Key criteria for the short-circuit capability of IGBTs

V. van Treek, H.-J. Schulze, R. Baburske, F. Hille, F.-J. Niedernostheide, and F. Pfirsch,
Infineon Technologies AG
Neubiberg, Germany

Abstract
Short-circuit behavior and capability are investigated and optimized during IGBT development. Thereby, knowledge about destruction and high-frequency short-circuit oscillation mechanisms is needed. For the thermal destruction mechanism, filaments are formed shortly before destruction during the thermal run-away itself, whereas for the electrical destruction mechanism strong current filaments are formed by an electrical mechanism, before the self-heating in the filaments leads to a thermal run-away. At low collector-emitter voltages, weak non-destructive filaments exist for a large current range. For both the filament formation and short-circuit oscillations (SCOs), an electric-field peak in the field-stop layer and a quasi-plasma layer beneath the MOS cells are mandatory. For SCOs, which are caused by a periodic storage and release of charge carriers inside the device, additionally, a weak electrical field at the beginning of the drift zone is necessary. Weak, non-destructive filaments and SCOs are likely to occur simultaneously. An increase of the bipolar current gain reduces the operating area with SCOs and increases the electrical short-circuit capability. A simultaneous reduction of the thermal short-circuit robustness can be avoided by advanced p-emitter concepts or (over-)compensated by an improved thermal setup.

Keywords: IGBT, short circuit, thermal short-circuit robustness, electrical short-circuit robustness, short-circuit destruction, high-frequency short-circuit oscillations, SOA

INTRODUCTION

For many industrial applications, IGBT chips and modules are required that have a certain short-circuit capability. During short-circuit operation, almost the entire DC-link voltage drops between collector and emitter, while the IGBT carries the driver and DC-link voltage-dependent saturation current, which is several times larger than the nominal current of the IGBT. IGBTs have to withstand a short-circuit incident for the time needed to detect the incident, and to turn-off the IGBT. The short-circuit robustness depends on parameters such as stray inductance during the short-circuit incident, emitter inductance, gate resistance, initial junction temperature and possible clamping circuits.

Short-circuit behavior and capability are investigated and optimized during IGBT development. Particular attention is paid on the thermal and electrical short-circuit robustness, and the occurrence of high-frequency short-circuit oscillations (SCOs). The optimization of the short-circuit capability can be challenging, because thermal short-circuit robustness, electrical short-circuit robustness and performance properties, such as conduction losses, switching losses and turn-off softness, are competing objectives. It can even be more challenging if undesired high-frequency SCOs need to be minimized in the application-relevant $I_C$-$V_{CE}$ phase space.

For the short-circuit optimization of IGBTs, knowledge about destruction and high-frequency SCO mechanisms is necessary for identifying improvement measures. Therefore, this work gives an overview of the thermal and electrical short-circuit destruction mechanisms and the SCO mechanism. Improvement measures and interdependencies of these three criteria of the short-circuit capability of IGBTs are discussed.

SHORT-CIRCUIT TYPES

For the same DC-link voltage and collector current, the operating state of the IGBT at the moment of a short circuit can determine whether the IGBT is destroyed or not. Different short-circuit types can be distinguished (e.g. [1]). The two most important types for IGBTs are described subsequently.

Short-circuit type I

Characteristic for short-circuit type I (SC1) is the turn-on of the IGBT during an existing short circuit: the DC-link voltage is initially applied across the turned-off IGBT. When the gate voltage exceeds the threshold voltage, the collector current increases, causing a voltage drop across the stray inductance, and thus, a $di/dt$-dependent reduction of the collector-emitter voltage below the DC-link voltage. If a certain switching speed is exceeded, a collector-current overshwing and a subsequent over-voltage occur during the transition to the short-circuit operating point. As long as the IGBT remains in the active region, both the collector-current overshwing and
the overvoltage are significantly smaller for SC1 than for short-circuit type II.

