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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 March 2002 (07.03.2002)

PCT

(10) International Publication Number
WO 02/19396 A1

(51) International Patent Classification[7]: H01L 21/00, 21/84, 21/336

(21) International Application Number: PCT/US01/26920

(22) International Filing Date: 29 August 2001 (29.08.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/229,552	29 August 2000 (29.08.2000)	US
60/287,226	27 April 2001 (27.04.2001)	US
09/	29 August 2001 (29.08.2001)	US

(71) Applicant *(for all designated States except US)*: **WIRELESS SYSTEMS, INC.** [US/US]; 8808 133rd Avenue NE, Redmond, WA 98052 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **PARKE, Stephen, A.** [—/US]; 3604 Brenan Drive, Nampa, ID 83686 (US).

(74) Agents: **LAURENCE, Kevin, B.** et al.; Stoel Rives LLP, Standard Insurance Center, 900 SW Fifth Avenue, Suite 2300, Portland, OR 97204-1268 (US).

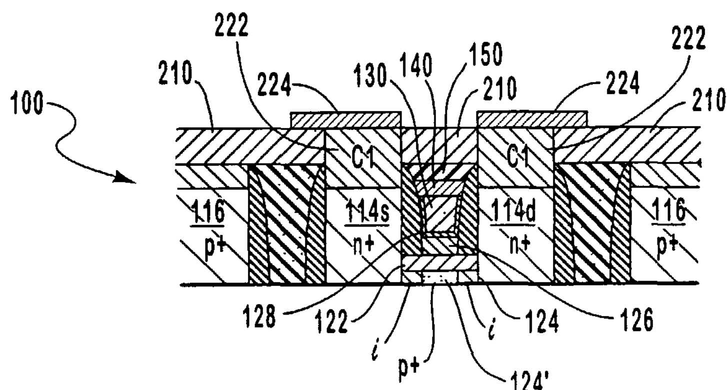
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

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.

(54) **Title:** DAMASCENE DOUBLE GATED TRANSISTORS AND RELATED MANUFACTURING METHODS



(57) Abstract: This invention provides the structure and fabrication process of a completely planar, Damascene double gated transistor (100). The structure has a novel self-aligned, hyper-abrupt retrograde body (130) and a zero-parasitic, endwall gate-body connection. The structure (100) provides for increased density and enables ultra low power to be utilized. The methods also provide for simultaneously making both four-terminal and dynamic threshold MOSFET devices.

WO 02/19396 A1

DAMASCENE DOUBLE GATED TRANSISTORS And RELATED MANUFACTURING METHODS

Technical Field

[0001] This invention relates generally to transistors. More particularly, this invention relates to a transistor that is fully planar and operates at low voltages.

Background of the Invention

[0002] The shrinking of silicon-based CMOS transistors from 10um to sub-100nm channel lengths has been the driving force behind the incredible advances in integration density of digital circuits. This exponential growth in circuit density (Moore's Law) has enabled increasingly sophisticated and fast microprocessors and digital signal processors. However, in this "post-PC age", the markets driving the semiconductor industry are handheld consumer electronics, wireless communications, and Internet infrastructure. The ITRS Roadmap for the next ten years calls as much or more for increased diversity of integrated functions as for increased density.

[0003] CMOS switching frequencies (f_T) exceeding 150GHz have recently been achieved. Because this high-performance can be achieved using a relatively low cost manufacturing process, CMOS is rapidly becoming a serious option for many wireless Radio Frequency (RF) applications that were previously considered to be the exclusive domain of more expensive technologies, such as bipolar and gallium arsenide. This has led to a research focus shift to highly-integrated, highly-diverse CMOS System-On-a-Chip (SOC) technologies, integrating entire wireless mixed-signal systems on a single chip. Innovative device and technology changes are needed to provide high quality passive elements and high-gain, high-linearity, low-noise, low-power, RF capable transistors on carefully engineered substrates which reduce coupling and crosstalk noise of digital circuits from affecting sensitive on-chip RF receivers.

[0004] RF SOC's with the lowest possible energy consumption will use advanced spread-spectrum communication algorithms. Current research spans the circuit, architecture, and algorithm issues which impact the design of SOC CMOS wireless transceivers. Research activities also span bulk Si, SiGe, and SOI deep-sub-micron CMOS designs, the design of analog RF front ends, the design of A/D interface

circuitry, and digital baseband signal processing. However, it is clear that single chip RF SOC solutions will require more than simply innovative circuit designs in order to be optimized.

[0005] The most serious limitation of these portable SOCs is power dissipation. The system battery typically comprises a large percentage of the cost, size, weight, and reliability problems in current handheld RF devices. As these systems become more highly integrated by utilizing sub-100nm RF-CMOS devices, the power dissipation will become an even greater concern. Therefore, in order for the low-cost and high-integration density of RF-CMOS to be fully realized, it is imperative that new ultra-low power device structures and circuit design techniques be developed. These devices should preferably maintain multi-GHz level performance while operating with power supply voltages down to 0.25 volts. They also should preferably maintain high transconductance and low noise, and be easily integrated with on-chip, high-Q, tunable passive elements such as spiral inductors.

[0006] In summary, there is a need in the art for devices that operate at ultra low power. There is a need in the art for innovative transistors that provide for increased density. Additionally, there is a need to simplify the process for manufacturing transistors. Finally, there is a need to reduce the cost of manufacturing transistors.

Summary of the Invention

[0007] An object of the present invention is provide unique transistors that enable increased density. Another object of the present invention is provide transistors that operate at ultra low power. An additional object of the present invention is to simplify the process for manufacturing transistors. Finally, an object is to reduce the cost of manufacturing transistors.

[0008] The above and other objects and advantages, which will be apparent to one of skill in the art, are achieved in the present invention which is directed to, in the first aspect, a method of manufacturing a Damascene double gated transistor, and the structure made by the method. More particularly, the invention provides a method for simultaneously making both four-terminal and dynamic threshold MOSFET devices. The method starts with either 1) a high-resistivity bulk silicon substrate (with or without an epitaxial layer) or a silicon-on-insulator substrate or similar substrate

having an isolated semiconductor region defined on its surface. These substrates may be created by any of the existing methods for making 1)high-resistivity bulk silicon or 2) SOI i.e. SIMOX, BESOI, SmartCut, Metal-induced recrystallization, Laser-induced recrystallization, etc.

[0009] In one embodiment, a silicon trench etch hard mask and polish stop "pad" layer (e.g. silicon nitride) is first deposited on the top silicon layer. A merged isolation/gate trench is then etched partially through this top silicon layer. A second mask and etch are used to completely remove the remaining silicon from the bottom of the isolation trenches, while leaving a thin active region of silicon under the gate trench which will become the channel. The remaining silicon "island" is barbell shaped, with large, thick source/drain regions on either side of the narrow, thin channel region. The gate trench is used for subsequently forming the bottom gate, channel, and top gate which are all self-aligned to this trench opening. A complementary set of shallow and deep implants is used to form a counter-doped channel and to heavily, uniformly dope the source/drain regions, respectively. This counter-doped channel is especially useful for maximizing the dynamic threshold swing, since it enhances body effect. Sidewall spacer material is then deposited, both to narrow the gate opening for sub-lithographic channel length and to isolate the sidewalls of the source/drain regions. Following the spacer formation, a highly-doped buried bottom gate is formed by ion implantation. This implant is masked by both the pad and the spacers, which space it away from both the heavily doped source and drain regions to prevent leakage and capacitance. The energy of this implant is high enough to place it below the counter-doped channel layer. High atomic weight ions are used (Indium for p+ and arsenic for n+) in order to obtain a super-steep retrograde implant profile between this buried layer and the channel above. Optionally, a thin undoped selective epitaxial silicon layer may be grown on top of the counter-doped channel to enhance the mobility of the device. At this point in the process, all high temperature steps are complete, meaning that temperature-sensitive materials (such as hafnium oxide and copper) may be used for the gate stack. Polysilicon or poly-silicon/germaninum with an adjustable percentage of Ge may be used as the first gate conductor layer, in order to achieve additional threshold voltage design control. Alternately, a metal or silicide may also be used for the first gate conductor layer.

The isolation/gate trenches are overfilled with this first gate conductor material, planarized back to the pad layer, and plasma recessed. Next, the entire channel and gate stack is patterned and etched, stopping on the bottom gate layer. The alignment of this mask to the original channel mask determines whether the resulting device is four-terminal, DTMOS, or floating body. All three types of devices are fabricated simultaneously on the same substrate and may be used as needed for various circuit applications. Next, the trenches are refilled with isolation dielectric (e.g. silicon dioxide) and planarized. This trench dielectric is then patterned and etched out in order to form a "Damascene" metal local interconnect (LI or M0) for wiring the top and bottom gates to adjacent circuits. This metal (second gate conductor) is deposited to overfill the trenches, planarized back to the pad layer, and plasma recessed slightly. This recess is then filled with a "cap" dielectric (e.g. silicon dioxide) and planarized. This cap prevents the subsequent contact etch from reaching the buried local interconnect metal. Finally, an interlevel dielectric is deposited and either plasma etched or Dual-Damascene metal one and contact one levels are patterned, etched, filled and planarized. The contact etch is selective, penetrating the pad dielectric in order to contact the top of the junctions, but not penetrating the trench cap dielectric. Thus the source/drain contacts are self-aligned with respect to the gate and isolation regions. Typical interconnect processing is used for all subsequent levels.

[0010] This invention describes the structure and fabrication process of a novel completely planar, DTMOS device structure which overcomes several key existing problems and extends the DTMOS advantages to both bulk and SOI sub-100nm CMOS technologies. This process permits easy incorporation of high-K gate dielectrics, such as tantalum pentoxide, and low resistance metal gates, such as copper, with a novel self-aligned, hyper-abrupt retrograde body implant. Current drive/speed, power-delay product, transconductance, output resistance, linearity, noise, hot-electron reliability, and device threshold matching are all superior to existing MOSFET designs. This invention's applications are far-reaching, but are particularly enabling for 0.25V-1.0V mixed-signal & RF CMOS SOC building block circuits. Between the large power supply reduction and superior analog/RF performance, order of magnitude power reductions can be achieved with this approach. This invention's application primary focus is on highly integrated CMOS

system-on-a-chip (SOC) implementations, which have the lowest possible energy consumption while using advanced spread-spectrum communication algorithms. This includes multi-GHz wireless transceivers. Applications span bulk and SOI deep-sub-micron (DSM) and sub-100nm CMOS technologies. It is expected that the application of this invention will permit single-chip digital pico-radios capable of supporting a wide variety of data types ranging from low bandwidth data, control, voice, and text messaging, up to full streaming video in a variety of pocket and handheld devices.

[0011] In one embodiment, a DTMOS structure comprises a silicon transistor in which the gate is constructed of two conducting layers and a capping insulating layer that are all embedded in a groove etched into the original SOI or epitaxial silicon layer. The result is a completely planar structure with low gate, body, and S/D resistances. Both the sequence of process steps and the resulting transistor structure are novel. The gate layers are all deposited after the high-temperature source/drain processing is complete, thus protecting them from possible high-temperature damage. A compact, low capacitance endwall (sometimes referred to as a sidewall) contact is formed between the top gate material and the self-aligned hyper-abrupt retrograde body. In one embodiment, this structure is built on a high-resistivity SOI wafer. In another embodiment, it is built on high-resistivity epitaxial bulk silicon wafers, permitting the typical triple well processing to be eliminated. These devices can also be integrated with conventional CMOS devices (non-DTMOS) on the same chip to achieve tradeoffs.

[0012] A novel zero-area, zero capacitance, low-resistance endwall contact is formed between the gate second layer and the buried low-resistance hyper-retrograde body region. The invention also provides a new hyper-abrupt retrograde well structure formed by counter-doping a pre-existing S/D implant only in and below the channel region, which is self-aligned to the channel and does not touch the source/drain junctions. This structure especially benefits bulk silicon implementation of DTMOS, achieving most of the benefits of SOI DTMOS implementations.

[0013] This invention further provides a method for eliminating conventional well processing (Well-less CMOS), while maintaining sufficient isolation, by using high-resistivity epitaxial isolation coupled with sub-volt biases (and optional VDD/2 epi substrate bias), thereby dramatically reducing the complexity/cost of bulk silicon

DTMOS implementations. This bulk silicon "semi-insulating" substrate design also decreases the cross-talk, noise coupling, and passive component Q losses typical of existing bulk CMOS processes.

[0014] Conventional DTMOS processes typically include a typical plasma-etched polysilicon gate process, which leads to a highly-non-planar gate stack, high-resistance gate stack, plasma etch damage to the gate oxide, depletion of the poly gate, boron penetration through the gate dielectric, and incompatibility with high-temperature-intolerant gate stack materials. This last problem is due to the source/drain implant thermal processing, which usually follows the gate formation, leading to this process flow being labeled "gate first". The processes disclosed herein includes a novel "recessed-channel" type of Damascene metal "gate-last" process that results in effectively raised source/drain regions for reduced S/D resistance. The processes also result in a novel dielectrically-capped, planar gate that permits self-aligned, borderless S/D contacts that will not short the gate even with complete overlap.

[0015] Additional aspects and advantages of this invention will be apparent from the following detailed description of preferred embodiments thereof, which proceeds with reference to the accompanying drawings.

Brief Description of the Drawings

[0016] FIG. 1A is a cross-sectional view a high-resistivity bulk starting wafer.

[0017] FIG. 1B is a cross-sectional view of a low-resistivity bulk starting wafer, with a high-resistivity epitaxial layer deposited on top.

[0018] FIG. 2A is a cross-sectional view of a high-resistivity SOI starting wafer.

[0019] FIG. 2B is a cross-sectional view of the high-resistivity SOI starting wafer shown in FIG. 2A after pad 20 has been deposited on top.

[0020] FIG. 3A is a plan view of the device after trenches have been etched into the top silicon layer.

[0021] FIG. 3B is a cross sectional view taken along the width of the device after trenches have been etched into the top silicon layer.

[0022] FIG. 3C is a cross sectional view along the length of the device after trenches have been etched into the top silicon layer.

[0023] FIG. 3D is an isometric view of the device taken along cutting line 3D after trenches have been etched into the top silicon layer.

[0024] FIG. 4A is a plan view of the device after a second silicon etch.

[0025] FIG. 4B is a cross sectional view taken along the width of the device after a second silicon etch.

[0026] FIG. 4C is a cross sectional view along the length of the device after a second silicon etch.

[0027] FIG. 4D is an isometric view of the device taken along cutting line 4D after a second silicon etch.

[0028] FIG. 5A is a plan view of the device after mask 3 has been implanted.

[0029] FIG 5B is a cross sectional view taken along the width of the device after mask 3 has been implanted.

[0030] FIG. 5C is a cross sectional view along the length of the device after mask 3 has been implanted.

[0031] FIG. 5D is an isometric view of the device taken along cutting line 5D after mask 3 has been implanted.

[0032] FIG. 6A is a plan view of the device after formation of sidewall spacers and after implant 119.

[0033] FIG. 6B is a cross sectional view taken along the width of the device after formation of sidewall spacers and after implant 119.

[0034] FIG. 6C is a cross sectional view along the length of the device after formation of sidewall spacers and after implant 119.

[0035] FIG. 6D is an isometric view of the device taken along cutting line 6D after formation of sidewall spacers and after implant 119.

[0036] FIG. 7A is a plan view of the device after the formation of a first gate conductor layer.

[0037] FIG. 7B is a cross sectional view taken along the width of the device after the formation of a first gate conductor layer.

[0038] FIG. 7C is a cross sectional view along the length of the device after the formation of a first gate conductor layer.

[0039] FIG. 7D is an isometric view of the device taken along cutting line 7D after the formation of a first gate conductor layer.

[0040] FIG. 8A is a plan view of the device after application of top gate mask (mask 4).

[0041] FIG. 8B is a cross sectional view taken along the width of the device after application of top gate mask (mask 4).

[0042] FIG. 8C is a cross sectional view along the length of the device after application of top gate mask (mask 4).

[0043] FIG. 8D is an isometric view of the device taken along cutting line 8D after application of top gate mask (mask 4).

[0044] FIG. 9A is a plan view of the device after the trenches have been refilled with isolation dielectric.

