POWER DEVICE RELIABILITY

Some types of semiconductor devices handle high current (I) and high voltage (V) simultaneously. The product of V and I is equivalent to electrical power (P), hence the name "power" device. These devices can be singular (discrete) or multiple in form. They can also be integrated into a conventional integrated circuit (IC) process to create a power IC (PIC) technology. Power device reliability and robustness during normal and transient operating modes are extremely important considerations, to avoid unwanted failure resulting from the high power levels.

There are many power device types. We here concentrate on those most frequently encountered. Power devices can be either n type (electron carrier), p type (hole carrier), or a combination of both. They can also conduct current in the vertical and/or lateral dimension of the semiconductor. This generally leads to the prefix of "V" for vertical or "L" for lateral. One of the most common power devices is the insulated gate field-effect transistors (IGFETs). IGFET was originally termed MOSFET for metal oxide semiconductor, a terminology still commonly used despite most gates being constructed of polysilicon, not metal. While being displaced by MOS devices in some applications, bipolar junction transistors (BJTs) are still widely used in power applications. Some devices, such as the insulated gate bipolar transistor (IGBT), are a combination of MOSFET and BJT. Some devices combine both n- and p-type carriers, such as thyristors, which are a combination of n- and p-type BJTs. A very important type of power device used in PICs is the double diffused MOS or DMOS, which when fabricated for lateral current flow is called an LDMOS. A new variant of this type of device combines lateral channel current flow and vertical body current flow with a reduced surface field (RESURF). The RESURF device is becoming very popular for radio frequency (RF) applications (1). As a result, RF reliability issues are beginning to be investigated. It is known that the mechanisms which create device damage are mismatched antenna load conditions, which generate the same combinations of high I and V transients that affect other power device types.

Typical power devices have current handling capability maximums from 1 to greater than 1000 A and are able to withstand voltages in some cases of higher than 1500 V. When an application requires device operation at high power for a period of time (t), energy (E) is created as heat. Too much energy will destroy the power device. The energy the device can handle during normal modes of operation is measured to determine the current and voltage limits. The resulting graph of the current versus voltage limits defines the shape of a box. The area inside the box is termed the safe operating area (SOA). It is safe to operate the device within the limits defining the SOA.

SAFE OPERATING AREAS

The SOA is the acceptable region of operation of a bipolar or metal oxide semiconductor (MOS) transistor. SOA is defined in terms of voltage, current, and time. The most common forms of SOA are forward bias SOA (FBSOA), reverse bias SOA (RBSOA), and commutating SOA (CSOA) (2–4). Re-

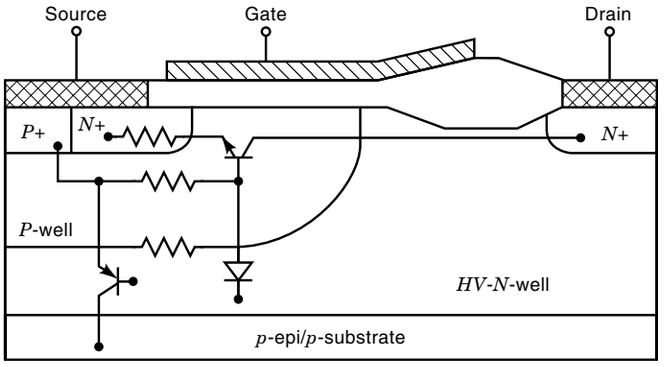

Figure 1. Cross section of lateral DMOS with schematic representation of intrinsic parasitic components, including source-drain *npn*, which can be turned on in the SOA limits.

cently a new form of SOA has been discovered that affects only MOS devices. Hot electron or HESOA is a result of prolonged high drain voltage at high gate voltage (5). The predominant power device used in modern integrated circuit (IC) processes is the lateral D-well metal oxide semiconductor field effect transistor (LDMOS). Figure 1 shows an LDMOS including its intrinsic parasitics. The components that have SOA dependency are the on-state metal oxide semiconductor field effect transistor (MOSFET), parasitic *npn* transistor, and body diode. The parasitics do not normally conduct current, but may in the SOA limits.