**Short-circuit type II**

Characteristic for short-circuit type II (SC2) is the occurrence of the short circuit while the IGBT is turned on. During SC2, collector current and collector-emitter voltage increase initially according to the corresponding output characteristic. The displacement current through the voltage-dependent feedback capacitance causes a voltage drop across the gate resistances. Therefore, the gate-emitter voltage increases above the driving voltage, and with that, a relatively large collector current over-shoot can be caused. The subsequent overvoltage during the negative \( \frac{dV}{dt} \) can be significantly above the DC-link voltage. Accordingly, the stress and the dissipated energy during the transition into the short-circuit operating point is usually larger for SC2 than for SC1.

**THERMAL SHORT-CIRCUIT CAPABILITY**

**Characteristics of thermal short-circuit destructions**

For long enough short-circuit pulses, a thermal destruction can occur within the pulse. Nevertheless, the thermal or energy short-circuit robustness is usually related to the destruction of an IGBT after its turn-off. The IGBT can withstand the short-circuit pulse and turn-off; however, some hundreds of microseconds up to milliseconds after the pulse, leakage currents increase and the IGBT is destroyed. For a given setup, the thermal destruction depends mainly on the dissipated energy during the short-circuit pulse [2]-[3], and is more critical for low-voltage IGBTs and relatively thin IGBTs.

**Measurement of the critical short-circuit energy**

By increasing the short-circuit pulse width in small increments (e.g. of 0.2 \( \mu s \)) until destruction, the critical short-circuit energy can be measured for defined test conditions. Usually, the device under test (DUT) is measured at a high initial device temperature, a high DC-link voltage, a typical driver voltage, and typical stray inductance and gate resistance.

**Simulation setup and destruction mechanism**

The thermal short-circuit capability of IGBT chips can be investigated by an electro-thermal device simulation which takes the electrical and thermal setup into account [4]. The thermal setup enables the simulation of the heat transport out of the silicon chip. It could e.g. consider the front-side metallization, the solder beneath the collector and the direct bonded copper (DBC) as insulator with their specific heat conductivities and heat capacitances. Effects due to an inhomogeneous current distribution are of minor importance, because once they have an impact on the critical short-circuit energy, the destruction occurs within and not after the short-circuit pulse [11]. Therewith, half-cell simulations are sufficient, since homogeneous self-heating can be assumed in the active area [5].

Such electro-thermal device simulations show that during the short-circuit pulse, the temperature increases mainly in the drift zone. After the pulse, the heat diffuses to the chip’s front and back side. There, the rising temperature causes increasing leakage currents, which are amplified by the bipolar current gain of the collector-sided \( p-n-p \) transistor \( \alpha_{pp} \). The leakage currents depend on the dissipated energy during the short-circuit pulse, the thermal setup and the IGBT chip. Up to a certain energy dissipation, the transient increase of the leakage current is followed by a decrease of the leakage current (Fig. 1). However, beyond a critical short-circuit energy, the leakage current keeps increasing. Excessive leakage currents lead to an almost homogeneous self-heating of the chip. After the critical temperature is reached, a local thermal runaway due to latch-up occurs at a position with a slightly higher temperature. The corresponding current crowding destroys the chip.

![Fig. 1: Electro-thermal simulation. Left picture: the collector emitter current (log scale) initially increases after the short-circuit pulse, and leads to device destruction if a critical energy is exceeded. Right picture: IGBT cell and mounting setup. (picture from [4])](image_url)
The bipolar current gain of the collector-sided $p$-$n$-$p$ transistor $\beta_{n-p}$ can e.g. be reduced by a decrease of the $p$-emitter dose or an increase of the field-stop dose. For a conventional soft solder die attach with a thickness of e.g. fifty microns, the heat capacitance at the chip front side can be significantly increased if an aluminum front-side metallization with a thickness of a few microns is replaced by a copper metallization that is thicker by a factor of two or five. Since diffusion solder is almost one order of magnitude thinner than the conventional soft solder, the thermal conductivity at the chip’s back side can e.g. be improved, if a conventional soft solder is replaced with a diffusion solder. An enormous increase of the critical short-circuit energy can only be achieved if cooling measures minimize leakage currents at both chip sides (Fig. 2). As soon as the front side is able to heat up, the corresponding leakage-current increase is amplified by the bipolar current gain. A temperature increase and the corresponding leakage-current increase at the back side cause an increasing temperature and leakage current at the emitter, which is again amplified by the bipolar current gain. Hence, improving the thermal setup only at one side yields in a limited improvement of the critical short-circuit energy.