[0045] FIG. 9B is a cross sectional view taken along the width of the device after the trenches have been refilled with isolation dielectric.

[0046] FIG. 9C is a cross sectional view along the length of the device after the trenches have been refilled with isolation dielectric.

[0047] FIG. 9D is an isometric view of the device taken along cutting line 9D after the trenches have been refilled with isolation dielectric.

[0048] FIG. 10A is a plan view of the device after the isolation dielectric has been patterned by local interconnect mask (mask 5).

[0049] FIG. 10B is a cross sectional view taken along the width of the device after the isolation dielectric has been patterned by local interconnect mask (mask 5).

[0050] FIG. 10C is a cross sectional view along the length of the device after the isolation dielectric has been patterned by local interconnect mask (mask 5).

[0051] FIG. 10D is an isometric view of the device taken along cutting line 10D after the isolation dielectric has been patterned by local interconnect mask (mask 5).

[0052] FIG. 11A is a plan view of the device after the gate structure has been completed.

[0053] FIG. 11B is a cross sectional view along the width of the device after the gate structure has been completed.

[0054] FIG. 11C is a cross sectional view along the length of the device after the gate structure has been completed.

[0055] FIG. 11D is an isometric view of the device taken along cutting line 11D after the gate structure has been completed.

[0056] FIG. 12A is a plan view of the device following conventional contact and second metal processing.

[0057] FIG. 12B is a cross sectional view along the width of the device following conventional contact and second metal processing.

[0058] FIG. 12C is cross sectional view along the length of the device following conventional contact and second metal processing.

[0059] FIG. 12D is an isometric view of the device taken along cutting line 12D following conventional contact and second metal processing.

[0060] FIG. 13A is a plan view of device with contact holes misaligned.

[0061] FIG. 13B is a cross sectional view along the width of the device with contact holes misaligned.

[0062] FIG. 13C is a cross sectional view along the length of the device with contact holes misaligned.

[0063] FIG. 13D is an isometric view of the device taken along cutting line 13D with contact holes misaligned.

[0064] FIG. 14A is provided for comparison to FIG. 14A and depicts a cross sectional view along the width of the device that is identical to FIG. 12B.

[0065] FIG. 14B schematically depicts the terminal connections of the device.

[0066] FIG. 15A shows the floating body configuration of the device.

[0067] FIG. 15B schematically depicts the terminal connections of the device.

[0068] FIG. 16A shows the dynamic threshold configuration of the device.

[0069] FIG. 16B schematically depicts a cross sectional view of the device showing the terminal connections of the device.

[0070] FIG. 17A depicts a cross sectional view of substrate of embodiment 100''.

[0071] FIG. 17B and FIG. 18 are cross sectional views depicting embodiment 100'' before and after etching.

[0072] FIG. 19A depicts a cross sectional view of the device after completion of the step of depositing and etching sidewall spacer for embodiment 100''.

[0073] FIG. 19B depicts a cross sectional view of the device after completion of the step of directing complementary, counter-doping ion implants for embodiment 100''.

[0074] FIG. 19C depicts a cross sectional view of the device after the step of directing complementary, counter-doping ion implants for embodiment 100''.

[0075] FIG. 20 depicts a cross sectional view of the device after the step of depositing a gate dielectric for embodiment 100''.

[0076] FIG. 21 depicts a cross sectional view of the device after the step of plasma etching first gate conductor for embodiment 100''.

[0077] FIG. 22A depicts a cross sectional view along the width of the device after the step of filling isolation trenches along the channel length direction of embodiment 100''.

[0078] FIG. 22B depicts a cross sectional view along the length of the device the step of filling isolation trenches along the channel width direction of embodiment 100''.

[0079] FIG. 23A depicts a cross sectional view along the width of the device after the step of etching oxide down to BOX layer 13.

[0080] FIG. 23B depicts a cross sectional view along the length of the device after the step of etching oxide down to BOX layer 13.

[0081] FIG. 24A depicts a cross sectional view of the Damascene double gated transistor taken along the channel length direction.

[0082] FIG. 24B depicts a cross sectional view of the Damascene double gated transistor taken along the channel width direction.

[0083] FIG. 25A is a cross sectional view of another embodiment of the Damascene double gated transistor.

[0084] FIG. 25B depicts the doping profile in the channel and body/well of embodiment shown in FIG. 25A.

Detailed Description of Preferred Embodiments

[0085] In describing an embodiment of the present invention, reference will be made in the drawings to elements of an embodiment by numerals that refer to like elements in another embodiment of the invention. Note that features of the invention are not necessarily shown to scale in the drawings. The device is described and drawn in an N-channel configuration, but the process provides complementary implants to simultaneously build P-channel devices as well.

[0086] Several methods are disclosed herein for forming Damascene double gated transistors. Before explaining the steps of the method shown in FIGS. 2-12 that are used to achieve the final structure, it is useful to appreciate the configuration of the final structure shown in FIGS. 11A-D and FIGS. 12A-D. Another method is depicted in FIGS. 17-24 with the final structure shown in FIGS. 24A-24B. FIG. 25A depicts an additional embodiment.

[0087] FIGS. 1A-1B and 2A depict some of the various substrates that can be used in accordance with the inventive methods and structures. Substrate 10 shown in FIG. 1A is bulk silicon 11 that preferably has a high resistivity such as 1k ohm-cm. The resistivity may range from 100-10k ohm-cm. Alternatively, a triple-well bulk silicon substrate may be used instead. Substrate 10' shown in FIG. 1B has an epitaxial layer 12 on bulk silicon 11' which has a low resistivity. The low resistivity of silicon 11' may range from 0.1 to 10 ohm-cm.

[0088] FIG. 2A depicts a starting silicon-on-insulator (SOI) substrate 10''. Layer 11 is the silicon substrate. Substrate 10'' has a layer 13 that is under a SOI layer or wafer 14 and above silicon 11. Layer 13 is a buried insulator (typically silicon dioxide referred to as a BOX layer). The thickness of BOX layer 13 may vary from several microns down to 100nm, however, the thickness is preferably in a range from about 100 nm to about 500 nm. Layer 14 is the thin crystalline silicon layer where devices are built. Layer or SOI wafer 14 may be formed by any desirable available method such as SIMOX, SmartCut bonded, Metal-induced poly recrystallization, etc. Layer 14 preferably has a high resistivity of 1k ohm-cm. The initial crystalline SOI film thickness may be any available thickness but is preferably either thinned by etching or thickened by epi-growth to a range from about 200 nm to about 300 nm. Note that substrate 10'' is particularly preferred for a RF-CMOS application.

[0089] Another substrate 10''' (not shown) may also be utilized that has four layers. The surface layer is an epitaxial layer 12'. Epitaxial layer 12' preferably has a thickness in a range from about 100 nm to about 200 nm. Epitaxial layer 12' is on a layer of silicon 11' that has a low resistivity and preferably a thickness of about 100 nm. Silicon layer 11' is on BOX layer 14. BOX layer 14 preferably has the same thickness range as BOX layer 14 described above in reference to substrate 10''. BOX

layer 14 is on a bottom layer of silicon 11' that has low resistivity, preferably about 10 ohm-cm.

[0100] As shown in FIG. 2B, once an appropriate substrate such as substrate 10'' has been obtained then the pad 20 is formed. Pad 20 preferably includes a silicon nitride layer 22 over a thin silicon oxide stress relief layer 24. Pad 20 is the silicon etch hard mask and the chemical-mechanical polish (CMP) stop layer. Silicon nitride (Si_3N_4) layer 22 may have a thickness ranging from about 50 nm to about 200nm. Layer 22 may be deposited or grown to be used as both a silicon etch hardmask and as a chemical-mechanical polish (CMP) stopping layer without an oxide layer. The oxide layer 24 is a very thin layer of thermal oxide 24. For example layer 24 may have a thickness of about 20 to about 40 nm.

[0090] Pad 20 is formed from a different material than the dielectric materials that surround the gate, namely spacers 30 and cap 150, which are discussed below. For a pad that comprises a silicon nitride layer 22 over a thin silicon oxide stress relief layer 24, spacers 30 and cap 150 comprise silicon dioxide. Conversely, for a pad 20 that comprises merely silicon dioxide, spacers 30 and cap 150 comprise silicon nitride.

[0091] FIGS. 3-12 provide views of the various steps involve in forming the final Damascene double gated transistor referred to herein as device 100 throughout the description of the process. At each step, four views are provided including a plan view referenced by A, a cross-sectional view taken along the width of the device referenced by B, a cross-sectional view taken along the length of the device referenced by C and an isometric view of the device along cutting line 3D in FIG. 3A.

[0092] FIGS. 3A-D provide various views of the device after the first masking/ patterning step has been completed utilizing mask 1. Merged gate trench 110 and isolation trench 112 are etched into the top silicon layer, leaving a 50-200 nm thick silicon layer 14 remaining on BOX layer 13. Note that the different portions of isolation trench 112 are individually identified as 112a-d. This mask is the union shape of typical CMOS Active Area & Gate masks. The silicon substrate layer 11 and BOX layer 13 are omitted from the subsequent drawings for clarity, however, it should be understood that layer 14 remains on BOX layer 13 which is on silicon substrate layer 11.

[0093] FIGS. 4A-4D show the result of the second mask identified as mask 2, which protects the channel region 120 from a second silicon etch. This second etch removes the rest of the silicon from the bottom of the trenches 112 in the isolation regions around the channel, source, and drain, stopping on the BOX layer 13.

[0094] FIGS. 5A-D shows the results of two sets, one deep 117 and one shallow 118, of masked, complementary ion implants utilizing mask 3. These implants are used to form the source/drain junctions or regions and a channel region 120. As discussed below, channel region 120 has a counter doped channel layer or portion 122 and a remaining bottom intrinsic layer or portion 124. The n+ regions correspond with the source and drain pedestals respectively identified in FIG. 5D as 114s and 114d. The p+ regions correspond with a surrounding p+ source/drain pedestal identified in FIG. 5D as 116. Note that p+ source/drain pedestal 116 belongs to adjacent p+ devices so it is not a part of device 100.

[0095] The deep n+ implant 117 may be a series of implants of various energies and doses in order to form a uniformly, heavily doped source/drain all the way down to the BOX layer 13. The counter doped channel layer 122 has a very abrupt profile in order to leave the bottom intrinsic layer 124 undoped, so heavy ion species are used such as Indium and Arsenic. Since BOX layer 13 and pad 20 are dielectrics, they permit use of these "shallow/deep" pair of implants. More particularly, because these regions are dielectric the dopants are implantable into these regions without significantly affecting the device structure or operation.

[0096] FIGS. 6A-6D shows the formation of sidewall spacers 30 followed by a high-dose, relatively high-energy p+ implant 119, such as Indium, that forms the buried low-resistance bottom gate 124' of device 100. Note that sidewall spacers 30 comprise silicon dioxide in this embodiment. It should be noted that the bottom gate 124' is self-aligned to the bottom of the gate trench 110 between the two spacers 30 and these spacers 30 prevent bottom gate 124' from abutting either the heavily-doped source or drain regions or pedestals 114s and 114d, leaving regions of bottom intrinsic layer 124 to the left and right of the bottom gate 124' that are undoped or at least lightly doped. These intrinsic regions of bottom intrinsic layer 124 are identified with the letter i.

[0097] FIGS. 7A-7D depict the formation of a first gate conductor layer 130 over channel region 120. A gate dielectric layer 128, preferably having a high dielectric constant, such as hafnium oxide, underlies first gate conductor layer. Additionally, it is preferable to form a thin undoped selective epitaxial silicon layer on top of the counter-doped channel layer 122 to enhance the mobility of the device as shown in FIGS. 7B-7C at 126 before forming the gate dielectric layer 128 and first gate conductor layer 130. After the selective epitaxial silicon layer 126 is formed, all high temperature steps are complete, meaning that temperature-sensitive materials, such as hafnium oxide and copper may respectively be used for the gate dielectric and conductor material.

[0098] First gate conductor layer 130 may be any suitable conductor material that enables the desired threshold voltage. Examples of suitable materials include tungsten, polysilicon, poly-SiGe or a silicide. As shown in FIG. 7D, the trenches 110 and 112 are overfilled with the material that forms the first gate conductor material 130. The materials is then planarized back to the pad 20, and then recessed, preferably by plasma etching, leaving about 50-200 nm of a first gate conductor layer 130 over the channel region 120.

[0099] FIGS. 8A-8D shows the results of the top gate mask, identified as mask 4, being applied to the device. The entire channel 110 and gate stack is patterned and etched, stopping on the bottom gate layer 124'. Mask 4 may fully or partially cover the bottom gate mask, which is mask 3. Additionally, mask 4 may be offset to the left or right of the bottom gate mask, simultaneously creating four-terminal as discussed below in reference to FIGS. 14A-B, floating body as discussed below in reference to FIGS. 15A-B, and dynamic threshold devices as discussed below in reference to FIGS. 16A-B. The final device length and width are respectively identified in FIG. 8B at l and in FIG. 8C at w.

[00100] FIGS. 9A-9D shows the trenches 110 and 112 refilled with isolation dielectric 40. Isolation dielectric is preferably silicon dioxide. After the trenches 110 and 112 have been filled then the isolation dielectric 40 is planarized, as shown.

[00101] FIG. 10A-10D shows isolation dielectric 40 after it has been patterned by the local interconnect (LI or M0) mask identified as mask 5. Mask 5 is used to remove portions of isolation dielectric 40 in the isolation trenches 112.

[00102] FIGS. 11A-11D show the completed gate structure. The second gate conductor material is deposited in a manner such that the trench is overfilled. After being overfilled, the second gate conductor material is planarized and then recessed. The second gate conductor material is preferably a metal. The resulting second gate conductor layer 140 serves as a first interconnect or metal level. A first metal level is used to locally wire the bottom and top gates to adjacent circuits and to create a zero-area DTMOS top gate-to-bottom gate connection. A cap 150 is then formed. Cap 150 is formed by refilling recessed top portions of the second gate conductor layer 140 with a dielectric such as silicon dioxide. Cap 150 or capping layer prevents contacts from subsequently shorting to the top of the M0 metal, second gate conductor layer 140.

[00103] FIGS. 12A-12D show device 100 following conventional contact and second metal processing. An interlevel dielectric 210 also identified as ILD is deposited, followed by contact patterning and etching through this ILD 210 and through the pad dielectric 20. This etch does not etch the cap 150 or isolation dielectric 40. The contact holes 220 are then filled with metal 222 and planarized, followed by deposition and patterning of second metal 224. Alternatively, Dual Damascene processing could be used for contact and second metal processing.

[00104] FIG. 12A-D shows the final completed device structure in a four-terminal configuration where the bottom gate is wired out to the left and the top gate is wired out to the right, the source is wire to the bottom and the drain is wire to the top. This configuration enables all four terminals 230, 232, 234 and 236 to be biased independently.

[00105] FIGS. 13A-13D show that even if contact holes 220 are misaligned the contact holes still provide a via between the second level of metal 224 and the pedestals 114s or 114d. Shorting is avoided as there is no connection to second gate conductor layer 140, metal M0, due to the isolating configuration of cap 150.

[00106] FIG. 14A is identical to FIG. 12B and is included for convenient comparison with FIG. 14B which schematically depict the terminal connections of the device. Note that the configuration shown in FIGS. 14A and 12B is enabled through the offset positioning of masks 3 and 4.

[00107] FIG. 15A shows the floating body configuration of this device, where the bottom gate is electrically isolated by the spacers and gate dielectric from the top gate or M0. Top gate encloses the bottom gate due to the use of top gate mask that is larger than the bottom gate mask. FIG. 15 B schematically depicts the terminal connections of the device.

[00108] FIG. 16A shows the dynamic threshold (DTMOS) configuration of this device, where the top and bottom gates are shorted at both the left and right sides of the channel. Bottom gate encloses the top gate due to the use of a bottom gate mask that is larger than the top gate mask. FIG. 16 B schematically depict the terminal connections of the device.