Forward-Biased Safe Operating Area

The FBSOA is restricted by the maximum voltage and current for which the part has been designed. Additional operational restrictions limit the voltage/current curve, including thermal considerations and, for bipolar devices, second breakdown. FBSOA defines the maximum useful on-state operating range. If the limits are exceeded, damage or device destruction can occur. Device design, package characteristics, and thermal capacity all affect the FBSOA. A typical curve is shown in Fig. 2. Limit *BC* represents the maximum current rating as stated in the device specification. Sometimes maximum current is increased under transient conditions, extending line *BC* to *B′C′* for pulsed operation. The region defining the direct current (dc) constant power limit is marked by line *CD*. Typically, maximum transient power is increased, leading to line *C′D′*. Maximum dc power capability is defined by the heat-sinking capability of the packaged device. *DE* is the second breakdown region, which is typically only significant in bipolar devices. The conditions for second breakdown are defined by the internal temperature of the device at which the voltage begins to collapse (6) where hot spots on the active silicon area are more prone to occur. Region *EF* is the maximum operating voltage as defined by the manufacturer's data sheet.

Line *AB* of Fig. 2 shows the on-resistance limit for the maximum gate drive, V_{gs}. This approximates to drain-source voltage (V_{ds})/drain current (I_d). This is not part of the intrinsic SOA curve, as it does not indicate a device destruction limit but shows the minimum on-resistance operating limit.

Reverse-Bias Safe Operating Area

RBSOA is the range of allowed operation when the device is off but forced to conduct current. This is a function of voltage, current, and time. It is typically defined for a particular circuit and specific operating conditions. RBSOA is important during inductive turn-off, where current initially remains constant, increasing voltage across the switching transistor.

RBSOA is the term typically reserved for bipolar devices. The MOSFET equivalent of RBSOA is unclamped inductive switching (UIS) since a reverse-biased gate is unnecessary for rapid FET turn-off. The UIS breakdown condition for a FET occurs through its parasitic *npn* (Fig. 1). The RBSOA and UIS failure mechanism occurs following avalanche breakdown. This breakdown mechanism is prone to instability, which can cause current to focus to a spot on the transistor. This is followed by rapid collector-emitter or drain-source voltage reduction. Transistor current increases, limited only by external circuit impedances. Without circuit current limiting the device overheats and ultimately degrades.

A typical RBSOA test circuit and waveforms are shown in Fig. 3. When the device under test (DUT) turns off, V_{ds} increases until the body diode is forced into avalanche multiplication, to maintain the inductor current. Other major MOSFET failure mechanisms that are grouped together under UIS failures include the following, which cause excessive power dissipation:

1. Parasitic *npn* turn-on due to high dV_{ds}/d_t
2. *npn* debiasing in the MOS device (turning the *npn* on)
3. dV_{ds}/d_t sensitive transient lateral voltage debiasing allowing partial re-turn-on

A typical method to improve FET UIS is to include a snubstack clamp (Fig. 4), which is designed to turn on the FET before device breakdown. Once the device is turned on, it operates in the FBSOA mode, where energy capability is considerably higher.

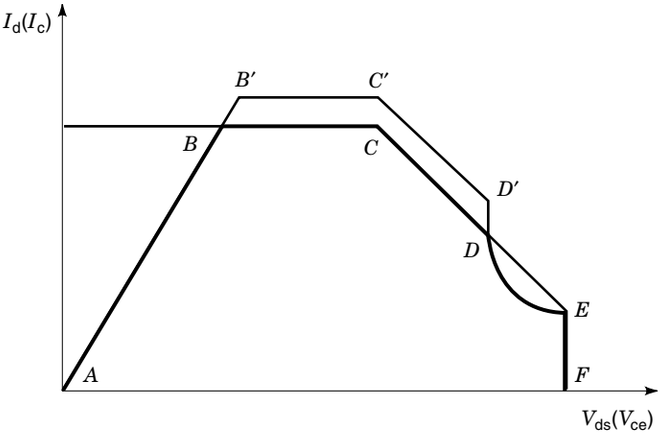

Figure 2. Forward bias safe operating area: *AB*—On-resistance limited operation. *BC*—Current-limited operation. *B′C′*—Transient current-limited operation. *CD*—Thermal-limited operation. *C′D′*—Transient thermal-limited operation. *DE*—Second breakdown limited operation (chiefly bipolar). *EF*—Voltage-limited operation.

