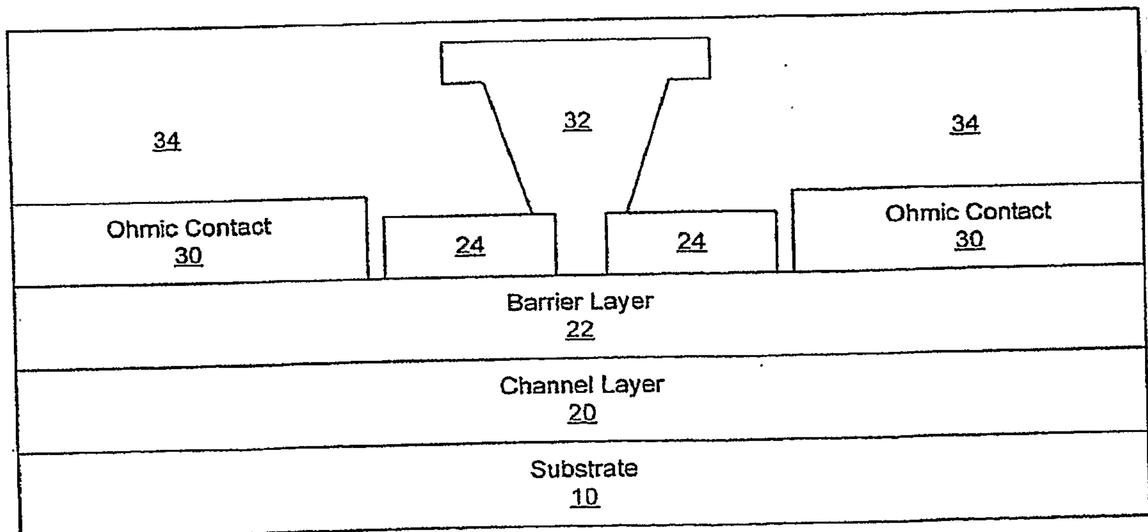


FIG. 9

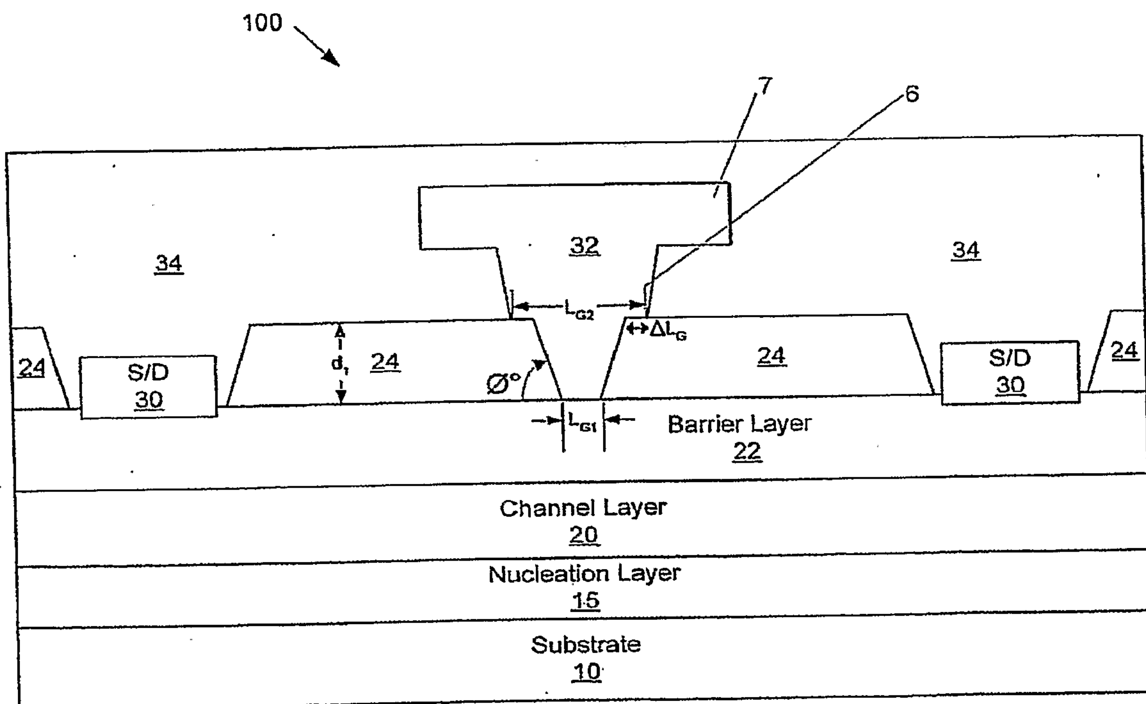


FIG. 10

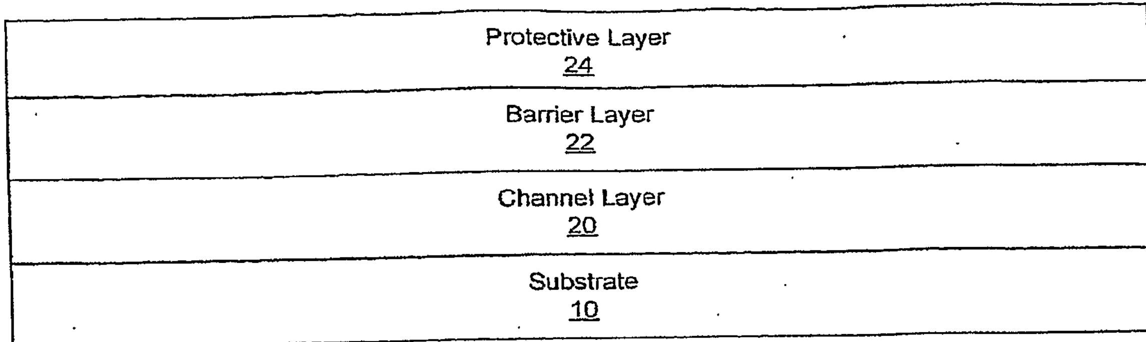


FIG. 11

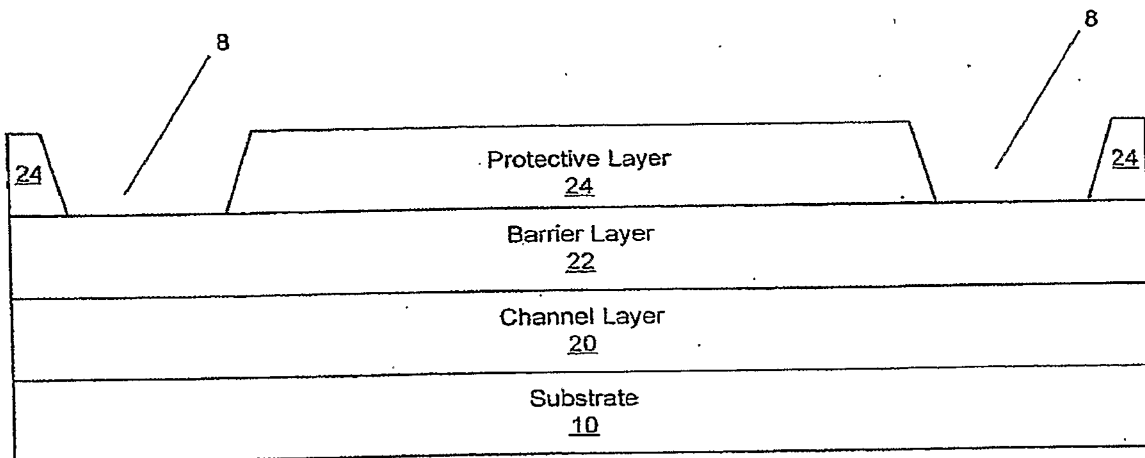


FIG. 12

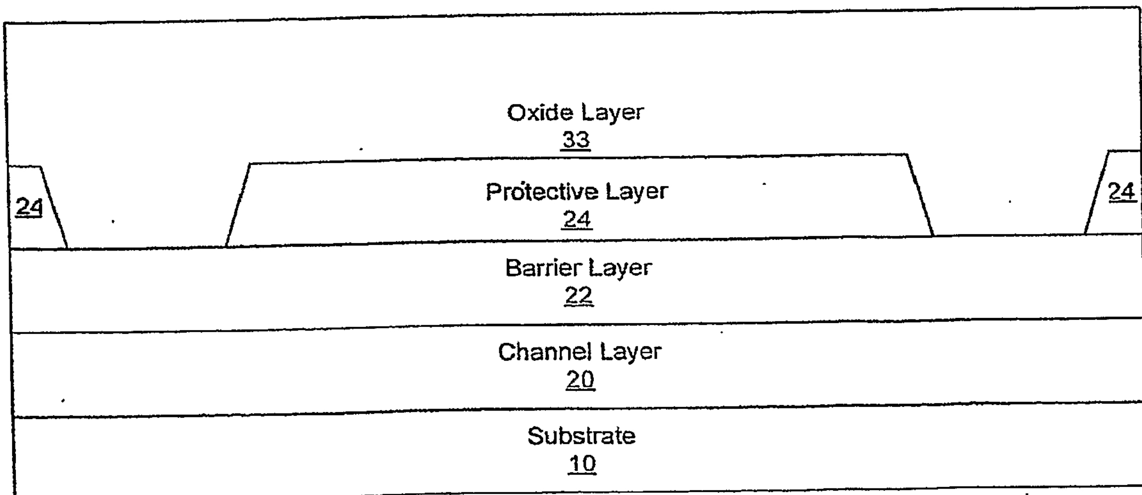


FIG. 13

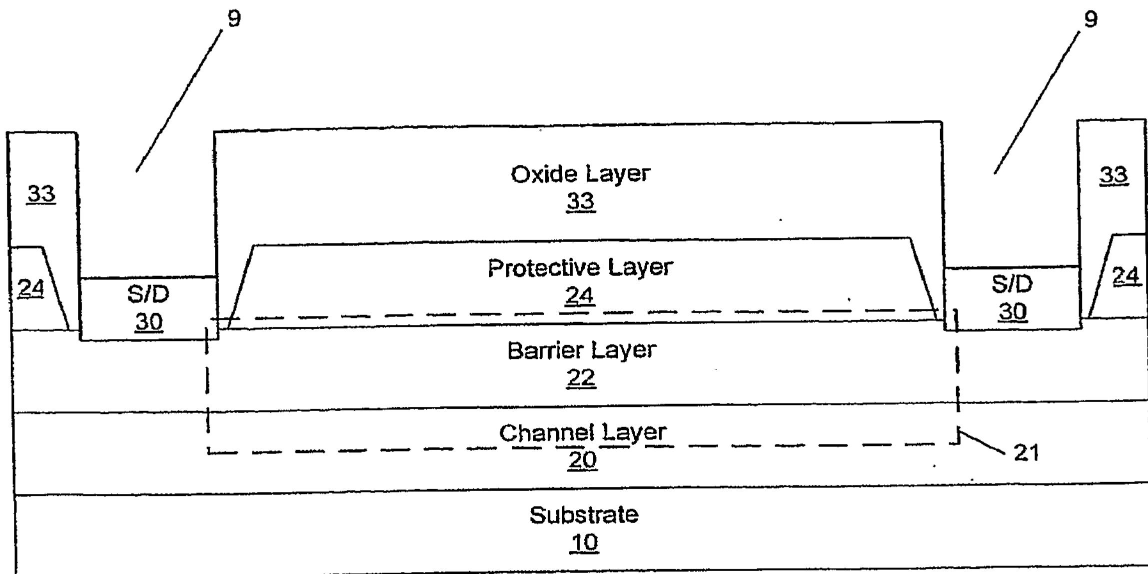


FIG. 14

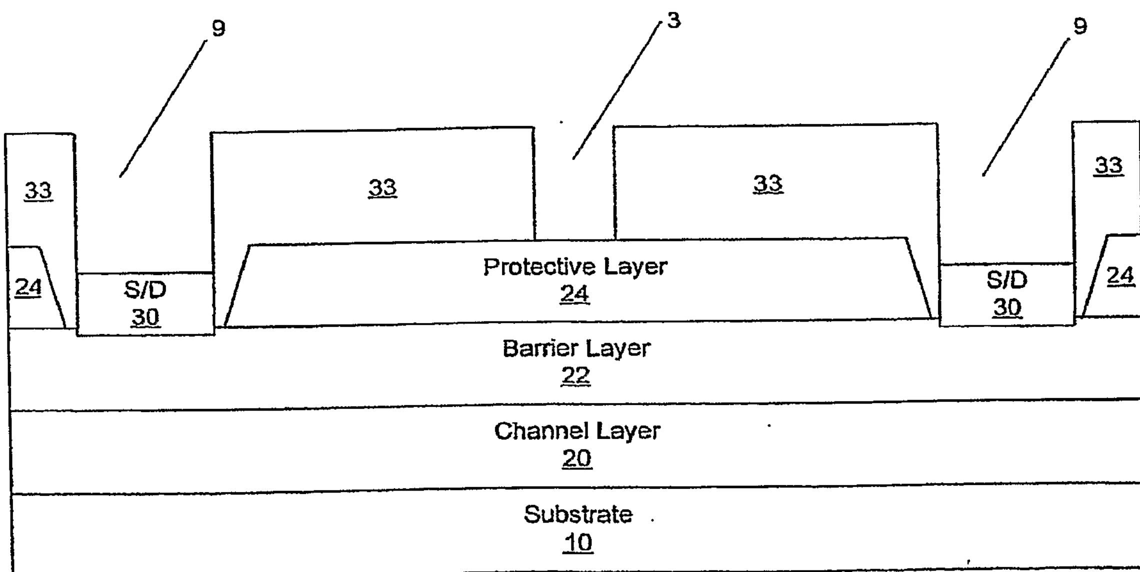


FIG. 15

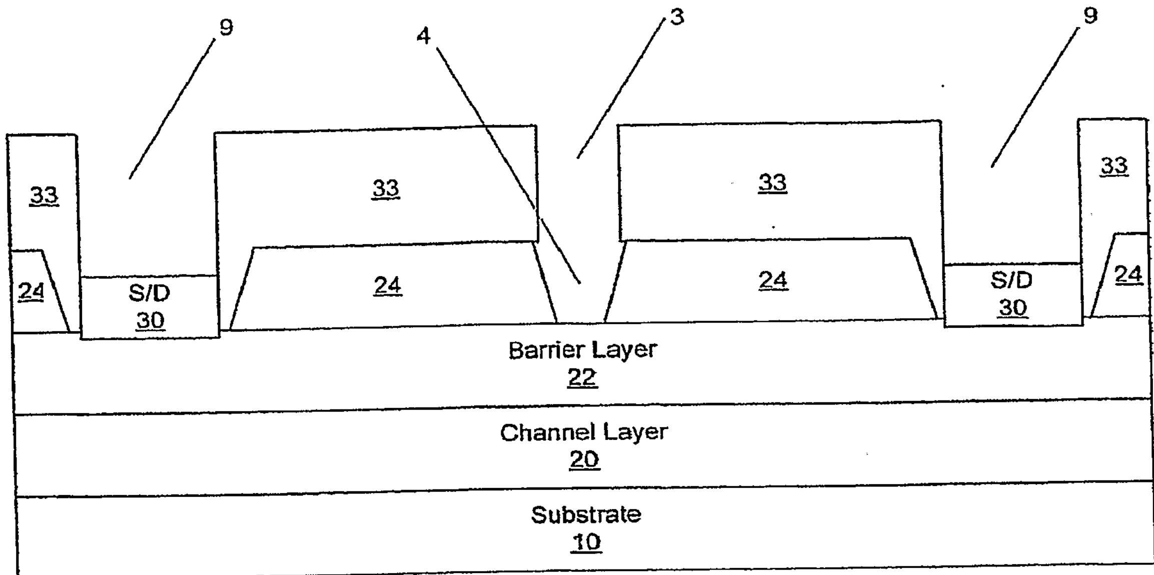


FIG. 16

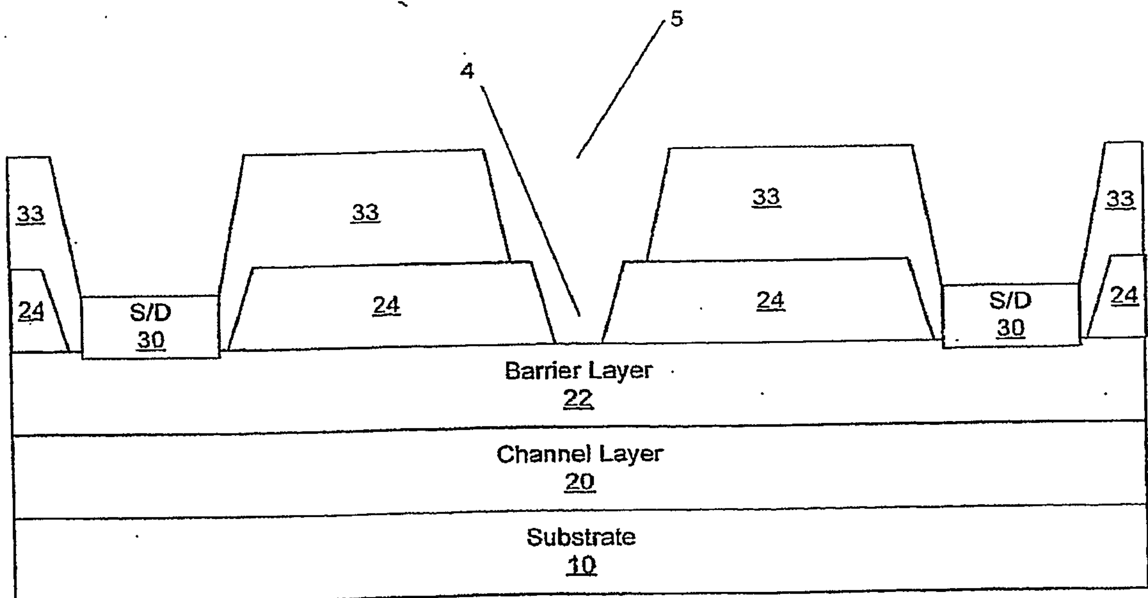