[00109] FIGS. 17-24 depict another embodiment for forming a Damascene double gate transistor. FIG. 17A depicts a substrate like that shown in FIG. 2B with a pad 20 on substrate 10''. In contrast to the embodiment described above in reference to FIGS. 2-12, the source/drain implants are formed before etching the trenches. The doping may be followed by rapid thermal activation. The depth of these top-side junctions can be adjusted as needed for subsequent silicidation and sheet resistance, without impacting the short-channel behavior of the device. Since these source and drain regions will effectively be "raised" with respect to the channel, they may be made thicker and lower in resistance than typical CMOS processes.

[00110] FIG. 17B depicts device 100'' after gate trench 110 and isolation trench 112 have been formed. This first photolithography level defines the source and drain regions of the device. The etching is preferably a silicon plasma etch. About 100nm of SOI film is preferably left under these trenches, as shown in the sideview in Fig. 17B.

[00111] FIG. 18 shows the formation of a doped channel layer or portion 122'. Doped channel layer 122' may be formed by any suitable method. For example, layer 122 may be formed through the use of large angle tilted implantation or alternatively through the use of dummy doping with a sacrificial layer that is subsequently removed. Preferably, complementary high-dose, low-energy, large-angle-tilted ion implants are performed to form the trench sidewall and bottom S/D junctions. Indium can be used for P-channel and arsenic for N-channel devices, followed by rapid thermal activation. Sidewall junctions are preferably made shallow to control short

channel effects. An optional thin sacrificial oxide can be performed to remove any plasma etch or implantation damage from the bottom of the trench where the channel will be and to prevent any stress damage from the coming nitride sidewall spacers 30.

[00112] As shown in FIG. 19A, nitride sidewall spacers 30 are then deposited and etched, providing an opportunity for optional halo (not shown) and S/D extension implants before and/or after the spacer formation. Nitride sidewall spacers 30 are preferably about 20 nm thick. Spacers 30 also protect the S/D regions from subsequent shorting to the gate during the gate to body contact process. Anisotropic etching may be used to remove the Si_3N_4 from the top surface and the bottom of the gate. As shown in FIG. 19B, complementary, counter-doping ion implants are preferably then directed into the open bottom of gate trench 110 between the sidewall spacers 30. These high-dose, higher-energy implants (preferably indium for N-channel and preferably arsenic for P-channel) compensate the previous shallow S/D implants, resulting in low net doping down into the body in a range of about 30 nm to about 50 nm where the inversion channel will be located, with a hyper-abrupt transition to high concentration in the bottom portion of the body where the body contact will be formed and body current conducted. The resulting configuration is shown in FIG. 19C with the p+ region indicated at 124'. The resulting heavily doped body "stripe" is self-aligned to the future gate "stripe" directly above it but not underneath the S/D regions, thus minimizing capacitance.. These implants are rapid thermal annealed and the sacrificial oxide is removed from the trench bottoms. Note that the rapid thermal annealing is the last high temperature step in the process. The counter-doping can be achieved by doping from ion implantation with or without tilt for symmetrical or asymmetrical structures. Counterdoping can also be achieved through solid source doping from a sacrificial layer.

[00113] As previously indicated, the implant is preferably indium for the N-channel and preferably arsenic for the P-channel devices. The resulting doping profile in the channel and body/well is shown in FIG. 25B, as discussed below. Advantages of the hyper-abrupt retrograde profile below the channel achieved through the critical body counter-doping ion implantation are explained below after the overall method has been fully set forth.

[00114] Next a gate dielectric 128 is deposited as shown in FIG. 20. Any appropriate gate dielectric may be used including temperature-sensitive high-k dielectrics such as hafnium oxide. Deposition of the gate dielectric is followed by a trench-filling deposition of the first gate conductor material 130. First gate conductor material 130 is preferably either doped or undoped polysilicon or a polysilicon-germanium alloy where the percent germanium content is used to adjust the gate workfunction and therefore the device threshold voltage. First gate conductor material 130 may also be a metal such as tungsten or titanium nitride. Chemical-mechanical polish (CMP) of this layer back to the pad then follows.

[00115] First gate conductor material 130 is plasma-etch recessed to about 30 nm in the trench bottom, followed by an optional temporary oxide trench fill deposition and CMP, as shown in FIG. 21.

[00116] Isolation trenches 112 are then filled with oxide 40, as is the region above first gate conductor material 130, and planarized back to nitride pad 20, resulting in the sideviews shown in FIGS. 22A. FIG. 22A depicts the channel length direction, while FIG. 22B shows the channel width direction, including the endwall where the first gate conductor material 130 will be shorted to the retrograde body implant 122 by the upcoming second gate conductor material 140.

[00117] Next, M0 (first level, local interconnect) lithography is performed utilizing photoresists PR, followed by etching a trench through oxide 40 to the bottom of the STI and trench cap oxides at the same time all the way to BOX layer 14 when the substrate is the SOI embodiment, substrate 10''). FIGS. 23A-23B depict the resulting configuration in views that are respectively, the channel length direction and the channel width direction. FIG. 23B depicts the removal of oxide 40 down to BOX layer 13. FIG. 23A depicts the removal down to first gate conductor material 130. This M0 trench is filled with the second gate material 140 (by deposition of a low-resistance metal such as copper or other suitable material, and then planarizing by a process such as CMP. Second gate material 140 is positioned to short the top and ends of the first gate conductor material 130 to the buried heavily doped body at both endwalls of the device, but does not short to the S/D implants due to the protecting nitride sidewall spacers. This second gate conductor material 140 is plasma recessed to a thickness of about 30 nm and then capped with an oxide fill 150 and then

planarized. The resulting final transistor structure is completely planar, as shown in FIG. 24A. The zero-area, zero-capacitance endwall gate-body DTMOS connection is identified in FIG 24B at 70.

[00118] A standard CMOS multi-level, dual-Damascene metal interconnect process is then completed, like that shown in FIGS. 12A-12D. Notice that the spacing of S/D metal contacts to the gate and to each other is reduced compared to prior art device. This condensed configuration increases layout density without causing S/D to gate shorts due to the tolerance of the structure to misalignment. These contacts can be "borderless" since they are self-aligned to the gate cap insulator. This improves the density of memory cells such as compact DRAM, SRAM, and FLASH cell designs.

[00119] Advantages of the unique connection structure and the hyper-abrupt retrograde well are described below. Other advantages of the method disclosed above should also be appreciated. Most DTMOS processes have included a typical plasma-etched polysilicon gate process, which leads a gate stack that is highly non-planar and has high resistance. Such processes also result in depletion of the poly gate and boron penetration through the gate dielectric. Further such processes result in incompatibility with high temperature intolerant gate stack materials due to the source/drain implant thermal processing, which usually follows the gate formation. Due to this sequence, such methods are referred to as "gate first" processes. Such processes involve building a dummy poly gate and sidewall spacers, performing all S/D implants and anneals, planarizing the structure with a CMP polish insulator, removing the dummy gate, depositing a high k gate dielectric and a low resistance gate metal, and then planarizing the surface to leave a Damascene metal gate in the groove where the dummy poly gate had been. Disadvantages of such process include a high S/D resistance due to shallow S/D junctions or the S/D contacts being non selfaligned to the gate. The first problem is particularly relevant to sub-100 nm devices and the second problem is particularly relevant to compact DRAM, SRAM and FLASH cells. In contrast, the present invention is a Damascene metal "gate last" process that results in effectively raised source/drain regions for reduced S/D resistance. The inventive process also provides self-aligned, borderless S/D contacts that will not short the gate even with a complete overlap.

[00120] Another advantage of the inventive method is the elimination of the wells typically formed bulk silicon by conventional implementations of DTMOS. Current DTMOS implementations require a "triple-well" structure with a shallow N-well, shallow P-well and a deep N- well to isolate individual body nodes from each other. The present method eliminates this well processing and is accordingly referred to as "Well-less CMOS". The well-less CMOS methodology results in sufficient isolation by using high resistivity epitaxial isolation coupled with sub-volt biases and optional VDD/2 epi substrate bias. The novel bulk silicon "semi-insulating" substrate design will also decrease the cross-talk, noise coupling, and passive component Q losses typical of existing bulk CMOS processes.

[00121] FIG. 25A is a side view of a transistor that has a counter doped region that was formed by a different method than that set forth above in relation to FIGS. 17-24. The surface of the substrate is first doped and the trench is formed on top of the substrate by depositing the insulative material around a dummy material that is subsequently removed. After removal of the dummy material then the spacers are deposited in preparation for counterdoping the bottom of the trench. After the counterdoping has been achieved then the subsequently layers can be deposited. Note that selective epitaxy, counter-doping , and indium/arsenide retrograde doping of the body region below the channel have all been proposed/demonstrated in order to achieve a hyper-abrupt retrograde well profile such as the so called "Super Steep Retrograde (SSR) and Delta-Doped (DD) profiles that will optimize/maximize the body effect, while at the same time providing a buried low-resistance body path. However, these reported methods cause excessive junction capacitance and/or leakage when the retrograde well extends under the source/grading (S/D) junctions (in bulk CMOS) or butts against them in SOI CMOS). The novel hyper-abrupt retrograde well structure formed by counter-doping a preexisting S/D implant only in and below the channel region, enables the channel to be self-aligned and prevents contacts with the source/drain junctions. Such novel structures are especially beneficial to bulk silicon implementations as it achieves most of the benefits of SOI DTMOS implementations.

[00122] As mentioned above, FIG. 25B illustrates the hyper-retrograde doping profile achieved in the substrate and the depth of the different layers of the transistor above the counterdoped region. FIG. 25B is shown as depicting the profile for the

device shown in FIG. 25A, however, the same profile results for the devices shown in FIGS. 11B and 24A.

[00123] It will be obvious to those having skill in the art that many changes may be made to the details of the above-described embodiments of this invention without departing from the underlying principles thereof. The scope of the present invention should, therefore, be determined only by the following claims.

Claims

1. A method for manufacturing a transistor comprising the steps of:

obtaining a silicon substrate having a high resistivity surface,

forming a layer on the surface of the silicon substrate adapted for use as a planarizing stop layer,

doping the surface of the silicon substrate to obtain source and drain regions,

creating doped channel regions,

forming gate trenches in the silicon substrate, each gate trench having sidewalls and a bottom,

forming spacers on the sidewalls of each gate trench,

depositing a thin layer of a gate dielectric in the gate trenches,

partially filling the gate trenches with a first gate conductor material,

performing gate photolithography while masking at least a portion of at least one gate trench to protect an intrinsic device region such that the first gate conductor material is removed from the gate trenches around the intrinsic device region and the surface of the silicon substrate is exposed,

filling the gate trenches around the intrinsic device region with silicon dioxide,

shorting the top and ends of the of the first gate conductor material in the gate trench in the intrinsic device region to the doped surface of the silicon substrate by depositing a second gate conductor material in the gate trench in the intrinsic device region such that it covers the top and ends of the first gate conductor material.

2. A method for manufacturing a transistor comprising the steps of:

obtaining a silicon substrate having a high resistivity surface,

forming a layer on the surface of the silicon substrate adapted for use as a planarizing stop layer,

doping the surface of the silicon substrate to obtain source and drain regions,

forming gate trenches in the silicon substrate, each gate trench having sidewalls and a bottom,

forming spacers on the sidewalls of each gate trench,

counter doping the bottoms of the gate trenches,

depositing a thin layer of a gate dielectric in the gate trenches,

partially filling the gate trenches with a first gate conductor material,

performing gate photolithography while masking at least a portion of at least one gate trench to protect an intrinsic device region such that the first gate conductor material is removed from the gate trenches around the intrinsic device region and the surface of the silicon substrate is exposed,

filling the gate trenches around the intrinsic device region with silicon dioxide,

shorting the top and ends of the of the first gate conductor material in the gate trench in the intrinsic device region to the doped surface of the silicon substrate by depositing a second gate conductor material in the gate trench in the intrinsic device region such that it covers the top and ends of the first gate conductor material.