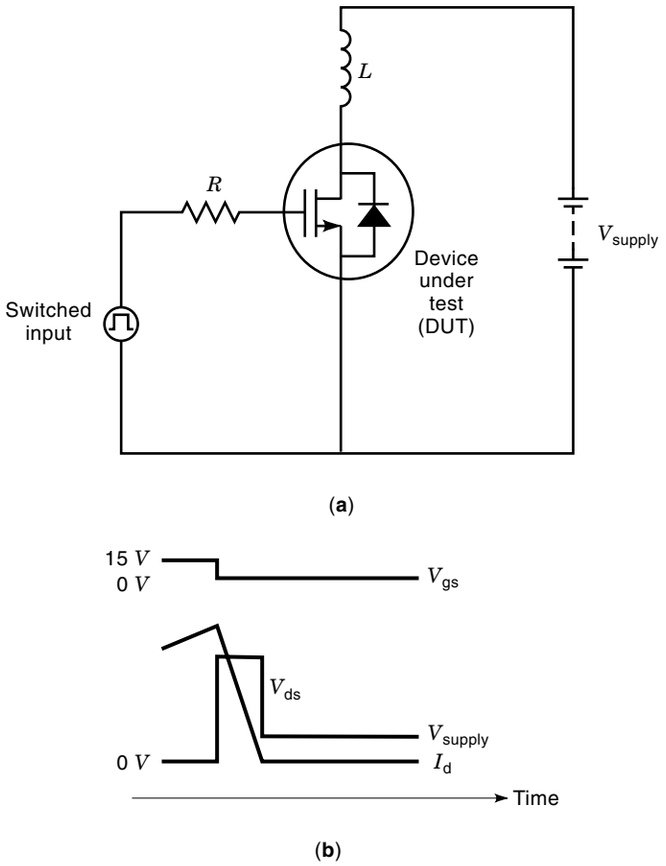

Figure 3. (a) Unclamped inductive switching test circuit showing pulsed switched input and inductive load. (b) Unclamped inductive switching waveforms for a DMOS-type MOSFET. The input and supply voltages can be adjusted to determine maximum SOA limits.

Commutating Safe Operating Area

The CSOA mechanism is similar to RBSOA. The CSOA boundaries are defined by the safe operation conditions for a diode. This is applicable in situations where MOS devices switch inductive loads, causing current flow through the body diode. This can occur in commutating conditions, such as when a bridge is used to drive a motor. Here a diode is typically forced to recover rapidly by instantaneous application of a large reverse voltage while it is conducting in forward mode. This reverse recovery phenomenon is well understood in discrete diodes and has recently become of interest in integrated diodes and especially FET body diodes.

Diode recovery is not instantaneous, and conduction occurs for a short time. The result is a "shoot-through" current. If recovery time is too long, or if repetition frequency is too high, power losses can cause operational inefficiencies or device destruction. The circuit in Fig. 5(a) shows a CSOA test configuration, and Fig. 5(b) its typical current and voltage recovery characteristics. Reverse recovery current can be many times higher than the forward current at turn-off. The higher the d_v/d_t turn-off rate, the larger the peak reverse current.

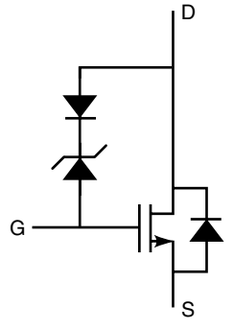

Figure 4. Snub-stack clamp used to improve UIS performance of a FET. The snub-stack consists of one or more zener diodes and forward biased diodes to determine the turn-on as drain voltage increases. The diodes which are in forward bias during snub-stack clamping block current flow when the gate is higher than the drain voltage (normal on-condition).