The chip’s current destruction limit, which is well above the save operating area (SOA), depends on the collector-emitter voltage (Fig. 3). For medium and high-voltage IGBTs, the current destruction limit has a minimum value at medium collector-emitter voltages [8]-[10].

![Fig. 3: Schematical view of the $I_C-V_{CE}$ phase diagram with the 15 V- and 0 V-output characteristics, the current destruction limit and $I_C-V_{CE}$ trajectories of the different non-destructive SC1 incidents close to the four failure modes (picture from [9])](image)

### Failure modes and electrical destruction mechanism

According to Fig. 3, four different failure modes can be distinguished [8]:

- pulse failure (failure mode 1),
- turn-off failure (failure mode 2),
- static clamping turn-off failure (failure mode 3) and
- turn-off in a transient low-voltage stage failure (failure mode 4).

However, only failure modes 1 and 2 are related to the current destruction limit. Failure mode 3 is linked to the blocking capability, and failure mode 4 to the reverse-biased safe operating area. Hence, failure modes 3 and 4 are not further discussed in this paper. Failure modes 1 and 2 correlate to current filaments, which are initiated at the drift-zone field-stop junction. If the IGBT stays too long in the area above the chip’s current destruction limit, the current filaments become strong enough to provoke a local thermal runaway in the current filament(s) with the strongest self-heating [10]. Due to the dynamics of the thermal runaway, the destruction often occurs when the area with destructive filaments in the phase diagram is already passed. However, the time delay is significantly smaller than in the case of the energy destruction.

The $I_C-V_{CE}$ trajectories, the duration in the area above the chip’s current destruction limit and, thus, the critical short-circuit current depend strongly on the stray inductance, the turn-on and turn-off gate resistance, the pulse width [9] and the operating condition of the IGBT at the moment of the short circuit. For SC1, failure mode 1 is likely to occur for higher DC-link voltages, where the critical short-circuit current increases with the collector-emitter voltage, as the critical area is passed during the transition to the short-circuit operation. In contrast, failure mode 2 can be expected for lower DC-link voltages, where the current destruction limit decreases with the collector-emitter voltage, as the critical area is passed.
during the IGBT’s turn-off. At lower DC-link voltages, the IGBT can endure current filaments relatively long due to their limited energy dissipation. A high enough overvoltage during the IGBT’s turn-off and the corresponding energy dissipation might also be sufficient for the destruction even though the critical area is not passed.

TCAD simulation of the filamentation border

TCAD simulations show that the relocation of the electric-field peak from the chip front side towards the field-stop layer [6]-[7], and the occurrence of a low-field/quasi-plasma region beneath the MOS cells [8], are necessary for the filament formation. The short-circuit current/gate-emitter voltage at which these two conditions are fulfilled increases with the collector-emitter voltage (Fig. 4).

![Simulated electric-field distributions for different collector-emitter voltages and different gate voltages](image)

Fig. 4: Simulated electric-field distributions for different collector-emitter voltages and different gate voltages (left: emitter, right: collector). Schematic IGBT output characteristics are displayed in the background. (picture from [10])

In [10], the ratio of the gradients of the electric-field strength in front of the field-stop layer and below the cell field at the beginning of the drift zone,

\[ c = \frac{(dE/dy)_{\text{near field stop}}}{(dE/dy)_{\text{near front side}}} \]

was introduced for estimating the appearance of a quasi-plasma at the beginning of the drift zone (Fig. 5). For different collector-emitter voltages, the ratio is calculated during the simulation of the transfer characteristic. For ratios significantly above 1, a quasi-plasma layer builds up and filaments may occur. This simple and phenomenological criterion can be adjusted to the increasing branch of the measured destruction limit (Fig. 6). A large collector-current range between the measured destruction limit and the estimated filamentation border exists at lower collector-emitter voltages. In this area, at a certain current significantly below the current destruction limit, the measured critical energy starts to decline, and the failure mode changes from a thermal to an electrical destruction [11]. The change in the failure mode at this transition current can be explained by weak current filaments that remain non-destructive due to their relatively low energy dissipation. The existence of non-destructive current filaments in the area between the filamentation border and the destruction limit could also be confirmed by images of thermal-reflectance measurements, which show a regular pattern above the estimated filamentation border, for different points in the \( I_c-V_{CE} \) phase space above the safe operating area (SOA) [12].

