No. 2021-1709

United States Court of Appeals for the Federal Circuit

ROHM SEMICONDUCTOR USA, LLC,
Plaintiff-Appellant,

v.

MAXPOWER SEMICONDUCTOR, INC.,
Defendant-Appellee.

Appeal from the United States District Court for the Northern District of California in Case No. 1:20-cv-06686-VC, Judge Vince Chhabria

---

ROHM SEMICONDUCTOR USA, LLC'S OPENING BRIEF

Lisa Kobialka
James Hannah
KRAMER LEVIN NAFTALIS & FRANKEL LLP
990 Marsh Road
Menlo Park, California 94025
Telephone: (650) 752-1700

Aaron M. Frankel
Cristina Martinez
Shannon H. Hedvat
KRAMER LEVIN NAFTALIS & FRANKEL LLP
1177 Avenue of the Americas
New York, New York 10036
Telephone: (212) 715-9100

COUNSEL FOR PLAINTIFF-APPELLANT ROHM SEMICONDUCTOR USA, LLC

CERTIFICATE OF INTEREST

1. The full name of every party represented by me is:

   - Rohm Semiconductor USA, LLC

2. The name of the real party in interest represented by me is:

   - Rohm Semiconductor USA, LLC
   - Rohm Co., Ltd.

3. All parent corporations and any publicly held companies that own 10 percent of the stock of the party or amicus curia represented by me are listed below:

   - Rohm USA, Inc.
   - Rohm Co., Ltd.

4. The names of all law firms, partners and associates that have not entered an appearance in the appeal, and (a) appeared for the entity in the lower tribunal; or (b) are expected to appear for the entity in this court:

   - None.

5. Other than the originating case numbers(s), the title and number of any case known to counsel to be pending in this or any other court or agency that will directly affect or be directly affected by this court's decision in the pending appeal:

   - *MaxPower Semiconductor, Inc. v. Rohm Co., Ltd. and Rohm Semiconductor USA, LLC*, Case No. 21-cv-375899 (Sup. Ct. Cal. Feb. 10, 2021).

6. All information required by Fed. R. App. P. 26.1(b) and (c) in criminal cases and in bankruptcy cases.

   - None.

Dated: May 3, 2021

/s/ Lisa Kobialka
Lisa Kobialka

TABLE OF CONTENTS

Page

TABLE OF AUTHORITIES

Page(s)

Federal Cases

California Cases

Other State Cases

California Statutes

STATEMENT OF RELATED CASES

Pursuant to Fed. Cir. Rule 47.5, Appellant Rohm Semiconductor USA, LLC ("Rohm USA"), states that:

1) The following appeals have been taken from the following lower courts or bodies:

- None.

2) Appellant Rohm USA is a party to the following additional cases, which may be directly affected by the Court's decision in this appeal:

- *MaxPower Semiconductor, Inc. v. Rohm Co., Ltd. and Rohm Semiconductor USA, LLC*, Case No. 21-cv-375899 (Sup. Ct. Cal. Feb. 10, 2021).

JURISDICTIONAL STATEMENT

The District Court had jurisdiction over this action under 28 U.S.C. §1331 and §1338 because this action arose under the Patent Act of 35 U.S.C. § 101 *et seq*.

The District Court dismissed this action in favor of Defendant-Appellee MaxPower Semiconductor, Inc. ("MaxPower") on February 4, 2021. Appx1-2.

Rohm USA timely filed a notice of appeal pursuant to Federal Rule of Appellate Procedure 4(a)(2) on February 23, 2021. Appx853-854.

This Court has jurisdiction under 28 U.S.C. § 1295(a).

STATEMENT OF THE ISSUE ON APPEAL

1. Whether the District Court erred in concluding that the parties' agreement to arbitrate "in accordance with the provisions of the California Code of Civil Procedure" represented a "clear and unmistakable" agreement to have an arbitrator decide questions of arbitrability instead of a court?

STATEMENT OF THE CASE

Rohm USA filed suit against MaxPower seeking a declaration that Rohm USA does not infringe U.S. Patent Nos. 7,843,004, 8,076,719, 8,466,025, and 8,659,076 (together, the "MaxPower Patents"). *Rohm Semiconductor USA, LLC v. MaxPower Semiconductor, Inc.*, Case No. 20-cv-06686-VC (N.D. Cal. Sep. 23, 2020 (the "DJ Action"); Appx156-163. The DJ Action was ultimately assigned to Judge Vince Chhabria (the "District Court") of the United States District Court for the District of Northern California.

MaxPower filed a Second Renewed Motion to Compel Arbitration and Stay Claims for Declaratory Judgment (the "Motion to Compel"), contending that Rohm USA's request for a declaration of non-infringement was subject to arbitration. Appx708-725. The District Court granted MaxPower's Motion to Compel on February 4, 2021, and dismissed the DJ Action without prejudice. Appx1-2. The District Court did not make any finding on the ultimate question of the arbitrability of Rohm USA's claim, but found that the question of arbitrability must be resolved by an arbitrator. *Id.*

Rohm USA timely filed a notice of appeal on February 23, 2021. Appx853-854.

STATEMENT OF THE FACTS

A. MaxPower Accused Rohm USA of Patent Infringement

Appellant Rohm USA sells a variety of products, including silicon carbide metal oxide semiconductor field effect transistors ("SiC MOSFETs"). Appx158 at ¶¶ 9-10, 12. Rohm USA's SiC MOSFETs are used to control the flow of electrical current in electronic devices, and were developed with leading-edge technology to provide superior performance in high-voltage applications. *Id.* Non-party Rohm Co., Ltd. ("Rohm Japan"), Rohm USA's ultimate corporate parent, developed Rohm USA's SiC MOSFETs. Appx315-316; Appx158 at ¶ 12.

MaxPower accused Rohm USA of infringing the four MaxPower Patents based on Rohm USA's sales in the United States of SiC MOSFETs. Appx157-159 ¶¶ 1, 13. Rohm USA denies that its SiC MOSFETs infringe any of the MaxPower Patents. Rohm USA filed the underlying declaratory judgment action to obtain a declaration of non-infringement and to clear the cloud of MaxPower's infringement allegations. Appx157 at ¶¶ 1-3.

Rohm USA also filed petitions for *inter partes* review against each of the four MaxPower Patents. As these petitions are not at issue in this appeal, they will not be further discussed herein.

B. Rohm Japan and MaxPower Entered Into a Technology License Agreement Thirteen Years Ago

In 2007, Rohm Japan and MaxPower entered into a Technology License Agreement, which was subsequently amended various times. Appx651-664 (the "TLA"). Under the TLA, MaxPower agreed to provide Rohm Japan with certain limited and now-obsolete disclosures in exchange for Rohm Japan's payment of royalties on products developed based on those disclosures, specifically silicon MOSFETs. Appx651-655 (TLA at §§ 1.2, 1.8, 4); Appx635-644 at Appx646.

The TLA does not cover SiC MOSFETs and MaxPower never provided SiC MOSFET plans, designs, or technology to Rohm Japan during the course of either agreement. *See, e.g.,* Appx646 (2007 Target Specification). In addition, under the TLA, Rohm Japan does not have any royalty obligations for any independently developed technology — that is, any developments without the use of MaxPower's technology disclosed under the TLA. Appx652 (TLA at § 1.8).

C. The Technology License Agreement's Arbitration Provision is Narrow and Limited to Contractual Disputes

While Rohm Japan and MaxPower agreed to arbitrate disputes "arising out of or in relation" to the TLA, they did not agree to arbitrate any other disputes, such as patent disputes:

> Arbitration: Any dispute, controversy, or claim ***arising out of or in relation to this Agreement*** or at law, or the breach, termination or invalidity thereof, that cannot be settled amicably by agreement of the parties hereto, shall

> be finally settled by arbitration in Santa Clara County, California, USA in accordance with the provisions of the California Code of Civil Procedure by one or more arbitrators appointed in accordance with said Code . . .

Appx672-678 (2nd Amend. TLA at ¶ 10) (amending § 13.6 of the TLA) (emphasis added).

Notably, the TLA's arbitration provision does not reference arbitrating disputes of patent infringement issues. The TLA is also silent as to the question of who should decide any disputes over arbitrability.

D. The District Court Granted MaxPower's Motion to Compel Arbitration

In lieu of answering Rohm USA's declaratory judgment complaint, MaxPower filed a first motion to compel arbitration, which the District Court dismissed on procedural grounds not at issue on appeal. Appx497-498; Appx567-568. MaxPower filed a renewed motion to compel, which it then withdrew to file the Motion to Compel now on appeal. Appx679-680; Appx704-705; Appx706-707; Appx708-725; Appx753-767; Appx834-846. The premise of MaxPower's Motion to Compel is MaxPower's allegation that the parties specifically agreed to arbitrate Rohm USA's claim for a declaration of non-infringement.

The District Court issued a one paragraph order granting MaxPower's Motion to Compel (the "Order"). Appx1-2. The District Court made no finding on the question of if the parties had agreed to arbitrate requests for a declaration of patent

non-infringement. Instead, the District Court found that this threshold question of arbitrability should be decided by an arbitrator and not by the District Court. *Id.*

The District Court noted that, "[t]he TLA incorporates the California Code of Civil Procedure [("CCCP")] into its arbitration provision." *Id.* The District Court then quoted the CCCP's add on provisions for international arbitration which state that "[t]he arbitral tribunal ***may*** rule on its own jurisdiction." *Id.* (emphasis added) (quoting Cal. Civ. Proc. Code § 1297.161). The District Court found this reference to the CCCP to have "clearly and unmistakably delegate[d] the question of arbitrability to the arbitrator," and dismissed the DJ Action. The District Court did not identify, nor did MaxPower offer, any other evidence of an agreement to delegate arbitrability to an arbitrator. *Id*.

SUMMARY OF THE ARGUMENT

The District Court erred in dismissing the DJ Action because the parties never "clearly and unmistakably" agreed to have an arbitrator decide disputes over arbitrability.

Under the Supreme Court's controlling *First Options* decision, in the absence of such a clear and unmistakable agreement, questions of arbitrability are reserved to the courts. *First Options of Chicago, Inc. v. Kaplan*, 514 U.S. 938, 944-45 (1995). Such a "clear and unmistakable delegation" would usually consist of an express statement in the arbitration provision agreeing to arbitrate any questions of

arbitrability. Here, in contrast, the parties' agreement is completely silent on the issue of who decides questions of arbitrability. Thus, because there was no clear and unmistakable delegation of arbitrability, the District Court should have itself decided the threshold question of arbitrability. For this reason, reversal and remand are warranted.

In *First Options*, the Supreme Court recognized the fundamental unfairness in having arbitrators decide questions of arbitrability, when arbitrators have every incentive to determine that a dispute is subject to arbitration. To prevent parties from being forced to arbitrate disputes they never agreed to arbitrate, the courts (not arbitrators) are supposed to be the gatekeepers of arbitrability unless the parties expressly agree to the contrary. *See id.*

In finding that Rohm USA had clearly and unmistakably agreed to have only an arbitrator decide arbitrability disputes, the District Court erroneously relied on the Ninth Circuit's decision in *Oracle America, Inc. v. Myriad Group A.G.*, 724 F.3d 1069 (9th Cir. 2013). In *Oracle*, the incorporation of UNCITRAL's rules was deemed an agreement to delegate arbitrability disputes to an arbitrator. Indeed, the UNCITRAL rules provided for arbitrators to decide arbitrability questions. Here, in contrast, the parties' agreement references the CCCP. The CCCP expressly provides that ***courts*** may decide questions of arbitrability, while also permitting arbitrators to decide such questions in limited circumstances. *See* CCCP § 1281.2. Thus,

incorporation of the CCCP is not a clear and unmistakable agreement to delegate arbitrability.

For this reason, *Oracle* does not apply here. But to the extent *Oracle would* require an arbitrator to decide arbitrability under these facts, it must be set aside as contrary to the Supreme Court's *First Options* decision. *Oracle* and its progeny cannot abrogate the Supreme Court's requirement that arbitrability questions can only be delegated to an arbitrator if there is a clear and unmistakable agreement to do so. Indeed, during oral argument, the District Court acknowledged that *Oracle* is inconsistent with the clear and unmistakable standard, acknowledging that "one could say that clear and unmistakable is no longer clear and unmistakable after *Oracle*." Appx887-889 at 9:6-16.

MaxPower is seeking to force an unwilling party to arbitrate an issue that was never contemplated under the TLA. Reversal is warranted under *First Options* so that the District Court may decide the threshold question of arbitrability.

ARGUMENT

I. The District Court Erred in Delegating to an Arbitrator the Threshold Question of Arbitrability

A. This Court Applies De Novo Review to the District Court's Dismissal

This Court applies de novo review to the District Court's dismissal of the DJ Action. *Verinata Health Inc. v. Ariosa Diagnostics Inc.*, 830 F.3d 1335, 1338 (Fed. Cir. 2016).

As the regional circuit, Ninth Circuit law applies. *Microchip Tech. Inc. v. U.S. Philips Corp.,* 367 F.3d 1350, 1356 (Fed. Cir. 2004) ("We are obligated to follow regional circuit law on questions of arbitrability that are not 'intimately involved in the substance of enforcement of a patent right.'") (citation omitted).

The District Court did not make any underlying factual determinations in support of its Order, which would have been subject to "clear error" review. *Verinata Health*, 830 F.3d at 1338 (citing *Promega Corp. v. Life Techs. Corp.*, 674 F.3d 1352, 1355 (Fed. Cir. 2012)). Specifically, the District Court's interpretation of the scope of the TLA's arbitration provision, which did not rest on any extrinsic evidence, is a question of law to which de novo review is applied. *See French v. Merrill Lynch, Pierce, Fenner & Smith, Inc.*, 784 F.2d 902, 908 (9th Cir. 1986).

B. *First Options* Requires Courts to Resolve Questions of Arbitrability Themselves Unless the Parties Clearly and Unmistakably Agreed to Delegate That Question to Arbitrators

The Supreme Court dictated that the courts, not arbitrators, should be the gatekeepers to decide if a dispute is subject to arbitration, unless the parties have a clear and unmistakable agreement to the contrary. Here, the District Court erred in finding that only an arbitrator should decide the threshold question of arbitrability because Rohm and MaxPower never expressly agreed to send arbitrability questions to an arbitrator.

In *First Options*, 514 U.S. 938 at 944, the Supreme Court held that courts should decide arbitrability questions and should ***not*** "assume that the parties agreed to arbitrate arbitrability unless there is 'clea[r] and unmistakabl[e]' evidence that they did so." *Id.* at 944 (citations omitted); *see Henry Schein, Inc. v. Archer & White Sales, Inc.*, 139 S. Ct. 524, 531 (2019) ("courts 'should not assume that the parties agreed to arbitrate arbitrability unless there is a clear and unmistakable evidence that they did so'"); *see also AT&T Techs., Inc. v. Communications Workers of Am.*, 475 U.S. 643, 649 (1986) ("Unless the parties clearly and unmistakably provide otherwise, the question of whether the parties agreed to arbitrate is to be decided by the court, not the arbitrator."). Here, as discussed in the following section, Rohm never agreed to delegate arbitrability, so it was legal error for the District Court to not itself decide the issue.

The Supreme Court set forth a basic presumption in favor of having courts decide if a dispute is arbitrable. When there is "silence or ambiguity about the question '*who* . . . should decide arbitrability,'" as is the case here, the courts must do so themselves. *First Options*, 514 U.S. at 944-45 (emphasis in original) (citation omitted). This is because the general presumption in favor of arbitration does not apply to the question of who should decide a dispute over arbitrability. *Id.* Thus, under *First Options*, if a contract includes an arbitration provision, it will be interpreted to require courts to decide questions of arbitrability, unless the

parties have a clear and unmistakable agreement to delegate that question to arbitrators.

The Supreme Court explained the policy behind this presumption against delegation of arbitrability to be a matter of fundamental fairness, given the "arcane" nature of the question and the likelihood that parties did not focus on it in negotiating their arbitration provision:

> the "who (primarily) should decide arbitrability" question—is rather arcane. A party often might not focus upon that question or upon the significance of having arbitrators decide the scope of their own powers.

First Options, 514 U.S. at 945 (citations omitted).

The Supreme Court further recognized the risk of giving arbitrators the power to decide if a dispute is subject to arbitration, given the incentive that arbitrators have to find in favor of arbitrability, causing parties such as Rohm USA to arbitrate disputes they never agreed to arbitrate:

> [G]iven the principle that a party can be forced to arbitrate only those issues it specifically has agreed to submit to arbitration, one can understand why courts might hesitate to interpret silence or ambiguity on the "who should decide arbitrability" point as giving the arbitrators that power, for doing so might too often ***force unwilling parties to arbitrate a matter they reasonably would have thought a judge, not an arbitrator, would decide***.

Id. (citations omitted) (emphasis added).

Thus, the District Court was to presume that Rohm USA did not agree to delegate arbitrability questions, and should have decided the question itself

because there is no clear and unmistakable agreement between Rohm and MaxPower to only allow arbitrators to decide that question.

C. The TLA's Reference to California Law is Not a Clear and Unmistakable Agreement to Delegate All Arbitrability Disputes to Arbitrators

The District Court erred in dismissing the DJ Action because the TLA does not include a clear and unmistakable agreement to exclusively delegate arbitrability disputes to arbitrators. To the contrary, it is entirely silent on the issue of who will decide questions of arbitrability. Appx677-678 (2nd Amend. TLA at § 10). Such silence cannot be a clear and unmistakable agreement to delegate arbitrability questions. Therefore, applying de novo review, the Court should reverse under *First Options*.

The only purported evidence offered by MaxPower of an agreement to delegate arbitrability, and the sole basis relied upon by the District Court, was the statement in the TLA's arbitration provision that "Any . . . claim arising out of or in relation to this Agreement . . . shall be finally settled by arbitration . . . in accordance with the provisions of the California Code of Civil Procedure [CCCP]." Appx1-2; Appx718-720; Appx840-841.

The District Court, applying the Ninth Circuit's *Oracle* decision (discussed in the following section), erroneously found that this single, generic reference to the CCCP was a "clear and unmistakable" agreement to have arbitrators — but not

courts — resolve arbitrability disputes because the CCCP's international arbitration subsection provides that an "arbitral tribunal ***may*** rule on its own jurisdiction." Appx1 (citing Cal. Civ. Proc. Code § 1297.161) (emphasis added).

The error in the District Court's analysis is that the CCCP does ***not*** require arbitrators to decide all questions of arbitrability, so a reference to the CCCP is not equivalent to an agreement to arbitrate arbitrability disputes. In particular, the CCCP's arbitration provisions expressly provide that ***courts*** should decide questions of arbitrability, providing a specific procedure for the resolution of such disputes: "the ***court*** shall order the petitioner and the respondent to arbitrate the controversy ***if it determines that an agreement to arbitrate the controversy exists.***" CCCP § 1281.2 (emphasis added). This general provision applies to all arbitrations under the CCCP. *Id.*

The provision of the CCCP upon which the District Court relied, § 1297.161, is permissive in permitting international arbitrators to resolve questions of arbitrability, but does not negate the courts' statutory authority under the CCCP to also determine questions of arbitrability. *See* CCCP § 1297.161. Thus, the most that can be said about the CCCP on this issue is that a court is always empowered to decide questions of arbitrability and, for international arbitrations, an arbitral tribunal may also decide such questions. Given that the CCCP does not exclusively delegate to arbitrators questions of arbitrability, a

reference to the CCCP, without more, is not a clear agreement to have only arbitrators decide arbitrability.

This should have ended the inquiry for the District Court in this case because the incorporation of the CCCP into the TLA cannot have been the result of a "clear and unmistakable" agreement to strip courts of the statutory ability to resolve arbitrability, and, therefore, "the question of whether the parties agreed to arbitrate [wa]s to be decided by the [C]ourt." *AT&T Techs.*, 475 U.S. at 649.

For the District Court to find otherwise, it had to find that, when the parties negotiated the TLA and its arbitration provision in 2007, they specifically agreed that if they later disagreed on an arbitrability question that they would take that question away from the courts and have it be decided only by an arbitrator. The District Court would further need to find that, in order to record the agreement, the parties specifically chose to incorporate the CCCP as a shorthand for expressly stating that they intended to have only arbitrators decide arbitrability. Moreover, the District Court would need to find that the parties appreciated the "arcane" nuances of the multiple provisions in the CCCP that permit both courts and arbitrators to decide questions of arbitrability and, yet, assumed that incorporation of the CCCP would be sufficiently clear to reflect their otherwise unexpressed agreement to delegate arbitrability questions.

The District Court made no such findings. Nor could it have. To the contrary, the District Court recognized that applying *Oracle* here effectively ignored the *First Options* clear and unmistakable standard and that there was no clear and unmistakable agreement here. Appx887-889 at 9:6-16. Reading an agreement to delegate arbitrability disputes into the TLA where none exists is legal error and does not reflect an actual agreement between the parties on this issue. Therefore, under *First Options*, the District Court erred in granting MaxPower's Motion to Compel in the absence of such evidence.

D. *Oracle* and the Other Cases Cited By the District Court Do Not Apply

In erroneously finding that the parties agreed to arbitrate the question of arbitrability, the District Court relied on the Ninth Circuit's *Oracle* decision and several similar cases. Those cases are distinguishable because of the differences in the parties' agreements they considered. In those cases, unlike this case, the parties agreed to arbitrate under the rules of a specific organization that ***only*** permitted arbitrators to decide arbitrability questions. The arbitration rules incorporated into the agreements in those cases did not include a provision equivalent to the CCCP's provision providing for ***courts*** to decide arbitrability disputes.

For example, in *Oracle America, Inc. v. Myriad Group A.G.*, 724 F.3d 1069 (9th Cir. 2013), upon which the District Court relied, the agreement at issue

incorporated UNCITRAL rules which ***expressly*** "delegate[] questions of arbitrability" to an arbitrator. *Id.* at 1073. Similarly, in *Dream Theater, Inc. v. Dream Theater,* 124 Cal. App. 4th 547 (2004), also cited in the Order, the court held that the "[AAA Commercial Arbitration Rules] . . . ***specify*** that the arbitrator will decide disputes over the scope of the arbitration agreement," deeming the incorporation of AAA rules to, therefore, be an unmistakable agree to delegate arbitrability to arbitrators. *Id.* at 557 (emphasis added). None of the rules incorporated into the arbitration provisions in these cases relied upon by the District Court expressly provide for courts to resolve questions of arbitrability, as the CCCP does.

An additional reason *Oracle* and the other cases are distinguishable is that the arbitration agreements in those cases incorporated a specific set of rules which, those courts reasoned, was an unmistakable agreement to incorporate the delegation of arbitrability unambiguously provided by those rules. *See, e.g.*, *Oracle*, 724 F.3d at 1074 (expressly incorporating UNCITRAL rules which is "consistent with the majority view regarding the effect of incorporating the AAA rules into an agreement").

Here, in contrast, the TLA's reference to the entirety of the CCCP's sprawling arbitration provisions, sweeps in general arbitration provisions, as well as a variety of rules that apply in different circumstances, such as the consolidation

of separate arbitration proceedings, the selection of the procedure to be followed by an arbitral tribunal to conduct the proceedings, and international arbitration. *See* CCCP at *e.g.*, §§ 1281.3, 1297.191, 1297.161. The TLA did not limit its incorporation of the CCCP to any particular subsections and, in particular, did not expressly provide for arbitration specifically under the international arbitration provisions that MaxPower relies on. Appx677-678 (2nd Amend. TLA at § 10). Therefore, general reference to the CCCP is not a clear an unmistakable reference to the specific arbitrability provision of the international sub-section of the CCCP. This is especially the case here, where the underlying case is a domestic dispute between Rohm USA and MaxPower — both California entities, for a declaratory judgment of non-infringement of MaxPower's United States patents relating to Rohm USA's sales of SiC MOSFETs in the United States. Appx157 at ¶¶ 1-3.

Even if the dispute between Rohm USA and MaxPower were interpreted to be an international one, CCCP § 1297.161 simply sets forth that an international arbitral tribunal "**may** rule on its own jurisdiction," not that it **must** or exclusively can do so. CCCP § 1297.161 (emphasis added). California courts have held that may has a specific legal meaning (in contrast to shall), writing that "the usual rule with California codes is that 'shall' is mandatory and 'may' is permissive unless the context requires otherwise." *Walt Rankin & Assocs., Inc. v. City of Murrieta*, 84 Cal. App. 4th 605, 614 (2000) (citation omitted).

This distinction between "shall" and "may" extends to the CCCP, as courts interpreting sections of the CCCP have held that, "in the absence of such special circumstances, ["may"] should be interpreted as permissive or conferring discretion." *Rhyne v. Mun. Ct.*, 113 Cal. App. 3d 807, 817 (Ct. App. 1980); *see also Gipson v. Davis Realty Co.*, 215 Cal. App. 2d 190, 202 (Ct. App. 1963) (holding that "may" "is primarily and ordinarily a permissive term").

The permissive and non-exclusive nature of the statement that, in international cases, an arbitrator "may" rule on arbitrability, is confirmed by other provisions in the CCCP providing that the courts may also decide this question. CCCP § 1281.2 (authorizing court to "determine[] that a written agreement to arbitrate a controversy exists").

The District Court, therefore, was ***not*** precluded by the CCCP from resolving the issue of arbitrability. The permissive language of the CCCP, solely pertaining to a small subset of disputes (that is, international ones), contrasts with the provisions at issue in the cases the District Court relied upon, which incorporate rules expressly delegating arbitrability. Appx 1. Therefore, the general reference to the CCCP is not a clear and unmistakable agreement to have only arbitrators decide arbitrability questions.

The premise that incorporation of rules for arbitration (especially rules that are ambiguous on arbitrability) is sufficient to delegate arbitrability violates the

public policy embedded in *First Options* that parties should be presumed not to have agreed to delegate arbitrability. *First Options*, 514 U.S. at 944-45. When Rohm Japan and MaxPower incorporated the CCCP, there is no reason to believe that they actually had in mind this "arcane" issue of the delegation of arbitrability and that they appreciated the interplay of various provisions among the voluminous CCCP that might relate to this issue. *Id.* at 945 ("[a] party often might not focus upon that question or upon the significance of having arbitrators decide the scope of their own powers . . . [thus] one can understand why courts might hesitate to interpret silence or ambiguity on the 'who should decide arbitrability' point as giving the arbitrators that power, for doing so might too often force unwilling parties to arbitrate a matter they reasonably would have thought a judge, not an arbitrator, would decide") (citations omitted). And, if Rohm and MaxPower truly did have a meeting of the mind on this issue, they would have said so expressly in the TLA, not indirectly and ambiguously through reference to the CCCP. Thus, the incorporation of the CCCP was not a clear and unmistakable agreement that only arbitrators should decide arbitrability.

E. *Oracle* Cannot Overturn the Supreme Court's Decision in *First Options*

As discussed above, *First Options* dictates reversal, and the Ninth Circuit's *Oracle* case is not to the contrary because it is distinguishable. Courts have recognized that while "a question over arbitrability [and] who should decide the

answer—the arbitrator or the court—can pose something of an analytical challenge," the *Supreme Court* has "provided a framework to resolve that first order issue in *First Options*." *Doe v. Natt*, 299 So. 3d 599, 602 (Fla. Dist. Ct. App. 2020). And that controlling framework must be applied in every instance, including here.

To the extent *Oracle* dictates a different result under these facts, it must be set aside as inconsistent with the Supreme Court's controlling *First Options* decision. It is pure legal fiction to read the parties' incorporation of the CCCP into the TLA to reflect a negotiated agreement between the parties expressly intended to waive having courts decide arbitrability questions. This fiction cannot be reconciled with *First Options*.

At least one court has criticized *Oracle* as contrary to *First Options* for this reason. In a state court decision declining to find that incorporation of AAA arbitration rules is a clear and unmistakable delegation of arbitrability, *Doe* recognizes that, in *Oracle,* the Ninth Circuit failed to "examine[] how or why the mere 'incorporation' of an arbitration rule . . . satisfied the heightened standard the Supreme Court set in *First Options*, nor how it overcomes the 'strong pro-court presumption' that is supposed to attend this inquiry." *Doe*, 299 So. 3d at 608 (citing *Howsam v. Dean Witter Reynolds, Inc.*, 537 U.S. 79, 86 (2002)).

The District Court here also acknowledged the inconsistency between *Oracle* and *First Options* during oral argument, but took the view that its hands were tied and it was obligated to follow *Oracle* notwithstanding this inconsistency:

> ***I think one could say that clear and unmistakable is no longer clear and unmistakable after Oracle*** and all of the cases that – all of the other cases that said that you delegate that to the arbitrator, with language like that. . . It's not even in the contract, but it's incorporated in the rules. The language itself is not super specific. ***I mean I think it is fair to say that clear and unmistakable is no longer just clear and unmistakable***, but that's the case law and I'm just not sure how you can get around that.

Appx888 at 9:7-16 (emphasis added).

In following *Oracle*, not the Supreme Court's controlling *First Options* decision, the District Court failed to heed the Ninth Circuit's recent acknowledgment that its own decisions cannot be inconsistent with *First Options*. *See SEIU Local 121RN v. Los Robles Reg'l Med. Ctr.*, 976 F.3d 849, 861 (9th Cir. 2020). In *SEIU Local*, the Ninth Circuit reversed prior Ninth Circuit precedent finding a presumption in favor of delegation of arbitrability in labor cases because those cases were inconsistent with *First Options*' general presumption against delegation. *Id.* ("[a]s *First Options* instructs us, ***'[c]ourts should not assume that the parties agreed to arbitrate arbitrability' in the face of 'silence or ambiguity' with respect to the Delegation Question***" (emphasis added) (citing *First Options*, 514 U.S. at 944-45)). Here, to the extent *Oracle* applies to the TLA's reference to

the CCCP's arbitration rules, which are at most ambiguous as to the question of delegation, it cannot stand in view of *First Options.*

Thus, under de novo review, the Court should find that *First Options* applies and, due to the absence of a "clear and unmistakable" agreement to delegate arbitrability, the Order should be reversed.

CONCLUSION

For the foregoing reasons, Rohm USA respectfully requests that this Court reverse the District Court's grant of the Motion to Compel, and remand this matter to the District Court to decide the question of the arbitrability of Rohm USA's request for a declaration of non-infringement.

Respectfully submitted,

Dated: May 3, 2021

By: */s/ Lisa Kobialka*

Lisa Kobialka
James Hannah
Kramer Levin Naftalis
& Frankel LLP
990 Marsh Road
Menlo Park, CA 94025
Tel: 650.752.1700
Fax: 650.752.1810
lkobialka@kramerlevin.com
jhannah@kramerlevin.com

Aaron M. Frankel
Cristina Martinez
Shannon H. Hedvat
Kramer Levin Naftalis
& Frankel LLP

1177 Avenue of the Americas
New York, NY 10036
Tel: 212.715.9100
Fax: 212.715.8000
afrankel@kramrelevin.com
cmartinez@kramerlevin.com
shedvat@kramerlevin.com

Attorneys for Plaintiff-Appellant
Rohm Semiconductor USA, LLC

CERTIFICATE OF COMPLIANCE

1. This brief complies with the type-volume limitation of Fed. Cir. R. 32(b) because this brief contains 4,792 words, exclusive of the certificate of interest, table of contents, table of citations, statement of related cases, addendum and this certificate of compliance as exempted by Fed. R. App. 32(f) and Fed. Cir. R. 32(b)(2).

2. This brief complies with the typeface requirements of Fed. R. App. P. 32(a)(5) and the type style requirements of Fed. R. App. P. 32(a)(6) because this brief has been prepared in a proportionally-spaced typeface using Microsoft Word 2010 in Times New Roman 14-point font.

/s/ Lisa Kobialka
Lisa Kobialka

ADDENDUM

Date	Description	Appx No.
02/04/2021	Order Granting Motion to Compel Arbitration	Appx1
	U.S. Patent No. 7,843,004 (the ’004 Patent)	Appx3
	U.S. Patent No. 8,076,719 (the ’719 Patent)	Appx39
	U.S. Patent No. 8,466,025 (the ’025 Patent)	Appx74
	U.S. Patent No. 8,659,076 (the ’076 Patent)	Appx110

UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA

ROHM SEMICONDUCTOR USA, LLC, Plaintiff, v. MAXPOWER SEMICONDUCTOR, INC., Defendant.	Case No. 20-cv-06686-VC **ORDER GRANTING MOTION TO COMPEL ARBITRATION** Re: Dkt. No. 41

MaxPower's motion to compel arbitration is granted. MaxPower's and Rohm Co.'s Technology License Agreement (TLA) binds Rohm Co.'s subsidiaries, including Rohm USA. The TLA incorporates the California Code of Civil Procedure into its arbitration provision. The applicable section provides: "The arbitral tribunal may rule on its own jurisdiction, including ruling on any objections with respect to the existence or validity of the arbitration agreement . . . ." Cal. Civ. Proc. Code § 1297.161. That type of language has repeatedly been held to clearly and unmistakably delegate the question of arbitrability to the arbitrator. *See Oracle America, Inc. v. Myriad Group A.G.*, 724 F.3d 1069 (9th Cir. 2013); *Loewen v. McDonnell*, 403 F. Supp. 3d 832 (N.D. Cal. 2019); *Dream Theater, Inc. v. Dream Theater*, 124 Cal. App. 4th 547 (2004); *see also Henry Schein, Inc. v. Archer and White Sales, Inc.*, 139 S. Ct. 524 (2019). MaxPower's motion is thus granted, and the case is dismissed without prejudice. *See Johnmohammadi v. Bloomingdale's, Inc.*, 755 F.3d 1072, 1073-74 (9th Cir. 2014).[1]

[1] MaxPower's motion to seal, filed in conjunction with its motion to compel, is also granted. *See* Docket No. 42.

IT IS SO ORDERED.

Dated: February 4, 2021

VINCE CHHABRIA
United States District Judge

(12) **United States Patent**
Darwish

(10) **Patent No.: US 7,843,004 B2**
(45) **Date of Patent: Nov. 30, 2010**

(54) **POWER MOSFET WITH RECESSED FIELD PLATE**

(75) Inventor: **Mohamed N. Darwish**, Campbell, CA (US)

(73) Assignee: **MaxPower Semiconductor Inc.**, Santa Clara, CA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 388 days.

(21) Appl. No.: **11/903,972**

(22) Filed: **Sep. 25, 2007**

(65) **Prior Publication Data**

US 2008/0073707 A1 Mar. 27, 2008

Related U.S. Application Data

(60) Provisional application No. 60/847,551, filed on Sep. 27, 2006.

(51) **Int. Cl.**
H01L 29/66 (2006.01)

(52) **U.S. Cl.** **257/341**; 257/329; 257/E29.131; 257/E29.183; 257/E29.189; 370/466

(58) **Field of Classification Search** 257/328, 257/330, E29.131, E29.183, E29.189, E29.201, 257/E29.257; 370/466
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,879,980	A	3/1999	Selcuk et al.	438/238
5,907,776	A	5/1999	Hshieh et al.	
6,388,286	B1	5/2002	Baliga	257/330
6,429,481	B1	8/2002	Mo et al.	
6,521,497	B2	2/2003	Mo	
6,710,403	B2	3/2004	Sapp	
6,710,406	B2	3/2004	Mo et al.	
6,828,195	B2	12/2004	Mo et al.	
6,833,584	B2	12/2004	Henninger et al.	257/334
2005/0167742	A1 *	8/2005	Challa et al.	257/328
2005/0208722	A1	9/2005	Peake et al.	438/259
2006/0060916	A1 *	3/2006	Girdhar et al.	257/330
2006/0209887	A1 *	9/2006	Bhalla et al.	370/466

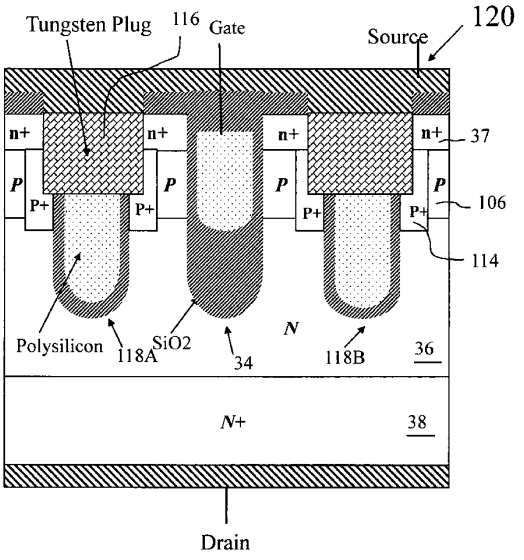
* cited by examiner

Primary Examiner—Kenneth A Parker
Assistant Examiner—Fang-Xing Jiang
(74) *Attorney, Agent, or Firm*—Groover & Associates

(57) **ABSTRACT**

A trench MOSFET contains a recessed field plate (RFP) trench adjacent the gate trench. The RFP trench contains an RFP electrode insulated from the die by a dielectric layer along the walls of the RFP trench. The gate trench has a thick bottom oxide layer, and the gate and RFP trenches are preferably formed in the same processing step and are of substantially the same depth. When the MOSFET operates in the third quadrant (with the source/body-to-drain junction forward-biased), the combined effect of the RFP and gate electrodes significantly reduces in the minority carrier diffusion current and reverse-recovery charge. The RFP electrode also functions as a recessed field plates to reduce the electric field in the channel regions when the MOSFET source/body to-drain junction reverse-biased.

27 Claims, 29 Drawing Sheets



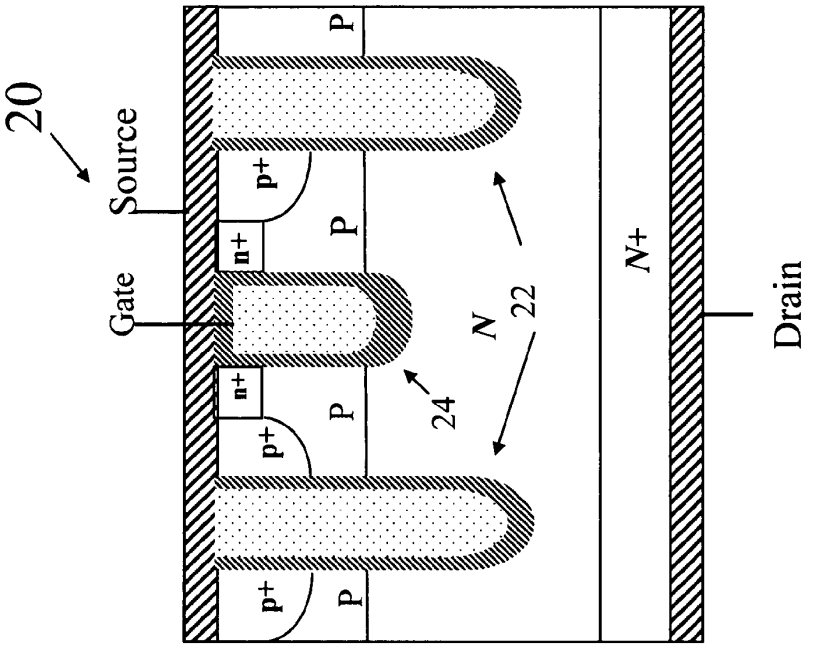

20
Source
Gate
P
p+
n+
P
N
22
24
N+
Drain

Fig. 2

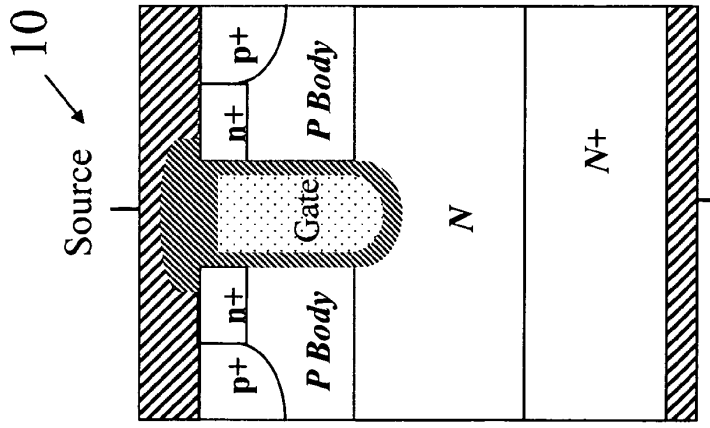

10
Source
p+
n+
P Body
Gate
N
N+
Drain

Fig.1

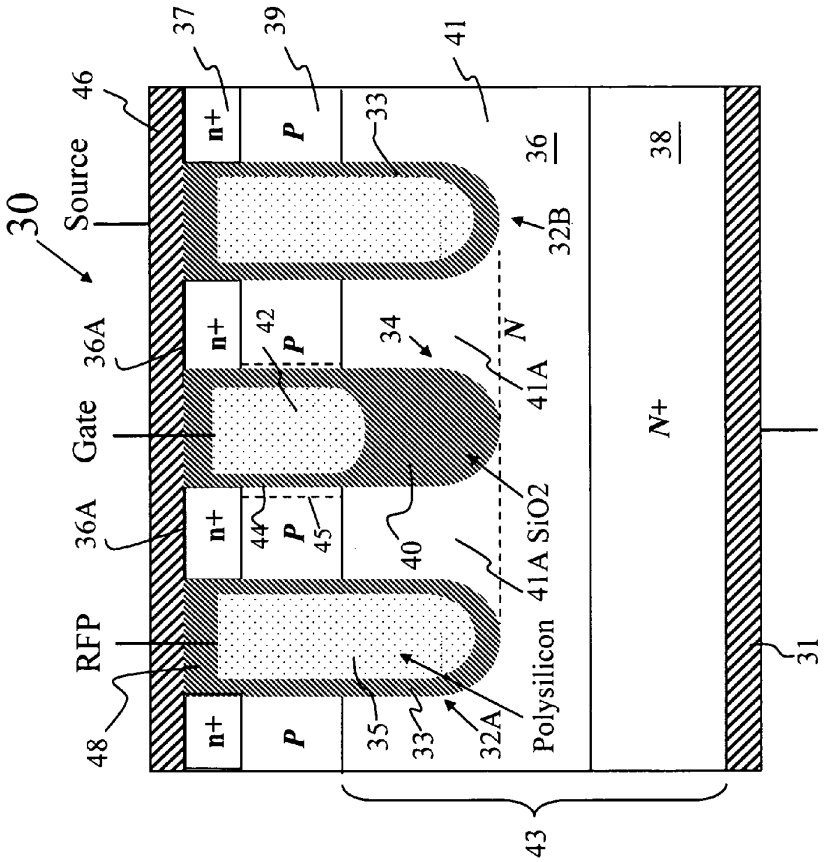

30
Source
Gate
RFP
Drain
46
37
39
41
33
36
38
32B
36A
42
34
41A
44
45
40
41A SiO2
48
35
33
32A
Polysilicon
31
43
n+
P
N
N+

Fig. 3A

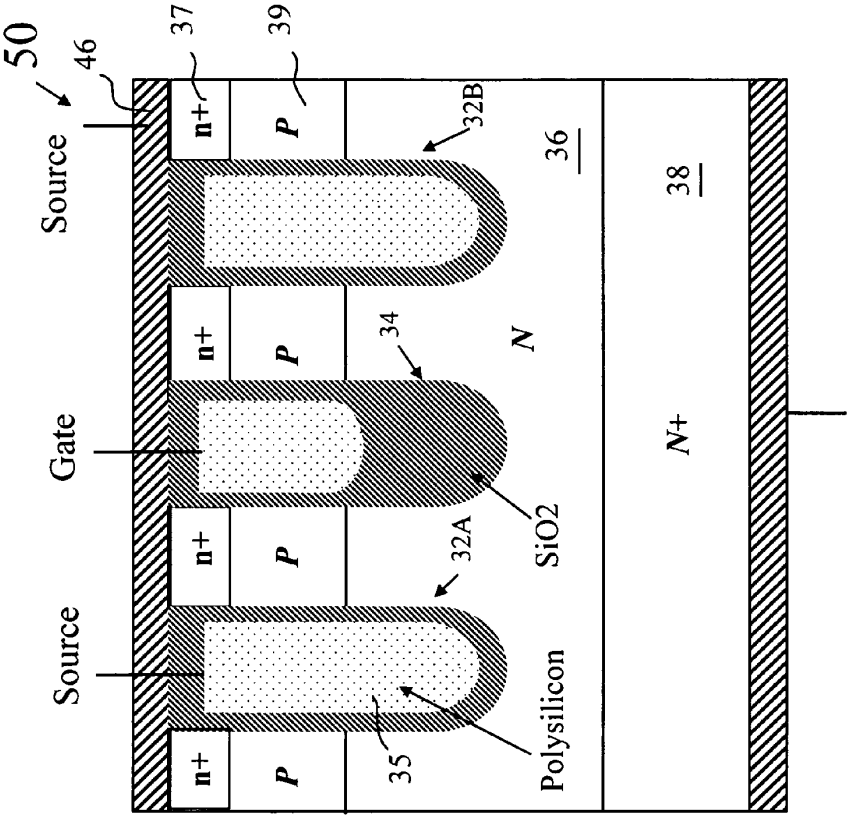

50
46
37
39
Source
n+
P
32B
36
38
n+
P
34
N
Gate
n+
P
32A
SiO2
N+
Drain
Source
n+
P
35
Polysilicon

Fig. 3B

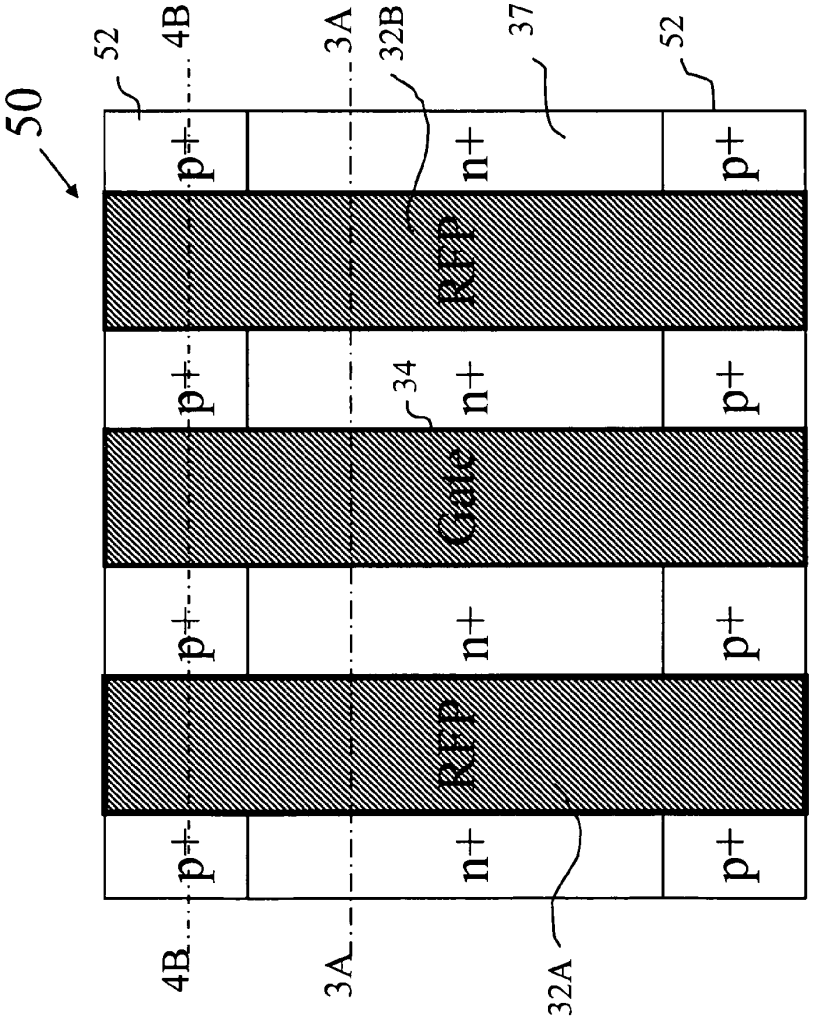


Fig. 4A

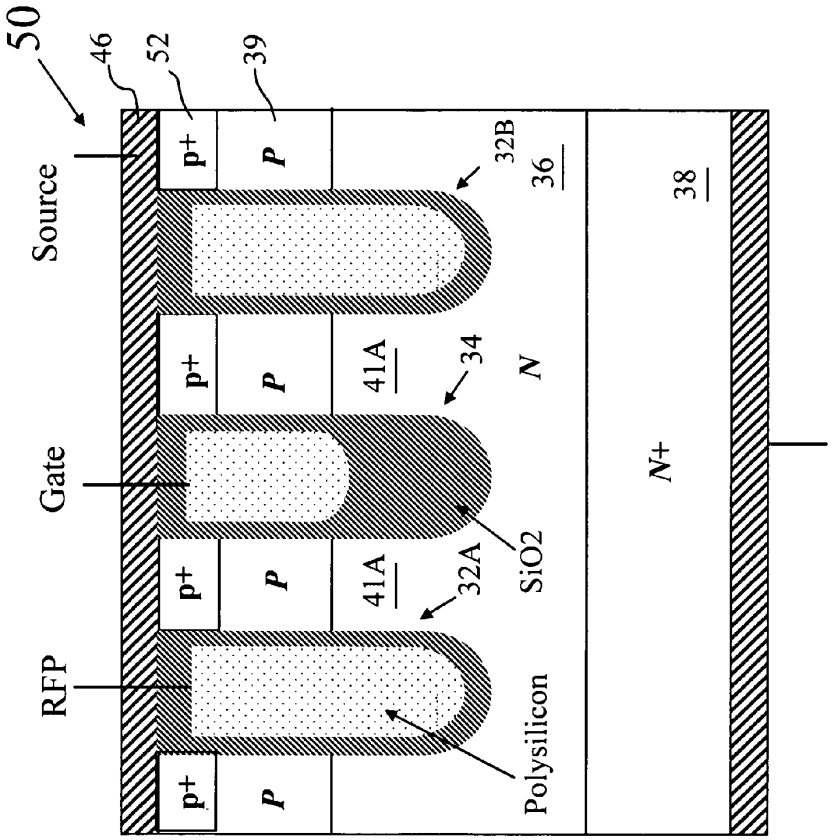

50
46
52
39
Source
Gate
RFP
p+
P
32B
36
38
41A
34
N
N+
SiO2
32A
Polysilicon
Drain