POWER MOSFET ELECTROSTATIC DISCHARGE RELIABILITY

Smart power chips built with low/high-voltage CMOS are a significant part of the semiconductor technology for automotive applications, where the harsh environment demands robust protection against the threat of electrostatic discharge (ESD) or other transient pulses, such as load dump (7). For the standard low-voltage CMOS logic chips, ESD protection

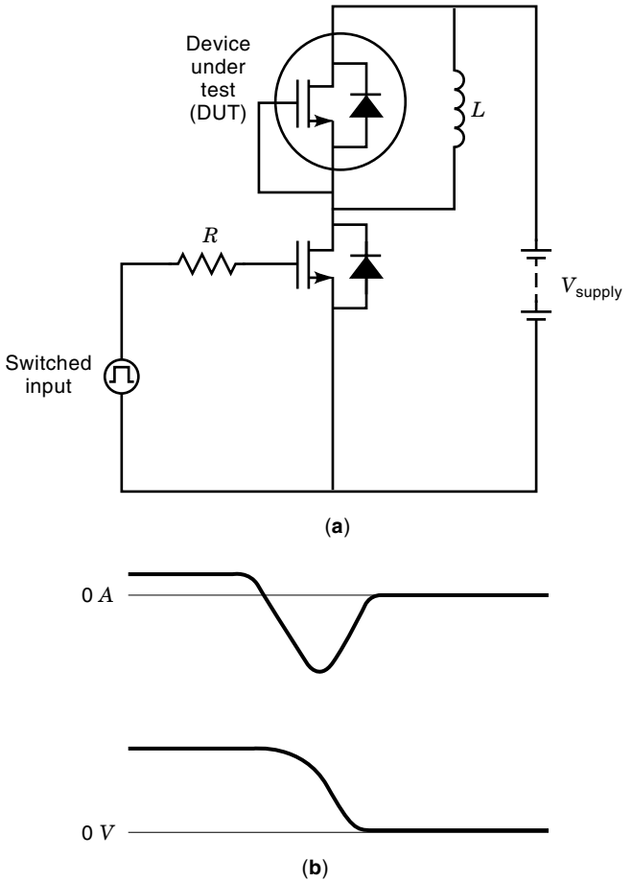

Figure 5. (a) Commutating SOA test circuit. (b) V_{ds} and I_d recovery characteristics for a MOSFET body-drain diode.

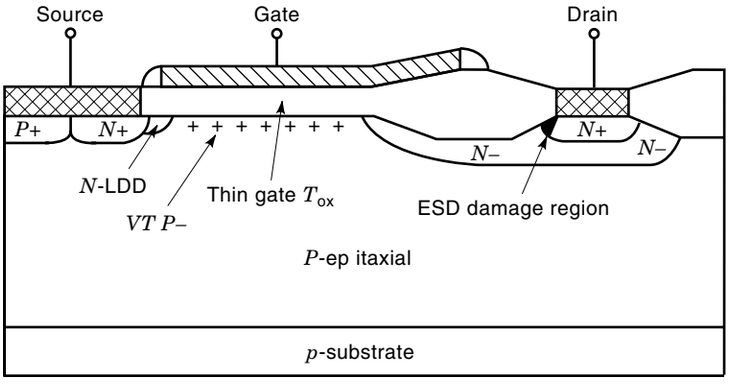

Figure 6. Cross section of drain extended NMOS (DENMOS) showing damage area, caused by ESD overstress.

against the human body model (HBM) stress levels of ±2 kV is commonly required. An immunity to the machine model (MM) stress model for ±200 V is also required from most customers of the power devices. In addition, the automotive applications in some cases require a higher HBM level of protection at least ±4 kV. However, achieving even a minimum 2 kV HBM protection level is a challenge for the power MOSFET structures. Usually achieving 3 kV of HBM protection ensures that 200 V for the MM can be met.

The high-voltage-power MOSFETs, somewhat contrary to their name, are relatively ineffective for the ESD protection design. As a result, an understanding of their behavior under ESD conditions and device optimization to counter this becomes necessary. In the following sections the ESD phenomena and protection design for two different high-voltage MOSFETs will be discussed. The first is the drain extended NMOS (DENMOS), where ESD performance can be optimized through device structure. The second is the LDMOS, which can be effectively used as an MOS clamp to carry ESD current while minimizing its power dissipation.