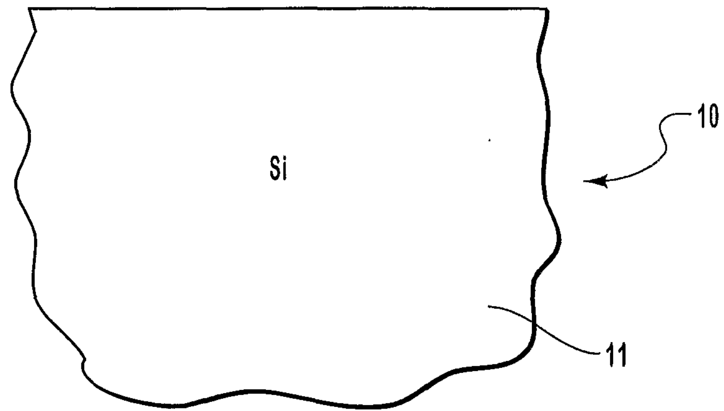


FIG. 1A

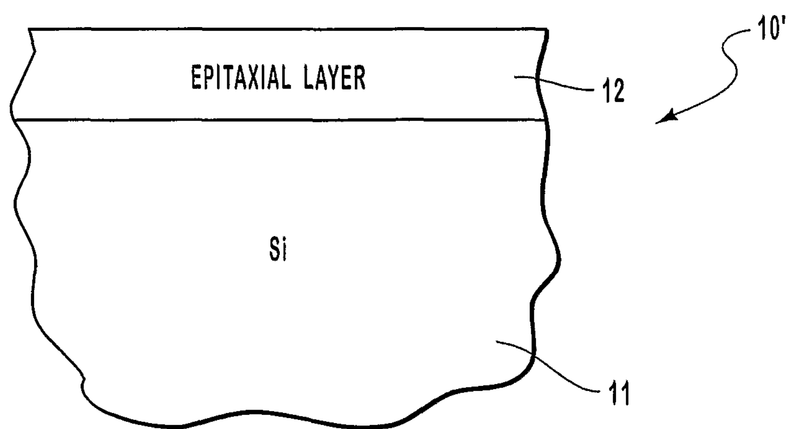


FIG. 1B

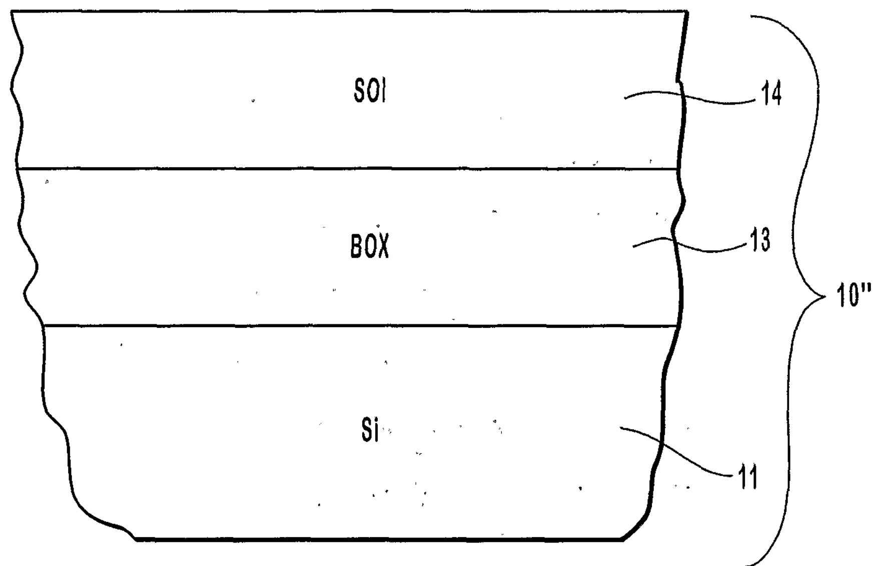


FIG. 2A

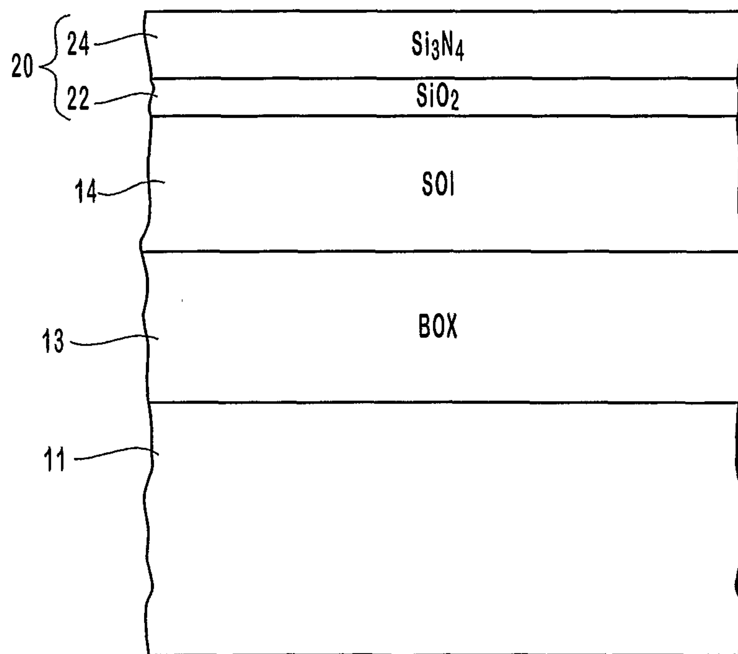


FIG. 2B

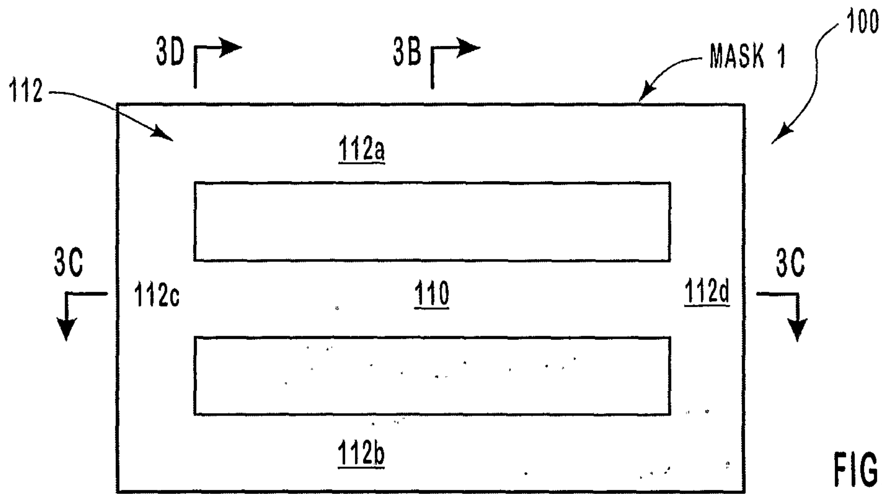


FIG. 3A

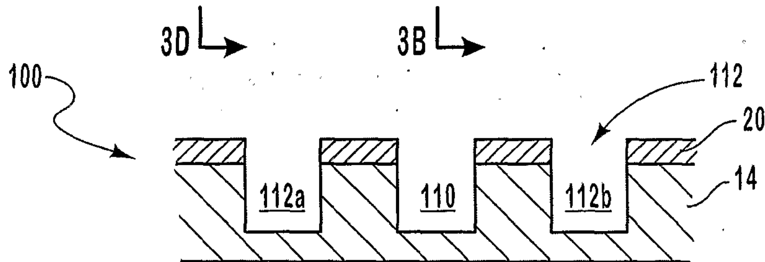


FIG. 3B

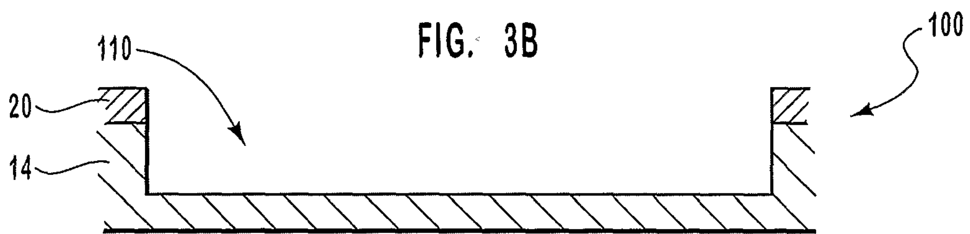


FIG. 3C

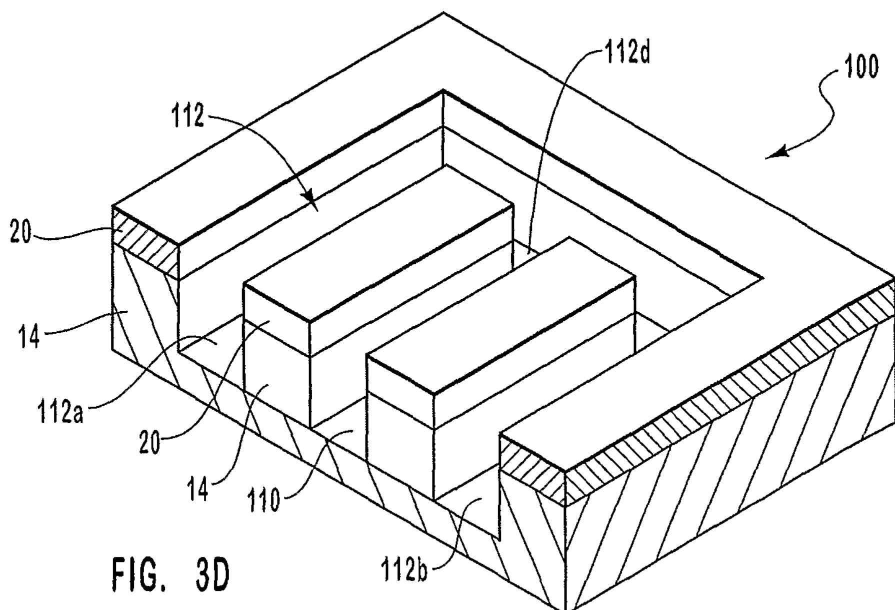


FIG. 3D

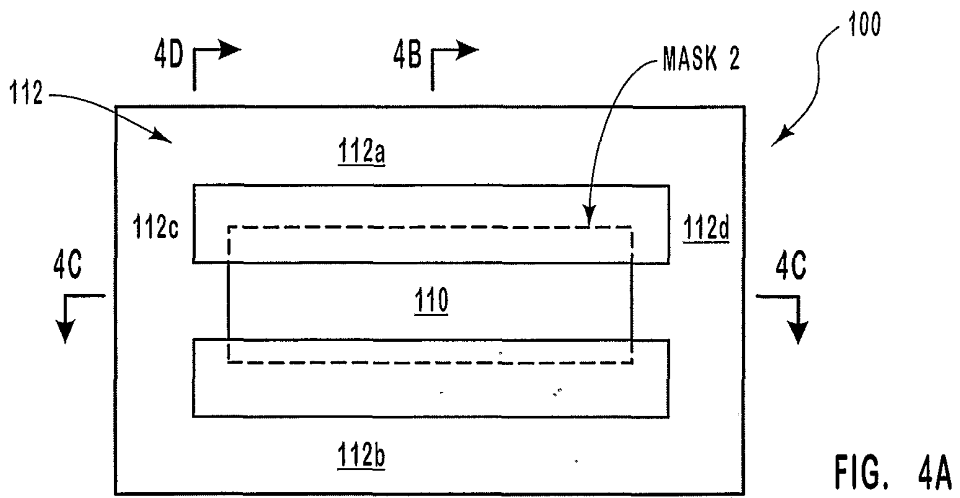


FIG. 4A

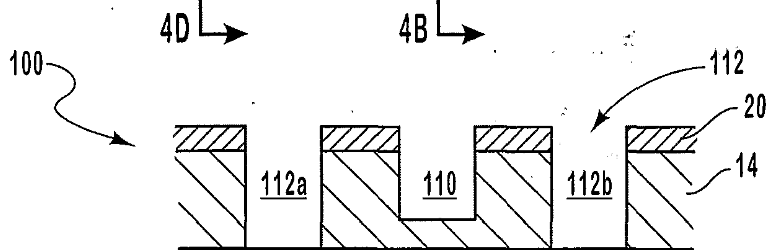


FIG. 4B

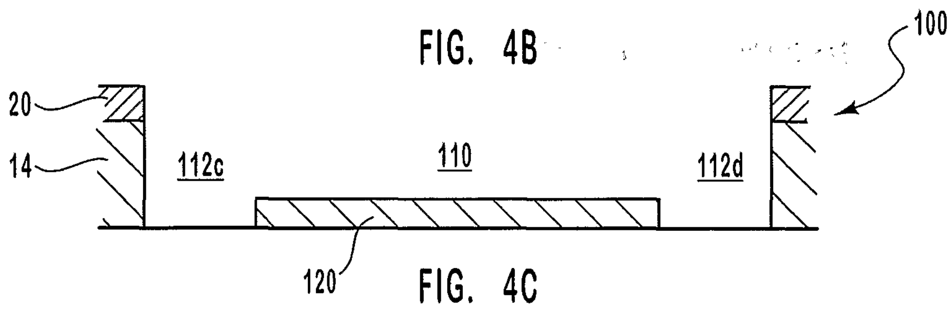


FIG. 4C

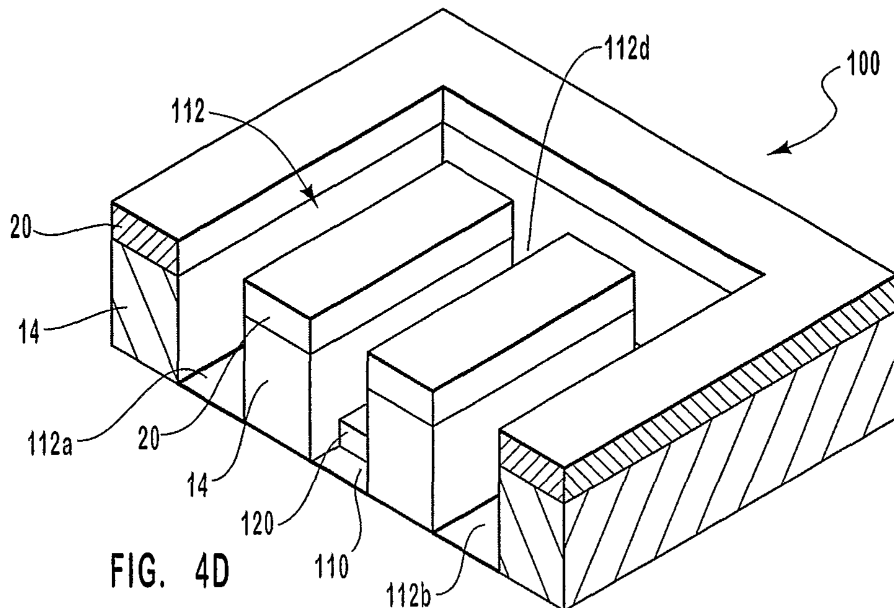


FIG. 4D

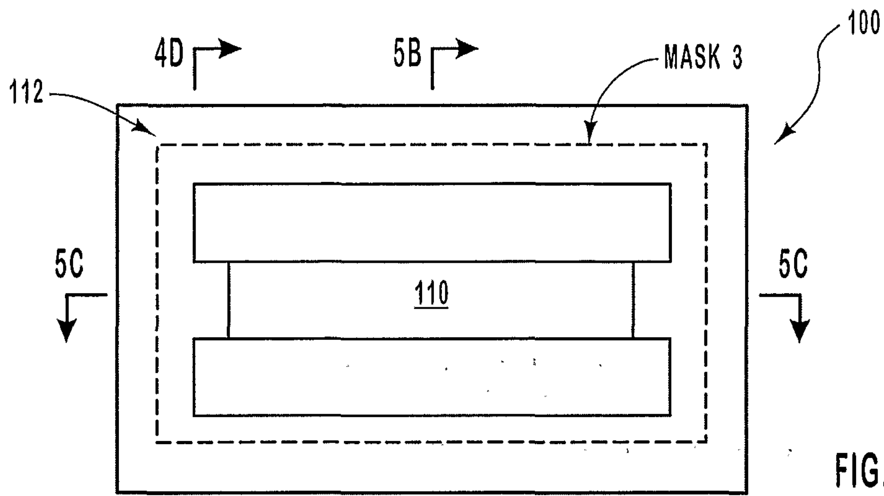


FIG. 5A

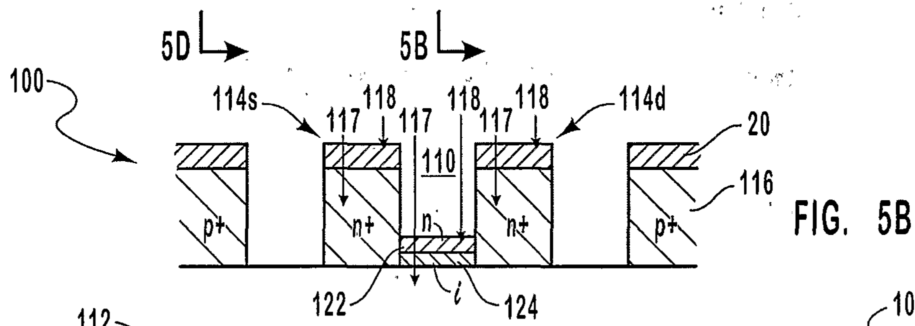


FIG. 5B

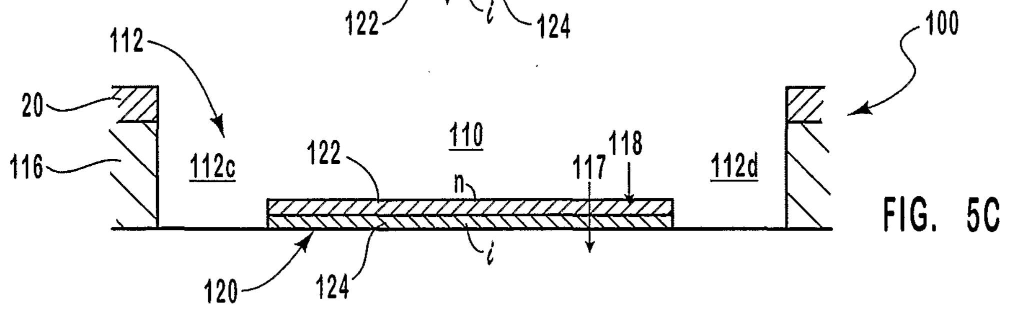


FIG. 5C

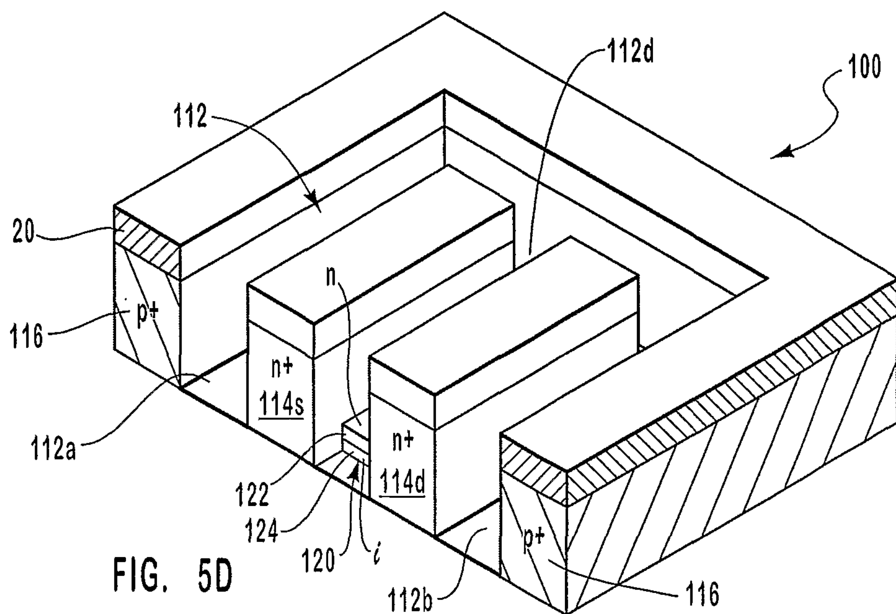


FIG. 5D

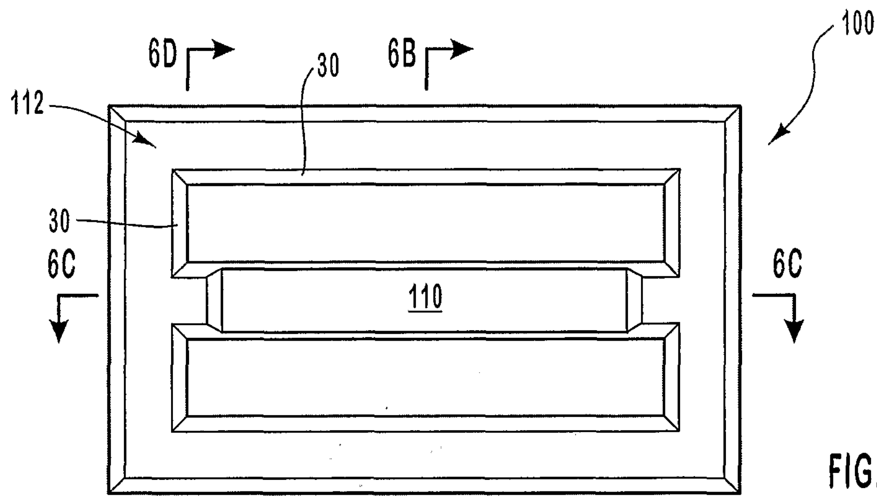


FIG. 6A

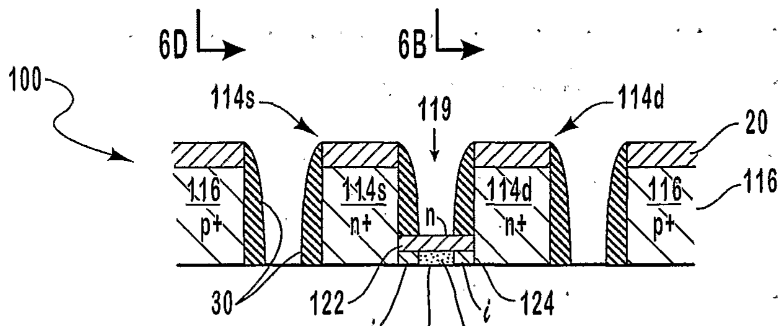


FIG. 6B

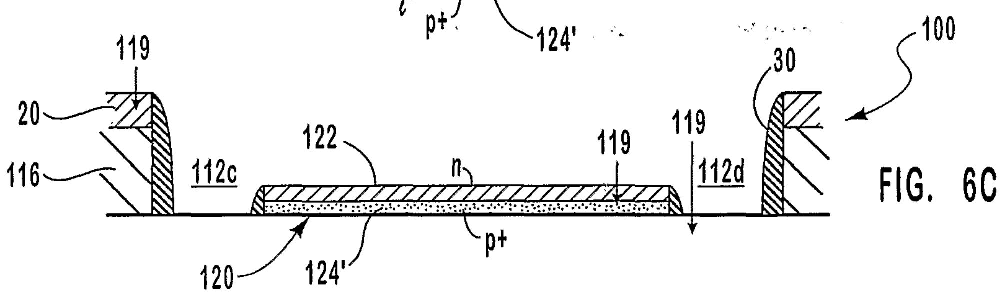


FIG. 6C