Fig. 4B

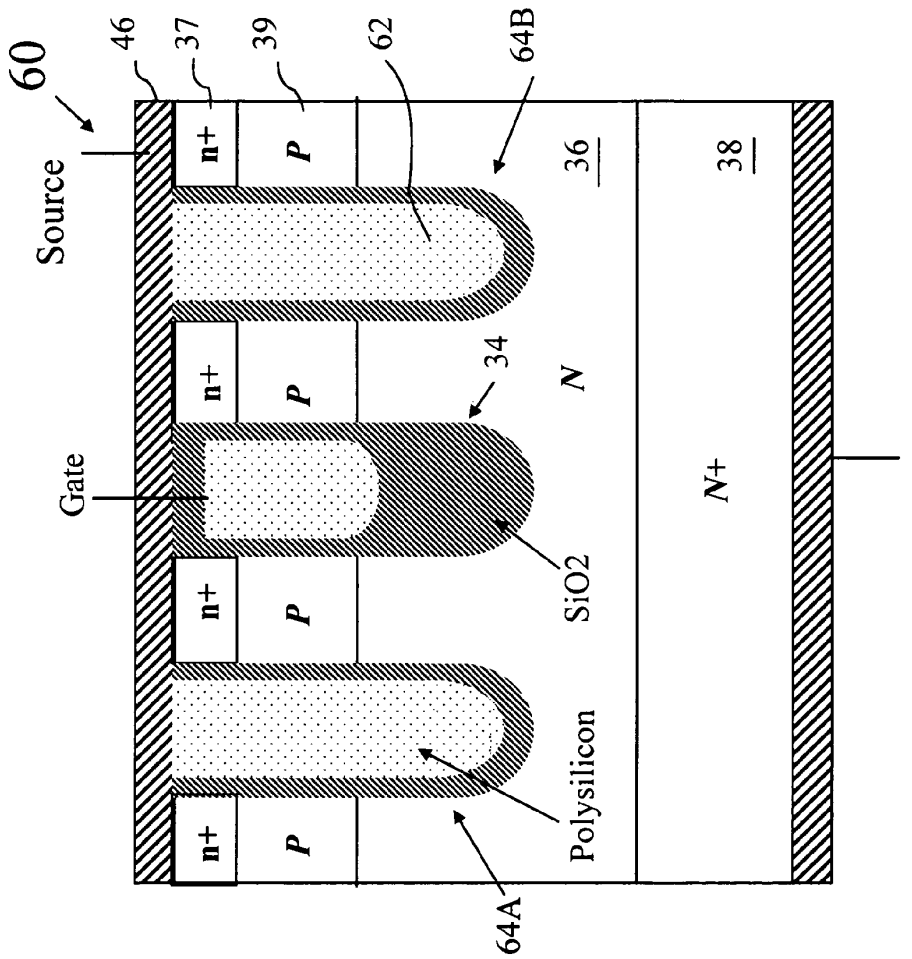

60
46
37
39
62
64B
Source
n+
P
36
38
34
N
Gate
N+
Drain
SiO2
Polysilicon
64A

Fig. 5A

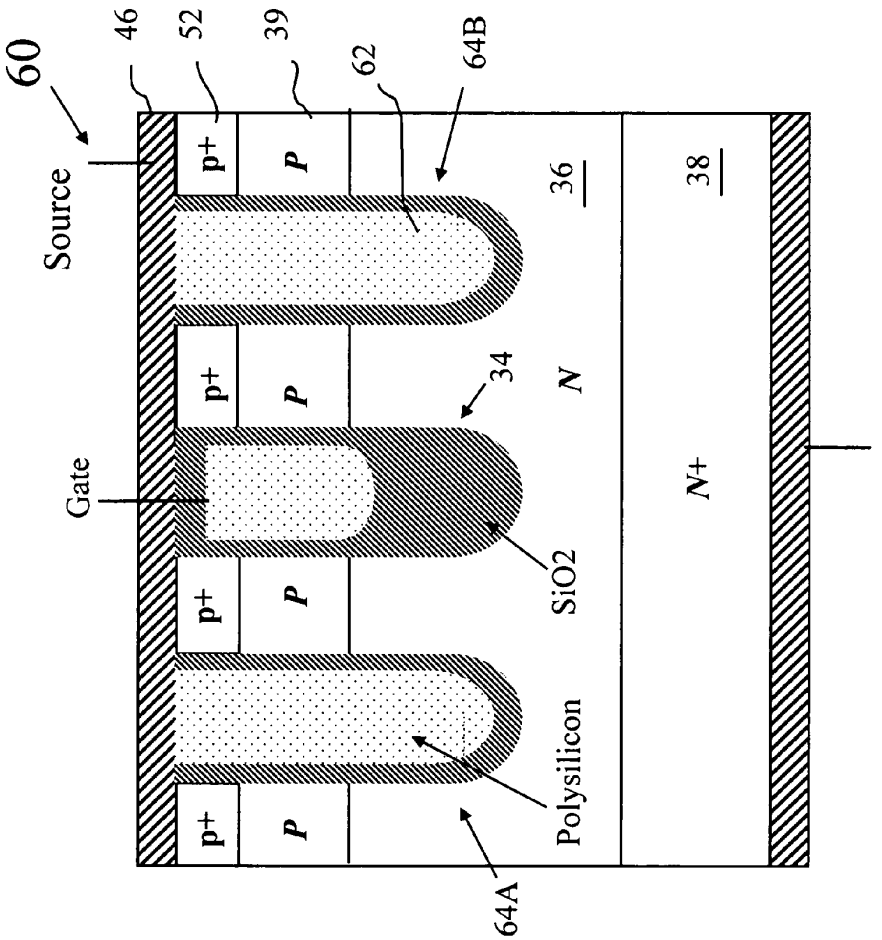

60
46
52
39
62
64B
Source
p+
P
36
38
34
N
Gate
N+
Drain
SiO2
Polysilicon
64A

Fig. 5B

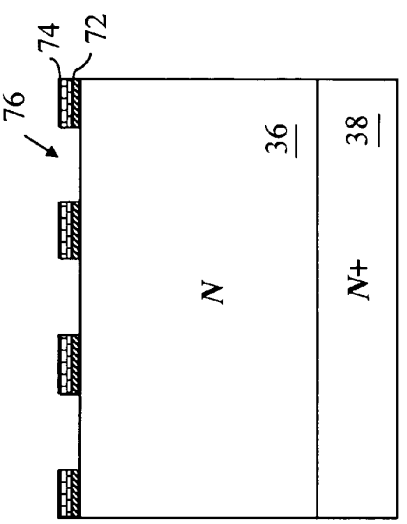


Fig. 6B

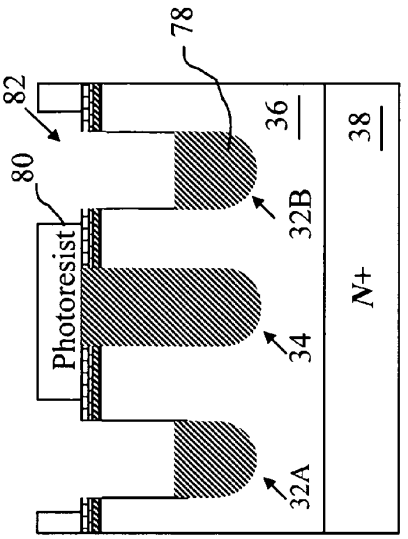


Fig. 6D

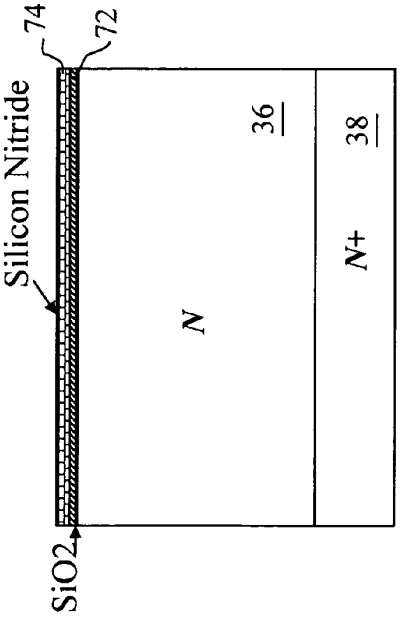


Fig.6A

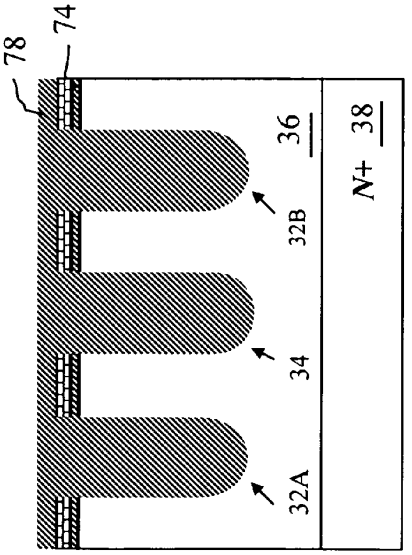


Fig. 6C

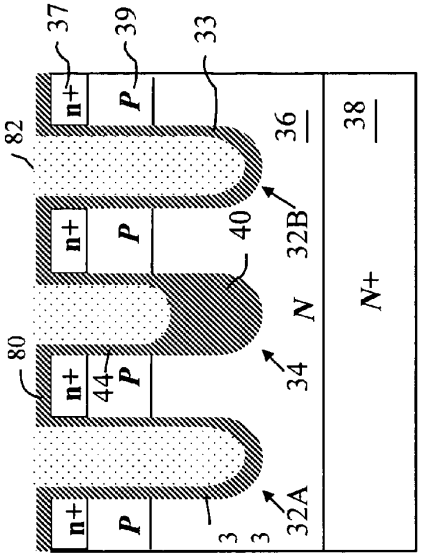


Fig. 6F

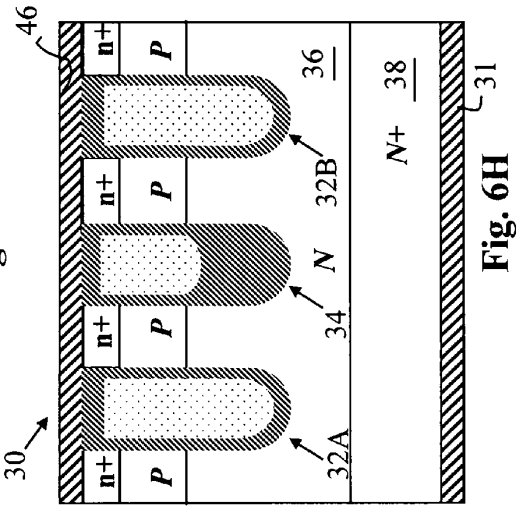


Fig. 6H

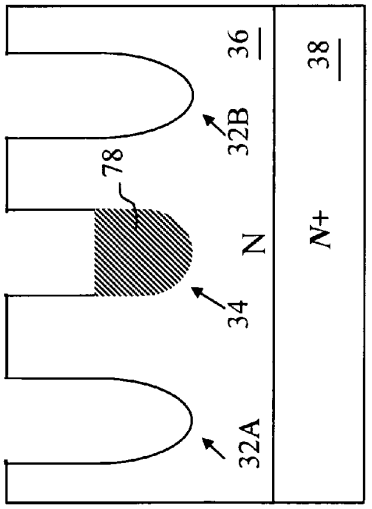


Fig.6E

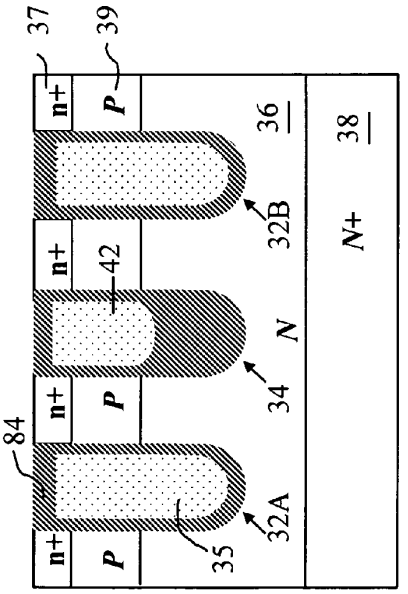


Fig. 6G

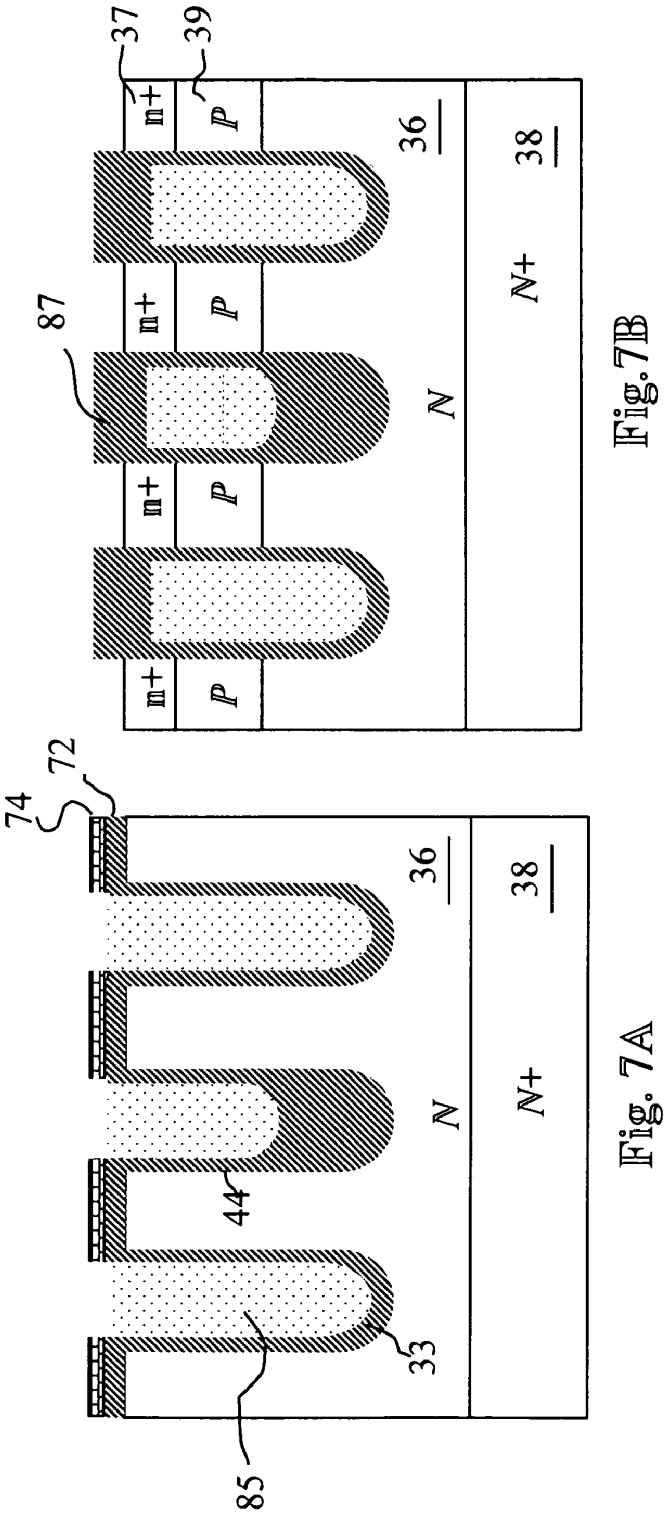


Fig. 7B

Fig. 7A

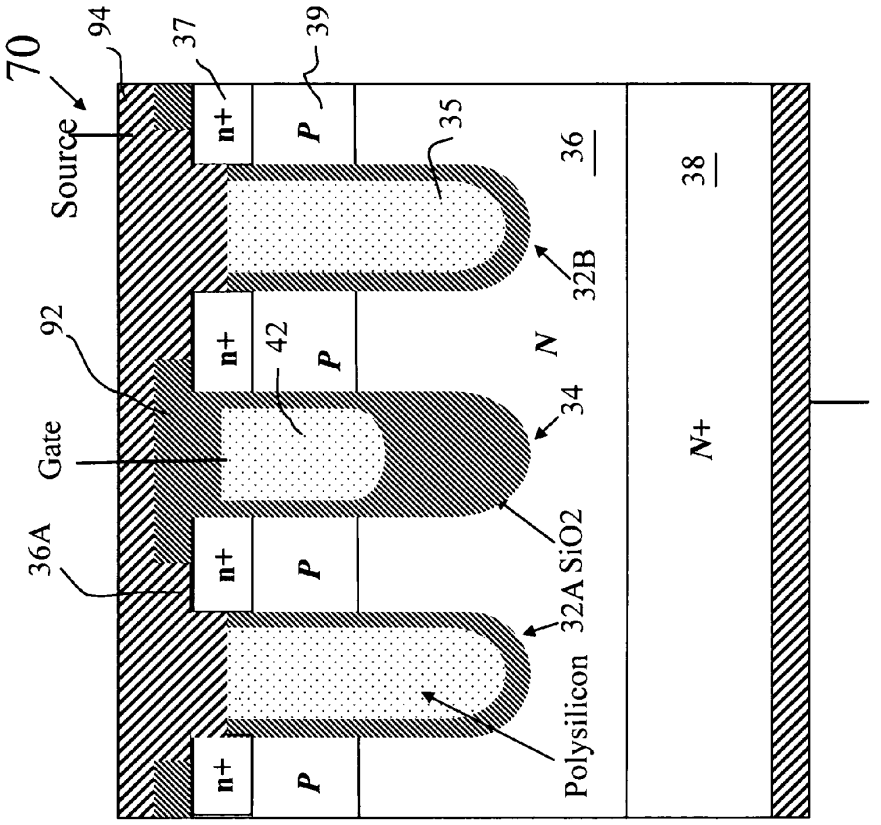

70
94
37
39
Source
n+
P
35
36
38
32B
92
n+
42
P
N
Gate
34
N+
Drain
36A
n+
P
32A SiO2
Polysilicon
n+
P

Fig. 8A

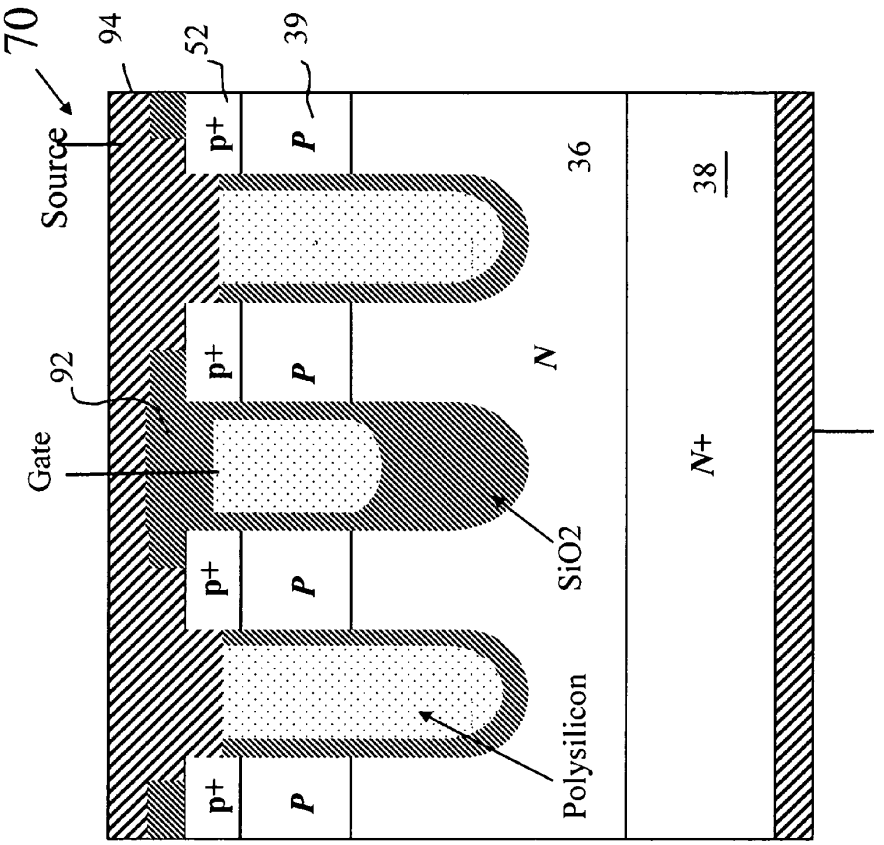


Fig. 8B

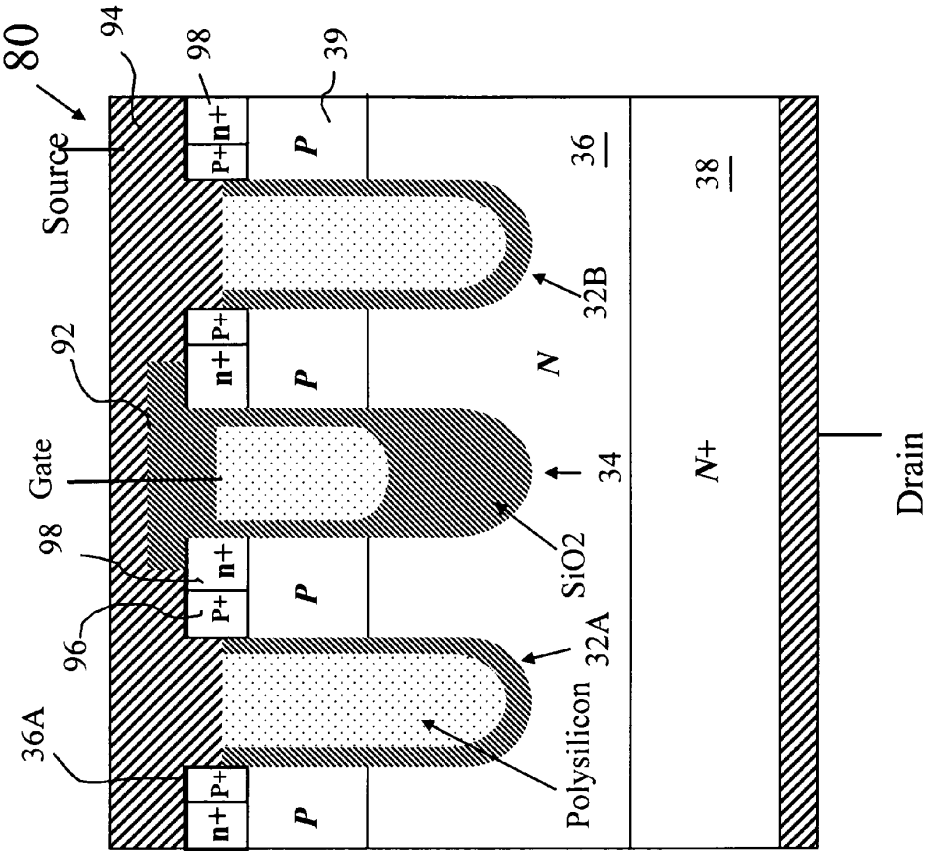

80
94
98
39
Source
P+
n+
P
36
38
32B
92
N
Gate
34
N+
Drain
98
SiO2
96
32A
36A
Polysilicon

Fig. 9

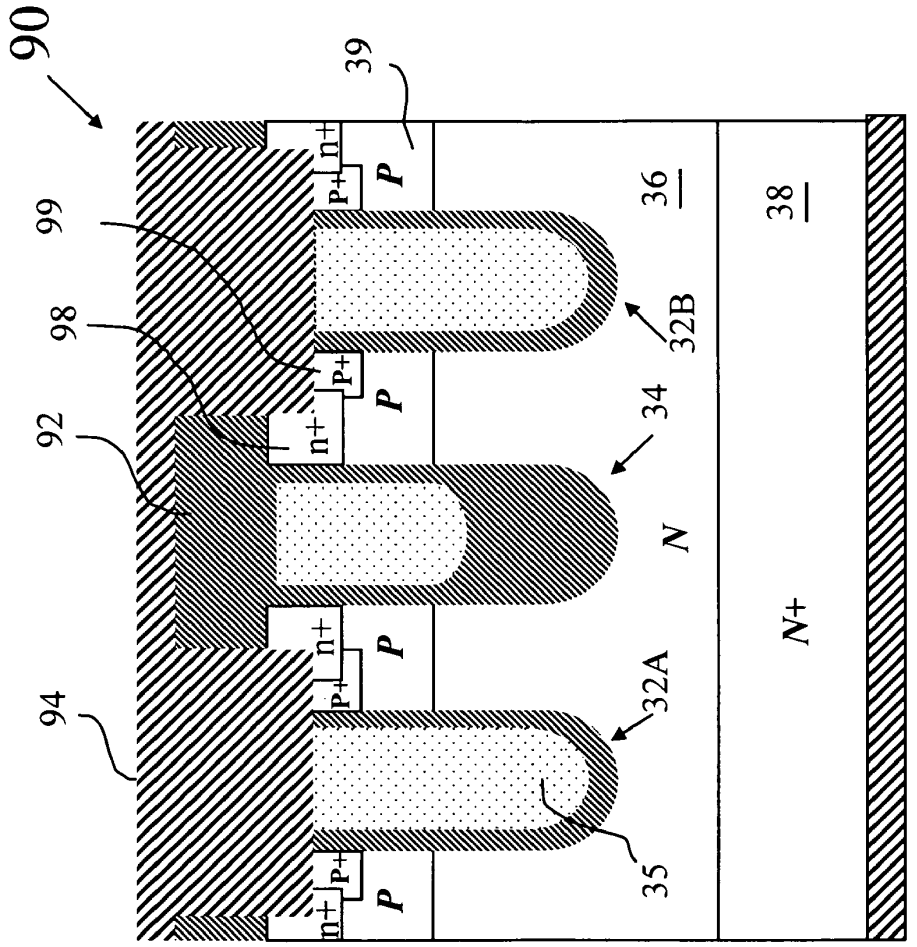


Fig. 10

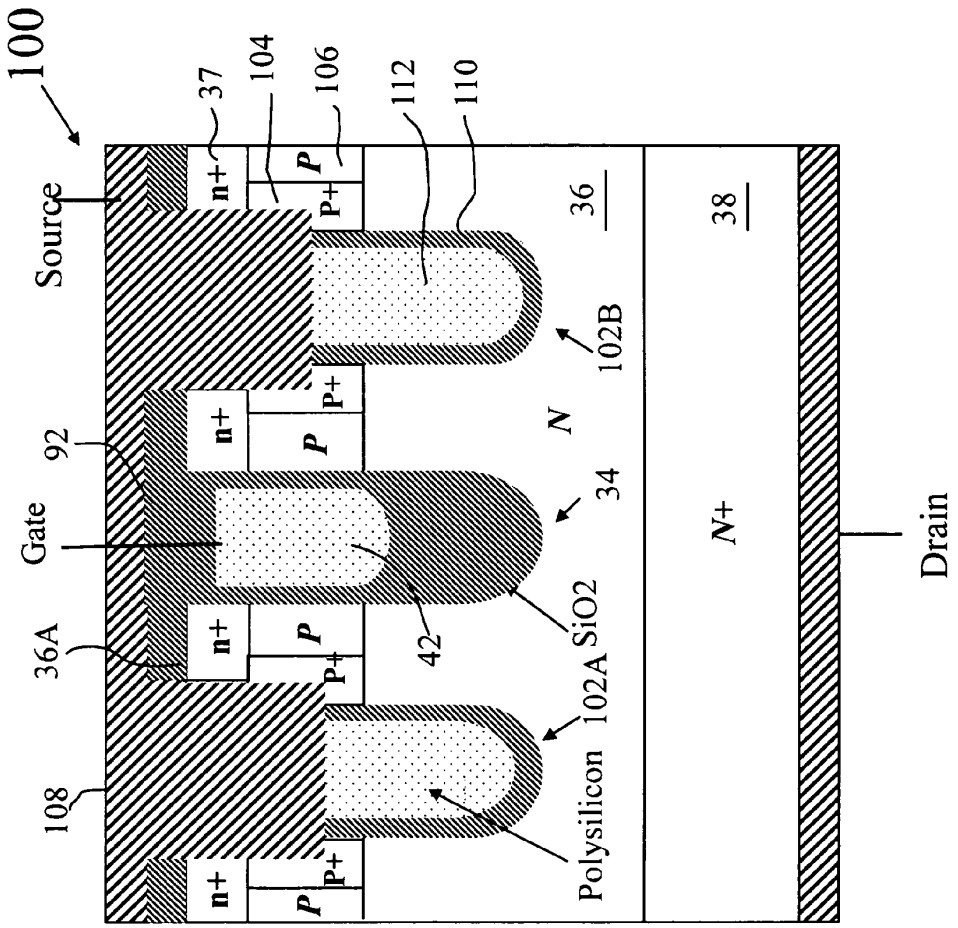

100
Source
37
104
106
112
110
36
38
102B
N
34
N+
Drain
92
Gate
36A
108
42
SiO2
102A
Polysilicon
n+
P
P+

Fig. 11

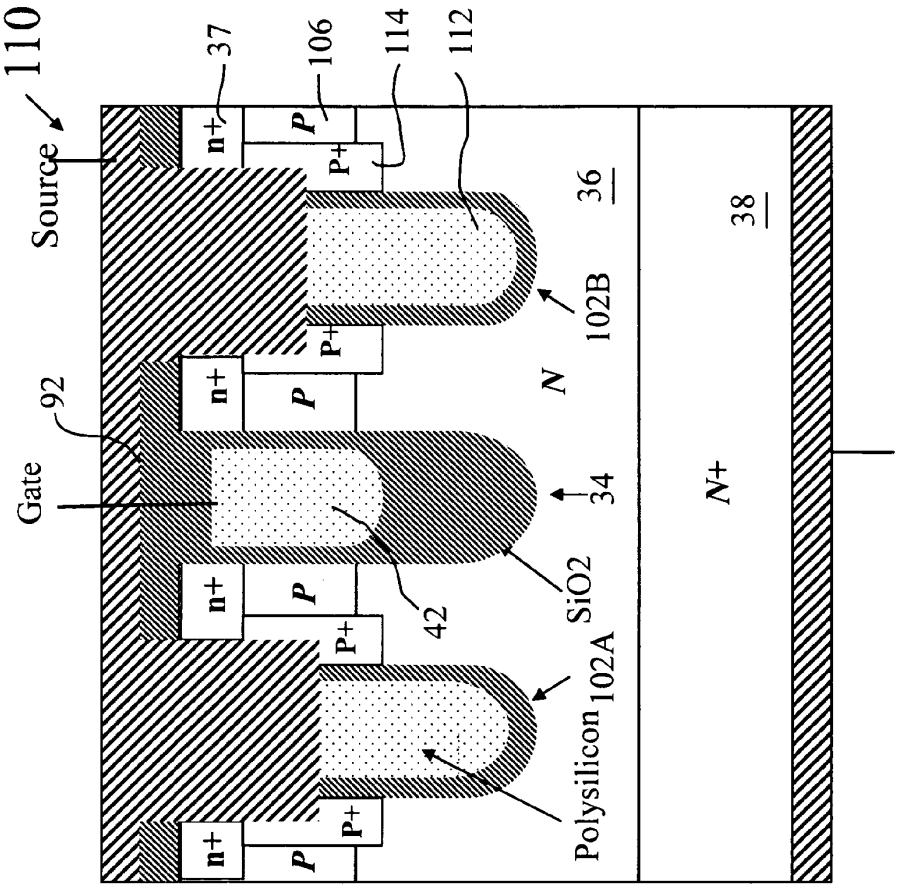

110
Source
37
106
114
112
n+
P
P+
36
38
102B
N
92
Gate
34
N+
Drain
SiO2
42
102A
Polysilicon

Fig. 12

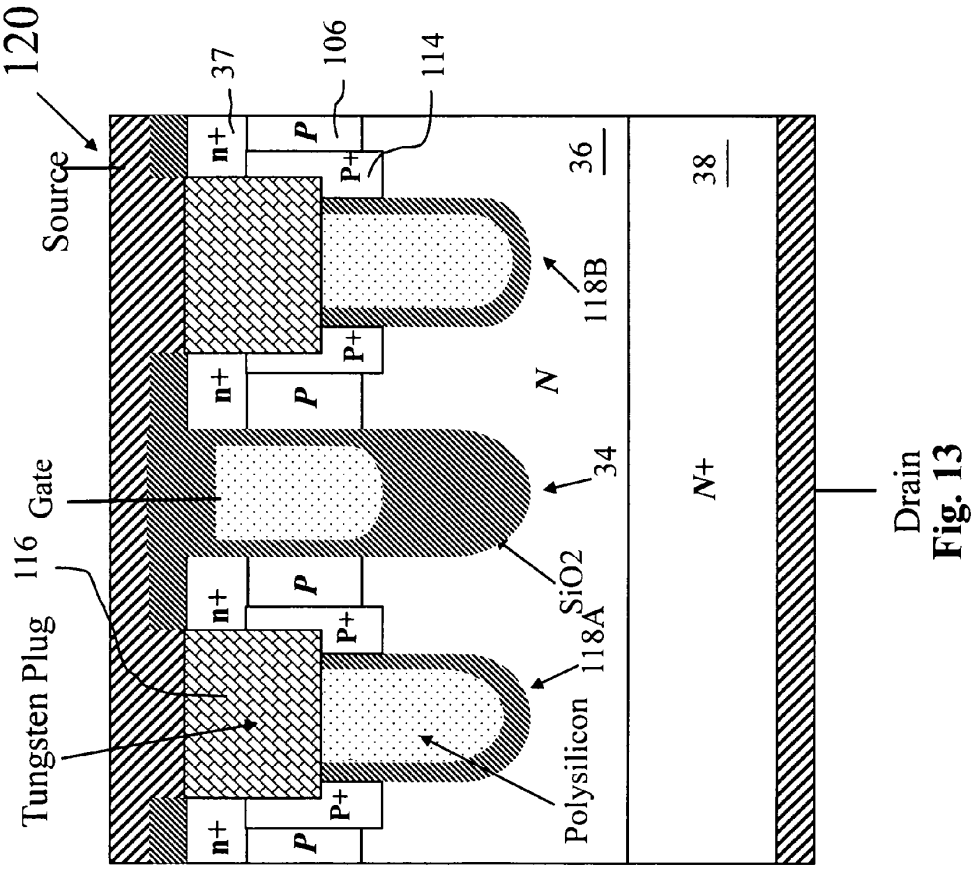
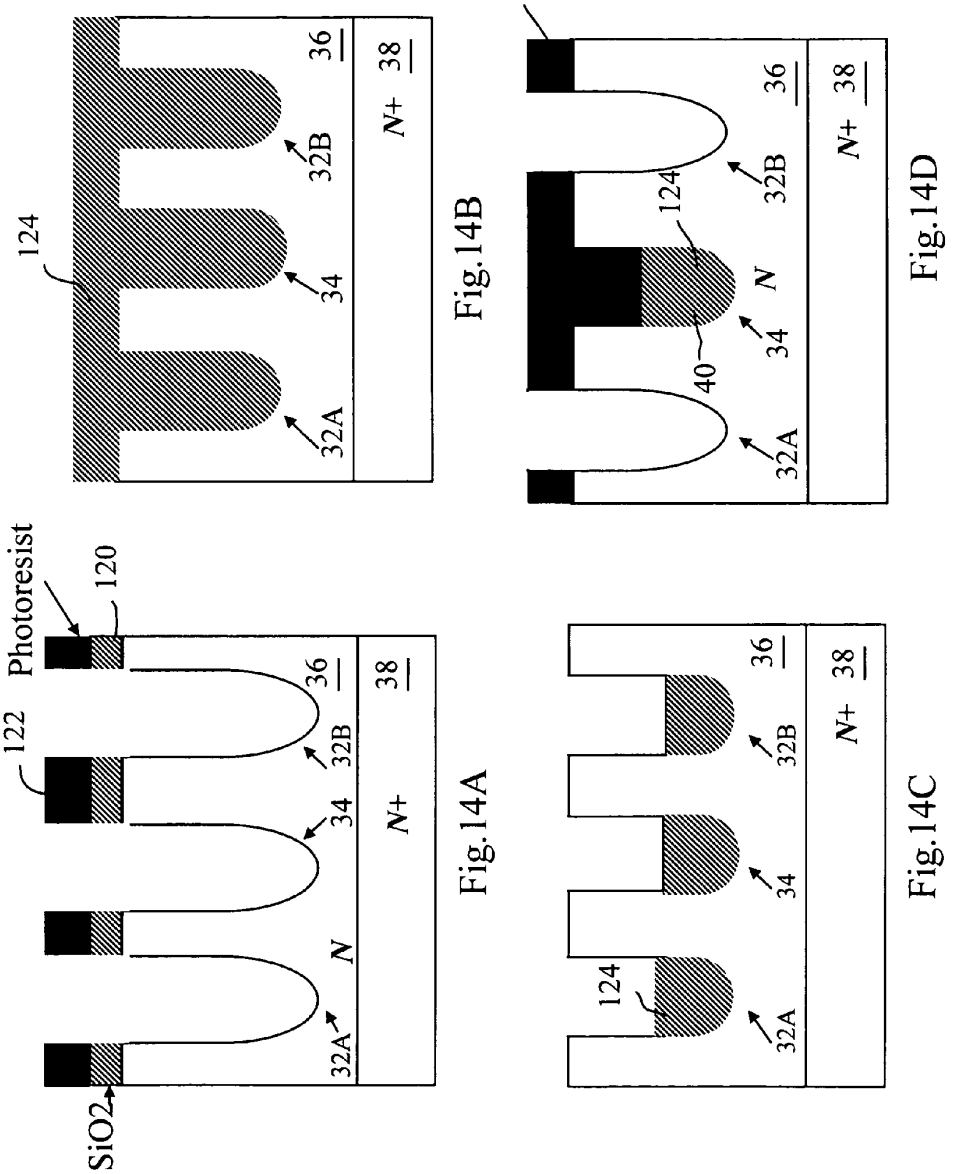

120
Source
37
106
114
n+
P
P+
36
38
118B
N
Gate
34
N+
Drain
116
Tungsten Plug
SiO2
118A
Polysilicon

Fig. 13


SiO2
Photoresist
120
122
32A
N
34
32B
36
N+
38
Fig.14A
124
Fig.14B
124
Fig.14C
126
40
Fig.14D

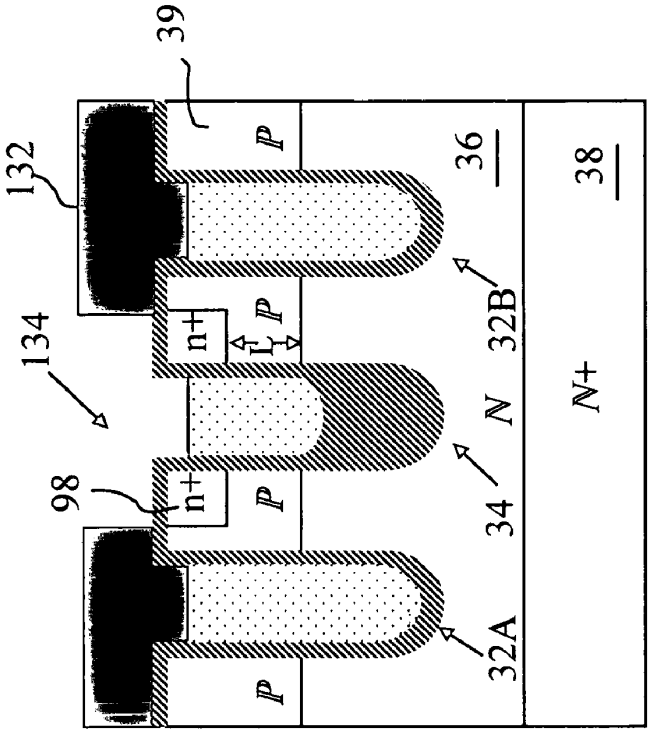


Fig. 14F

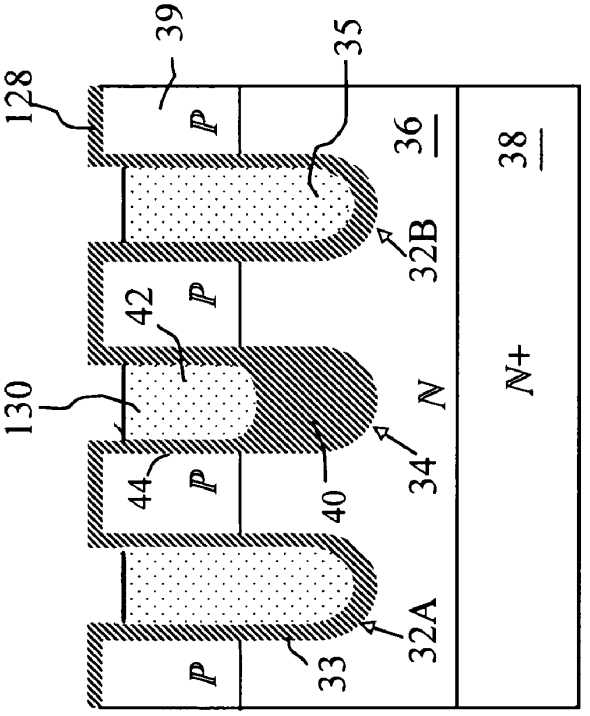


Fig. 14E

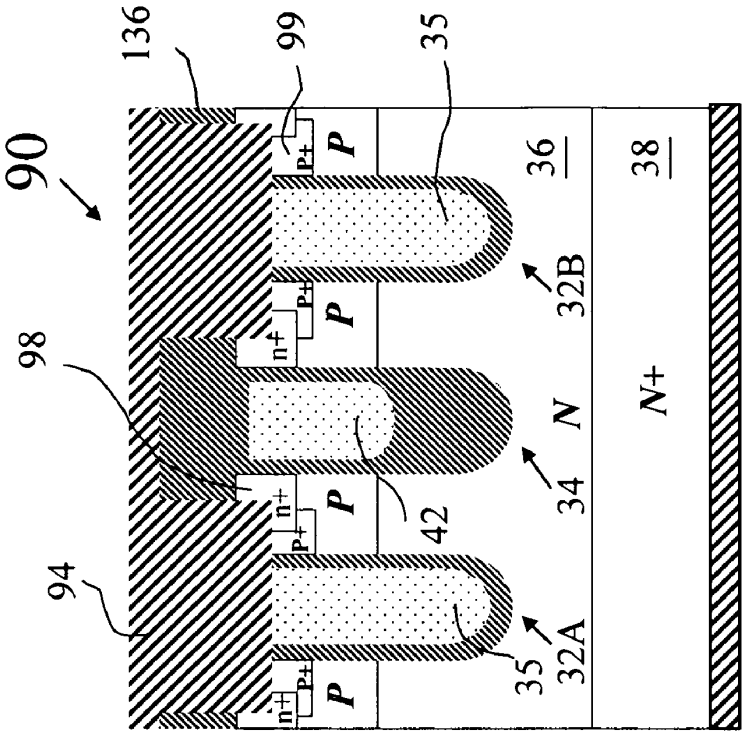


Fig. 14H

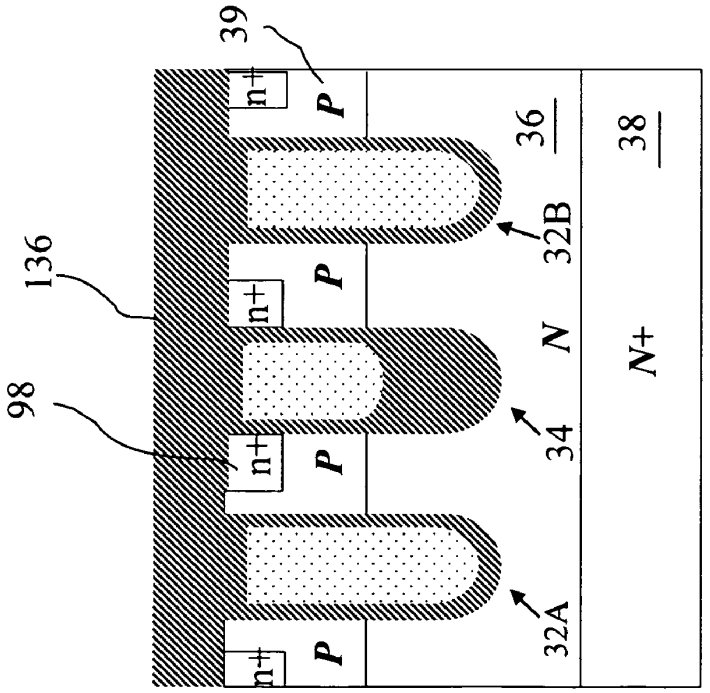


Fig. 14G

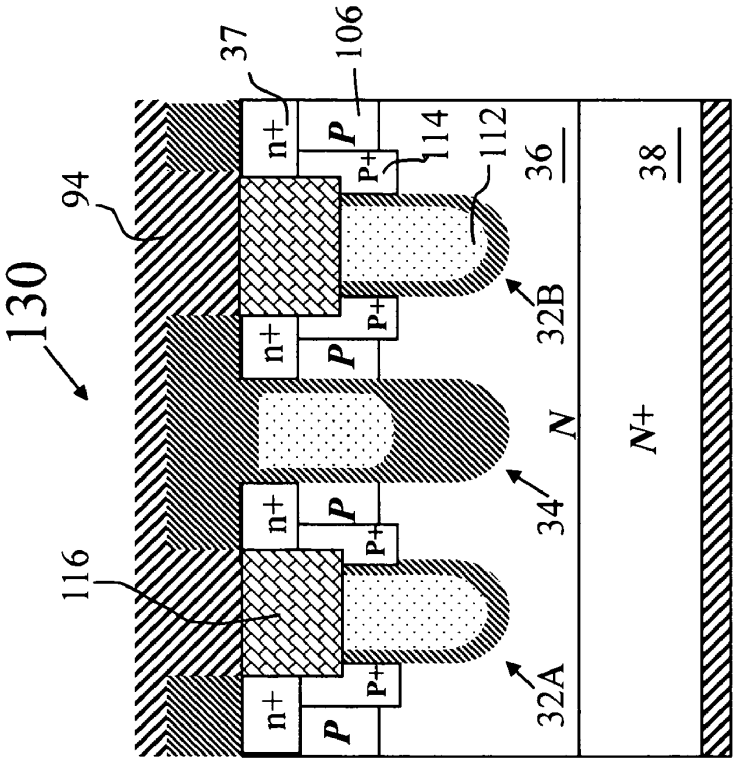


Fig. 15B

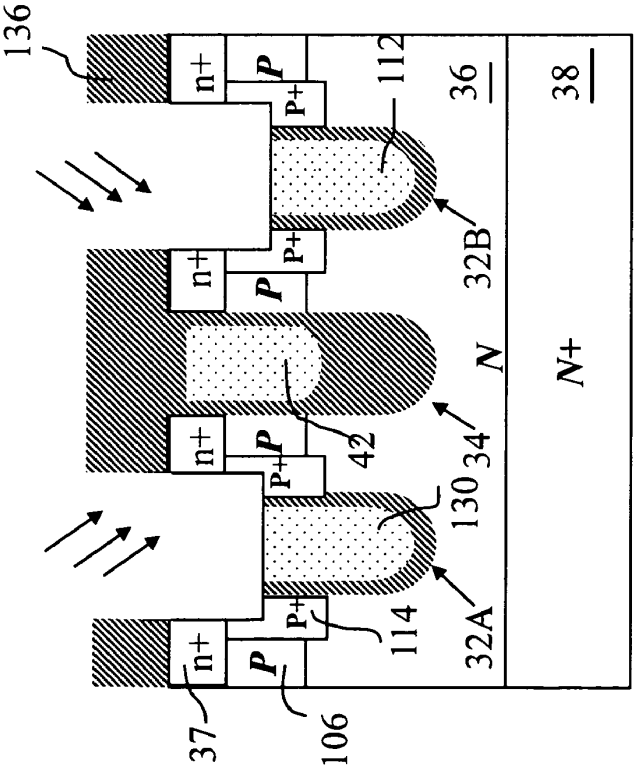


Fig. 15A

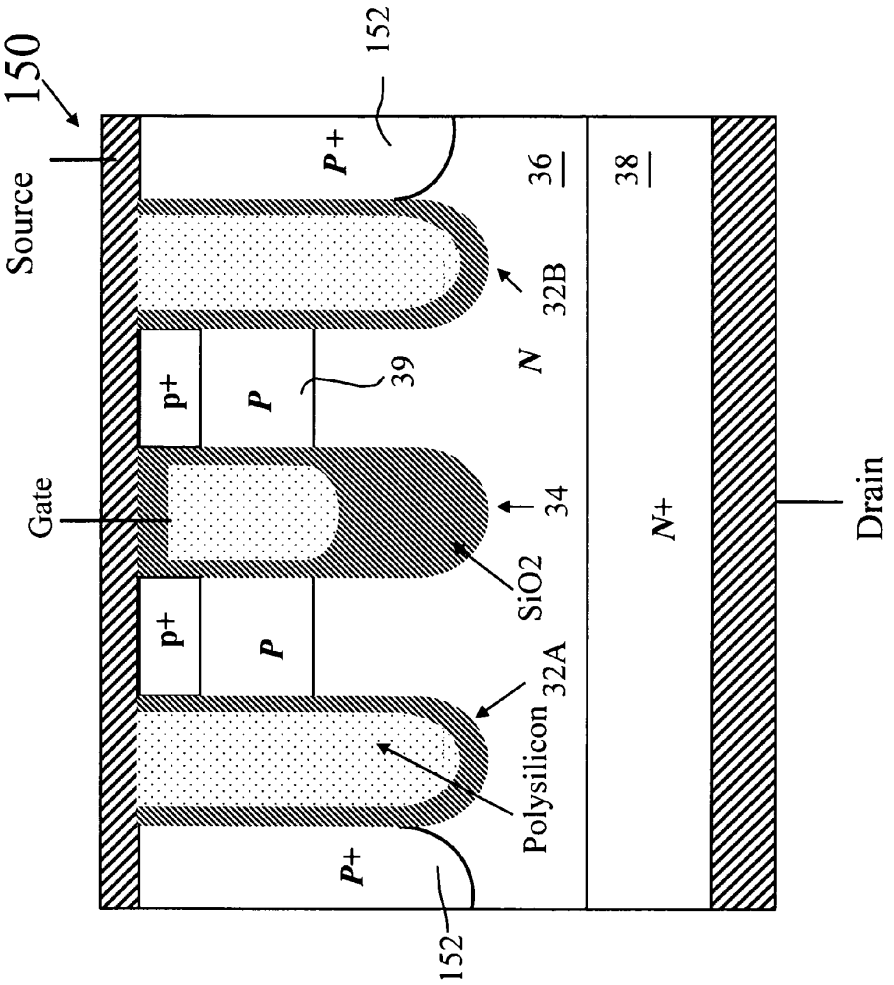

150
Source
152
P+
36
38
32B
N
39
p+
P
Gate
34
N+
Drain
SiO2
32A
Polysilicon
P+
152

Fig.16A

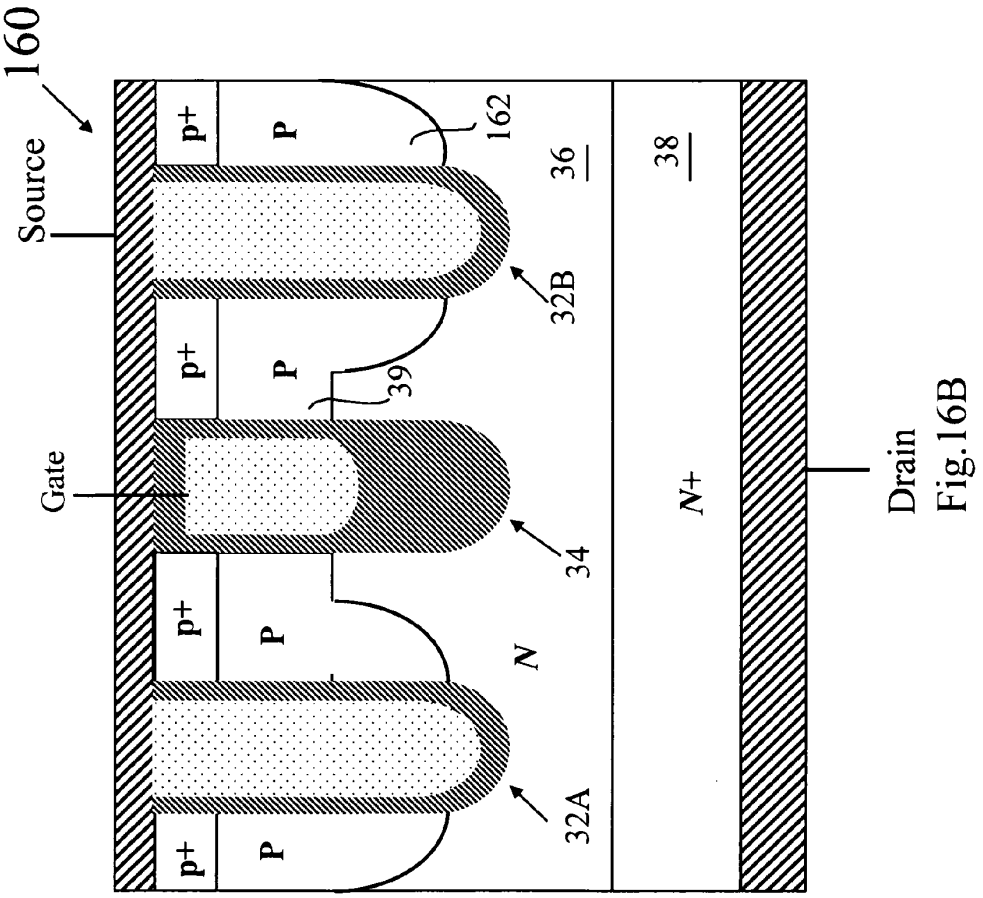


Fig.16B

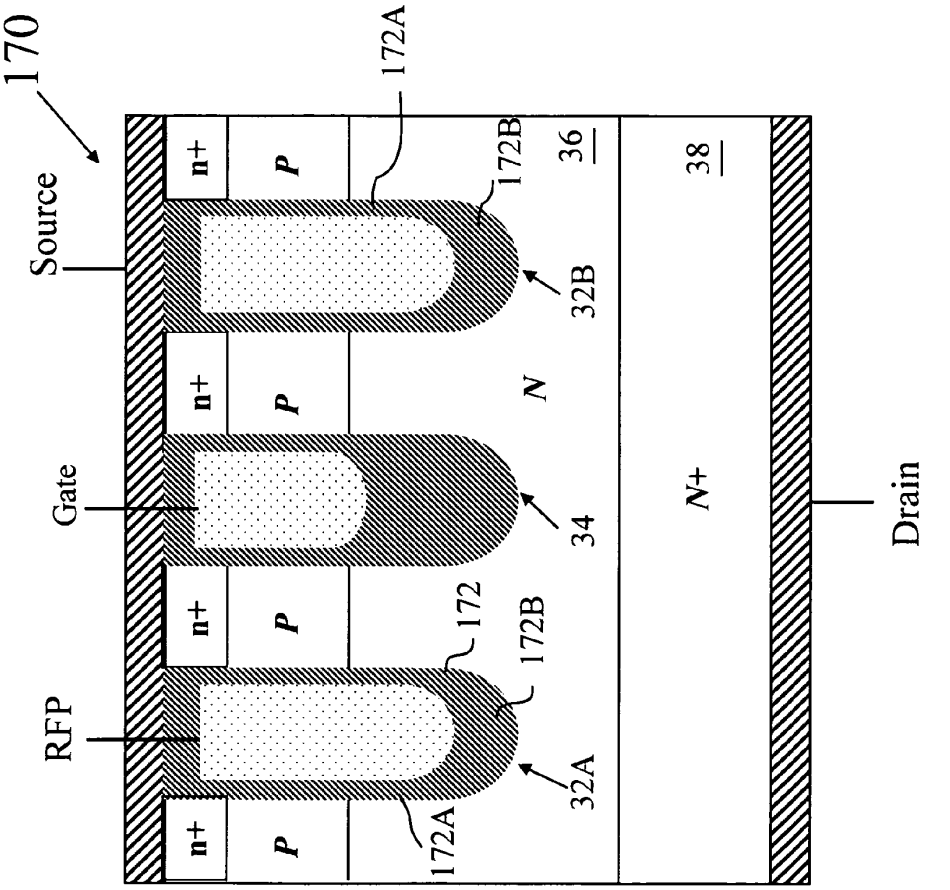


Fig. 17

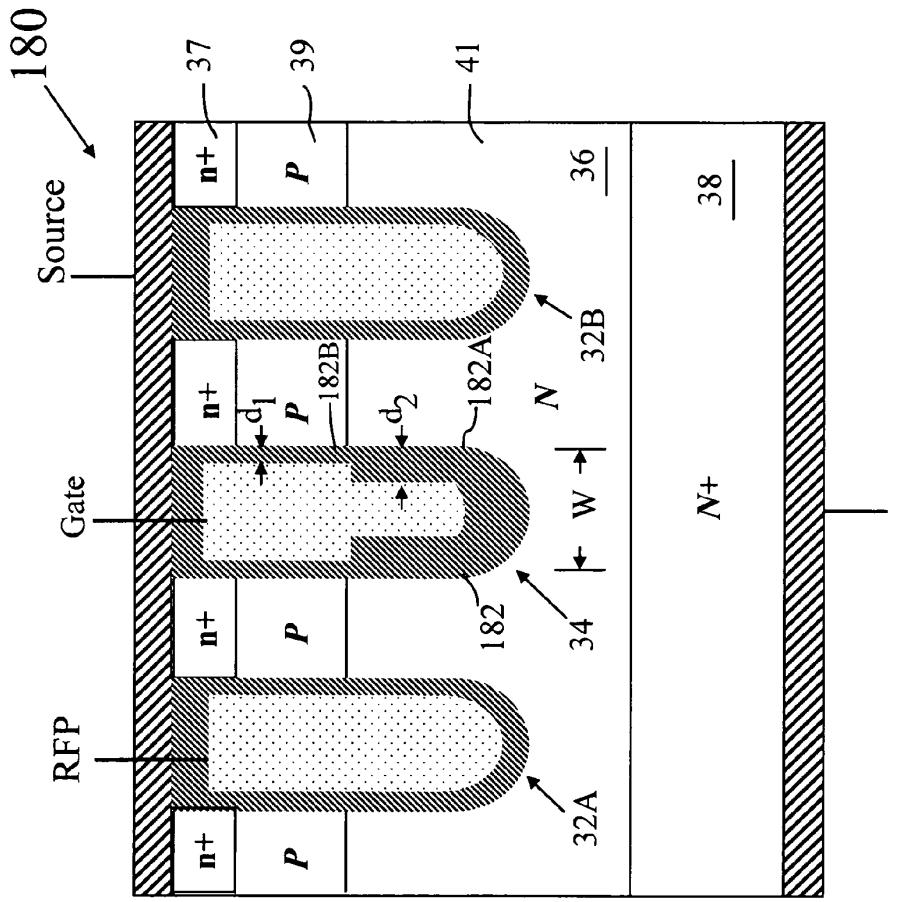

180
37
39
41
n+
P
36
38
Source
32B
182A
n+
P
182B
d_1
d_2
N
Gate
W
N+
Drain
182
34
n+
P
RFP
32A
n+
P

Fig. 18A

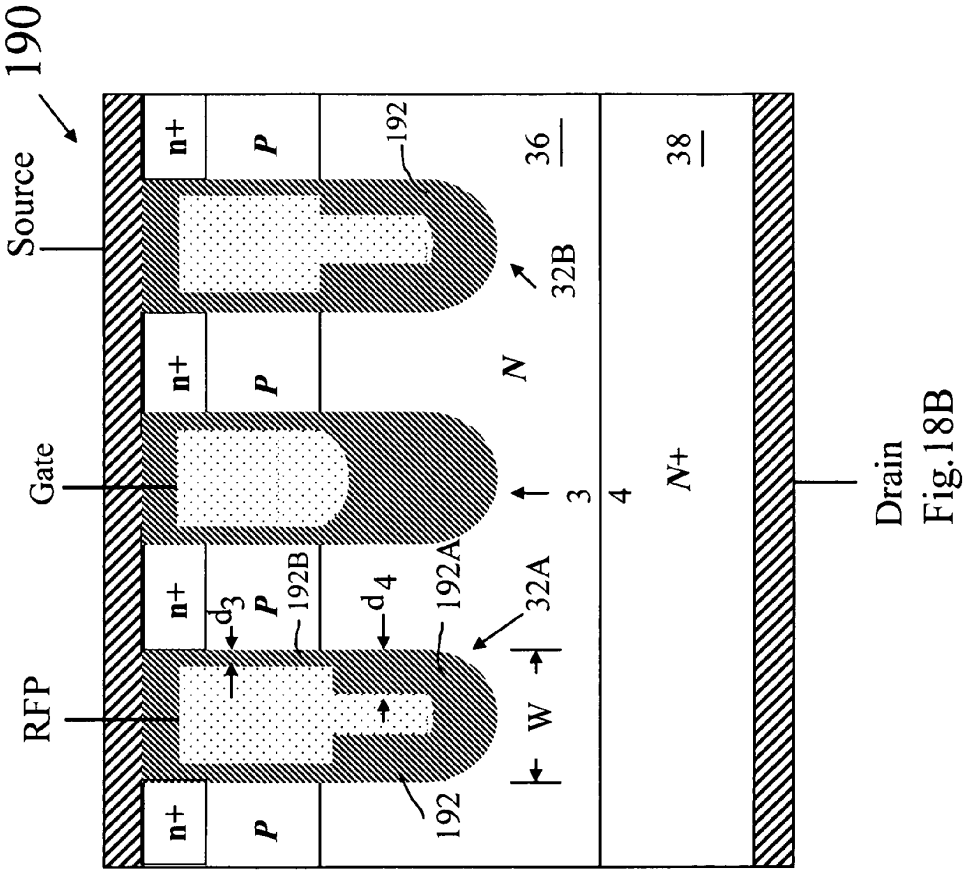

190
Source
Gate
RFP
Drain
n+
P
N
N+
192
192A
192B
32A
32B
36
38
3
4
W
d_3
d_4

Fig.18B

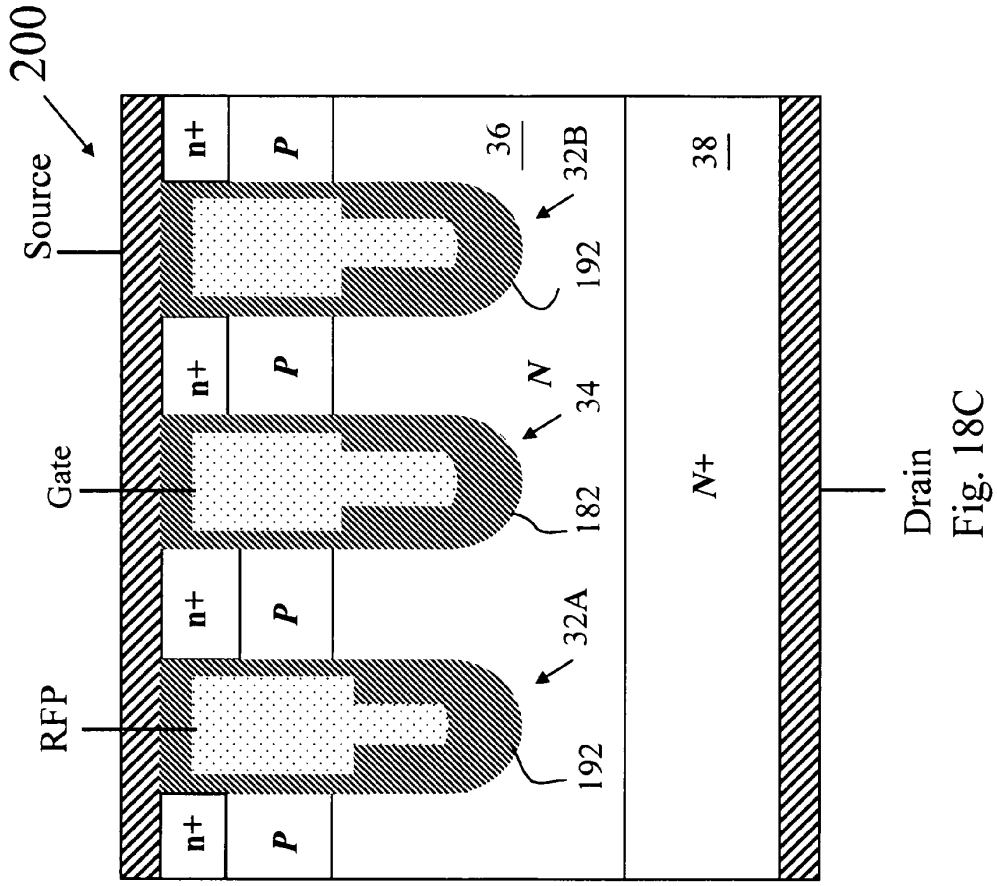


Fig. 18C

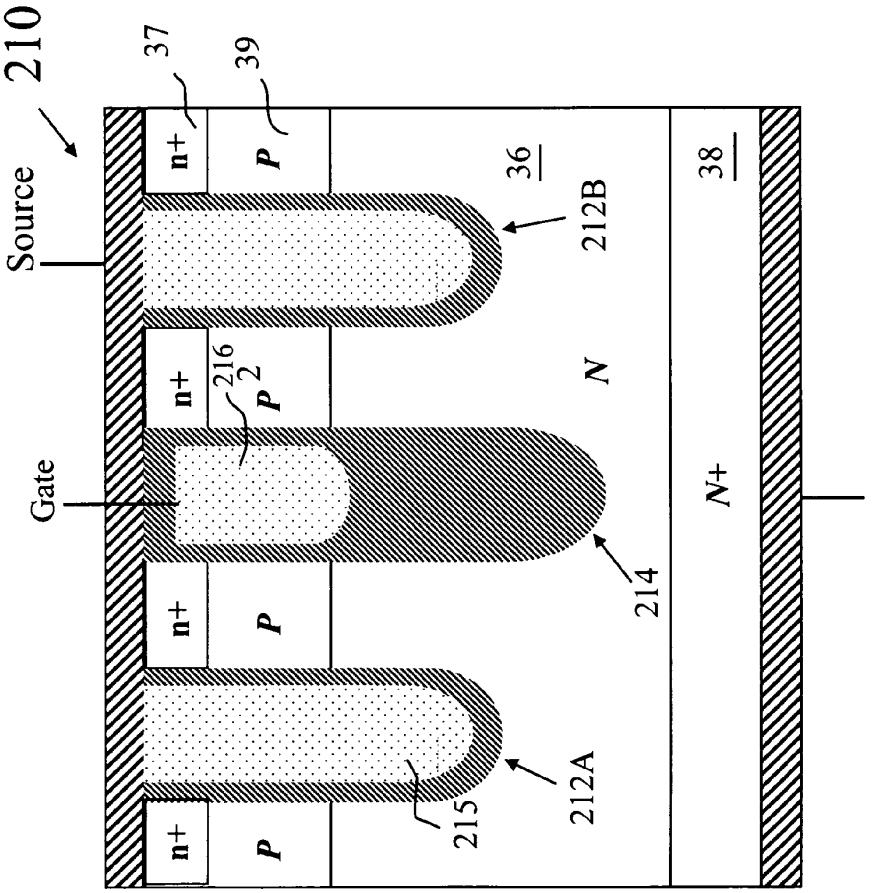

210
37
39
Source
n+
P
36
212B
38
n+
216
2
P
N
Gate
N+
Drain
n+
P
214
n+
P
215
212A

Fig. 19

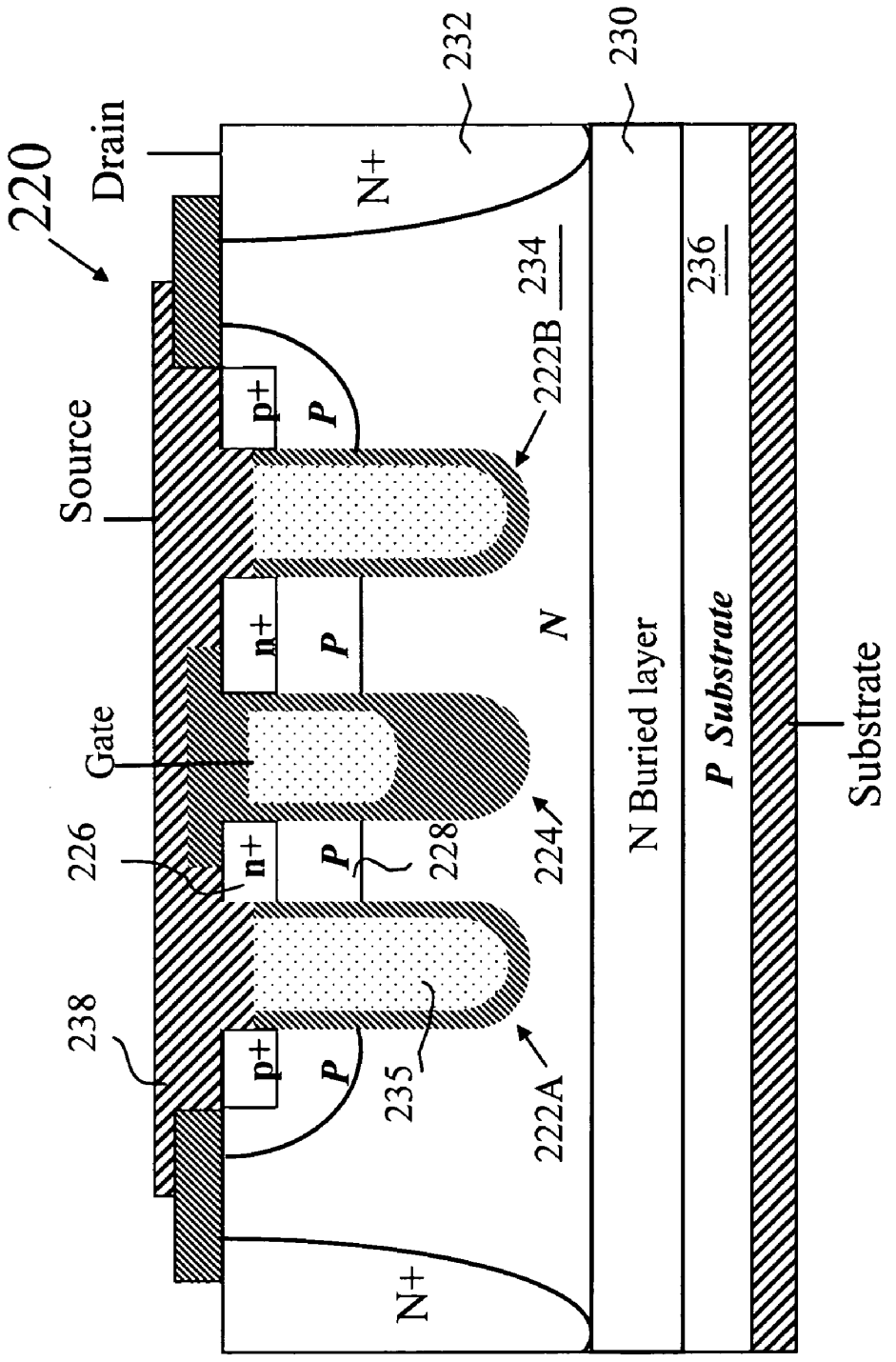

220
Drain
Source
Gate
N+
232
230
234
222B
236
p+
P
n+
P
N
226
228
224
N Buried layer
P Substrate
Substrate
238
235
222A
N+

Fig. 20

POWER MOSFET WITH RECESSED FIELD PLATE

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the priority of Provisional Application No. 60/847,551, filed Sep. 27, 2006, which is incorporated herein by reference in its entirety.