ESD PROTECTION USING DRAIN EXTENDED NMOS

ESD Phenomena in DENMOS

The typical cross section of a DENMOS is shown in Fig. 6. For integration with 5 V and 1 μm CMOS technologies, the DENMOS channel length range is 3 to 4 μm and the gate oxide is about 500 Å. Under drain avalanche, the thin oxide region near the gate as indicated in Fig. 6 does not break down since part of the voltage is supported by the depletion of the n-tank. With ESD pulse the voltage at the drain reaches avalanche

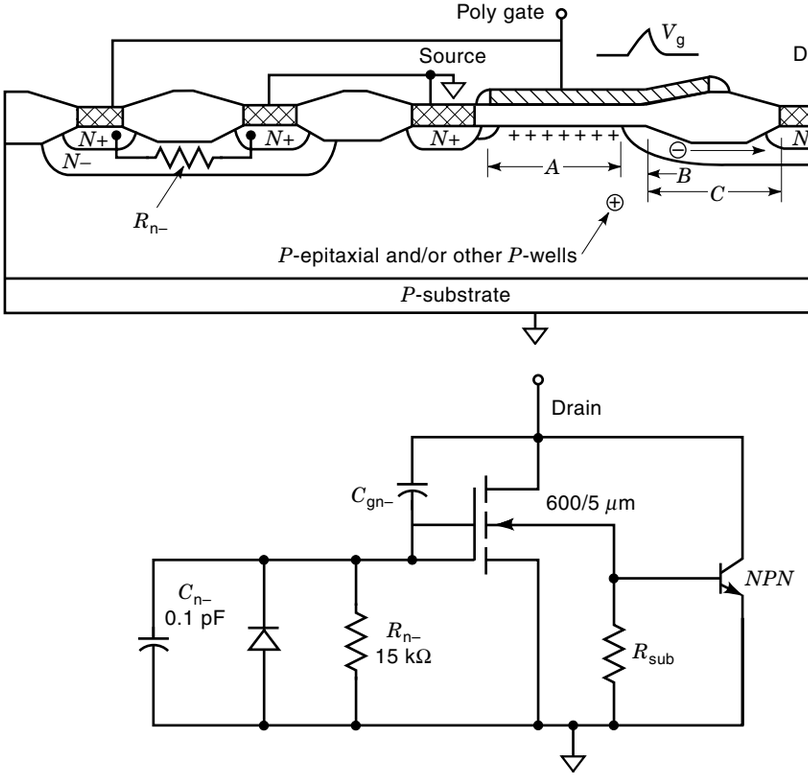

Figure 7. Cross section of gate coupled drain extended NMOS type ESD structure (top) and its equivalent circuit (C_n and the diode are parasitics of tank resistor R_n, designed to inhibit MOS turn-on during fast transients.

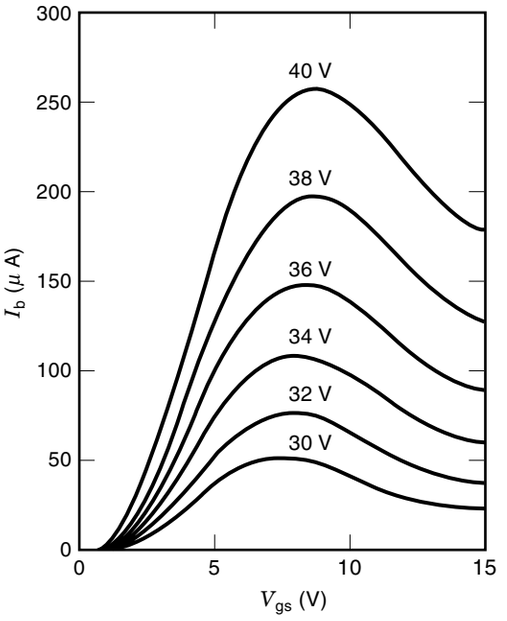

Figure 8. Substrate current characteristics (I_b) of a 25/5 μm DENMOS as a function of different drain bias values. This plot shows that maximum substrate current occurs for gate voltages around 8 V.