![Schematic view of the hole density](image)

Fig. 5: Schematic view of the hole density (gray) and the electric-field strength (red) in an IGBT during a short circuit with a high current (picture taken from [5])

![I_c-V_{CE} phase diagram](image)

Fig. 6: \( I_c-V_{CE} \) phase diagram with critical SC current (of last non-destructive pulse) of a 1200 V IGBT (black line), the borderline for the appearance of a front-side plasma layer estimated by TCAD simulations with \( c = 30 \) (blue) [10] and transition current between energy and electrical destruction (red) together with thermal reflectance measurements for certain points [12] (picture taken from [11])

Measures to improve the electrical short-circuit robustness

By increasing the value of \( \ell_{\text{gap}} \), the relocation of the electric-field peak, the occurrence of a quasi-plasma layer, and thus, the electrical short-circuit destruction can be shifted to higher \( I_c \) values. However, it should be kept in mind that for very high \( \ell_{\text{gap}} \) values, dynamic avalanche is triggered [7]. Advanced concepts, such as the injection enhanced floating emitter, enable the improvement of the electrical short-circuit capability without reducing the thermal short-circuit capability as the hole injection is only enhanced during short-circuit conditions [13].

HIGH-FREQUENCY SHORT-CIRCUIT OSCILLATIONS

During short-circuit operation high-frequency short-circuit oscillations (SCOs) may be present. Their occurrence depends on both the chip and the circuit design. Typically, SCOs occur at low junction temperatures in a DC-link voltage interval well below the nominal voltage.
Measurement of SCOs

A sample measurement with SCOs is shown in Fig. 7. SCOs do not only depend on the DC-link voltage $V_{DC\text{-}link}$, they also depend strongly on the driver voltage $V_{Dr}$. For a test-vehicle chip, which was intentionally designed with a large operating area with SCOs and strong oscillation amplitudes, this is shown in Fig. 8. For determining of the $V_{Dr}$-$V_{DC\text{-}link}$ phase space with SCOs, the minimum and maximum DC-link voltage and the minimum driver voltage with SCOs were determined by means of test measurements. Afterwards, a measurement matrix was specified. Within this matrix, the DC-link voltage was varied in 25 V steps and the driver voltage in 100 mV steps. The measured short-circuit conditions were then assigned to one of the following categories, whereby, the second half of the short-circuit pulse was considered as quasi-stationary (Fig. 7).

- The red category is characterized by SCOs during the quasi-stationary short-circuit.
- The green category is characterized by no SCOs during the quasi-stationary short-circuit.

With their corresponding color code, the operating points were drawn into the $V_{Dr}$-$V_{DC\text{-}link}$ phase space (Fig. 8, top). Additionally, the maximum gate voltage variation (Fig. 8, middle) and the dominant oscillation frequency (Fig. 8, bottom) were evaluated. Although the device under test was intentionally designed with a large operating area with SCOs and strong oscillation amplitudes, it was not destroyed during the measurements.

The collector and emitter currents of a transient simulation of an operating point with SCOs are depicted in Fig. 9. After a transition time with increasing oscillation amplitudes, permanent oscillations with an almost constant oscillation amplitude occur. At the front side, the electron current is almost constant due to the fixed gate-emitter voltage. However, at the back side the electron current has a significant oscillation amplitude. The hole current has a smaller oscillation amplitude at the collector than at the emitter. The different electron- and hole-current amplitudes at the emitter and the collector indicate a periodic storage and release of charge carriers in the IGBT cell during the intrinsic oscillation.