BACKGROUND OF THE INVENTION

Power MOSFETs are widely used as switching devices in many electronic applications. In order to minimize conduction power loss it is desirable that power MOSFETs have a low specific on-resistance, which is defined as the product of the on-resistance (Ron) of the MOSFET multiplied by the active die area (A) of the MOSFET (Ron*A). A trench-type MOSFET, as shown in the schematic cross-sectional view of a MOSFET **10** in FIG. **1**, provides a low specific on-resistance because of its high packing density or number of cells per unit area. As the cell density increases, the associated capacitances, such as the gate-to-source capacitance (Cgs), the gate-to-drain capacitance (Cgd), and the drain-to-source capacitance (Cds), also increase. In many switching applications such as the synchronous buck dc-dc converters used in mobile products, MOSFETs with a breakdown voltage in the range of 12 to 30V are required to operate at switching frequencies approaching 1 MHz. Therefore, it is desirable to minimize the switching or dynamic power loss caused by these capacitances. The magnitudes of these capacitances are directly proportional to the gate charge (Qg), the gate-drain charge (Qgd), and output charge (Qoss). Furthermore, when these devices operate in the third quadrant, i.e. where the drain-body junction is forward-biased, charge is stored as a result of minority carrier injection, and this stored charge causes a delay in switching speed of the device. It is therefore critical that a MOSFET switch have a low reverse recovery charge (Qrr).

U.S. Pat. No. 6,710,403 to Sapp proposes a dual-trench power MOSFET, as shown in FIG. **2**, with two deeper polysilicon-filled trenches **22** on either side of an active trench **24**, to lower the levels of Ron, Cgs and Cgd. However, MOSFET **20** does not lower the reverse recovery charge Qrr and requires the fabrication of trenches having two different depths. Furthermore, in MOSFET **20** the deep and shallow trenches are not self-aligned, which causes variations in mesa widths and hence in breakdown voltages.

As the switching-speed requirements increase to 1 MHz and above, driven by new applications such as CPU voltage regulator module (VRM), power MOSFETs are becoming increasingly unable to operate with satisfactory efficiency performance and power loss. Therefore, there is a clear need for a power MOS transistor that has low gate charges Qg and Qgd, a low output charge Qoss and a low reverse-recovery charge Qrr, in addition to having a low specific on-resistance (Ron*A).

BRIEF SUMMARY OF THE INVENTION

A MOSFET according to the invention is formed in a semiconductor die and comprises a gate trench and a recessed field plate (RFP) trench that are self-aligned, both trenches extending from a surface of the die and forming a mesa between them. The gate trench comprises a gate electrode separated from the die by a first dielectric layer having a thick section at the bottom of the gate trench and extends to substantially the same depth as the RFP trench. The RFP trench contains an RFP electrode separated from the die by a second dielectric layer. The MOSFET also comprises a source region of a first conductivity type adjacent the surface of the die and a sidewall of the gate trench and adjacent to the RFP electrode trench in some areas of the MOSFET and a body region of a second conductivity type opposite to the first conductivity type adjacent the sidewall of the gate trench and the source region. In some areas of the MOSFET, a p+ body contact region may be placed laterally adjacent to the P− body. The RFP electrode may be independently biased or may be biased at the source potential In one embodiment, the respective depths of the gate and RFP trenches are substantially the same.

The invention also comprises a method of fabricating a MOSFET. The method comprises providing a semiconductor die; etching the die so as to form a gate trench and a recessed field plate (RFP) trench, the gate trench and the RFP trench extending from a surface of the die and being of substantially equal depth; forming an insulating layer at a bottom of the gate trench; forming a gate dielectric layer on a sidewall of the gate trench above the insulating layer; forming a second dielectric layer along the walls of the RFP trench; introducing conductive material into the gate trench to form a gate electrode; introducing conductive material into the RFP trench to form an RFP electrode; implanting dopant of a second conductivity type opposite to the first conductivity type to form a body region in the mesa adjacent the sidewall of the gate trench; implanting dopant of a first conductivity type to form a source region adjacent the surface of the die in the mesa; and depositing a source contact layer on the surface of the die in contact with the source region.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWING

FIG. **1** is a cross-sectional view of a conventional trench-type MOSFET.

FIG. **2** is a cross-sectional view of a known dual-trench MOSFET.

FIG. **3**A is a cross-sectional view of a MOSFET having a recessed field plate (RFP) in accordance with the invention, with the recessed field plate (RFP) electrode independently biased.

FIG. **3**B is a cross-sectional view of a MOSFET having a recessed field plate (RFP) in accordance with the invention, with the RFP electrode biased at the same potential as the source.

FIG. **4**A is a top view of the MOSFET shown in FIG. **3**A.

FIG. **4**B is a cross-sectional view of the MOSFET shown in FIG. **3**A, taken at cross-section **4**B-**4**B in FIG. **4**A.

FIG. **5**A is a cross-sectional view of an alternative embodiment wherein the RFP electrode is in contact with the source region.

FIG. **5**B is a cross-sectional view of the MOSFET of FIG. **5**A taken at a cross-section corresponding to cross-section **4**B-**4**B in FIG. **4**A.

FIGS. **6**A-**6**H illustrate a process of fabricating the MOSFET shown in FIG. **3**A.

FIGS. **7**A and B illustrate an alternative version of a portion of the process shown in FIGS. **6**A-**6**H.

FIGS. **8**A and **8**B are cross-sectional views of a MOSFET according to the invention wherein an insulating layer above the gate trench overlaps a portion of the source regions and the source contact layer contacts the RFP electrodes.

FIG. **9** is a cross-sectional view of a MOSFET similar to the MOSFET of FIGS. **8**A and **8**B wherein body contact regions are formed laterally adjacent to the source regions.

FIG. **10** is a cross-sectional view of an MOSFET similar to the MOSFET of FIG. **9** wherein the body contact regions extend to a level below the source regions.

FIG. **11** is a cross-sectional view of a MOSFET according to the invention wherein the RFP electrodes are recessed and body contact regions are formed laterally adjacent to the body regions.

FIG. **12** is a cross-sectional view of a MOSFET similar to the MOSFET of FIG. **11** wherein the body contact regions extend to a level below the body regions.

FIG. **13** is a cross-sectional view of a MOSFET similar to the MOSFET of FIG. **12** wherein metal plugs are formed in the upper portions of the RFP trenches.

FIGS. **14**A-**14**H illustrate a process of fabricating the MOSFET shown in FIG. **10**.

FIGS. **15**A and **15**B illustrate a variation of the process shown in FIGS. **14**A-**14**H.

FIGS. **16**A and **16**B are cross-sectional views of MOSFETs according to the invention that contain deep p-type regions to limit the breakdown voltage of the MOSFET.

FIG. **17** is a cross-sectional view of a MOSFETs according to the invention wherein the RFP trenches contain a thick bottom oxide layer.

FIGS. **18**A-**18**C are cross-sectional views of MOSFETs containing a stepped oxide layer in the gate trench and/or the RFP trenches.

FIG. **19** is a cross-sectional view of a MOSFET wherein the gate trench is deeper than the RFP trenches.

FIG. **20** is a cross-sectional view of a quasi-vertical MOSFET in accordance with the invention.

DETAILED DESCRIPTION OF THE INVENTION

A basic cell of an n-channel MOSFET **30** in accordance with this invention is shown in FIG. **3**A. MOSFET **30** is formed in a semiconductor die comprising an n-type epitaxial layer **36** that is grown over a heavily doped n+ substrate **38**. MOSFET **30** includes recessed field plate (RFP) trenches **32**A and **32**B located on either side a gate trench **34** in the n-type epitaxial layer **36**. Unlike the trenches in the MOSFET **20** shown in FIG. **2**, the thick bottom oxide of the gate trench **34** extends to substantially the same depth as the RFP trenches **32**A and **32**B. Furthermore, the RFP trenches **32**A and **32**B and the gate trench **34** are preferably formed in the same processing step and therefore are self-aligned (i.e., the RFP trenches **32**A and **32**B are equally spaced from gate trench **34** regardless of processing and alignment variations), and RFP trenches **32**A and **32**B and gate trench **34** are of equal or substantially equal depth (e.g., the respective depths of RFP trenches **32**A and **32**B are within +/−10%, or preferably +/−5%, of the depth of gate trench **32**). The mesas between the trenches contain n+ source regions **37** and p-body regions **39**, and below p-body regions **39** is an n-type drain-drift region **41** of epitaxial layer **36**. The drain-drift region **41** includes areas **41**A between RFP trenches **32**A and **32**B and gate trench **34**. The drain-drift region **41** and the n+ substrate **38** together form the drain **43** of MOSFET **30**.

In the prior art MOSFET **20** shown in FIG. **2**, when the body-drain junction is reverse-biased, the electric field strength is reduced and hence the breakdown voltage is increased due to the spread of the depletion region between the two deep trenches **22**, mainly in the drift region below the gate trench **24**. In MOSFET **30**, under the same conditions the electric field is further reduced because the depletion layer is initially limited to the narrower areas **41**A of drain-drift region **41** located between the thick oxide of the gate trench **34** and RFP trenches **32**A and **32**B. Moreover, since the trenches **22** in the prior art MOSFET **20** are deeper than the gate trench **24**, the electric field at the bottom of trenches **22** is higher, which sets a lower limit on the thickness of the oxide layer in trenches **22**. This limitation degrades the effectiveness of the trenches **22** in reducing the electric field in the channel of MOSFET **20**.

Referring again to FIG. **3**A, the walls of trenches **32**A and **32**B are lined with a layer **33** of an insulating material such as silicon dioxide (SiO_2) and trenches **32**A and **32**B contain RFP electrodes **35**, each of which comprises a layer of a conductive material such as n-type or p-type doped polysilicon. The oxide layer **33** preferably has a breakdown voltage that exceeds the breakdown voltage of MOSFET **30**. The gate trench **34** is filled with a layer **40** of insulating material up to the level of the p-n junctions between p-body regions **39** and the drain-drift region **41** of epitaxial layer **36**. Above insulating layer **40** is a gate electrode **42**, which may be formed of n-type doped polysilicon and which is separated from epitaxial layer **36** by a gate dielectric layer **44**. Channel regions **45** (represented by the dashed lines) lie within the p-body regions **39** adjacent the gate dielectric layer **44**. The thickness of insulating layer **40** is preferably set so as to minimize the overlap of the gate electrode **42** and the drain-drift region **41**. An overlying source contact layer **46** contacts n+ source regions **37**, and a drain contact layer **31** contacts n+ substrate **38**. Contact layers **46** and **31** are typically formed of a metal although they could also be formed of another conductive material.

The RFP electrodes **35** in RFP trenches **32**A and **32**B extend to a level that is deeper below the surface **36**A of epitaxial layer **36** than the bottom of gate electrode **42**. RFP electrodes **35** may be independently biased or, as shown in FIG. **3**B, may be connected to n+ source regions **37** outside the plane of the drawing.

In some embodiments, the doping concentration of the drain-drift region **41** in the areas between the RFP trenches **32**A and **32**B and the gate trench **34** is less (e.g., 5×10^{15} to 1.5×10^{16} cm^{-3}) than the doping concentration of the drain-drift region **41** in the area below the gate trench **34** (e.g., 2×10^{16} to 3×10^{16} cm^{-3}). This structure improves depletion-spreading in the areas of drain-drift region **41**A between the RFP trenches **32**A and **32**B and the gate trench **34** when the PN junction between the body region **39** and the drain-drift region **41** is reverse-biased due to the limited depletion charge in region **41**A. This can be further improved, for example, by varying the doping of epitaxial layer **36** as it is being formed. It also results in shorter channel length, which in turn provides a lower Ron, a lower gate-to-source capacitance (Cgs), and a lower gate-to-drain capacitance (Cgd)

In the embodiments shown in FIGS. **3**A and **3**B, n+ source regions **37** extend between the gate trench **34** and the RFP trenches **32**A and **32**B. Alternatively, a MOSFET **50**, shown in the top view of FIG. **4**A and the cross-sectional view of FIG. **4**B, is similar to MOSFET **30**, but in selected areas the n+ source regions **37** are replaced by p+ body contact regions **52** so as to avoid snap-back or second breakdown characteristics. In FIG. **4**A, the view at cross-section **3**A-**3**A is identical to cross-sectional view shown in FIG. **3**A, and the view at cross-section **4**B-**4**B is shown in FIG. **4**B.

When a conventional N-channel MOSFET operates in the third quadrant, its drain is negatively biased with respect to its source-body electrode, and the diffusion current results in minority carrier injection and a high Qrr. In MOSFETs **30** and **50**, because the n+ source regions **37** extend all the way across

the mesa between the gate trench **34** and the RFP trenches **32**A and **32**B, the RFP electrodes **35** provide a majority carrier channel current path from drain to source in addition to that provided by the gate electrode in the conventional structure. The combined effect of the RFP and gate electrodes results in a significant reduction in the minority carrier diffusion current and reverse recovery charge Qrr than in conventional structures. In other words, in the third quadrant operation the RFP electrode acts as an additional gate without the penalty of the added gate-drain capacitance (Cgd) and gate-source capacitance (Cgs).

The RFP electrodes **35** also function as a recessed field plates to reduce the electric field in the channel regions **45** when MOSFETs **30** and **50** are reverse-biased. This effect enables the use of shorter channel lengths, without concern about punchthrough breakdown, and this results in a lower specific on-resistance (Ron*A) and a lower gate charge (Qg). Unlike MOSFET **20**, the drift regions **41**A below p– body regions **39** are constrained between the thick gate bottom oxide and the RFP electrode and are therefore more effectively depleted. Therefore, a deeper depletion layer results for the same reverse bias body-drain junction conditions, and a shorter channel can be used, resulting in a lower on-resistance. Furthermore, because the gate-to-drain capacitance (Cgd) drops at a faster rate with applied drain-source voltage Vds, a lower gate-drain charge Qgd and a better Ron-Qgd trade-off are realized. In other words, the combined effect of the RFP electrodes **35** and of having the gate trench **34** filled with insulating layer **40** up to the junctions between the p-body regions **39** and the drain-drift region **41** helps to deplete area of the drain-drift region **36** between the trenches **32**A, **32**B and **34** at a higher rate as the drain-source voltage Vds is increased while MOSFETs **30** and **50** are in the off-state. Therefore, a lower gate-drain charge (Qgd) results because of the low gate-to-drain capacitance (Cgd) and its fast falling rate with increasing Vds. Furthermore, the doping in the p-body regions **39** can be adjusted to obtain a lower threshold voltage at the same breakdown or punchthrough voltage. The doping in the p-body regions **39** can be further adjusted so that the p-body regions **39** are fully depleted, which significantly reduces gate charge Qg.

In MOSFET **30** shown in FIGS. **3**A and **3**B, the RFP electrodes **35** are separated from the source contact layer **46** by an insulating layer **48**. Alternatively, in a MOSFET **60**, shown in FIGS. **5**A and **5**B, there is no insulating layer **48**, and RFP electrodes **62** in RFP trenches **64**A and **64**B extend upward to electrically contact the source contact layer **46**. FIG. **5**A is a view taken at a cross-section corresponding to cross-section **3**A-**3**A in FIG. **4**A; FIG. **5**B is a view taken at a cross-section corresponding to cross-section **4**B-**4**B in FIG. **4**A

An exemplary process for fabricating MOSFET **30** of FIG. **3**A is shown in FIGS. **6**A-**6**H. As shown in FIG. **6**A, the starting material is the heavily-doped n+ substrate **38**, which may be doped, for example, with phosphorus or arsenic. The n-type epitaxial layer **36** is grown on top of the n+ substrate **38**. A thin oxide layer **72** is grown over the n-type epitaxial layer **36**, and a silicon nitride layer **74** is deposited on top of the oxide layer **72**. For example, the oxide layer **72** can be 200-300 A thick, and the silicon nitride layer **74** can be 1000 A thick.

A photoresist mask (not shown) is used to pattern the silicon nitride layer **74** and oxide layer **72** with openings **76**, as shown in FIG. **6**B.

The epitaxial layer **36** is etched through the openings **76** to form the RFP trenches **32**A and **32**B and the gate trench **34**, as shown in FIG. **6**C. It will be understood that trenches **32**A, **32**B and **34** are illustrative only; typically numerous trenches would be formed. A thin thermal oxide layer (not shown), for example, 300 A thick, is grown on the walls of trenches **32**A, **32**B and **34**, and trenches **32**A, **32**B and **34** are then filled with a layer **78** of deposited silicon dioxide such as LTO or TEOS or high density plasma (HDP) oxide. The resulting structure is shown in FIG. **6**C.

The oxide layer **78** is etched back using a dry plasma etch or chemical-mechanical polishing (CMP) technique to a level the same as or slightly below the top surface of the silicon nitride layer **74**. A photoresist mask layer **80** is deposited and patterned to form openings **82** over the RFP trenches **32**A and **32**B, as shown in FIG. **6**D, and the oxide layer **78** in the RFP trenches **32**A and **32**B is then completely or partially etched through the openings **82** by a wet or a dry plasma etch or a combination of wet and dry etching to yield the structure shown in FIG. **6**D. Silicon nitride layer **74** prevents epitaxial layer **36** from being etched.

The photoresist mask layer **80** is removed and the portion of oxide layer **78** remaining in the gate trench **34** is partially etched by wet etching. At the same time, the remaining portion of oxide layer **78** at the bottom of the RFP trenches **32**A and **32**B is removed. This is followed by an etch to remove the silicon nitride layer **74** and a wet oxide etch to remove the oxide layer **72** and any remaining oxide on the walls of the RFP trenches **32**A and **32**B. The resulting structure is shown in FIG. **6**E.

As shown in FIG. **6**F, a thermal process is then performed to form oxide layer **33** on the walls of the RFP trenches **32**A and **32**B and gate dielectric layer **44** on the walls of the gate trench **34** above the remaining portion of the oxide layer **78** (which forms the insulating layer **40** at the bottom of the gate trench **34**). This process also creates an oxide layer **80** on the top surface of the epitaxial layer **36**. These oxide layers could have a thickness, for example, in the range of 200 A to 1000 A. A layer **82** of polysilicon is then deposited and then doped, for example, by implanting with an n-type dopant such as phosphorus and then etched back to a level at or below that of the surface of oxide layer **80** by a plasma dry etching or CMP technique. Several masking and implant steps are performed to form the p-body regions **39** and p+ body contact regions **52** (shown in FIGS. **4**A and **4**B), using a p-type dopant such as boron, and to form the n+ source regions **37**, using an n-type dopant such as arsenic or phosphorus or a combination thereof. A deep p layer (not shown) may also be implanted at this step or earlier in the process.

The polysilicon layer **82** is etched back and an oxide layer is deposited and later etched, leaving the RFP electrodes **35** and the gate electrode **42** covered by an oxide layer **84**, as shown in FIG. **6**G

The source contact layer **46** is deposited and defined by a photoresist mask (not shown). The back side of the n+ substrate **38** is thinned and the drain metal layer **31** deposited, resulting in the MOSFET **30** shown in FIGS. **3**A and **6**H.

In an alternative method, after the step shown in FIG. **6**D the portions of oxide layer **78** remaining in RFP trenches **32**A and **32**B and in gate trench **34** are etched, and the steps shown in FIGS. **7**A and **7**B are performed. As shown in FIG. **7**A, with oxide layer **72** and nitride layer **74** still in place oxide layers **33** and **44** are thermally grown on the walls of the RFP and gate trenches **32**A/**32**B and **34**, respectively. A polysilicon layer **85** is deposited and partially removed by etching or CMP so that the surface of the polysilicon layer **85** is level with or below the surface of the nitride layer **74**. The resulting structure is shown in FIG. **7**A.

The polysilicon layer **85** is etched (e.g., by about 0.1 μm) and an oxide layer **87** is deposited and etched back so that the

surface of oxide layer **87** is level with or below the nitride layer **74**. Nitride layer **74** is removed by etching and optionally oxide layer **72** may be removed and a sacrificial oxide layer (not shown) may be re-grown. P-body regions **39** and n+ source regions **37** are implanted. Oxide layer **87** is then removed from the areas over the n+ source regions **37**, leaving the structure shown in FIG. **7**B. A source contact layer (not shown) is then deposited on top of the oxide layer **87** and n+ source regions **37**.

Numerous variations of the MOSFETs described above are within the scope of this invention. In a MOSFET **70** shown in FIG. **8**A, an insulating layer **92** over the gate electrode **42** extends above the surface **36**A of epitaxial layer **36** and partially covers the n+ source regions **37**, while the RFP electrodes **35** remain recessed below the surface **36**A. A source contact layer **94** contacts the RFP electrodes **35** so that the RFP electrodes **35** are biased at the source potential. FIG. **8**B is another view of MOSFET **70**, taken at a cross-section similar to cross-section **4**B-**4**B shown in FIG. **4**A, where the n+ source regions are replaced by p+ body contact regions **52**.

A MOSFET **80**, shown in FIG. **9**, is similar to MOSFET **70** except that p+ body contact regions **96** are formed adjacent to n+ source regions **98** at the surface **36**A of epitaxial layer **36** to provide a source-body short via source contact layer **94**. In MOSFET **80**, it would not be necessary to interrupt the n+ source regions with p+ body contact regions **52** as shown in FIG. **4**A to provide a source-body short. MOSFET **90**, shown in FIG. **10**, is similar to MOSFET **80** except that p+ body contact regions **99** extend to a level below the bottoms of n+ source regions to provide a larger surface along the walls of RFP trenches **32**A and **32**B for contact with source contact layer **94**. A process for making MOSFET **90** is shown in FIGS. **14**A-**14**H.

In MOSFET **100**, shown in FIG. **11**, p+ body contact regions **104** are formed below the n+ source regions **37** and adjacent to the p-body regions **106**. The bottom junctions of p-body regions **106** and p+ body contact regions **104** are at approximately the same level in epitaxial layer **36**. In RFP trenches **102**A and **102**B, the RFP electrodes **112** and the oxide layers **110** are recessed sufficiently to permit the source contact layer **108** to contact the p+ body contact regions **104**, thereby providing a source-body short. In the gate trench **34**, insulating layer **92** extends above the surface **36**A of epitaxial layer **36** and covers the n+ source regions **37**. MOSFET **100** may allow for a higher cell density and therefore a lower specific on-resistance. MOSFET **110**, shown in FIG. **12**, is similar to MOSFET **100** except that p+ body contact regions **114** extend to a level below the bottom of p-body regions **106**. MOSFET **120**, shown in FIG. **13**, is similar to MOSFET **110** except that RFP trenches **118**A and **118** B contain metal plugs **116**, comprising for example tungsten, which contact both the source regions **37** and the p+ body contact regions **114** to provide a very low-resistance conduction path between source regions **37** and p+ body contact regions **114**.

FIGS. **14**A-**14**H illustrate a process of fabricating MOSFET **90**, shown in FIG. **10**. The process starts with growing n-type epitaxial layer **36** on top of n+ substrate **38**. As shown in FIG. **14**A, an oxide layer **120** (e.g., 0.5 μm thick) and a photoresist trench mask layer **122** are deposited over the top surface of epitaxial layer **36**. Mask layer **122** is patterned to form openings and oxide layer **120** and epitaxial layer **36** are etched through the openings to form RFP trenches **32**A and **32**B and gate trench **34**. The mask layer **122** and oxide layer **120** are removed, and a sacrificial oxide layer and a pad oxide layer (not shown) are grown.

As shown in FIG. **14**B, an oxide layer **124** is deposited in trenches **32**A, **32**B and **34**. Oxide layer **124** is preferably a high quality oxide, such as a high-density plasma oxide. An active mask (not shown) is deposited and patterned with an opening over the active areas of the die, and oxide layer **124** is etched down into trenches **32**A, **32**B and **34**, as shown in FIG. **14**C. The active mask prevents optional field termination oxide (not shown) or oxide layer **124** from being etched in the termination areas of the die.

A photoresist bottom oxide (BOX) mask layer **126** is deposited and patterned with openings over the RFP trenches **32**A and **32**B, and the remains of oxide layer **124** is removed from RFP trenches **32**A and **32**B, leaving the remains of oxide layer **124** (which becomes insulating layer **40**) in the bottom of gate trench **34**. The resulting structure is shown in FIG. **14**D.

BOX mask layer **126** is removed and, as shown in FIG. **14**E, an oxide layer **128** is grown, resulting in the formation of oxide layer **33** on the walls of RFP trenches **32**A and **32**B and oxide layer **44** on the walls of gate trench **34** above insulating layer **40**. A polysilicon layer **130** is deposited (e.g., 7000 A thick) and a mask layer (not shown) is deposited and patterned, and polysilicon layer **130** is etched back into trenches **32**A, **32**B and **34** using CMP and/or a dry etch process, thereby forming the gate electrode **42** in gate trench **34** and the RFP electrodes **35** in RFP trenches **32**A and **32**B. A p-type dopant is implanted to form p-body regions **39** (e.g., boron at a dose of 5×10^{12} cm^{-2} and an energy of 100 keV). A rapid thermal anneal (RTA) process may be performed at a temperature of 1025° C. for 30 seconds, for example, resulting in a junction depth of 0.5 μm for p-body regions **39**.

As shown in FIG. **14**F, a photoresist source mask layer **132** is deposited and patterned to form an opening, and an n-type dopant is implanted to form n+ source regions **98** at the surface of epitaxial layer **36** adjacent to gate trench **34**. For example, arsenic may be implanted at a dose of 2×10^{15} cm^{-2} and an energy of 80 keV to provide a junction depth of 0.2 μm for n+ source regions **98** and a channel length (L) of 0.25-0.3 μm.

Source mask layer **132** is removed and, as shown in FIG. **14**G, an inter-level dielectric (ILD) layer **136**—e.g., low-temperature oxide (LTO) and borophosphosilicate glass (BPSG)—is deposited to a thickness of 0.5-1.5 μm, for example. ILD layer **136** may then be densified.

As shown in FIG. **14**H, a contact mask layer (not shown) is deposited on ILD layer **136** and patterned with openings over RFP trenches **32**A and **32**B, and ILD layer **136** and portions of RFP electrodes **35** are dry or wet etched through the openings in the contact mask layer. A reflow process may be performed on ILD layer **136**. A p-type dopant is implanted to form p+ body contact regions **99**. For example, boron may be implanted at a dose of 1×10^{15} cm^{-2} to 4×10^{15} cm^{-2} and an energy of 20-60 keV to produce a junction depth of 0.4 μm for p+ body contact regions **99**. Source contact layer **94** is then deposited in contact with the RFP electrodes **35**, resulting in MOSFET **90**. Optionally, tungsten plugs may be formed in the upper portions of RFP trenches **32**A and **32**B in contact with RFP electrodes **35** before source contact layer **94** is deposited.

In a variation of the process, as shown in FIG. **15**A, the polysilicon layer **130** is etched further down into RFP trenches **32**A and **32**B through the openings in the contact mask layer, forming RFP electrodes **112**. A p-type dopant such as boron is then implanted at an angle, forming p+ body contact regions **114** that are located below n+ source regions **37** and may extend further into the epitaxial layer **36** than the bottoms of p-body regions **106**. As shown in FIG. **15**B, metal plugs **116**, comprising tungsten, for example, may be formed in the upper portions of RFP trenches **32**A and **32**B before

source contact layer **94** is deposited. Except for the presence of the metal plugs **116**, MOSFET **130** is similar to MOSFET **110**, shown in FIG. **12**.

Numerous other embodiments are possible within the broad scope of this invention. In some embodiments, a deeper p region is formed in selected areas of the MOSFET to clamp its breakdown voltage to a predetermined value that is lower than the breakdown voltage of the RFP trench or other points in the device. In MOSFET **150**, shown in FIG. **16**A, deep p+ regions **152** are located outside of the areas between the RFP trenches **32**A and **32**B and the gate trench **34**. In MOSFET **160**, shown in FIG. **16**B, deep p+ regions **162** are located on both sides of each of the RFP trenches **32**A and **32**B. P+ regions **152** and **162** extend deeper into the epitaxial layer **36** than the p-body regions **39** and may be as deep as the RFP trenches **32**A and **32**B and the gate trench **34**. It should be noted that FIGS. **16**A and **16**B are taken at cross-sections where there is no n+ source region.

In another embodiment, the insulating layers lining the RFP trenches are thicker at the bottoms of the trenches than on the sides of the trenches. In MOSFET **170**, shown in FIG. **17**, the insulating layer **172** lining the walls of RFP trenches **32**A and **32**B includes a section **172**B at the bottom of the trench that is thicker than a section **172**A along the sidewalls of the trench.

In other embodiments, the doping of the N-epitaxial drain-drift region **41** of the epitaxial layer **36** is non-uniform (see FIG. **3**A). For example, the doping of drain-drift region **41** may be non-uniform, with the doping concentration increasing with increasing depth in the epitaxial layer **36** so that the doping concentration in the areas **41**A of drain-drift region **41** is less than the doping concentration in the portion of drain-drift region **41** below areas **41**A.

Other variations of the new structure include a stepped oxide lining the gate trench and/or the RFP trench. In MOSFET **180**, shown in FIG. **18**A, the gate oxide layer **182** in gate trench **34** includes a thinner section **182**B, having a thickness d_1, on the sidewalls of the trench **34** adjacent to the p-body regions **39** and a thicker section **182**A, having a thickness d_2, along the lower sidewalls and bottom of the trench **34**. The thickness d_2 is less than one-half the width W of the trench **34**, so that the gate oxide layer **182** forms a "keyhole" shape. In MOSFET **180** the thinner section **182**B and the thicker section **182**A are joined at a location adjacent to the junction between the p-body regions **39** and the drain-drift region **41**.

Similarly, in MOSFET **190**, shown in FIG. **18**B, the oxide layer **192** in each of the RFP trenches **32**A and **32**B includes a thinner section **192**B, having a thickness d_3, on the upper sidewalls of the trenches **32**A and **32**B and a thicker section **192**A, having a thickness d_4, along the lower sidewalls and bottom of the trenches **32**A and **32**B. The thickness d_4 is less than one-half the width W of the trenches **32**A and **32**B, so that the oxide layer **192** forms a "keyhole" shape.

In MOSFET **200**, shown in FIG. **18**C, the gate trench **34** contains the gate oxide layer **182** (as described above) and the RFP trenches **32**A and **32**B contain the oxide layer **192** (as described above).

In MOSFET **210**, shown in FIG. **19**, the gate trench **214** is deeper than the RFP trenches **212**A and **212**B to reduce the electric field at the RFP trenches **212**A and **212**B, while the gate electrode **216** is shallower than the RFP electrodes **215**. For example, gate trench **214** extends to a deeper level in epitaxial layer **36** than the bottoms of RFP trenches **212**A and **212**B but owing to the thickness of insulating layer **40** at the bottom of the gate trench **216** RFP electrodes **215** extend to a deeper level than gate electrode **216**.

The principles of this invention are applicable to quasi-vertical as well as vertical MOSFETs. FIG. **20** is a cross-sectional view of a quasi-vertical MOSFET **220**. MOSFET **220** includes a gate trench **224**, RFP trenches **222**A and **222**B, n+ source regions **226** and p-body regions **228**. An n-buried layer **230** is formed at the interface between a p-type substrate **236** and an n-epitaxial layer **234**. N-buried layer **230** is contacted from the top surface of n-epitaxial layer **234** via n+ sinker regions **232**. RFP electrodes **235** in RFP trenches **222**A and **222**B are contacted by a source contact layer **238**. When MOSFET **220** is turned on, a current flows from n+ source regions **226**, through p-body regions **228** to n-buried layer **230** and back up to the surface of n-epitaxial layer **234** via n+ sinker regions **232**.

The embodiments described above are illustrative only and not limiting. Many additional and alternative embodiments in accordance with the broad principles of this invention will be obvious to persons of skill in the art from the above descriptions. For example, devices in accordance with this invention may be fabricated in various layouts, including "stripe" and "cellular" layouts. While the embodiments described above have generally been n-channel MOSFETs, the principles of this invention are equally applicable to p-channel MOSFETs. While the embodiments described above include an epitaxial layer grown on a substrate, in some embodiments the epitaxial layer may be omitted. It should also be noted that various combinations of the above embodiments can be realized and are included within the scope of this disclosure.

I claim:

1. A MOSFET formed in a semiconductor die comprising:

a gate trench extending from a surface of the die, the gate trench comprising a gate electrode, the gate electrode being insulated from the die by a first dielectric layer, the first dielectric layer comprising a first section at a bottom of the gate trench and a second section at a sidewall of the gate trench, the first section being thicker than the second section;

a first recessed field plate (RFP) trench extending from the surface of the die, the first RFP trench containing a first RFP electrode having an upper surface which is recessed below an upper surface of said gate electrode and which is insulated from the die by a second dielectric layer;

a second RFP trench extending from the surface of the die, the second RFP trench containing a second RFP electrode which is also insulated from the die, the gate trench being located between the first and second RFP trenches;

a mesa of the die between the gate trench and the first RFP trench;

a source region of a first conductivity type in the mesa adjacent to the surface of the die and a sidewall of the gate trench;

a body region, having a second conductivity type opposite to the first conductivity type, adjacent to the sidewall of the gate trench and to the source region;

a body contact region, abutting said first and second RFP trenches and said body region,

which is at least partly self-aligned to at least one of said RFP electrodes,

which is doped with said second conductivity type,

which has a doping concentration greater than that of said body region, and

which defines a bottom junction depth which is deeper than a bottom junction of said body region; and

a drain-drift region of the first conductivity type adjacent to the body region;

wherein the respective bottoms of the first and second RFP electrodes are located at a level deeper below the surface of the die than a bottom of the gate electrode.

2. The MOSFET of claim **1**, wherein the gate trench is equidistant from the first and second RFP trenches.

3. The MOSFET of claim **1**, wherein a doping concentration of the drain-drift region in an area between the first RFP trench and the gate trench is less than a doping concentration of the drain-drift region in an area below the gate trench.

4. The MOSFET of claim **1**, wherein the first and second RFP electrodes comprise polysilicon doped with a dopant of the first conductivity type.

5. The MOSFET of claim **1**, wherein the first and second RFP electrodes comprise polysilicon doped with a dopant of the second conductivity type.

6. The MOSFET of claim **1**, wherein said body contact region has a junction contour which corresponds to implantation and outdiffusion from ions introduced near the tops of said RFP electrodes.

7. The MOSFET of claim **1**, further comprising metal plugs which overlie said RFP electrodes, and also a metal layer overlying said metal plugs.

8. The MOSFET of claim **1**, wherein said first conductivity type is n type.

9. The MOSFET of claim **3**, wherein said body contact region has a junction contour which corresponds to implantation and outdiffusion from ions introduced near the tops of said RFP electrodes.

10. The MOSFET of claim **3**, further comprising metal plugs which overlie said RFP electrodes, and also a metal layer overlying said metal plugs.

11. The MOSFET of claim **3**, wherein said first conductivity type is n type.

12. The MOSFET of claim **4**, wherein said body contact region has a junction contour which corresponds to implantation and outdiffusion from ions introduced near the tops of said RFP electrodes.

13. The MOSFET of claim **4**, further comprising metal plugs which overlie said RFP electrodes, and also a metal layer overlying said metal plugs.

14. The MOSFET of claim **4**, wherein said first conductivity type is n type.

15. A MOSFET formed in a semiconductor die comprising:

a gate trench extending from a surface of the die, the gate trench comprising a gate electrode, the gate electrode being insulated from the die by a first dielectric layer, the first dielectric layer comprising a first section at a bottom of the gate trench and a second section at a sidewall of the gate trench, the first section being thicker than the second section;

a recessed field plate (RFP) trench extending from the surface of the die, the RFP trench containing an RFP electrode having an upper surface which is recessed below an upper surface of said gate electrode, the RFP electrode being insulated from the die by a second dielectric layer, a bottom of the RFP electrode being located at a level deeper below the surface of the die than a bottom of the gate electrode, the RFP electrode being electrically isolated from the gate electrode;

a mesa of the die between the gate trench and the RFP trench;

a source region of a first conductivity type in the mesa adjacent to the surface of the die, the source region extending across the mesa between a sidewall of the RFP trench and a sidewall of the gate trench;

a body region of a second conductivity type opposite to the first conductivity type in the mesa, the body region being adjacent to the source region;

a body contact region, abutting said RFP trench and said body region,

which is at least partly self-aligned to said RFP electrodes,

which is doped with said second conductivity type,

which has a doping concentration greater than that of said body region, and

which defines a bottom junction depth which is deeper than a bottom junction of said body region; and

a drain-drift region of the first conductivity type adjacent to the body region.

16. The MOSFET of claim **15**, wherein a depth of the RFP trench is substantially equal to a depth of the gate trench.

17. The MOSFET of claim **15**, comprising:

a second RFP trench extending from the surface of the die, the second RFP trench containing a second RFP electrode, the second RFP electrode being insulated from the die by a third dielectric layer, a bottom of the second RFP electrode being located at a level deeper below the surface of the die than a bottom of the gate electrode;

a second mesa of the die between the gate trench and the second RFP trench;

a second source region of the first conductivity type in the second mesa adjacent to the surface of the die, the second source region extending across the second mesa between a sidewall of the second RFP trench and a second sidewall of the gate trench;

a second body region of the second conductivity type in the mesa, the second body region being adjacent to the second source region and extending across the second mesa between the sidewall of the RFP trench and the second sidewall of the gate trench.

18. The MOSFET of claim **17**, wherein a depth of each of the RFP trench and the second RFP trench is substantially equal to a depth of the gate trench.

19. The MOSFET of claim **15**, wherein said body contact region has a junction contour which corresponds to implantation and outdiffusion from ions introduced near the tops of said RFP electrodes.

20. The MOSFET of claim **15**, further comprising metal plugs which overlie said RFP electrodes, and also a metal layer overlying said metal plugs.

21. The MOSFET of claim **15**, wherein said first conductivity type is n type.

22. The MOSFET of claim **16**, wherein said body contact region has a junction contour which corresponds to implantation and outdiffusion from ions introduced near the tops of said RFP electrodes.

23. The MOSFET of claim **16**, further comprising metal plugs which overlie said RFP electrodes, and also a metal layer overlying said metal plugs.

24. The MOSFET of claim **16**, wherein said first conductivity type is n type.

25. The MOSFET of claim **17**, wherein said body contact region has a junction contour which corresponds to implantation and outdiffusion from ions introduced near the tops of said RFP electrodes.

26. The MOSFET of claim **17**, further comprising metal plugs which overlie said RFP electrodes, and also a metal layer overlying said metal plugs.

27. The MOSFET of claim **17**, wherein said first conductivity type is n type.

* * * * *

US008076719B2

(12) **United States Patent**
Zeng et al.

(10) **Patent No.: US 8,076,719 B2**
(45) **Date of Patent: Dec. 13, 2011**

(54) **SEMICONDUCTOR DEVICE STRUCTURES AND RELATED PROCESSES**

(75) Inventors: **Jun Zeng**, Torrance, CA (US); **Mohamed N. Darwish**, Campbell, CA (US)

(73) Assignee: **MaxPower Semiconductor, Inc.**, Santa Clara, CA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **12/368,399**

(22) Filed: **Feb. 10, 2009**

(65) **Prior Publication Data**

US 2009/0206924 A1 Aug. 20, 2009

Related U.S. Application Data

(60) Provisional application No. 61/065,759, filed on Feb. 14, 2008.

(51) **Int. Cl.**
H01L 29/78 (2006.01)

(52) **U.S. Cl.** .. **257/330**; 257/334; 257/341; 257/E29.257

(58) **Field of Classification Search** 257/330, 257/334, 341, E29.257
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,168,331	A	12/1992	Yilmaz	
5,282,018	A	1/1994	Hiraki et al.	
5,525,821	A	6/1996	Harada	
5,637,898	A	6/1997	Baliga	
5,864,159	A	1/1999	Takahashi	
5,973,359	A	10/1999	Kobayashi	
5,998,833	A	12/1999	Baliga	
6,069,372	A	5/2000	Uenishi	
6,114,727	A	9/2000	Ogura et al.	
6,191,447	B1	2/2001	Baliga	
6,251,730	B1	6/2001	Luo	
6,388,286	B1	5/2002	Baliga	
6,468,878	B1 *	10/2002	Petruzzello et al.	438/454
6,525,373	B1	2/2003	Kim	
6,534,828	B1	3/2003	Kocon	
6,541,820	B1	4/2003	Bol	
6,649,975	B2	11/2003	Baliga	
6,686,244	B2	2/2004	Blanchard	
6,710,403	B2	3/2004	Sapp	
6,803,627	B2	10/2004	Pfirsch	
2001/0001494	A1	5/2001	Kocon	

(Continued)

FOREIGN PATENT DOCUMENTS

WO	97/33320	A1	9/1997

(Continued)

OTHER PUBLICATIONS

J. T. Watt, B. J. Fishbein & J. D. Plummer; Low-Temperature NMOS Technology with Cesium-Implanted Load Devices; IEEE Trans. Electron Devices, vol. 34, # 1, Jan. 1987; p. 28-38.

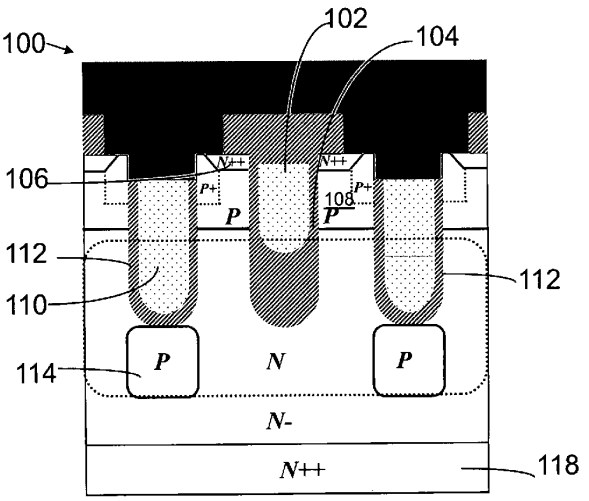
(Continued)

Primary Examiner — Long Pham
(74) *Attorney, Agent, or Firm* — Robert Groover

(57) **ABSTRACT**

Improved highly reliable power RFP structures and fabrication and operation processes. The structure includes plurality of localized dopant concentrated zones beneath the trenches of RFPs, either floating or extending and merging with the body layer of the MOSFET or connecting with the source layer through a region of vertical doped region. This local dopant zone decreases the minority carrier injection efficiency of the body diode of the device and alters the electric field distribution during the body diode reverse recovery.

6 Claims, 27 Drawing Sheets



U.S. PATENT DOCUMENTS

2001/0041407	A1	11/2001	Brown	
2003/0203576	A1	10/2003	Kitada et al.	
2006/0060916	A1	3/2006	Girdhar et al.	
2007/0004116	A1	1/2007	Hshieh	
2007/0013000	A1	1/2007	Shiraishi	
2008/0099837	A1*	5/2008	Akiyama et al.	257/341
2010/0084706	A1*	4/2010	Kocon	257/330

FOREIGN PATENT DOCUMENTS

WO	2006027739	3/2006

OTHER PUBLICATIONS

J.T.Watt,B.J.Fishbein & J.D.Plummer;Characterization of Surface Mobility in MOS Structures Containing Interfacial Cesium Ions;IEEE Trans.Electron Devices,V36,Jan. 1989; p. 96-100.

J.R.Pfiester, J.R.Alvis & C.D.Gunderson; Gain-Enhanced LDD NMOS Device Using Cesium Implantation; IEEE Trans.Electron Devices, V39, #6, Jun. 1992; p. 1469-1476.

* cited by examiner

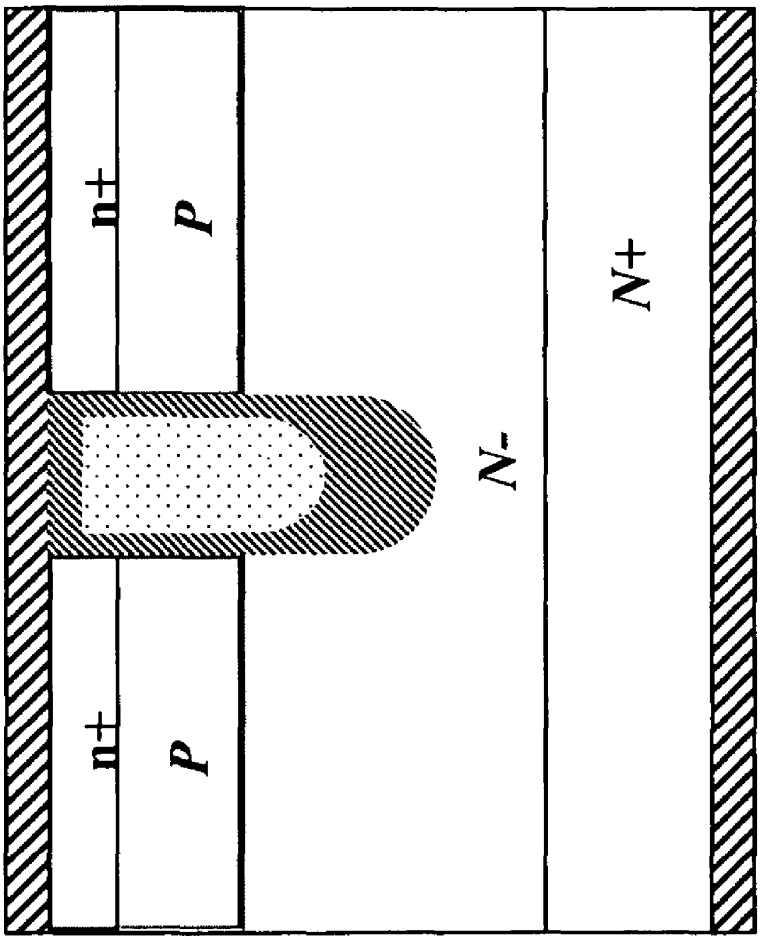


FIG. 1
(Prior Art)

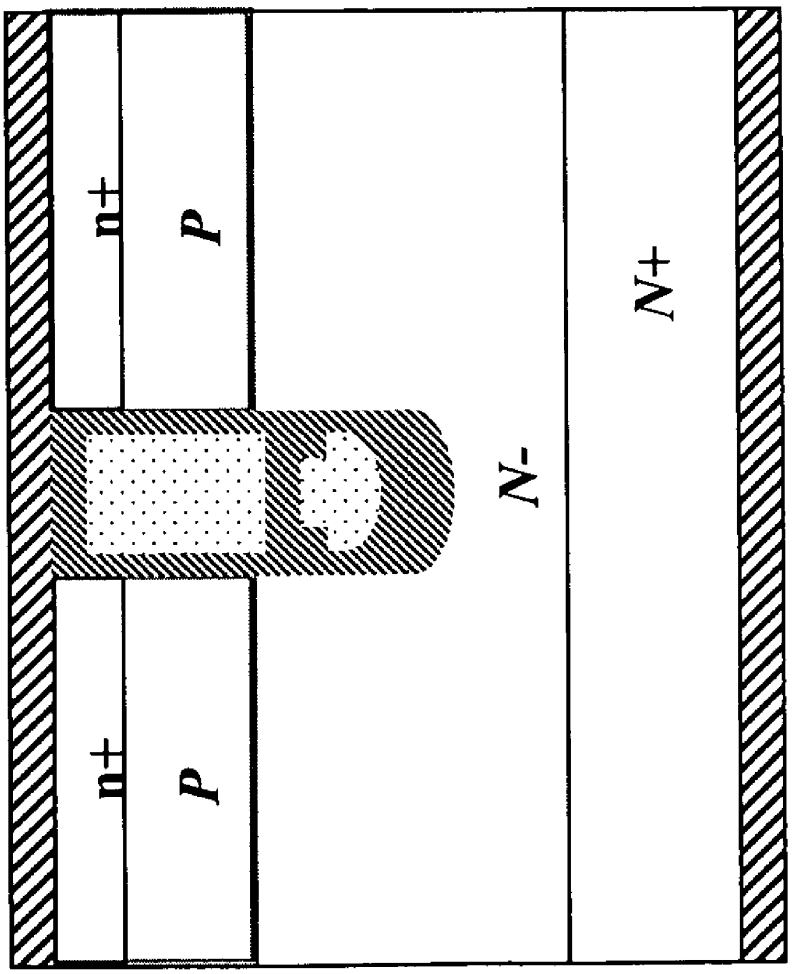


FIG. 2
(Prior Art)

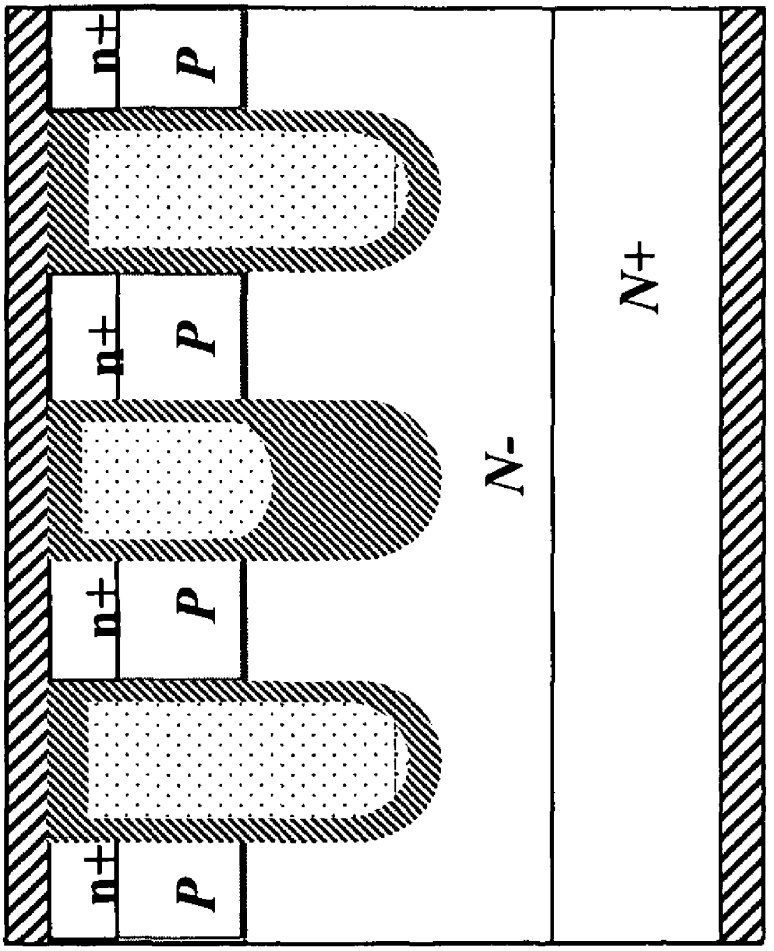


FIG. 3
(Prior Art)