breakdown and the device fails even before the parasitic *npn* fully turns on. This parasitic *npn*, formed with the drain as the collector, substrate as the base, and source as the emitter, is the most important device effect for any NMOS under ESD stress. For the standard CMOS transistors with channel lengths of 1 μm or smaller, the parasitic *npn* effectively turns on and robust ESD levels can be designed. Full ESD protection is achieved when the *npn* operates below the second breakdown condition. However, in the DENMOS because of the relatively longer channel length, as the bipolar attempts to turn on, the *n*-region of the drain depletes, exposing the LDD junction near the drain. The $\boldsymbol{J} \cdot \boldsymbol{E}$ (where $\boldsymbol{J}$ is the current density and $\boldsymbol{E}$ is the peak electric field) at this point leads to local melt filament and leakage. The failure in device occurs at or slightly below 2 kV for a 400 μm wide device, and this does not improve much when a larger multifinger device is used since the first finger that attempts to turn on as a bipolar would fail first. This weakness is only for positive stress, and for negative stress the n^+/n^- region of the drain to p^+ (substrate) acts as a forward-biased diode to give good ESD performance to the p^+ substrate. There are two different ways to improve the ESD performance of this device: (1) generate local substrate bias during the ESD event, or (2) integrate a thyristor or semiconductor-controlled-rectifier (SCR) latchup device into its structure to clamp the voltage to a low level.

Gate-Coupled DENMOS

For using the DENMOS as an ESD protection device, the gate is usually tied to ground to prevent leakage. However, as mentioned, the lateral *npn* is hard to turn on because of the large base width. If this device does not turn on, the drain junction becomes a reverse-biased diode to substrate and the resulting ESD protection level is low. One way to improve the *npn* turn on is to bias the substrate during the ESD event (8). This can be done using a gate coupled technique, where a large resistor of 10 to 15 kΩ is connected to the gate of the DENMOS. The cross section and equivalent circuit schematic of this device are shown in Figs. 7(a) and 7(b), respectively. During the ESD pulse the gate is dynamically coupled due to the overlap capacitance between the drain and gate. In Fig. 7(a), spacing A refers to the channel length, and $B + C$ refers to the *n*-tank extension of drain. The actual gate to drain overlap capacitance (C_{gn-}) is determined by spacing B. Under ESD, with the drain going high, the gate is pulled up through C_{gn-} and is discharged to ground through R_{n-} (15 kΩ). For typical spacing of $B = 2$ μm, the gate couples up to 9 V for a fast transient on the drain from 0 to 45 V corresponding to the drain avalanche condition. Now considering the typical substrate current versus gate voltage plot shown in Fig. 8, it can be seen that maximum substrate current occurs for a gate voltage range of 8 to 10 V. Indeed these conditions can generate enough substrate bias near the source to turn on the bipolar *npn*. For example, if the resistivity of the p^+ high-voltage tank region near the source is 30 kΩ-μm, then 20 μA/μm of substrate current can generate sufficient potential to forward bias the source-substrate junction. Combined with the high bipolar beta of the lightly doped *p*-region, this allows an effective turn-on of the lateral *npn*, resulting in excellent ESD protection for the power IC.

SCR Integrated DENMOS

An SCR is one of the most powerful and efficient ESD protection devices available in a CMOS technology. Its trigger voltage is both process and design dependent. To protect a DENMOS the SCR can be integrated into its structure, as shown in Fig. 9 (9). Note here that the SCR is formed with n^+ and p^+ regions, placed in an *n*-well and connected to the

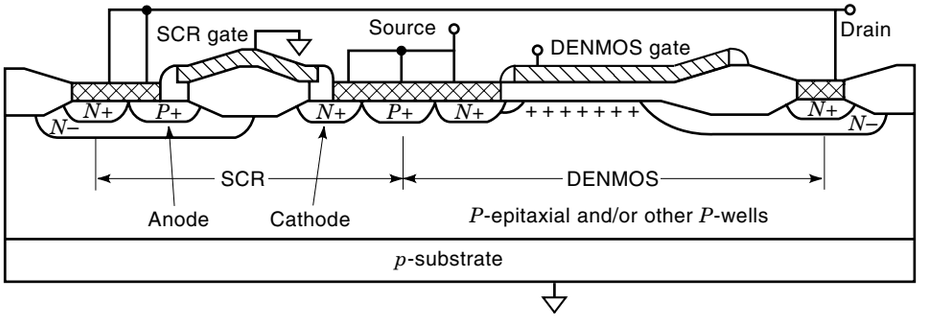

Figure 9. Cross section of DENMOS with lateral SCR integrated into the DENMOS structure.