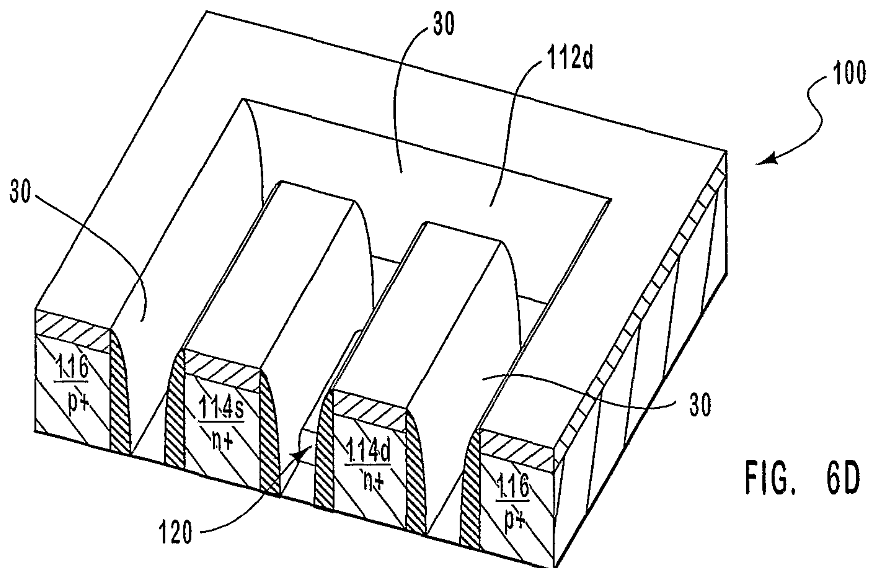


FIG. 6D

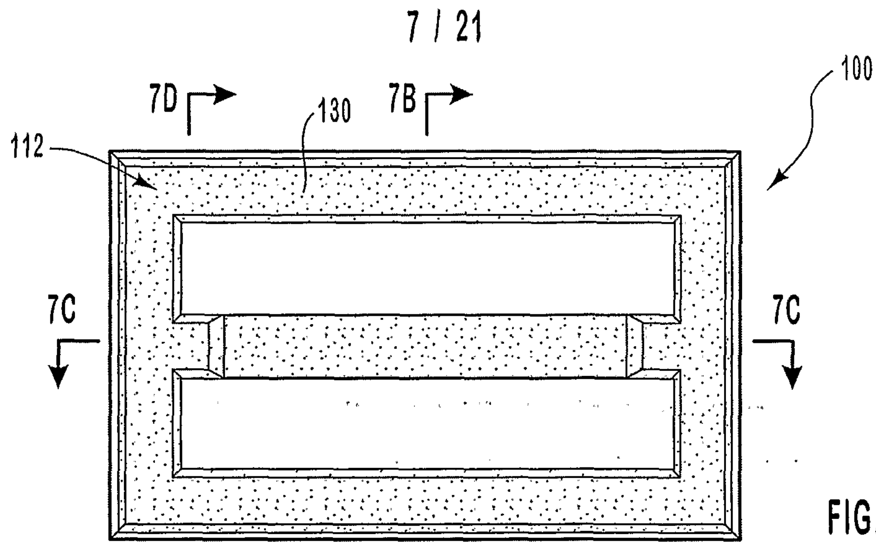


FIG. 7A

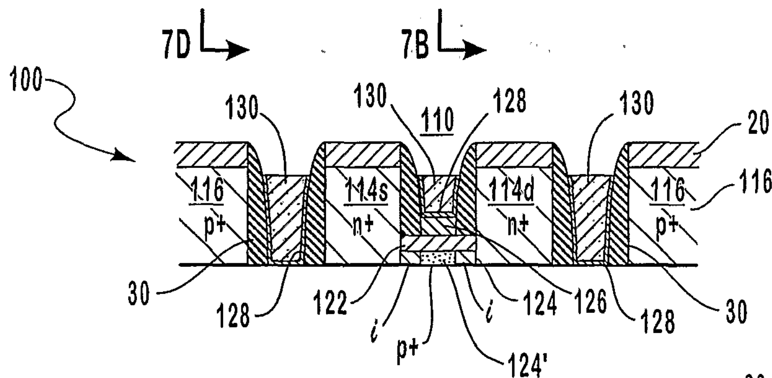


FIG. 7B

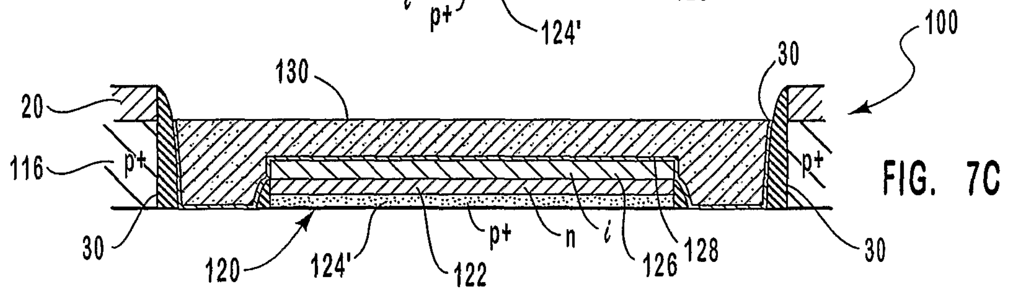


FIG. 7C

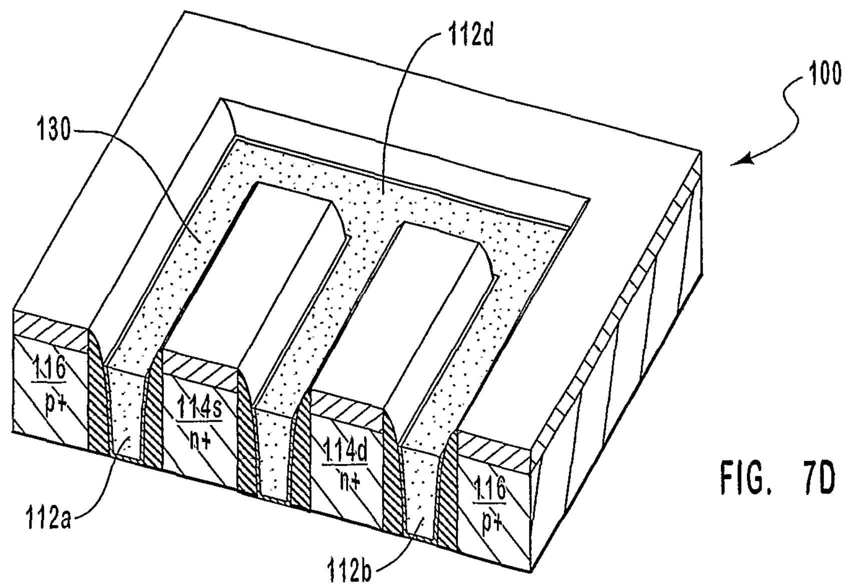


FIG. 7D

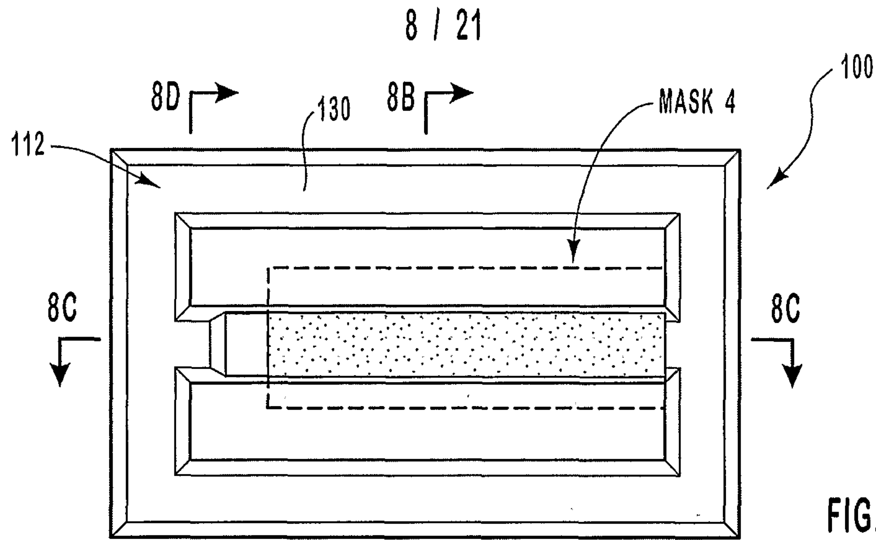


FIG. 8A

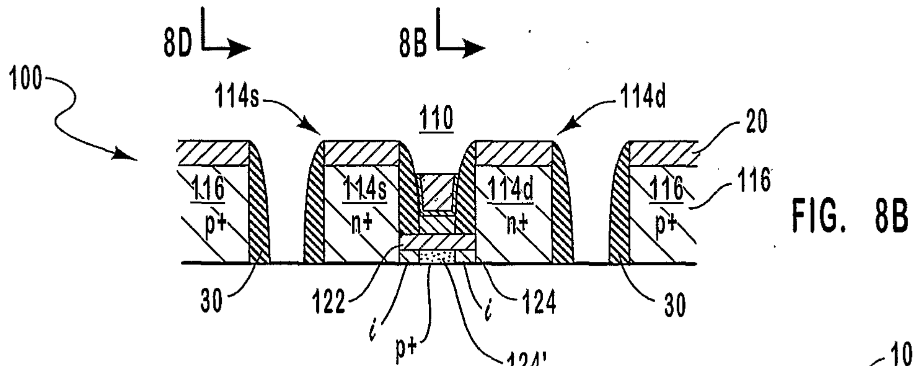


FIG. 8B

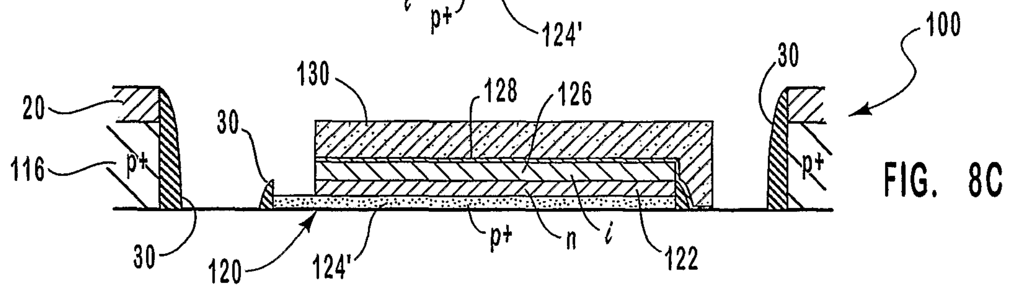


FIG. 8C

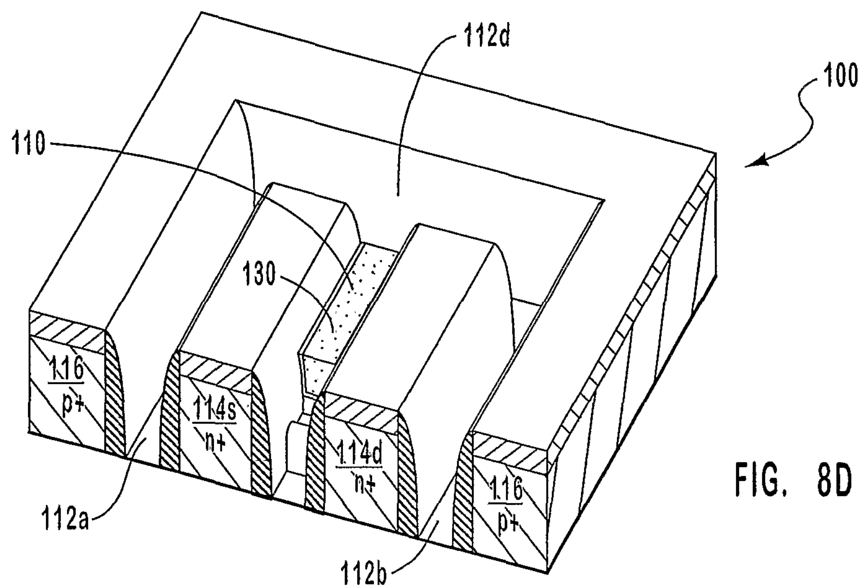


FIG. 8D

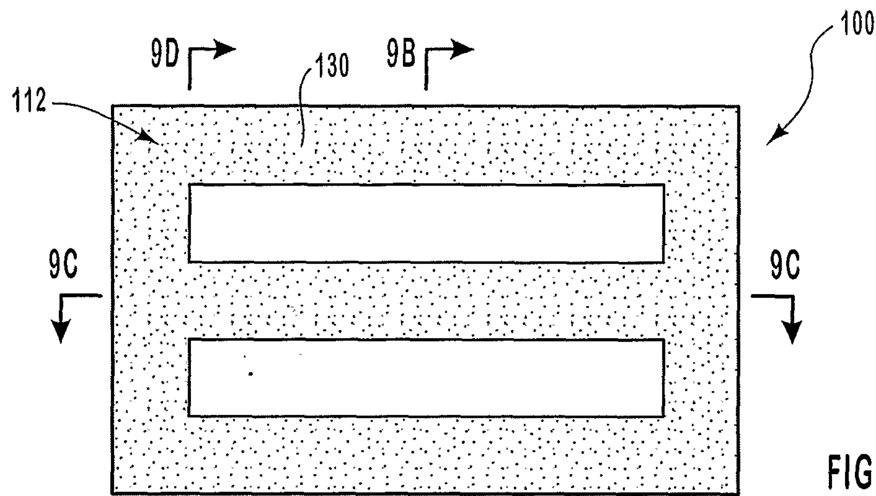


FIG. 9A

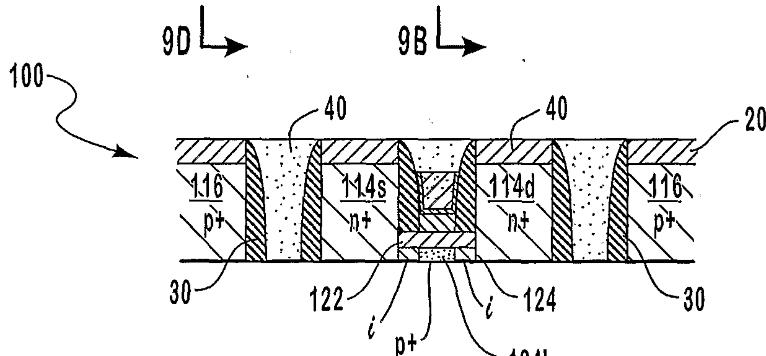


FIG. 9B

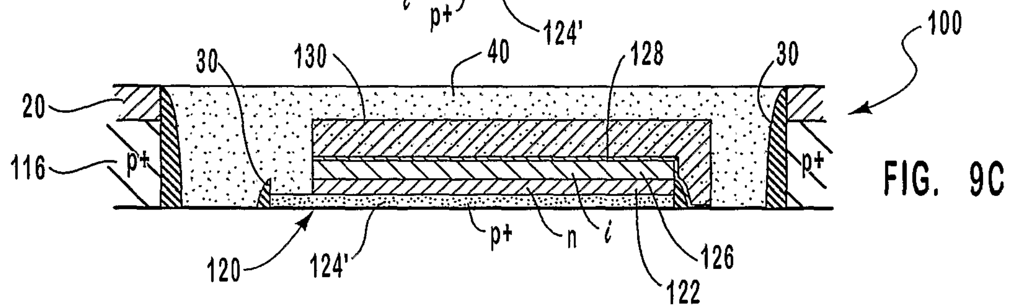


FIG. 9C

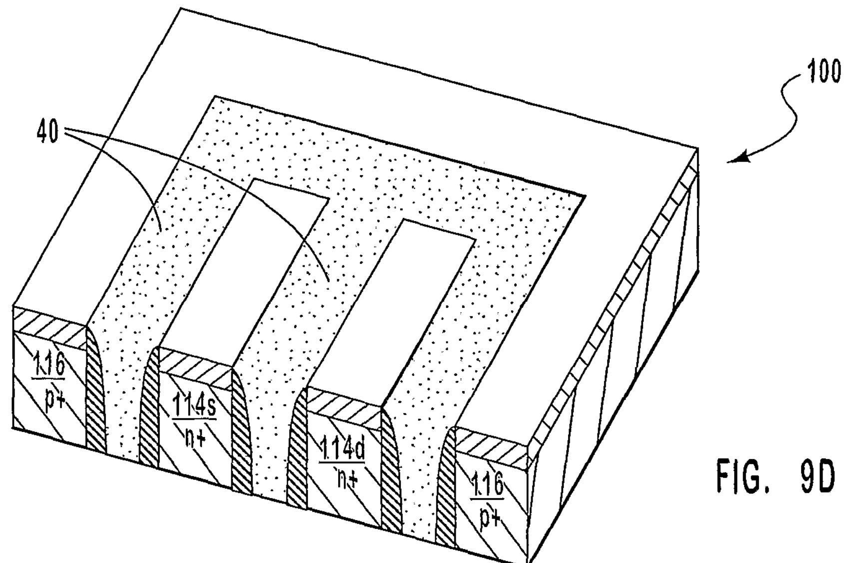


FIG. 9D

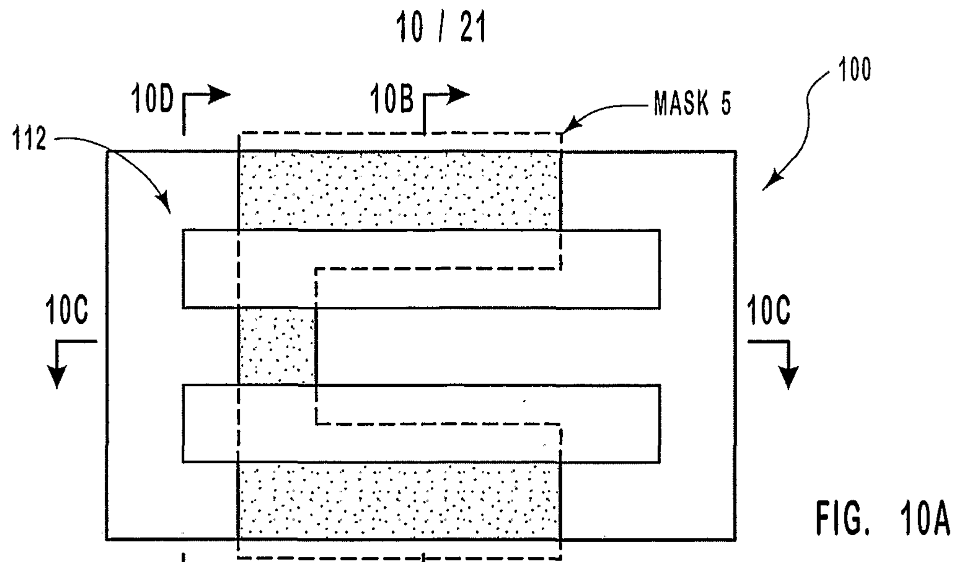


FIG. 10A

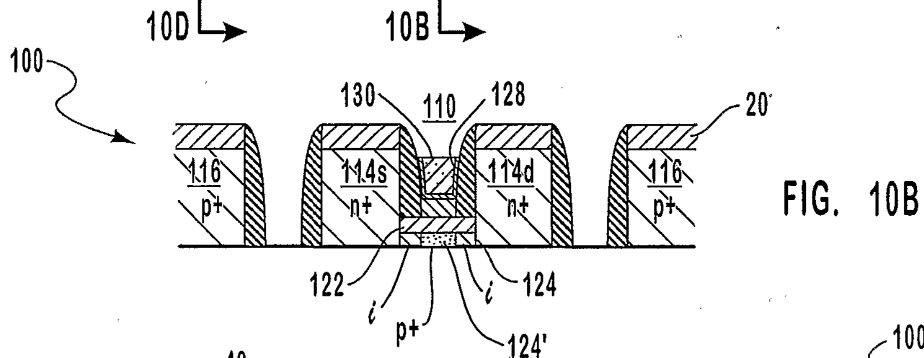


FIG. 10B

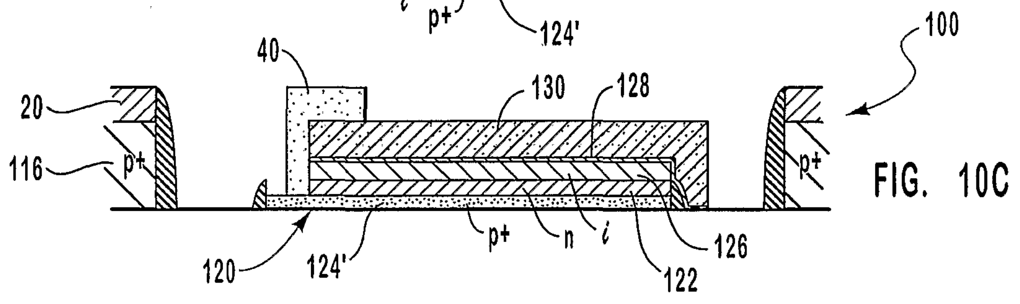


FIG. 10C

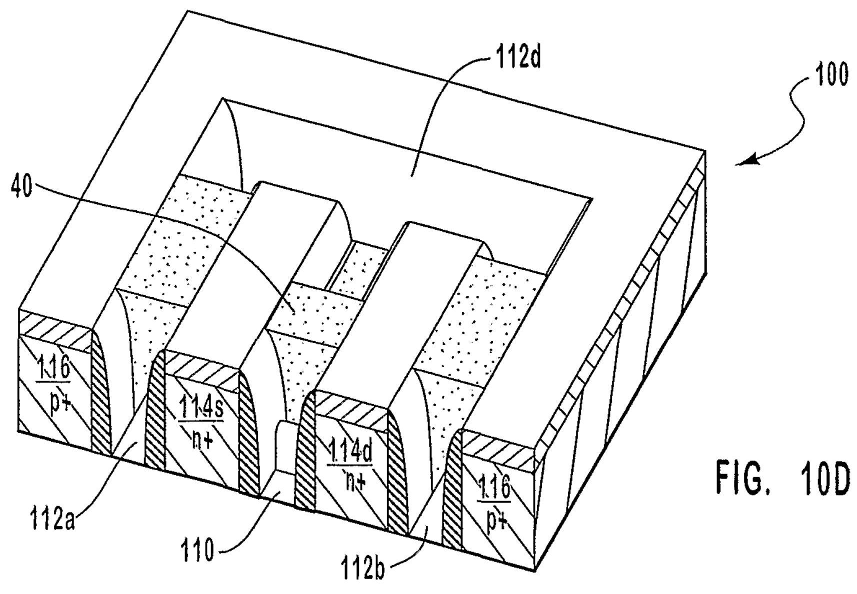


FIG. 10D

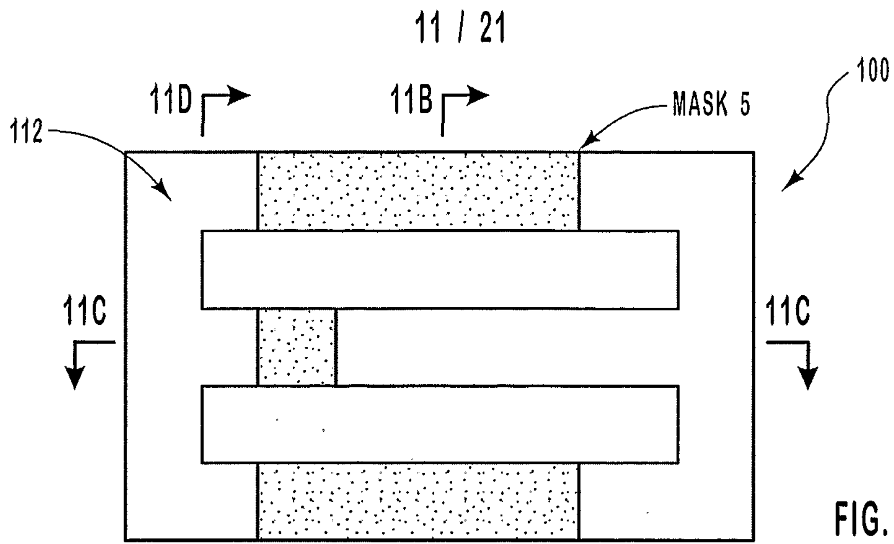


FIG. 11A

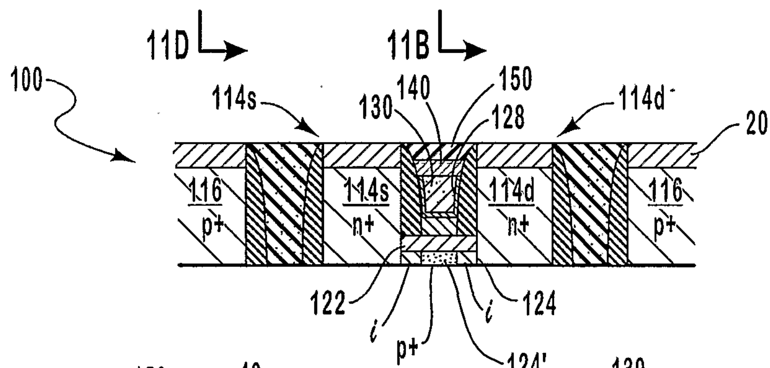