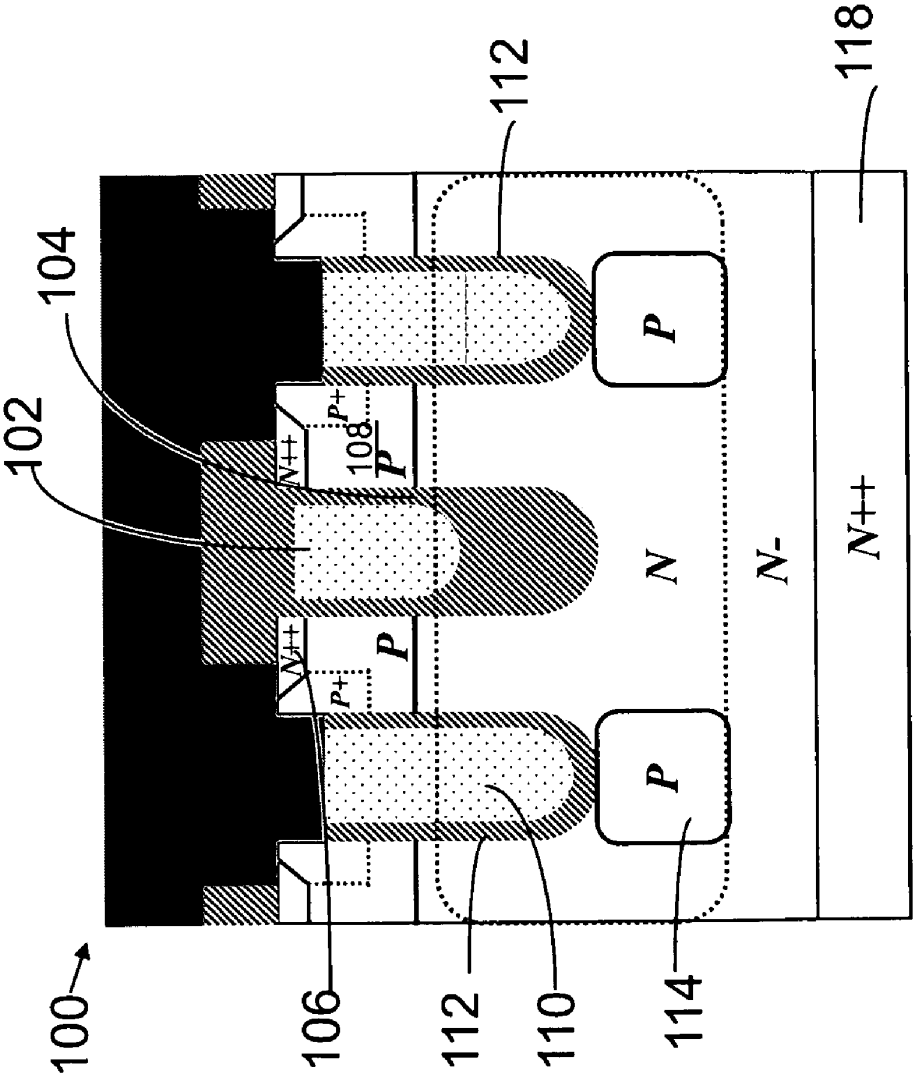


FIG. 4a

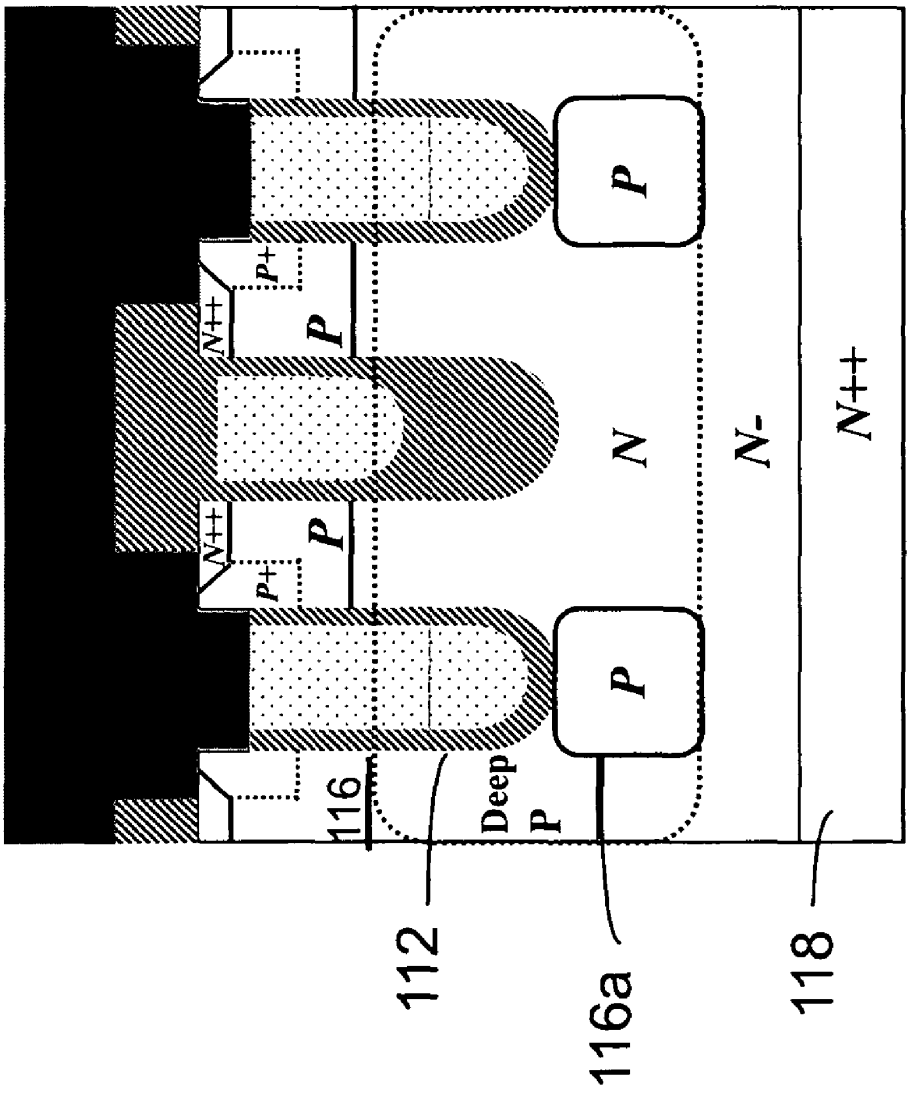


FIG. 4b

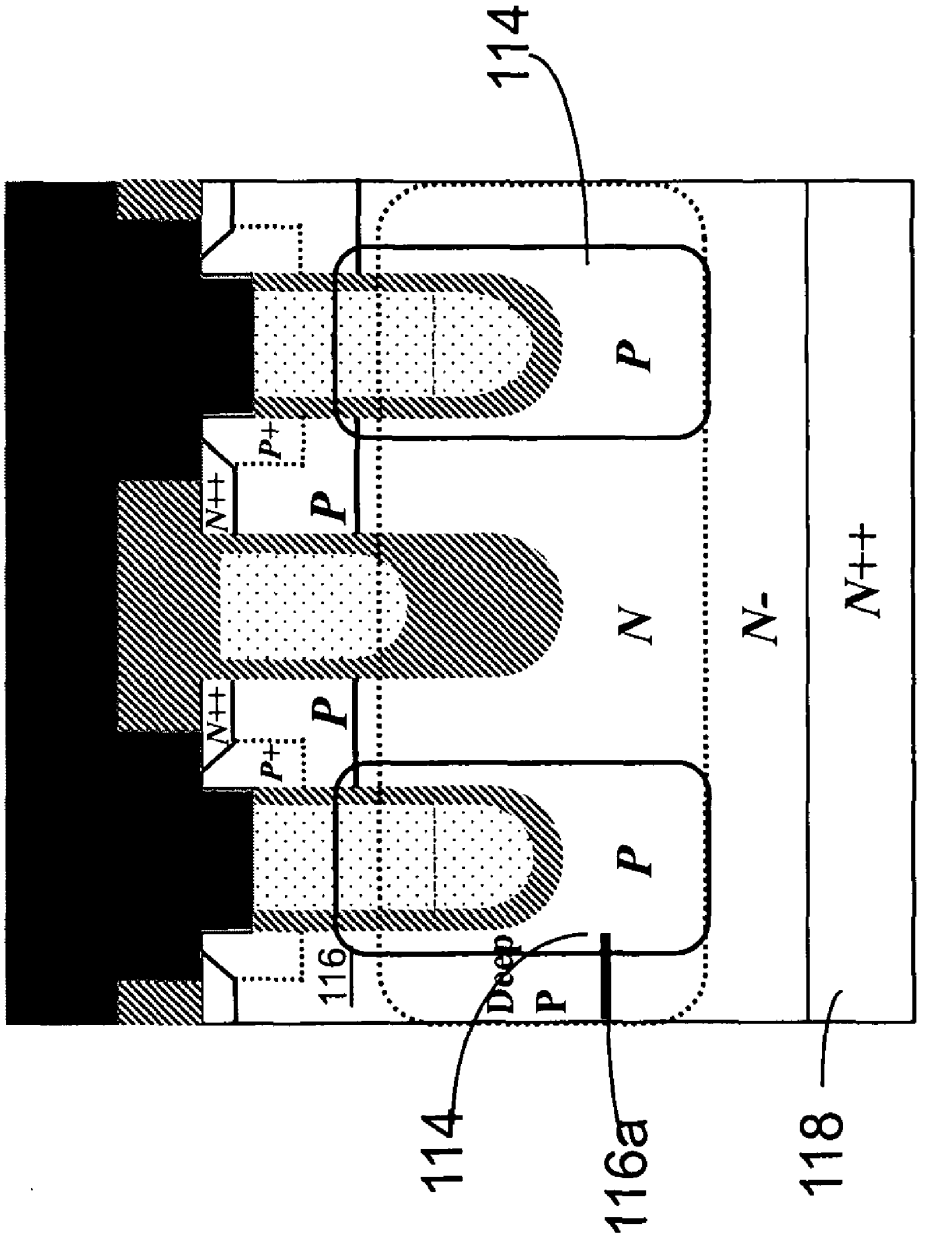


FIG. 4c

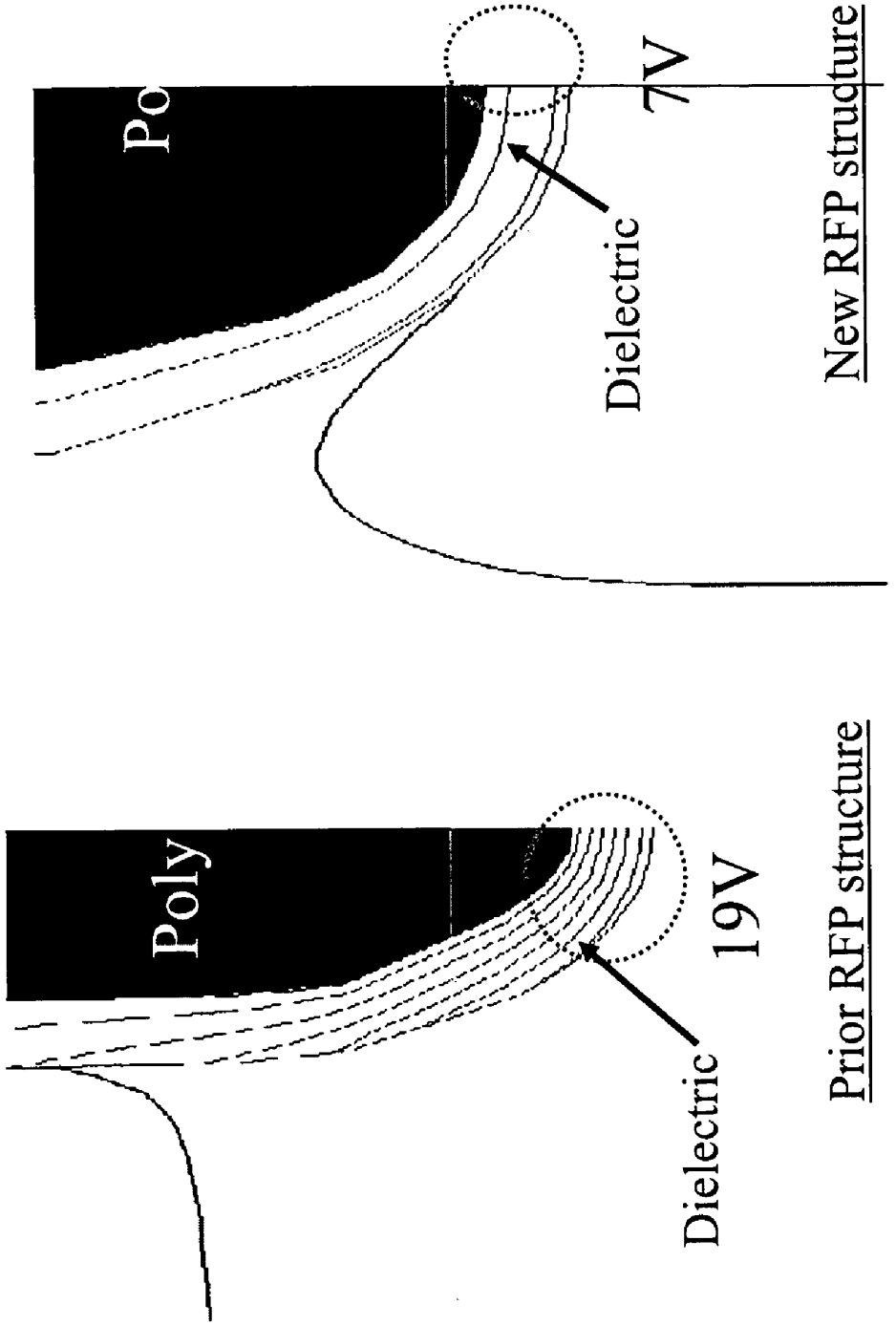


FIG. 5

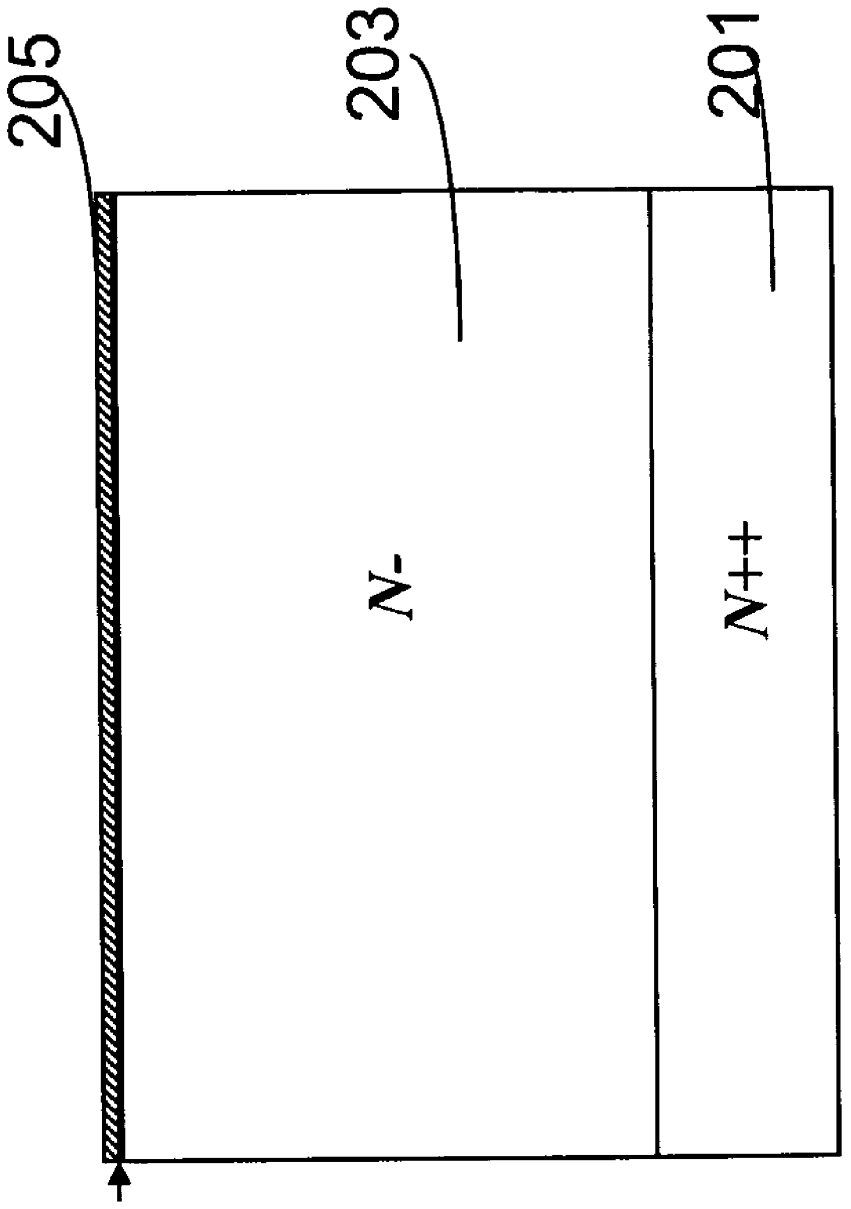


FIG. 6

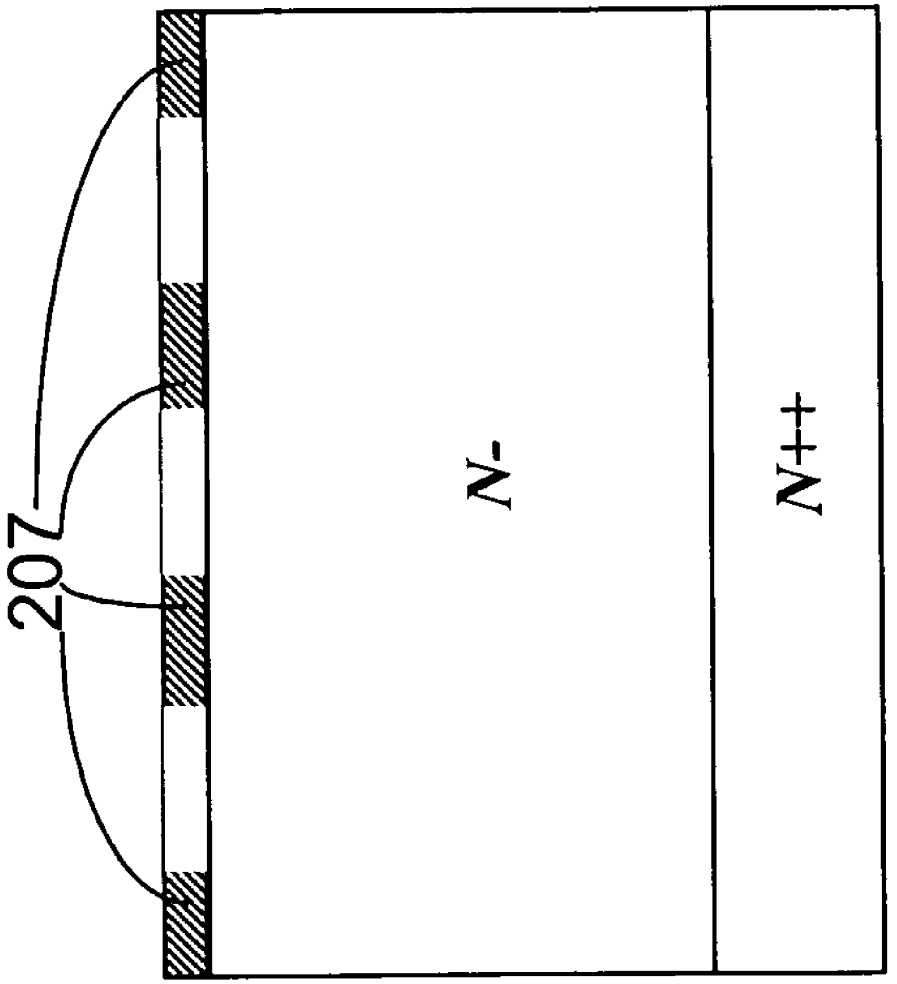


FIG. 7

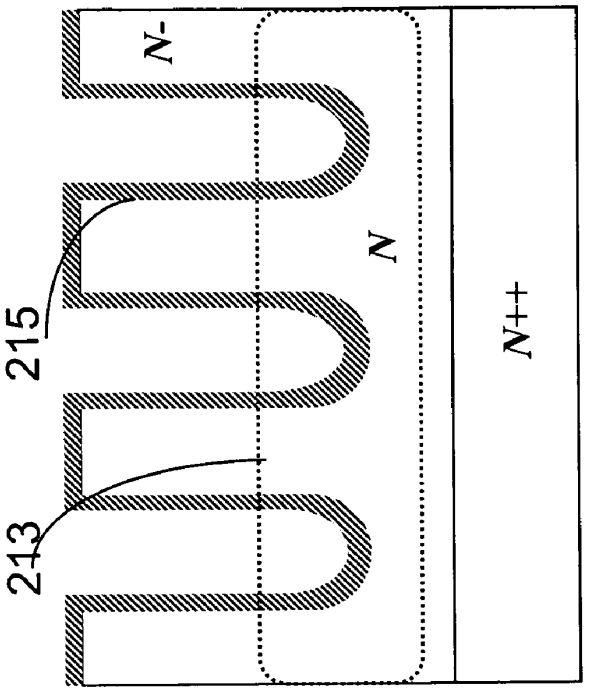


FIG. 9

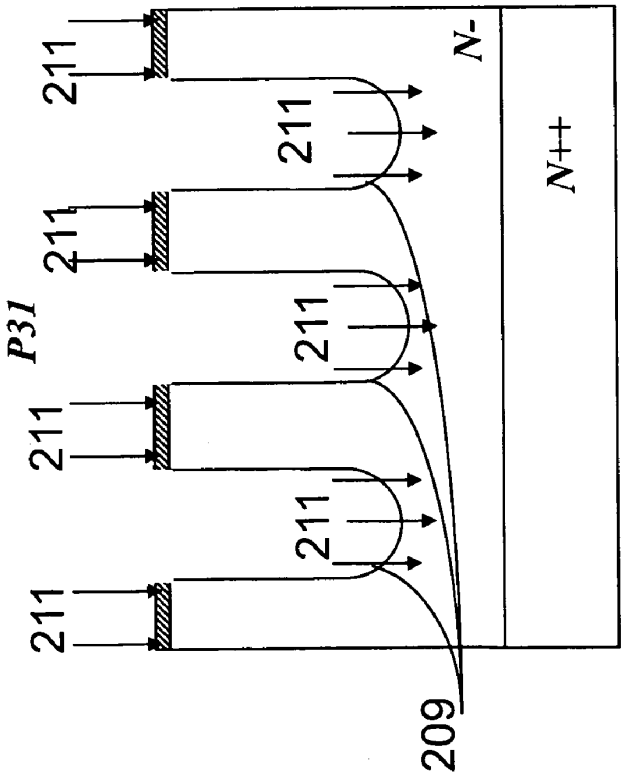


FIG. 8

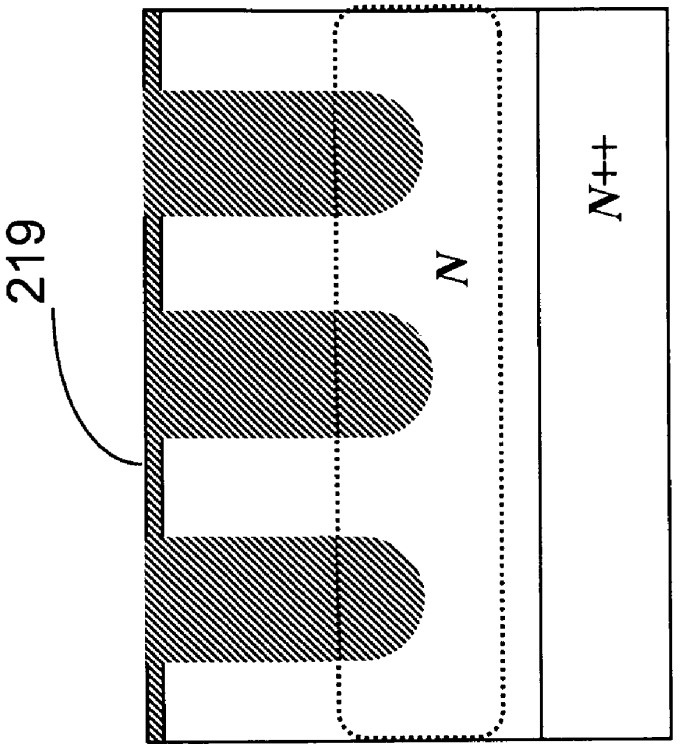


FIG. 11

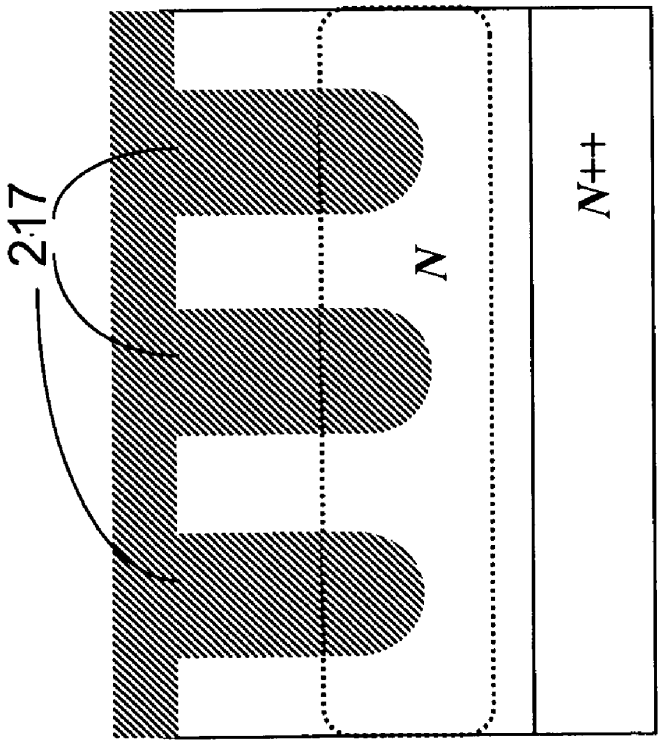


FIG. 10

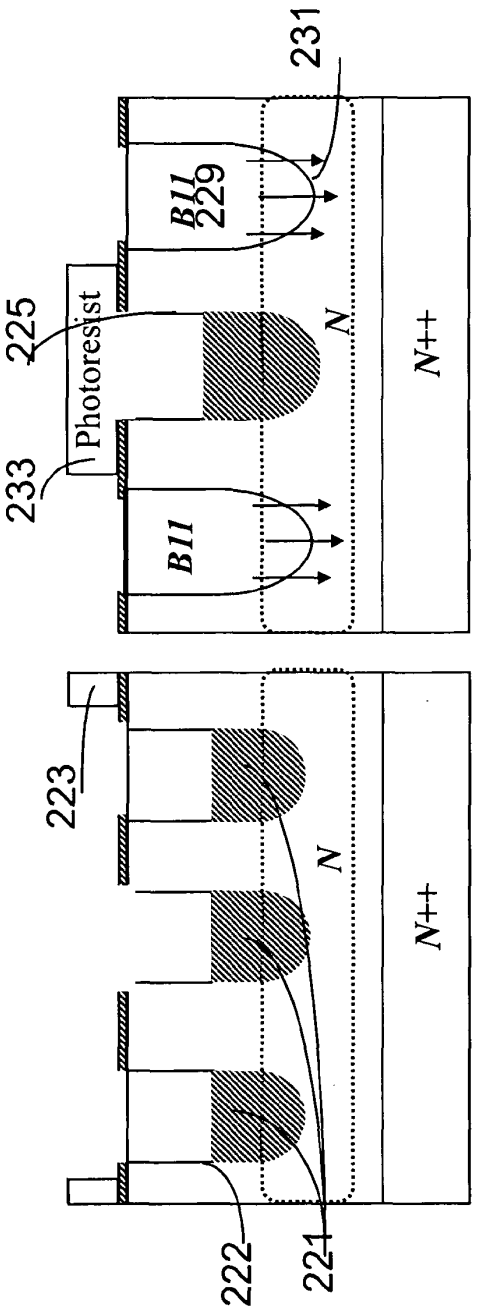


FIG.13

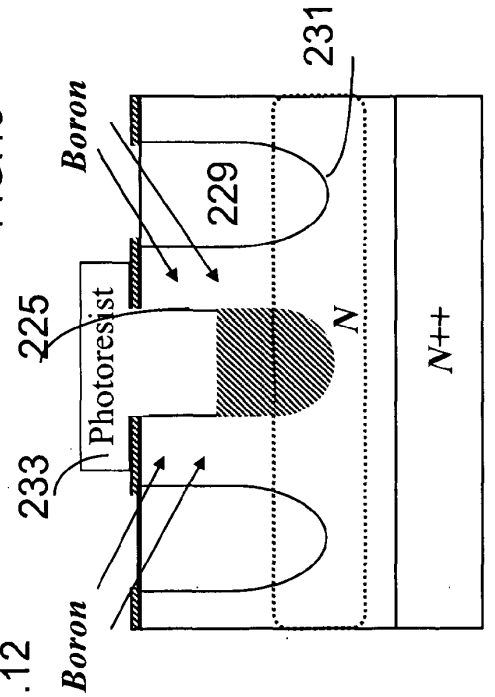


FIG.12

FIG.13+ (optional)