drain, forming the anode. The cathode is an n^+ region connected to ground that is shared with the source of the DENMOS. The SCR trigger is determined by the n-well overlap of P^+ and the spacing from well to n^+ cathode. At typical 3 μm for these spacings, the SCR trigger can be in the 45 V range. This value actually depends on the doping of the n-well, but since the DENMOS is also built in a similar n-well, the SCR trigger would be close to the drain avalanche breakdown of the DENMOS. A spacing of 3 μm, lower than the channel length of 5 μm for the DENMOS, will ensure that the SCR will trigger first. To ensure that the SCR will trigger first, a polysilicon grounded gate is added to the SCR. The I–V characteristics of this hybrid device are shown in Fig. 10. Note that the SCR clearly triggers before the drain avalanche of the DENMOS. The holding voltage of the SCR is about 3 V, leading to low power dissipation and improved ESD performance. The structure of Fig. 9 can be used as an output buffer transistor, with the DENMOS gate tied to the appropriate predrive circuitry. As long as the device is operated in the region indicated in Fig. 10, the integrated SCR/DENMOS device will serve as a normal power high-voltage transistor for circuit applications.

ESD Protection Using Lateral DMOS

The LDMOS is a common high-voltage transistor that is used for power circuit applications. As shown with its cross section of Fig. 1, the drain is built inside a high-voltage n-tank and the source inside a p-tank. In the usual circuit applications it operates as an MOS device. As shown in the typical schematic of a gauge driver in Fig. 11, the gate and drain are clamped with zener diode stacks. When an ESD pulse is applied to the output pin, the LDMOS cannot go into bipolar breakdown but simply operates as a MOS device even under ESD. Since the surface MOS current density is much lower than for the parasitic bulk bipolar *npn*, the device width has to be appropriately sized to dissipate the ESD current levels of 1–2 A. These issues are discussed here.

In an LDMOS the self-heating effect reduces the channel mobility. But under ESD events, which are too short at the order of 100 ns, the heating effect is negligible and the drain current is relatively larger. This is seen in Fig. 12, where the drain current as a function of gate bias for a 2400 μm wide, 1 μm channel length LDMOS is shown for both ESD pulse conditions and based on a normal curve trace method. These curves are for a drain clamp of 20 V. It can be seen that with no channel heating and with only mobility degradation due to the vertical electric field, a current of 1.6 A can be handled at a gate bias of 18 V. Thus this corresponds to 0.67 mA/μm or 1.0 V/μm of ESD capability, and hence a 2000 μm wide device could provide 2 kV ESD level for HBM under this clamp design.

At high drain potentials the power dissipation in the device increases and reduces the aforementioned ESD capability. Thus an ideal LDMOS device design for ESD would have a high gate clamp voltage to obtain highest possible MOS conduction, and a low drain clamp voltage to minimize the power dissipation. For any given technology, this type of analysis can be done to optimize the LDMOS design to obtain the desired ESD levels, as given by the following equation (10):

$$V_{\mathrm{ESD(HBM)}} = [0.5W(V_{\mathrm{GC}} - V_t)^2]/[(1 + 0.1(V_{\mathrm{GC}} - V_t))V_{\mathrm{DC}}]$$

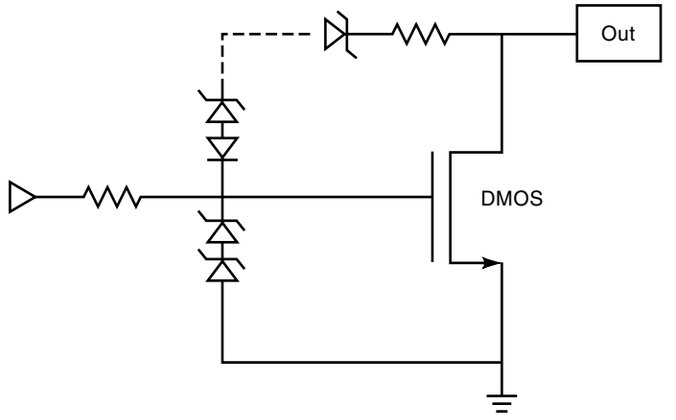

Figure 11. Circuit and ESD application of an LDMOS design. This shows full protection of an LDMOS output structure including gate-source protection clamp and SOA breakdown protection.