FIG. 11B

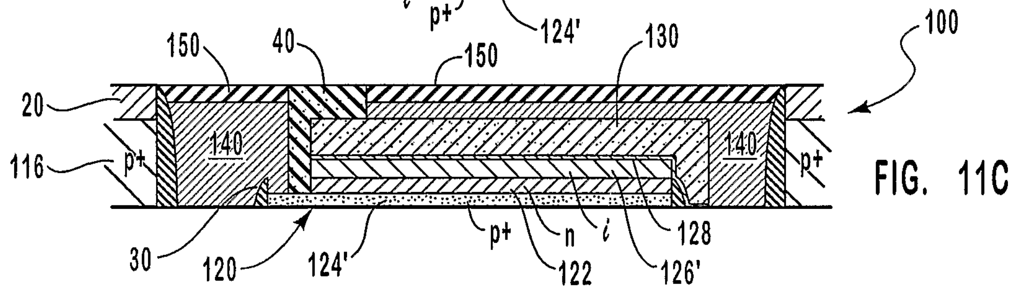


FIG. 11C

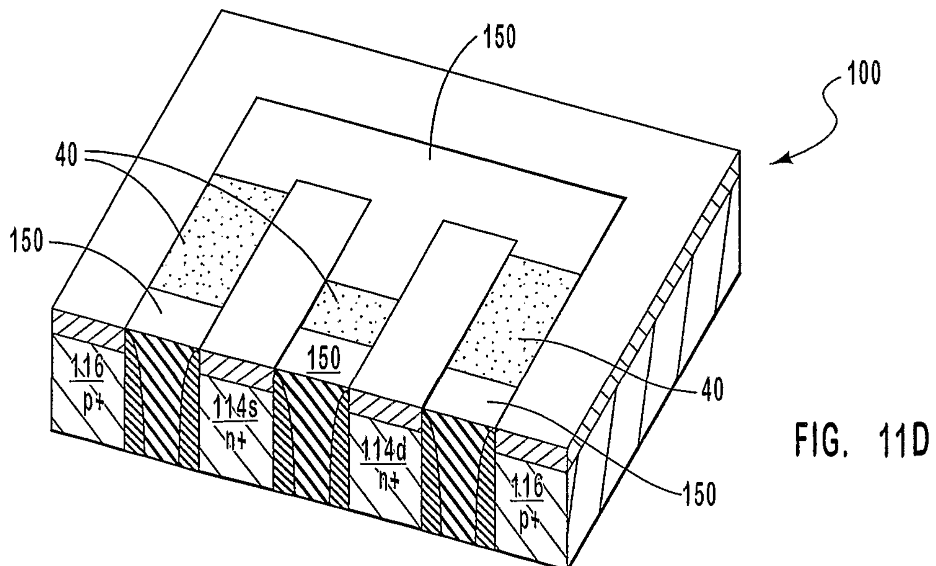


FIG. 11D

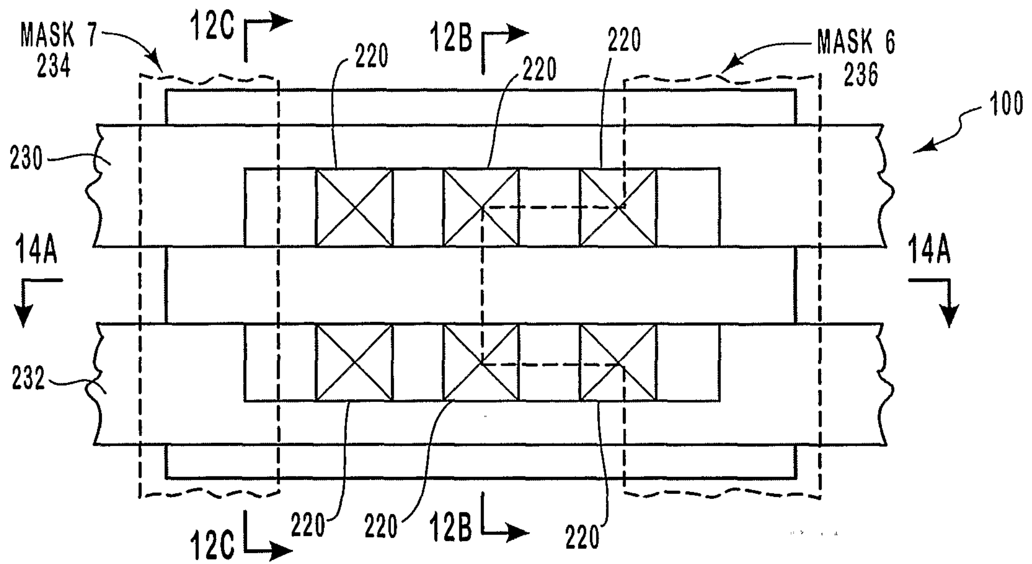


FIG. 12A

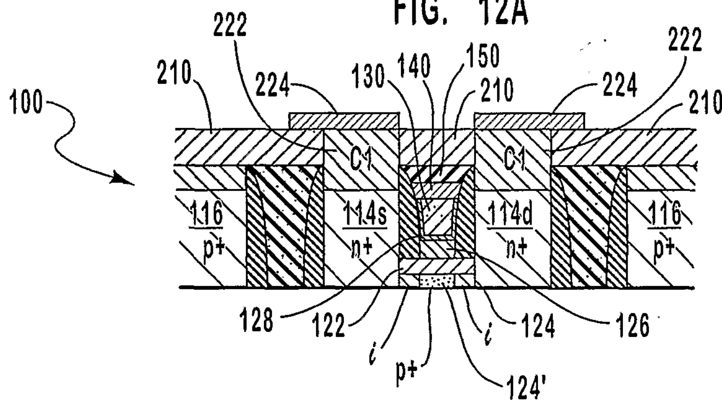


FIG. 12B

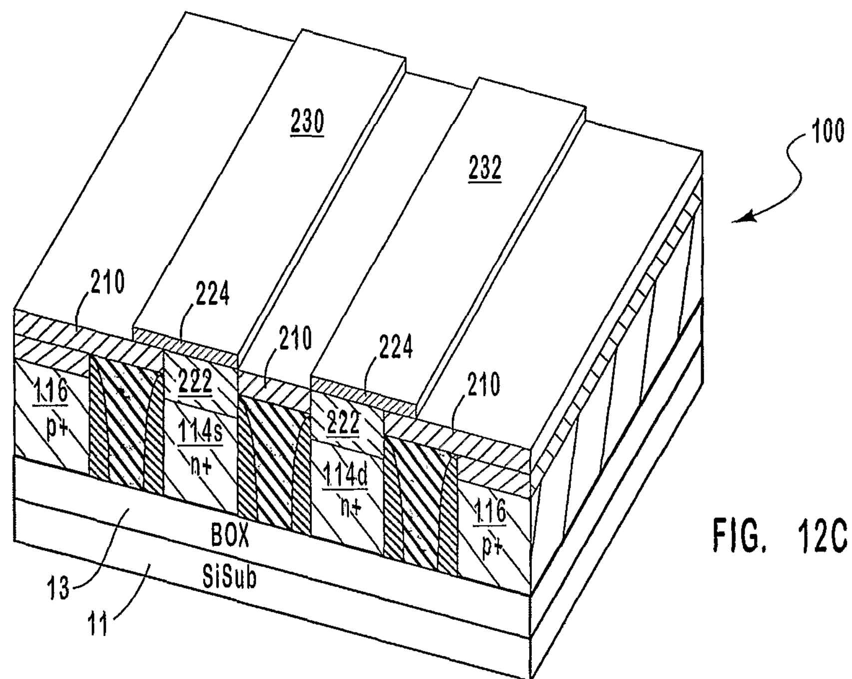


FIG. 12C

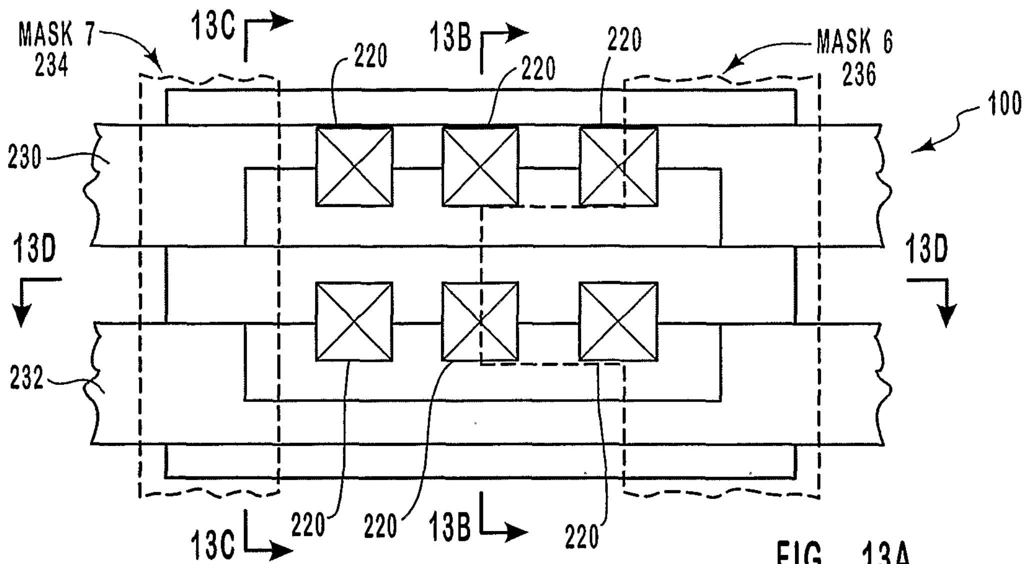


FIG. 13A

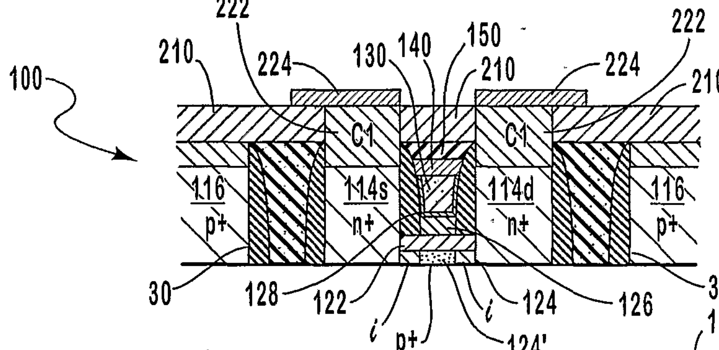


FIG. 13B

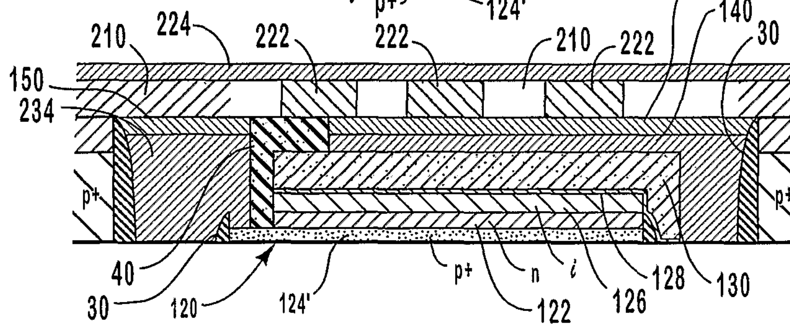


FIG. 13C

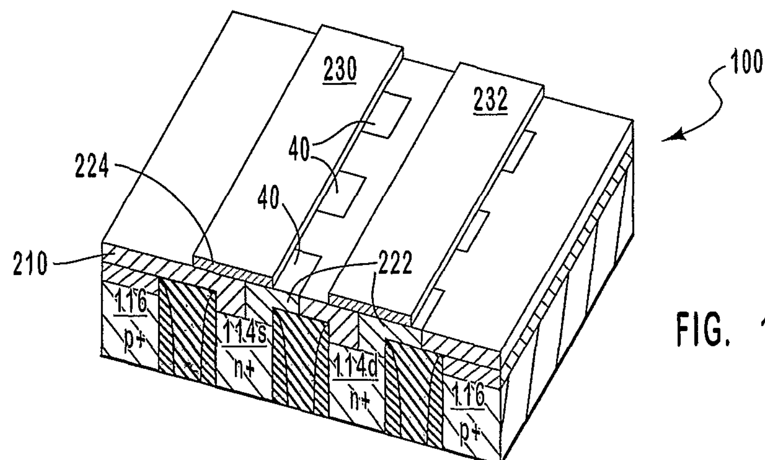


FIG. 13D

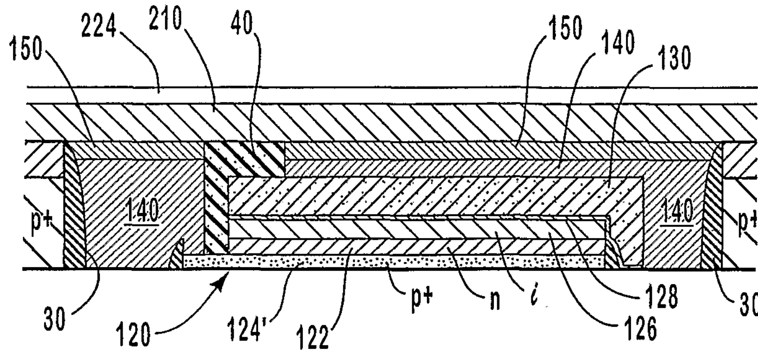


FIG. 14A

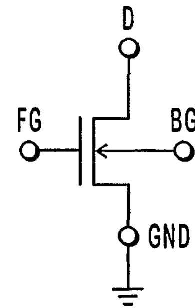


FIG. 14B

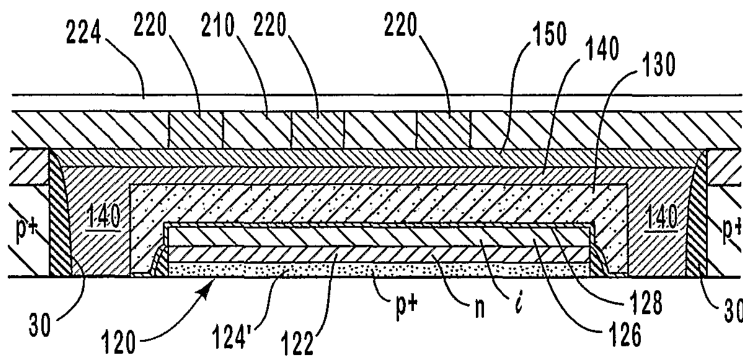


FIG. 15A

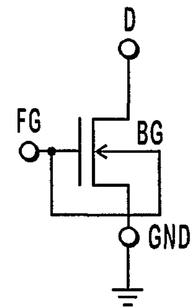


FIG. 15B

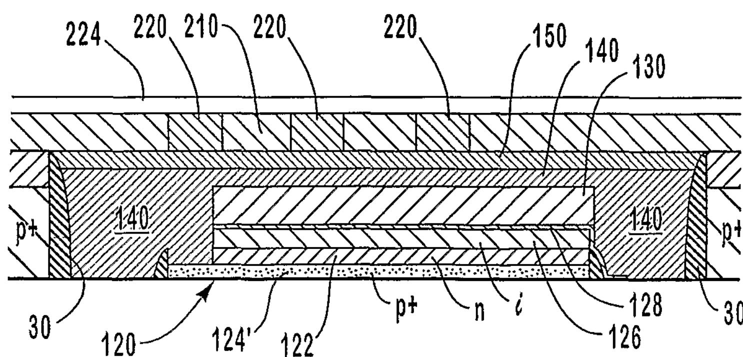


FIG. 16A

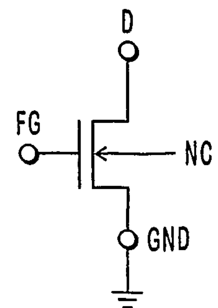


FIG. 16B

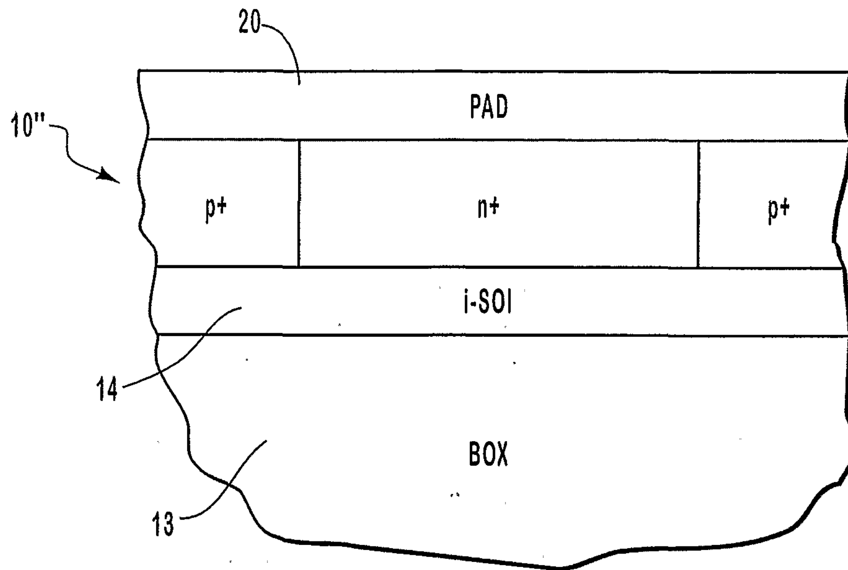


FIG. 17A

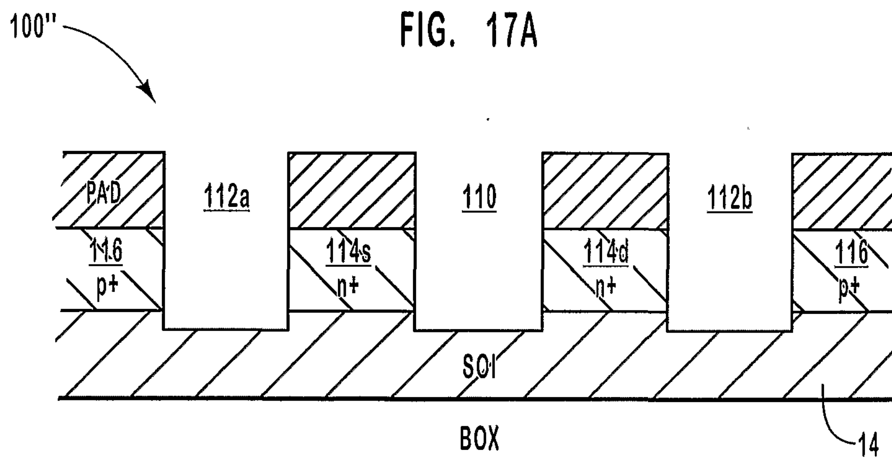


FIG. 17B

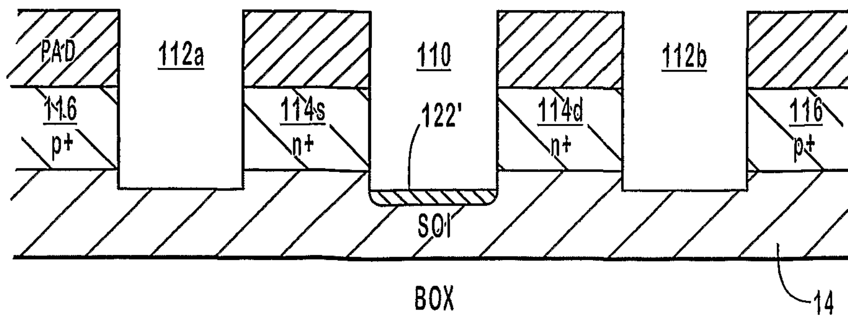


FIG. 18