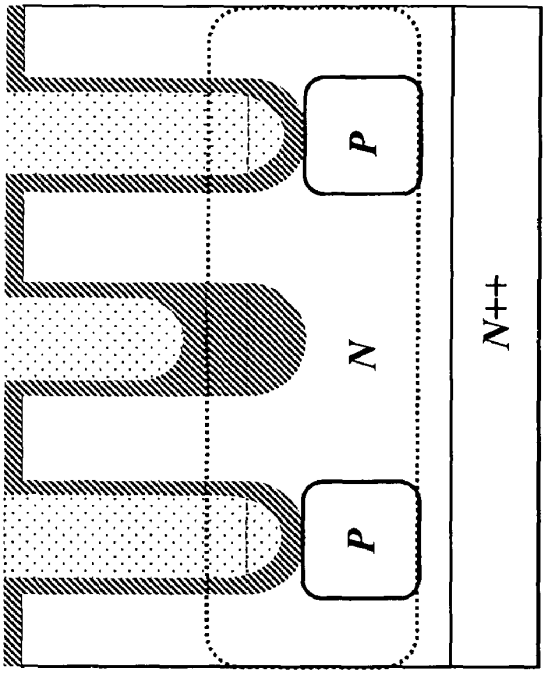


FIG. 15

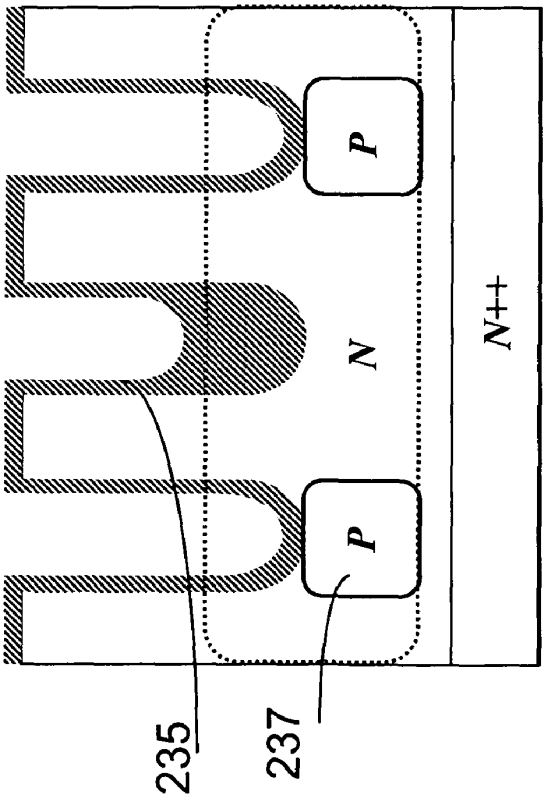


FIG. 14

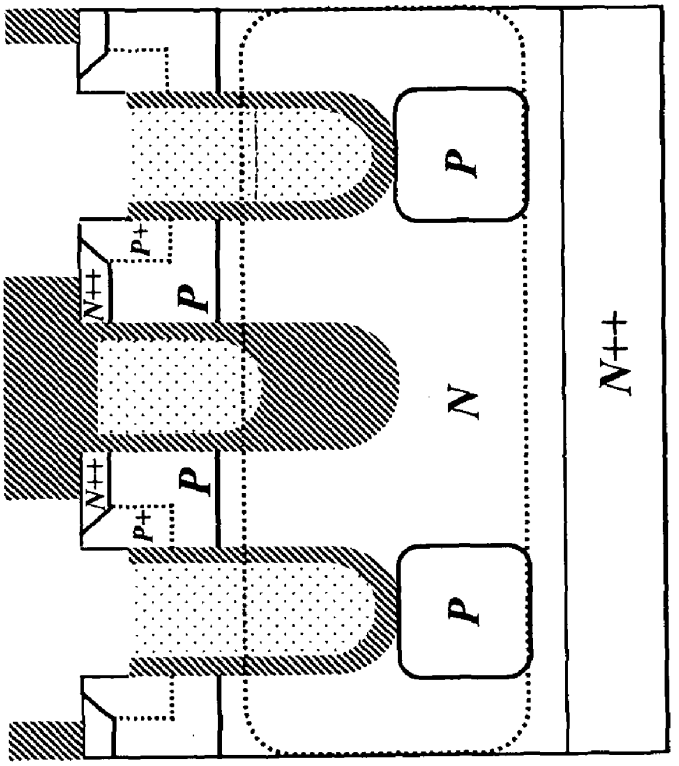


FIG. 17

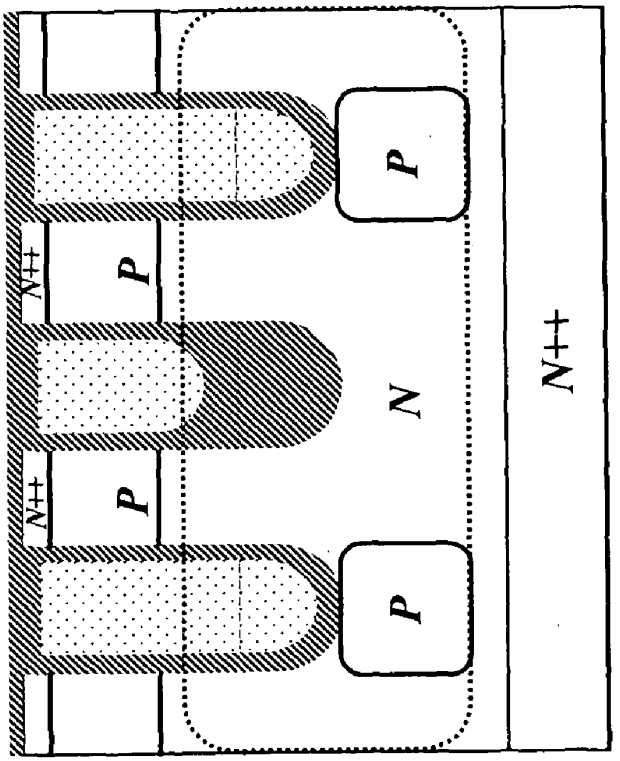


FIG. 16

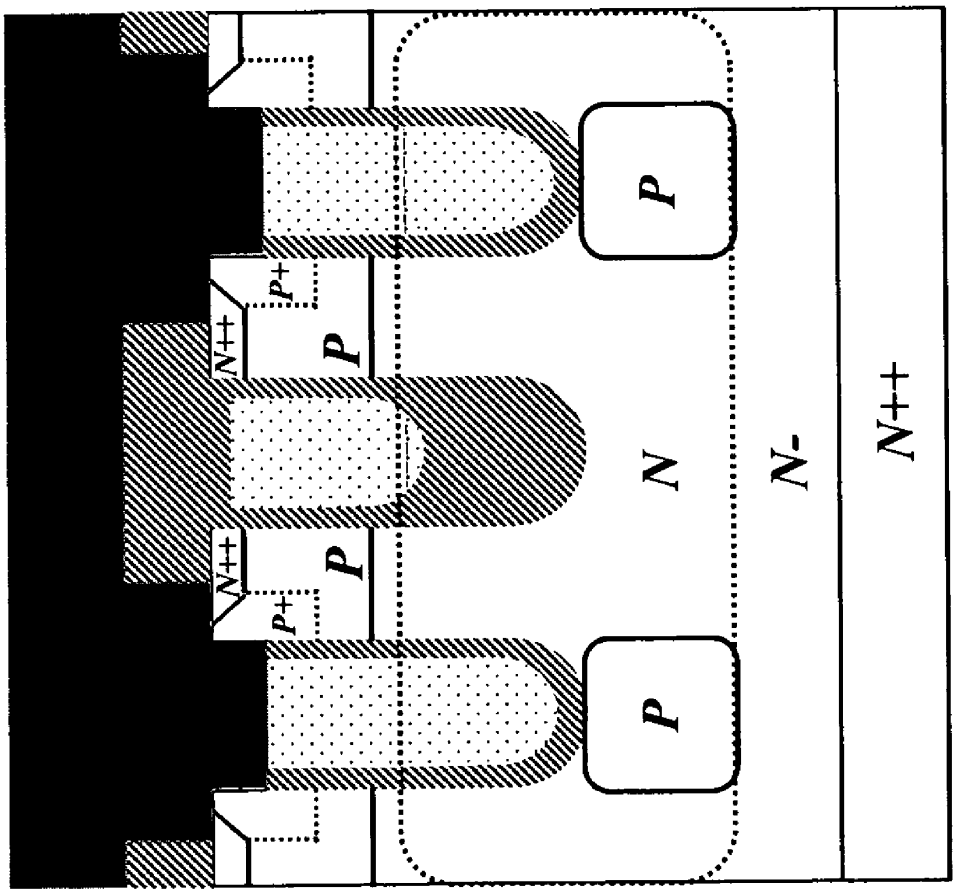


FIG.18

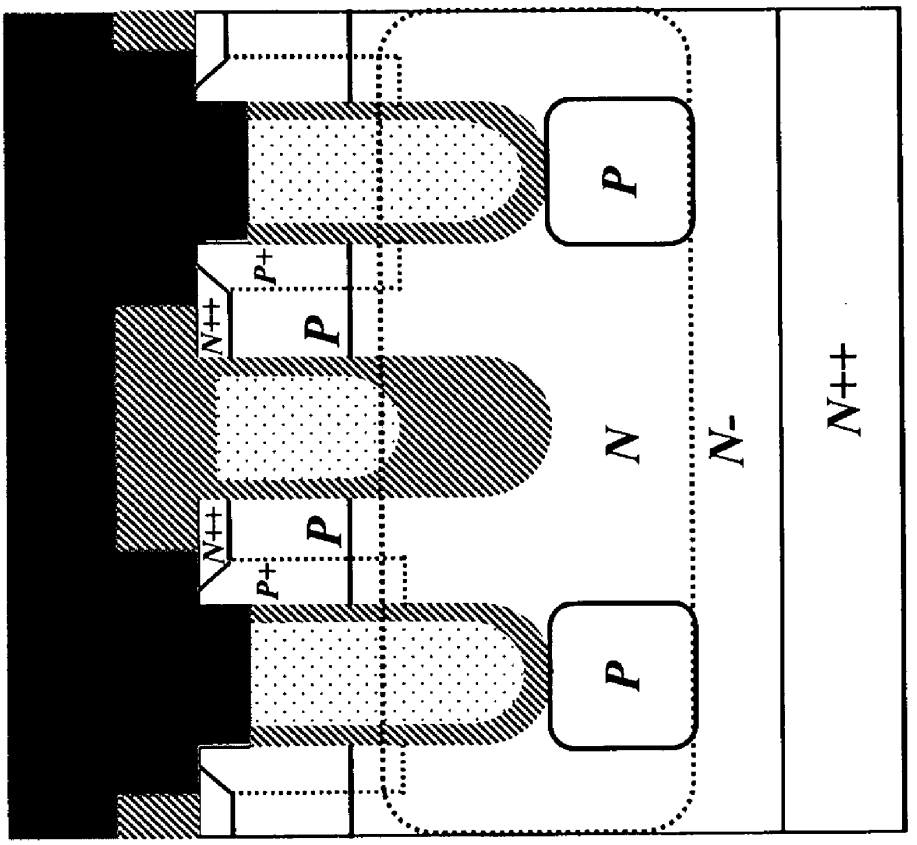


FIG. 19A

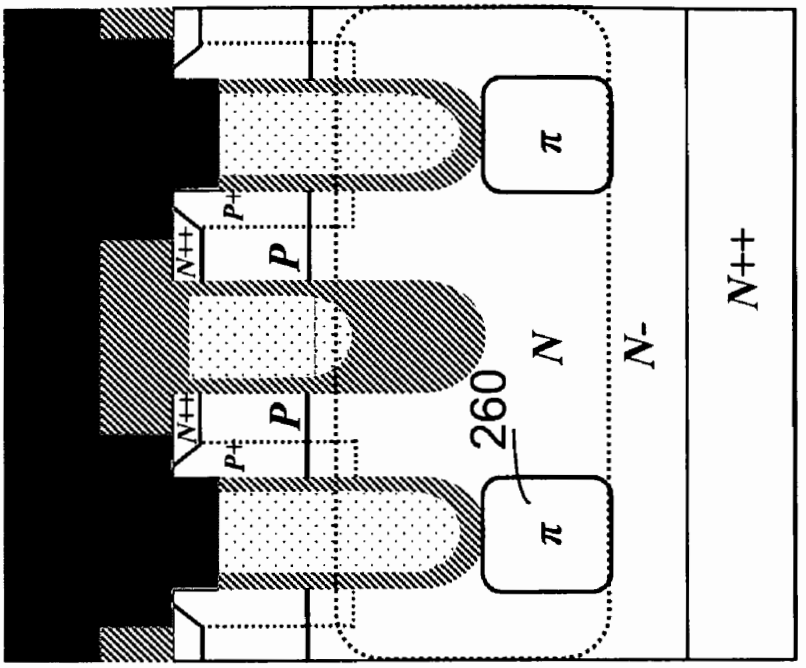


FIG.19C

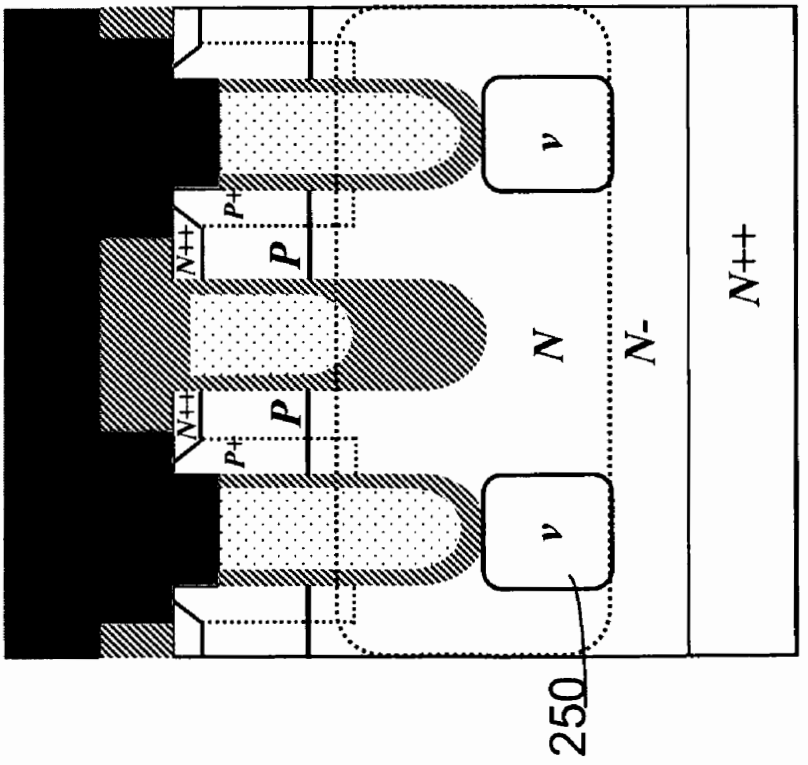


FIG. 19B

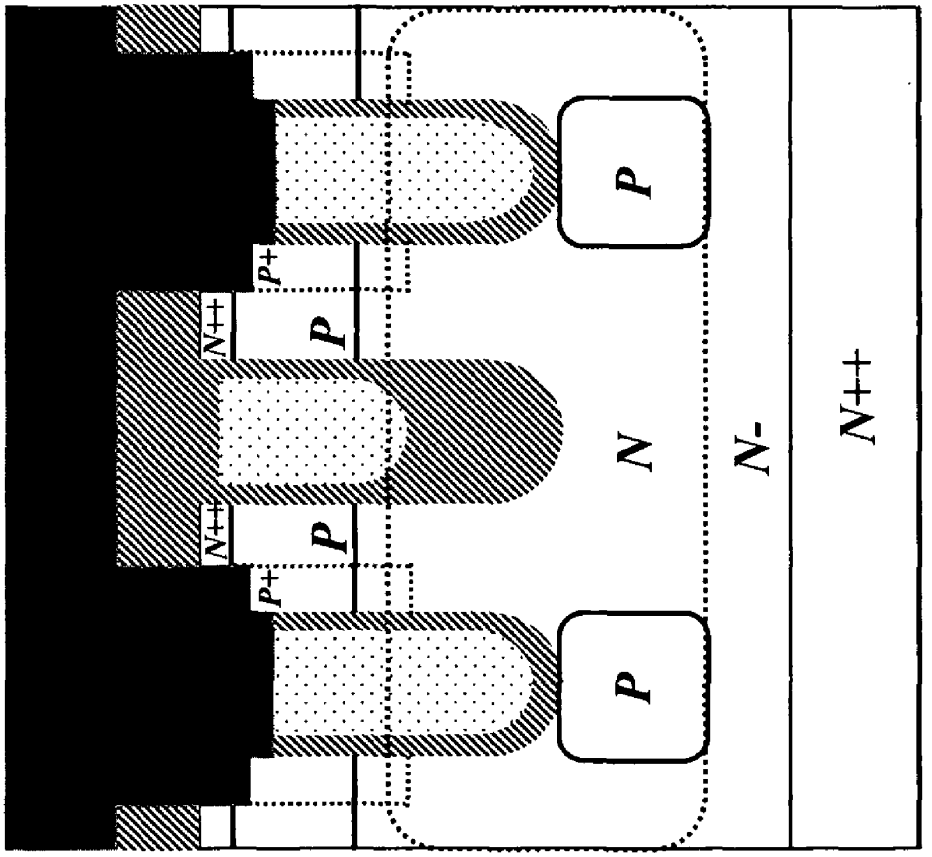


FIG. 20

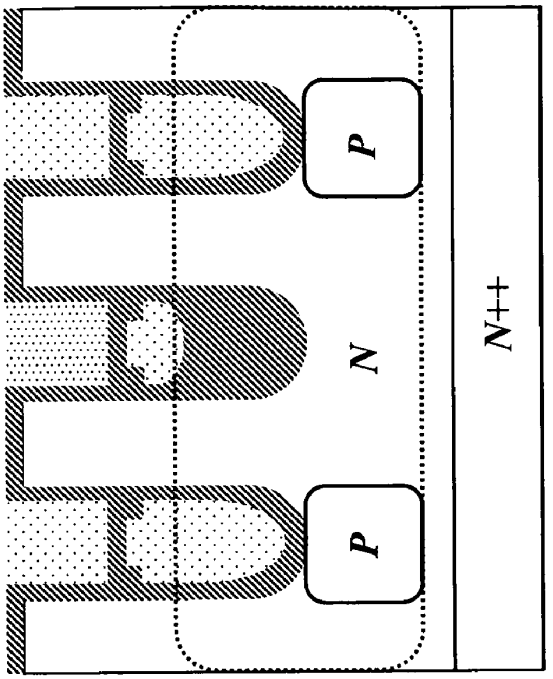


FIG. 22

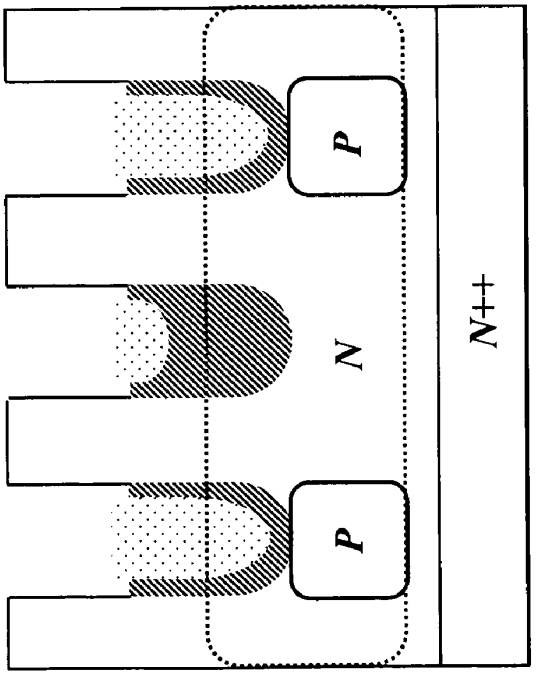


FIG. 21

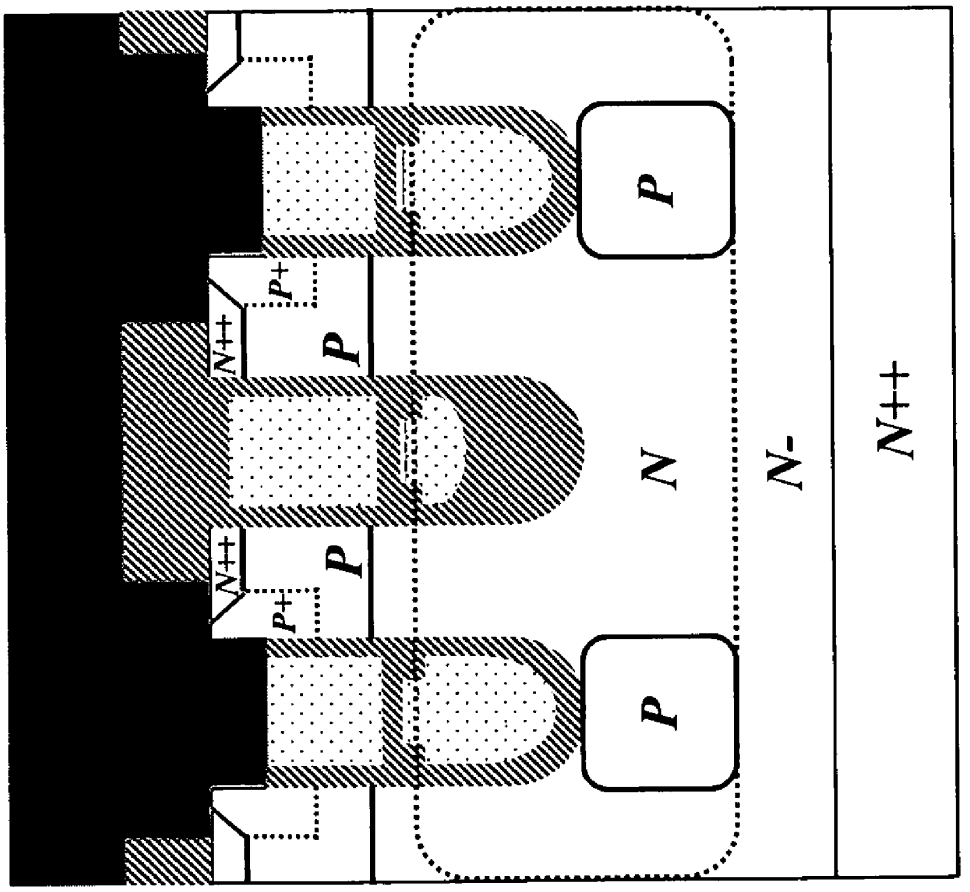


FIG. 23

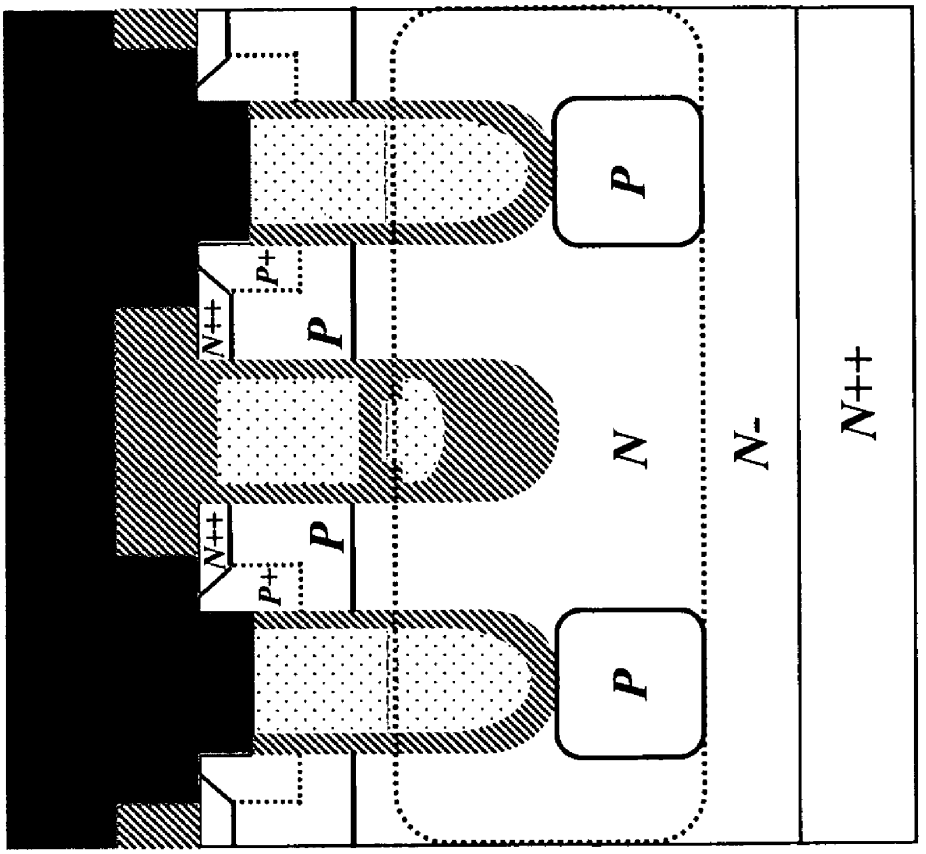


FIG. 24A

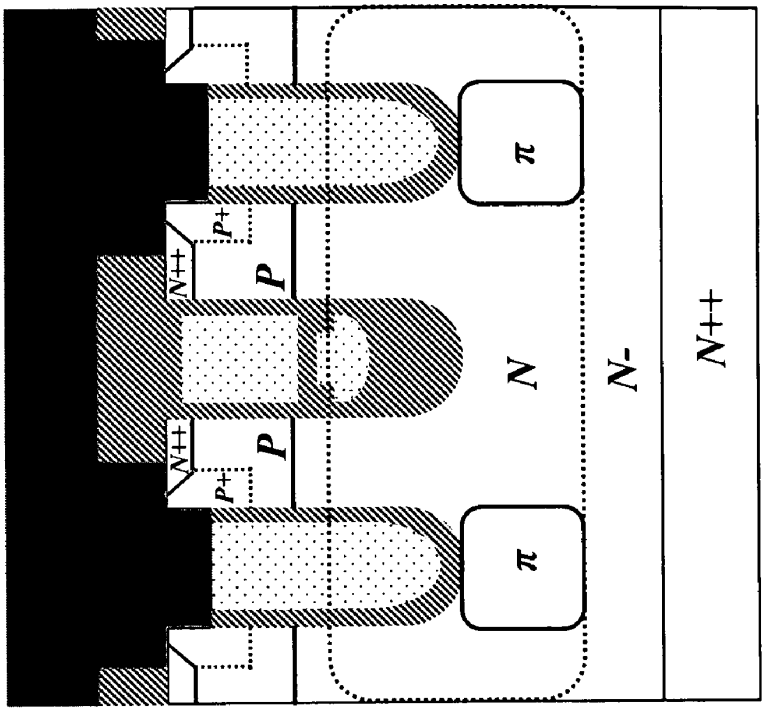


FIG. 24C

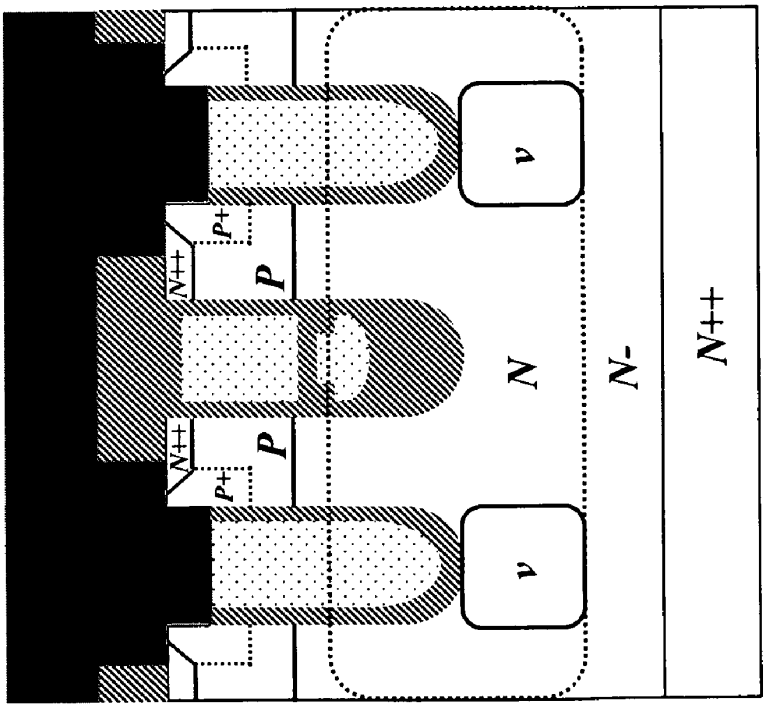


FIG. 24B

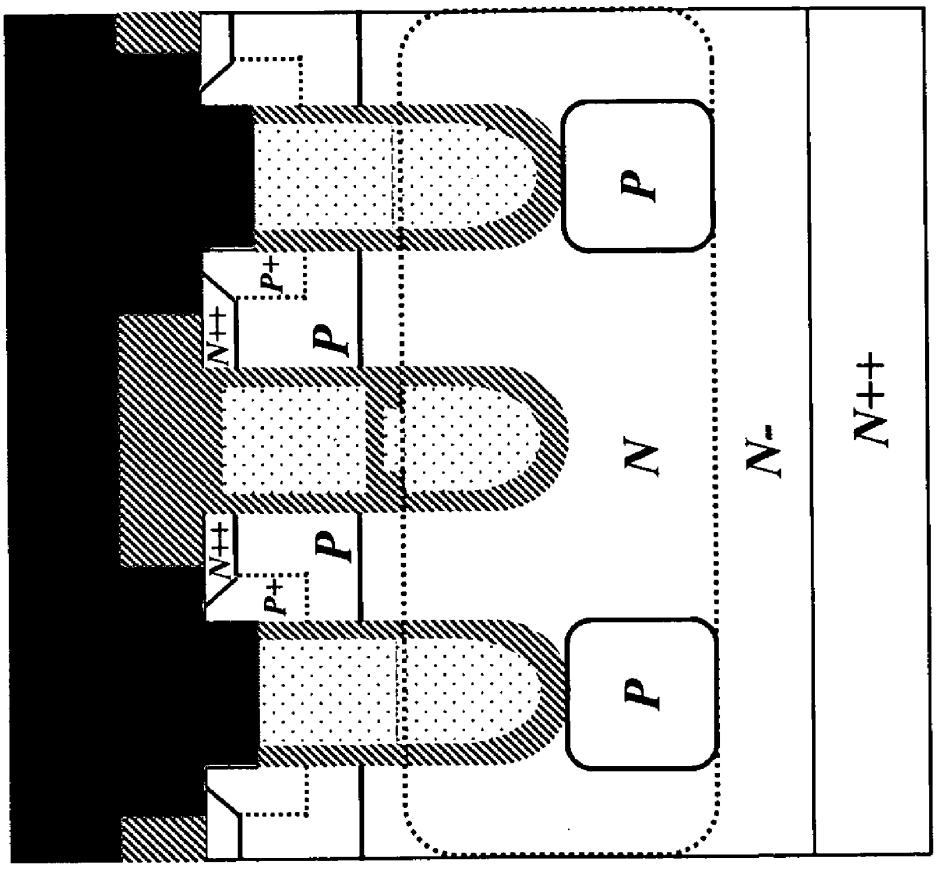


FIG. 25A

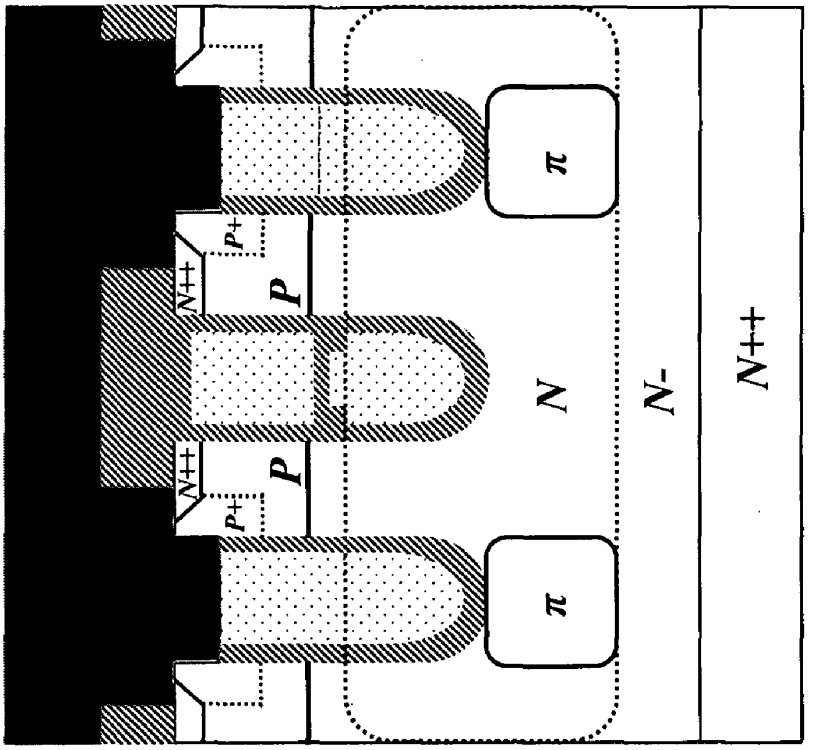


FIG. 25C

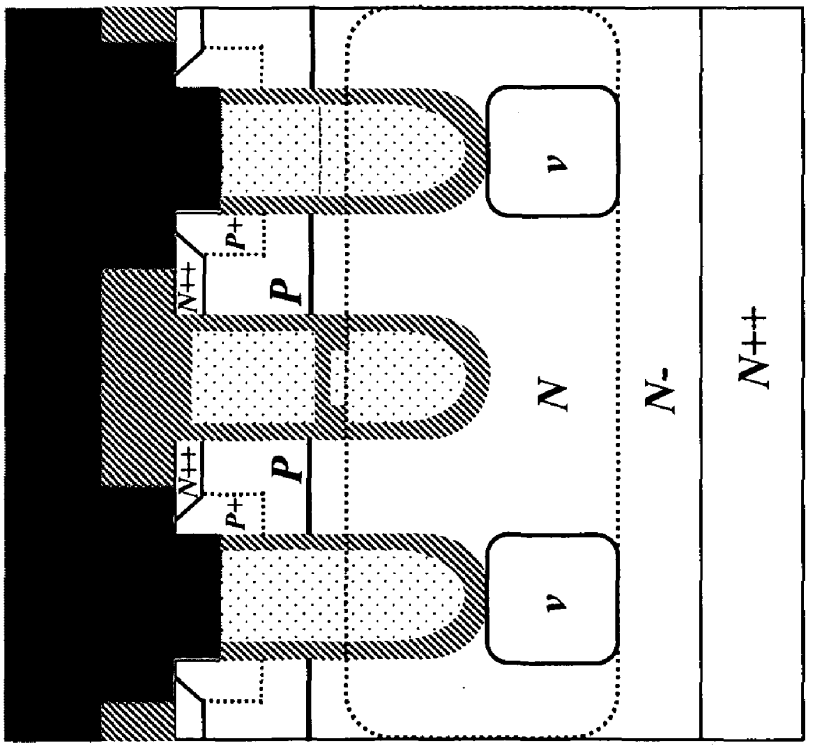


FIG. 25B

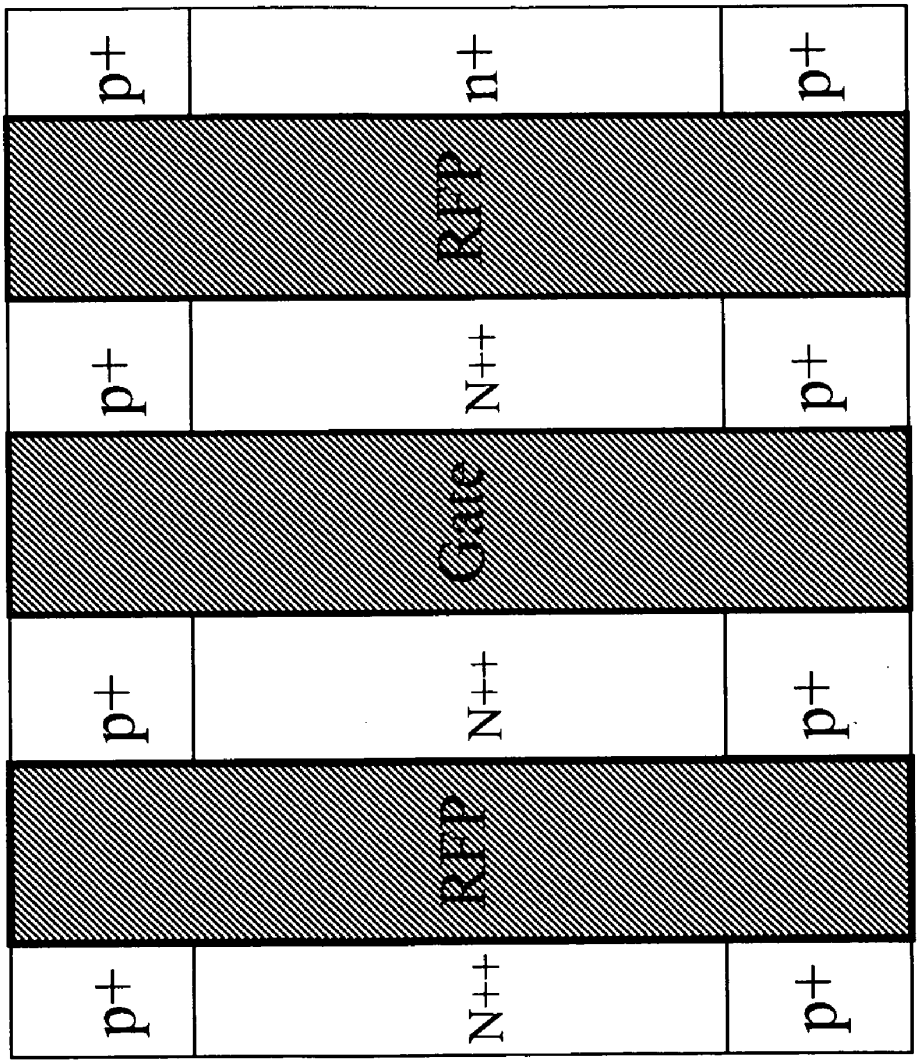


FIG. 26

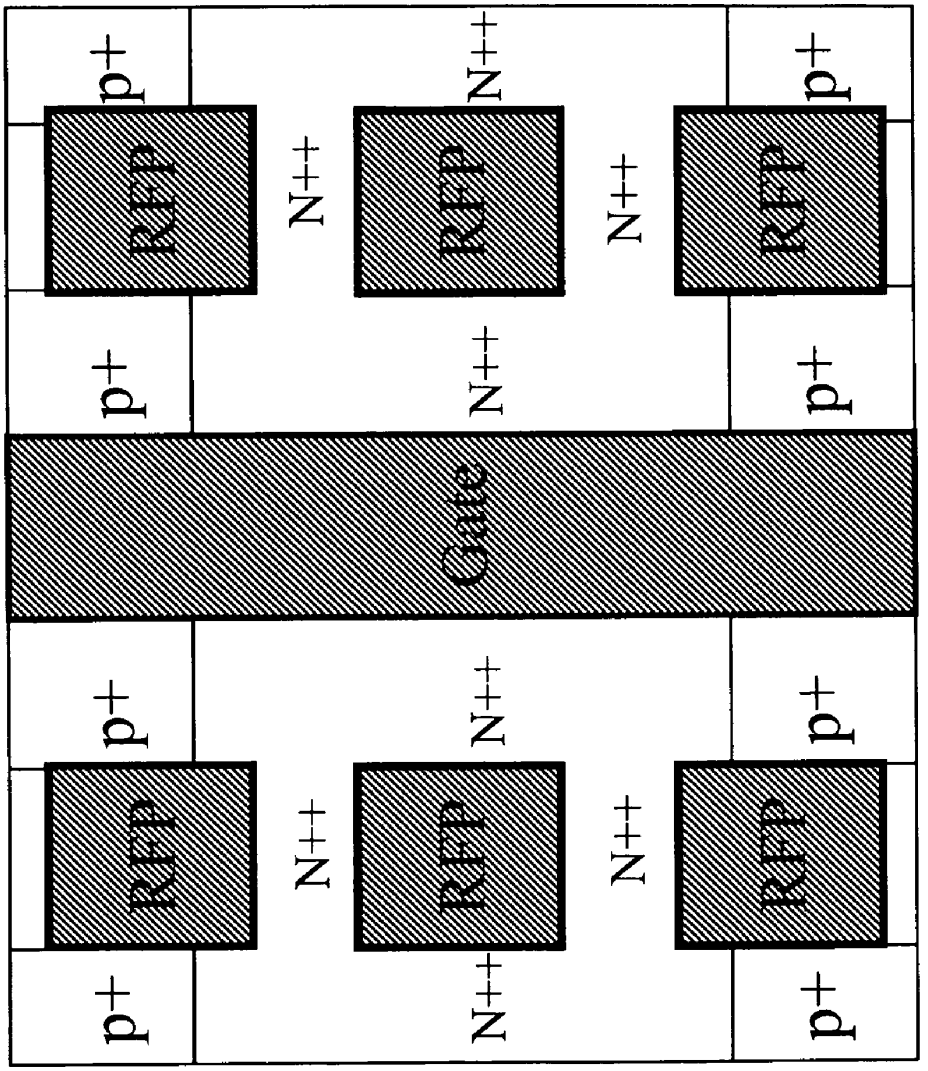


FIG. 27

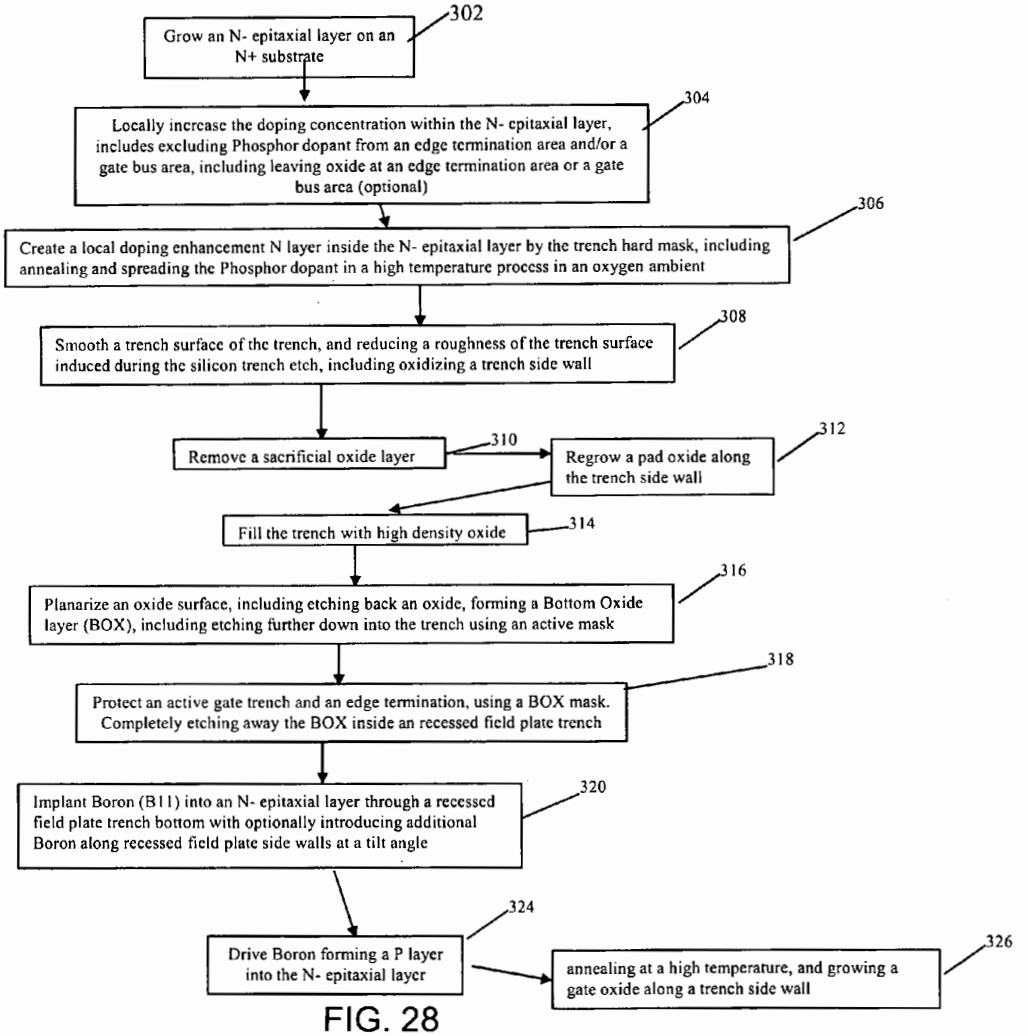

302
Grow an N- epitaxial layer on an N+ substrate
304
Locally increase the doping concentration within the N- epitaxial layer, includes excluding Phosphor dopant from an edge termination area and/or a gate bus area, including leaving oxide at an edge termination area or a gate bus area (optional)
306
Create a local doping enhancement N layer inside the N- epitaxial layer by the trench hard mask, including annealing and spreading the Phosphor dopant in a high temperature process in an oxygen ambient
308
Smooth a trench surface of the trench, and reducing a roughness of the trench surface induced during the silicon trench etch, including oxidizing a trench side wall
310
Remove a sacrificial oxide layer
312
Regrow a pad oxide along the trench side wall
314
Fill the trench with high density oxide
316
Planarize an oxide surface, including etching back an oxide, forming a Bottom Oxide layer (BOX), including etching further down into the trench using an active mask
318
Protect an active gate trench and an edge termination, using a BOX mask. Completely etching away the BOX inside an recessed field plate trench
320
Implant Boron (B11) into an N- epitaxial layer through a recessed field plate trench bottom with optionally introducing additional Boron along recessed field plate side walls at a tilt angle
324
Drive Boron forming a P layer into the N- epitaxial layer
326
annealing at a high temperature, and growing a gate oxide along a trench side wall

FIG. 28

SEMICONDUCTOR DEVICE STRUCTURES AND RELATED PROCESSES

CROSS REFERENCE TO RELATED APPLICATIONS

This application claims priority to U.S. Provisional Application Ser. No. 61/065,759 filed Feb. 14, 2008, which is incorporated by reference herein in its entirety.

BACKGROUND

The present invention relates to field effect transistors and methods, and more particularly to highly reliable power insulated-gate field effect transistors (MOSFET) with a Recessed Field Plate (RFP) and related techniques.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** is a cross-sectional view of a prior art MOSFET with a trench gate having a thick bottom oxide structure.

FIG. **2** is a cross-sectional view of a prior art MOSFET with a split poly gate structure.

FIG. **3** is a cross-sectional view of a prior art MOSFET with RFPs in parallel with the gate trench.

FIG. **4**(*a*) is a cross-sectional view of an RFP containing MOSFET structure with a floating Deep compensated zone.

FIG. **4**(*b*) is a cross-sectional view of an RFP containing MOSFET structure with a Deep compensated zone that extends to and is connected to the source electrode.

FIG. **4**(*c*) is a cross-sectional view of an RFP containing MOSFET structure with a Deep compensated zone that extends to the P body region.

FIG. **5** shows a two-dimensional electrical voltage simulation comparison between the prior RFP-MOSFET structure and a MOSFET containing a Deep compensated zone.

FIGS. **6-18** show successive steps in a sample process for making the sample structure depicted in FIG. **4**(*a*).

FIG. **19**A is a cross-sectional view of a RFP containing MOSFET structure with a Deep compensated zone and a P+ implant region extending beyond the P-N junction and into the N drift region.

FIG. **19**B is a cross-sectional view of an RFP containing MOSFET structure with a Deep compensated zone that is a lightly doped p region, and a P+ implant region extending beyond the P-N junction and into the N drift region.

FIG. **19**C is a cross-sectional view of an RFP containing MOSFET structure with a Deep compensated zone that is a lightly doped n region, and a P+ implant region extending beyond P-N junction and into the N drift region.

FIG. **20** is a cross-sectional view of an RFP containing MOSFET structure with a Deep compensated zone, and a P+ implant region extending beyond P-N junction and into the N drift region and the N++ source region that is completely recessed down.

FIGS. **21-23** show a process for fabrication of an embodiment of FIG. **4**(*a*) structure implemented in the split poly gated structure with split poly layer structure in the RFP trench.

FIGS. **24**A, **24**B, **24**C, **25**A, **25**B and **25**C show cross-sectional views of RFP containing MOSFET structures with a Deep compensated zone (lightly doped p regions in **24**B and **25**B and lightly doped n regions in **24**C and **25**C), implemented in the split poly gated structure with a single poly layer structure in the RFP trench.

FIG. **26** shows a top view of an embodiment of FIG. **4**(*a*) wherein the RFP region is a continuous strip in the horizontal direction.

FIG. **27** shows a top view of an embodiment of FIG. **4**(*a*) wherein the RFP region is divided into several columns in the horizontal direction.

FIG. **28** shows a schematic flow chart for a sample fabrication process.

DETAILED DESCRIPTION OF SAMPLE EMBODIMENTS

Power MOSFETs are widely used as switching devices in many electronic applications. In order to minimize conduction power loss, it is desirable that MOSFETs have low specific on-resistance, which is defined as the on-resistance area product (R_{on}*A), where R_{on} is a MOSFET resistance when the MOSFET is in an ON state, where A is the area of the MOSFET. Trench MOSFETs provide low specific on-resistance, particularly in the 10-100 voltage range. As cell density increases, any associated capacitances such as a gate-to-source capacitance Cgs, a gate-to-drain capacitance Cgd and/or a drain-to-source capacitance Cds also increase. In many switching applications such as synchronous buck dc-dc converters in mobile products, MOSFETs with breakdown voltages of 30 V often operate at higher speeds approaching 1 MHz. Therefore, it may be desirable to minimize switching or dynamic power loss caused by these capacitances. The magnitude of these capacitances are directly proportional to gate charge Qg, gate-drain charge Qgd and output charge Qoss. Furthermore, for a device that operates in the third quadrant (i.e., when a drain-body junction becomes forward biased), minority charge is stored in the device during its forward conduction. This stored charge causes a delay in switching from conducting to non-conducting. To overcome this delay, a body diode with fast reverse recovery is desirable. However, a fast recovery body diode often causes high electromagnetic interference (EMI). This means that during diode recovery, the ratio between the negative going waveform (t_a) and the positive going waveform (t_b) must be less than one for a soft recovery which avoids EMI problem.

As switching speed requirements increase to 1 MHz and beyond with new applications, state of the art power MOSFETs are increasingly unable to operate at such high speeds with satisfactory efficiency. A power MOS transistor that has low charges Qg, Qgd, Qoss and Qrr in addition to having a low specific on-resistance (Ron*A), is desirable.

There are two common techniques to improve the switching performance of power MOSFETs. The first one is the trench-gated MOSFET with thick bottom oxide, as shown in FIG. **1** (U.S. Pat. No. 6,849,898). The second one is the split poly gated MOSFET structure, in which the first poly gate is electrically shorted to the source electrode (U.S. Pat. Nos. 5,998,833, 6,683,346), as illustrated in FIG. **2**.

Recently, as shown in FIG. **3**, US Patent Application No. 2008/0073707 A1 to Darwish disclosed a power MOSFET with the recessed field plate (RFP) structure which realize a very short channel region (~0.25 um) for further reducing the gate-source capacitance and the gate-drain capacitance, and consequently, the total gate charge (Qg) and the "Miller" charge (Qgd). The RFP structure additionally improves the body diode reverse recovery speed due to providing an additional path for current and the enhanced depletion of the drift region induced by the RFP.

The present application discloses improvements to power insulated-gate field effect transistors with Recessed Field Plate (RFP) and similar structures. The inventors have real-

ized that the performance of RFP-type power MOSFETs can be improved by performing a compensating implant into the RFP trench. This compensating implant helps to shape the depletion boundaries in the OFF state, and thus helps to avoid punchthrough. Because of this, a local enhancement can also be added to the doping between channel and drain, in the drift or spreading region. This provides a synergistic combination, wherein the on-resistance can be improved with no degradation in breakdown voltage.

The disclosed innovations, in various embodiments, provide one or more of at least the following advantages. However, not all of these advantages result from every one of the innovations disclosed, and this list of advantages does not limit the various claimed inventions.

- Improved (reduced) on-resistance;
- Improved (increased) breakdown;
- Reduced electrical stress on any dielectric layer at the bottom of the RFP trench;
- Higher reliability and longer operation life; and/or
- Increased ability to increase local doping concentration in the drift region.

The numerous innovative teachings of the present application will be described with particular reference to presently preferred embodiments (by way of example, and not of limitation). The present application describes several embodiments, and none of the statements below should be taken as limiting the claims generally.

For simplicity and clarity of illustration, the drawing figures illustrate the general manner of construction, and description and details of well-known features and techniques may be omitted to avoid unnecessarily obscuring the invention. Additionally, elements in the drawing figures are not necessarily drawn to scale, some areas or elements may be expanded to help improve understanding of embodiments of the invention.

The terms "first," "second," "third," "fourth," and the like in the description and the claims, if any, may be used for distinguishing between similar elements and not necessarily for describing a particular sequential or chronological order. It is to be understood that the terms so used are interchangeable. Furthermore, the terms "comprise," "include," "have," and any variations thereof, are intended to cover non-exclusive inclusions, such that a process, method, article, apparatus, or composition that comprises a list of elements is not necessarily limited to those elements, but may include other elements not expressly listed or inherent to such process, method, article, apparatus, or composition.

It is contemplated and intended that the design apply to both n-type and p-type MOSFETs; for clarity reason, the examples are given based on n-channel MOSFET structure, but an ordinary person in the art would know the variations to modify the design to make a similar p-channel device.

The present application discloses improvements to power insulated-gate field effect transistors with Recessed Field Plate (RFP) and similar structures. The inventors have realized that the performance of RFP-type power MOSFETs can be improved by performing a compensating implant into the RFP trench. This compensating implant helps to shape the depletion boundaries in the OFF state, and thus helps to avoid punchthrough. Because of this, a local enhancement can also be added to the doping between channel and drain, in the drift or spreading region. This provides a synergistic combination, wherein the on-resistance can be improved with no degradation in breakdown voltage.

In one sample embodiment, the RFP containing MOSFET has a buried Deep compensated zone floating in the N body region underneath the RFP trench. The Deep compensated zone reduces the voltage across the dielectric layer between the RFP and the N-epitaxial layer when a high drain-source voltage is applied.

In one sample embodiment, the RFP containing MOSFET has a buried Deep compensated zone floating in the N body region underneath the RFP trench and a local enhancement to the doping between channel and drain, in the drift or spreading region in the N epitaxial layer.

In one embodiment, the RFP containing MOSFET also has a deep P+ region in the P body in contact with the RFP trench wall that extends from the P body into the N epitaxial layer.

In one embodiment, the Deep compensated zone underneath the RFP trench extends to and is connected to the source electrode.

In one embodiment, the Deep compensated zone underneath the RFP trench extends to P body region that is in contact with the side walls of the RFP trench.

In one embodiment, the Deep compensated zone is a very lightly doped p region; while in another embodiment, the Deep compensated zone is a very lightly doped n region.

Referring now to FIG. **4**(*a*), a semiconductor device structure **100** comprises a gate **102** which is positioned in a first trench **104**, also described herein as a gate trench **104**. The first trench **104** containing the gate **102** may be one of many gate trenches within the semiconductor device structure **100**. The semiconductor device structure **100** is capacitively coupled to control vertical conduction from a source region **106**, having a first-conductivity-type, through semiconductor material **108**, which is adjacent to the first trench **104**.

As shown in FIG. **4**(*a*), the gate **102** has a gate electrode comprising conductive gate material having a width approximately equal to a width of the gate trench **104**. It will be appreciated that, although the gate electrode may have a width that is approximately equal to the width of the gate trench **104**, the gate electrode may alternatively be contacted using a wider gate trench and a smaller gate electrode, allowing the gate trench to be insulated from the gate conductor.

The semiconductor device structure **100** also includes recessed field plates **110**, which are positioned in proximity to and capacitively coupled to the semiconductor material **108**. The recessed field plates **110** are positioned in respective second trenches **112**, also described herein as RFP trenches **112**. Each of the trenches (i.e., the respective second trenches **112** and the gate trenches) has trench walls that are coated with an insulating material, such as silicon dioxide (SiO_2). The RFP trenches **112** contain an insulating material having a breakdown voltage which preferably exceeds the breakdown voltage of the semiconductor device structure **100**. The gate trench **104** preferably contains an insulating material up to the p-body drain junction, minimizing any overlap of the gate electrode (connected to the gate **102**) with the drain or drift region.

In one embodiment, the gate trench contains thick bottomed insulation dielectric material such as silicon dioxide. In another embodiment, the insulating material within the RFP trench and/or the gate trench **104** has a stepped thickness. Providing a stepped thickness can help shape the channel and can help control "hot" electron effects.

Conductive material, such as n-type doped polysilicon, forms a gate electrode electrically separated from the gate trench **104** by the insulating material. The conductive material may be silicided to reduce its resistance. Conductive material also fills the RFP trenches **112**, electrically separated from the gate trench **104** by the insulating material, and extending above the RFP trenches to form a plurality of RFP electrodes. Each of the trenches may be of substantially equal depth, or may differ in depth, and may be self-aligned by

being etched at the same processing step, although the RFP electrode is deeper than the gate electrode and is either independently biased or connected to a source electrode (i.e., the source **106**), and a source region (including the source electrode) may extend between the gate **106** and the RFP trenches **112**.

In one embodiment, the n-epitaxial drift region is uniformly doped. In another embodiment, the n-epitaxial drift region is not uniformly doped. Specifically, the doping is graded to have a doping concentration that is higher at an interface with the underlying **118** substrate and decrease toward the surface. Non-uniform doping of the drain drift region allows for greater shaping of the channel and for control over "hot" electron injection.

The source region may be doped n+. The gate trench **104** and the RFP trench may have a thin layer of the insulating material, reducing on-resistance, or a thick layer of the insulating material, providing greater electrical isolation increasing reverse-bias breakdown voltage. In the embodiment depicted, the RFP electrodes have a uniform depth. In another embodiment, at least one of the RFP electrodes extends up and contacts the source **106**.

Advantageously, the semiconductor device structure **100** also includes deep compensated zones **114** of either p-type or n-type lying at least partially beneath the respective RFP trenches **112**. The deep compensated zone **114** may be floating islands of either p-type dopant concentration regions (as shown in FIG. **4***a*), or lightly doped n-type dopant concentration regions in the N-drift region underneath the RFP trenches. The drawing shows the boundaries of this compensated zone **114** as if it had been fully counterdoped, but those of ordinary skill will understand that the boundaries of a compensated but not counterdoped zone can be similarly envisioned, using e.g. the concentration contours of a single dopant species.

The deep compensated zones **114** also reduce the voltage across the dielectric layer between the RFP and the N-epitaxial layer when a high drain-source voltage is applied.

As shown in FIG. **4**(*b*), device **100** also contains a deep p-body region **116** that is in contact with the side walls of the RFP trench **112**. The deep p-body region **116** with boundary at **116***a* can be in connection with the source electrode and also be in connection with the deep compensated zones **114**. The deep P-N junction in the edge termination can be formed by the Deep compensated implant and its related annealing without adding new mask. Therefore, the disclosed structure can offer a more reliable edge termination.

Alternatively, as shown in FIG. **4**(*c*), the deep compensated zones **114** can extend vertically and merge with the p-body region.

The two-dimensional electric voltage simulation shown in FIG. **5**, reveals that under the same bias conditions the conventional device of structure shown in FIG. **3** sees about 19V across the bottom dielectric layer between RFP and the N-epitaxial layer, while an embodiment of FIG. **4**(*a*)-(*c*) only shows 7V across the bottom dielectric layer between RFP and the N-epitaxial layer due to the protection from the Deep compensated zone **114**.

As the electrical stress on the bottom dielectric layer between RFP and drain is reduced significantly, the device structures of FIG. **4**(*a*)-(*c*) will offer a higher reliability and longer operation life. In addition, the Deep compensated zones **114** enhance the lateral and vertical depletion of N-epitaxial layer, which provides room for higher local doping concentration in the epitaxial layer without degrading the device breakdown voltage.

The increase in local doping concentration in the epitaxial layer further reduces the on-resistance of drift region. By properly adjusting the doping concentration of P and N regions in the N-epitaxial layer, the total on-resistance of device can be lowered without reduction of the breakdown voltage. Furthermore, the local doping enhanced N layer also decreases the minority carrier injection efficiency of the body diode of the device and alters the electric field distribution during the body diode reverse recovery. Thus, the reverse recovery of the body diode is improved, resulting in a device having lower reverse recovery charge and soft recovery features.

Since the doping enhancement only occurs in the active region, the termination efficiency of the improved device edge junction termination region will not be degraded.

The recessed field plates **110** may be positioned in multiple respective trenches **112**, which are separate from the gate trench **104**. Accordingly, the semiconductor device structure **100** may be, for example, an n-channel MOSFET with a recessed field plate (RFP) trench **112** and a gate trench **104** formed on an N-type epitaxial layer grown over a heavily doped N+ substrate.

In third quadrant operation, in which where the drain **118** is negatively biased with respect to a source-body electrode (i.e. the source **106**), and in which diffusion current results in minority carrier injection and a high reverse recovery charge Qrr, the plurality of RFP electrodes form majority carrier channel current path from drain to source in addition to that provided by the gate electrode in a conventional structure. The combined effect of the RFP electrodes and the gate electrode is both reduced minority carrier diffusion current and reduced recovery charge Qrr. Accordingly, in the third quadrant operation, the RFP electrodes act as an additional gate without any penalty of an added gate-drain capacitance Cgd.

In reverse-biased operation, the RFP also reduces any electric field in a channel region. Accordingly, shorter channel lengths are possible, without substantial risk of punch-through breakdown, further allowing reduction in Ron*A and Qg. The capacitive coupling between the gate trench **104**, the RFP trenches **112** and the drain region further deplete the drain drift region, at a higher rate as a drain-source voltage VDS is increased in an off-state. The low Cgd and its fast falling rate, combined with increasing drain-source voltage VDS, provides a lower gate-drain charge.

The semiconductor device structure **100** can have a quasi-vertical or lateral configuration. Ensuring that the semiconductor device structure **100** has a quasi-vertical or lateral configuration can help shape the channel, and can reduce hot electron effects.

Various variations in gate conductor and RFP conductors may be used. Various combinations have been shown in US Application No. 2008/0073707 A1 to Darwish, the entirety of which is thereby incorporated by reference. Polysilicon may be used as the conductive material. Example variations in structural designs of gate conductor and RFP conductors include split poly configurations and single poly configurations (FIGS. **21**-**25**), thick bottom oxide, and step shaped bottom oxide and the combinations of the various forms.

Referring to FIG. **26**, each of the foregoing embodiments may be implemented in a single configuration, a multi-stripe configuration, a cellular layout configuration, or a combination of the foregoing. Moreover, the polarity and conductivity type may be reversed.

Referring to FIG. **27**, each of the foregoing embodiment RFP may also be implemented in an interrupted manner where the RFP trenches and conductors form columns in the

source-body-drain layers of the device. With this interrupted scheme, more N++ surface area can be provided, reducing the N++ resistance, and lowering the total on-resistance.

A fabrication process for making the described embodiment is detailed in FIG. **6-18**. In FIG. **6**, starting with N++ substrate **201**, the N-epitaxial layer **203** is grown followed by forming a thin layer of silicon oxide layer **205**. Substrate **201** may have been doped with phosphorus or arsenic. The preferred thickness for oxide layer **205**, for example, can be 200-300 Å. In FIG. **7**, the trench mask **207** is applied to form the hard mask for trench etching and the oxide layer is etched

Then a standard silicon etch step is carried out to form the plurality of trenches **209** according to the mask. In FIG. **8**, blanket implanting of phosphorus ions **211** (e.g. P^{31}) to the whole device may be performed to locally increase the doping concentration of N-epitaxial layer. The implantation is preferred to be done at tilt of 0 degree. The trench mask around the edge termination area or the gate bus area (not shown in Figures) prevents the phosphor dopant getting into these areas. Therefore, only the active region of the device receives the doping enhancement implant.

After implantation a high temperature process in an oxygen-containing ambient is used to anneal and diffuse the phosphorus dopant. Consequently, a doping enhancement N layer **213** is formed inside N-epitaxial layer as shown in FIG. **9**. The trench walls may then be oxidized first using a sacrificial oxidation. After removing the sacrificial oxide layer a pad oxide is re-grown along trench side wall. In FIG. **10**, the trenches are filled with high density oxide **217**. Oxide **217** may include silicon dioxide, or other types of deposited oxide, such as LTO or TEOS or High Density Plasma (HDP) oxide. The oxide is then thinned as shown in FIG. **11** using a dry plasma etch or CMP technique to planarize the oxide surface **219**.

In FIG. **12**, after active mask **223** has been applied with openings over trenches **222**, the oxide is etched further down into trench forming the trench Bottom Oxide layer (BOX) **221**. Then, in FIG. **13**, the BOX mask is used to protect the active gate trench **225** and the edge termination areas (not shown). The oxide removal step is carried out to completely etch away the BOX inside the RFP trench. Before removing the BOX mask, boron-11 ions **229** are implanted into N/N-epitaxial layer through RFP trench bottom **231**, forming P layer or isolation zones **237** shown in FIG. **14**.

In one embodiment, to implement the structure shown in FIG. **4**(*c*), a tilt angle implant is used to introduce boron along the RFP side walls. After removal of BOX photoresist **233** an optional high temperature anneal is employed to diffuse the boron, forming P layer or isolation zones **237** inside N-epitaxial region. Then the gate oxide **235** is grown along the trench sidewall in FIG. **14**.

The rest of process steps shown in FIG. **15** to FIG. **17** are similar to the one described in FIGS. 14-17 in US patent application No. 2008/0073707, which is herein incorporated by reference. The final device structure is shown in FIG. **18**. It is essential to point out that, by properly choosing RFP poly recess depth combined with the implant energy of P+ implant, the P+ region can be made deeper than P body, as shown in FIG. **19**A. Depending on the doping concentration of P shield region (or isolation zone), the P shield zone could be a "π" region **260** (a very lightly doped P region) shown in FIG. **19**C or a "ν" region **250** (a very lightly doped n region) shown in FIG. **19**B. A deeper P+ region is desired in order to improve the device ruggedness and connect the buried P region to the source electrode. In addition, the N++ source region can also be recessed completely as shown in FIG. **20**, so that the N++ source mask-photo step can be eliminated.

Furthermore, the techniques proposed in this invention may also be implemented using split poly gated device structures. One of implementation schemes are briefly demonstrated in the FIG. **21** to FIG. **23**. The process includes deposition of a first poly layer in the trenches, poly etch-back, and oxide removal, gate oxidation, second poly layer deposition, and CMP and/or poly etch-back. The split gated double poly configuration shown in FIG. **21-23** is used to replace the single poly layer in the active trench gate and the RFP trench shown in FIG. **18**. In this case, the bottom poly layer and the upper poly layer in RFP trench are both electrically shorted to the source metal. In addition, the split poly layers in the RFP region of device in FIG. **23** can be directly replaced by the single RFP poly layer as demonstrated by FIG. **24**A and FIG. **25**A. Depending on the doping concentration of P shield region (or isolation zone), at very light concentration, the P shield zone could be a "π" region (a very lightly doped P region) or a "ν" region (a very lightly doped n region) as shown in FIG. **24**B, **24**C, **25**B and **25**C.

FIG. **28** is a flow chart depicting a fabrication process for making a MOSFET in accordance with one embodiment of the present invention. The fabrication process includes growing **302** an N-epitaxial layer on an N+ substrate. The fabrication process also includes locally increasing **304** the doping concentration within the N-epitaxial layer. Locally increasing **304** the doping concentration within the N-epitaxial layer includes blanket implanting phosphorous. The blanket implanting of phosphorous may be at a tilt angle of zero degrees, or may be at some other tilt angle. Locally increasing **304** the doping concentration within the N-epitaxial layer also includes excluding phosphorous dopant from an edge termination area and/or a gate bus area, including leaving oxide at an edge termination area or a gate bus area.

The fabrication process for making a MOSFET also includes creating **306** a doping enhancement N layer inside the N-epitaxial layer, including annealing and spreading the phosphorous dopant using a high temperature thermal process in an oxygen ambient. The fabrication process for making a MOSFET also includes smoothing **308** the trench surface of the trench, and reducing the roughness of the trench surface induced during the silicon trench etch, including oxidizing the trench side wall, removing **310** the sacrificial oxide layer, and regrowing **312** the pad oxide along the trench side wall.

The fabrication process for making a MOSFET also includes filling **314** the trench with high density oxide, planarizing oxide surface **316**, including etching back an oxide, forming a Bottom Oxide layer (BOX), including etching further down into the trench using an active mask, and protecting **318** an active gate trench and an edge termination, using a BOX mask. The fabrication process for making a MOSFET also includes implanting **320** boron (B11) into an N-epitaxial layer through a recessed field plate trench bottom, including introducing boron along recessed field plate side walls at a tilt angle, completely etching away **322** the BOX inside an recessed field plate trench, including removing the oxide, optionally driving **324** boron forming a P layer into the N-epitaxial layer, including annealing at a high temperature, and growing **324** a gate oxide along a trench side wall.

For a sample 40V embodiment, preferred parameters are as follows. However, it must be understood that these parameters would be scaled for different operating voltages, and of course they can also be adapted for use with many other processes. In this sample embodiment, the trenches are 0.3 microns wide, about 1.0 micron deep, and are laid out on a one micron pitch. (The cell pitch is two microns, since there are two types of trench present.) In this sample embodiment, the

starting material is 0.35 ohm-cm n-on-n+ epi, about 5.5 microns thick. A blanket n-enhancement implant is performed, e.g. with phosphorus at 3 E12/cm^2 (i.e. 3×10^{12} cm^{-2}). The trenches are then etched. After a sacrificial oxidation and trench fill (preferably using a deposited oxide plus oxidation), an etchback is preferably performed to clear the trenches to about half their depth. Photoresist is then patterned, to expose the RFP trenches but not the gate trenches, and the oxide plugs are removed from the RFP trenches. A P-type implant is then performed to form the P-isolation regions; in this example, a combination of two boron implants, one at 2.5 E12/cm^2 at 30 keV plus another 2 E12 at 120 keV. This will produce a counterdoped or compensated isolation region **114** below the RFP trenches, of about 0.7 micron depth. The remaining process steps then proceed conventionally, with formation of gate, body, source, contacts, etc.

As mentioned above, the locally enhanced n-doping which connects gate to drain, in various embodiments described above, reduces on-resistance. However, it is the improved off-state behavior provided by the added isolation regions which makes this enhanced n-doping possible.

In alternative embodiments, the depth of the isolation region can be e.g. from 0.25 micron to 2.5 micron, and scaled accordingly for operating voltages other than 40V. Similarly, the isolation implant, in alternative embodiments, can use a dose from 2 E12 cm^{-2} to 1 E13 at 20-320 keV, or even higher or lower doses and/or energies, plus allowance for scaling.

It will be appreciated that the foregoing is merely a description of some specific illustrative and exemplary embodiments of the present invention, and should not be considered as descriptive of the entire gamut of embodiments that fall within the scope of the present invention.

According to various embodiments, there is provided: a semiconductor device structure, comprising a gate which is positioned in a first trench, and capacitively coupled to control vertical conduction from a first-conductivity-type source through semiconductor material which is adjacent to said trench; recessed field plates, positioned in proximity to and capacitively coupled to said semiconductor material; said recessed field plates being positioned in respective second trenches; and diffusions of a second conductivity type lying at least partially beneath said respective second trenches.

According to various embodiments, there is provided: a semiconductor device structure, comprising a semiconductor layer; a gate which is positioned in a first trench within said semiconductor layer, and is capacitively coupled to control vertical conduction from a first-conductivity-type source through second-conductivity-type portions of said layer near said trench; recessed field plates, positioned in proximity to and capacitively coupled to said semiconductor material; said recessed field plates being positioned in respective second trenches; diffusion components of a second conductivity type lying at least partially beneath said respective second trenches; whereby said diffusion components reduce depletion of said second-conductivity-type portions of said layer in the OFF state.

According to various embodiments, there is provided: a semiconductor device structure, comprising a semiconductor layer; a gate which is positioned in a first trench within said semiconductor layer, and is capacitively coupled to control vertical conduction from a first-conductivity-type source through second-conductivity-type portions of said layer near said trench; recessed field plates, positioned in proximity to and capacitively coupled to said semiconductor material; said recessed field plates being positioned in respective second trenches; a first additional diffusion component of a second conductivity type lying at least partially beneath said respective second trenches; and a second additional diffusion component of said first conductivity type lying at least partially within said second-conductivity-type portions of said layer; whereby said first additional diffusion component reduces depletion of said second-conductivity-type portions of said layer in the OFF state; and whereby said second additional diffusion component reduces the on-resistance of the device in the ON state.

According to various embodiments, there is provided: an improved RFP transistor structure having (a) low total on-resistance, (b) reduced minority carrier injection efficiency (of body diode), (c) improved reverse recovery (of body diode), (c) lower reverse recovery charge, (d) soft recovery characteristic, (e) as reliable edge termination, without either reduction of breakdown voltage or degradation of the termination efficiency of device edge junction termination region, the improved structure comprising: an RFP transistor structure, including at least one or more gate trenches adjoined by one or more recessed-field-plate trenches; and respective deep compensated zones underneath said recessed-field-plate trenches.

According to various embodiments, there is provided: a method for operating a semiconductor device structure, comprising: controlling conduction between first and second source/drain electrodes through a channel location in semiconductor material using a gate electrode positioned in a first trench to provide at least ON and OFF states; and avoiding punchthrough of said channel location, using both one or more recessed field plates, positioned in proximity to and capacitively coupled to said semiconductor material; said recessed field plates being positioned in respective second trenches, and one or more diffusion components of a second conductivity type lying at least partially beneath said respective second trenches; whereby said diffusion components reduce depletion spreading in the OFF state.

According to various embodiments, there is provided: a fabrication process for making a MOSFET, comprising the actions, in any order, of: a) providing an n-type semiconductor layer; b) forming a p-type body in said layer; c) forming an n-type source, which is isolated by said body, in said layer; d) forming an insulated gate trench in said layer, and a gate electrode in said gate trench; said gate electrode being capacitively coupled to at least a portion of said body; e) forming a second insulated trench in said layer, providing an additional dose of acceptor dopants below said trench, and forming a Recessed Field Plate electrode in said second trench; and f) providing an additional dose of donor dopant atoms in said portion of said body, to thereby reduce the on-resistance.

According to various embodiments, there is provided: improved highly reliable power RFP structures and fabrication and operation processes. The structure includes plurality of localized dopant concentrated zones beneath the trenches of RFPs, either floating or extending and merging with the body layer of the MOSFET or connecting with the source layer through a region of vertical doped region. This local dopant zone decreases the minority carrier injection efficiency of the body diode of the device and alters the electric field distribution during the body diode reverse recovery.

Modifications and Variations

As will be recognized by those skilled in the art, the innovative concepts described in the present application can be modified and varied over a tremendous range of applications, and accordingly the scope of patented subject matter is not limited by any of the specific exemplary teachings given. It is

intended to embrace all such alternatives, modifications and variations that fall within the spirit and broad scope of the appended claims.

The device may be fabricated in various layouts, including "stripe" and "cellular" layouts. The layers of source, body, and drain regions can be configured vertically, quasi-vertically as well as laterally. The epitaxial drift region can be either uniformly or non-uniformly doped. While the embodiments described above include an epitaxial layer grown on a substrate, the epitaxial layer may be omitted in some applications. Various features of different embodiments may be combined and recombined for various applications.

For example, the region between channel and drain does not have to be uniformly doped, neither vertically nor laterally. The improvements in drift or spreading region doping provided by the disclosed inventions can be combined with a wide variety of other device improvements and features.

For another example, the RFP and gate trenches do not necessarily have to have the same width.

The design could be applied to IGBTs or other devices which include bipolar conduction. The bottom of the gate trench can be modified with dopant; the design can also vary at the source structure and at the drain structure; and alternative body structure may be used; contact trench may be produced first, then cut gate trench, and construct the source and drain structure.

Of course, the n-type dopant, in silicon, can be phosphorus, antimony, or arsenic, or combinations of these. Appropriate donor dopants can be used in other semiconductor materials.

As the disclosed process is scaled to other operating voltages, it is expected that predictable scaling of dimensions and dopants may allow the same synergy. For example, in a 200V embodiment, the inventors contemplate that the trench depth would be slightly deeper (e.g. 1.5 to 2.5 micron), and the compensation implant energy and dose would be about the same. Of course the epi layer doping would be substantially less, and the epi layer thickness greater, as is well understood by those of ordinary skill. The n-enhancement doping (which is preferably blocked from the termination) can have a distribution, after drive-in, which reaches to the upper boundary of the compensation implant, but preferably not to the lower boundary of the compensation implant.

The following applications may contain additional information and alternative modifications: Ser. No. 61/125,892 filed Apr. 29, 2008; Ser. No. 61/058,069 filed Jun. 2, 2008 and entitled "Edge Termination for Devices Containing Permanent Charge"; Ser. No. 61/060,488 filed Jun. 11, 2008 and entitled "MOSFET Switch"; Ser. No. 61/074,162 filed Jun. 20, 2008 and entitled "MOSFET Switch"; Ser. No. 61/076,767 filed Jun. 30, 2008 and entitled "Trench-Gate Power Device"; Ser. No. 61/080,702 filed Jul. 15, 2008 and entitled "A MOSFET Switch"; Ser. No. 61/084,639 filed Jul. 30, 2008 and entitled "Lateral Devices Containing Permanent Charge"; Ser. No. 61/084,642 filed Jul. 30, 2008 and entitled "Silicon on Insulator Devices Containing Permanent Charge"; Ser. No. 61/027,699 filed Feb. 11, 2008 and entitled "Use of Permanent Charge in Trench Sidewalls to Fabricate Un-Gated Current Sources, Gate Current Sources, and Schottky Diodes"; Ser. No. 61/028,790 filed Feb. 14, 2008 and entitled "Trench MOSFET Structure and Fabrication Technique that Uses Implantation Through the Trench Sidewall to Form the Active Body Region and the Source Region"; Ser. No. 61/028,783 filed Feb. 14, 2008 and entitled "Techniques for Introducing and Adjusting the Dopant Distribution in a Trench MOSFET to Obtain Improved Device Characteristics"; Ser. No. 61/091,442 filed Aug. 25, 2008 and entitled "Devices Containing Permanent Charge"; Ser. No. 61/118,664 filed Dec. 1, 2008 and entitled "An Improved Power MOSFET and Its Edge Termination"; and Ser. No. 61/122,794 filed Dec. 16, 2008 and entitled "A Power MOSFET Transistor".

None of the description in the present application should be read as implying that any particular element, step, or function is an essential element which must be included in the claim scope: THE SCOPE OF PATENTED SUBJECT MATTER IS DEFINED ONLY BY THE ALLOWED CLAIMS. Moreover, none of these claims are intended to invoke paragraph six of 35 USC section 112 unless the exact words "means for" are followed by a participle.

The claims as filed are intended to be as comprehensive as possible, and NO subject matter is intentionally relinquished, dedicated, or abandoned.

What is claimed is:

1. A semiconductor device structure, comprising:

a semiconductor layer;

a gate which is positioned in a first trench within said semiconductor layer, and is capacitively coupled to control vertical conduction from a first-conductivity-type source through second-conductivity-type portions of said layer near said trench;

recessed field plates, positioned in proximity to and capacitively coupled to said semiconductor material; said recessed field plates being positioned in respective second trenches;

a first additional diffusion component of a second conductivity type lying at least partially beneath said respective second trenches; and

a second additional diffusion component of said first conductivity type lying at least partially within said second-conductivity-type portions of said layer;

whereby said first additional diffusion component reduces depletion of said second-conductivity-type portions of said layer in the OFF state;

and whereby said second additional diffusion component reduces the on-resistance of the device in the ON state.

2. The semiconductor device structure of claim **1**, wherein said gate has a split poly configuration.

3. The semiconductor device structure of claim **1**, wherein at least one of said recessed field plates has a split poly configuration.

4. The semiconductor device structure of claim **1**, wherein both the gate and at least one of the recessed field plates have a split poly configuration.

5. The semiconductor device structure of claim **1**, wherein said diffusion components have a concentration sufficiently high to locally counterdope said semiconductor layer and thereby produce a second conductivity-type region below said second trench.

6. The semiconductor device structure of claim **1**, wherein said semiconductor layer is an epitaxial layer.

* * * * *

US008466025B2

(12) **United States Patent**
Zeng et al.

(10) **Patent No.: US 8,466,025 B2**
(45) **Date of Patent: Jun. 18, 2013**

(54) **SEMICONDUCTOR DEVICE STRUCTURES AND RELATED PROCESSES**

(75) Inventors: **Jun Zeng**, Torrance, CA (US); **Mohamed N. Darwish**, Campbell, CA (US)

(73) Assignee: **MaxPower Semiconductor, Inc.**, Santa Clara, CA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **13/195,154**

(22) Filed: **Aug. 1, 2011**

(65) **Prior Publication Data**

US 2011/0298043 A1 Dec. 8, 2011

Related U.S. Application Data

(63) Continuation of application No. 12/368,399, filed on Feb. 10, 2009, now Pat. No. 8,076,719.

(60) Provisional application No. 61/065,759, filed on Feb. 14, 2008.

(51) **Int. Cl.**
H01L 21/336 (2006.01)

(52) **U.S. Cl.**
USPC **438/270**; 438/589; 438/197; 438/466

(58) **Field of Classification Search**
USPC 438/270, 197, 466, 589
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,168,331	A	12/1992	Yilmaz	
5,282,018	A	1/1994	Hiraki et al.	
5,525,821	A	6/1996	Harada	
5,637,898	A	6/1997	Baliga	
5,864,159	A	1/1999	Takahashi	
5,973,359	A	10/1999	Kobayashi	
5,998,833	A	12/1999	Baliga	
6,069,372	A	5/2000	Uenishi	
6,114,727	A	9/2000	Ogura et al.	
6,191,447	B1	2/2001	Baliga	
6,251,730	B1	6/2001	Luo	
6,388,286	B1	5/2002	Baliga	
6,468,878	B1 *	10/2002	Petruzzello et al.	438/454

(Continued)

FOREIGN PATENT DOCUMENTS

WO	97/33320	A1	9/1997
WO	2006027739		3/2006

OTHER PUBLICATIONS

J. T. Watt, B. J. Fishbein & J. D. Plummer; Low-Temperature NMOS Technology with Cesium-Implanted Load Devices; IEEE Trans. Electron Devices, vol. 34, # 1, Jan. 1987; p. 28-38.

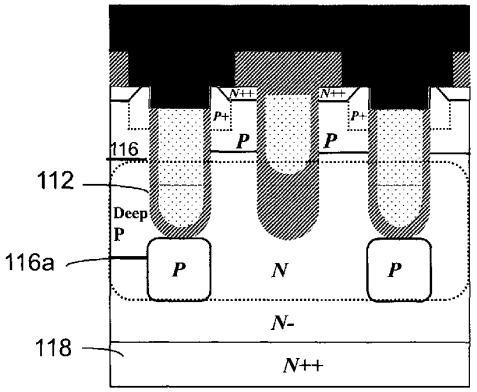
(Continued)

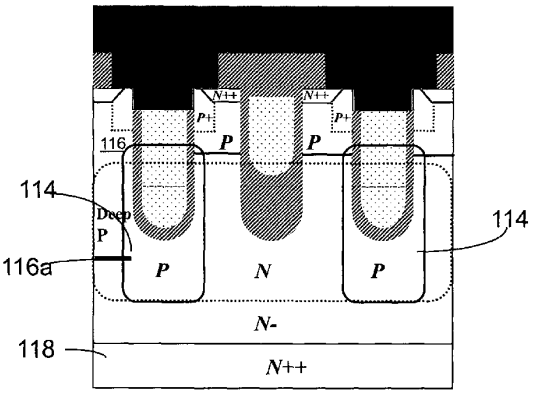
Primary Examiner — Long Pham
(74) *Attorney, Agent, or Firm* — Robert O. Groover, III; Gwendolyn S. S. Groover; Groover & Associates PLLC

(57) **ABSTRACT**

Improved highly reliable power RFP structures and fabrication and operation processes. The structure includes plurality of localized dopant concentrated zones beneath the trenches of RFPs, either floating or extending and merging with the body layer of the MOSFET or connecting with the source layer through a region of vertical doped region. This local dopant zone decreases the minority carrier injection efficiency of the body diode of the device and alters the electric field distribution during the body diode reverse recovery.

7 Claims, 27 Drawing Sheets





U.S. PATENT DOCUMENTS

6,525,373	B1	2/2003	Kim	
6,534,828	B1	3/2003	Kocon	
6,541,820	B1	4/2003	Bol	
6,649,975	B2	11/2003	Baliga	
6,686,244	B2	2/2004	Blanchard	
6,710,403	B2	3/2004	Sapp	
6,803,627	B2	10/2004	Pfirsch	
2001/0001494	A1	5/2001	Kocon	
2001/0041407	A1	11/2001	Brown	
2003/0203576	A1	10/2003	Kitada et al.	
2006/0060916	A1	3/2006	Girdhar et al.	
2007/0004116	A1	1/2007	Hshieh	
2007/0013000	A1	1/2007	Shiraishi	
2008/0099837	A1 *	5/2008	Akiyama et al.	257/341
2009/0309156	A1 *	12/2009	Darwish et al.	257/332

OTHER PUBLICATIONS

J.T.Watt,B.J.Fishbein & J.D.Plummer;Characterization of Surface Mobility in MOS Structures Containing Interfacial Cesium Ions;IEEE Trans.Electron Devices,V36,Jan. 1989; p. 96-100.

J.R.Pfiester, J.R.Alvis & C.D.Gunderson; Gain-Enhanced LDD NMOS Device Using Cesium Implantation; IEEE Trans.Electron Devices, V39, #6, Jun. 1992; p. 1469-1476.