FIG. 17

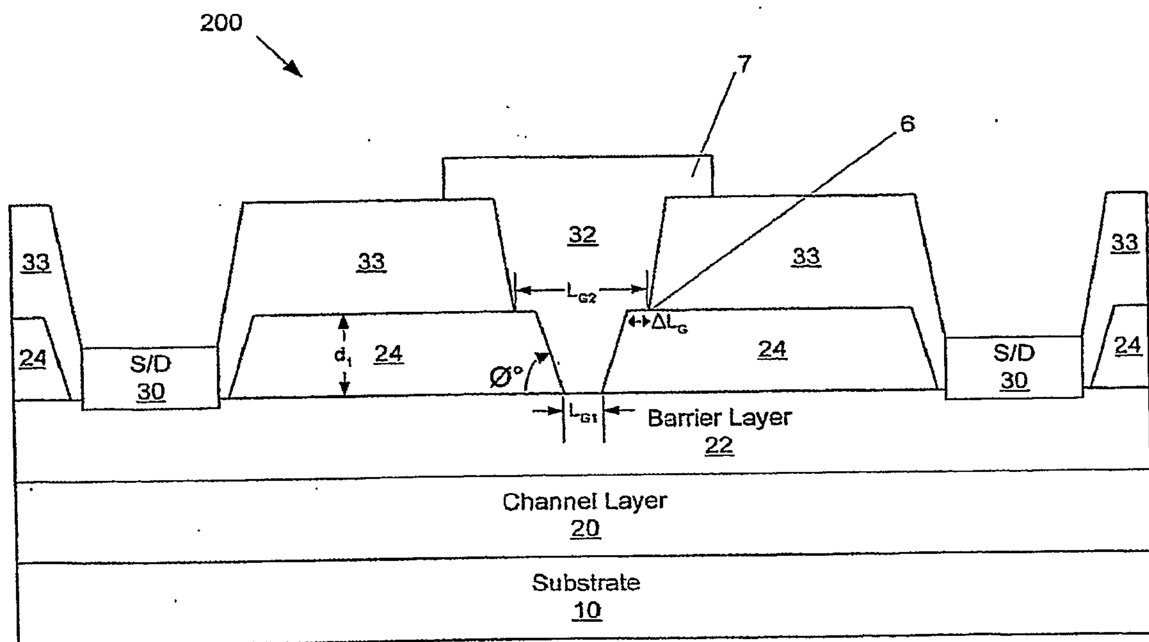


FIG. 18

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2007/001045

A. CLASSIFICATION OF SUBJECT MATTER
INV. H01L21/338 H01L29/423 H01L29/778
ADD. H01L21/335

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
H01L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2005/029589 A (CREE INC [US]) 31 March 2005 (2005-03-31)	1-3, 8-25, 41-61