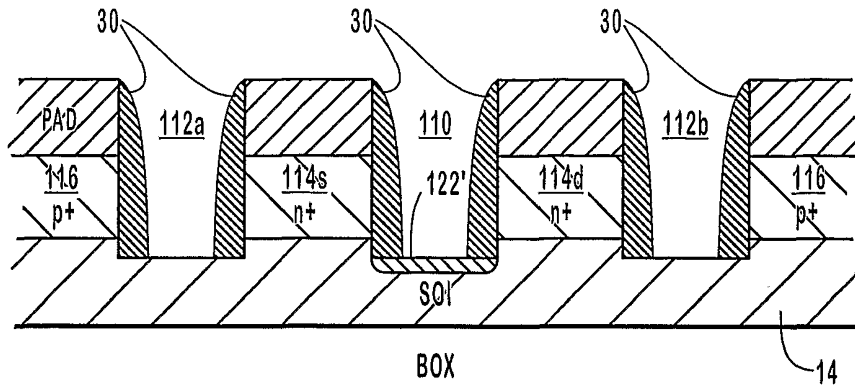


FIG. 19A

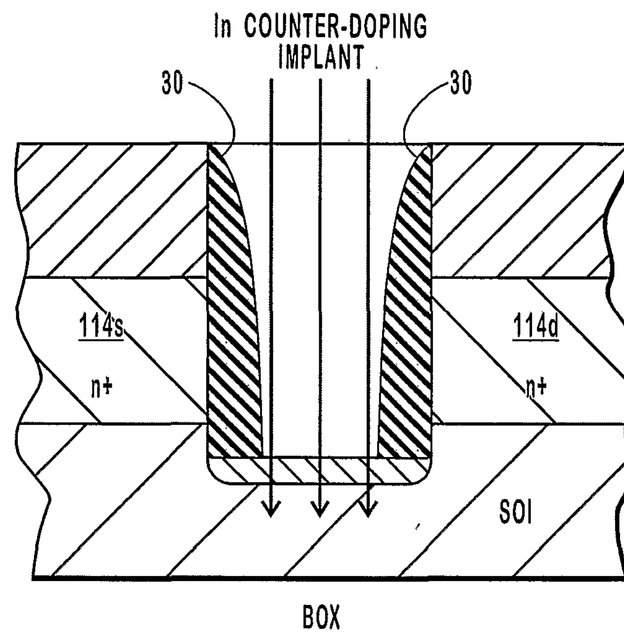


FIG. 19B

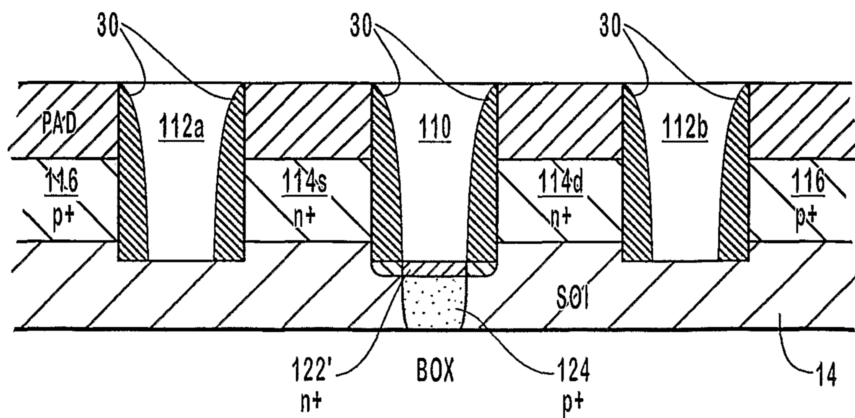


FIG. 19C

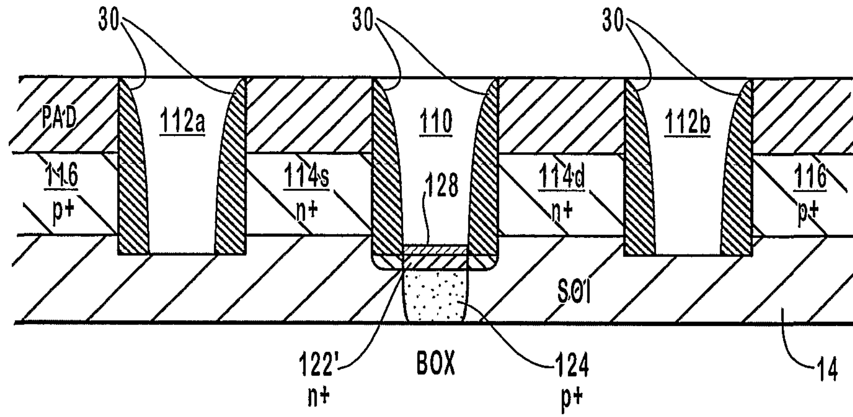


FIG. 20

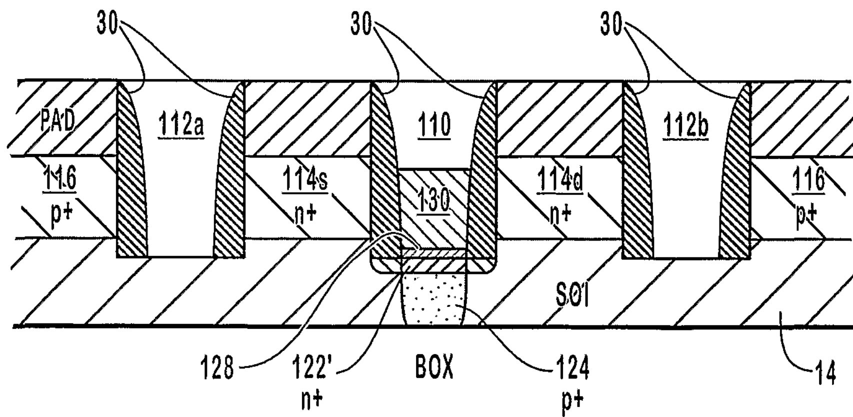


FIG. 21

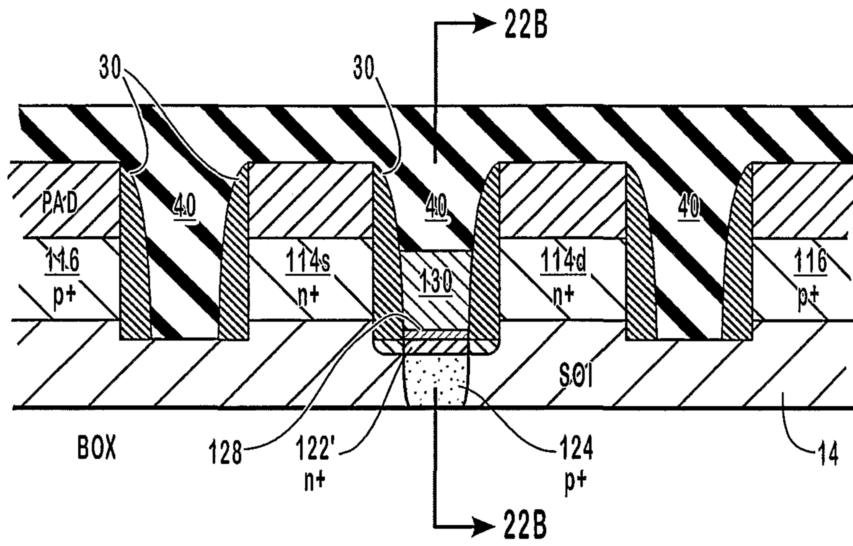


FIG. 22A

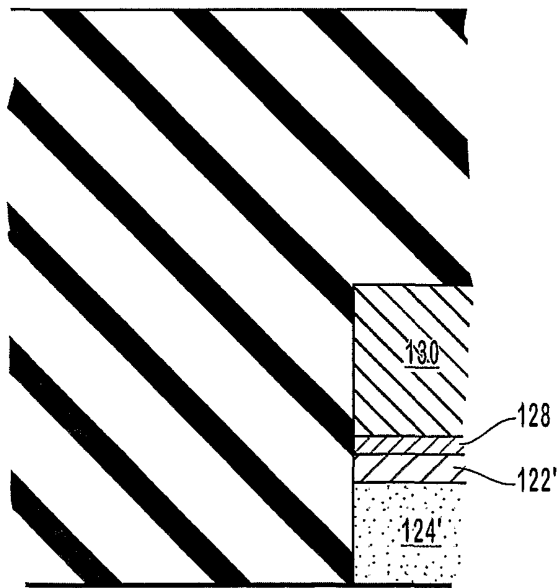


FIG. 22B

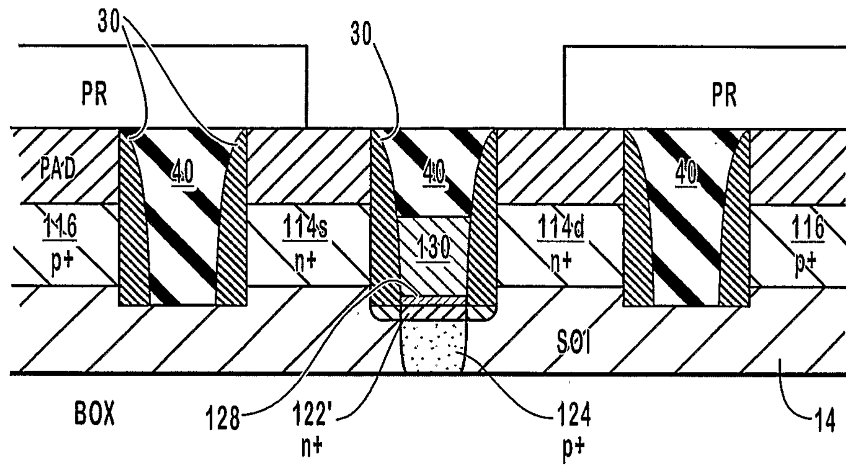


FIG. 23A

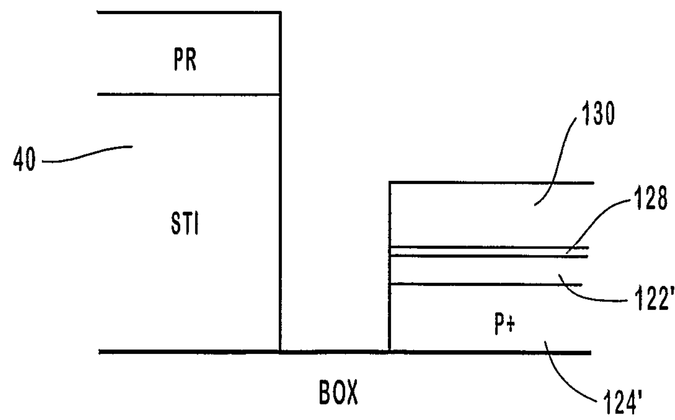


FIG. 23B

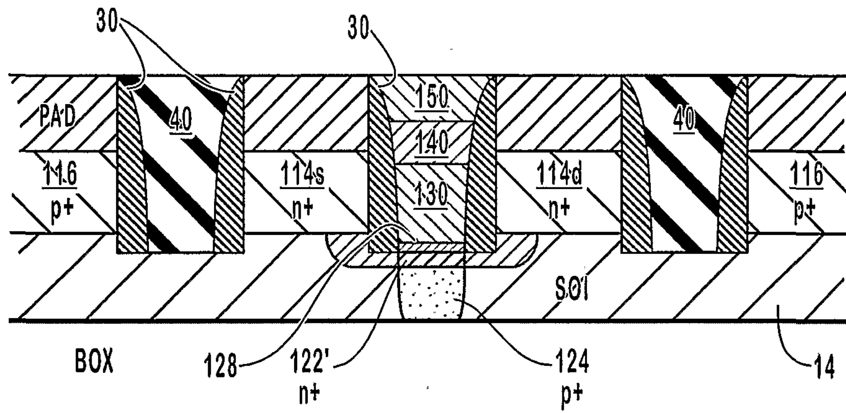


FIG. 24A

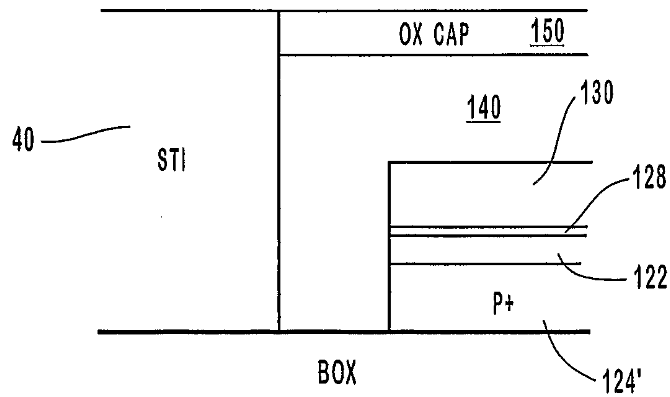


FIG. 24B

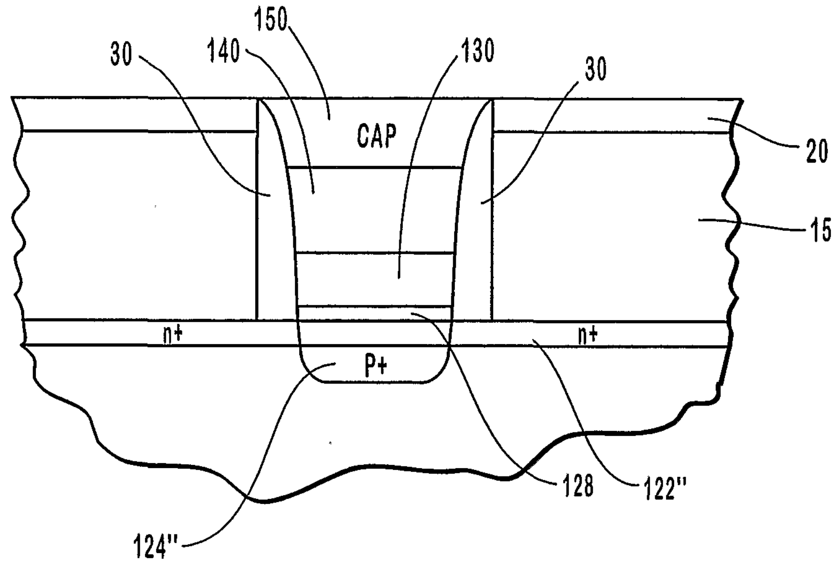


FIG. 25A

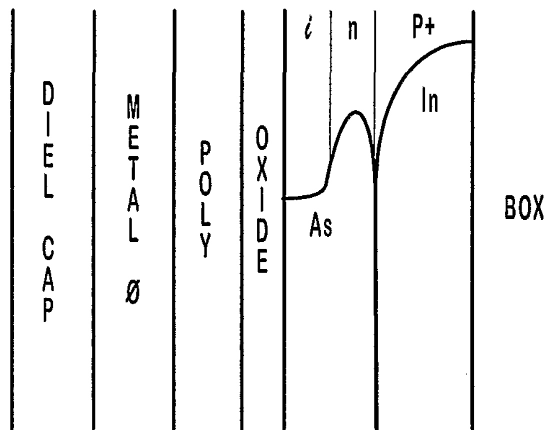


FIG. 25B

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/26920

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : H01L 21/00, H01L 21/84, H01L 21/336

US CL : 438/157, 438/164, 438/283, 438/215, 438/221, 438/231, 438/305, 438/527, 438/586, 438/FOR191, 438/FOR193, 438/FOR363; 257/523

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : Please See Continuation Sheet

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 practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 4,488,162 A (Jambotkar) 11 December 1984 (11.12.1984), column 8, lines 25+ and figs. 13-20.	1-2
A	US 6,255,171 B1 (Noble) 3 July 2001 (03.07.2001), column 1, lines 24- 27 and column 6, lines 30+.	1-2
A	US 5,091,324 A (Hsu et al.) 25 February 1992 (25.02.1992), column 4, lines 20+.	1-2

☐ Further documents are listed in the continuation of Box C. ☐ 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 application or patent 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
29 October 2001 (29.10.2001)	17 DEC 2001

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703)305-3230

Authorized officer
Wael Fahmy
Telephone No. 703-308-4918

Continuation of B. FIELDS SEARCHED Item 1:
438/157, 438/164, 438/283, 438/215, 438/221, 438/231, 438/305, 438/527, 438/586, 438/FOR191, 438/FOR193, 438/FOR363; 257/523