* cited by examiner

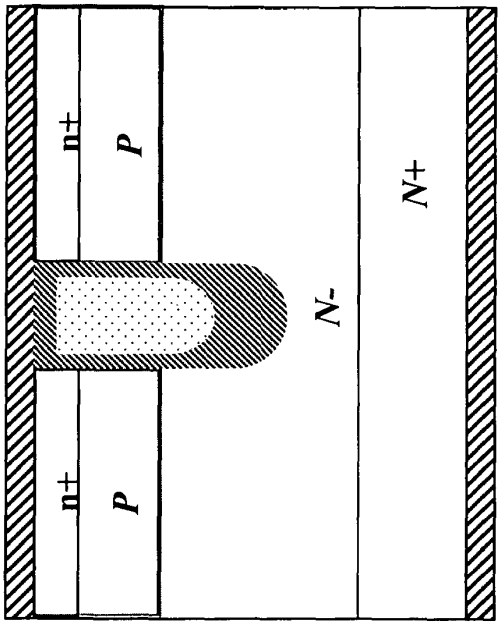


FIG. 1
(Prior Art)

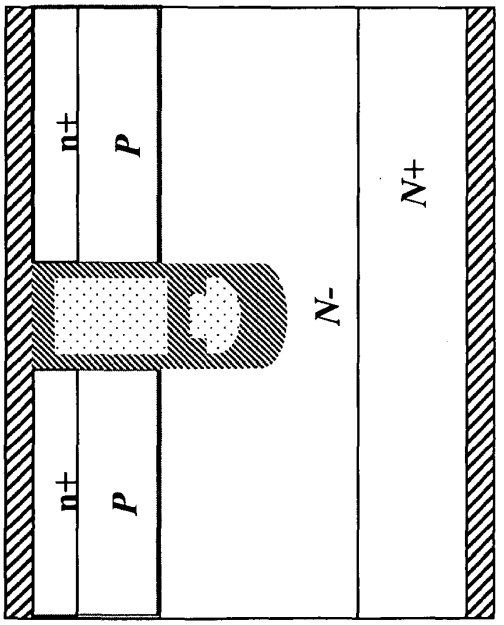


FIG. 2
(Prior Art)

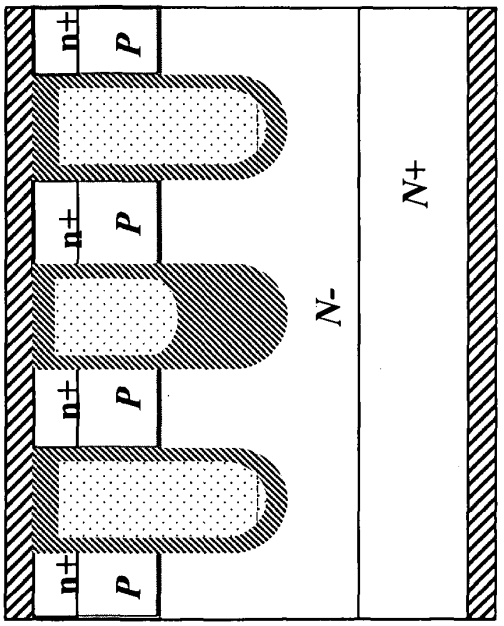


FIG. 3
(Prior Art)

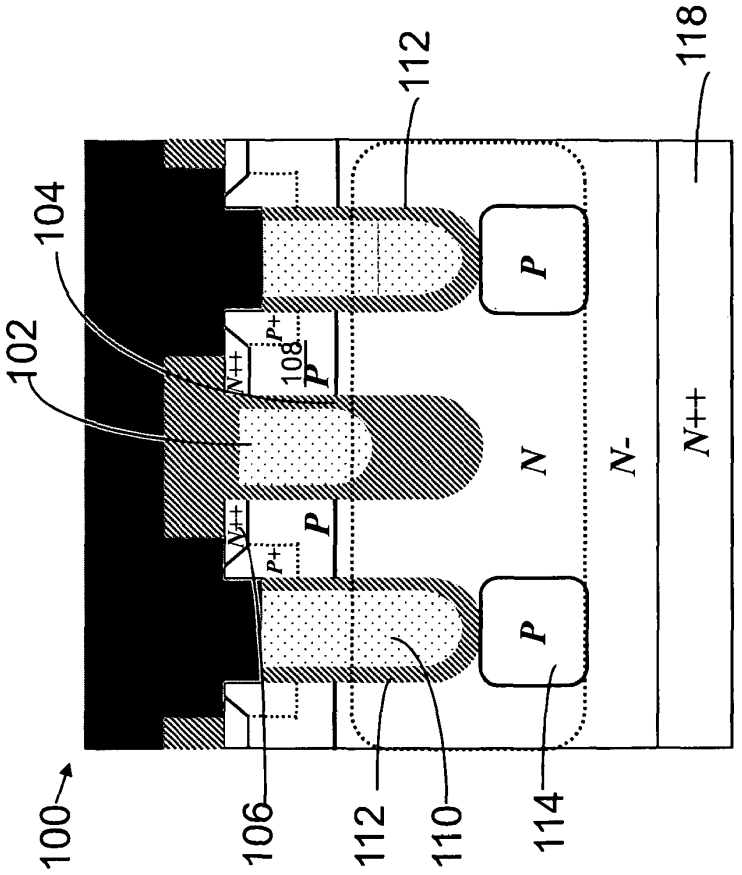


FIG. 4a

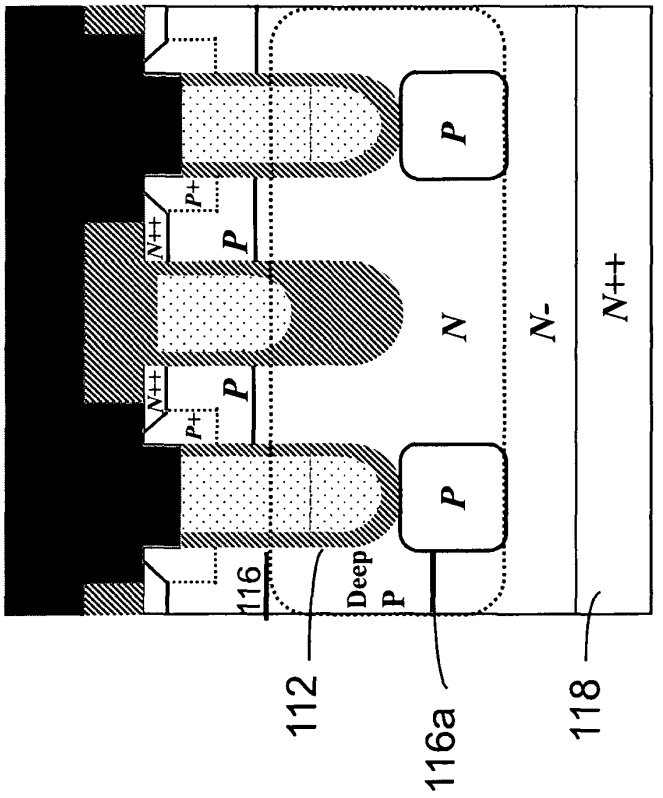


FIG. 4b

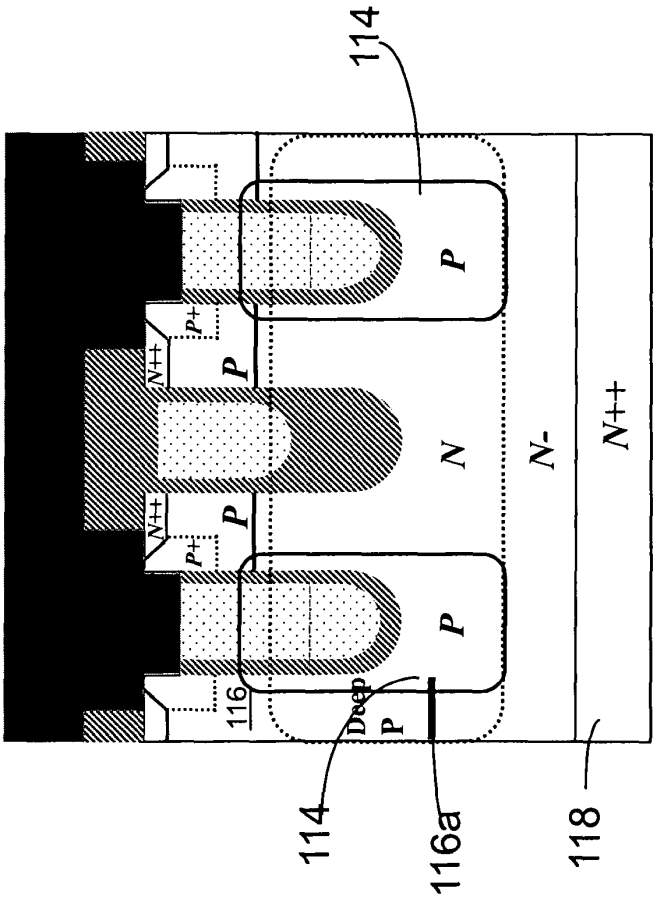


FIG. 4c

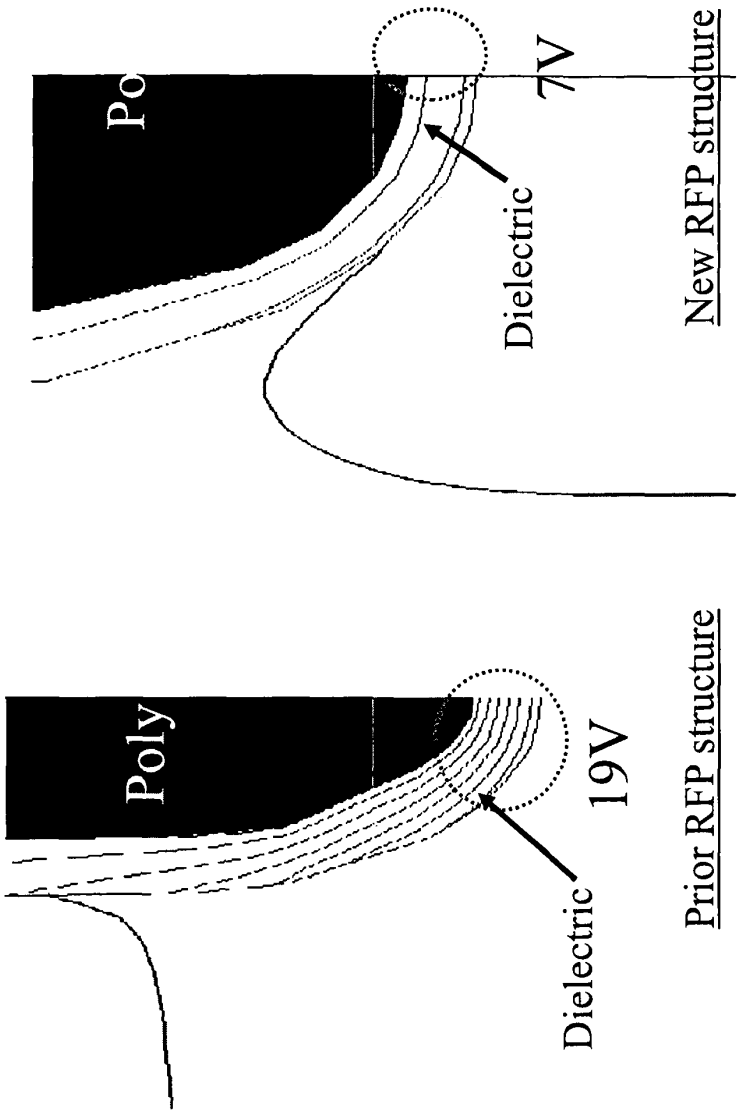


FIG. 5

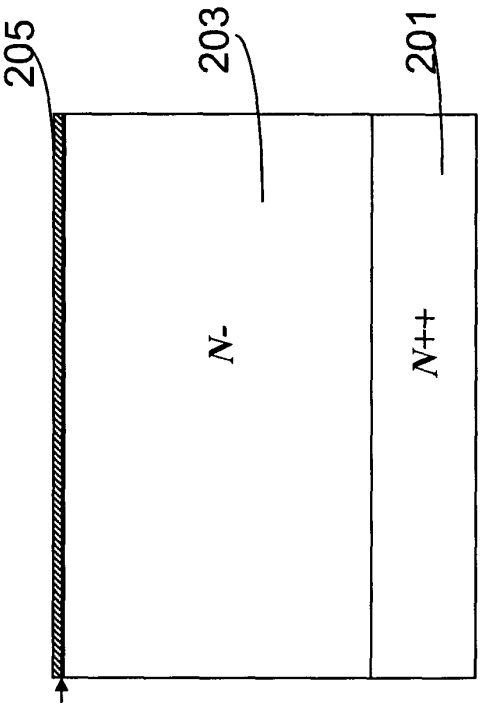


FIG. 6

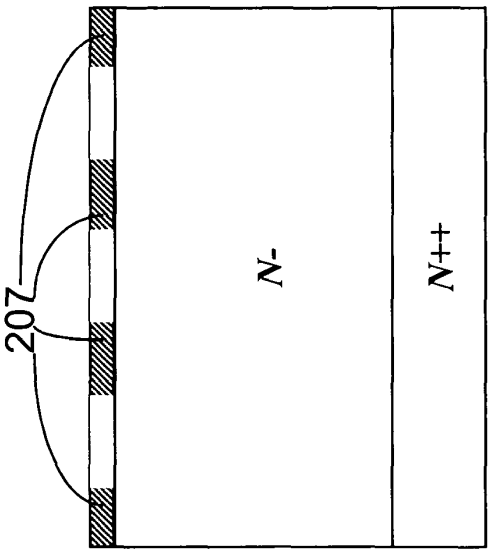


FIG. 7

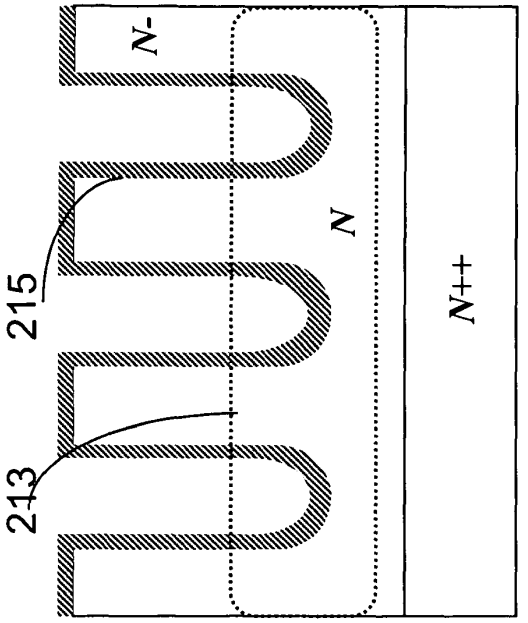


FIG. 9

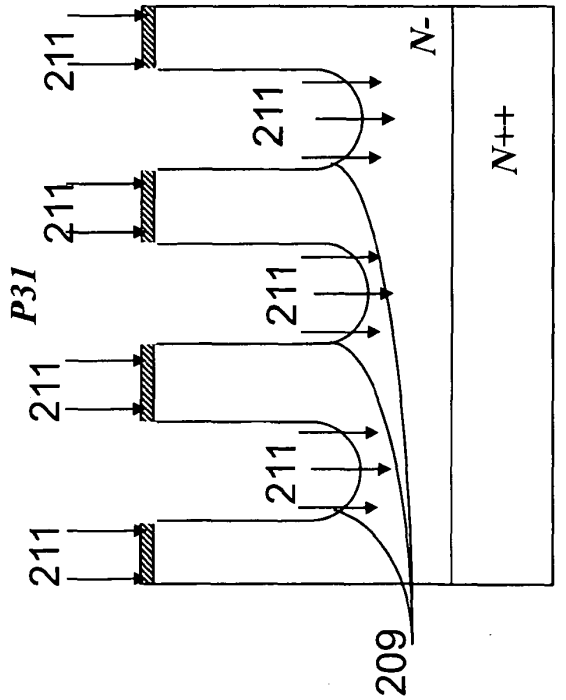


FIG. 8

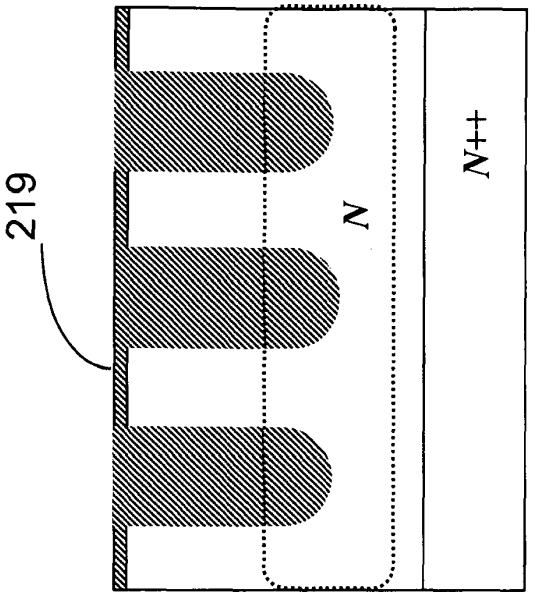


FIG. 11

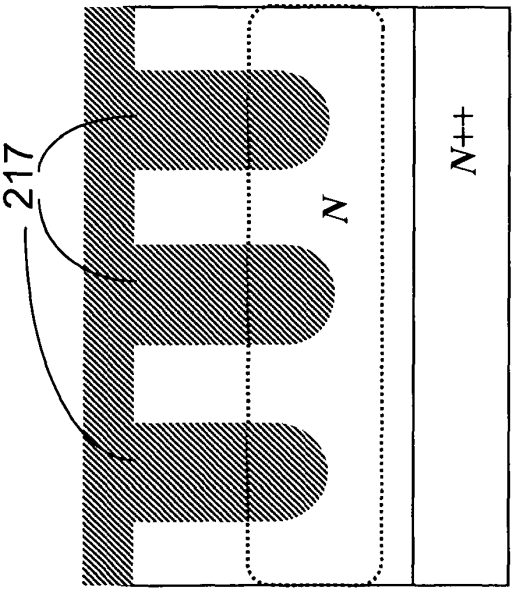


FIG. 10

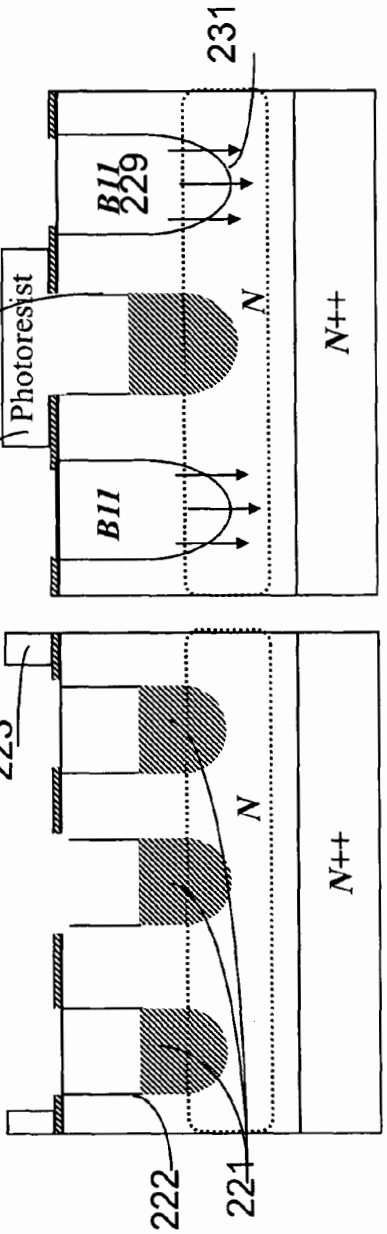


FIG.13

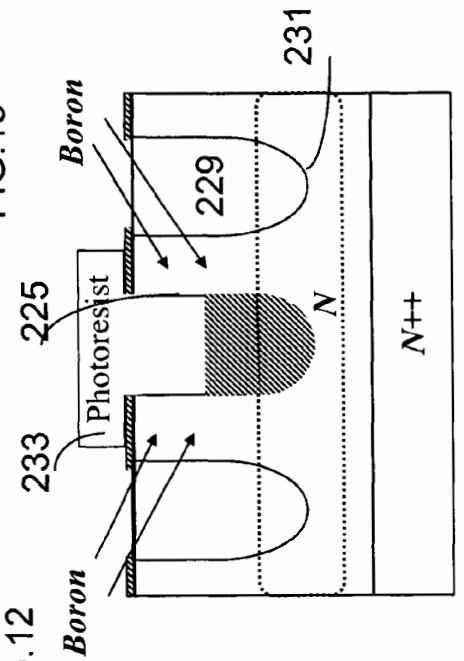


FIG.12

FIG.13+ (optional)

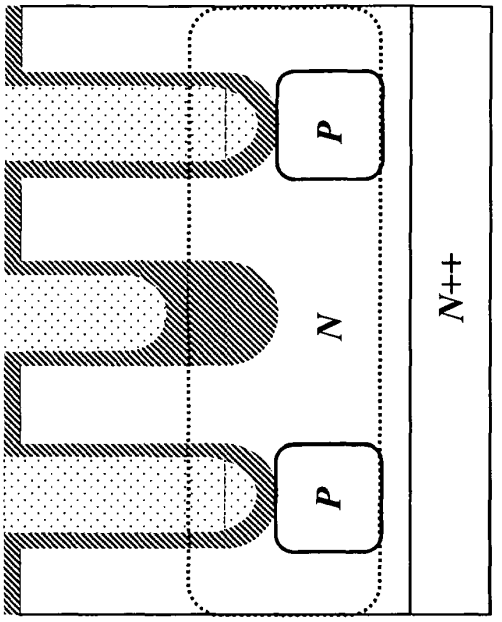


FIG. 15

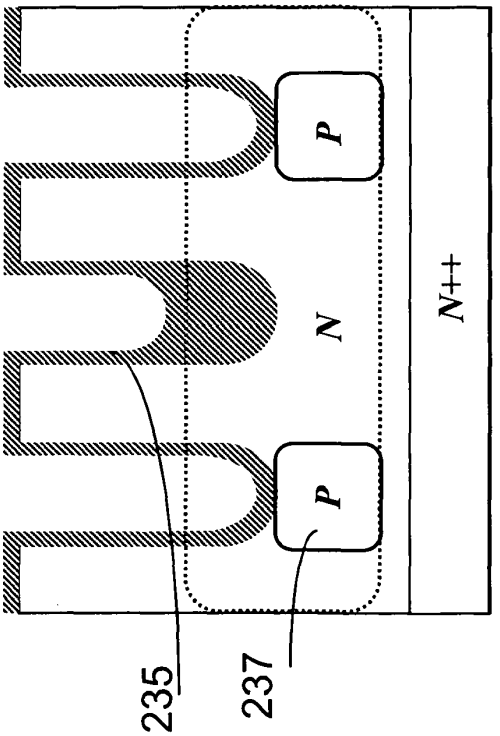


FIG. 14

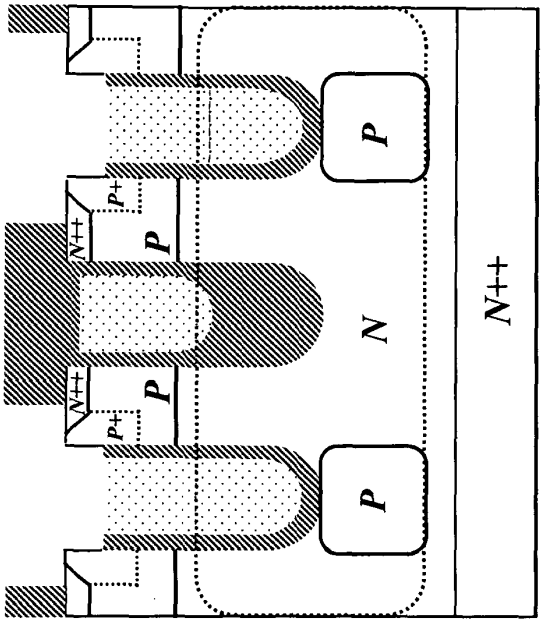


FIG. 17

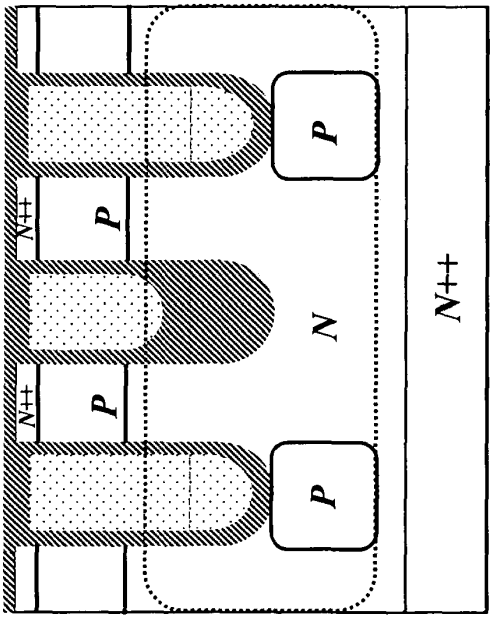


FIG. 16

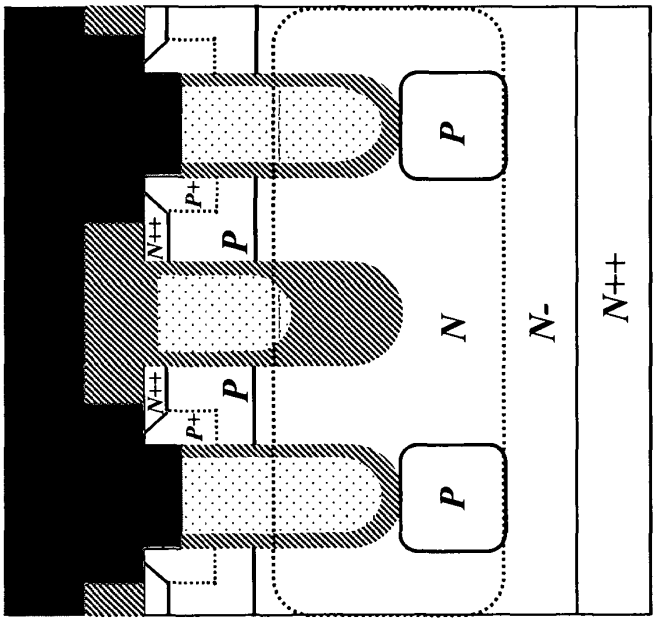


FIG.18

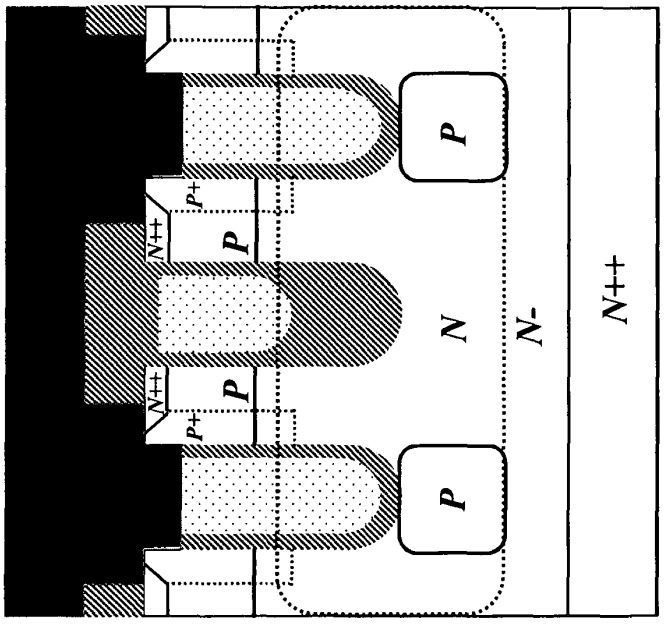


FIG. 19A

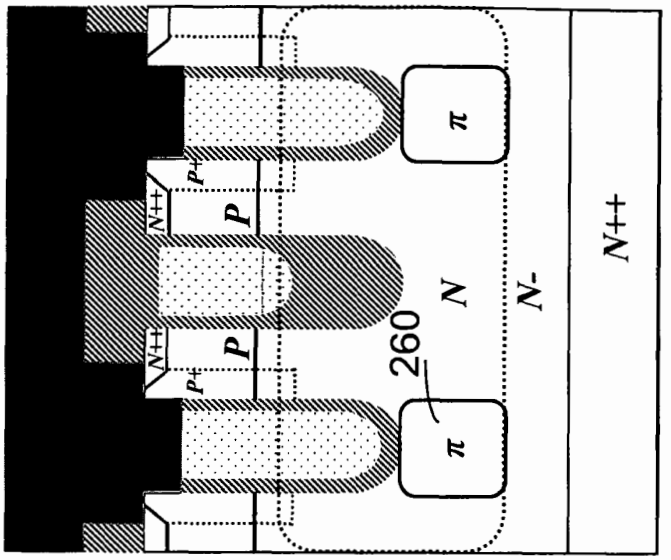


FIG.19C

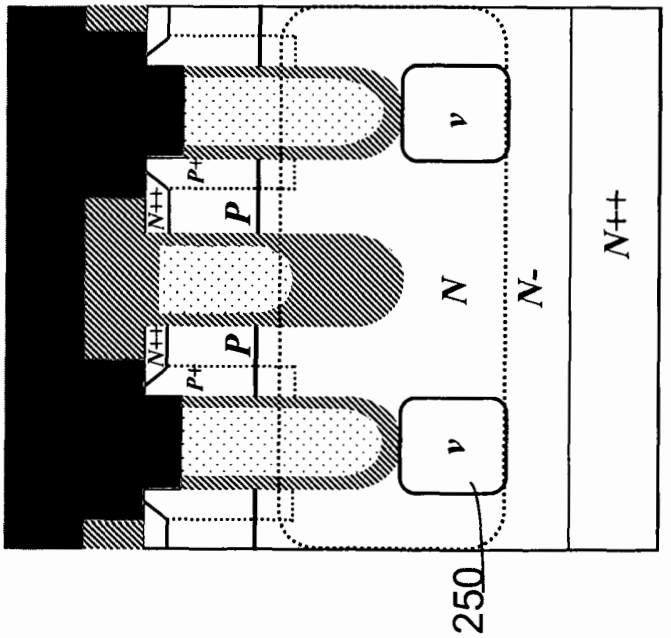


FIG. 19B

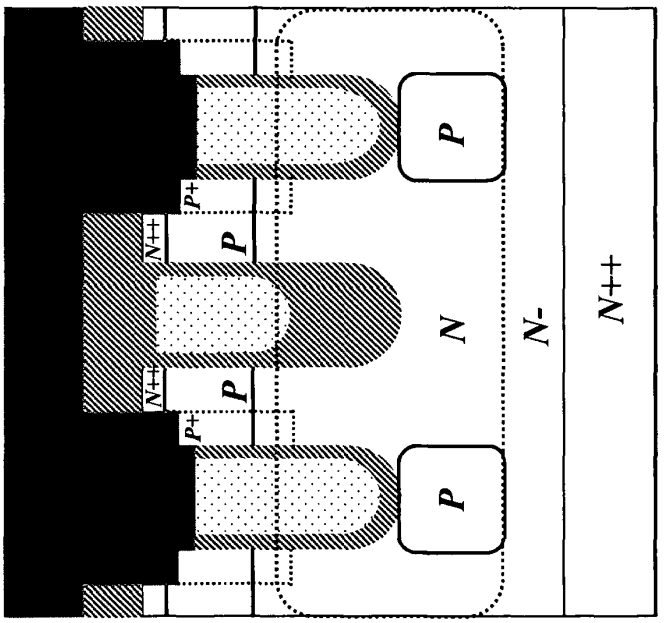


FIG. 20

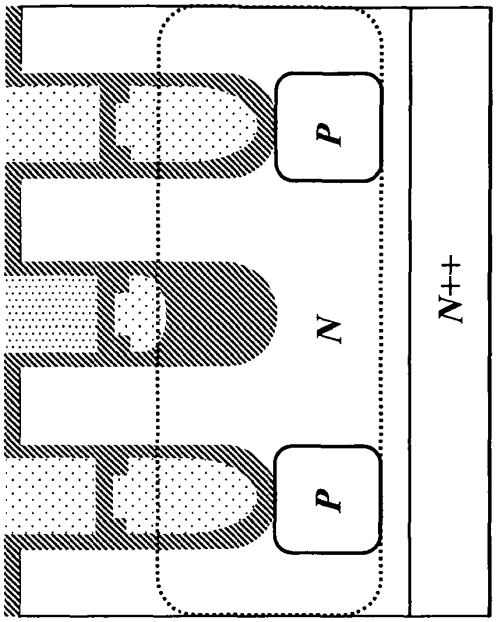


FIG. 22

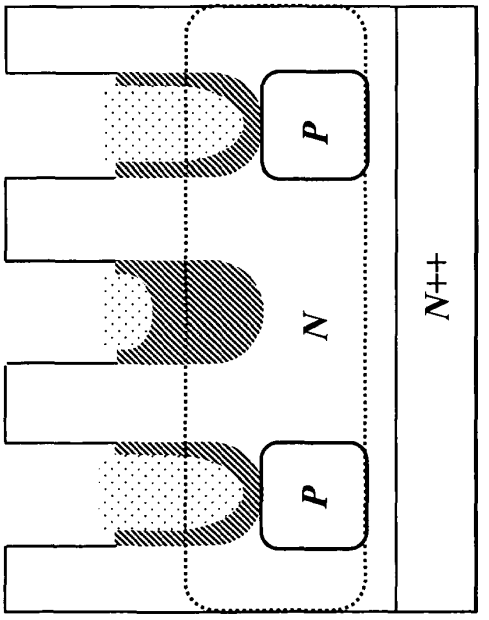


FIG. 21

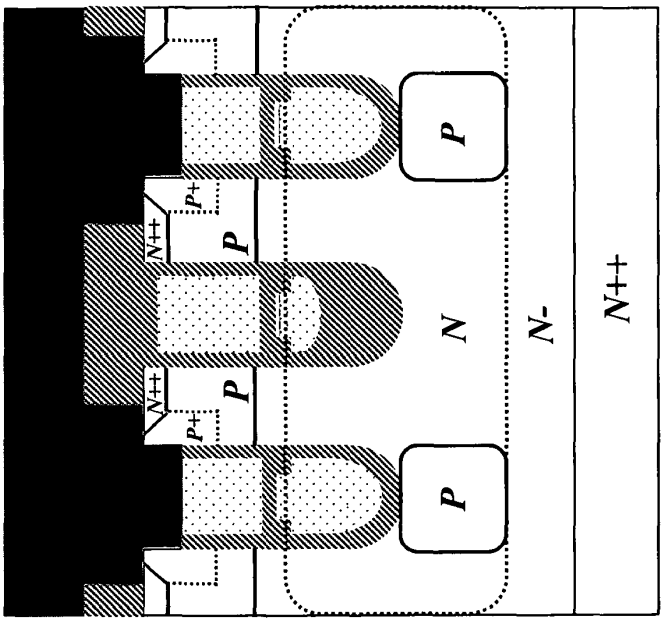


FIG. 23

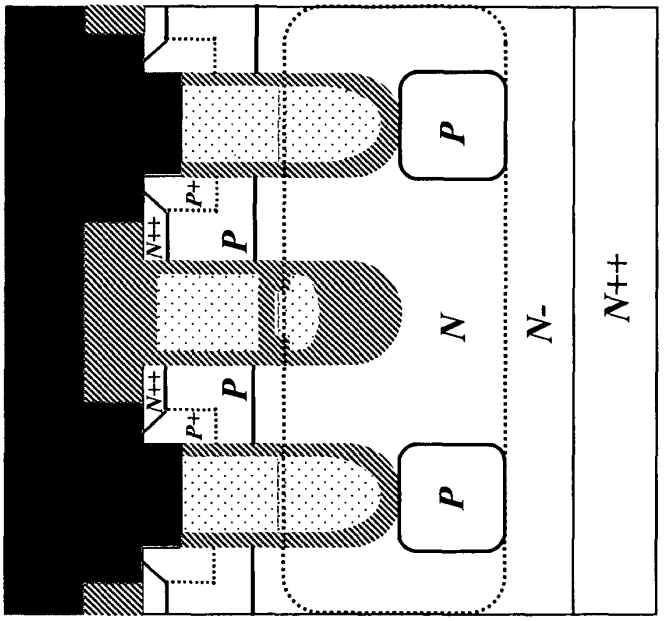


FIG. 24A

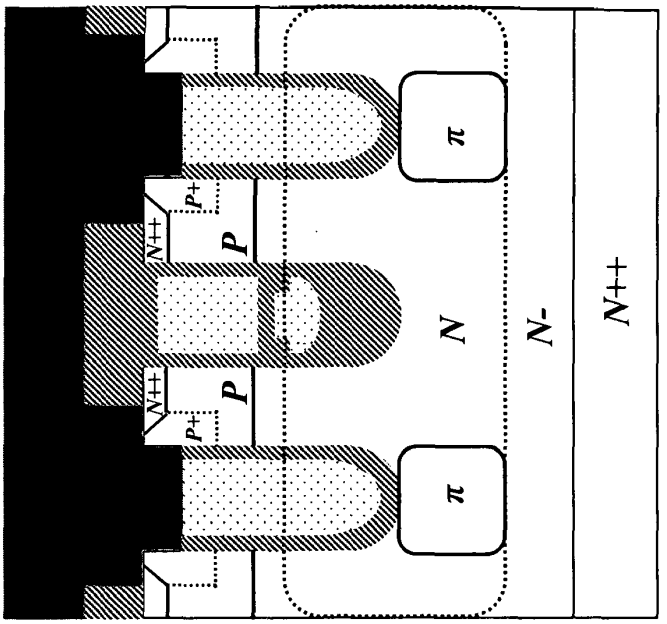


FIG. 24C

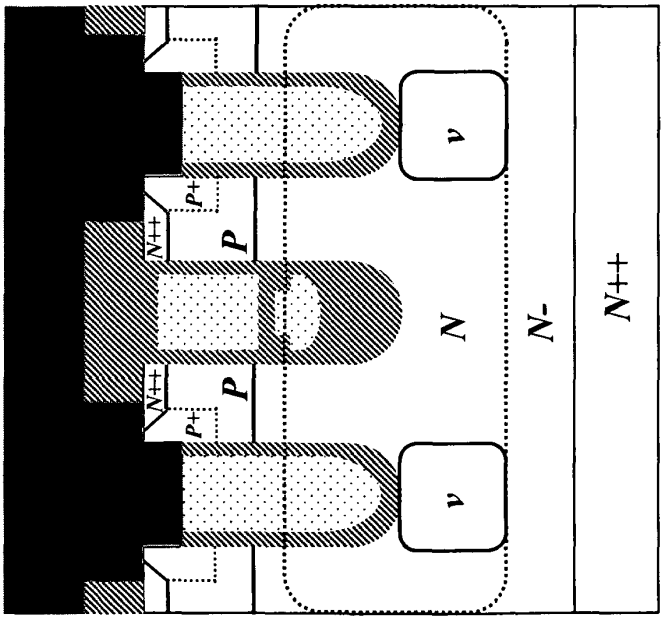


FIG. 24B

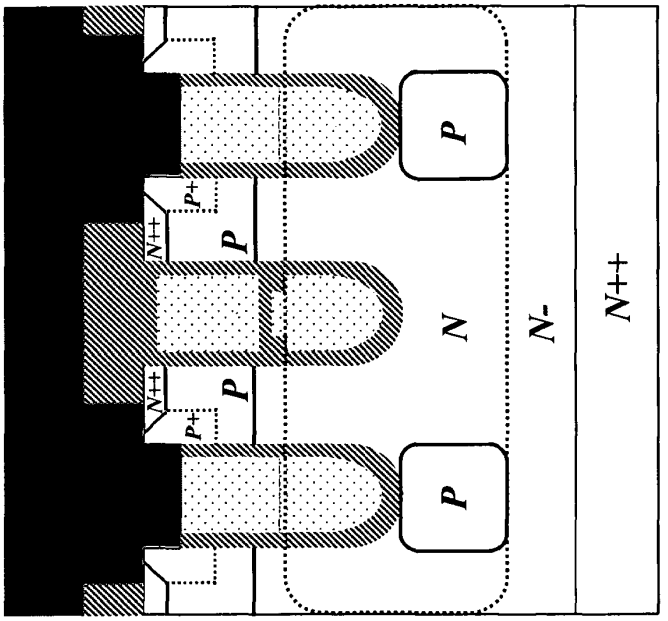


FIG. 25A

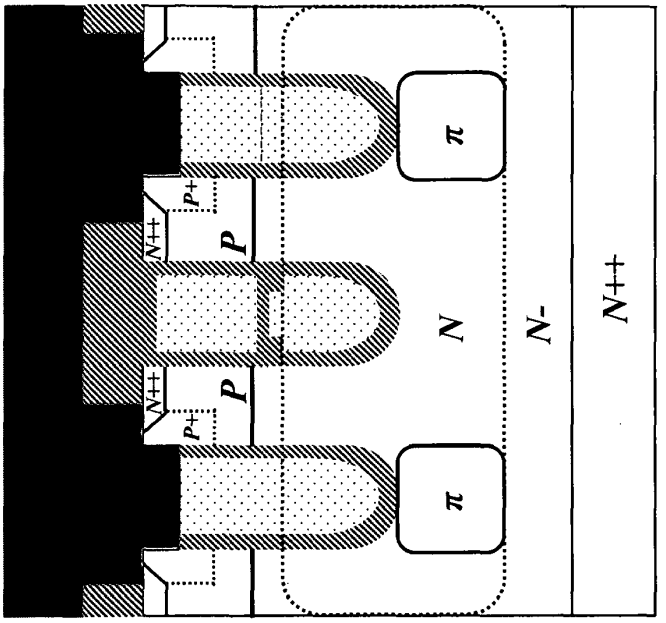


FIG. 25C

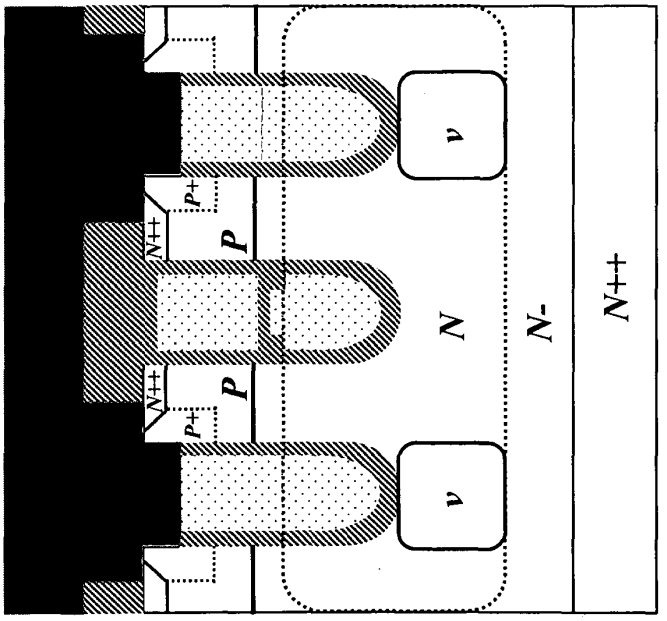


FIG. 25B

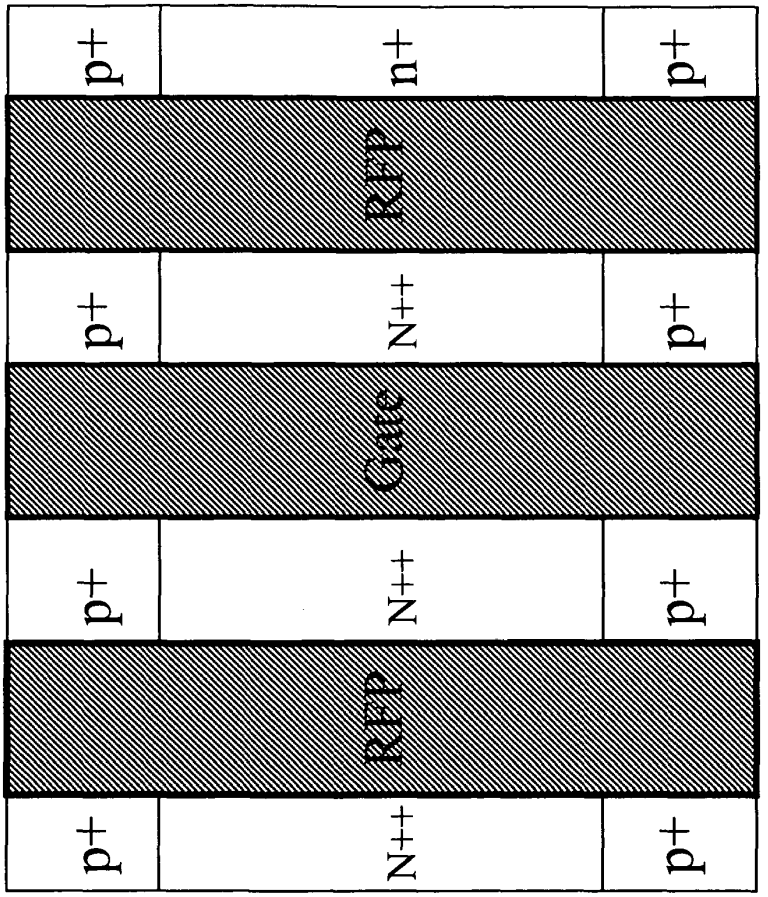


FIG. 26

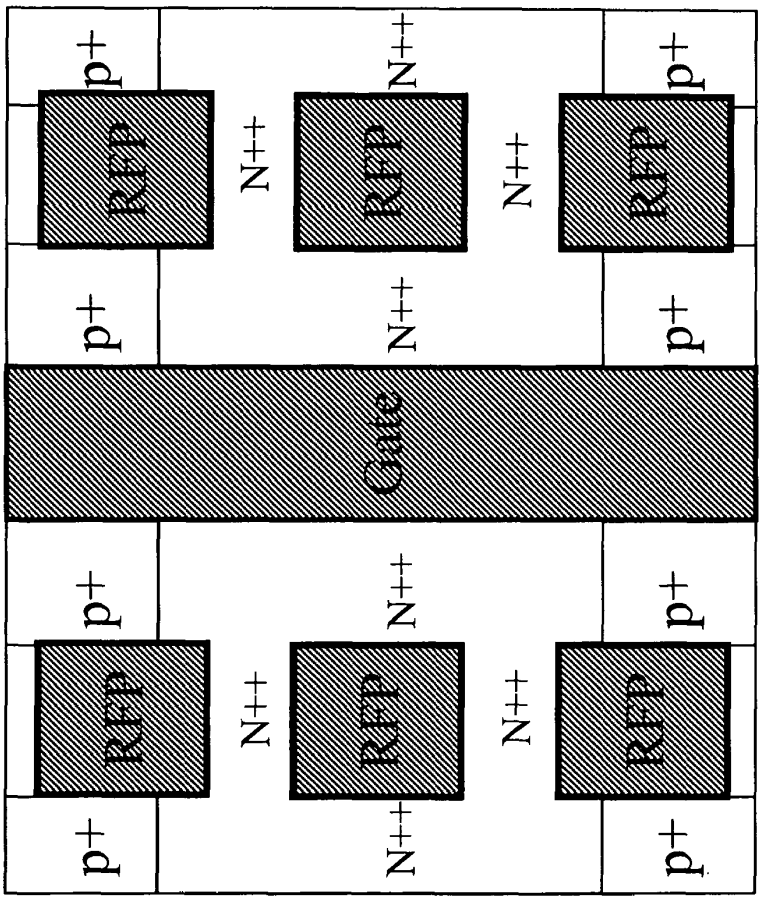


FIG. 27

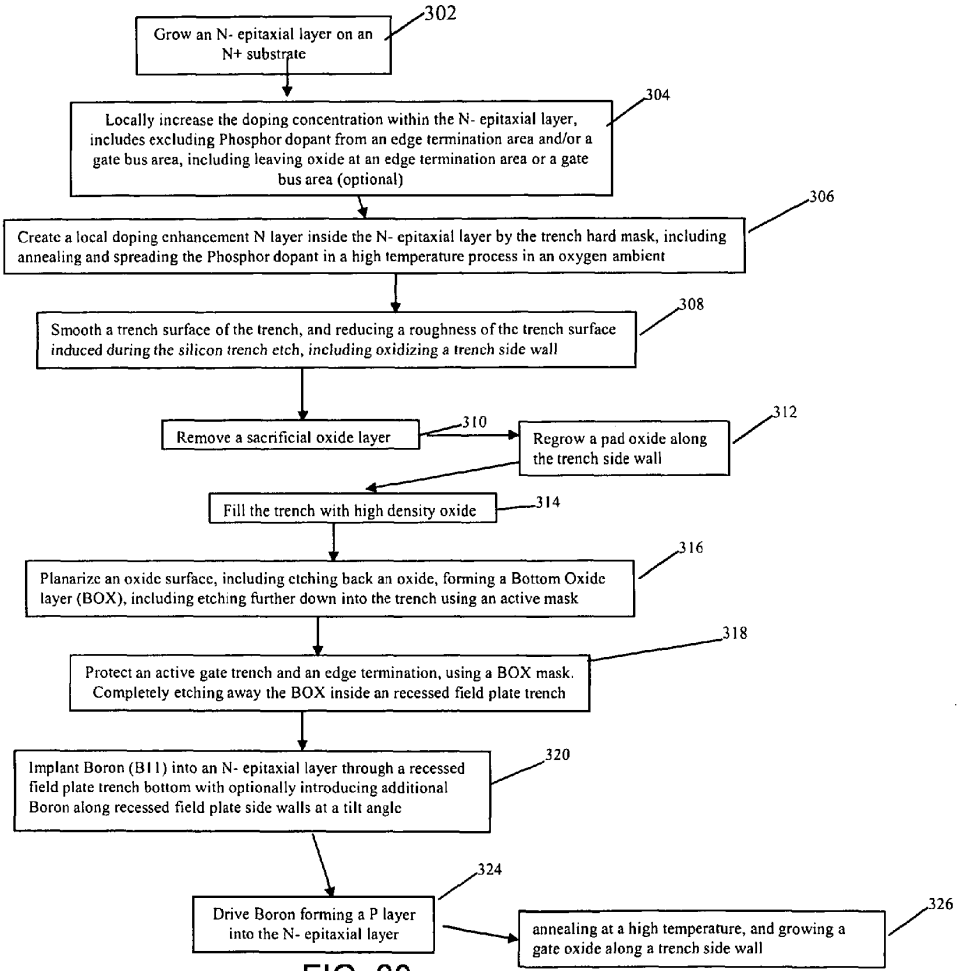

Grow an N- epitaxial layer on an N+ substrate
302
Locally increase the doping concentration within the N- epitaxial layer, includes excluding Phosphor dopant from an edge termination area and/or a gate bus area, including leaving oxide at an edge termination area or a gate bus area (optional)
304
Create a local doping enhancement N layer inside the N- epitaxial layer by the trench hard mask, including annealing and spreading the Phosphor dopant in a high temperature process in an oxygen ambient
306
Smooth a trench surface of the trench, and reducing a roughness of the trench surface induced during the silicon trench etch, including oxidizing a trench side wall
308
Remove a sacrificial oxide layer
310
Regrow a pad oxide along the trench side wall
312
Fill the trench with high density oxide
314
Planarize an oxide surface, including etching back an oxide, forming a Bottom Oxide layer (BOX), including etching further down into the trench using an active mask
316
Protect an active gate trench and an edge termination, using a BOX mask. Completely etching away the BOX inside an recessed field plate trench
318
Implant Boron (B11) into an N- epitaxial layer through a recessed field plate trench bottom with optionally introducing additional Boron along recessed field plate side walls at a tilt angle
320
Drive Boron forming a P layer into the N- epitaxial layer
324
annealing at a high temperature, and growing a gate oxide along a trench side wall
326

FIG. 28

SEMICONDUCTOR DEVICE STRUCTURES AND RELATED PROCESSES

CROSS REFERENCE TO RELATED APPLICATIONS

This application claims priority to U.S. Provisional Application Ser. No. 61/065,759 filed Feb. 14, 2008, which is incorporated by reference herein in its entirety.

BACKGROUND

The present invention relates to field effect transistors and methods, and more particularly to highly reliable power insulated-gate field effect transistors (MOSFET) with a Recessed Field Plate (RFP) and related techniques.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** is a cross-sectional view of a prior art MOSFET with a trench gate having a thick bottom oxide structure.

FIG. **2** is a cross-sectional view of a prior art MOSFET with a split poly gate structure.

FIG. **3** is a cross-sectional view of a prior art MOSFET with RFPs in parallel with the gate trench.

FIG. **4**(*a*) is a cross-sectional view of an RFP containing MOSFET structure with a floating Deep compensated zone.

FIG. **4**(*b*) is a cross-sectional view of an RFP containing MOSFET structure with a Deep compensated zone that extends to and is connected to the source electrode.

FIG. **4**(*c*) is a cross-sectional view of an RFP containing MOSFET structure with a Deep compensated zone that extends to the P body region.

FIG. **5** shows a two-dimensional electrical voltage simulation comparison between the prior RFP-MOSFET structure and a MOSFET containing a Deep compensated zone.

FIGS. **6-18** show successive steps in a sample process for making the sample structure depicted in FIG. **4**(*a*).

FIG. **19**A is a cross-sectional view of a RFP containing MOSFET structure with a Deep compensated zone and a P+ implant region extending beyond the P-N junction and into the N drift region.

FIG. **19**B is a cross-sectional view of an RFP containing MOSFET structure with a Deep compensated zone that is a lightly doped p region, and a P+ implant region extending beyond the P-N junction and into the N drift region.

FIG. **19**C is a cross-sectional view of an RFP containing MOSFET structure with a Deep compensated zone that is a lightly doped n region, and a P+ implant region extending beyond P-N junction and into the N drift region.

FIG. **20** is a cross-sectional view of an RFP containing MOSFET structure with a Deep compensated zone, and a P+ implant region extending beyond P-N junction and into the N drift region and the N++ source region that is completely recessed down.

FIGS. **21-23** show a process for fabrication of an embodiment of FIG. **4**(*a*) structure implemented in the split poly gated structure with split poly layer structure in the RFP trench.

FIGS. **24**A, **24**B, **24**C, **25**A, **25**B and **25**C show cross-sectional views of RFP containing MOSFET structures with a Deep compensated zone (lightly doped p regions in **24**B and **25**B and lightly doped n regions in **24**C and **25**C), implemented in the split poly gated structure with a single poly layer structure in the RFP trench.

FIG. **26** shows a top view of an embodiment of FIG. **4**(*a*) wherein the RFP region is a continuous strip in the horizontal direction.

FIG. **27** shows a top view of an embodiment of FIG. **4**(*a*) wherein the RFP region is divided into several columns in the horizontal direction.

FIG. **28** shows a schematic flow chart for a sample fabrication process.

DETAILED DESCRIPTION OF SAMPLE EMBODIMENTS

Power MOSFETs are widely used as switching devices in many electronic applications. In order to minimize conduction power loss, it is desirable that MOSFETs have low specific on-resistance, which is defined as the on-resistance area product (R_{on}*A), where R_{on} is a MOSFET resistance when the MOSFET is in an ON state, where A is the area of the MOSFET. Trench MOSFETs provide low specific on-resistance, particularly in the 10-100 voltage range. As cell density increases, any associated capacitances such as a gate-to-source capacitance Cgs, a gate-to-drain capacitance Cgd and/or a drain-to-source capacitance Cds also increase. In many switching applications such as synchronous buck dc-dc converters in mobile products, MOSFETs with breakdown voltages of 30 V often operate at higher speeds approaching 1 MHz. Therefore, it may be desirable to minimize switching or dynamic power loss caused by these capacitances. The magnitude of these capacitances are directly proportional to gate charge Qg, gate-drain charge Qgd and output charge Qoss. Furthermore, for a device that operates in the third quadrant (i.e., when a drain-body junction becomes forward biased), minority charge is stored in the device during its forward conduction. This stored charge causes a delay in switching from conducting to non-conducting. To overcome this delay, a body diode with fast reverse recovery is desirable. However, a fast recovery body diode often causes high electromagnetic interference (EMI). This means that during diode recovery, the ratio between the negative going waveform (t_a) and the positive going waveform (t_b) must be less than one for a soft recovery which avoids EMI problem.

As switching speed requirements increase to 1 MHz and beyond with new applications, state of the art power MOSFETs are increasingly unable to operate at such high speeds with satisfactory efficiency. A power MOS transistor that has low charges Qg, Qgd, Qoss and Qrr in addition to having a low specific on-resistance (Ron*A), is desirable.

There are two common techniques to improve the switching performance of power MOSFETs. The first one is the trench-gated MOSFET with thick bottom oxide, as shown in FIG. **1** (U.S. Pat. No. 6,849,898). The second one is the split poly gated MOSFET structure, in which the first poly gate is electrically shorted to the source electrode (U.S. Pat. Nos. 5,998,833, 6,683,346), as illustrated in FIG. **2**.

Recently, as shown in FIG. **3**, US Patent Application No. 2008/0073707 A1 to Darwish disclosed a power MOSFET with the recessed field plate (RFP) structure which realize a very short channel region (~0.25 um) for further reducing the gate-source capacitance and the gate-drain capacitance, and consequently, the total gate charge (Qg) and the "Miller" charge (Qgd). The RFP structure additionally improves the body diode reverse recovery speed due to providing an additional path for current and the enhanced depletion of the drift region induced by the RFP.

The present application discloses improvements to power insulated-gate field effect transistors with Recessed Field Plate (RFP) and similar structures. The inventors have real-

ized that the performance of RFP-type power MOSFETs can be improved by performing a compensating implant into the RFP trench. This compensating implant helps to shape the depletion boundaries in the OFF state, and thus helps to avoid punchthrough. Because of this, a local enhancement can also be added to the doping between channel and drain, in the drift or spreading region. This provides a synergistic combination, wherein the on-resistance can be improved with no degradation in breakdown voltage.

The disclosed innovations, in various embodiments, provide one or more of at least the following advantages. However, not all of these advantages result from every one of the innovations disclosed, and this list of advantages does not limit the various claimed inventions.

- Improved (reduced) on-resistance;
- Improved (increased) breakdown;
- Reduced electrical stress on any dielectric layer at the bottom of the RFP trench;
- Higher reliability and longer operation life; and/or
- Increased ability to increase local doping concentration in the drift region.

The numerous innovative teachings of the present application will be described with particular reference to presently preferred embodiments (by way of example, and not of limitation). The present application describes several embodiments, and none of the statements below should be taken as limiting the claims generally.

For simplicity and clarity of illustration, the drawing figures illustrate the general manner of construction, and description and details of well-known features and techniques may be omitted to avoid unnecessarily obscuring the invention. Additionally, elements in the drawing figures are not necessarily drawn to scale, some areas or elements may be expanded to help improve understanding of embodiments of the invention.

The terms "first," "second," "third," "fourth," and the like in the description and the claims, if any, may be used for distinguishing between similar elements and not necessarily for describing a particular sequential or chronological order. It is to be understood that the terms so used are interchangeable. Furthermore, the terms "comprise," "include," "have," and any variations thereof, are intended to cover non-exclusive inclusions, such that a process, method, article, apparatus, or composition that comprises a list of elements is not necessarily limited to those elements, but may include other elements not expressly listed or inherent to such process, method, article, apparatus, or composition.

It is contemplated and intended that the design apply to both n-type and p-type MOSFETs; for clarity reason, the examples are given based on n-channel MOSFET structure, but an ordinary person in the art would know the variations to modify the design to make a similar p-channel device.

The present application discloses improvements to power insulated-gate field effect transistors with Recessed Field Plate (RFP) and similar structures. The inventors have realized that the performance of RFP-type power MOSFETs can be improved by performing a compensating implant into the RFP trench. This compensating implant helps to shape the depletion boundaries in the OFF state, and thus helps to avoid punchthrough. Because of this, a local enhancement can also be added to the doping between channel and drain, in the drift or spreading region. This provides a synergistic combination, wherein the on-resistance can be improved with no degradation in breakdown voltage.

In one sample embodiment, the RFP containing MOSFET has a buried Deep compensated zone floating in the N body region underneath the RFP trench. The Deep compensated zone reduces the voltage across the dielectric layer between the RFP and the N^- epitaxial layer when a high drain-source voltage is applied.

In one sample embodiment, the RFP containing MOSFET has a buried Deep compensated zone floating in the N body region underneath the RFP trench and a local enhancement to the doping between channel and drain, in the drift or spreading region in the N epitaxial layer

In one embodiment, the RFP containing MOSFET also has a deep P+ region in the P body in contact with the RFP trench wall that extends from the P body into the N epitaxial layer.

In one embodiment, the Deep compensated zone underneath the RFP trench extends to and is connected to the source electrode.

In one embodiment, the Deep compensated zone underneath the RFP trench extends to P body region that is in contact with the side walls of the RFP trench.

In one embodiment, the Deep compensated zone is a very lightly doped p region; while in another embodiment, the Deep compensated zone is a very lightly doped n region.

Referring now to FIG. **4**(*a*), a semiconductor device structure **100** comprises a gate **102** which is positioned in a first trench **104**, also described herein as a gate trench **104**. The first trench **104** containing the gate **102** may be one of many gate trenches within the semiconductor device structure **100**. The semiconductor device structure **100** is capacitively coupled to control vertical conduction from a source region **106**, having a first-conductivity-type, through semiconductor material **108**, which is adjacent to the first trench **104**.

As shown in FIG. **4**(*a*), the gate **102** has a gate electrode comprising conductive gate material having a width approximately equal to a width of the gate trench **104**. It will be appreciated that, although the gate electrode may have a width that is approximately equal to the width of the gate trench **104**, the gate electrode may alternatively be contacted using a wider gate trench and a smaller gate electrode, allowing the gate trench to be insulated from the gate conductor.

The semiconductor device structure **100** also includes recessed field plates **110**, which are positioned in proximity to and capacitively coupled to the semiconductor material **108**. The recessed field plates **110** are positioned in respective second trenches **112**, also described herein as RFP trenches **112**. Each of the trenches (i.e., the respective second trenches **112** and the gate trenches) has trench walls that are coated with an insulating material, such as silicon dioxide (SiO_2). The RFP trenches **112** contain an insulating material having a breakdown voltage which preferably exceeds the breakdown voltage of the semiconductor device structure **100**. The gate trench **104** preferably contains an insulating material up to the p-body drain junction, minimizing any overlap of the gate electrode (connected to the gate **102**) with the drain or drift region.

In one embodiment, the gate trench contains thick bottomed insulation dielectric material such as silicon dioxide. In another embodiment, the insulating material within the RFP trench and/or the gate trench **104** has a stepped thickness. Providing a stepped thickness can help shape the channel and can help control "hot" electron effects.

Conductive material, such as n-type doped polysilicon, forms a gate electrode electrically separated from the gate trench **104** by the insulating material. The conductive material may be silicided to reduce its resistance. Conductive material also fills the RFP trenches **112**, electrically separated from the gate trench **104** by the insulating material, and extending above the RFP trenches to form a plurality of RFP electrodes. Each of the trenches may be of substantially equal depth, or may differ in depth, and may be self-aligned by

being etched at the same processing step, although the RFP electrode is deeper than the gate electrode and is either independently biased or connected to a source electrode (i.e., the source **106**), and a source region (including the source electrode) may extend between the gate **106** and the RFP trenches **112**.

In one embodiment, the n-epitaxial drift region is uniformly doped. In another embodiment, the n-epitaxial drift region is not uniformly doped. Specifically, the doping is graded to have a doping concentration that is higher at an interface with the underlying **118** substrate and decrease toward the surface. Non-uniform doping of the drain drift region allows for greater shaping of the channel and for control over "hot" electron injection.

The source region may be doped n+. The gate trench **104** and the RFP trench may have a thin layer of the insulating material, reducing on-resistance, or a thick layer of the insulating material, providing greater electrical isolation increasing reverse-bias breakdown voltage. In the embodiment depicted, the RFP electrodes have a uniform depth. In another embodiment, at least one of the RFP electrodes extends up and contacts the source **106**.