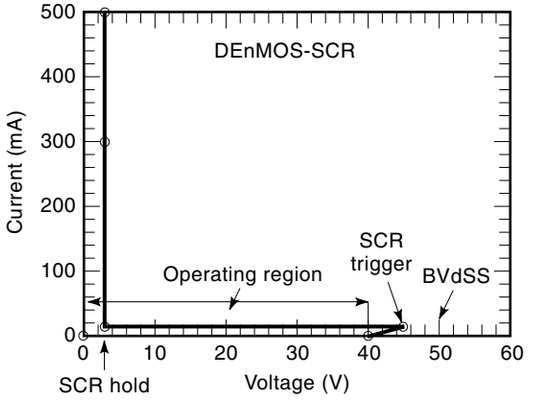

Figure 10. I–V characteristics of the integrated DENMOS-SCR. Note the SCR is designed to trigger before the BV_{dss} of the DENMOS is reached. If the design does not permit SCR triggering, catastrophic DENMOS breakdown is liable to occur.

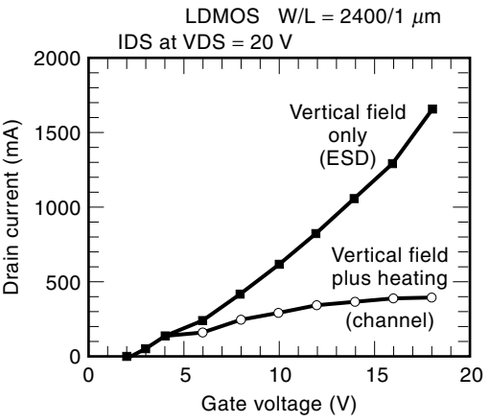

Figure 12. High-current I–V data for an LDMOS under 200 ns ESD pulses (filled squares) and standard non-ESD curve trace (open circles). This shows the detrimental effects of heating on the device, and can lead to falsely high R_{dson} results.

where $V_{ESD(HBM)}$ is the ESD level in volts that can be obtained for the human body model. W is the device width, V_{GC} and V_{DC} are the gate and drain clamps in volts, and V_t is the threshold voltage in volts.

BIBLIOGRAPHY

1. I. Yoshida, 2-GHz Si power MOSFET technology, *IEDM Tech. Dig.,* 51–54, Dec. 1997.
2. J. Devore et al., *Safe Operating Area—Interfacing to the Real World,* Analog and Mixed Signal Conference, Santa Clara, CA, Oct. 30–Nov. 1, 1991, paper 27.
3. C. Hu and M.-H. Chi, Second breakdown of vertical power MOSFETs, *IEEE Trans. Electron. Devices,* ED-29, 1287–1293, 1982.
4. D. L. Blackburn, Turn-off failure of power MOSFETs, PESC 85, *16th Annu. IEEE Power Electronics Specialist Conf.,* Toulouse, France, June 24–28, 1985.
5. S. Manzini and C. Contiero, Hot electron induced degradation in high-voltage submicron DMOS transistors, *ISPSD Proc.,* 1996, pp. 65–68.
6. H. A. Schafft and J. C. French, Second breakdown in transistors, *IEEE Trans. Electron Devices,* ED-13, 613–618, 1962.
7. T. Efland et al., The load dump, *IEEE Workshop on Electronics Applications in Transportation,* Dearborn, MI, Oct. 18–19. IEEE Catalog No. 90TH0310-3, 1990, 73–78.
8. C. Duvvury et al., Efficient *npn* operation in high voltage NMOSFET for ESD robustness. *IEDM Tech. Digest,* 345–348, 1995.
9. C. Duvvury et al., Device integration for ESD robustness of high voltage power MOSFETs, *IEDM Tech. Digest,* 407–410, 1994.
10. C. Duvvury et al., Lateral DMOS design for ESD robustness, *IEDM Tech. Digest,* 1997.

ANDREW MARSHALL
CHARVAKA DUVVURY
TAYLOR EFLAND
Texas Instruments Incorporated