Y	page 16, line 23 - page 19, line 25; figures 7,8	4-7, 26-40,62
X	US 2003/006437 A1 (MIZUTA MASASHI [JP] ET AL) 9 January 2003 (2003-01-09) paragraphs [0098] - [0106]; figures 6,7	1,2, 12-14, 17,19, 41-43, 47-49, 52,61

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
5 June 2007	15/06/2007

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Dauw, Xavier

2

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2007/001045

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2005/059197 A1 (YAMASHITA YOSHIMI [JP] ET AL) 17 March 2005 (2005-03-17) abstract; figure 8	1,2, 12-14, 17,19, 41-43, 47-49, 52,61
Y	US 6 403 456 B1 (PLAT MARINA [US] ET AL) 11 June 2002 (2002-06-11) abstract; figures 7-12	4-7, 26-29, 31-40
X	US 6 867 078 B1 (GREEN BRUCE M [US] ET AL) 15 March 2005 (2005-03-15) abstract; figures 7,8	1,41
Y	US 5 181 087 A (USAGAWA TOSHIYUKI [JP] ET AL) 19 January 1993 (1993-01-19) figures 6a,b	30,62
A	US 5 364 816 A (BOOS J BRAD [US] ET AL) 15 November 1994 (1994-11-15) abstract; figures 4-8	1-62
A	US 6 087 256 A (WADA SHIGEKI [JP]) 11 July 2000 (2000-07-11) abstract; figures 6,11	1-62

2

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2007/001045

Patent document cited in search report		Publication date	Patent family member(s)			Publication date
WO 2005029589	A	31-03-2005	CA	2536030	A1	31-03-2005
			EP	1665385	A1	07-06-2006
			JP	2007505501	T	08-03-2007
			KR	20060071415	A	26-06-2006
			US	2005051796	A1	10-03-2005
US 2003006437	A1	09-01-2003	NONE			
US 2005059197	A1	17-03-2005	JP	2005086171	A	31-03-2005
US 6403456	B1	11-06-2002	NONE			
US 6867078	B1	15-03-2005	NONE			
US 5181087	A	19-01-1993	US	5373191	A	13-12-1994
US 5364816	A	15-11-1994	NONE			
US 6087256	A	11-07-2000	JP	2904167	B2	14-06-1999
			JP	10178025	A	30-06-1998