Advantageously, the semiconductor device structure **100** also includes deep compensated zones **114** of either p-type or n-type lying at least partially beneath the respective RFP trenches **112**. The deep compensated zone **114** may be floating islands of either p-type dopant concentration regions (as shown in FIG. **4***a*), or lightly doped n-type dopant concentration regions in the N-drift region underneath the RFP trenches. The drawing shows the boundaries of this compensated zone **114** as if it had been fully counterdoped, but those of ordinary skill will understand that the boundaries of a compensated but not counterdoped zone can be similarly envisioned, using e.g. the concentration contours of a single dopant species.

The deep compensated zones **114** also reduce the voltage across the dielectric layer between the RFP and the N− epitaxial layer when a high drain-source voltage is applied.

As shown in FIG. **4**(*b*), device **100** also contains a deep p-body region **116** that is in contact with the side walls of the RFP trench **112**. The deep p-body region **116** with boundary at **116***a* can be in connection with the source electrode and also be in connection with the deep compensated zones **114**. The deep P-N junction in the edge termination can be formed by the Deep compensated implant and its related annealing without adding new mask. Therefore, the disclosed structure can offer a more reliable edge termination.

Alternatively, as shown in FIG. **4**(*c*), the deep compensated zones **114** can extend vertically and merge with the p-body region.

The two-dimensional electric voltage simulation shown in FIG. **5**, reveals that under the same bias conditions the conventional device of structure shown in FIG. **3** sees about 19V across the bottom dielectric layer between RFP and the N− epitaxial layer, while an embodiment of FIG. **4**(*a*)-(*c*) only shows 7V across the bottom dielectric layer between RFP and the N− epitaxial layer due to the protection from the Deep compensated zone **114**.

As the electrical stress on the bottom dielectric layer between RFP and drain is reduced significantly, the device structures of FIG. **4**(*a*)-(*c*) will offer a higher reliability and longer operation life. In addition, the Deep compensated zones **114** enhance the lateral and vertical depletion of N− epitaxial layer, which provides room for higher local doping concentration in the epitaxial layer without degrading the device breakdown voltage.

The increase in local doping concentration in the epitaxial layer further reduces the on-resistance of drift region. By properly adjusting the doping concentration of P and N regions in the N− epitaxial layer, the total on-resistance of device can be lowered without reduction of the breakdown voltage. Furthermore, the local doping enhanced N layer also decreases the minority carrier injection efficiency of the body diode of the device and alters the electric field distribution during the body diode reverse recovery. Thus, the reverse recovery of the body diode is improved, resulting in a device having lower reverse recovery charge and soft recovery features.

Since the doping enhancement only occurs in the active region, the termination efficiency of the improved device edge junction termination region will not be degraded.

The recessed field plates **110** may be positioned in multiple respective trenches **112**, which are separate from the gate trench **104**. Accordingly, the semiconductor device structure **100** may be, for example, an n-channel MOSFET with a recessed field plate (RFP) trench **112** and a gate trench **104** formed on an N-type epitaxial layer grown over a heavily doped N+ substrate.

In third quadrant operation, in which where the drain **118** is negatively biased with respect to a source-body electrode (i.e. the source **106**), and in which diffusion current results in minority carrier injection and a high reverse recovery charge Qrr, the plurality of RFP electrodes form majority carrier channel current path from drain to source in addition to that provided by the gate electrode in a conventional structure. The combined effect of the RFP electrodes and the gate electrode is both reduced minority carrier diffusion current and reduced recovery charge Qrr. Accordingly, in the third quadrant operation, the RFP electrodes act as an additional gate without any penalty of an added gate-drain capacitance Cgd.

In reverse-biased operation, the RFP also reduces any electric field in a channel region. Accordingly, shorter channel lengths are possible, without substantial risk of punch-through breakdown, further allowing reduction in Ron*A and Qg. The capacitive coupling between the gate trench **104**, the RFP trenches **112** and the drain region further deplete the drain drift region, at a higher rate as a drain-source voltage VDS is increased in an off-state. The low Cgd and its fast falling rate, combined with increasing drain-source voltage VDS, provides a lower gate-drain charge.

The semiconductor device structure **100** can have a quasi-vertical or lateral configuration. Ensuring that the semiconductor device structure **100** has a quasi-vertical or lateral configuration can help shape the channel, and can reduce hot electron effects.

Various variations in gate conductor and RFP conductors may be used. Various combinations have been shown in US Application No. 2008/0073707 A1 to Darwish, the entirety of which is thereby incorporated by reference. Polysilicon may be used as the conductive material. Example variations in structural designs of gate conductor and RFP conductors include split poly configurations and single poly configurations (FIGS. **21**-**25**), thick bottom oxide, and step shaped bottom oxide and the combinations of the various forms.

Referring to FIG. **26**, each of the foregoing embodiments may be implemented in a single configuration, a multi-stripe configuration, a cellular layout configuration, or a combination of the foregoing. Moreover, the polarity and conductivity type may be reversed.

Referring to FIG. **27**, each of the foregoing embodiment RFP may also be implemented in an interrupted manner where the RFP trenches and conductors form columns in the

source-body-drain layers of the device. With this interrupted scheme, more N++ surface area can be provided, reducing the N++ resistance, and lowering the total on-resistance.

A fabrication process for making the described embodiment is detailed in FIG. **6-18**. In FIG. **6**, starting with N++ substrate **201**, the N-epitaxial layer **203** is grown followed by forming a thin layer of silicon oxide layer **205**. Substrate **201** may have been doped with phosphorus or arsenic. The preferred thickness for oxide layer **205**, for example, can be 200-300 Å. In FIG. **7**, the trench mask **207** is applied to form the hard mask for trench etching and the oxide layer is etched

Then a standard silicon etch step is carried out to form the plurality of trenches **209** according to the mask. In FIG. **8**, blanket implanting of phosphorus ions **211** (e.g. P^{31}) to the whole device may be performed to locally increase the doping concentration of N− epitaxial layer. The implantation is preferred to be done at tilt of 0 degree. The trench mask around the edge termination area or the gate bus area (not shown in Figures) prevents the phosphor dopant getting into these areas. Therefore, only the active region of the device receives the doping enhancement implant.

After implantation a high temperature process in an oxygen-containing ambient is used to anneal and diffuse the phosphorus dopant. Consequently, a doping enhancement N layer **213** is formed inside N− epitaxial layer as shown in FIG. **9**. The trench walls may then be oxidized first using a sacrificial oxidation. After removing the sacrificial oxide layer a pad oxide is re-grown along trench side wall. In FIG. **10**, the trenches are filled with high density oxide **217**. Oxide **217** may include silicon dioxide, or other types of deposited oxide, such as LTO or TEOS or High Density Plasma (HDP) oxide. The oxide is then thinned as shown in FIG. **11** using a dry plasma etch or CMP technique to planarize the oxide surface **219**.

In FIG. **12**, after active mask **223** has been applied with openings over trenches **222**, the oxide is etched further down into trench forming the trench Bottom Oxide layer (BOX) **221**. Then, in FIG. **13**, the BOX mask is used to protect the active gate trench **225** and the edge termination areas (not shown). The oxide removal step is carried out to completely etch away the BOX inside the RFP trench. Before removing the BOX mask, boron-11 ions **229** are implanted into N/N− epitaxial layer through RFP trench bottom **231**, forming P layer or isolation zones **237** shown in FIG. **14**.

In one embodiment, to implement the structure shown in FIG. **4**(*c*), a tilt angle implant is used to introduce boron along the RFP side walls. After removal of BOX photoresist **233** an optional high temperature anneal is employed to diffuse the boron, forming P layer or isolation zones **237** inside N− epitaxial region. Then the gate oxide **235** is grown along the trench sidewall in FIG. **14**.

The rest of process steps shown in FIG. **15** to FIG. **17** are similar to the one described in FIGS. 14-17 in US patent application No. 2008/0073707, which is herein incorporated by reference. The final device structure is shown in FIG. **18**. It is essential to point out that, by properly choosing RFP poly recess depth combined with the implant energy of P+ implant, the P+ region can be made deeper than P body, as shown in FIG. **19**A. Depending on the doping concentration of P shield region (or isolation zone), the P shield zone could be a "π" region **260** (a very lightly doped P region) shown in FIG. **19**C or a "ν" region **250** (a very lightly doped n region) shown in FIG. **19**B. A deeper P+ region is desired in order to improve the device ruggedness and connect the buried P region to the source electrode. In addition, the N++ source region can also be recessed completely as shown in FIG. **20**, so that the N++ source mask-photo step can be eliminated.

Furthermore, the techniques proposed in this invention may also be implemented using split poly gated device structures. One of implementation schemes are briefly demonstrated in the FIG. **21** to FIG. **23**. The process includes deposition of a first poly layer in the trenches, poly etch-back, and oxide removal, gate oxidation, second poly layer deposition, and CMP and/or poly etch-back. The split gated double poly configuration shown in FIG. **21-23** is used to replace the single poly layer in the active trench gate and the RFP trench shown in FIG. **18**. In this case, the bottom poly layer and the upper poly layer in RFP trench are both electrically shorted to the source metal. In addition, the split poly layers in the RFP region of device in FIG. **23** can be directly replaced by the single RFP poly layer as demonstrated by FIG. **24**A and FIG. **25**A. Depending on the doping concentration of P shield region (or isolation zone), at very light concentration, the P shield zone could be a "π" region (a very lightly doped P region) or a "ν" region (a very lightly doped n region) as shown in FIGS. **24**B, **24**C, **25**B and **25**C.

FIG. **28** is a flow chart depicting a fabrication process for making a MOSFET in accordance with one embodiment of the present invention. The fabrication process includes growing **302** an N− epitaxial layer on an N+ substrate. The fabrication process also includes locally increasing **304** the doping concentration within the N− epitaxial layer. Locally increasing **304** the doping concentration within the N− epitaxial layer includes blanket implanting phosphorous. The blanket implanting of phosphorous may be at a tilt angle of zero degrees, or may be at some other tilt angle. Locally increasing **304** the doping concentration within the N− epitaxial layer also includes excluding phosphorous dopant from an edge termination area and/or a gate bus area, including leaving oxide at an edge termination area or a gate bus area.

The fabrication process for making a MOSFET also includes creating **306** a doping enhancement N layer inside the N− epitaxial layer, including annealing and spreading the phosphorous dopant using a high temperature thermal process in an oxygen ambient. The fabrication process for making a MOSFET also includes smoothing **308** the trench surface of the trench, and reducing the roughness of the trench surface induced during the silicon trench etch, including oxidizing the trench side wall, removing **310** the sacrificial oxide layer, and regrowing **312** the pad oxide along the trench side wall.

The fabrication process for making a MOSFET also includes filling **314** the trench with high density oxide, planarizing oxide surface **316**, including etching back an oxide, forming a Bottom Oxide layer (BOX), including etching further down into the trench using an active mask, and protecting **318** an active gate trench and an edge termination, using a BOX mask. The fabrication process for making a MOSFET also includes implanting **320** boron (B11) into an N− epitaxial layer through a recessed field plate trench bottom, including introducing boron along recessed field plate side walls at a tilt angle, completely etching away **322** the BOX inside an recessed field plate trench, including removing the oxide, optionally driving **324** boron forming a P layer into the N− epitaxial layer, including annealing at a high temperature, and growing **324** a gate oxide along a trench side wall.

For a sample 40V embodiment, preferred parameters are as follows. However, it must be understood that these parameters would be scaled for different operating voltages, and of course they can also be adapted for use with many other processes. In this sample embodiment, the trenches are 0.3 microns wide, about 1.0 micron deep, and are laid out on a one micron pitch. (The cell pitch is two microns, since there are two types of trench present.) In this sample embodiment, the

starting material is 0.35 ohm-cm n-on-n+ epi, about 5.5 microns thick. A blanket n-enhancement implant is performed, e.g. with phosphorus at 3E12/cm^2 (i.e. 3×10^{12} cm^{-2}). The trenches are then etched. After a sacrificial oxidation and trench fill (preferably using a deposited oxide plus oxidation), an etchback is preferably performed to clear the trenches to about half their depth. Photoresist is then patterned, to expose the RFP trenches but not the gate trenches, and the oxide plugs are removed from the RFP trenches. A P-type implant is then performed to form the P-isolation regions; in this example, a combination of two boron implants, one at 2.5E12/cm^2 at 30 keV plus another 2E12 at 120 keV. This will produce a counterdoped or compensated isolation region **114** below the RFP trenches, of about 0.7 micron depth. The remaining process steps then proceed conventionally, with formation of gate, body, source, contacts, etc.

As mentioned above, the locally enhanced n-doping which connects gate to drain, in various embodiments described above, reduces on-resistance. However, it is the improved off-state behavior provided by the added isolation regions which makes this enhanced n-doping possible.

In alternative embodiments, the depth of the isolation region can be e.g. from 0.25 micron to 2.5 micron, and scaled accordingly for operating voltages other than 40V. Similarly, the isolation implant, in alternative embodiments, can use a dose from 2E12 cm^{-2} to 1E13 at 20-320 keV, or even higher or lower doses and/or energies, plus allowance for scaling.

It will be appreciated that the foregoing is merely a description of some specific illustrative and exemplary embodiments of the present invention, and should not be considered as descriptive of the entire gamut of embodiments that fall within the scope of the present invention.

According to various embodiments, there is provided: a semiconductor device structure, comprising a gate which is positioned in a first trench, and capacitively coupled to control vertical conduction from a first-conductivity-type source through semiconductor material which is adjacent to said trench; recessed field plates, positioned in proximity to and capacitively coupled to said semiconductor material; said recessed field plates being positioned in respective second trenches; and diffusions of a second conductivity type lying at least partially beneath said respective second trenches.

According to various embodiments, there is provided: a semiconductor device structure, comprising a semiconductor layer; a gate which is positioned in a first trench within said semiconductor layer, and is capacitively coupled to control vertical conduction from a first-conductivity-type source through second-conductivity-type portions of said layer near said trench; recessed field plates, positioned in proximity to and capacitively coupled to said semiconductor material; said recessed field plates being positioned in respective second trenches; diffusion components of a second conductivity type lying at least partially beneath said respective second trenches; whereby said diffusion components reduce depletion of said second-conductivity-type portions of said layer in the OFF state.

According to various embodiments, there is provided: a semiconductor device structure, comprising a semiconductor layer; a gate which is positioned in a first trench within said semiconductor layer, and is capacitively coupled to control vertical conduction from a first-conductivity-type source through second-conductivity-type portions of said layer near said trench; recessed field plates, positioned in proximity to and capacitively coupled to said semiconductor material; said recessed field plates being positioned in respective second trenches; a first additional diffusion component of a second conductivity type lying at least partially beneath said respective second trenches; and a second additional diffusion component of said first conductivity type lying at least partially within said second-conductivity-type portions of said layer; whereby said first additional diffusion component reduces depletion of said second-conductivity-type portions of said layer in the OFF state; and whereby said second additional diffusion component reduces the on-resistance of the device in the ON state.

According to various embodiments, there is provided: an improved RFP transistor structure having (a) low total on-resistance, (b) reduced minority carrier injection efficiency (of body diode), (c) improved reverse recovery (of body diode), (c) lower reverse recovery charge, (d) soft recovery characteristic, (e) as reliable edge termination, without either reduction of breakdown voltage or degradation of the termination efficiency of device edge junction termination region, the improved structure comprising: an RFP transistor structure, including at least one or more gate trenches adjoined by one or more recessed-field-plate trenches; and respective deep compensated zones underneath said recessed-field-plate trenches.

According to various embodiments, there is provided: a method for operating a semiconductor device structure, comprising: controlling conduction between first and second source/drain electrodes through a channel location in semiconductor material using a gate electrode positioned in a first trench to provide at least ON and OFF states; and avoiding punchthrough of said channel location, using both one or more recessed field plates, positioned in proximity to and capacitively coupled to said semiconductor material; said recessed field plates being positioned in respective second trenches, and one or more diffusion components of a second conductivity type lying at least partially beneath said respective second trenches; whereby said diffusion components reduce depletion spreading in the OFF state.

According to various embodiments, there is provided: a fabrication process for making a MOSFET, comprising the actions, in any order, of: a) providing an n-type semiconductor layer; b) forming a p-type body in said layer; c) forming an n-type source, which is isolated by said body, in said layer; d) forming an insulated gate trench in said layer, and a gate electrode in said gate trench; said gate electrode being capacitively coupled to at least a portion of said body; e) forming a second insulated trench in said layer, providing an additional dose of acceptor dopants below said trench, and forming a Recessed Field Plate electrode in said second trench; and f) providing an additional dose of donor dopant atoms in said portion of said body, to thereby reduce the on-resistance.

According to various embodiments, there is provided: improved highly reliable power RFP structures and fabrication and operation processes. The structure includes plurality of localized dopant concentrated zones beneath the trenches of RFPs, either floating or extending and merging with the body layer of the MOSFET or connecting with the source layer through a region of vertical doped region. This local dopant zone decreases the minority carrier injection efficiency of the body diode of the device and alters the electric field distribution during the body diode reverse recovery.

Modifications and Variations

As will be recognized by those skilled in the art, the innovative concepts described in the present application can be modified and varied over a tremendous range of applications, and accordingly the scope of patented subject matter is not limited by any of the specific exemplary teachings given. It is

intended to embrace all such alternatives, modifications and variations that fall within the spirit and broad scope of the appended claims.

The device may be fabricated in various layouts, including "stripe" and "cellular" layouts. The layers of source, body, and drain regions can be configured vertically, quasi-vertically as well as laterally. The epitaxial drift region can be either uniformly or non-uniformly doped. While the embodiments described above include an epitaxial layer grown on a substrate, the epitaxial layer may be omitted in some applications. Various features of different embodiments may be combined and recombined for various applications.

For example, the region between channel and drain does not have to be uniformly doped, neither vertically nor laterally. The improvements in drift or spreading region doping provided by the disclosed inventions can be combined with a wide variety of other device improvements and features.

For another example, the RFP and gate trenches do not necessarily have to have the same width.

The design could be applied to IGBTs or other devices which include bipolar conduction. The bottom of the gate trench can be modified with dopant; the design can also vary at the source structure and at the drain structure; and alternative body structure may be used; contact trench may be produced first, then cut gate trench, and construct the source and drain structure.

Of course, the n-type dopant, in silicon, can be phosphorus, antimony, or arsenic, or combinations of these. Appropriate donor dopants can be used in other semiconductor materials.

As the disclosed process is scaled to other operating voltages, it is expected that predictable scaling of dimensions and dopants may allow the same synergy. For example, in a 200V embodiment, the inventors contemplate that the trench depth would be slightly deeper (e.g. 1.5 to 2.5 micron), and the compensation implant energy and dose would be about the same. Of course the epi layer doping would be substantially less, and the epi layer thickness greater, as is well understood by those of ordinary skill. The n-enhancement doping (which is preferably blocked from the termination) can have a distribution, after drive-in, which reaches to the upper boundary of the compensation implant, but preferably not to the lower boundary of the compensation implant.

The following applications may contain additional information and alternative modifications: Ser. No. 61/125,892 filed Apr. 29, 2008; Ser. No. 61/058,069 filed Jun. 2, 2008 and entitled "Edge Termination for Devices Containing Permanent Charge"; Ser. No. 61/060,488 filed Jun. 11, 2008 and entitled "MOSFET Switch"; Ser. No. 61/074,162 filed Jun. 20, 2008 and entitled "MOSFET Switch"; Ser. No. 61/076,767 filed Jun. 30, 2008 and entitled "Trench-Gate Power Device"; Ser. No. 61/080,702 filed Jul. 15, 2008 and entitled "A MOSFET Switch"; Ser. No. 61/084,639 filed Jul. 30, 2008 and entitled "Lateral Devices Containing Permanent Charge"; Ser. No. 61/084,642 filed Jul. 30, 2008 and entitled "Silicon on Insulator Devices Containing Permanent Charge"; Ser. No. 61/027,699 filed Feb. 11, 2008 and entitled "Use of Permanent Charge in Trench Sidewalls to Fabricate Un-Gated Current Sources, Gate Current Sources, and Schottky Diodes"; Ser. No. 61/028,790 filed Feb. 14, 2008 and entitled "Trench MOSFET Structure and Fabrication Technique that Uses Implantation Through the Trench Sidewall to Form the Active Body Region and the Source Region"; Ser. No. 61/028,783 filed Feb. 14, 2008 and entitled "Techniques for Introducing and Adjusting the Dopant Distribution in a Trench MOSFET to Obtain Improved Device Characteristics"; Ser. No. 61/091,442 filed Aug. 25, 2008 and entitled "Devices Containing Permanent Charge"; Ser. No. 61/118,664 filed Dec. 1, 2008 and entitled "An Improved Power MOSFET and Its Edge Termination"; and Ser. No. 61/122,794 filed Dec. 16, 2008 and entitled "A Power MOSFET Transistor".

None of the description in the present application should be read as implying that any particular element, step, or function is an essential element which must be included in the claim scope: THE SCOPE OF PATENTED SUBJECT MATTER IS DEFINED ONLY BY THE ALLOWED CLAIMS. Moreover, none of these claims are intended to invoke paragraph six of 35 USC section 112 unless the exact words "means for" are followed by a participle.

The claims as filed are intended to be as comprehensive as possible, and NO subject matter is intentionally relinquished, dedicated, or abandoned.

What is claimed is:

1. A method for operating a semiconductor device structure, comprising:

controlling conduction between first and second source/drain electrodes through a channel location in semiconductor material using a gate electrode positioned in a first trench to provide at least ON and OFF states; and

avoiding punchthrough of said channel location, using both

one or more recessed field plates, positioned in proximity to and capacitively coupled to said semiconductor material; said recessed field plates being positioned in respective second trenches, and

one or more diffusion components of a second conductivity type lying at least partially directly beneath said respective second trenches;

whereby said diffusion components reduce depletion spreading in the OFF state.

2. A method for operating a semiconductor device structure, comprising:

controlling conduction between first and second source/drain electrodes through a channel location in semiconductor material using a gate electrode positioned in a first trench to provide at least ON and OFF states; and

avoiding punchthrough of said channel location, using both

one or more recessed field plates, positioned in proximity to and capacitively coupled to said semiconductor material; said recessed field plates being positioned in respective second trenches, and

one or more diffusion components of a second conductivity type lying at least partially directly beneath said respective second trenches;

whereby said diffusion components reduce depletion spreading in the OFF state;

wherein said device further includes a layer of dopant concentration region of second-conductivity-type that extends from a source layer to at least one location of said diffusion components.

3. The method of claim **1**, wherein the gate has a split poly configuration.

4. The method of claim **1**, wherein at least one of the recessed field plates has a split poly configuration.

5. The method of claim **1**, wherein both the gate and at least one of the recessed field plates have a split poly configuration.

6. The method of claim **1**, wherein said gate is capacitively coupled to control vertical conduction to a drain diffusion of a first conductivity type.

7. A fabrication process for making a MOSFET, comprising the actions, in any order, of:

a) providing an n-type semiconductor layer;

b) forming a p-type body in said layer;

c) forming an n-type source, which is isolated by said body, in said layer;

d) forming an insulated gate trench in said layer, and a gate electrode in said gate trench; said gate electrode being capacitively coupled to at least a portion of said body;

e) forming a second insulated trench in said layer, providing an additional dose of acceptor dopants below said trench, and forming a Recessed Field Plate electrode in said second trench; and

f) providing an additional dose of donor dopant atoms in said portion of said body, to thereby reduce the on-resistance.

* * * * *

US008659076B2

(12) **United States Patent**
Zeng et al.

(10) **Patent No.: US 8,659,076 B2**
(45) **Date of Patent: *Feb. 25, 2014**

(54) **SEMICONDUCTOR DEVICE STRUCTURES AND RELATED PROCESSES**

(75) Inventors: **Jun Zeng**, Torrance, CA (US); **Mohamed N. Darwish**, Campbell, CA (US)

(73) Assignee: **MaxPower Semiconductor, Inc.**, Santa Clara, CA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 319 days.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: **13/212,747**

(22) Filed: **Aug. 18, 2011**

(65) **Prior Publication Data**

US 2012/0032258 A1 Feb. 9, 2012

Related U.S. Application Data

(63) Continuation of application No. 12/368,399, filed on Feb. 10, 2009, now Pat. No. 8,076,719.

(60) Provisional application No. 61/065,759, filed on Feb. 14, 2008.

(51) **Int. Cl.**
H01L 29/78 (2006.01)

(52) **U.S. Cl.**
USPC **257/330**; 257/341; 257/E29.262

(58) **Field of Classification Search**
USPC 257/330, 341, E29.262
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,168,331	A	12/1992	Yilmaz
5,282,018	A	1/1994	Hiraki et al.
5,525,821	A	6/1996	Harada
5,637,898	A	6/1997	Baliga
5,864,159	A	1/1999	Takahashi
5,973,359	A	10/1999	Kobayashi
5,998,833	A	12/1999	Baliga
6,069,372	A	5/2000	Uenishi
6,114,727	A	9/2000	Ogura et al.
6,191,447	B1	2/2001	Baliga
6,251,730	B1	6/2001	Luo

(Continued)

FOREIGN PATENT DOCUMENTS

WO	97/33320 A1	9/1997
WO	2006027739	3/2006

OTHER PUBLICATIONS

J. T. Watt, B. J. Fishbein & J. D. Plummer; Low-Temperature NMOS Technology with Cesium-Implanted Load Devices; IEEE Trans. Electron Devices, vol. 34, # 1, Jan. 1987; p. 28-38.

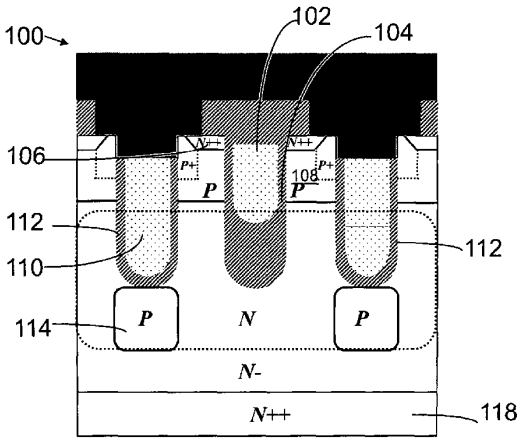
(Continued)

Primary Examiner — Long Pham
(74) *Attorney, Agent, or Firm* — Robert O. Groover, III; Gwendolyn S. S. Groover; Groover & Associates, PLLC

(57) **ABSTRACT**

Improved highly reliable power RFP structures and fabrication and operation processes. The structure includes plurality of localized dopant concentrated zones beneath the trenches of RFPs, either floating or extending and merging with the body layer of the MOSFET or connecting with the source layer through a region of vertical doped region. This local dopant zone decreases the minority carrier injection efficiency of the body diode of the device and alters the electric field distribution during the body diode reverse recovery.

15 Claims, 27 Drawing Sheets



(56) **References Cited**

U.S. PATENT DOCUMENTS

6,388,286	B1	5/2002	Baliga	
6,468,878	B1 *	10/2002	Petruzzello et al.	438/454
6,525,373	B1	2/2003	Kim	
6,534,828	B1	3/2003	Kocon	
6,541,820	B1	4/2003	Bol	
6,649,975	B2	11/2003	Baliga	
6,686,244	B2	2/2004	Blanchard	
6,710,403	B2	3/2004	Sapp	
6,803,627	B2	10/2004	Pfirsch	
2001/0001494	A1	5/2001	Kocon	
2001/0041407	A1	11/2001	Brown	
2003/0203576	A1	10/2003	Kitada et al.	
2006/0060916	A1	3/2006	Girdhar	
2007/0004116	A1	1/2007	Hshieh	
2007/0013000	A1	1/2007	Shiraishi	
2008/0099837	A1 *	5/2008	Akiyama et al.	257/341

OTHER PUBLICATIONS

J.T.Watt,B.J.Fishbein & J.D.Plummer;Characterization of Surface Mobility in MOS Structures Containing Interfacial Cesium Ions;IEEE Trans.Electron Devices,V36,Jan. 1989; p. 96-100.

J.R.Pfiester, J.R.Alvis & C.D.Gunderson; Gain-Enhanced LDD NMOS Device Using Cesium Implantation; IEEE Trans.Electron Devices, V39, #6, Jun. 1992; p. 1469-1476.

* cited by examiner

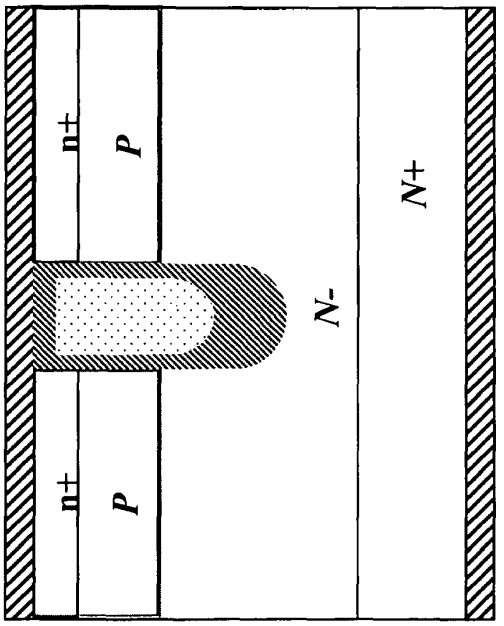


FIG. 1
(Prior Art)

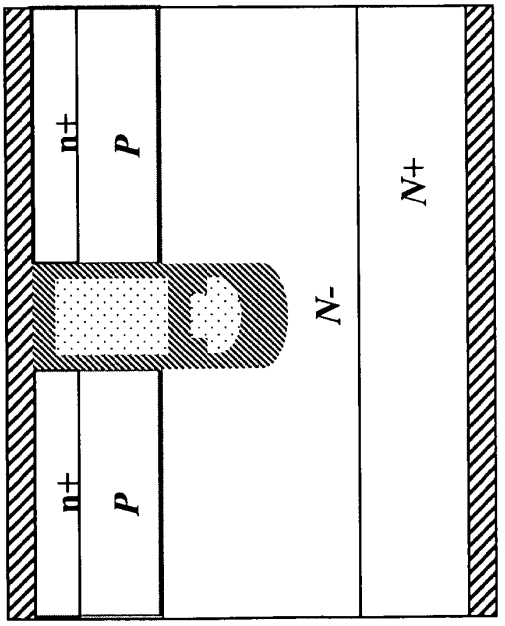


FIG. 2
(Prior Art)

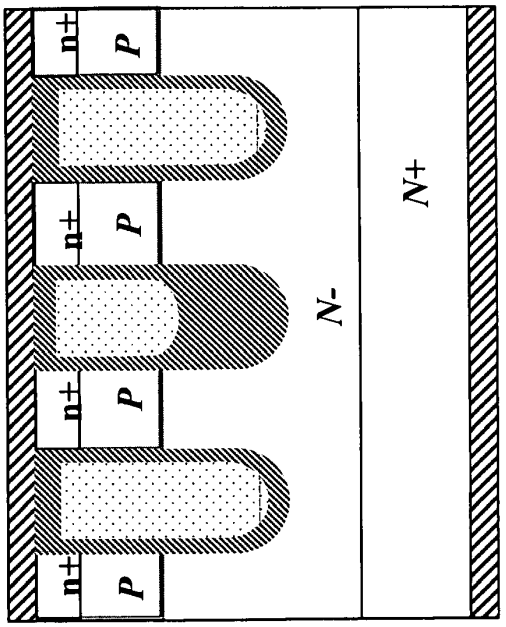


FIG. 3
(Prior Art)

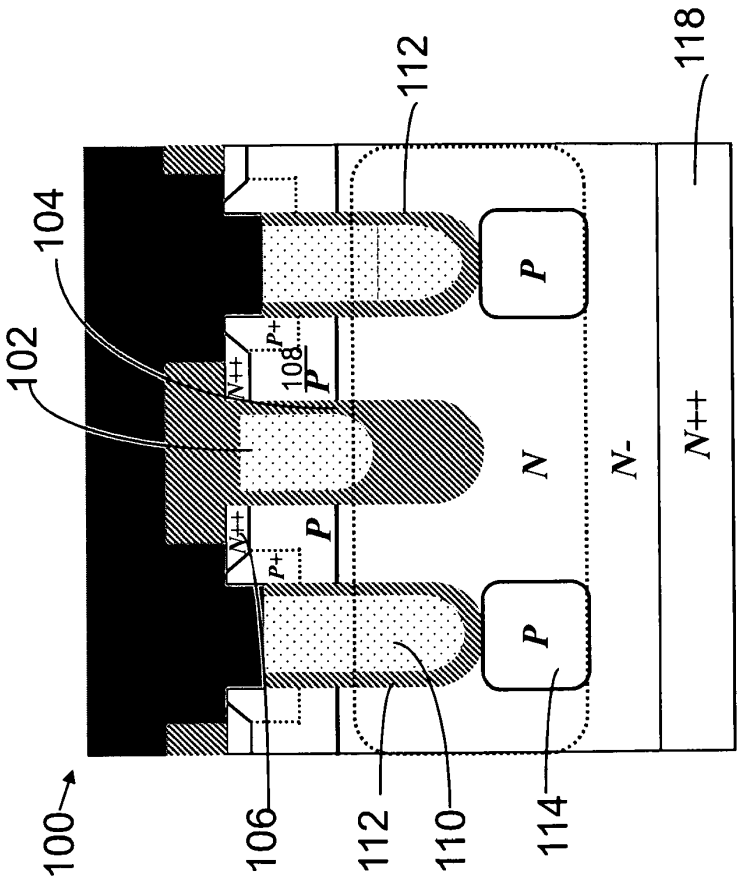


FIG. 4a

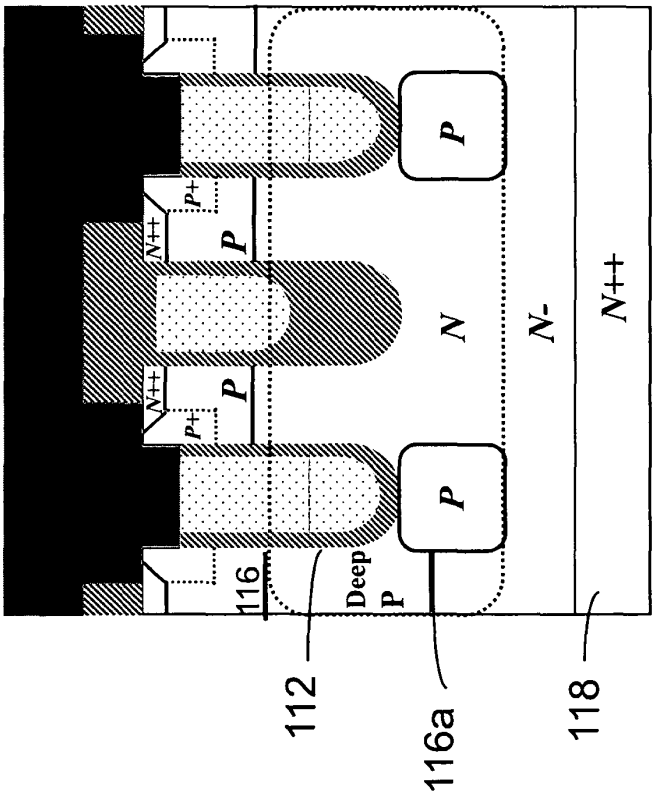


FIG. 4b

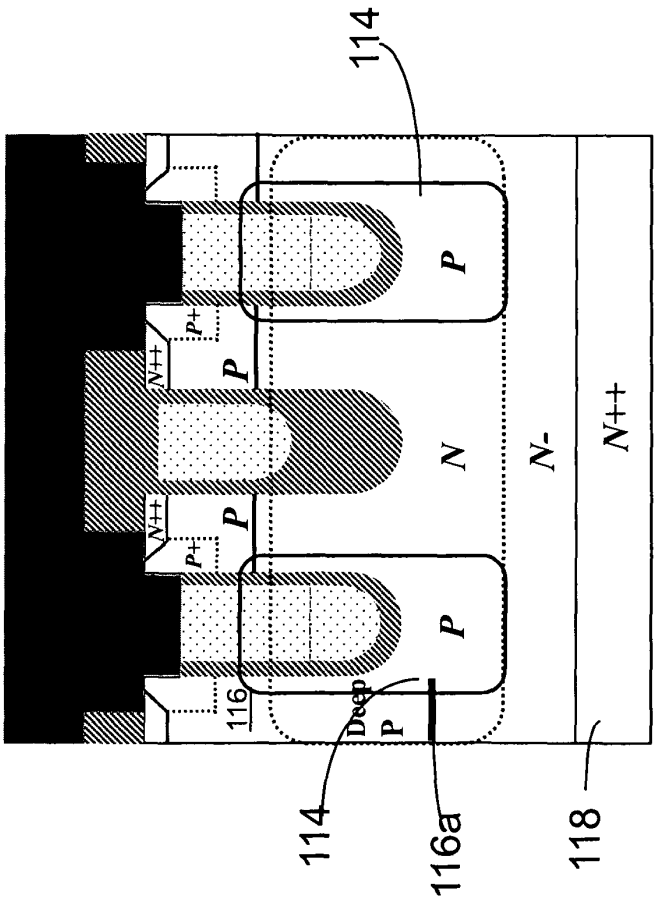


FIG. 4c

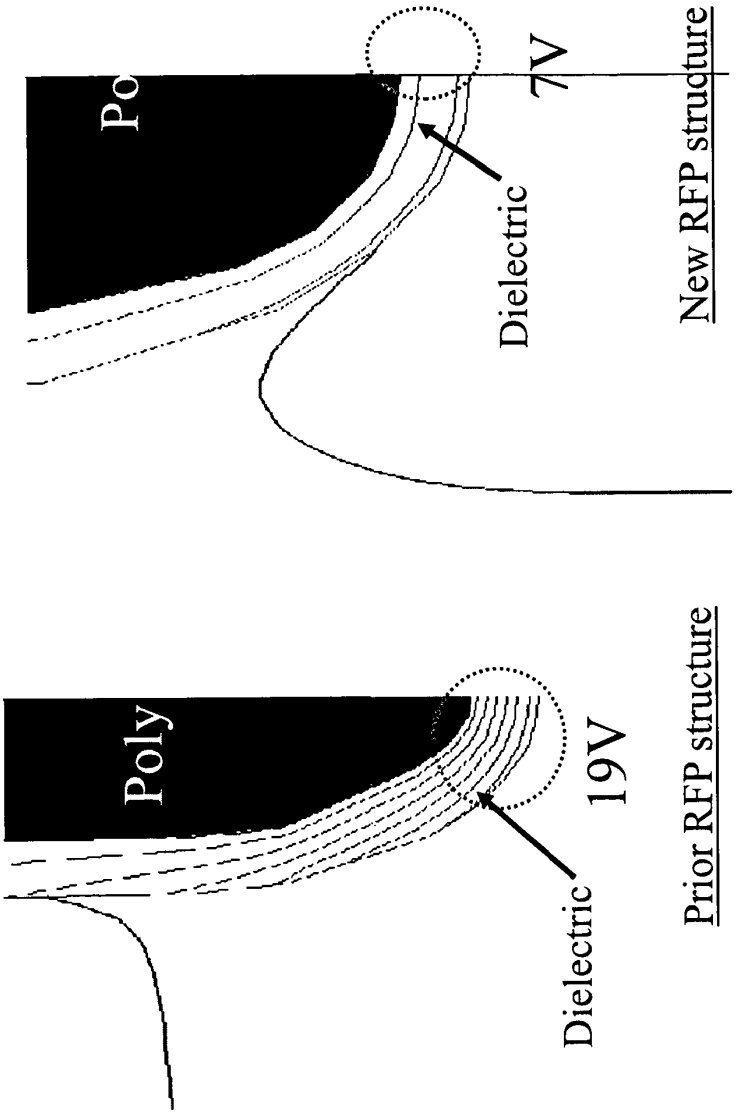


FIG. 5

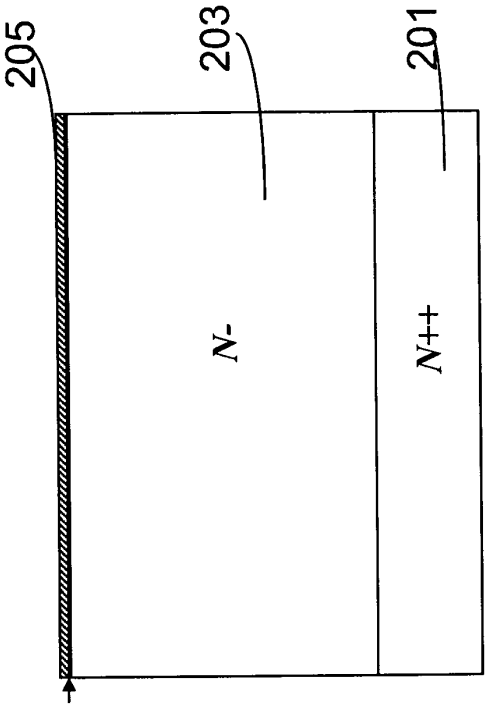


FIG. 6

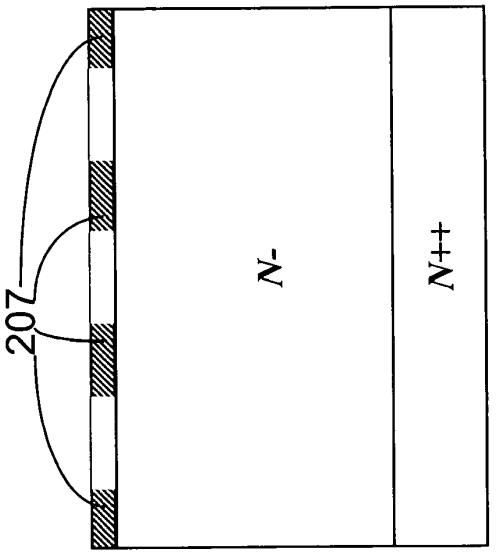


FIG. 7

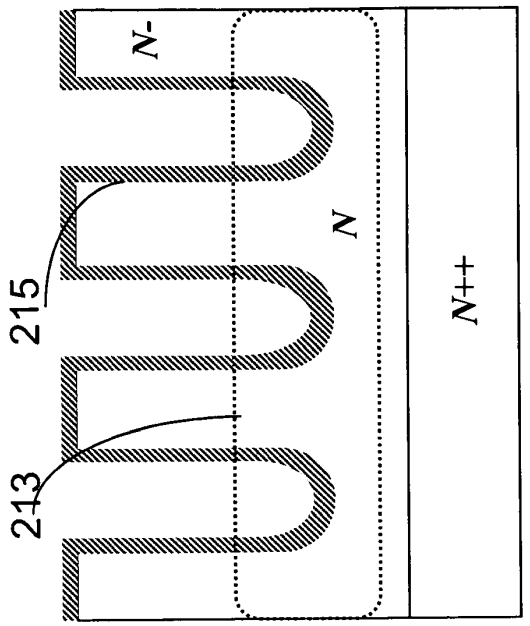


FIG. 9

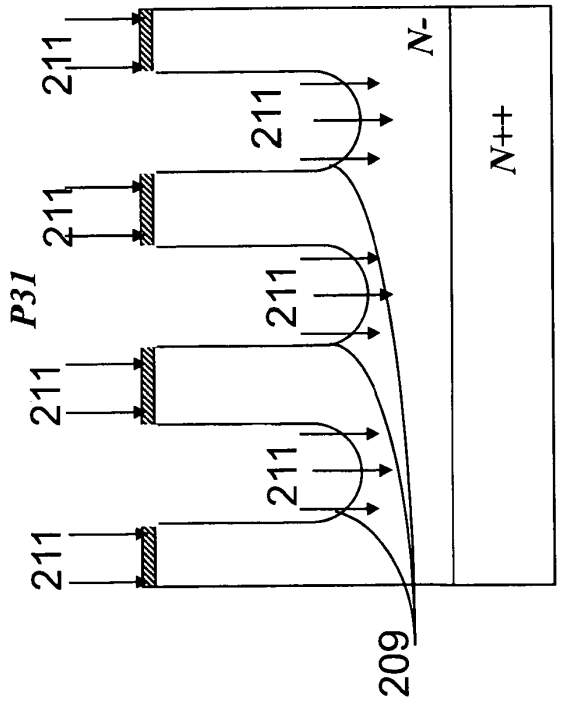


FIG. 8

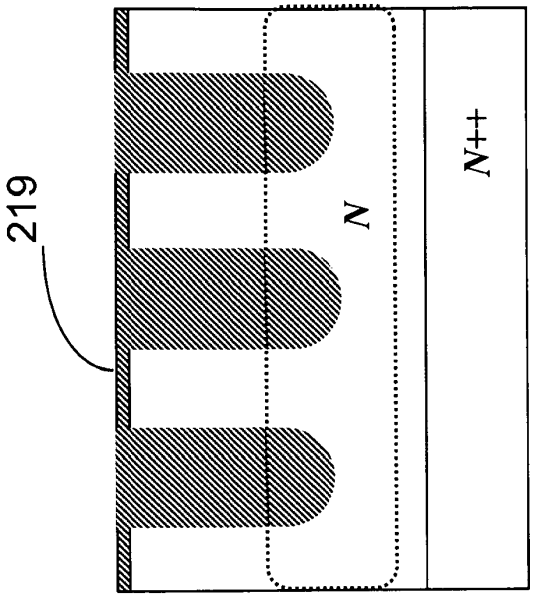


FIG. 11

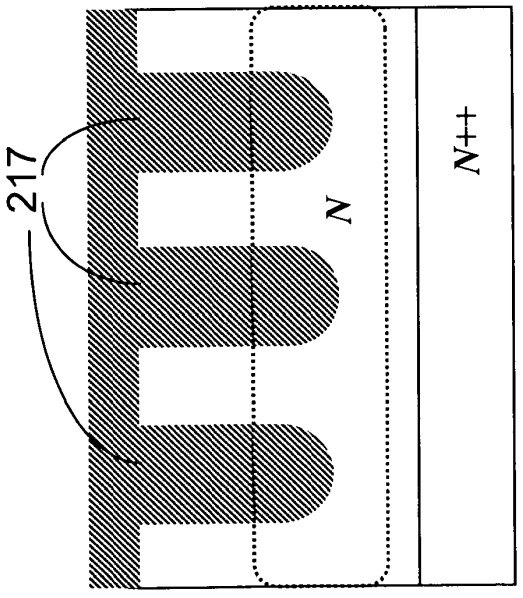


FIG. 10

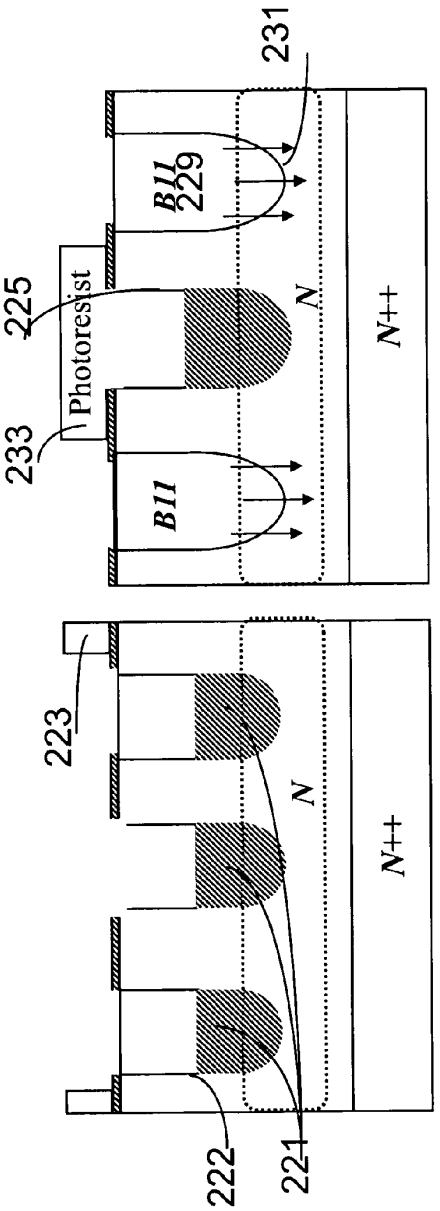


FIG.13

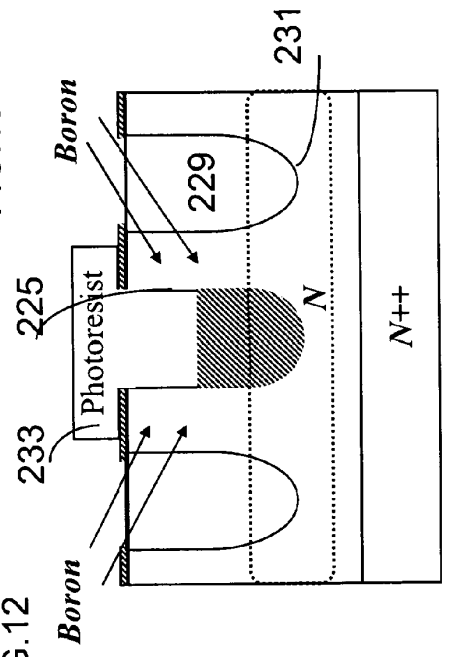


FIG.12

FIG.13+ (optional)