Simulation setup and intrinsic SCO mechanism

TCAD simulations of a circuit with just one IGBT cell, a driver and a DC-link voltage source demonstrate that IGBTs can already have a built-in ability to SCOs for some operating points [15]. For the TCAD simulation of intrinsic SCOs, a two-step simulation procedure was proposed [15]. First, the operating point was simulated in a quasi-stationary manner by simulating a transfer characteristic or output characteristic up to the operating point. The result of the quasi-stationary simulation was then used as initial condition for the second step: the transient isothermal simulation of the operating point.

![Image](https://example.com/image.png)

**Fig. 7**: Measured transients $V_{CE}(t)$, $i_{C}(t)$, and $v_{CE}(t)$ for a short-circuit pulse with $V_{Dr} = 15$ V, $V_{DC\text{-}link} = 250$ V and $T_I = 300$ K of a 1200 V IGBT test-vehicle chip

![Image](https://example.com/image.png)

**Fig. 8**: Measured $V_{Dr}$-$V_{DC\text{-}link}$ phase space with SCOs at room temperature (top), the corresponding contour plots of the maximum gate-voltage variation $Δv_{GE}$ (middle), and the dominant oscillation frequency $f$ (bottom) during the quasi-stationary short-circuit of a 1200 V 100 A IGBT test-vehicle chip (picture from [15])

![Image](https://example.com/image.png)

**Fig. 9**: Simulated total, electron and hole current at collector and emitter of an operating point with high-frequency SCOs ($V_{Dr} = v_{GE}(t) = 13.5$ V, $V_{DC\text{-}link} = v_{CE}(t) = 200$ V, $T_I = 300$ K)
and holes are released. The electrons flow to the back side and increase the electron current at the collector. The holes flow to the front side and increase the hole current at the emitter. Since most of the carriers are stored at the beginning of the drift zone, fewer carriers are released at the front. The carrier confinement in the drift zone causes the characteristic current shape between \( t_2 \) and \( t_6 \). At \( t_6 \), the collector current of the quasi-stationary simulation is reached again, and the charge-carrier storage and release start again. Although most of the carriers are stored at the beginning of the drift zone, and thus, close to the gate trench, MOS capacitances are not necessary for the intrinsic SCO mechanism [15].

An electric-field peak in the field-stop layer and a quasi-plasma layer beneath the MOS cells are mandatory for the occurrence of high-frequency SCOs. Additionally, a weak electrical field beneath the MOS cells is necessary [14], [15]. Since the electric field builds up in the quasi-plasma layer with increasing \( V_{CE} \) and \( I_C \) values, SCOs usually disappear in well-designed IGBTs until approximately one-third of the nominal voltage, and well before the electrical destruction current. Accordingly, we have not seen a destruction due to SCOs in our extensive SCO measurements of IGBT chips of different generations and voltage classes.

The electric-field peak in the field-stop layer and a quasi-plasma layer beneath the MOS cells are also necessary for the filament formation. Under conditions where the electric-field strength in the quasi-plasma layer at the beginning of the drift zone is not too large, SCOs and non-destructive filaments are therefore likely to occur simultaneously near the filamentation border (Fig. 11).

The distributions of carrier densities and the electric-field strength at selected points in time of one oscillation period during permanent SCOs are shown in Fig. 10. At \( t_0 \), the transient collector current equals the collector current of the corresponding quasi-stationary simulation of the operating point. However, compared to the corresponding quasi-stationary simulation, fewer carriers are present in the drift-zone at \( t_0 \). Between \( t_0 \) and \( t_2 \), the collector current is smaller than in the quasi-stationary simulation, as a significant number of charge carriers is stored in the drift zone. Since electrons are stored in the drift zone, the electron current at the collector is smaller than at the emitter. The storage of holes in the drift zone results in a reduced hole current at the emitter (Fig. 9). At \( t_2 \), the collector current of the quasi-stationary simulation is again reached. However, many more carriers are present in the drift zone at \( t_2 \). After \( t_2 \), the transient collector current is larger than the quasi-stationary current. Between \( t_2 \) and \( t_5 \), a carrier front develops and propagates towards the collector. At the front, electrons

![Fig. 10: Vertical distributions of electrons (solid lines) and holes (dashed lines) (middle), and of the electric-field strength (bottom) at the indicated points of one oscillation period (top). The gray characteristics show the result of the corresponding quasi-stationary simulation.](image)

![Fig. 11: Critical SC current (of last non-destructive pulse) depending on \( V_{CE} \) of a 1200 V IGBT for \( L_{diff} = 70 \) nH and 25°C (black line), borderline for the appearance of a front-side plasma layer estimated by TCAD simulations with \( c = 30 \) (blue), transition current between electrical and energy destruction (brown) as well as measured operating area with SCOs (red)](image)
Measures to reduce high-frequency SCOs

An increased $\alpha_{\text{top}}$ strengthens the electric field beneath the MOS cells, and thus, reduces SCOs. Accordingly, both the area with SCOs and the oscillation amplitudes decrease for a decreasing field-stop dose (Fig. 12, top) and an increasing p-emitter dose (Fig. 12, bottom). The reduction of the operating area with SCOs goes along with an increased electrical short-circuit robustness. However, unless advanced concepts, such as the injection enhanced floating p-emitter, are used, an increase of the bipolar current gain results in increased leakage currents and a reduced thermal short-circuit capability. An improved thermal setup could (over-)compensate an increased bipolar current gain.

Fig. 12. Comparison of the simulated SCO OA border of the 1200 V IGBT structure for a variation of the initial field-stop dose by the factor $d_{\text{fss}}$ (top) and for a variation of the initial p-emitter dose by the factor $d_{\text{pem}}$ (bottom) (picture from [15])

The occurrence of SCOs does not only depend on the chip design, the circuit design has also an impact as e.g. shown in [16].

CONCLUSIONS

Current filaments and local thermal runaway are characteristic of both the thermal and the electrical short-circuit destruction mechanism. Accordingly, in both cases, the failure picture shows a local melting of metallization and silicon in the active area [11]. However, for the thermal destruction mechanism, filaments are formed shortly before destruction during the thermal runaway itself, whereas for the electrical destruction mechanism strong filaments are first formed above the SOA by an electrical mechanism before the self-heating in the destructive current filament leads to a thermal runaway.

At low collector-emitter voltages, weak, non-destructive current filaments may exist for a large collector current range. For the filament formation and the occurrence of SCOs, both an electric-field peak in the field-stop layer and a quasi-plasma layer beneath the MOS cells are mandatory. Accordingly, weak, non-destructive current filaments and SCOs are likely to occur simultaneously. However, for SCOs, which are caused by a periodic storage and release of charge carriers in the device, additionally, a weak electrical field beneath the MOS cells is necessary.

An increase of the bipolar current gain reduces the operating area with SCOs and increases the electrical short-circuit capability. The simultaneous reduction of the thermal short-circuit robustness and the impact on performance properties, such as conduction and switching losses and turn-off softness, can be avoided by advanced concepts, such as the injection enhanced floating p-emitter, or compensated by an improved thermal setup. The thermal short-circuit capability can be improved significantly by improved double-sided cooling without affecting the electrical short-circuit capability and performance properties.

REFERENCES


Addresses of the authors

Vera van Treek, Am Campeon 1-15, 85579 Neubiberg, Germany, Vera.vanTreek@infineon.com

Roman Baburske, Am Campeon 1-15, 85579 Neubiberg, Germany, Roman.Baburske@infineon.com

Frank Hille, Am Campeon 1-15, 85579 Neubiberg, Germany, Frank.Hille@infineon.com

Hans-Joachim Schulze, Am Campeon 1-15, 85579 Neubiberg, Germany, Hans-Joachim.Schulze@infineon.com

Franz-Josef Niedernostheide, Am Campeon 1-15, 85579 Neubiberg, Germany, Franz-Josef.Niedernostheide@infineon.com

Frank Pfirsch, Am Campeon 1-15, 85579 Neubiberg, Germany, Frank.Pfirsch@infineon.com

DOI: https://doi.org/10.14311/ISPS.2021.001