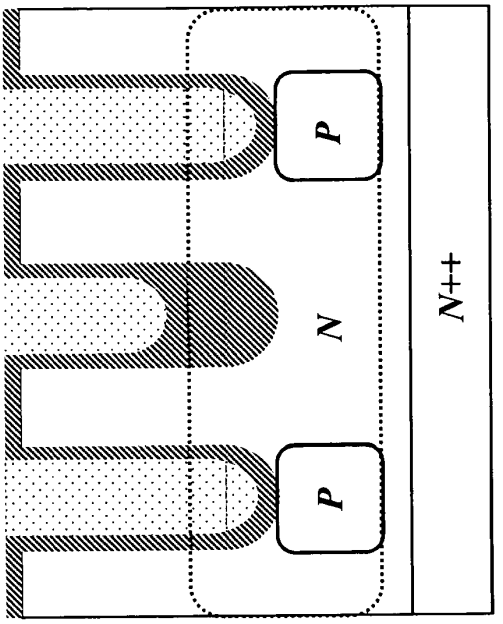


FIG. 15

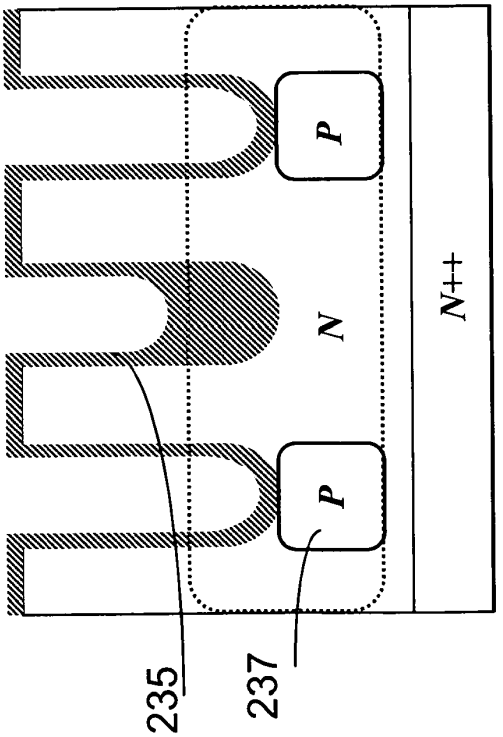


FIG. 14

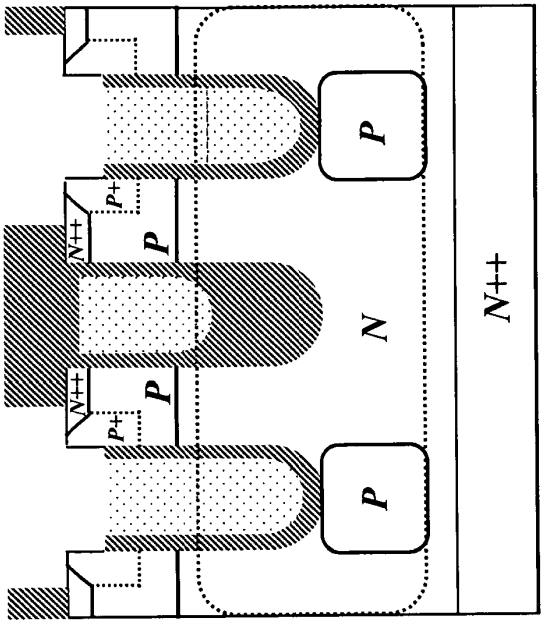


FIG. 17

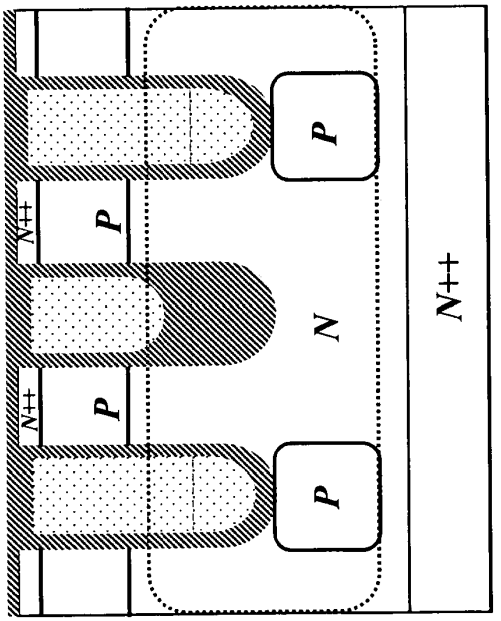


FIG. 16

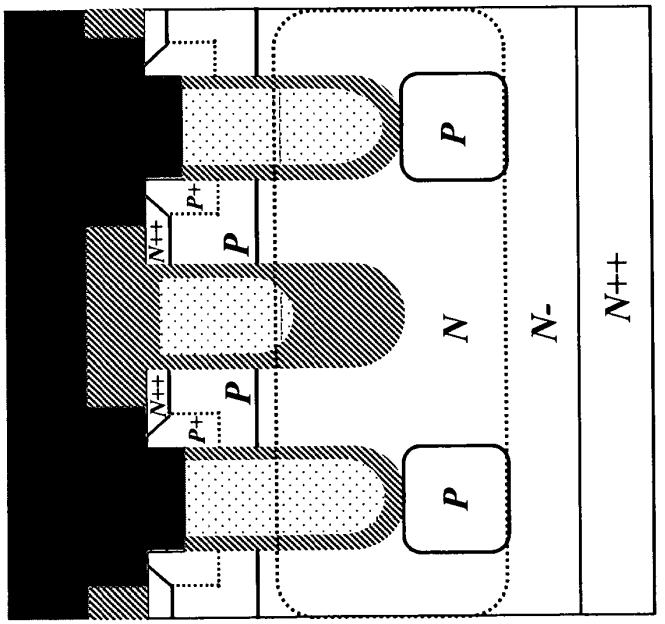


FIG.18

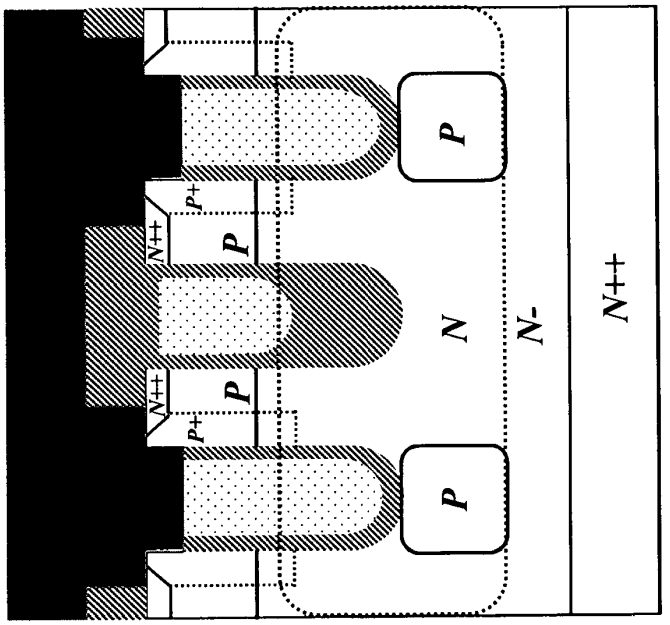


FIG. 19A

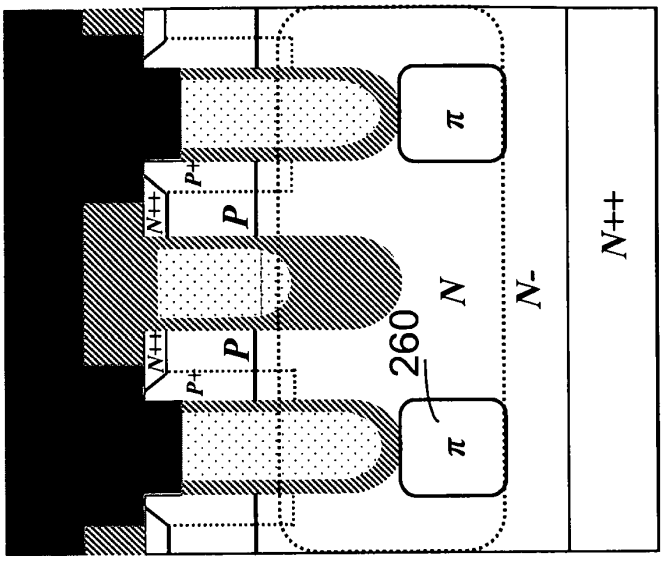


FIG.19C

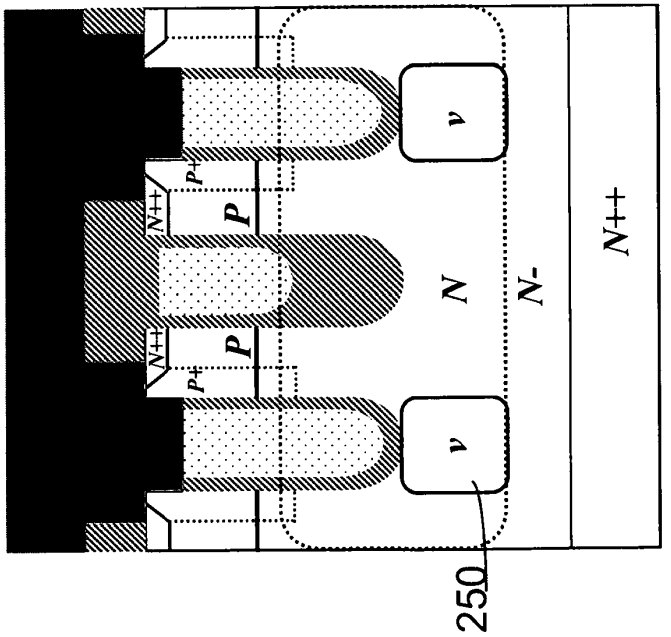


FIG. 19B

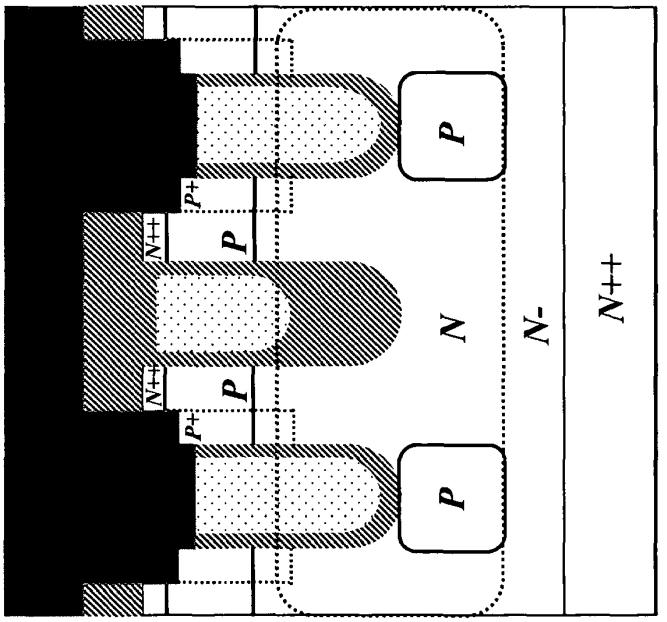


FIG. 20

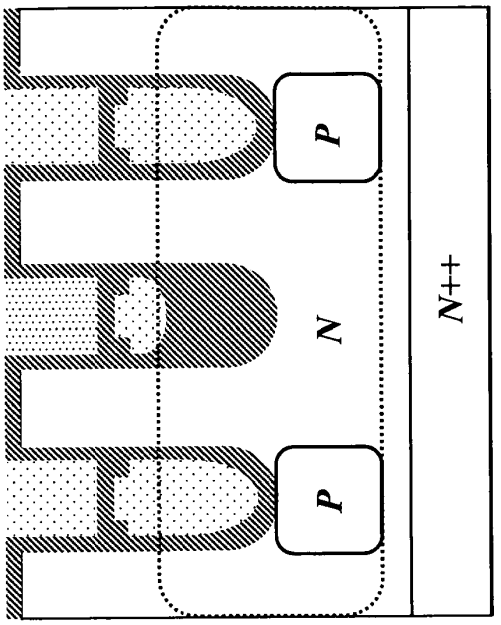


FIG. 22

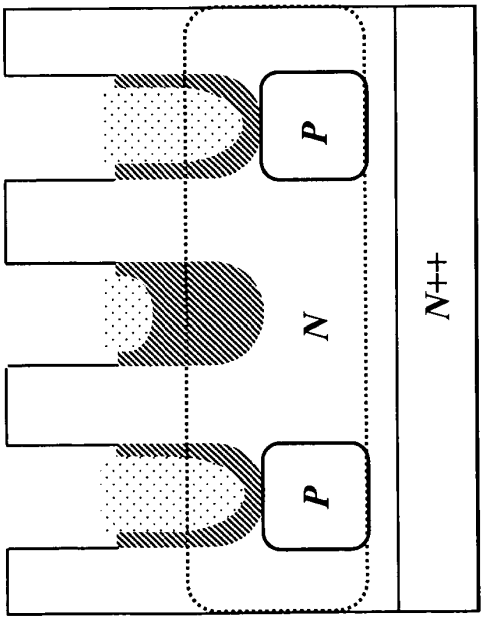


FIG. 21

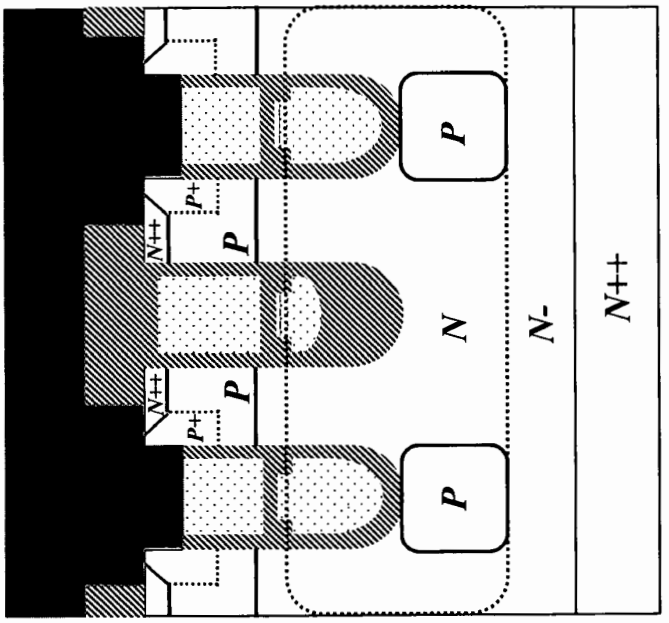


FIG. 23

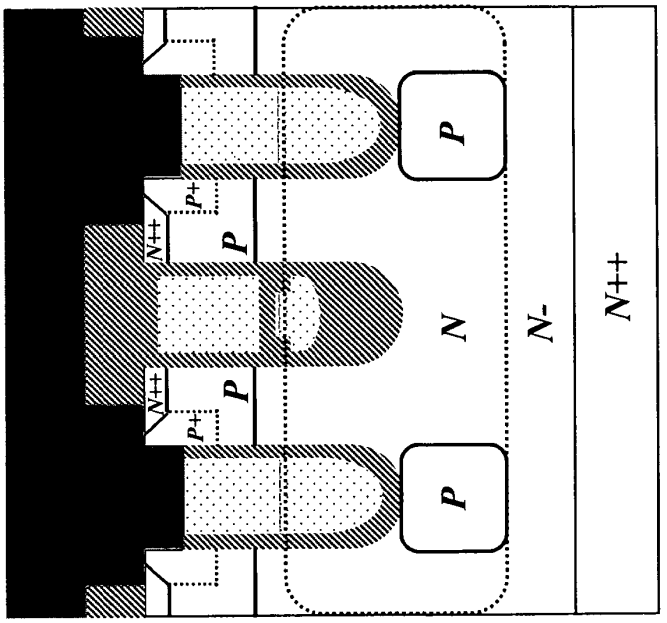


FIG. 24A

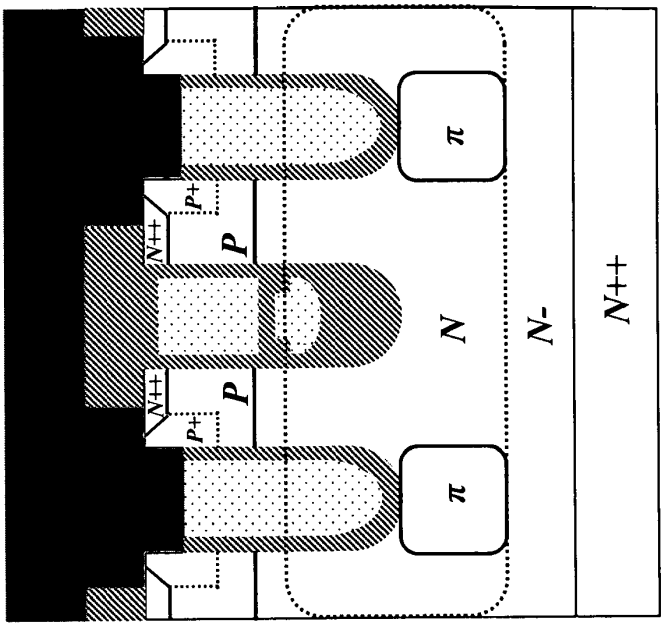


FIG. 24C

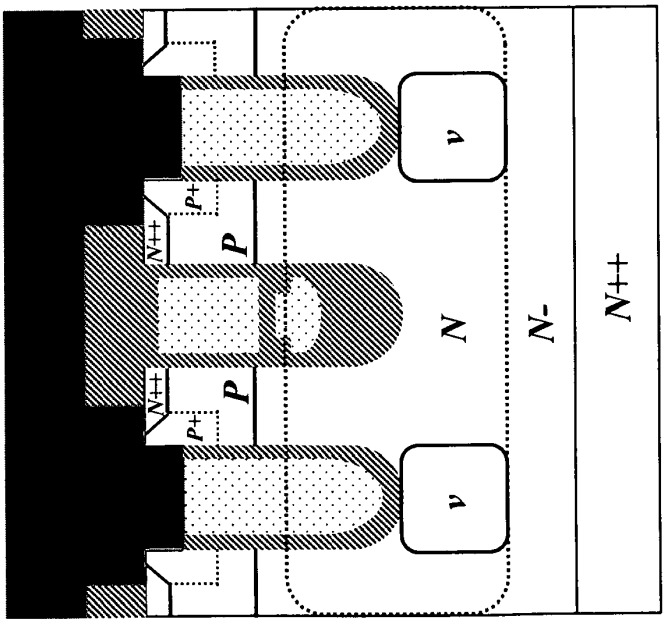


FIG. 24B

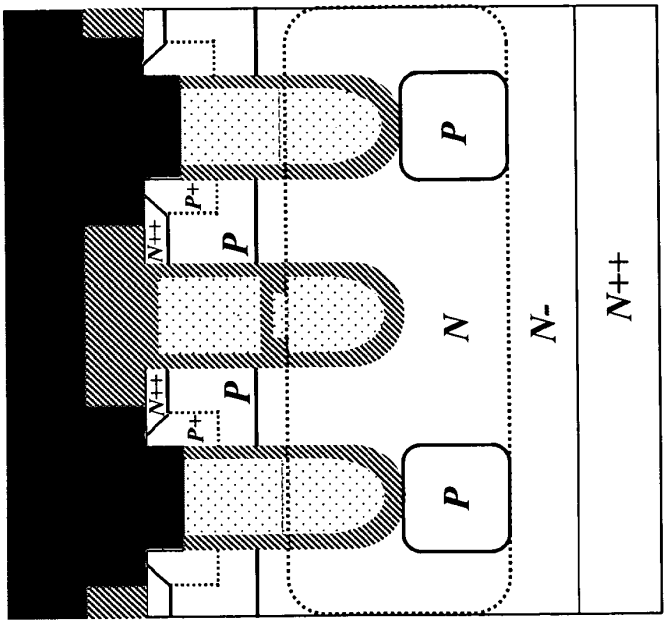


FIG. 25A

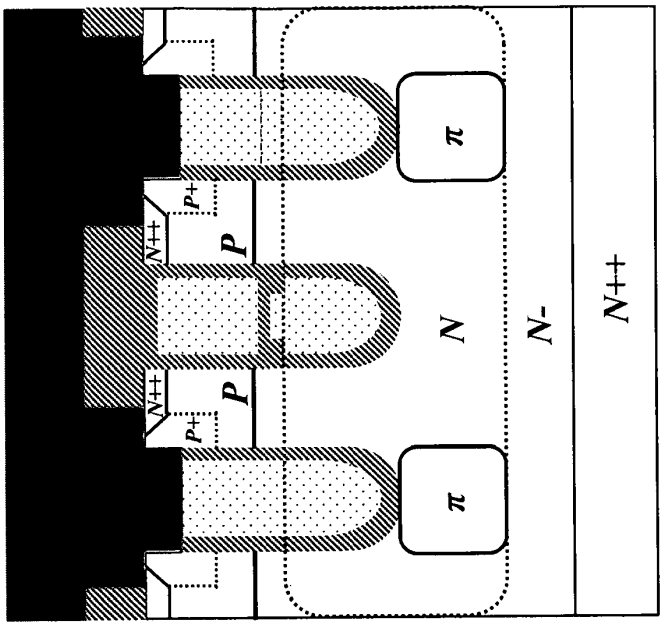


FIG. 25C

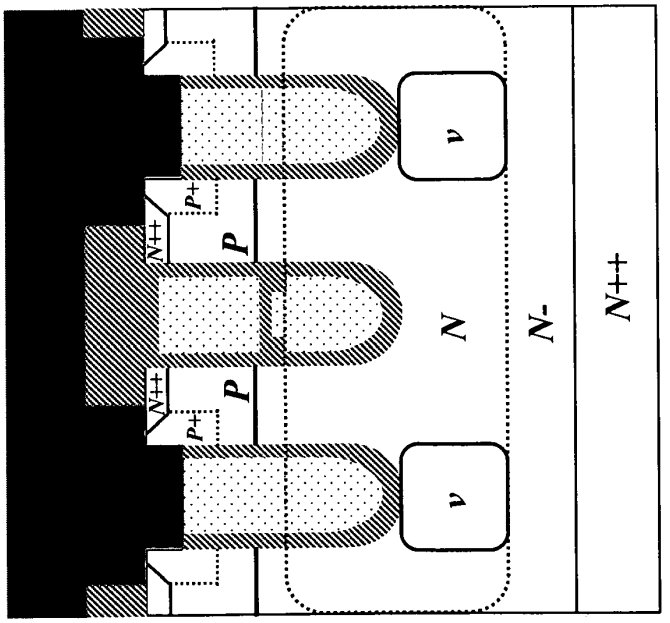


FIG. 25B

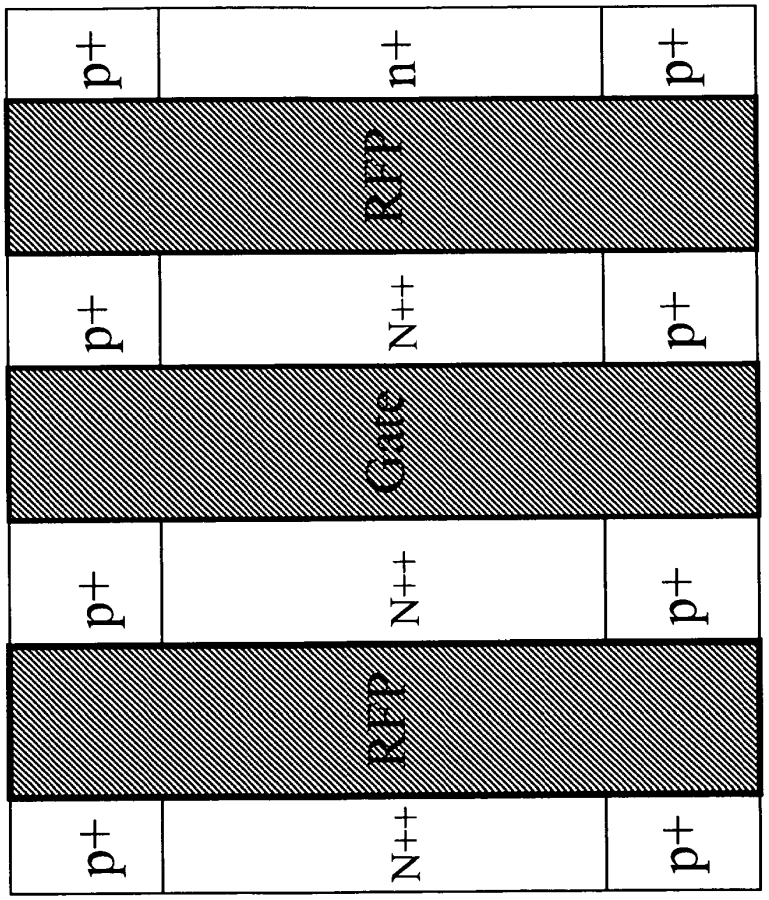


FIG. 26

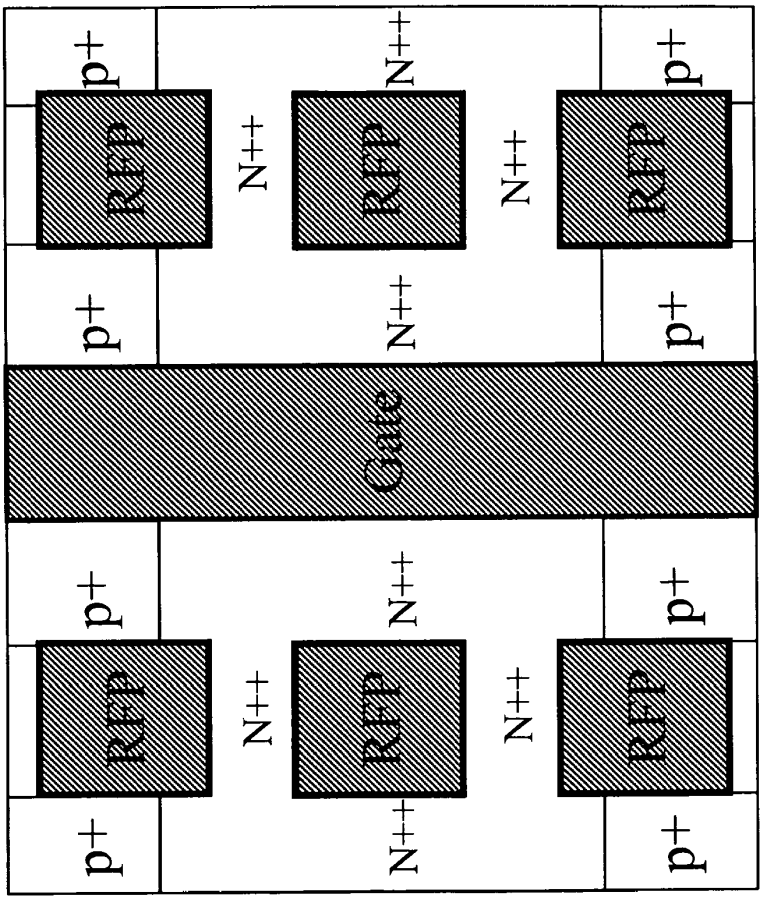


FIG. 27

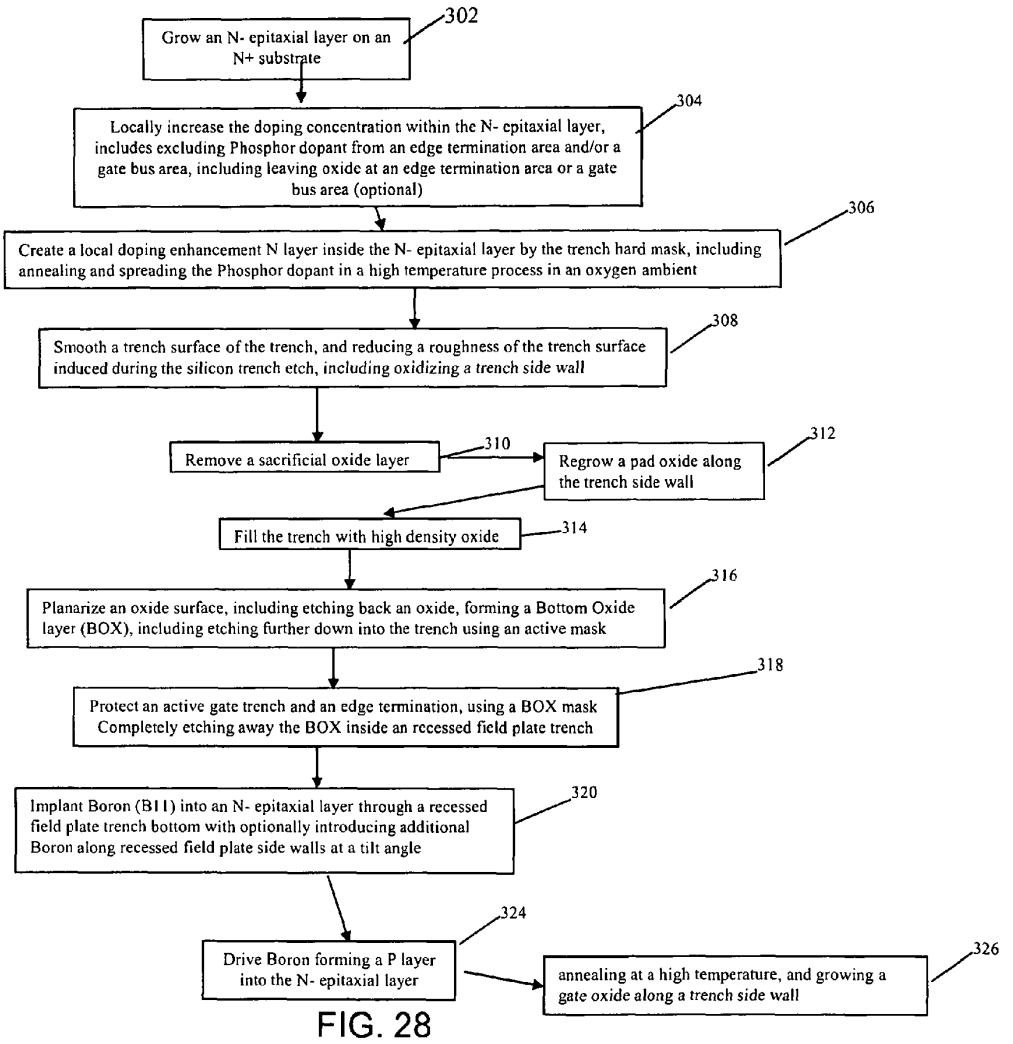

Grow an N- epitaxial layer on an N+ substrate
302
Locally increase the doping concentration within the N- epitaxial layer, includes excluding Phosphor dopant from an edge termination area and/or a gate bus area, including leaving oxide at an edge termination area or a gate bus area (optional)
304
Create a local doping enhancement N layer inside the N- epitaxial layer by the trench hard mask, including annealing and spreading the Phosphor dopant in a high temperature process in an oxygen ambient
306
Smooth a trench surface of the trench, and reducing a roughness of the trench surface induced during the silicon trench etch, including oxidizing a trench side wall
308
Remove a sacrificial oxide layer
310
Regrow a pad oxide along the trench side wall
312
Fill the trench with high density oxide
314
Planarize an oxide surface, including etching back an oxide, forming a Bottom Oxide layer (BOX), including etching further down into the trench using an active mask
316
Protect an active gate trench and an edge termination, using a BOX mask Completely etching away the BOX inside an recessed field plate trench
318
Implant Boron (B11) into an N- epitaxial layer through a recessed field plate trench bottom with optionally introducing additional Boron along recessed field plate side walls at a tilt angle
320
Drive Boron forming a P layer into the N- epitaxial layer
324
annealing at a high temperature, and growing a gate oxide along a trench side wall
326

FIG. 28

SEMICONDUCTOR DEVICE STRUCTURES AND RELATED PROCESSES

CROSS REFERENCE TO RELATED APPLICATIONS

This application claims priority to U.S. Provisional Application Ser. No. 61/065,759 filed Feb. 14, 2008, which is incorporated by reference herein in its entirety.

BACKGROUND

The present invention relates to field effect transistors and methods, and more particularly to highly reliable power insulated-gate field effect transistors (MOSFET) with a Recessed Field Plate (RFP) and related techniques.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** is a cross-sectional view of a prior art MOSFET with a trench gate having a thick bottom oxide structure.

FIG. **2** is a cross-sectional view of a prior art MOSFET with a split poly gate structure.

FIG. **3** is a cross-sectional view of a prior art MOSFET with RFPs in parallel with the gate trench.

FIG. **4**(*a*) is a cross-sectional view of an RFP containing MOSFET structure with a floating Deep compensated zone.

FIG. **4**(*b*) is a cross-sectional view of an RFP containing MOSFET structure with a Deep compensated zone that extends to and is connected to the source electrode.

FIG. **4**(*c*) is a cross-sectional view of an RFP containing MOSFET structure with a Deep compensated zone that extends to the P body region.

FIG. **5** shows a two-dimensional electrical voltage simulation comparison between the prior RFP-MOSFET structure and a MOSFET containing a Deep compensated zone.

FIGS. **6-18** show successive steps in a sample process for making the sample structure depicted in FIG. **4**(*a*).

FIG. **19**A is a cross-sectional view of a RFP containing MOSFET structure with a Deep compensated zone and a P+ implant region extending beyond the P-N junction and into the N drift region.

FIG. **19**B is a cross-sectional view of an RFP containing MOSFET structure with a Deep compensated zone that is a lightly doped p region, and a P+ implant region extending beyond the P-N junction and into the N drift region.

FIG. **19**C is a cross-sectional view of an RFP containing MOSFET structure with a Deep compensated zone that is a lightly doped n region, and a P+ implant region extending beyond P-N junction and into the N drift region.

FIG. **20** is a cross-sectional view of an RFP containing MOSFET structure with a Deep compensated zone, and a P+ implant region extending beyond P-N junction and into the N drift region and the N++ source region that is completely recessed down.

FIGS. **21-23** show a process for fabrication of an embodiment of FIG. **4**(*a*) structure implemented in the split poly gated structure with split poly layer structure in the RFP trench.

FIGS. **24**A, **24**B, **24**C, **25**A, **25**B and **25**C show cross-sectional views of RFP containing MOSFET structures with a Deep compensated zone (lightly doped p regions in **24**B and **25**B and lightly doped n regions in **24**C and **25**C), implemented in the split poly gated structure with a single poly layer structure in the RFP trench.

FIG. **26** shows a top view of an embodiment of FIG. **4**(*a*) wherein the RFP region is a continuous strip in the horizontal direction.

FIG. **27** shows a top view of an embodiment of FIG. **4**(*a*) wherein the RFP region is divided into several columns in the horizontal direction.

FIG. **28** shows a schematic flow chart for a sample fabrication process.

DETAILED DESCRIPTION OF SAMPLE EMBODIMENTS

Power MOSFETs are widely used as switching devices in many electronic applications. In order to minimize conduction power loss, it is desirable that MOSFETs have low specific on-resistance, which is defined as the on-resistance area product (R_{on}*A), where R_{on} is a MOSFET resistance when the MOSFET is in an ON state, where A is the area of the MOSFET. Trench MOSFETs provide low specific on-resistance, particularly in the 10-100 voltage range. As cell density increases, any associated capacitances such as a gate-to-source capacitance Cgs, a gate-to-drain capacitance Cgd and/or a drain-to-source capacitance Cds also increase. In many switching applications such as synchronous buck dc-dc converters in mobile products, MOSFETs with breakdown voltages of 30 V often operate at higher speeds approaching 1 MHz. Therefore, it may be desirable to minimize switching or dynamic power loss caused by these capacitances. The magnitude of these capacitances are directly proportional to gate charge Qg, gate-drain charge Qgd and output charge Qoss. Furthermore, for a device that operates in the third quadrant (i.e., when a drain-body junction becomes forward biased), minority charge is stored in the device during its forward conduction. This stored charge causes a delay in switching from conducting to non-conducting. To overcome this delay, a body diode with fast reverse recovery is desirable. However, a fast recovery body diode often causes high electromagnetic interference (EMI). This means that during diode recovery, the ratio between the negative going waveform (t_a) and the positive going waveform (t_b) must be less than one for a soft recovery which avoids EMI problem.

As switching speed requirements increase to 1 MHz and beyond with new applications, state of the art power MOSFETs are increasingly unable to operate at such high speeds with satisfactory efficiency. A power MOS transistor that has low charges Qg, Qgd, Qoss and Qrr in addition to having a low specific on-resistance (Ron*A), is desirable.

There are two common techniques to improve the switching performance of power MOSFETs. The first one is the trench-gated MOSFET with thick bottom oxide, as shown in FIG. **1** (U.S. Pat. No. 6,849,898). The second one is the split poly gated MOSFET structure, in which the first poly gate is electrically shorted to the source electrode (U.S. Pat. No. 5,998,833, 6,683,346), as illustrated in FIG. **2**.

Recently, as shown in FIG. **3**, U.S. Patent Application No. 2008/0073707 A1 to Darwish disclosed a power MOSFET with the recessed field plate (RFP) structure which realize a very short channel region (~0.25 μm) for further reducing the gate-source capacitance and the gate-drain capacitance, and consequently, the total gate charge (Qg) and the "Miller" charge (Qgd). The RFP structure additionally improves the body diode reverse recovery speed due to providing an additional path for current and the enhanced depletion of the drift region induced by the RFP.

The present application discloses improvements to power insulated-gate field effect transistors with Recessed Field Plate (RFP) and similar structures. The inventors have real-

ized that the performance of RFP-type power MOSFETs can be improved by performing a compensating implant into the RFP trench. This compensating implant helps to shape the depletion boundaries in the OFF state, and thus helps to avoid punchthrough. Because of this, a local enhancement can also be added to the doping between channel and drain, in the drift or spreading region. This provides a synergistic combination, wherein the on-resistance can be improved with no degradation in breakdown voltage.

The disclosed innovations, in various embodiments, provide one or more of at least the following advantages. However, not all of these advantages result from every one of the innovations disclosed, and this list of advantages does not limit the various claimed inventions.

- Improved (reduced) on-resistance;
- Improved (increased) breakdown;
- Reduced electrical stress on any dielectric layer at the bottom of the RFP trench;
- Higher reliability and longer operation life; and/or
- Increased ability to increase local doping concentration in the drift region.

The numerous innovative teachings of the present application will be described with particular reference to presently preferred embodiments (by way of example, and not of limitation). The present application describes several embodiments, and none of the statements below should be taken as limiting the claims generally.

For simplicity and clarity of illustration, the drawing figures illustrate the general manner of construction, and description and details of well-known features and techniques may be omitted to avoid unnecessarily obscuring the invention. Additionally, elements in the drawing figures are not necessarily drawn to scale, some areas or elements may be expanded to help improve understanding of embodiments of the invention.

The terms "first," "second," "third," "fourth," and the like in the description and the claims, if any, may be used for distinguishing between similar elements and not necessarily for describing a particular sequential or chronological order. It is to be understood that the terms so used are interchangeable. Furthermore, the terms "comprise," "include," "have," and any variations thereof, are intended to cover non-exclusive inclusions, such that a process, method, article, apparatus, or composition that comprises a list of elements is not necessarily limited to those elements, but may include other elements not expressly listed or inherent to such process, method, article, apparatus, or composition.

It is contemplated and intended that the design apply to both n-type and p-type MOSFETs; for clarity reason, the examples are given based on n-channel MOSFET structure, but an ordinary person in the art would know the variations to modify the design to make a similar p-channel device.

The present application discloses improvements to power insulated-gate field effect transistors with Recessed Field Plate (RFP) and similar structures. The inventors have realized that the performance of RFP-type power MOSFETs can be improved by performing a compensating implant into the RFP trench. This compensating implant helps to shape the depletion boundaries in the OFF state, and thus helps to avoid punchthrough. Because of this, a local enhancement can also be added to the doping between channel and drain, in the drift or spreading region. This provides a synergistic combination, wherein the on-resistance can be improved with no degradation in breakdown voltage.

In one sample embodiment, the RFP containing MOSFET has a buried Deep compensated zone floating in the N body region underneath the RFP trench. The Deep compensated zone reduces the voltage across the dielectric layer between the RFP and the N^- epitaxial layer when a high drain-source voltage is applied.

In one sample embodiment, the RFP containing MOSFET has a buried Deep compensated zone floating in the N body region underneath the RFP trench and a local enhancement to the doping between channel and drain, in the drift or spreading region in the N epitaxial layer

In one embodiment, the RFP containing MOSFET also has a deep P+ region in the P body in contact with the RFP trench wall that extends from the P body into the N epitaxial layer.

In one embodiment, the Deep compensated zone underneath the RFP trench extends to and is connected to the source electrode.

In one embodiment, the Deep compensated zone underneath the RFP trench extends to P body region that is in contact with the side walls of the RFP trench.

In one embodiment, the Deep compensated zone is a very lightly doped p region; while in another embodiment, the Deep compensated zone is a very lightly doped n region.

Referring now to FIG. **4**(*a*), a semiconductor device structure **100** comprises a gate **102** which is positioned in a first trench **104**, also described herein as a gate trench **104**. The first trench **104** containing the gate **102** may be one of many gate trenches within the semiconductor device structure **100**. The semiconductor device structure **100** is capacitively coupled to control vertical conduction from a source region **106**, having a first-conductivity-type, through semiconductor material **108**, which is adjacent to the first trench **104**.

As shown in FIG. **4**(*a*), the gate **102** has a gate electrode comprising conductive gate material having a width approximately equal to a width of the gate trench **104**. It will be appreciated that, although the gate electrode may have a width that is approximately equal to the width of the gate trench **104**, the gate electrode may alternatively be contacted using a wider gate trench and a smaller gate electrode, allowing the gate trench to be insulated from the gate conductor.

The semiconductor device structure **100** also includes recessed field plates **110**, which are positioned in proximity to and capacitively coupled to the semiconductor material **108**. The recessed field plates **110** are positioned in respective second trenches **112**, also described herein as RFP trenches **112**. Each of the trenches (i.e., the respective second trenches **112** and the gate trenches) has trench walls that are coated with an insulating material, such as silicon dioxide (SiO_2). The RFP trenches **112** contain an insulating material having a breakdown voltage which preferably exceeds the breakdown voltage of the semiconductor device structure **100**. The gate trench **104** preferably contains an insulating material up to the p-body drain junction, minimizing any overlap of the gate electrode (connected to the gate **102**) with the drain or drift region.

In one embodiment, the gate trench contains thick bottomed insulation dielectric material such as silicon dioxide. In another embodiment, the insulating material within the RFP trench and/or the gate trench **104** has a stepped thickness. Providing a stepped thickness can help shape the channel and can help control "hot" electron effects.

Conductive material, such as n-type doped polysilicon, forms a gate electrode electrically separated from the gate trench **104** by the insulating material. The conductive material may be silicided to reduce its resistance. Conductive material also fills the RFP trenches **112**, electrically separated from the gate trench **104** by the insulating material, and extending above the RFP trenches to form a plurality of RFP electrodes. Each of the trenches may be of substantially equal depth, or may differ in depth, and may be self-aligned by

being etched at the same processing step, although the RFP electrode is deeper than the gate electrode and is either independently biased or connected to a source electrode (i.e., the source **106**), and a source region (including the source electrode) may extend between the gate **106** and the RFP trenches **112**.

In one embodiment, the n-epitaxial drift region is uniformly doped. In another embodiment, the n-epitaxial drift region is not uniformly doped. Specifically, the doping is graded to have a doping concentration that is higher at an interface with the underlying **118** substrate and decrease toward the surface. Non-uniform doping of the drain drift region allows for greater shaping of the channel and for control over "hot" electron injection.

The source region may be doped n+. The gate trench **104** and the RFP trench may have a thin layer of the insulating material, reducing on-resistance, or a thick layer of the insulating material, providing greater electrical isolation increasing reverse-bias breakdown voltage. In the embodiment depicted, the RFP electrodes have a uniform depth. In another embodiment, at least one of the RFP electrodes extends up and contacts the source **106**.

Advantageously, the semiconductor device structure **100** also includes deep compensated zones **114** of either p-type or n-type lying at least partially beneath the respective RFP trenches **112**. The deep compensated zone **114** may be floating islands of either p-type dopant concentration regions (as shown in FIG. **4***a*), or lightly doped n-type dopant concentration regions in the N-drift region underneath the RFP trenches. The drawing shows the boundaries of this compensated zone **114** as if it had been fully counterdoped, but those of ordinary skill will understand that the boundaries of a compensated but not counterdoped zone can be similarly envisioned, using e.g. the concentration contours of a single dopant species.

The deep compensated zones **114** also reduce the voltage across the dielectric layer between the RFP and the N− epitaxial layer when a high drain-source voltage is applied.

As shown in FIG. **4**(*b*), device **100** also contains a deep p-body region **116** that is in contact with the side walls of the RFP trench **112**. The deep p-body region **116** with boundary at **116***a* can be in connection with the source electrode and also be in connection with the deep compensated zones **114**. The deep P-N junction in the edge termination can be formed by the Deep compensated implant and its related annealing without adding new mask. Therefore, the disclosed structure can offer a more reliable edge termination.

Alternatively, as shown in FIG. **4**(*c*), the deep compensated zones **114** can extend vertically and merge with the p-body region.

The two-dimensional electric voltage simulation shown in FIG. **5**, reveals that under the same bias conditions the conventional device of structure shown in FIG. **3** sees about 19V across the bottom dielectric layer between RFP and the N− epitaxial layer, while an embodiment of FIG. **4**(*a*)-(*c*) only shows 7V across the bottom dielectric layer between RFP and the N− epitaxial layer due to the protection from the Deep compensated zone **114**.

As the electrical stress on the bottom dielectric layer between RFP and drain is reduced significantly, the device structures of FIG. **4**(*a*)-(*c*) will offer a higher reliability and longer operation life. In addition, the Deep compensated zones **114** enhance the lateral and vertical depletion of N− epitaxial layer, which provides room for higher local doping concentration in the epitaxial layer without degrading the device breakdown voltage.

The increase in local doping concentration in the epitaxial layer further reduces the on-resistance of drift region. By properly adjusting the doping concentration of P and N regions in the N− epitaxial layer, the total on-resistance of device can be lowered without reduction of the breakdown voltage. Furthermore, the local doping enhanced N layer also decreases the minority carrier injection efficiency of the body diode of the device and alters the electric field distribution during the body diode reverse recovery. Thus, the reverse recovery of the body diode is improved, resulting in a device having lower reverse recovery charge and soft recovery features.

Since the doping enhancement only occurs in the active region, the termination efficiency of the improved device edge junction termination region will not be degraded.

The recessed field plates **110** may be positioned in multiple respective trenches **112**, which are separate from the gate trench **104**. Accordingly, the semiconductor device structure **100** may be, for example, an n-channel MOSFET with a recessed field plate (RFP) trench **112** and a gate trench **104** formed on an N-type epitaxial layer grown over a heavily doped N+ substrate.

In third quadrant operation, in which where the drain **118** is negatively biased with respect to a source-body electrode (i.e. the source **106**), and in which diffusion current results in minority carrier injection and a high reverse recovery charge Qrr, the plurality of RFP electrodes form majority carrier channel current path from drain to source in addition to that provided by the gate electrode in a conventional structure. The combined effect of the RFP electrodes and the gate electrode is both reduced minority carrier diffusion current and reduced recovery charge Qrr. Accordingly, in the third quadrant operation, the RFP electrodes act as an additional gate without any penalty of an added gate-drain capacitance Cgd.

In reverse-biased operation, the RFP also reduces any electric field in a channel region. Accordingly, shorter channel lengths are possible, without substantial risk of punch-through breakdown, further allowing reduction in Ron*A and Qg. The capacitive coupling between the gate trench **104**, the RFP trenches **112** and the drain region further deplete the drain drift region, at a higher rate as a drain-source voltage VDS is increased in an off-state. The low Cgd and its fast falling rate, combined with increasing drain-source voltage VDS, provides a lower gate-drain charge.

The semiconductor device structure **100** can have a quasi-vertical or lateral configuration. Ensuring that the semiconductor device structure **100** has a quasi-vertical or lateral configuration can help shape the channel, and can reduce hot electron effects.

Various variations in gate conductor and RFP conductors may be used. Various combinations have been shown in U.S. Application No. 2008/0073707 A1 to Darwish, the entirety of which is thereby incorporated by reference. Polysilicon may be used as the conductive material. Example variations in structural designs of gate conductor and RFP conductors include split poly configurations and single poly configurations (FIGS. **21**-**25**), thick bottom oxide, and step shaped bottom oxide and the combinations of the various forms.

Referring to FIG. **26**, each of the foregoing embodiments may be implemented in a single configuration, a multi-stripe configuration, a cellular layout configuration, or a combination of the foregoing. Moreover, the polarity and conductivity type may be reversed.

Referring to FIG. **27**, each of the foregoing embodiment RFP may also be implemented in an interrupted manner where the RFP trenches and conductors form columns in the

source-body-drain layers of the device. With this interrupted scheme, more N++ surface area can be provided, reducing the N++ resistance, and lowering the total on-resistance.

A fabrication process for making the described embodiment is detailed in FIG. **6-18**. In FIG. **6**, starting with N++ substrate **201**, the N-epitaxial layer **203** is grown followed by forming a thin layer of silicon oxide layer **205**. Substrate **201** may have been doped with phosphorus or arsenic. The preferred thickness for oxide layer **205**, for example, can be 200-300 Å. In FIG. **7**, the trench mask **207** is applied to form the hard mask for trench etching and the oxide layer is etched

Then a standard silicon etch step is carried out to form the plurality of trenches **209** according to the mask. In FIG. **8**, blanket implanting of phosphorus ions **211** (e.g. P^{31}) to the whole device may be performed to locally increase the doping concentration of N− epitaxial layer. The implantation is preferred to be done at tilt of 0 degree. The trench mask around the edge termination area or the gate bus area (not shown in Figures) prevents the phosphor dopant getting into these areas. Therefore, only the active region of the device receives the doping enhancement implant.

After implantation a high temperature process in an oxygen-containing ambient is used to anneal and diffuse the phosphorus dopant. Consequently, a doping enhancement N layer **213** is formed inside N− epitaxial layer as shown in FIG. **9**. The trench walls may then be oxidized first using a sacrificial oxidation. After removing the sacrificial oxide layer a pad oxide is re-grown along trench side wall. In FIG. **10**, the trenches are filled with high density oxide **217**. Oxide **217** may include silicon dioxide, or other types of deposited oxide, such as LTO or TEOS or High Density Plasma (HDP) oxide. The oxide is then thinned as shown in FIG. **11** using a dry plasma etch or CMP technique to planarize the oxide surface **219**.

In FIG. **12**, after active mask **223** has been applied with openings over trenches **222**, the oxide is etched further down into trench forming the trench Bottom Oxide layer (BOX) **221**. Then, in FIG. **13**, the BOX mask is used to protect the active gate trench **225** and the edge termination areas (not shown). The oxide removal step is carried out to completely etch away the BOX inside the RFP trench. Before removing the BOX mask, boron-11 ions **229** are implanted into N/N− epitaxial layer through RFP trench bottom **231**, forming P layer or isolation zones **237** shown in FIG. **14**.

In one embodiment, to implement the structure shown in FIG. **4**(*c*), a tilt angle implant is used to introduce boron along the RFP side walls. After removal of BOX photoresist **233** an optional high temperature anneal is employed to diffuse the boron, forming P layer or isolation zones **237** inside N− epitaxial region. Then the gate oxide **235** is grown along the trench sidewall in FIG. **14**.

The rest of process steps shown in FIG. **15** to FIG. **17** are similar to the one described in FIGS. 14-17 in U.S. patent application No. 2008/0073707, which is herein incorporated by reference. The final device structure is shown in FIG. **18**. It is essential to point out that, by properly choosing RFP poly recess depth combined with the implant energy of P+ implant, the P+ region can be made deeper than P body, as shown in FIG. **19**A. Depending on the doping concentration of P shield region (or isolation zone), the P shield zone could be a "π" region **260** (a very lightly doped P region) shown in FIG. **19**C or a "ν" region **250** (a very lightly doped n region) shown in FIG. **19**B. A deeper P+ region is desired in order to improve the device ruggedness and connect the buried P region to the source electrode. In addition, the N++ source region can also be recessed completely as shown in FIG. **20**, so that the N++ source mask-photo step can be eliminated.

Furthermore, the techniques proposed in this invention may also be implemented using split poly gated device structures. One of implementation schemes are briefly demonstrated in the FIG. **21** to FIG. **23**. The process includes deposition of a first poly layer in the trenches, poly etch-back, and oxide removal, gate oxidation, second poly layer deposition, and CMP and/or poly etch-back. The split gated double poly configuration shown in FIG. **21-23** is used to replace the single poly layer in the active trench gate and the RFP trench shown in FIG. **18**. In this case, the bottom poly layer and the upper poly layer in RFP trench are both electrically shorted to the source metal. In addition, the split poly layers in the RFP region of device in FIG. **23** can be directly replaced by the single RFP poly layer as demonstrated by FIG. **24**A and FIG. **25**A. Depending on the doping concentration of P shield region (or isolation zone), at very light concentration, the P shield zone could be a "π" region (a very lightly doped P region) or a "ν" region (a very lightly doped n region) as shown in FIGS. **24**B, **24**C, **25**B and **25**C.

FIG. **28** is a flow chart depicting a fabrication process for making a MOSFET in accordance with one embodiment of the present invention. The fabrication process includes growing **302** an N− epitaxial layer on an N+ substrate. The fabrication process also includes locally increasing **304** the doping concentration within the N− epitaxial layer. Locally increasing **304** the doping concentration within the N− epitaxial layer includes blanket implanting phosphorous. The blanket implanting of phosphorous may be at a tilt angle of zero degrees, or may be at some other tilt angle. Locally increasing **304** the doping concentration within the N− epitaxial layer also includes excluding phosphorous dopant from an edge termination area and/or a gate bus area, including leaving oxide at an edge termination area or a gate bus area.

The fabrication process for making a MOSFET also includes creating **306** a doping enhancement N layer inside the N− epitaxial layer, including annealing and spreading the phosphorous dopant using a high temperature thermal process in an oxygen ambient. The fabrication process for making a MOSFET also includes smoothing **308** the trench surface of the trench, and reducing the roughness of the trench surface induced during the silicon trench etch, including oxidizing the trench side wall, removing **310** the sacrificial oxide layer, and regrowing **312** the pad oxide along the trench side wall.

The fabrication process for making a MOSFET also includes filling **314** the trench with high density oxide, planarizing oxide surface **316**, including etching back an oxide, forming a Bottom Oxide layer (BOX), including etching further down into the trench using an active mask, and protecting **318** an active gate trench and an edge termination, using a BOX mask. The fabrication process for making a MOSFET also includes implanting **320** boron (B11) into an N− epitaxial layer through a recessed field plate trench bottom, including introducing boron along recessed field plate side walls at a tilt angle, completely etching away **322** the BOX inside an recessed field plate trench, including removing the oxide, optionally driving **324** boron forming a P layer into the N− epitaxial layer, including annealing at a high temperature, and growing **324** a gate oxide along a trench side wall.

For a sample 40V embodiment, preferred parameters are as follows. However, it must be understood that these parameters would be scaled for different operating voltages, and of course they can also be adapted for use with many other processes. In this sample embodiment, the trenches are 0.3 microns wide, about 1.0 micron deep, and are laid out on a one micron pitch. (The cell pitch is two microns, since there are two types of trench present.) In this sample embodiment, the

starting material is 0.35 ohm-cm n-on-n+ epi, about 5.5 microns thick. A blanket n-enhancement implant is performed, e.g. with phosphorus at 3E12/cm^2 (i.e. 3×10^{12} cm^{-2}). The trenches are then etched. After a sacrificial oxidation and trench fill (preferably using a deposited oxide plus oxidation), an etchback is preferably performed to clear the trenches to about half their depth. Photoresist is then patterned, to expose the RFP trenches but not the gate trenches, and the oxide plugs are removed from the RFP trenches. A P-type implant is then performed to form the P-isolation regions; in this example, a combination of two boron implants, one at 2.5E12/cm^2 at 30 keV plus another 2E12 at 120 keV. This will produce a counterdoped or compensated isolation region **114** below the RFP trenches, of about 0.7 micron depth. The remaining process steps then proceed conventionally, with formation of gate, body, source, contacts, etc.

As mentioned above, the locally enhanced n-doping which connects gate to drain, in various embodiments described above, reduces on-resistance. However, it is the improved off-state behavior provided by the added isolation regions which makes this enhanced n-doping possible.

In alternative embodiments, the depth of the isolation region can be e.g. from 0.25 micron to 2.5 micron, and scaled accordingly for operating voltages other than 40V. Similarly, the isolation implant, in alternative embodiments, can use a dose from 2E12 cm^{-2} to 1E13 at 20-320 keV, or even higher or lower doses and/or energies, plus allowance for scaling.

It will be appreciated that the foregoing is merely a description of some specific illustrative and exemplary embodiments of the present invention, and should not be considered as descriptive of the entire gamut of embodiments that fall within the scope of the present invention.

According to various embodiments, there is provided: a semiconductor device structure, comprising a gate which is positioned in a first trench, and capacitively coupled to control vertical conduction from a first-conductivity-type source through semiconductor material which is adjacent to said trench; recessed field plates, positioned in proximity to and capacitively coupled to said semiconductor material; said recessed field plates being positioned in respective second trenches; and diffusions of a second conductivity type lying at least partially beneath said respective second trenches.

According to various embodiments, there is provided: a semiconductor device structure, comprising a semiconductor layer; a gate which is positioned in a first trench within said semiconductor layer, and is capacitively coupled to control vertical conduction from a first-conductivity-type source through second-conductivity-type portions of said layer near said trench; recessed field plates, positioned in proximity to and capacitively coupled to said semiconductor material; said recessed field plates being positioned in respective second trenches; diffusion components of a second conductivity type lying at least partially beneath said respective second trenches; whereby said diffusion components reduce depletion of said second-conductivity-type portions of said layer in the OFF state.

According to various embodiments, there is provided: a semiconductor device structure, comprising a semiconductor layer; a gate which is positioned in a first trench within said semiconductor layer, and is capacitively coupled to control vertical conduction from a first-conductivity-type source through second-conductivity-type portions of said layer near said trench; recessed field plates, positioned in proximity to and capacitively coupled to said semiconductor material; said recessed field plates being positioned in respective second trenches; a first additional diffusion component of a second conductivity type lying at least partially beneath said respective second trenches; and a second additional diffusion component of said first conductivity type lying at least partially within said second-conductivity-type portions of said layer; whereby said first additional diffusion component reduces depletion of said second-conductivity-type portions of said layer in the OFF state; and whereby said second additional diffusion component reduces the on-resistance of the device in the ON state.

According to various embodiments, there is provided: an improved RFP transistor structure having (a) low total on-resistance, (b) reduced minority carrier injection efficiency (of body diode), (c) improved reverse recovery (of body diode), (c) lower reverse recovery charge, (d) soft recovery characteristic, (e) as reliable edge termination, without either reduction of breakdown voltage or degradation of the termination efficiency of device edge junction termination region, the improved structure comprising: an RFP transistor structure, including at least one or more gate trenches adjoined by one or more recessed-field-plate trenches; and respective deep compensated zones underneath said recessed-field-plate trenches.

According to various embodiments, there is provided: a method for operating a semiconductor device structure, comprising: controlling conduction between first and second source/drain electrodes through a channel location in semiconductor material using a gate electrode positioned in a first trench to provide at least ON and OFF states; and avoiding punchthrough of said channel location, using both one or more recessed field plates, positioned in proximity to and capacitively coupled to said semiconductor material; said recessed field plates being positioned in respective second trenches, and one or more diffusion components of a second conductivity type lying at least partially beneath said respective second trenches; whereby said diffusion components reduce depletion spreading in the OFF state.

According to various embodiments, there is provided: a fabrication process for making a MOSFET, comprising the actions, in any order, of: a) providing an n-type semiconductor layer; b) forming a p-type body in said layer; c) forming an n-type source, which is isolated by said body, in said layer; d) forming an insulated gate trench in said layer, and a gate electrode in said gate trench; said gate electrode being capacitively coupled to at least a portion of said body; e) forming a second insulated trench in said layer, providing an additional dose of acceptor dopants below said trench, and forming a Recessed Field Plate electrode in said second trench; and f) providing an additional dose of donor dopant atoms in said portion of said body, to thereby reduce the on-resistance.

According to various embodiments, there is provided: improved highly reliable power RFP structures and fabrication and operation processes. The structure includes plurality of localized dopant concentrated zones beneath the trenches of RFPs, either floating or extending and merging with the body layer of the MOSFET or connecting with the source layer through a region of vertical doped region. This local dopant zone decreases the minority carrier injection efficiency of the body diode of the device and alters the electric field distribution during the body diode reverse recovery.

Modifications and Variations

As will be recognized by those skilled in the art, the innovative concepts described in the present application can be modified and varied over a tremendous range of applications, and accordingly the scope of patented subject matter is not limited by any of the specific exemplary teachings given. It is

intended to embrace all such alternatives, modifications and variations that fall within the spirit and broad scope of the appended claims.

The device may be fabricated in various layouts, including "stripe" and "cellular" layouts. The layers of source, body, and drain regions can be configured vertically, quasi-vertically as well as laterally. The epitaxial drift region can be either uniformly or non-uniformly doped. While the embodiments described above include an epitaxial layer grown on a substrate, the epitaxial layer may be omitted in some applications. Various features of different embodiments may be combined and recombined for various applications.

For example, the region between channel and drain does not have to be uniformly doped, neither vertically nor laterally. The improvements in drift or spreading region doping provided by the disclosed inventions can be combined with a wide variety of other device improvements and features.

For another example, the RFP and gate trenches do not necessarily have to have the same width.

The design could be applied to IGBTs or other devices which include bipolar conduction. The bottom of the gate trench can be modified with dopant; the design can also vary at the source structure and at the drain structure; and alternative body structure may be used; contact trench may be produced first, then cut gate trench, and construct the source and drain structure.

Of course, the n-type dopant, in silicon, can be phosphorus, antimony, or arsenic, or combinations of these. Appropriate donor dopants can be used in other semiconductor materials.

As the disclosed process is scaled to other operating voltages, it is expected that predictable scaling of dimensions and dopants may allow the same synergy. For example, in a 200V embodiment, the inventors contemplate that the trench depth would be slightly deeper (e.g. 1.5 to 2.5 micron), and the compensation implant energy and dose would be about the same. Of course the epi layer doping would be substantially less, and the epi layer thickness greater, as is well understood by those of ordinary skill. The n-enhancement doping (which is preferably blocked from the termination) can have a distribution, after drive-in, which reaches to the upper boundary of the compensation implant, but preferably not to the lower boundary of the compensation implant.

The following applications may contain additional information and alternative modifications: Ser. No. 61/125,892 filed Apr. 29, 2008; Ser. No. 61/058,069 filed Jun. 2, 2008 and entitled "Edge Termination for Devices Containing Permanent Charge"; Ser. No. 61/060,488 filed Jun. 11, 2008 and entitled "MOSFET Switch"; Ser. No. 61/074,162 filed Jun. 20, 2008 and entitled "MOSFET Switch"; Ser. No. 61/076,767 filed Jun. 30, 2008 and entitled "Trench-Gate Power Device"; Ser. No. 61/080,702 filed Jul. 15, 2008 and entitled "A MOSFET Switch"; Ser. No. 61/084,639 filed Jul. 30, 2008 and entitled "Lateral Devices Containing Permanent Charge"; Ser. No. 61/084,642 filed Jul. 30, 2008 and entitled "Silicon on Insulator Devices Containing Permanent Charge"; Ser. No. 61/027,699 filed Feb. 11, 2008 and entitled "Use of Permanent Charge in Trench Sidewalls to Fabricate Un-Gated Current Sources, Gate Current Sources, and Schottky Diodes"; Ser. No. 61/028,790 filed Feb. 14, 2008 and entitled "Trench MOSFET Structure and Fabrication Technique that Uses Implantation Through the Trench Sidewall to Form the Active Body Region and the Source Region"; Ser. No. 61/028,783 filed Feb. 14, 2008 and entitled "Techniques for Introducing and Adjusting the Dopant Distribution in a Trench MOSFET to Obtain Improved Device Characteristics"; Ser. No. 61/091,442 filed Aug. 25, 2008 and entitled "Devices Containing Permanent Charge"; Ser. No. 61/118,664 filed Dec. 1, 2008 and entitled "An Improved Power MOSFET and Its Edge Termination"; and Ser. No. 61/122,794 filed Dec. 16, 2008 and entitled "A Power MOSFET Transistor".

None of the description in the present application should be read as implying that any particular element, step, or function is an essential element which must be included in the claim scope: THE SCOPE OF PATENTED SUBJECT MATTER IS DEFINED ONLY BY THE ALLOWED CLAIMS. Moreover, none of these claims are intended to invoke paragraph six of 35 USC section 112 unless the exact words "means for" are followed by a participle.

The claims as filed are intended to be as comprehensive as possible, and NO subject matter is intentionally relinquished, dedicated, or abandoned.

What is claimed is:

1. A semiconductor device structure, comprising:

a gate which is positioned in a first trench, and capacitively coupled to control vertical conduction from a first-conductivity-type source through semiconductor material which is adjacent to said trench;

recessed field plates, positioned in proximity to and capacitively coupled to said semiconductor material; said recessed field plates being positioned in respective second trenches; and

diffusions of a second conductivity type lying at least partially directly below said respective second trenches.

2. A semiconductor device structure, comprising:

a gate which is positioned in a first trench, and capacitively coupled to control vertical conduction from a first-conductivity-type source through semiconductor material which is adjacent to said trench;

recessed field plates, positioned in proximity to and capacitively coupled to said semiconductor material; said recessed field plates being positioned in respective second trenches; and

diffusions of a second conductivity type lying at least partially beneath said respective second trenches;

wherein said device further includes a layer of dopant concentration region of second-conductivity-type that extends from a source layer to at least one of said diffusions.

3. The semiconductor device structure of claim **1**, wherein at least one of said diffusions extends vertically and merges with a body layer of second-conductivity-type.

4. A semiconductor device structure, comprising:

a gate which is positioned in a first trench, and capacitively coupled to control vertical conduction from a first-conductivity-type source through semiconductor material which is adjacent to said trench;

recessed field plates, positioned in proximity to and capacitively coupled to said semiconductor material; said recessed field plates being positioned in respective second trenches; and

diffusions of a second conductivity type lying at least partially beneath said respective second trenches;

wherein said device further includes a layer of dopant concentration region of second-conductivity-type that extends from a source layer to at least one of said diffusions and at least one of said diffusions extends vertically and merges with a body layer of second-conductivity-type.

5. The semiconductor device structure of claim **1**, wherein the gate has a split poly configuration.

6. The semiconductor device structure of claim **1**, wherein at least one of the recessed field plates has a split poly configuration.

7. The semiconductor device structure of claim **1**, wherein both the gate and at least one of the recessed field plates have a split poly configuration.

8. The semiconductor device structure of claim **1**, wherein said first conductivity type is n-type.

9. The semiconductor device structure of claim **1**, wherein said gate is capacitively coupled to control vertical conduction to a drain diffusion of said first conductivity type.

10. A semiconductor device structure, comprising:

a semiconductor layer;

a gate which is positioned in a first trench within said semiconductor layer, and is capacitively coupled to control vertical conduction from a first-conductivity-type source through second-conductivity-type portions of said layer near said trench;

recessed field plates, positioned in proximity to and capacitively coupled to said semiconductor material; said recessed field plates being positioned in respective second trenches;

diffusion components of a second conductivity type lying at least partially directly below said respective second trenches;

whereby said diffusion components reduce depletion of said second-conductivity-type portions of said layer in the OFF state.

11. The semiconductor device structure of claim **10**, wherein said gate has a split poly configuration.

12. The semiconductor device structure of claim **10**, wherein at least one of said recessed field plates has a split poly configuration.

13. The semiconductor device structure of claim **10**, wherein both the gate and at least one of the recessed field plates have a split poly configuration.

14. The semiconductor device structure of claim **10**, wherein said diffusion components have a concentration sufficiently high to locally counterdope said semiconductor layer and thereby produce a second conductivity-type region below said second trench.

15. The semiconductor device structure of claim **10**, wherein said semiconductor layer is an epitaxial layer.

* * * * *